



Cooperation over finite horizons: A theory and experiments

Attila Ambrus^a, Parag A. Pathak^{b,c,*}

^a Department of Economics, Harvard University, Cambridge, MA 02138, United States

^b Society of Fellows, Harvard University, Cambridge, MA 02138, United States

^c MIT, United States

ARTICLE INFO

Article history:

Received 1 August 2008

Received in revised form 5 September 2010

Accepted 19 November 2010

Available online 7 December 2010

Keywords:

Public goods

Restart effect

ABSTRACT

This paper shows that the presence of different types of players – those who only care about their own material payoffs and those who reciprocate others' contributions – can explain the robust features of observed contribution patterns in public good contribution games, even without the presence of asymmetric information. We show what conditions on reciprocity are sufficient for a unique perfect equilibrium, in which contributions are decreasing. Under these conditions, selfish players have enough future benefits to induce subsequent contributions by reciprocal players, and this incentive diminishes as the end of the game approaches. The model explains the puzzling restart effect and is consistent with various other empirical findings. We also report the results of a series of experiments, using a probabilistic continuation design in which after each set of 10-period games, the group is restarted with low probability. We find specific support for the theory in our data, including that selfish players (identified exogenously) stop contributing earlier than reciprocal players, as directly implied by the model.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Finitely repeated public good contribution games are popular models of social dilemma situations with a fixed time-horizon.¹ Players face myopic incentives not to contribute to the public good, even though contributing is socially efficient. Standard game theoretic solution concepts predict that players should not contribute.

In contrast to this prediction, there is considerable experimental evidence, starting with Isaac et al. (1984) and Kim and Walker (1984), that finitely repeated interaction facilitates significant levels of cooperation. Moreover, across experiments, there is a robust pattern: a relatively large amount of cooperation in early periods eventually breaks down and by the end of the interaction, there is little cooperation.² These observations also hold for games played by experienced subjects.³ Furthermore, they prevail even in games that follow a “surprise restart”

announcement. Namely, at the end of a repeated public good contribution game, if a surprise announcement is made that the same group of subjects will play another repeated game, contributions initially jump back to a relatively high level and then decrease again over time. This “restart effect,” which was first reported by Andreoni (1988), suggests that the decreasing contributions phenomenon is not the result of a simple learning procedure through which players learn to play the standard equilibrium prediction.

In this paper we show that these documented patterns of cooperation can be explained by a model involving heterogeneous social preferences, even without incorporating asymmetric information.⁴ The basic feature of the model is that we assume the presence of both selfish players (in the sense of maximizing only their own material payoffs) and players who reciprocate contributions by others. This assumption is in accordance with the observations of Fischbacher et al. (2001), Brandts and Schram (2001), Palfrey and Prisbey (1997), Ledyard (1995), and Saijo and Yamaguchi (1992) that roughly half of the subjects in public good experiments maximize individual payoffs, while 40–50% of them are conditional cooperators.⁵ Also motivated by existing experimental evidence, we assume that the reciprocal players

* Corresponding author. Society of Fellows, Harvard University, Cambridge, MA 02138, United States.

E-mail address: ppathak@mit.edu (P.A. Pathak).

URLs: <http://post.economics.harvard.edu/faculty/ambrus/ambrus.html> (A. Ambrus), ppathak@fas.harvard.edu, <http://econ-www.mit.edu/faculty/ppathak> (P.A. Pathak).

¹ These games are also widely studied in other social sciences. Fiske (1992) and Field (2002) are references in sociology and anthropology, respectively. See also the references in Chapter 6.1 of Plott and Smith (2008). For a survey of the early economics literature, see Ledyard (1995).

² According to Fehr and Schmidt (2002), in the final period, roughly 75% of the subjects contribute nothing to the public good, and the rest contribute very little. Andreoni (1988) and Andreoni and Petrie (2004) find that roughly 70% of the subjects contribute zero, while the rest of the subjects contribute about 28% of their endowment, on average.

³ See Isaac and Walker (1988).

⁴ Although we do not extend our analysis to incorporate asymmetric information, we expect that asymmetric information regarding players' types could lead to initial high initial and low later contributions under much weaker assumptions than in the current complete information setting. This is because reputational concerns in such models would provide another incentive for selfish players to start out with positive contributions.

⁵ The idea of conditionally cooperating behavior originated in social psychology (see Kelley and Stahelski, 1970). The first related work in economics that we are aware of is Guttman (1978).

reciprocate both past realized contributions and current expected contributions by others.⁶ The reciprocal preferences we adopt can be derived from various underlying motives, including fairness considerations, conditional altruism, or following some social norm.

We show that a set of conditions on the reciprocity functions imply the existing experimental findings as an equilibrium phenomena. The conditions on reciprocity functions that we identify are arguably strong. However, they can be substantially relaxed to obtain weaker results (such as that ultimately the contribution pattern becomes decreasing). The main driving force in our model is that selfish players can influence future contributions of reciprocal players, and the more periods are left, the higher the increment they can induce on these contributions. As a result, it is worthwhile for them to contribute more of their endowment to the public good at the beginning of the game. In equilibrium, reciprocal players correctly anticipate these high contribution levels in early periods, which induces them to also contribute. As the game progresses, selfish players have less incentive to contribute, and in equilibrium, their contributions to the public good decrease. Lastly, decreasing contributions by the selfish players imply decreasing contributions by the reciprocal players. The same logic can be used to explain the restart effect: since a major factor in determining equilibrium contributions is the number of remaining periods in the game, a surprise announcement of playing additional periods increases equilibrium contributions.⁷

The second part of this paper reports the results of a series of experiments designed to test various assumptions and implications of the model. Each experiment is based on repeated sets of 10-period linear public good contribution games. To be able to conduct multiple “surprise” restarts, we use a design in which after each set of 10-period games, it is randomly decided whether the group stays together and plays a “restarted” 10-period game (25% probability) or whether players in the group are randomly reshuffled and play the next set of 10-period games with a new group (75% probability). To approximate the complete information assumption of our model, we focus on repeated games in which players are experienced, either because the game is a restarted one with the same participants or because they are shown the previous choices of each other in a decision revealing reciprocal behavior.

First, we investigate whether the behavior of experienced players can be approximated as an equilibrium of the game, by testing whether experienced players correctly foresee the contribution pattern in the game. We conduct a treatment where we solicit player's forecasts of others' subsequent contributions before the start of a 10-period game. Experienced players' forecasts on average closely track the average of actual play, with the median forecasted average contribution near the median of average contribution. In particular, before the start of the restarted game, 69% of subjects anticipate that there will be no contribution in the tenth game, yet each of these subjects contribute a positive amount in the first game, contributing more than half their endowment on average.

Next, we show that selfish players could find it worth to contribute positive amounts, because it induces future contributions by others. To demonstrate this, we regress players' second-period contributions on others' first-period contributions. In restarted games, a unit increase in average contributions of others increased a player's second-period contribution by a statistically significant 0.73 units.

Finally, we test a direct implication of our model, namely that selfish players stop contributing earlier than reciprocal players. We do this using data from sessions in which the 10-period public good contribution games are preceded by a sequence of gift-giving games.

We use data from these gift-giving games to classify players as selfish or reciprocal based on their offers as second proposers. The average reciprocal player stops contributing between period 7 and 8, while the average selfish player stops contributing between period 5 and 6, a statistically significant difference.

Our theoretical work complements Fischbacher and Gächter (2010), who empirically connect decreasing contribution patterns to the presence of both players who are selfish and players who are conditional cooperators. They do not provide a formal theoretical framework for their findings, but their interpretation of their experiments has similarities with the equilibrium dynamics in our model.

2. Related literature

A number of papers examine strategic interaction between selfish and reciprocal players. Offerman et al. (1996), Anderson et al. (1998), and Cox (2007) analyze public good contribution games, but without dynamic strategic considerations.⁸ Fehr and Schmidt (1999), and Andreoni and Samuelson (2006) investigate two-period games with multiple types. Brown et al. (2004) study a finite-horizon contracting game in which a high enough fraction of workers are fairness-minded. Duwfenberg and Kirchsteiger (2004) extend reciprocity equilibrium (Rabin, 1993) to extensive form games.

There are various explanations of cooperation over finite horizon in which the focus is not on the dynamic interaction of different types of players. Radner (1980, 1986) shows that cooperation can be maintained for a while in a repeated oligopoly game and in a repeated prisoner's dilemma if players only care about maximizing their payoff up to epsilon precision, even for small values of epsilon. A somewhat similar argument is presented by Klumpp (2010) for repeated public good contribution games. He shows that even a relatively small amount of altruism can generate large contributions at the beginning of the game. The scope of these explanations are limited by the fact that in games with discrete action spaces, small departures from maximizing individual payoffs cannot explain any amount of cooperation, unless the number of periods is very large, while large deviations do not explain the breakdown of cooperation in the end. Neyman (1985) shows that cooperation can be achieved in the equilibrium of a finitely repeated prisoner's game if players can only use strategies with bounded complexity. Jehiel (2005) presents a behavioral solution concept for multi-stage games, analogy-based expectation equilibrium, in which some players cannot distinguish between different stages of the game, and just best-respond to the average behavior of other players. The paper does not analyze repeated public good contribution games, but in somewhat similar contexts (like centipede games) it shows that the solution concept allows for initial cooperation between players. Two recent papers, Mengel (2009) and Ule (2010) consider models with players who are both backward-looking and limited forward-looking and show that this can lead to cooperation in finitely repeated prisoner's dilemma games.⁹

Finally, there are various explanations which relax the assumption that the fundamentals of the game are common knowledge among players. Kreps and Wilson (1982), Kreps et al. (1982), Sobel (1985), and Fudenberg and Maskin (1986) show how a small amount of uncertainty about payoffs (reputation) can induce cooperation in games with finite horizon, while Neyman (1999) points out that a small departure from the length of the game being common knowledge can lead to cooperative outcomes.

⁶ Sonnemans et al. (1999) and Keser and Winden (2000) find both forward-looking and backward-looking (adaptive) behavior in public good experiments.

⁷ Related to this point, the model we present implies that in a longer game, contributions to the public good are more persistent, another empirical result reported for example by Isaac et al. (1994).

⁸ See also Levine (1998). Outside the class of public good contribution games, Cox et al. (2007) introduces a parametric model of other-regarding preferences and structurally estimates it for simple games such as the ultimatum game.

⁹ The idea of limited forward-lookingness goes back to Stahl and Wilson (1995).

3. Model

We consider a T -period public good contribution game with $N \geq 2$ players. The model below could be extended in a straightforward manner to games in which for any strategy profile by others, a player's best response action and the action that maximizes the joint payoff of the players are always ordered the same way (for example, the former is always smaller than the latter). As a result, our qualitative results can be extended to a broad class of games, including repeated prisoner's dilemmas and repeated oligopoly games. Nevertheless, for ease of exposition we stick to the framework of public good contribution games.

Besides denoting the number of players, we also use N to denote the set of players whenever it does not cause confusion. In each period, each player has an endowment of 1 unit. Players in each period simultaneously decide how much of their endowment to contribute for public investment and how much to retain for private investment. Let $x_i^t \in [0, 1]$ denote player i 's contribution to the public investment in period t . After each period, players observe the contributions by all other players.

The material payoff of player i in period t is the amount of endowment she retains for herself plus her share from the aggregate returns to the public investment:

$$(1 - x_i^t) + \frac{A}{N} \sum_{j \in N} x_j^t, \quad \forall i \in R.$$

Public investment yields a constant marginal return A , which is divided equally to all players. We assume $A > 1$, but $\frac{A}{N} < 1$.

Players $i = 1, \dots, S$ maximize the sum of their per period material payoffs. From now on, we refer to them as selfish players. Let S also denote the set of selfish players and R denote the rest of the players. Players $R = \{S + 1, \dots, N\}$ are reciprocal. Their payoffs are determined through their reciprocity functions. It is convenient to think about these functions as specifying target contribution levels. The arguments of f_i^t , the period- t reciprocity function of player $i \in \{S + 1, \dots, N\}$, are past and current contributions to the public good by others. We assume that every reciprocity function is nondecreasing in all other players' contributions and takes values in $[0, 1]$.

To keep the analysis tractable, we only consider reciprocity functions which are additively separable with respect to contributions made at different periods and which, within the same period, are additively separable with respect to contributions made by different players:

$$f_i^t \left((x_j^1)_{j \in N/\{i\}}, \dots, (x_j^t)_{j \in N/\{i\}} \right) = \sum_{k=1}^t \sum_{j \in N/\{i\}} f_{i,j}^{t,k} (x_j^k),$$

where $f_{i,j}^{t,k}(\cdot)$ is nondecreasing for $i \neq j$ and $f_{i,j}^{t,k}$ is defined only for $k \leq t$. Furthermore, throughout the paper, we assume that $f_{i,j}^{t,k}$ is concave and differentiable for every $i \in R, j \in N/\{i\}$ and $t, k \in \{1, \dots, T\}$. The t -period payoff of player i is $g(x_i^t - f_i^t(\cdot)) = -(x_i^t - f_i^t(\cdot))^2$.¹⁰ Reciprocal players maximize the sum of these per period payoffs.

Note that in this specification the payoffs of reciprocal players depend only on the differences between realized and target contributions. We regard this a reduced-form representation: reciprocity functions provide a simple and tractable way to model players with social preferences. Our framework allows for the underlying motivation behind reciprocal preferences to come from many sources. One possibility is that reciprocal players are conditionally altruistic or exhibit conditional warm-glow effects (Andreoni, 1989). Another source can be a desire to follow social norms: a player with such considerations contributes more if she thinks others contribute more

¹⁰ Instead of the quadratic loss function specified above, we could consider any strictly quasiconcave function $g(\cdot)$ which attains its maximum at 0.

(or if she observes that others contributed a lot in past rounds). Yet another possibility is that reciprocal players care about fairness or equality, as in Fehr and Schmidt (1999) or Bolton and Ockenfels (2000).¹¹

Our model specification is also consistent with the consideration that reciprocal players care not only about how much others contribute, but also about the types, or preferences, of the other players. In particular, we allow reciprocity towards other players to be asymmetric. For example, reciprocity functions are allowed in our model to be more responsive to other reciprocal players' contributions than to contributions made by selfish players. Therefore, as long as intentions of a player depend only on his or her preferences, our model allows for reciprocity to depend on the intentions behind contributions.

The game is a standard extensive form game. Reciprocal players differ from the selfish players in that their payoffs are not simply the sum of monetary payments they receive. We assume that the game is of complete information. Players know how many selfish and how many reciprocal players there are in the group, and they know the reciprocity functions.¹² This corresponds to our intention to analyze play in games in which the players are experienced and become familiar with each other.

4. Theoretical results

In this section we impose some regularity conditions on reciprocity functions, and investigate the subgame perfect Nash equilibria of the game.

4.1. Assumptions on reciprocity

We impose the following five assumptions on reciprocity functions.

A1. Linear Reciprocity toward Reciprocal Players : $f_{i,j}^{t,k}(x_j^k) = \alpha_{i,j}^{t,k} x_j^k \quad \forall i \in R, j \in R/\{i\}$ and $t, k \in \{1, \dots, T\}$, where $k \leq t$.

A2. Nonincreasing Total Impact of Contributions over Time: Suppose $i \in R, j \in N/\{i\}, t \in \{1, \dots, T-1\}$, and x_j^1, \dots, x_j^{t+1} is such that $x_j^k \geq x_j^{k+1} \quad \forall k \in \{1, \dots, t\}$. Then $\sum_{k=1}^{t+1} f_{i,j}^{t+1,k}(x_j^k) \leq \sum_{k=1}^t f_{i,j}^{t,k}(x_j^k)$.

A3. Nonincreasing Marginal Impact of Contributions over Time: For any $x \in [0, 1], i \in R, j \in N/\{i\}, t < t', k \geq 0$, and $t' + k \leq T$, $\frac{\partial f_{i,j}^{t+k,t}(x)}{\partial x} \geq \frac{\partial f_{i,j}^{t+k,t'}(x)}{\partial x}$.

A4. Positive Initial Reciprocity: For every $i \in R, j \in N/\{i\}$, and $t \in \{1, \dots, T\}$, there exists $t' \in \{1, \dots, t\}$ such that $\left. \frac{\partial f_{i,j}^{t,t'}(x)}{\partial x} \right|_{x=0} > 0$.

A5. No Overreciprocation: $\sum_{k=1}^t \sum_{j \in N/\{i\}} \frac{\partial f_{i,j}^{t,k}(x_j^k)}{\partial x_j^k} \leq 1 \quad \forall t \in \{1, \dots, T\}, i \in R$, and $(x_j^k)_{j \in N/\{i\}}^{k=1, \dots, t} \in [0, 1]^{(N-1)t}$.

A benchmark case which satisfies these assumptions is when reciprocal players reciprocate the time average of others contribution. This example is analyzed in Appendix A1 guarantees that the impact of a marginal contribution by a selfish player on subsequent contributions by reciprocal players does not depend on the history of contributions. This both greatly simplifies the analysis and helps to avoid multiplicity of equilibria. A2 is a condition on the level of reciprocity: it states that a reciprocal player does not reciprocate in a

¹¹ A player who is concerned about fairness reciprocates others' contributions in a public good contribution game because a contribution increases others' flow payoff at the expense of one's own payoff.

¹² More precisely, all this information is common knowledge among players.

strictly increasing manner a nonincreasing sequence of contributions by any other player. This assumption implies that reciprocity toward a given (one-time) contribution weakly decreases over time. A3 is a statement on marginal reciprocity: it requires that the marginal impact of a contribution on reciprocity k periods later (e.g., a contribution at t on reciprocity at $t+k$, a contribution at $t+1$ on reciprocity at $t+k+1, \dots$) does not increase over time.¹³ A4 imposes that reciprocity is initially strictly positive toward every other player. A5 implies that at any period t , a unit increase in contributions by other players up until t increases the value of a reciprocity function by not more than a unit. This, besides being a natural requirement, is imposed in order to avoid multiplicity of equilibria resulting from reciprocal players either having optimistic expectations with respect to each others' contributions and contributing more, or having pessimistic expectations and contributing less.

The assumptions above are strong, but each is necessary to obtain the strong result we derive in the next section, namely that generically the resulting game has a unique subgame-perfect Nash equilibrium with a decreasing pattern of contributions. However, weaker results along the same line can be established even when dropping some of these assumptions. For example, since for any reciprocity functions it holds that selfish players do not contribute in the very last period, A2 and A3 can be substantially weakened to show that the equilibrium contribution pattern is ultimately decreasing (even if initially it might not be). Similarly, A5 can be dispensed with if the goal is to show the existence of some subgame-perfect Nash equilibrium implying a decreasing pattern of contributions (as opposed to the uniqueness of such equilibrium).

A key assumption in our model is A4, which is necessary to induce contributions from selfish players. That is, to explain positive contributions in a complete information setting, it is important that reciprocal players reciprocate contributions of a player, even when the motives of this player are known to be selfish. Note that the assumption allows for smaller amount of reciprocity towards selfish players than towards other reciprocal ones. The validity of the assumption is related to an ongoing discussion in the literature that to what extent reciprocity depends on outcomes (consequential reciprocity) vs. intentions (intentional reciprocity). Several papers, including Fischbacher et al. (2001) argue that people reciprocate intentions in public good contribution games and only reciprocate outcomes as long as they are signals of good intentions. McCabe et al. (2003) find supporting evidence for this, reporting that trustees in a trust game returned a greater amount of money if the truster could actually decide to trust or not, compared to a condition where the truster was forced to "trust" the trustee.¹⁴ However, existing evidence only establishes that intentions matter, and not that contributions without good intentions do not get reciprocated at all.¹⁵ In particular, in the experiments of McCabe et al. a positive amount of money is returned even in the forced trust condition. Our experiments provide further evidence: we find that selfish subjects start out contributing in the presence of at least one reciprocal player, even when the types of different group members are publicly revealed at the beginning of the game, and these contributions are reciprocated (see subsection 4.5).

4.2. Decreasing pattern of contributions in equilibrium

Our main result is that if the assumptions from the previous subsection hold and there is at least one selfish player in the game, then generically there is a unique subgame perfect Nash equilibrium, which exhibits a decreasing pattern of contributions.

Theorem 1. *If $S \geq 1$ and A1–A5 hold, then for generic A , the public good contribution game has a unique subgame perfect Nash equilibrium, and this equilibrium exhibits a weakly decreasing pattern of contributions. If reciprocity functions are strictly concave in selfish players' contributions, then the above statement holds for all A .*

For the formal proof of this result, as well as all formal statements in our paper, see Appendix A. The brief intuition is as follows. Concavity of the reciprocity functions, together with no overreciprocation, implies that the marginal impact of an extra unit of contribution on future contributions by reciprocal players is well-defined. Linearity in other reciprocal players' contributions implies that this impact is independent of contributions made in other periods or by other players. Then for generic values of A (if reciprocity functions are strictly concave in selfish players' contributions, then for all values of A), selfish players' contributions are uniquely determined in subgame perfect equilibrium at every period. Nonincreasing marginal impact of contributions over time then implies that the marginal return of contributions at earlier periods, when more periods are left to be played, is higher. This establishes that selfish players' contributions are weakly decreasing over time. Finally, nonincreasing total impact of contributions over time implies that the reciprocal players' contributions are also weakly decreasing over time.

4.3. Consistency with existing experimental findings

Our model is consistent with a series of results in the existing literature on public good contribution experiments. Here we summarize these implications of our model informally. For corresponding formal results, see Appendix A.

First, consider increasing the return of the public investment (A), which brings the individual return from contributing to the public good ($\frac{A}{N}$) closer to the return from private investment (1). It is well-documented in experimental settings that this increases players' contributions to the public good (see for example Isaac and Walker (1988), and Isaac et al. (1994)). This effect is captured by our model: increasing the marginal individual return from contributing increases contributions by the selfish players. Through positive reciprocity, this also increases contributions by the reciprocal players.

Another robust experimental finding is that increasing the number of periods in public good contribution games is shown to result in a longer period of positive contributions and in higher aggregate contribution levels (see Isaac et al. (1994)). This effect is also implied by our model. In a longer game, selfish players have more incentives to contribute, because there are more future periods in which reciprocal contributions are affected. In equilibrium, all players end up contributing more.

Our model is also consistent with the famous restart effect first shown in Andreoni (1988). If players treat the restarted game as a new game, then our model immediately explains the restart effect, since in this case it predicts first period contributions in the restarted game to jump back to the same level as in the first round of the game preceding the restart. Even if players do not treat the new session as a new game, but as an extension of the first game (and therefore aggregate contributions in the previous game become a history in a longer game), the model is compatible with the restart effect. If it is unexpectedly revealed that more periods are to be played than

¹³ We note that although the above assumptions on the time structure of reciprocity are sufficient to establish our main results, the reciprocity functions that are most appealing to us are ones in which reciprocity is in some sense constant over time. One way to formalize constant reciprocity over time is strengthening A2 by requiring that $\sum_{k=1}^{t+1} f_{ij}^{t+1,k}(x_j^k) = \sum_{k=1}^t f_{ij}^{t,k}(x_j^k)$ whenever $x_j^k = x_j^{k+1} \forall k \in \{1, \dots, t\}$ and $j \in N \setminus \{i\}$.

¹⁴ See also Blount (1995) for evidence that intentions matter in ultimatum games.

¹⁵ In fact, even that intentions matter is still not undisputed. Biele (2006) finds more evidence for consequential reciprocity than for intentional reciprocity in public good contribution games.

previously thought, selfish players have increased incentives to contribute. This unambiguously increases the contributions of selfish players.

Finally, in a one-shot game assumptions A1–A5 guarantee that if there is at least one selfish player then there is a unique Nash equilibrium, in which all players contribute zero. The result applies to restarted one-shot games as well. The intuition behind the result is simple. Since there is no continuation, all selfish players contribute zero. Then A4 and A5 imply that the only fixed point of the expectations of reciprocal players is when they expect zero contributions from each other. This result corresponds to the empirical finding that although initially subjects contribute significantly positive amounts in one-shot games in an experimental setting, contributions seem to go to zero with learning. In a setting in which after each round of play, subjects get randomly assigned to a new group (called the “strangers” treatment in the literature), contributions to the public good diminish over time (see Andreoni (1988), Croson (1996)).

4.4. Contributions by selfish vs. reciprocal players

Our model does not give a clear prediction for the relative magnitudes of contributions by a selfish and a reciprocal player. However, there is a general implication of the model for the timing of contributions of different players: every reciprocal player contributes positive amounts for at least as many periods as any selfish player. The reason is that selfish players have purely forward-looking considerations in contributing, while reciprocal players are partly backward-looking; hence selfish players' contributions tend to be relatively more concentrated on earlier periods than reciprocal players' contributions.

Theorem 2. Suppose assumptions A1–A5 hold and the game has a unique subgame perfect Nash equilibrium s . Then if $s_i^k > 0$ for some $i \in S$, then $s_j^k > 0 \forall j \in R$ and $k \in \{1, \dots, t\}$.

5. Experimental design

5.1. Hypotheses

Besides confronting our model with existing experimental evidence, we conduct a set of experiments to investigate some of the assumptions and predictions of the model.

Since the model assumes complete information, one goal of the experimental design is to allow subjects to have experience with both the game and their group members. In contrast to Andreoni (1988), where subjects are only surprised once, we adopt a probabilistic restart where after 10 periods, subjects are assigned to new groups with high probability. With the remaining probability, subjects remain in the same group and play for an additional 10 periods. To allow subjects to obtain experience, we examine multiple sets of these 10-period games. These design features allow us to investigate experienced play after a restart. We regard play in later sets of 10-period games in restarted groups as an approximation of the complete information assumption.

To allow subjects to learn about their group members, in another treatment, subjects first participate in gift-exchange games. Afterwards, they participate in public good contribution games. Once subjects have had an opportunity to obtain experience with the public good contribution game, we assign them to new groups and show them the histories of their group members' play as responders in the gift-exchange games. Subjects may use this information to form an assessment of the type of their opponents.

In games with experienced players, we investigate three hypotheses. The first is that positive contributions (and other robust features of contribution patterns) can be modeled as an equilibrium phenomenon. Since equilibrium implies that players' expectations are fulfilled, we

Table 1
Treatments.

1. Experienced Restart	2. Experienced Restart with Forecast
Six sets of 10-period public good contribution games with groups reshuffled with known probability 0.75 after each set	Same as 1. except at start of each set of 10-period games, solicit forecast of path of contributions
3. Identifying types	
Six gift exchange games (play as first and second proposer equal number of times)	
Two sets of 10-period public good contribution games, group composition changes randomly after first set	
Construct groups based on gift exchange play; inform subjects of gift exchange play of fellow group members	
One 10-period game	
Construct groups based on gift exchange play; inform subjects of gift exchange play of fellow group members	
One 10-period game	

examine how accurate experienced players' forecasts are about the average contribution patterns of their opponents before a 10-period game. Second, to explain positive contributions, our model requires that even completely selfish players have incentives to contribute at the beginning of the game. Therefore, we test how contributions in the first period of a game affect contributions of others in the next period. Finally, we test a direct prediction of our model, corresponding to Theorem 2, that selfish players stop contributing in an earlier period than reciprocal players. We use the treatment in which subjects first play gift-exchange games to classify players as selfish or reciprocal.

The three hypotheses we formally test are:

H1. Experienced players' forecasts are close to actual play in restarted 10 times repeated games.

H2. With experienced players, contributions positively affect other players' subsequent contributions.

H3. With experienced players, selfish players stop contributing earlier than reciprocal players.

5.2. Treatments

Table 1 provides a summary of the three treatments. In each treatment, subjects participate in a public good game in groups of four for ten periods with the same group members. The stage game shares features with other experiments examining repeated public good contribution games. Each player receives an endowment of 20 tokens and must simultaneously decide how many tokens to contribute to the public project. After subjects make their contributions, they are informed of each group member's contribution to the public project and their income from the stage game.¹⁶ The income of a player is the amount that is not contributed to the public project plus 1.6 times the average contribution of the group. That is, the stage game payoff for each subject is given by:

$$\pi_i = 20 - g_i + 0.4 \sum_{j=1}^4 g_j,$$

where g_i is the contribution of subject i . In a “10-period game,” the stage game is repeated 10 times with the same group of four players. We refer to the first set of 10-period games as games 1–10, the second

¹⁶ We also conduct versions of the treatments in which subjects are told the total contributions by the group rather than individual contributions after each period. The rationale for conducting these sessions is related to the model presented in working paper version of the paper, which had this information environment. The Supplementary Appendix describes the results from these experiments, which are similar to those reported here.

set as games 11–20, and so on.¹⁷ At the end of the session, tokens were converted to dollar amounts.

5.2.1. Experienced Restart

This treatment consists of six sets of 10-period public good games where after the tenth period, subjects are reshuffled with probability 0.75. If a subject is reshuffled, then she is randomly assigned to a new group. If a subject is not reshuffled, then she stays in the same group. A group is not allowed to stay together for more than two consecutive sets of 10-period games. Subjects are informed of how reshuffling takes place and the probability of being reshuffled. In a typical session with 32 subjects, on average 6 out of the 8 10-period games are reshuffled.

5.2.2. Experienced Restart with Forecast

The only difference with this treatment and the Experienced Restart treatment is that at the beginning of every set of 10-period games, subjects are asked to forecast the average contribution of their three opponents in the first, fifth, and tenth period. In this treatment, we do not pay an additional amount to subjects based on their forecasts to avoid interfering with the incentives provided by the public good game and possible interactions with contributions. Other experiments that elicit beliefs in linear public goods games incentivize the revelation of beliefs, but usually with small stakes to mitigate this concern (see e.g., Fischbacher and Gächter, 2010). Since conducting our experiments, Gächter and Renner (2010) report new evidence on the impact of incentivizing beliefs in repeated public good contribution games. They report that incentivizing beliefs increases their accuracy, but the average level of beliefs is not affected. On the other hand, if beliefs are incentivized, they also lead to higher contributions, in particular in the later half of a 10-period game. Since we are interested in whether players correctly anticipate the pattern of contributions, we expect that our estimate of the accuracy of forecasts is a lower bound relative to an incentivized forecast measure.

5.2.3. Identifying types

In this treatment, we use a gift-exchange game to identify player types and then sort subjects into groups for the 10-period repeated public good game. In Part I, subjects are randomly paired for each of six gift-exchange games. Each subject takes the role of first proposer and second proposer three times and no subjects are ever paired with the same opponent more than once. The first proposer starts the game by deciding what fraction of her endowment of 10 tokens to give to the second proposer. This amount is doubled and given to the second proposer. The remaining portion of the first proposer's endowment is retained and is not doubled. Following this first offer, the second proposer responds by deciding how much of her endowment of 10 tokens to give to the first proposer, who receives double this amount. The remaining portion is kept by the second proposer and is not doubled.

Play in the gift-exchange games is used to identify the degree of reciprocity in players. We construct a measure of reciprocity based on how subjects respond to positive proposals when playing as the second proposer. The index of reciprocity is the ratio of a subject's response to the first proposer's offer when the first proposer's offer was nonzero. The index is the average across up to three observations of this ratio per subject. Under this metric, a subject is more reciprocal when the reciprocity ratio is higher. We label the eight subjects with the highest reciprocity ratios as "reciprocal players" and label the eight subjects with the lowest reciprocity ratios as "selfish players." We label the rest of the players in the experiment "unidentified."

¹⁷ In each session, all subjects answered control questions which intended to verify their understanding of the instructions. The actual instructions and control questions are available from the authors upon request.

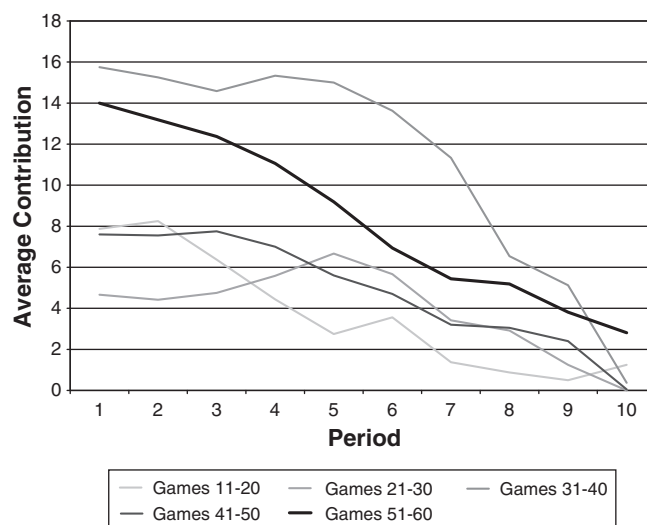


Fig. 1. Average contribution in restarted games for each set of 10-period games in experienced restart and experienced restart with forecast sessions.

Subjects are not informed of the second part of the session until after completion of the first part. In Part II, subjects play two sets of 10-period public good games in groups of four, where groups are randomly reshuffled, to allow them to obtain experience. Then, we assign players to groups with three reciprocal players and one selfish player and groups with one reciprocal player and three selfish players.¹⁸ Before the start of the 10-period game with these groups, subjects are informed of the histories of all their opponents in gift-exchange games when they play as second proposers. After the tenth period, subjects are re-assigned into different groups of three reciprocal players and one selfish player and groups with one reciprocal player and three selfish players. Before playing another 10-period game, subjects are again shown the histories of play for their opponents, in gift-exchange games when they are second proposers. In this treatment, there are two sets of observations for 10-period games with a measure of player type in the group.

5.3. Experimental results

All lab sessions were conducted with zTree (Fischbacher, 2007) at the Computer Lab for Experimental Research at Harvard Business School. Subjects were recruited from the greater Boston area, with a large fraction of university students. No subject was allowed to participate in multiple sessions. Each session lasted approximately 1.5 h and involved either 32 and 36 subjects. The average earnings across sessions were between \$21 and \$23. Here we utilize data from one session of Experienced Restart, one session of Experienced Restart with Forecast, and two sessions of Identifying Types.¹⁹ In total, 140 subjects participated in the sessions. Additional details of the sessions are in Table A1, in the Supplementary Appendix A.

5.3.1. Main features of contribution paths

Before presenting the results from testing the hypotheses, we describe the basic features of contribution paths in the games with

¹⁸ The unidentified players play among each other in the remaining groups.

¹⁹ In one of the Identifying Types sessions, after each public good stage game, we display the contributions of each of the individuals in the group, while in the other we display the group's contribution. The results are reported separately for each session in the Supplementary Appendix. The reason we pool observations from Identifying Types (IT) sessions, but not from the Experienced Restart (ER) and Experienced Restart with Forecast (ERF) sessions where only group contributions are revealed, is that the justification for pooling is weaker for those sessions. There is evidence, detailed in the Supplementary Appendix, that revealing only group contributions led to a different distribution of contributions in the ER and ERF sessions, but not in the IT session.

Table 2
Contribution path in restarted games.

Period	Experience Restart		Experienced Restart w/Forecast	
	Mean	Median	Mean	Median
<i>A: Last two sets of 10-period games (games 41–60)</i>				
1	10.94	10.00	10.05	8.50
2	11.81	11.00	8.65	5.00
3	11.38	10.50	8.55	8.00
4	10.25	8.50	7.65	3.50
5	8.63	7.50	6.05	1.00
6	5.88	4.00	5.55	0.50
7	2.81	0.00	5.30	0.00
8	1.63	0.00	5.90	0.50
9	1.25	0.00	4.45	0.00
10	0.00	0.00	2.30	0.00
N	16		20	
<i>B: Intermediate sets of 10-period games (games 21–40)</i>				
1	13.88	16.00	8.42	3.00
2	13.92	15.50	7.08	3.50
3	13.63	15.50	6.67	4.00
4	14.79	17.00	6.67	3.50
5	14.92	19.00	6.83	3.00
6	13.75	20.00	5.42	1.00
7	10.83	12.50	4.42	1.50
8	5.54	0.00	4.92	5.00
9	3.17	0.00	5.17	2.00
10	0.08	0.00	0.58	0.00
N	24		12	

Notes. *N* is the number of individual observations.

experienced players. The short summary is that we find exactly the same features with experienced players as in other studies: contributions are significantly different from zero at the beginning, contributions decrease over time, and there is a significant restart effect.

Fig. 1 plots the average contribution paths in restarted games for the Experienced Restart and Experienced Restart with Forecast sessions.²⁰ Darker lines in the figure refer to sets of 10-period games later in the sessions, so the darkest line corresponds to Games 51–60, the last set of restarted 10-period games. The figure shows that the path of contributions displays an overall downward trend, with the average contribution in the first period always much larger than the last period. Table 2 reports the mean and median contribution for each period in a 10-period restarted game for these two sessions. Panel A reports the distribution of the average time path of individual contributions from the last two sets of 10-period games. In the Experienced Restart session, there are 4 restarted groups (so the number of individual observations, *N*, is 16), playing a restarted game among Games 41–50 and Games 51–60, while in the Experienced Restart with Forecast treatment, there are 5 restarted groups (*N*=20). Players at this stage of a session have experienced at least 4 sets of 10-period games. For both sessions, the mean individual contribution in the first period is over half the endowment, while in the tenth period the mean individual contribution in the Experienced Restart session is 0 and in the Experienced Restart with Forecast session it is 2.3. Across both sessions, in the first period, 16% of subjects contribute zero, while in the tenth period 88% contribute zero.

In Panel B, we report mean and median individual contributions for subjects who have less experience than in Panel A. These observations are from the two sets of 10-period games prior to the last two sets of 10-period games. Subjects at this stage of a session

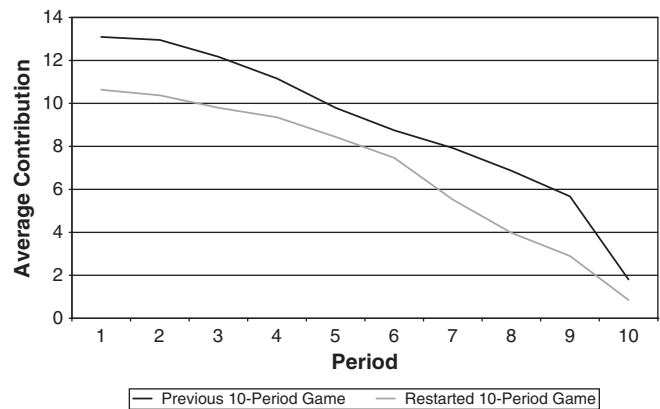


Fig. 2. Average contribution pattern within restarted groups in experienced restart and experienced restart with forecast sessions.

have participated in at least 2 sets of 10-period games. For both the sessions, the pattern is similar to the pattern for players with relatively more experience in Panel A: there is a downward path of contributions and the majority of individual contributions are zero in the last period of the restarted game.

Next, we examine the restart effect. Fig. 2 plots the average contribution for a group that is not reshuffled before the restart and the ten periods after the restart. The figure shows that there is a restart effect, as the average contribution in the last period of the previous 10-period game is much smaller than the average contribution in the first period of the restarted 10-period game. To verify the presence of a restart effect in our probabilistic continuation design, in Table 3, we report the average contribution immediately before and after restarts. That is, we compare the last period in the set of 10-period games with the first period in the restarted 10-period game.

Table 3 presents two types of comparisons: Panel A reports the last two sets of 10-period games (Games 40 vs. 41 and Games 50 vs. 51), while Panel B reports the intermediate set of 10-period games (Games 20 vs. 21 and Games 30 vs. 31). For instance, in the Experienced Restart session, all players in the tenth period prior to being restarted contribute 0, while in the first period of the restart with the same group, the mean individual contribution is 10.94 and only 19% of subjects contribute zero. Both the individual contribution level and number of players who contribute zero are statistically different, based on a Wilcoxon two-sided test ($p < 0.01$).²¹ The next two rows of the table report similar comparisons for the Experienced Restart with Forecast session. Here, 60% of subjects contribute 0 in the tenth period before the restart, while only 15% contribute 0 in the first restarted game. Both the individual contribution level and the number of players who contribute zero are statistically different ($p < 0.01$).

In Panel B of Table 3, we report the similar comparisons using data from subjects who are not as experienced as in Panel A. For the Experienced Restart session, the individual contribution amount and number who contribute zero are statistically different in the first restarted game ($p < 0.01$). For the Experienced Restart with Forecast session, the individual contribution levels are statistically different ($p = 0.03$), as are the fraction who contributed zero ($p = 0.03$). Note that for this comparison, there are 3 restarted groups and hence there are only 12 individual observations, so the comparisons are noisier.

The significance of the restart is also apparent comparing the average group contributions. To examine this formally, we pool the data from the two sessions presented in Panel A and run a regression of group contribution on indicators for the session and the first period. We find that the group average contribution in the first restarted

²⁰ In Fig. 1, we pool observations from restarted games played in both sessions because they follow the same qualitative pattern. When we formally test whether the distribution of individual contributions among experienced players in the first game and last game of the 10-period game are from the same distribution using an *F*-test, we find no difference for restarted Games 41–50 and Games 51–60 ($p = 0.51$).

²¹ All tests reported here, unless otherwise noted, are two-sided Wilcoxon rank sum tests.

Table 3
Restart effect in 10 period games across sessions.

Session	Games	N	Mean	Median	Fraction who contribute zero
<i>A: Last two sets of 10-period games</i>					
Experienced	Last before Restart (40,50)	16	0.00	0.00	1.00
Restart	First after Restart (41,51)	16	10.94	10.00	0.19
Experienced	Last before Restart (40,50)	20	5.60	0.00	0.60
Restart w/ Forecast	First after Restart (41,51)	20	10.05	8.50	0.15
<i>B: Intermediate sets of 10-period games</i>					
Experienced	Last before Restart (20,30)	24	0.75	0.00	0.83
Restart	First after Restart (21,31)	24	13.88	16.00	0.08
Experienced	Last before Restart (20,30)	12	1.67	0.00	0.92
Restart w/ Forecast	First after Restart (21,31)	12	8.42	3.00	0.42

Notes. *N* is the number of individual observations.

game is 7.33 higher than in the previous game (T -stat = 3.97). Likewise, when we pool the data presented in Panel B, we find that the group average contribution in the first restarted game is 11.00 higher than in the previous game (T -stat = 5.35).²²

In summary, Tables 2 and 3 as well as Figs. 1 and 2 demonstrate that there is a declining average contribution path even for experienced players in restarted games and that the restart effect is statistically significant for experienced players.

Next, we consider summary statistics from the Identifying Types sessions. From the gift-exchange game, the average index of reciprocity for the 8 players we classify as selfish is 0, while for the 8 players we classify as reciprocal it is 1.07 (median = 0.94). This means that a selfish player always returns 0 for each positive first proposer amount, while a reciprocal player returns about the same amount. Using these player types, we form 4 groups for each of the two 10-period games. That is, we have observations of 8 groups of 3 reciprocal players and 1 selfish player and 8 groups of 1 reciprocal player and 3 selfish players, for a total of 640 individual observations of contributions. The groups with three selfish players exhibit little cooperation relative to groups with three reciprocal players. The average contribution per player for the groups with 3 reciprocal players and 1 selfish player is 9.78, while the average contribution per player for the groups with 1 reciprocal player and 3 selfish players is 4.87 ($p < 0.01$). In the former group type, in each of the 8 groups, the last positive contribution is made by a reciprocal player. In the latter group type with 3 selfish players, in 3 out of 8 groups, the last positive contribution is made by a reciprocal player even though there are 3 times as many selfish players. Across groups, the average selfish player contributes 5.41, while the average reciprocal player contributes 9.24 ($p < 0.01$). In selfish groups, the average selfish player contributes 4.83, while the average reciprocal player contributes 5.01 ($p = 0.57$). In reciprocal groups, the average selfish player contributes 7.18, while the average reciprocal player contributes 10.65 ($p < 0.01$).

Fig. 3 shows the time path of average contributions for the Identifying Types sessions for the two sets of 10-period games after the gift exchange game histories are made public to group members. The lines report the average contribution for different group

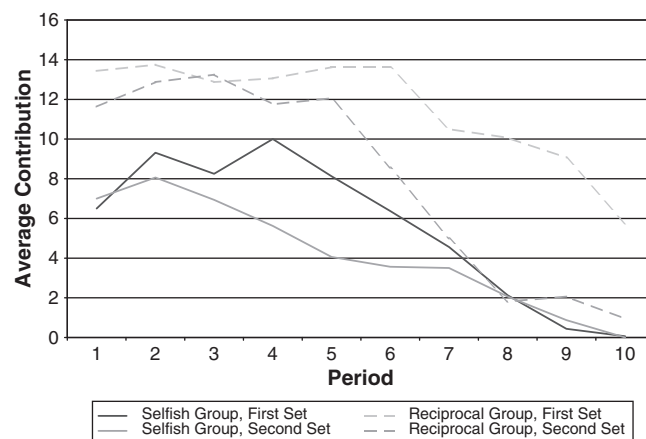


Fig. 3. Average contribution in identifying types treatment by group type.

compositions for each set of 10-period games. A selfish group in the figure is one with 3 selfish players and 1 reciprocal player, while a reciprocal group is one with 3 reciprocal players and 1 selfish player. For all groups, the average group contribution in the first game is much larger than the last game.²³ The figure also shows that the average group contribution in groups with 3 reciprocal players is larger than the average group contribution in groups with only 1 reciprocal player.

The summary statistics for the Identifying Types session and Fig. 3 suggest that heterogeneity in player types is an important consideration for understanding the dynamics of the average contribution path in public good contribution games.

5.3.2. Testing the hypotheses

Using data from the Experienced Restart with Forecast session, Table 4 examines how closely forecasts match the average contributions of the other group members. The table reports data on restarted games with subjects who have played at least 4 sets of 10-period games. In these games, subjects forecast the average contribution will be 10.87 in the first period and on average, the actual average contribution of the three opponents is 10.44. For the fifth period, the average forecast is 7.06 and the average actual contribution is 4.82. For the tenth period, the average forecast is 2.19 and the average actual contribution is 1.63. Note that for the tenth period, the averages are skewed by the presence of one subject who contributes his entire endowment and forecasts that the average contribution is 20 for the tenth period. The difference between the forecast and actual contributions for the median subject is zero, and half of the subject's forecasts are within one unit of the average contribution of their opponents. Moreover, 11 out of 16 subjects forecast no contribution in the tenth period. However, each of these subjects contribute a positive amount in the first period, contributing more than half their endowment on average.

It is also interesting to disaggregate forecasts based on subjects' contribution in the first period. Subjects who contribute more in the first period forecast that contributions to be greater in all periods. For instance, if a subject contributes more than 10 tokens in the first period, on average she forecasts 15 for the first period and 4 for the tenth period. On the other hand, if a subject contributes less than 10 tokens in the first period, she forecasts 8.4 tokens in the first period and 1.3 tokens in the tenth period. Both initially high contributors and low contributors anticipate that there is a decreasing path of contributions. The findings together indicate that a large fraction of subjects correctly anticipate at the beginning of the game that

²² We note that we find a significant restart effect even though our experimental design implies that we potentially underestimate the restart effect of a genuinely surprise restart. This is because in the last period of a newly started game our subjects know that the game continues with some probability (even though we specified a low probability in order to mitigate this effect).

²³ Although not shown, when we plot the average contribution for the groups with unidentified players, we also find a downward average contribution path and restart effect.

Table 4
Forecasts and actual play in restarted games.

Games	N	Forecasted play		Actual play	
		Mean	Median	Mean	Median
<i>First period and forecast</i>					
41,51	16	10.87	10.00	10.44	10.83
<i>Fifth period and forecast</i>					
41,51	16	7.06	5.00	4.82	4.17
<i>Tenth period and forecast</i>					
41,51	16	2.19	0.00	1.63	0.33

Notes. N is the number of individual observations.

although initial contributions in the game will be high, contributions at the end of the game will be close to zero. In short, subjects foresee the declining pattern of contributions.²⁴

To measure forecast accuracy, we regress contribution forecasts on average contribution of opponents using data from the last two sets of restarted games. For the first game, the coefficient on average contribution of opponents is 0.92, with *T*-statistic of 5.06 and $R^2 = 0.63$. This suggests that forecasts and actual contributions are highly correlated. When we regress forecasts on actual contributions for game five, the coefficient is 1.05 (*T*-stat = 4.12, $R^2 = 0.53$). In the tenth game, 11 out of 16 players forecast no contribution, and 6 subjects are paired with opponents whose average contribution is zero. The large fraction of zeros limits our ability to estimate a precise relationship, even though many correctly anticipate little or no contribution. These regressions show that there is no large systematic error in forecasts for experienced players in restarted games, which leads to our first conclusion:

5.3.2.1. Conclusion 1. Experienced players on average correctly anticipate the pattern of contributions in a 10-period game. In particular, they foresee that contributions will be close to zero by the tenth period.

Turning to the second hypothesis, Table 5 considers a measure of the importance of the strategic incentive to contribute. For the 10-period public good games in the Experienced Restart and Experienced Restart with Forecast session, the table reports estimates of the correlation of the average first period contribution on subsequent play. In the three columns, we regress an individual's contribution in the second period on the average contribution in the first period. The table reports the estimated coefficient and *T*-statistic in brackets of the impact of the average first period contribution. Since we examine only the impact of the first period contribution, for new groups, this corresponds to the impact of a randomly assigned group average contribution on a subject's second period contribution. In this analysis, two members of a given group may share portions of the randomly assigned group average because for they share two out of three of their opponents. This introduces potential dependencies for members of a given group, which motivate various estimation approaches in Table 5. Each row corresponds to a different specification: OLS is simply an ordinary least squares regression; group fixed effects include a dummy variable for each possible group; and lastly we report estimates from a random-effects tobit model. Group fixed effects are indicators for each particular group across sessions, so their inclusion precludes including session level controls.

For the OLS specification, we report test statistics based on three versions of standard errors. The first version is based on the

²⁴ It is also interesting to examine the interaction between beliefs and contributions in the dynamics of play, as in Fischbacher and Gächter (2010). In the Supplementary Appendix Table A6, we investigate how subsequent contributions react to forecasting errors. The analysis suggests that if a subject underforecasts her opponents contribution by 1 unit in the first period, then in the second period she contributes about 0.3 more tokens.

Table 5
Impact of opponents' first period average contribution on 2nd period contribution.

Dependent variable	Contributions in 2nd game		
	Pooled (1)	Not restarted (2)	Restarted (3)
Controls			
OLS	0.49 [6.45]	0.31 [3.18]	0.73 [6.00]
Eicker-White robust	[6.05]	[2.89]	[6.43]
Two-way cluster on group and individual	[4.52]	[2.85]	[6.55]
Group fixed effect with individual cluster	0.43 [4.86]	0.22 [1.88]	0.60 [4.51]
Random effects tobit	0.26 [4.00]	0.13 [1.75]	0.71 [5.79]
N	408	320	88

Notes. Table reports estimates of a player's contribution in the second period of a 10-period game on the average contribution of the other group members in the first period. *T*-statistics are in brackets under estimated coefficients. The OLS specification reports three versions of *T*-statistics: the first is the conventional estimate (assuming homoskedastic errors), the second is from Eicker-White robust standard errors, and the third is from two-way clustering on individual and the particular group, as in Cameron et al. (2011). The next estimate reports group fixed effects, with one way individual clustering. The final estimate is from a random effects tobit model, where the effect accounts for multiple observations for the same individual (implemented using STATA xtobit). Pooled includes both not restarted and restarted games. All 10-period games in the Experienced Restart and Experienced Restart with Forecast sessions are included. N is the number of individual observations.

conventional estimate, the second is based on Eicker-White robust standard errors, and the third allows for two-way clustering on individual and the particular group, following Cameron et al. (2011). The next point estimate presented in the Table is from a model with a group specific fixed effect and individual level clustering. Finally, we report estimates from a random effects tobit model, where the effects account for multiple observations for the same individual. Across these methods, we find a precisely estimated effect of first period average opponent contribution on a subject's second period individual contribution in column (1), ranging from 0.26 to 0.49. The magnitude of the effect is larger in restarted games, where using the estimate in the first cell in column (3), a unit increase in the first period contribution of a player increases others' contributions in the next period by 0.73 unit. For each of the models we estimate we find that the coefficient is larger in restarted games than not restarted games.

The positive and significant coefficient suggests the existence of conditional reciprocity even when players are experienced. The amount of responsiveness we find in second period contributions is not enough by itself to induce selfish players to contribute in the first period. However, presumably first-period contributions have an effect on contributions in periods after the second one, too, increasing the selfish players' incentives to contribute at the beginning. We do not estimate the latter effects from our data, because players' contributions are endogenous in all previous contributions of others and therefore in their own contributions up until two periods preceding the current period.²⁵

5.3.2.2. Conclusion 2. For experienced players, contributions in the first period positively affect contributions in the subsequent period.

Lastly, we examine the period in which selfish and reciprocal players stop contributing positive amounts. We have already mentioned that in the groups with a majority of reciprocal players, the last positive contribution is always by a reciprocal player. We can formally test Theorem 2 by comparing the last period in which selfish and reciprocal players give positive contributions. If a player never

²⁵ In restarted games, the correlation coefficient between the total number of units a player contributes in periods 2–10 and the average number of units the others contribute in the first game is 6.37.

gives a positive contribution, then the period where she last gives a positive contribution is defined to be 0. We find that the average reciprocal player stops contributing in period 7.41, while the selfish player stops contributing in period 5.19. The difference is statistically significant ($p < 0.01$) and supports the third hypothesis.

5.3.2.3. *Conclusion 3.* In 10-period games played by experienced players, selfish players stop contributing in an earlier period than reciprocal players.

6. Conclusion

This paper shows that all documented findings from finitely repeated public good contribution games can be captured by a model in which there are both selfish and conditionally reciprocal players, even in the absence of asymmetric information. We provide conditions on the preferences of reciprocal players for this to be the case and show that in the resulting equilibrium selfish players induce subsequent contributions from reciprocal players by contributing substantial amounts at the beginning of the interaction. Our paper therefore provides a rationale for the presence of both selfish and reciprocal behavior in a population, which is consistent with experimental findings. Namely, if reciprocal types reciprocate past contributions imperfectly, then forward-looking selfish types can induce reciprocal types to contribute more throughout the game than other reciprocal types.

The theoretical analysis in this paper provides an investigation into the dynamic incentives in the presence of conditional cooperators. Some implications of conditional cooperation for public policy and organization of the management are presented in Gächter (2007). Some areas for future work involve examining the extent to which a complete information model with heterogenous players can explain cooperation in other environments where cooperation emerges in finite horizons, such as in centipede games (McKelvey and Palfrey, 1992) and in situations outside of the laboratory (see e.g., Bandiera et al., 2005). Another direction involves further attempts to measure the evolution of beliefs and to more precisely characterize the nature of player heterogeneity (for interesting work in this direction, see Fischbacher and Gächter, 2010).

Acknowledgements

We would like to thank Satoru Takahashi for a careful reading of the paper, and Colin Camerer, Ernst Fehr, Drew Fudenberg, Thomas König, Imran Rasul, Al Roth, and participants of the SITE meeting on Experimental Economics for the useful comments. We would also like to thank Jim Andreoni and two anonymous referees for helping us improve the paper. The experiment was programmed and conducted with z-Tree; we are grateful to Urs Fischbacher for making this software available. Pablo Guillen and the staff of the Computer Lab for Experimental Research at Harvard Business School were helpful in conducting the laboratory sessions. For financial support, Pathak thanks the National Science Foundation and the Division of Research at Harvard Business School.

Appendix A

This appendix has three sections: 1) an example of the model with time average reciprocation, 2) proofs of the theorems stated in the main text, and 3) statements and proofs of the results described in Section 4.3 of the main text.

A.1. Example: Time average reciprocation

To provide intuition for the basic features of strategic interaction between selfish and reciprocal players, we present an example to

illustrate why there are decreasing contributions in equilibrium. In this specification reciprocal players reciprocate the simple time average of others' contributions²⁶:

$$f_i^t(\cdot) = \frac{1}{t(N-1)} \left(\sum_{k=1}^t \sum_{j \in N/\{i\}} x_j^k \right) \quad \forall i \in R \text{ and } t \in \{1, \dots, T\}.$$

It is convenient to define the following terms:

$$C(t) = \frac{A}{N} \sum_{k=t+1}^T c(t, k),$$

where

$$c(t, t+1) = \frac{N-S}{S+t(N-1)}$$

$$c(t, k) = \frac{N-S}{S+(k-1)(N-1)} \left[1 + \sum_{l=t+1}^{k-1} \left(\frac{N-S-1}{N-S} \right) c(t, l) \right] \text{ for } k > t+1.$$

The interpretation of these terms is the following: $C(t)$ is the marginal impact of a contribution at t on future payoffs by reciprocal players. This includes both the direct impact of a unit contribution for the future periods and the indirect effect of a unit contribution in subsequent periods on future periods. In period $t+1$, the impact is $c(t, t+1)$, while for periods $k > t+1$ it is $c(t, k)$.

Claim 1. Suppose $C(t) \neq 1 - \frac{A}{N} \forall t \in \{1, \dots, T\}$ and $S \geq 1$. Then the above game has a unique subgame perfect Nash equilibrium, which exhibits the following contribution path. If $C(1) < 1 - \frac{A}{N}$, then every player contributes 0 at every period. Otherwise, let T^* be the largest integer between 1 and $T-1$ such that $C(T^*) > 1 - \frac{A}{N}$. Then every selfish player contributes 1 in periods 1 to T^* and 0 afterwards. Meanwhile, reciprocal players contribute 1 until period T^* , and their contributions are strictly decreasing afterwards.

Proof of Claim 1. We begin with the following lemma.

Lemma 1. Let $t \in \{0, \dots, T-1\}$, $(x_i^1)_{i \in N}, \dots, (x_i^t)_{i \in N}$ be a length- t contribution history and $(x_i^{t+1})_{i \in S}, \dots, (x_i^T)_{i \in S}$ be a sequence of contributions by the selfish players after period t . Assume A1, A4 and A5 hold. Then there is a unique sequence of contributions by the reciprocal players after period t , $(x_i^{t+1})_{i \in R}, \dots, (x_i^T)_{i \in R}$ such that $x_i^k = f_i^k((x_j^l)_{j \in N/\{i\}}, \dots, (x_j^T)_{j \in N/\{i\}})$ $\forall i \in R$ and $k \in \{t+1, \dots, T\}$, and $x_R^k = (I - M^k)^{-1} \bar{x}_R^k$, where $\bar{x}_i^k = \sum_{j=1}^{k-1} f_{i,j}^{k,k'}(x_j^{k'}) + \sum_{j \in S} f_{i,j}^{k,k}(x_j^k) \quad \forall k \in \{t+1, \dots, T\}$ and $i \in R$, $\bar{x}_R^k = (\bar{x}_i^k)_{i \in R}$, and M^k is the $(N-S) \times (N-S)$ matrix whose diagonal elements are 0 and its (m, n) -th element is $\alpha_{S+m, S+n}^{k,k}$.

Proof of Lemma 1. Assume that $k \in \{t+1, \dots, T\}$ is such that after any length- k history $(x_i^1)_{i \in N}, \dots, (x_i^k)_{i \in N}$ and any sequence of period $k+1$ to period T contributions $(x_i^{k+1})_{i \in S}, \dots, (x_i^T)_{i \in S}$ by the selfish players there is a unique sequence of period $k+1$ to period T contributions $(x_i^{k+1})_{i \in S}, \dots, (x_i^T)_{i \in S}$ by the reciprocal players such that

$$x_i^l = f_i^l \left((x_j^1)_{j \in N/\{i\}}, \dots, (x_j^l)_{j \in N/\{i\}} \right)$$

$\forall i \in R$ and $l \in \{k+1, \dots, T\}$. Note that the above trivially holds for $k=T$.

²⁶ It is easy to verify that these reciprocity functions satisfy the assumptions made in Section 4.1.

Consider now any length length- $(k-1)$ history $(x_i^1)_{i \in N, \dots}, (x_i^{k-1})_{i \in N}$ and any sequence of period k to period T contributions $(x_i^k)_{i \in S}, \dots, (x_i^T)_{i \in S}$ by the selfish players. By definition, if for some $(y_i^k)_{i \in R}$ it holds that

$$y_i^k = f_i^k \left((x_j^1)_{j \in N \setminus \{i\}}, \dots, (x_j^{k-1})_{j \in N \setminus \{i\}}, \left((x_j^k)_{j \in S}, (y_j^k)_{j \in R} \right) \right)$$

$\forall i \in R$ then $x_i^k = \hat{x}_R^k + M^k x_i^k$. Assumptions A4, A5 and $S > 1$ imply that $\sum_{j \in R \setminus \{i\}} \alpha_{i,j}^{k,k} < 1 \forall i \in R$. Then by a well-known theorem (see Takayama, 1985, p. 381) $I - M^k$ is invertible and therefore the solution to $x_i^k = \hat{x}_R^k + M^k x_i^k$ is unique and satisfies $x_R^k = (I - M^k)^{-1} \hat{x}_R$. The claim then follows by induction. \square

First note that $\frac{A}{N} < 1$ implies that in any subgame perfect Nash equilibrium all selfish players contribute 0 after any $(T-1)$ -length history $(x_i^1)_{i \in N}, \dots, (x_i^{T-1})_{i \in N}$. Furthermore, in any subgame perfect Nash equilibrium $x_i^T = E_i f_i^T((x_i^1)_{i \in N}, \dots, (x_i^{T-1})_{i \in N}, (x_j^T)_{j \in N \setminus \{i\}}) \forall i \in R$, where the expectation is taken with respect to player i 's beliefs concerning $(x_j^T)_{j \in N \setminus \{i\}}$ after history $(x_i^1)_{i \in N}, \dots, (x_i^{T-1})_{i \in N}$. Lemma 1 then implies that there is a unique continuation strategy profile $(x_j^T)_{j \in N}$ after $(x_i^1)_{i \in N}, \dots, (x_i^{T-1})_{i \in N}$ in subgame perfect Nash equilibrium, and in this continuation profile

$$x_i^T = \frac{N-S}{S+(T-1)(N-1)} \sum_{t'=1}^{T-1} \sum_{j \in S} x_j^{t'} + \sum_{t'=1}^{T-1} \sum_{j \in R \setminus \{i\}} \frac{NT-T}{S-N+2T+NT+ST+NST-T^2-N^2T+N^2T^2+1} x_j^{t'} + \sum_{t'=1}^{T-1} \frac{N-1-S}{S-N+2T+NT+ST+NST-T^2-N^2T+N^2T^2+1} x_i^{t'}$$

$\forall i \in R$. Note that only the first term depends on selfish players' contributions.

Let $t \in \{1, \dots, T-1\}$ and assume that for every $k \in \{t, \dots, T-1\}$ and for every length- k history $(x_i^1)_{i \in N}, \dots, (x_i^k)_{i \in N}$ all subgame perfect Nash equilibria specify the same continuation profile, which is history-independent for selfish players and satisfies:

$$x_i^{k+1} = \frac{N-S}{S+k(N-1)} \sum_{t'=1}^k \sum_{j \in S} x_j^{t'} + \sum_{t'=1}^k \sum_{j \in R \setminus \{i\}} \frac{(N-1)(k+1)}{S-N+1+(k+1)(2+N+S+NS-N^2)+(k+1)^2(N^2-1)} x_j^{t'} + \sum_{t'=1}^k \frac{N-1-S}{S-N+1+(k+1)(2+N+S+NS-N^2)+(k+1)^2(N^2-1)} x_i^{t'}$$

for every $i \in R$. Then a marginal contribution by any $i \in S$ at t has zero impact on contributions of any $j \in S$, and its marginal impact on the total future contributions of any $j \in R$ is $C(t)$. By assumption $C(t) \neq 1 - \frac{A}{N}$. Then independently of history, in any subgame perfect Nash equilibrium all selfish players contribute 1 at t if $C(t) > 1 - \frac{A}{N}$ and 0 if $C(t) < 1 - \frac{A}{N}$. Note that the starting assumption implies that in any subgame perfect Nash equilibrium, after any length- t history all reciprocal players get a per period payoff of 0 in periods $t+1, \dots, T$. Then after any length- $(t-1)$ history $(x_i^1)_{i \in N}, \dots, (x_i^{t-1})_{i \in N}$, it has to hold that $x_i^t = E_i f_i^t((x_i^1)_{i \in N}, \dots, (x_i^{t-1})_{i \in N}, (x_j^t)_{j \in N \setminus \{i\}}) \forall i \in R$, where the expectation is taken with respect to player i 's beliefs concerning $(x_j^t)_{j \in N \setminus \{i\}}$ after history $(x_i^1)_{i \in N}, \dots, (x_i^{t-1})_{i \in N}$. Then Lemma 1, together with the starting assumption, implies that for any length- $(t-1)$ history, there is a unique continuation profile in subgame perfect Nash

equilibrium, which is history-independent for selfish players and satisfies:

$$x_i^k = \frac{N-S}{S+(k-1)(N-1)} \sum_{t'=1}^{k-1} \sum_{j \in S} x_j^{t'} + \sum_{t'=1}^{k-1} \sum_{j \in R \setminus \{i\}} \frac{(N-1)k}{S-N+1+k(2+N+S+NS-N^2)+k^2(N^2-1)} x_j^{t'} + \sum_{t'=1}^{k-1} \frac{N-1-S}{S-N+1+k(2+N+S+NS-N^2)+k^2(N^2-1)} x_i^{t'}$$

for every $k \in \{t, \dots, T\}$ and $i \in R$. The claim then follows by induction. \square

It is easy to see that the property $C(t) \neq 1 - \frac{A}{N} \forall t \in \{1, \dots, T\}$ in the claim holds generically for A , for every N and S .

A.2. Proofs of Theorems in Main Text

Theorem 1. If $S \geq 1$ and A1–A5 hold, then for generic A , the public good contribution game has a unique subgame perfect Nash equilibrium, and this equilibrium exhibits a weakly decreasing pattern of contributions. If reciprocity functions are strictly concave in selfish players' contributions, then the above statement holds for all A .

Proof of Theorem 1. The same arguments as in the proof of Claim 1 establish that in every subgame perfect Nash equilibrium at period T , all selfish players contribute 0 after any $(T-1)$ -length history, and that the continuation strategy of reciprocal players after any $(T-1)$ -length history $(x_i^1)_{i \in N}, \dots, (x_i^{T-1})_{i \in N}$ is uniquely determined in subgame perfect Nash equilibrium and satisfies $x_i^T = f_i^T((x_j^1)_{j \in N}, \dots, (x_j^{T-1})_{j \in N}, ((0)_{j \in S}, (x_j^T)_{j \in R \setminus \{i\}}) \forall i \in R$.

Let $t \in \{1, \dots, T-1\}$ and assume that for almost every A (if $f_{i,j}^{k,l}$ is strictly concave for every $i \in R, j \in S$ and $k, l \in \{1, \dots, T\}$ then for every A), the following hold: for every $k \in \{t, \dots, T-1\}$ and for every length- k history $(x_i^1)_{i \in N}, \dots, (x_i^k)_{i \in N}$, all subgame perfect Nash equilibria specify the same continuation profile, which is history-independent for selfish players and satisfies that $x_i^l = f_i^l((x_j^1)_{j \in N \setminus \{i\}}, \dots, (x_j^l)_{j \in N \setminus \{i\}}) \forall i \in R$ and $l \in \{k+1, \dots, T\}$. Consider now an arbitrary $(t-1)$ -length history $(x_i^1)_{i \in N}, \dots, (x_i^{t-1})_{i \in N}$. Note that any period- t contribution by a selfish player cannot influence period- t contributions by other players. Furthermore, by the starting assumption, selfish players' contributions from $t+1$ on are generically uniquely pinned down in subgame perfect Nash equilibrium. Therefore a period- t contribution by a selfish player can only influence reciprocal players' contributions from $t+1$ on. Let $M^{k,k'}$ be the $(N-S) \times (N-S)$ matrix whose diagonal components are 0 and whose (m, n) -th component is $\alpha_{S+m, S+n}^{k,k'}$. Then Lemma 1, together with A1, implies that contribution x_i^t by $i \in S$ induces a total of $\mathbf{1}(I - M^{t,t+1})^{-1} (f_{i,j}^{t+1,t}(x_i^t))_{j \in R}$ contributions at $t+1$, a total of $(I - M^{t+1})^{-1} [M^{t+2,t+1} ((I - M^{t+1})^{-1} (f_{i,j}^{t+1,t}(x_i^t))_{j \in R} + (f_{i,j}^{t+2,t}(x_i^t))_{j \in R})]$ contributions at $t+2$, and in general for $l \in \{t+1, \dots, T\}$ a total which is equal to a linear nonnegative combination of $(f_{i,j}^{t+1,t}(x_i^t))_{j \in R}, \dots, (f_{i,j}^{l,t}(x_i^t))_{j \in R}$ at l . The aggregate impact, and therefore the aggregate change in the flow future payoffs of i , $\Delta_i^{t+}(x_i^t)$ is given by a linear nonnegative combination of $(f_{i,j}^{t+1,t}(x_i^t))_{j \in R}, \dots, (f_{i,j}^{l,t}(x_i^t))_{j \in R}$. Since $f_{i,j}^{l,t}$ is concave, differentiable, and increasing for every $j \in R$ and $l \in \{t+1, \dots, T\}$, Δ_i^{t+} is also concave, differentiable, and increasing. Furthermore, if $f_{i,j}^{l,t}$ is strictly concave for every $j \in R$ then Δ_i^{t+} is strictly concave. Meanwhile, the net change in the flow payoff of i at t is $(\frac{A}{N} - 1)x_i^t$, a linear and decreasing function. Since Δ_i^{t+} is concave and increasing, there can only be a countable set of parameter values for A such that $\frac{d\Delta_i^{t+}}{dx_i^t} = 1 - \frac{A}{N}$ has multiple solutions in $[0, 1]$, and if Δ_i^{t+} is strictly concave then there are no parameter values like that. Therefore, for almost all values of A (for any A if $f_{i,j}^{l,t}$ is strictly concave

for every $j \in R$ and $l \in \{t+1, \dots, T\}$, the contribution of i is uniquely pinned down after $(x_j^1)_{j \in N, \dots}, (x_i^{t-1})_{j \in N}$. Since this holds for any $i \in S$ and length- $(t-1)$ history $(x_j^1)_{j \in N, \dots}, (x_i^{t-1})_{j \in N}$, the induction assumption implies that for almost all values of A (for any value of A if reciprocity functions towards selfish players are strictly concave), the continuation strategy after $t-1$ is uniquely pinned down for all selfish players in subgame perfect Nash equilibrium.

Note that the induction assumption implies that in any subgame perfect Nash equilibrium, after any length- t history, all reciprocal players get a flow payoff of 0 in periods $t+1, \dots, T$. Therefore, after any length- $(t-1)$ history $(x_j^1)_{j \in N, \dots}, (x_i^{t-1})_{j \in N}$ and any subgame perfect Nash equilibrium s , it holds that the action specified by s_i after $(x_j^1)_{j \in N, \dots}, (x_i^{t-1})_{j \in N}$ is

$$E_{ij} f_i^T \left((x_j^1)_{j \in N / \{i\}}, \dots, (x_j^T)_{j \in N / \{i\}} \right)$$

$\forall i \in R$, where the expectation is taken with respect to player i 's beliefs concerning $(x_j^T)_{j \in N / \{i\}}$ after history $(x_i^1)_{i \in N, \dots}, (x_i^{T-1})_{i \in N}$. Lemma 1 then implies that after $(x_j^1)_{j \in N, \dots}, (x_i^{t-1})_{j \in N}$ the period t contribution of i is uniquely pinned down in subgame perfect Nash equilibrium, and

$$x_i^t = f_i^T \left((x_j^1)_{j \in N / \{i\}}, \dots, (x_j^T)_{j \in N / \{i\}} \right).$$

This concludes that for almost every A (if $f_{i,j}^{k,l}$ is strictly concave for every $i \in R, j \in S$ and $k, l \in \{1, \dots, T\}$ then for every A), the following hold: for every $k \in \{t-1, \dots, T-1\}$ and for every length- k history $(x_i^1)_{i \in N, \dots}, (x_i^k)_{i \in N}$, all subgame perfect Nash equilibria specify the same continuation profile, which is history-independent for selfish players and satisfies:

$$x_i^l = f_i^l \left((x_j^1)_{j \in N / \{i\}}, \dots, (x_j^l)_{j \in N / \{i\}} \right)$$

$\forall i \in R$ and $l \in \{k+1, \dots, T\}$. By induction then, for almost every A (if $f_{i,j}^{k,l}$ is strictly concave for every $i \in R, j \in S$ and $k, l \in \{1, \dots, T\}$ then for every A), there is a unique subgame perfect Nash equilibrium, which is history-independent for selfish players and satisfies:

$$x_i^l = f_i^l \left((x_j^1)_{j \in N / \{i\}}, \dots, (x_j^l)_{j \in N / \{i\}} \right)$$

$\forall i \in R$ and $l \in \{1, \dots, T\}$.

A4 implies that $\frac{d\Delta_i^{t+1}(x)}{dx} \geq \frac{d\Delta_i^{(t+1)+}(x)}{dx}$ for every $i \in S, x \in [0, 1]$ and $t \in \{1, \dots, T-1\}$; therefore, for generic parameter values, the contributions of selfish players are weakly decreasing in subgame perfect Nash equilibrium. A3 then implies that along the equilibrium path, the contributions of all players are weakly decreasing. \square

Theorem 2. Suppose assumptions A1–A5 hold and the game has a unique subgame perfect Nash equilibrium s . Then if $s_i^t > 0$ for some $i \in S$, then $s_j^t > 0 \forall j \in R$ and $k \in \{1, \dots, t\}$.

Proof of Theorem 2. By Theorem 1, $s_i^t > 0$ for some $i \in S$ implies $s_j^t > 0 \forall j \in R, k \in \{1, \dots, t\}$. A4 then implies the claim. \square

A.3. Statements and proofs of additional results

A.3.1. Private return from contributing

Proposition 1. Assume $S \geq 1$. Let A and \hat{A} be such that that the games in which the return of private investment are A and \hat{A} have unique subgame perfect Nash equilibria s and \hat{s} . Then $A < \hat{A}$ implies $s_i^t \leq \hat{s}_i^t \forall t \in \{1, \dots, T\}$ and $i \in N$.

Proof of Proposition 1. We first begin with the following Lemma:

Lemma 2. Consider two contribution games \mathcal{G} and $\hat{\mathcal{G}}$ with individual contributions revealed after rounds, in which the players are the same: $N = \hat{N}, S = \hat{S}$, and $f_i^t = \hat{f}_i^t \forall i \in R$ and $t \in \{1, \dots, \min(T, \hat{T})\}$. Suppose \mathcal{G} and $\hat{\mathcal{G}}$ have unique subgame perfect Nash equilibria s and \hat{s} . If $s_i^t \leq \hat{s}_i^t \forall i \in S$, then $s_i^t \leq \hat{s}_i^t \forall i \in N$.

Proof of Lemma 2. Let $T^* = \min(T, \hat{T})$. Let $\hat{y}_i^k = \sum_{k'=1}^{k-1} \sum_{j \in N / \{i\}} f_{ij}^{k,k'}(\hat{s}_j^{k'}) + \sum_{j \in S} f_{ij}^{k,k}(\hat{s}_j^1)$, and let $y_i^k = \sum_{k'=1}^{k-1} \sum_{j \in N / \{i\}} f_{ij}^{k,k'}(s_j^{k'}) + \sum_{j \in S} f_{ij}^{k,k}(s_j^1) \forall i \in R$ and $k \in \{1, \dots, T^*\}$. Suppose that for some $k \in \{1, \dots, T^*\}$, it holds that $\hat{s}_i^t \geq s_i^t \forall i \in N$ and $t \in \{1, \dots, k-1\}$, and $\hat{s}_i^k \geq s_i^k \forall i \in S$. Note that this holds for $k=1$. Lemma 1 implies that $\hat{s}_R^k = (I-M^k)^{-1} \hat{y}_R^k$ and $s_R^k = (I-M^k)^{-1} y_R^k$, where $\hat{y}_R^k = (\hat{y}_i^k)_{i \in R}, y_R^k = (y_i^k)_{i \in R}, \hat{s}_R^k = (\hat{s}_i^k)_{i \in R}, s_R^k = (s_i^k)_{i \in R}$, and M^k is the $(N-S) \times (N-S)$ matrix whose diagonal elements are 0 and its (m, n) -th element is $\alpha_{m,n}^{k,k} s_{m,n} + n$. Since $\sum_{j \in S} f_{ij}^{k,k}(s_j^k)$ is increasing in $s_j^k \forall i \in R$ and $j \in S$, and $(I-M^k)^{-1} y_R^k$ is increasing in y_R^k in the relevant nonnegative range, the above establishes that $\hat{s}_i^k \geq s_i^k \forall i \in R$. Then $k < T$ implies $\hat{s}_i^t \geq s_i^t \forall i \in N$ and $t \in \{1, \dots, k\}$, and $\hat{s}_i^{k+1} \geq s_i^{k+1} \forall i \in S$. The claim then follows by induction. \square

The proof of Theorem 1 establishes that the equilibrium contribution of any player $i \in S$ in any period $t \in \{1, \dots, T\}$ is given by $x_i^t = \operatorname{argmax}_{x \in [0,1]} \left(\left(\frac{A}{N} - 1 \right) x + \Delta_i^{t+}(x) \right)$ if the return to contributing is

A , and it is given by $x_i^t = \operatorname{argmax}_{x \in [0,1]} \left(\left(\frac{\hat{A}}{N} - 1 \right) x + \Delta_i^{t+}(x) \right)$ if the return

to contributing is \hat{A} , where Δ_i^{t+} is a term that increases in x . $\hat{A} > A$ then implies that $\hat{s}_i^t \geq s_i^t \forall i \in S$ and $t \in \{1, \dots, T\}$, where \hat{s} is the unique subgame perfect Nash equilibria of the game in which the return to contributing is \hat{A} and s is the unique subgame perfect Nash equilibria of the game in which the return to contributing is A . The claim then follows from Lemma 2. \square

A.3.2. Number of periods

Proposition 2. Assume $S \geq 1$. Let A be such that the games with T and \hat{T} number of periods have unique subgame perfect Nash equilibria s and \hat{s} . Then $T < \hat{T}, f_i^t = \hat{f}_i^t$ implies $s_i^t \leq \hat{s}_i^t \forall t \in \{1, \dots, T\}$ such that $t \leq \min(T, \hat{T})$ and $i \in N$.

Proof of Proposition 2. Theorem 1 establishes that the equilibrium contribution of any player $i \in S$ in any period $t \in \{1, \dots, T\}$ is given by $x_i^t = \operatorname{argmax}_{x \in [0,1]} \left(\left(\frac{A}{N} - 1 \right) x + \Delta_i^{t+}(x) \right)$, where Δ_i^{t+} is a term that increases in x and increases in T . This implies $\hat{s}_i^t \geq s_i^t \forall i \in S$ and $t \in \{1, \dots, T\}$. The claim then follows from Lemma 2. \square

A.3.3. One-shot games

Proposition 3. If $T=1$ and $S \geq 1$, then the game has a unique Nash equilibrium, which involves all players contributing 0 to the public good.

Proof of Proposition 3. For any $i \in S$, contributing 0 is a dominant strategy, therefore $s_i^1 = 0$ for any Nash equilibrium s . By Lemma 1, there is a unique vector of contributions $x_{S+1, \dots, N}$ by the reciprocal players such that $f_i(0, \dots, 0, x_{S+1, \dots, N}) = x_i \forall i \in R$. Since $f_i(0, \dots, 0) = 0 \forall i \in R$, the unique Nash equilibrium of the game is then $s_i^1 = 0 \forall i \in N$. \square

Appendix B. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.jpubeco.2010.11.016.

References

- Anderson, S., Goeree, J., Holt, C., 1998. A theoretical analysis of altruism and decision error in public goods games. *Journal of Public Economics* 70 (2), 297–323.
- Andreoni, J., 1988. Why free ride? Strategies and learning in public goods experiments. *Journal of Public Economics* 37 (3), 291–304.
- Andreoni, J., 1989. Giving with impure altruism: applications to charity and Ricardian equivalence. *Journal of Political Economy* 97 (6), 1447–1458.
- Andreoni, J., Petrie, R., 2004. Public goods experiments without confidentiality: a glimpse into fund-raising. *Journal of Public Economics* 88, 1605–1623.
- Andreoni, J., Samuelson, L., 2006. Building rational cooperation. *Journal of Economic Theory* 127, 117–154.
- Bandiera, O., Barankay, I., Rasul, I., 2005. Social preferences and the response to incentives: evidence from personnel data. *Quarterly Journal of Economics* 120, 917–962.
- Biele, G., 2006. No man is an island: Cooperation in groups and social learning. PhD dissertation, Freie Universität Berlin.
- Blount, S., 1995. When social outcomes aren't fair: the effect of causal attributions on preferences. *Organizational Behavior and Human Decision Processes* 63, 131–144.
- Bolton, G.E., Ockenfels, A., 2000. ERC: a theory of equity, reciprocity, and competition. *American Economic Review* 90, 166–193.
- Brandts, J., Schram, A., 2001. Cooperation and noise in public goods experiments: applying the contribution function approach. *Journal of Public Economics* 79, 399–427.
- Brown, M., Falk, A., Fehr, E., 2004. Relational contracts and the nature of market interactions. *Econometrica* 72, 747–780.
- Cameron, C., Gelbach, J., Miller, D., 2011. Robust Inference with Multi-Way Clustering. *Journal of Business and Economic Statistics* 29 (2), 238–249.
- Cox, J., 2007. On modeling voluntary contributions to public goods. *Public Finance Review* 35, 311–332.
- Cox, J., Friedman, D., Gjestad, S., 2007. A tractable model of reciprocity and fairness. *Games and Economic Behavior* 59, 17–45.
- Crosno, R., 1996. Partners and strangers revisited. *Economics Letters* 25, 25–32.
- Duwfenberg, M., Kirchsteiger, G., 2004. A theory of sequential reciprocity. *Games and Economic Behavior* 47, 268–298.
- Fehr, E., Schmidt, K., 1999. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114, 817–868.
- Fehr, E., Schmidt, K., 2002. Theories of fairness and reciprocity – evidence and economic applications. *Advances in Economics and Econometrics – 8th World Congress, Econometric Society Monographs*. Cambridge University Press, Cambridge.
- Field, A., 2002. Altruistically inclined? The behavioral sciences, evolutionary theory and the origins of reciprocity. *UMichigan Press, Ann Arbor*.
- Fischbacher, U., 2007. z-Tree: Zurich Toolbox for Readymade Economic Experiments. *Experimental Economics* 10 (2), 171–178.
- Fischbacher, U., Gächter, S., 2010. Social preferences, beliefs, and the dynamics of free riding in public goods. *American Economic Review* 100, 541–556.
- Fischbacher, U., Gächter, S., Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters* 71, 397–404.
- Fiske, A.P., 1992. The four elementary forms of sociality: framework for unified theory of social relations. *Psychological Review* 99 (4), 689–723.
- Fudenberg, D., Maskin, E., 1986. The folk theorem in repeated games with discounting and incomplete information. *Econometrica* 54, 533–554.
- Gächter, S., 2007. Conditional cooperation. behavioral regularities from the lab and the field and their policy implications. *Economics and Psychology. A Promising New Cross-Disciplinary Field*. In: Frey, B., Stutzer, A. (Eds.), CESifo Seminar Series. The MIT Press.
- Gächter, S., Renner, E., 2010. The effects of (incentivized) belief elicitation in public goods experiments. *Experimental Economics* 13, 364–377.
- Guttman, J., 1978. Understanding collective action: matching behavior. *American Economic Review* 68, 251–255.
- Isaac, M., Walker, J., 1988. Group size effects in public goods provision: the voluntary contributions mechanism. *Quarterly Journal of Economics* 103 (1), 179–199.
- Isaac, M., Walker, J., Thomas, S., 1984. Divergent evidence on free riding: an experimental evidence of possible explanations. *Public Choice* 43 (1), 113–149.
- Isaac, M., Walker, J., Williams, A., 1994. Group size and the voluntary provision of public goods: experimental evidence utilizing large groups. *Journal of Public Economics* 54, 1–36.
- Jehiel, P., 2005. Analogy-based expectation equilibrium. *Journal of Economic Theory* 123, 81–104.
- Kelley, H., Stahelski, A., 1970. Social interaction basis of cooperators' and competitors' beliefs about others. *Journal of Personality and Social Psychology* 16, 190–197.
- Keser, C., Winden, F., 2000. Conditional cooperation and voluntary contributions to public goods. *Scandinavian Journal of Economics* 102 (1), 23–39.
- Kim, O., Walker, M., 1984. The free rider problem: experimental evidence. *Public Choice* 18, 3–24.
- KLUMPP, T., 2010. Finitely repeated provision of a public good. mimeo Emory University.
- Kreps, D., Wilson, R., 1982. Reputation and imperfect information. *Journal of Economic Theory* 27, 253–279.
- Kreps, D., Milgrom, P., Roberts, J., Wilson, R., 1982. Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory* 27, 245–252.
- Ledyard, J., 1995. Public goods. In: Kagel, J., Roth, A. (Eds.), *Handbook of Experimental Economics*. Princeton University Press.
- Levine, D., 1998. Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics* 1, 593–622.
- McCabe, K., Rigdon, M., Smith, V., 2003. Positive reciprocity and intentions in trust games. *Journal of Economic Behavior & Organization* 52, 267–275.
- Mckelvey, R., Palfrey, T., 1992. An experimental study of the centipede game. *Econometrica* 60, 803–836.
- Mengel, F., 2009. Learning by (limited) forward looking players. mimeo Maastricht University.
- Neyman, A., 1985. Bounded complexity justifies cooperation in the prisoner's dilemma. *Economic Letters* 19, 227–229.
- Neyman, A., 1999. Cooperation in repeated games when the number of stages is not commonly known. *Econometrica* 67, 45–64.
- Offerman, T., Sonnemans, J., Schram, A., 1996. Value orientations, expectations, and voluntary contributions in public goods. *Economic Journal* 106, 817–845.
- Palfrey, T., Prisbey, J., 1997. Anomalous behavior in linear public goods experiments: how much and why? *American Economic Review* 87 (5), 829–846.
- Plott, C., Smith, V., 2008. *Handbook of experimental economics results, Volume 1*. Elsevier North-Holland Publishing, New York.
- Rabin, M., 1993. Incorporating fairness into game theory and economics. *American Economic Review* 83, 1281–1302.
- Radner, R., 1980. Collusive behavior in noncooperative epsilon-equilibria of oligopolies with long but finite lives. *Journal of Economic Theory* 22, 136–154.
- Radner, R., 1986. Can bounded rationality resolve the prisoner's dilemma? In: Andreu, Mas-Colell, Hildenbrand, W. (Eds.), *Contributions to Mathematical Economics*. North-Holland, Amsterdam, pp. 387–399.
- Saijo, T. and Yamaguchi, H., 1992. The “spite” dilemma in voluntary contribution mechanism experiments. University of Tsukuba mimeo.
- Sobel, J., 1985. A theory of credibility. *Review of Economic Studies* 52, 557–573.
- Sonnemans, J., Schram, A., Offerman, T., 1999. Strategic behavior in public good games: when partners drift apart. *Economic Journal* 35–41.
- Stahl, D., Wilson, P., 1995. On players' models of other players: theory and experimental evidence. *Games and Economic Behavior* 10, 218–254.
- Takayama, A., 1985. *Mathematical Economics*. Cambridge University Press.
- Ule, A., 2010. Cooperation in finitely repeated games on networks. mimeo University of Amsterdam.