The strategic sophistication of conditional cooperators: Evidence from public goods games*

Francesco Fallucchi (LISER)

R. Andrew Luccasen, III (Mississippi University for Women)

Theodore L. Turocy (University of East Anglia)

October 18, 2017

Abstract

We develop the theory of behavioural types in social dilemma games by experimentally studying four public goods voluntary contributions games in which the incentive structures are varied. We replicate previous studies by identifying free-riders, as well as two distinct groups of conditional cooperators (strong and weak), as comprising a supermajority of participants when payoffs are linear. Strong conditional cooperators, who match contributions one-for-one in the linear game, display sophisticated behaviour in games with nonlinear payoffs: they match one-for-one only when doing so improves social welfare. This mode of conditional cooperation appears to be underpinned by a sophisticated understanding and assessment of the financial incentives presented by the game.

Keywords: public goods, behavioural types, conditional cooperation, sophistication, experiment.

JEL Classifications: C72, C92, D71.

1 Introduction

The extensive literature on laboratory experiments in social dilemma games reports frequent and often substantial levels of cooperative behaviour. The survey of Chaudhuri (2011) reviews the evidence in the case of public goods games. In the usual voluntary contributions mechanism (VCM)

*For many helpful comments, we thank seminar participants at University of Nottingham and Queen's University Belfast; conference participants at the European Meetings of the Economic Science Association; as well as Abigail Barr and Werner Güth. Fallucchi and Turocy acknowledge the support of the Network for Integrated Behavioural Science (Economic and Social Research Council Grant ES/K002201/1). The Network also provided the funds for participant payments in the reported experiments. All errors are the responsibility of the authors.

setup of public goods games, an individual can contribute none, some, or all of her private endowment to a "public account," which yields benefits to all members of a group. The dilemma arises when, as is usual in experiments, an individual's private marginal benefit of contribution is less than their cost of contribution, while the overall marginal group benefit exceeds the cost of contribution.

Emerging from the work of Fischbacher et al. (2001) and Fischbacher and Gächter (2010), among others, is an account in which many, but far from all, players in VCM games are conditional cooperators. Players are more inclined to make contributions when they have seen others contribute, or when they anticipate others will contribute. Such an account organises the behaviour commonly seen when a VCM game is played repeatedly among the same group of participants, in which contribution levels start out relatively high, and then decline with repetition. Conditional cooperators may be willing initially to contribute in anticipation of the contributions of others. The presence of free-riders would result in lower-than-hoped-for contributions, resulting in those conditional cooperators lowering their expectations for contributions of others, and therefore their own contributions as well.

Explaining the willingness of participants to contribute has been an important focus of the social dilemma literature. Andreoni (1995b) addressed the question by comparing behaviour in a standard linear VCM with related games in which motives for kind behaviour are absent. He finds evidence that potential confusion decreases quickly with experience, and therefore confusion cannot alone explain the persistence of positive contributions in the linear VCM. In contrast, Burton-Chellew et al. (2016) classify participants in a linear VCM experiment based on the method of Fischbacher et al. (2001), and report that most participants classified as conditional cooperators fail to answer correctly when asked about their understanding of the financial incentives of the game.

While the literature offers a variety of reasons to explain conditional cooperation, little is known about how robust this behaviour is across variations in the strategic structure of interactions. The linear VCM places the contribution level that maximises a participant's payoff at zero, which is on the boundary of the set of possible contributions. The pure altruist model of Andreoni (1990) allows for a general structure of preferences over private and public good consumption. For many specifications of this model, the contribution that maximises a player's earnings will be greater than zero, at least for some possible choices of the other players.

In particular, contributions to the public good in real-world settings might be strategic substitutes. Parents with school-aged children are often asked to volunteer to chaperone on a number of class field trips throughout the year. If other parents are not available or willing to volunteer, then the optimal response for a family is to volunteer as much as possible, as, in the most extreme case, an insufficient number of chaperones might lead to cancellation of trips, disappointing the children. However, the value of additional chaperones, beyond a certain number, is minimal. If a family anticipates there will be many volunteers among the other parents, then they would likely respond by putting themselves forward for fewer events.

Equally, contributions to the public good might be strategic complements. Online review sites aggregate the comments and evaluations of many people on restaurants, hotels, and other products. Using information about the reviews of others, these sites can offer generic recommendations. However, the more an individual submits reviews of their own experiences, the more it is possible for a site to offer customised recommendations, such as "Other users who liked the Red Lion Pub also recommended the Lamb Inn." In such a system, if others are contributing many recommendations, an individual might have incentives to contribute many recommendations themselves, to help "train" the system to give them good recommendations for new products to try.

Aside from capturing the rich possibilities for externalities found in real-world public goods, public goods games with nonlinear payoff structures can provide evidence for distinguishing among candidate explanations for behaviour. If conditional cooperators are following a generic heuristic, along the lines of Emily Post's "do as others do" advice, this should be observed irrespective of the strategic structure of the game, and by extension the financial incentives of its experimental implementation. However, other types of reasoning can generate conditional cooperation in the linear game, including a desire not to be exploited by free-riders, or as a type of focal point for selecting which socially-improving contribution level to choose from among many possibilities. Contribution patterns in nonlinear games can provide evidence as to the possible motivations of conditionally cooperative participants, and indeed whether they continue to be conditionally cooperative at all in games where positive contributions to the public good are expected even from participants who care only about their own earnings.

In this paper we find strong evidence that a substantial fraction of participants make a conscious and well-informed decision to behave in a conditionally cooperative way. We apply the contingency-choice framework of Fischbacher et al. (2001) to public goods games with nonlinear financial payoffs. Standard instructions for linear VCMs present a participant's endowment in the form of *tokens* which can be allocated either to the public good or kept for private consumption. We develop a new choice architecture, based on separately-identified tokens, which outlines the marginal financial incentives of allocating each token to both the participant and to the group. Following the methodology of Fallucchi et al. (2017), using data from a linear VCM we find evidence for a clear distinction between two types of conditional cooperators, which are conflated in the typology used by Burton-Chellew et al. (2016): strong conditional cooperators, who match anticipated contributions one-for-one, and weak conditional cooperators, who match at a lower rate. We find that when playing nonlinear VCMs, strong conditional cooperators predominantly follow a sophisticated rule: they continue to match contributions one-for-one only when this implies a contribution above the amount that would maximise their own earnings. If low contributions by other players are anticipated, strong conditional cooperators instead choose the contribution which maximises their own earnings. The sophistication of this response pattern, and the fact that it is adopted by a substantial number of participants, is not consistent with a misunderstanding of the financial incentives.

The workhorse linear payoff structure employed in a majority of VCM experiments is attractive due to its ease of explanation without requiring participants to deal with arithmetic to determine their responses. However, explaining incentives in nonlinear games is more difficult; Andreoni (1993), Keser (1996), and Chan et al. (2002) used payoff tables, and Gronberg et al. (2012) used a visualisation of the payoff surface; these methods generally focus on conveying how the participant's own earnings depend on the decisions of the participant and the others in the group. In our novel choice architecture, we represent the participant's endowment by separately-numbered tokens. Each token is labeled with its (marginal) earnings consequences if allocated to the public good, or if allocated to private consumption. This contrasts with most protocols which ask *how many* tokens to allocate. It is known that how this allocation question is framed can affect behaviour. (Andreoni, 1995a) Our design presents the public and private consumption options symmetrically. The labeling of tokens with their marginal consequences simplifies the presentation of nonlinear incentives, while communicating for each token the tradeoff between own earnings and group earnings. In addition, the use of this choice architecture provides a robustness check on contribution behaviour in the linear game. Using the classification method of Fallucchi et al. (2017), the proportions of behavioural types in our game are similar to those found in previous studies. The robustness of behavioural types to the choice architecture used to elicit them supports the hypothesis that, at least for many participants, the reported choices are based on an informed understanding of the structure of the game.[1]

The remainder of the paper is structured as follows. In Section 2 we introduce the theoretical framework. In Section 3 we provide a description the experimental design with a particular emphasis on the choice architecture. Section 4 contains the data analyses and results. We conclude in Section 5 with a discussion.

2 Theoretical framework

There are N players, $i = 1, \ldots, N$. Each player i has an endowment $\omega_i > 0$ of resources, which she can allocate between private consumption x_i and a contribution towards a public good g_i, where $x_i + g_i = \omega_i$. The total amount contributed towards the public good is $G = \sum_{j=1}^{N} g_j$. In a standard

[1]Prior to conducting the experiments, we were concerned that our explicit representation of the incentives in the linear VCM would lead to significantly higher rates of free-riding behaviour. Our data conclusively reject this.

abuse of notation we will write $G_{-i} = \sum_{j \neq i} g_j$ for the contributions of players other than i to the public good. For a given private consumption x_i and total contributions to the public good G, the payoff of player i is given by a function $\Pi_i(x_i, G)$.[2] We assume that Π_i is strictly increasing in both x_i and G.

Several studies have used nonlinear payoff functions to generate games in which the best response for player i is to contribute a strictly positive amount g_i to the project. Andreoni (1993) used a Cobb-Douglas payoff specification; Cason and Gangadharan (2015) a piecewise-linear specification; and Chan et al. (2002) a quadratic specification. Keser (1996), Sefton and Steinberg (1996), Willinger and Ziegelmeyer (1999) and Gronberg et al. (2012) likewise used quadratic specifications, such that a strictly positive contribution was a dominant strategy.

In our experiment, we will specialize to the case in which $\frac{\partial \Pi_i}{\partial G} = 0.4$. This derivative is often called the marginal per-capita return (MPCR). We consider payoff functions of the form

$$\Pi_i(x_i, G) = \alpha + (\beta_1 - \gamma G_{-i})\, x_i - \beta_2 x_i^2 + 0.4G. \tag{1}$$

where $\beta_1 > 0$ and $\beta_2 \geq 0$. When $\beta_2 = \gamma = 0$, this is the additively separable and linear model, and the best response is $(g_i^\star(G_{-i}), x_i^\star(G_{-i})) = (0, \omega)$.

When $\beta_2 > 0$, the payoff from private consumption is strictly concave in x_i and the reaction function for an interior solution is

$$x_i^\star(G_{-i}) = \frac{\beta_1 - \gamma G_{-i} - 0.4}{2\beta_2},$$

The slope of the reaction function is determined by the sign of γ. If $\gamma > 0$, then x_i is smaller if G_{-i} is larger; that is, player i wants to contribute more to the public good when others are making larger contributions, whereas when $\gamma < 0$ player i wants to contribute less to the public good when others' contributions are higher.

We use the form (1) to determine the *monetary* payoffs in our experiment. This functional form is additively separable in earnings received from resources x_i allocated to private consumption, and resources g_i allocated to the project. In our experiment, we set parameters such that full contribution of resources to the project maximises total earnings of the group.

[2] This is the general "pure altruist" model of Andreoni (1990).

3 Experimental design

3.1 Payoff structure treatments

Participants were allocated at random into groups of four. For identification purposes, the member IDs of the group were the four suits of a standard deck of cards (clubs, diamonds, hearts, and spades). The icons for these suits were used extensively in the instructions as well as the decision screens. Each participant's instructions were customised based on their suit identification. For example, the instructions for a participant with the identifier clubs (♣) consistently used phrasing like "your ID (♣)" and "the other members of your group (♢♡♠)."[3]

Participants played a series of four public goods games. Each participant was endowed with 20 *tokens*, to be allocated between a public good (*project*) benefiting all four members of the group, and a *private account* benefiting only the participant herself. In all games, the consequence of allocating any token to the project was a payoff of 40p to each member of the group. The games differed in the structure of the consequences to allocating tokens to the private account.

In the baseline game LINEAR, the value of each token allocated to the private account was £1.00, irrespective of how many tokens were allocated. This is the same payoff structure as used in Fischbacher et al. (2001), and results in a dominant strategy for own-earnings maximisers to allocate all tokens to the private account.

We contrast this with three treatments in which the value of tokens allocated to the private account depends on how many tokens are allocated. With the quadratic payoff specification, the marginal value to the private account changed linearly in the number of tokens. We consider three specifications:

- DOMINANT: The marginal value of a token allocated to the private account is decreasing in the number of tokens, but is independent of the allocations of other players.
- SUBSTITUTES: The marginal value of a token allocated to the private account is decreasing in the number of tokens, but is increasing in the number of tokens allocated to the project by other group members.
- COMPLEMENTS: The marginal value of a token allocated to the private account is decreasing in the number of tokens and in the number of tokens allocated to the project by other group members increases.

The description for each of the specifications and the tables with the value of the allocations to the private accounts are reported in Appendices A.1-A.4. The Nash equilibrium for allocations

[3] Complete instructions are available in Appendix A.

to the project, assuming own-earnings maximisation, is 7 tokens in each of the nonlinear games. Participants played all four games without feedback. After decisions for all four games were made, one of the games was selected at random to determine payment.

3.2 Timing of moves

Following Fischbacher et al. (2001), allocations in each game were realized in a two-stage process. In Stage 1, three of the four members of the group set their allocation. Then, in Stage 2, the fourth member of the group would learn the average of the three other members' contributions to the project, and then set their allocation. Choices were elicited using a variation on the strategy method. The identity of which member of the group would choose in Stage 2 was not determined until the end of the experiment. Therefore, participants were asked to make decisions both for the event in which they would make their allocation in Stage 1, and, respectively, in Stage 2.

Stage 1 consisted of a single decision about how many tokens to allocate to the project, which Fischbacher et al. call the *unconditional contribution*. The Stage 2 decision was a complete schedule of contributions, one for each possible average allocation by the other members of the group, from 0 tokens up through 20 tokens; Fischbacher et al. refer to this as the *contribution table*.

Our instructions refer to the decisions as Stage 1 and Stage 2 decisions, rather than unconditional contributions and contribution tables as in Fischbacher et al.. Numbering the stages helps to communicate the sequential realisation of the actual decisions, and eliminates the need for extra jargon to describe the decisions.

3.3 The choice architecture

We elicited decisions using the graphical device shown in Figure 1, which we referred to as the *allocation panel* in the instructions. Several considerations informed the design of this choice architecture. Our starting point was the observation that many experiments in public goods games describe the task as allocating tokens between a private account and the public good. Most experiments then elicit the choice as the *number* of tokens to allocate, and express the earnings consequences in terms of total earnings. When payoffs are not linear in the allocation, it is more complicated to explain how earnings depend on choices, in a way that is accessible to participants.

The allocation panel instead expresses the decision on a token-by-token basis. Each token is given a distinct number from 1 up to 20. Each token corresponds to two possible consequences, depending on the allocation decision. There is one consequence if the token is allocated to the project; in the case of our experiment, we fixed this as resulting in 40p of earnings to each of the four group members. There is another consequence if the token is allocated to the private account. For a game with nonlinear payoffs, the distinct numbering of the tokens allow us to specify different

Figure 1: Screenshot of the allocation panel, used by participants to indicate decisions in the experiment. Left: The panel at the start of a decision. Right: The panel with an allocation selected.

values for each token. Tokens were numbered such that tokens with higher numbers had higher values if allocated to the private account.

The participant allocated a token to the project by clicking on the box to the left of that token. Similarly, the participant allocated a token to the private account by clicking on the box to the right of that token. When a participant clicked to allocate token i, the device automatically allocated all tokens with numbers below i to the project, and all tokens with numbers above i to the private account.[4] Participants were able to adjust their allocations as many times as they wished before confirming. Colour-coding was used to indicate the currently-selected allocation, as shown in Figure 1.

Enforcing that the allocation to the project always consists of lower-numbered tokens does make the decision equivalent to asking how many tokens the participant wanted to allocate to the project and to the private account. Our design for the choice architecture has several features conceived with the research question in mind. This design invites the use of marginal reasoning; the information required to make calculations at the margin is exactly the comparison of the consequences on the left and the right of each token. In particular, this means that computing the own-earnings-maximising response is straightforward, as it involves only seeking which tokens would yield a return of greater than 40p if allocated to the private account. This feature suggests our choice architecture might result in a greater proportion of participants choosing to maximise their own earnings. However, unlike e.g. Gronberg et al. (2012), who use a different device for making own-earnings consequences easy to discover, our choice architecture also incorporates reminders of the social consequences of allocations to the project. Each consequence in the project column includes the word "each," and the symbols of all four group members are displayed atop the Project column, while only the participant's symbol appears over their own private account.

Our preference for a graphical device rather than typing in numbers is intended to reduce further any frictions in indicating decisions. "Accessible" numbers (such as 0, or 5) are observed more commonly in decisions when input by keyboard versus graphical devices, and, in the case of the allocation table, it would be easier to type the same number in repeatedly than to type different numbers for different hypothetical contributions of others group members. Our graphical device also allows us to present the allocations to the project and to the private account using entirely parallel language; text-based responses generally force a breaking of symmetry by asking either for an allocation to the project or an allocation to the private account.

In each game, the Stage 1 unconditional allocation was made first. Figure 2 displays the choice architecture for the Stage 2 contingency table, which requires the specification of 21 decisions. We referred to each possible realisation of the average Stage 1 allocation to the project as a *scenario*.

[4]Therefore the allocation panel did enforce efficiency in the sense that, if k tokens were allocated to the project, they were always the k tokens worth the least to the participant in their private account.

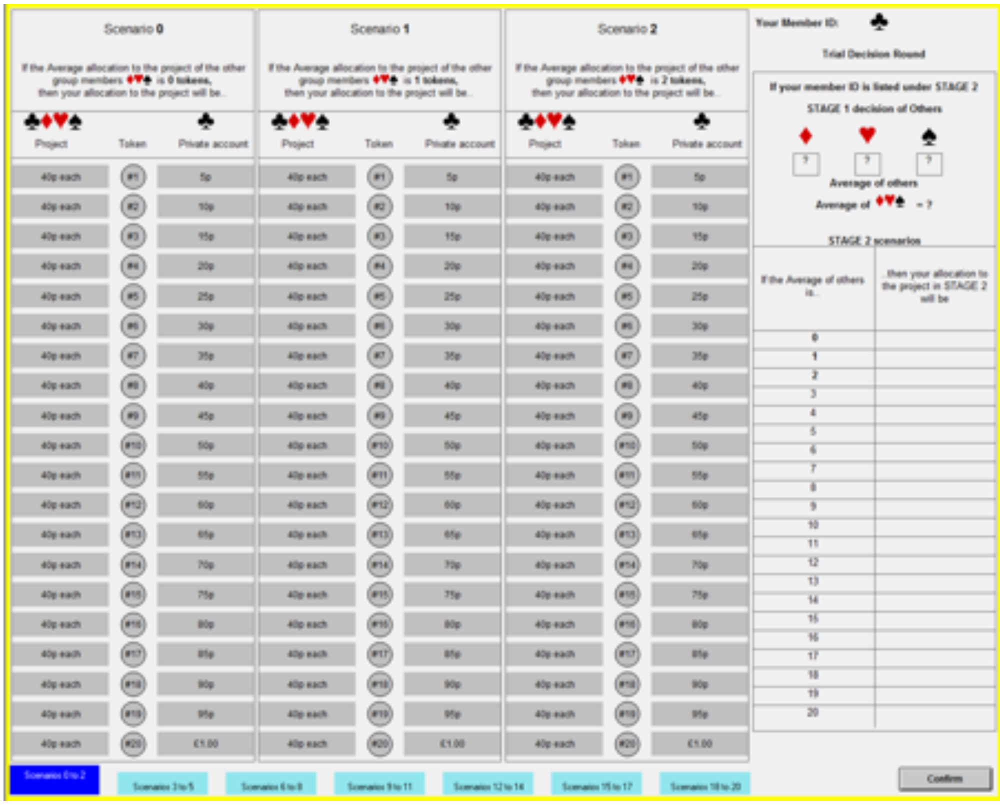


Figure 2: Indication of Stage 2 allocation table decisions. Three scenarios were available on the screen at a time. Navigation across scenarios was available using tabs at the bottom of the screen. A panel at the right summarised the allocation decisions made so far.

The allocation panels for three scenarios were available on the screen at any time, with a tabbed interface available to navigate among scenarios. A panel at the right of the screen summarised the allocations made by the participant so far. Allocations could be made in any order and changed as often as the participant liked, before confirming the decisions with the button at the bottom-right of the screen.

3.4 Experimental sessions

We conducted a total of 8 sessions at the laboratory of the Centre for Behavioural and Experimental Social Science (CBESS) at University of East Anglia, in April and May, 2016. 148 participants were recruited from the standing participant pool, maintained using the hRoot system (Bock et al., 2012). The treatments were presented to subjects in one of four orderings, which differed across sessions. Games 1 and 2 were always LINEAR and DOMINANT, in either order, and Games 3 and 4 where always COMPLEMENTS and SUBSTITUTES, again in either order. The experiment was programmed using zTree (Fischbacher, 2007). Sessions lasted on average 75 minutes, including instructions and control questions, and participants earned on average £23.39 with an inter-quartile range of £5.34.[5]

4 Results

4.1 Behavioural types in the linear game

Our experiment builds on the substantial literature which has used the contingency-choice protocol in linear VCMs. In our implementation we have developed a new design for the choice architecture to elicit decisions. A main objective of the design process was to represent the nonlinear payoff scheme to participants in the simplest possible way, avoiding resorting to mathematical formulas or calculations to focus on the tradeoff players face between own earnings and group earnings. This new design also provides a robustness test on contribution behaviour in the linear game. In particular, during the design process, we were aware of the possibility that the layout of the financial incentives in the linear game using the explicit labeled-token metaphor might make own-earnings-maximising behaviour more common, whether through a reduction in confusion or simple salience considerations.

As such, we begin by comparing our data to previous experiments with the same payoff structure. Fallucchi et al. (2017) re-analyse the data from six experiments using the two-stage protocol. They use hierarchical cluster analysis (Lloyd, 1982) to classify participant types based on their Stage 2 strategies, and find five behaviourally distinct groups. We briefly recap the methodology before reporting on the results as applied to our dataset.

Each participant's Stage 2 strategy can be thought of as a point in a 21-dimensional space, with each dimension corresponding to the contribution to the project for one of the possible realisations of the average contribution of the other group members. Write the strategy of a participant i as $\mathbf{g}_i = \left(g_i^{(0)}, g_i^{(1)}, \ldots, g_i^{(20)}\right)$, where $g_i^{(k)}$ is the contribution of participant i to the project if other group members contribute on average k tokens to the project.

The distance between the strategies of any two participants i and j is computed using the Manhattan distance,

$$d(\mathbf{g}_i, \mathbf{g}_j) = \sum_{k=0}^{20} \left| g_i^{(k)} - g_j^{(k)} \right|. \tag{2}$$

Economically, this metric can be interpreted as the expected difference between the Stage 2 contributions of participants i and j, if the average contribution k of other group members is chosen uniformly at random. Clusters are then formed using Ward's linkage method (Ward, 1963), which minimises the within-clusters sum of squares. We use both Caliński-Harabasz criterion (Caliński and Harabasz, 1974) and the Duda-Hart Je(2)/Je(1) index (Duda and Hart, 1973) to determine

[5] For comparison, the current living wage in the United Kingdom is £8.25 per hour.

	This paper re-analysis	Fallucchi et al. (2017)
Own-maximisers	31.1%	25.8%
Strong conditional cooperators	28.4%	38.8%
Weak conditional cooperators	25.7%	18.9%
Unconditional contributors	4.0%	4.7%
Mid-range or Various	10.8%	11.8%

Table 1: Summary of type classifications, compared to results in Fallucchi et al. (2017) based on six previous studies.

		Fallucchi et al. (2017)				
		OWN	SCC	WCC	VAR	UNC
This paper	OWN	**0.18**	8.63	3.71	8.62	15.90
	SCC	8.83	**0.36**	5.12	4.92	7.12
	WCC	5.17	3.98	**1.44**	3.52	10.75
	MID	9.64	4.60	5.92	**1.19**	6.27
	UNC	17.11	8.54	13.38	8.47	**1.99**

Table 2: Mean absolute distance of Stage 2 contributions between types in Fallucchi et al. (2017) and types in this paper.

the number of clusters to produce. In the event the criteria suggest different numbers of clusters, we choose the larger number of clusters, following the standard practice in applying hierarchical clustering.

Result 1. *Hierarchical cluster analysis divides participants into five behavioural types based on their Stage 2 strategies in* LINEAR*. The proportions of types are similar between our data and data from previously-reported experiments.*

Support. For each of the five types identified by clustering, we produce a heatmap of the Stage 2 strategies classified as that type. Let $T(i)$ denote the type to which a participant i is assigned. The heatmap for a type t is produced by taking the Stage 2 strategies of all participants assigned to that type, and constructing the set $\{(k, c^i_k)\}_{i:T(i)=t,k=0,\ldots,20}$. The frequencies of these ordered pairs are used to generate the heatmap, with darker shades corresponding to higher frequencies. The modal behaviour for a given average contribution k of other group members is identifiable by the darkest cells. The resulting heatmaps are presented in Figure 3, with the results from LINEAR in the left column, and the corresponding clusters from Fallucchi et al. (2017) in the right column. There is good behavioural agreement between the clustering generated from LINEAR and that of Fallucchi et al. (2017) in four of the five groups: *own-maximisers*, *strong conditional cooperators*, *weak conditional cooperators*, and *unconditional contributors*. The residual group, called "various" in

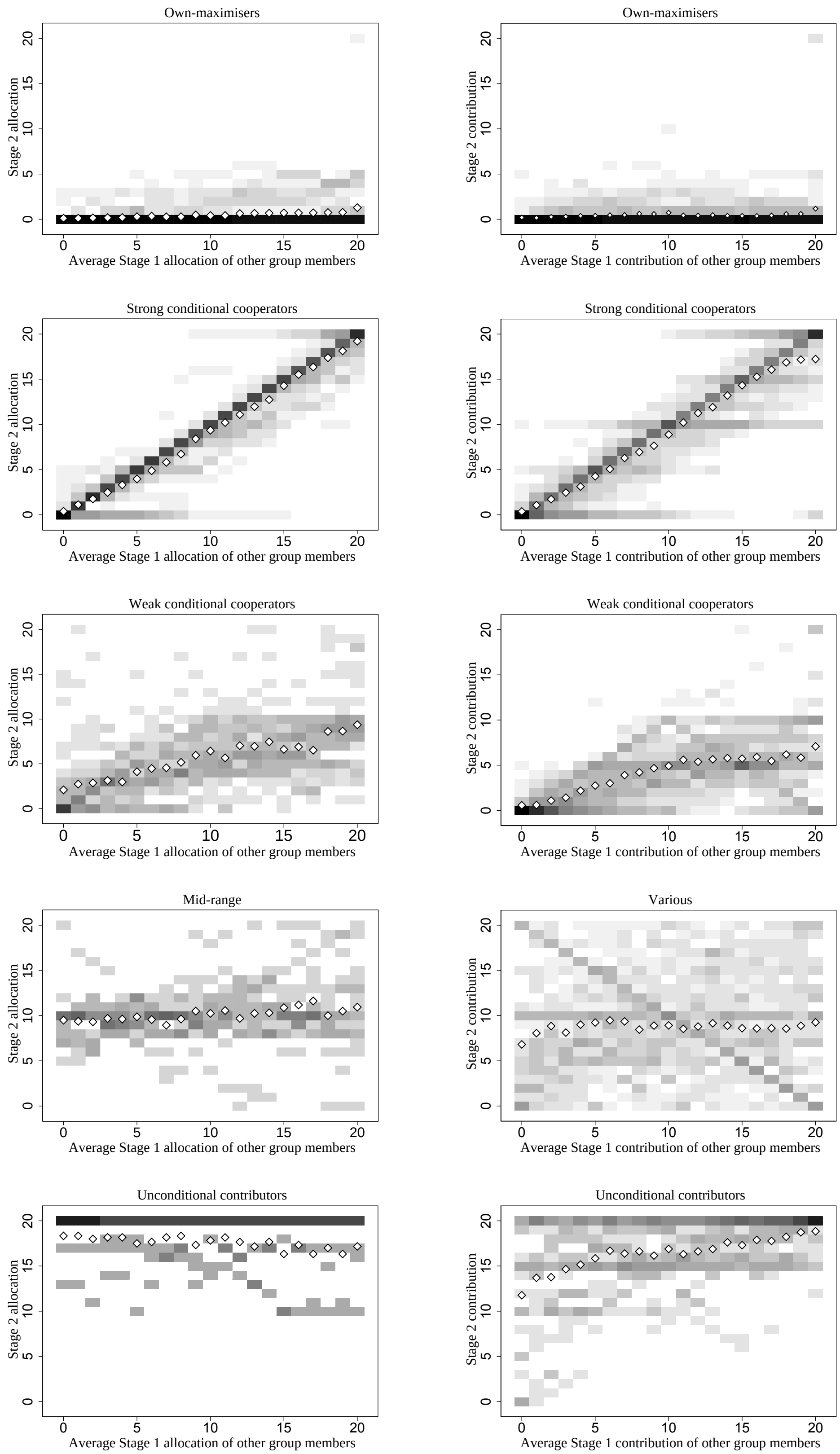


Figure 3: Heatmaps of Stage 2 contribution strategies, by behavioural type. The left column contains data from LINEAR. The right column contains the data from six previous experiments re-analysed by Fallucchi et al. (2017).

Fallucchi et al. (2017), has similar average behaviour in the two datasets, with both involving a contribution of about one-half the endowment irrespective of the contributions of the other group members.

Table 2 quantifies the similarity between the clusters in Fallucchi et al. (2017) in the clusters in our data. Each cell compares the Stage 2 contributions of one type in our data with the Stage 2 contributions of one type in Fallucchi et al. (2017). The entry is the mean absolute difference given by (2) of the average Stage 2 strategy of each type. The entries on the diagonal are much smaller than those off the diagonal. This gives a quantitative justification of our mapping between the types found in our data and those in Fallucchi et al. (2017), and shows that the average behaviour of each type in our experiment is similar to previous studies.

Table 1 compares the proportion of types produced by the cluster analysis in our study with the results in (Fallucchi et al., 2017). We observe a slightly higher proportion of own-maximisers (31.1% versus 25.8%), and slightly more weak conditional cooperators, and slightly fewer strong conditional cooperators. □

We compare the coherence of the clusters using silhouette plots (Rousseeuw, 1987) in Figure 4. This method assigns to each participant i an index in $[-1, +1]$. This number compares the average distance between participant i's strategy and the strategies of other participants of the same type, against the average distance to participants who are classified in the "next closest" type. A positive silhouette index means the participant's strategy is more similar to the strategies of others of the same type, whereas a negative silhouette index means the participant's strategy is more similar to the strategies of some other type. Participants are grouped by their type classification, and then sorted in decreasing order of the index to generate the plot. The silhouette indices for participants in our clusters are almost all positive and larger in magnitude than those in the clusters based on the previous studies. This indicates the coherence of the clusters identified in our data.

Cluster analysis also draws a distinction between strong conditional cooperators, who match the average contribution of other group members on a one-for-one basis, and weak conditional cooperators, who have Stage 2 strategies which are generally increasing, but at a rate of less than one-for-one. Weak conditional cooperators are heterogeneous, although the silhouette analysis confirms that they are well-distinguished from strong conditional cooperators on one side and own-maximisers on the other. This confirms the distinction previously proposed by by Chaudhuri et al. (2006) and Gächter et al. (2012).

Our Stage 2 differ from the surveyed previous experiments in two other respects. First, among the participants classified as weak conditional cooperators in previous studies are the so-called "hump-shaped" or "triangle" contributors, who match contributions roughly one-for-one up to 10, and then start to decrease contributions as the average contribution of other members increases. While a clear triangle pattern is evident in the heatmap of weak conditional cooperators derived

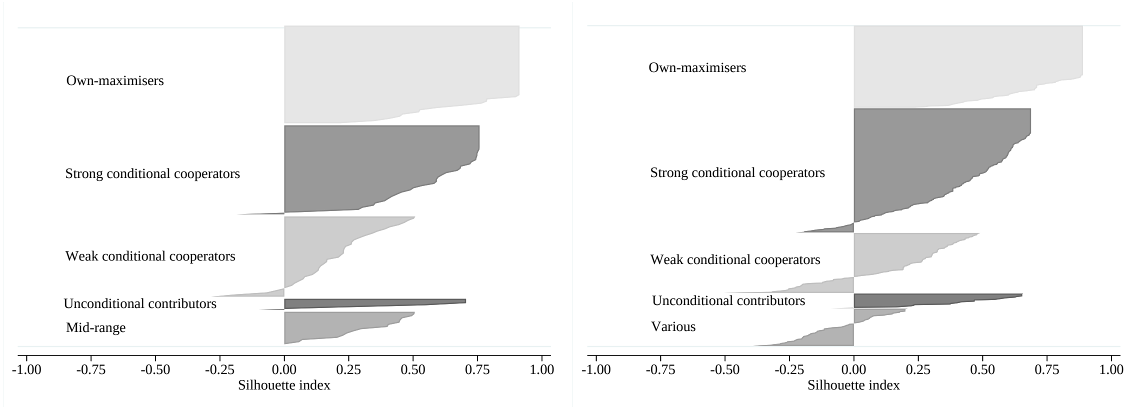


Figure 4: Silhouette plots of type clusters. Left column: our data. Right column: data from Fallucchi et al. (2017).

from previous studies, no similar pattern is found in our data.[6] Second, our residual cluster, which we label *mid-range*, is concentrated around contributions near one-half the endowment, whereas the analogous cluster *various* derived from previous studies includes contributions which span the whole of the set of possible contributions.

The Stage 2 strategies of participants classified as either strong or weak conditional cooperators suggest their Stage 1 contributions to the project depends on their beliefs about how the others in their group will play the game. The own-maximiser, mid-range, and full contributor types report Stage 2 strategies in which their intended contributions do not depend much on what others will do. We can therefore ask whether the Stage 1 decisions in LINEAR are broadly consistent with the clusters determined by the Stage 2 strategies.

Result 2. *Stage 1 contributions differ across the type clusters. In particular, own-maximisers contribute significantly less than all other types.*

Support. The average contribution in Stage 1 within each cluster reflects the differences found in their Stage 2 behaviour. Own-maximisers have the lowest level of contribution, on average of 0.39 tokens. They are followed by the two types of conditional cooperators, the weak conditional cooperators with 4.18 tokens, and the strong conditional cooperators with 5.93 tokens on average. Mid-range types contribute on average 8.31 tokens while the unconditional contributors contribute 12.17 tokens. One-way ANOVA to compare Stage 1 contributions across different clusters finds a statistically significant difference among the five clusters ($p < 0.001$).[7] The mean contribution of the own-maximisers is significantly different from all other clusters (all $p < 0.001$; significant

[6]Examining the individual Stage 2 strategies in our data likewise confirms an absence of such strategies.

[7]The non-parametric Kruskall-Wallis test also rejects the hypothesis that mean contributions are the same across different clusters ($p < 0.001$).

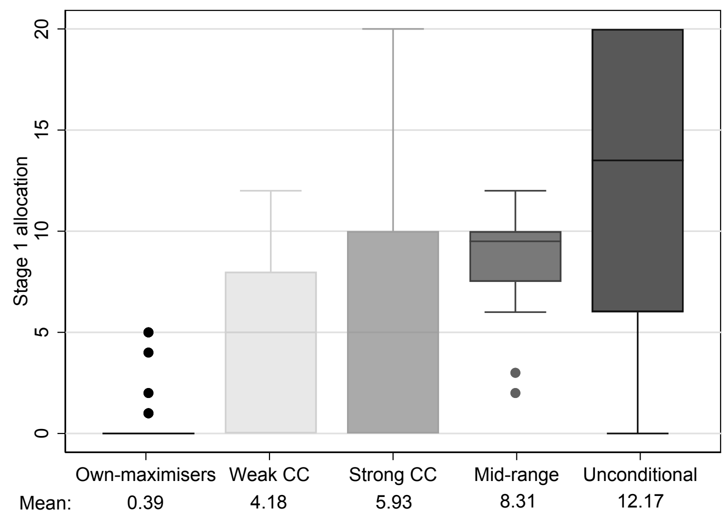


Figure 5: Distribution of Stage 1 allocations in LINEAR, grouped by the behavioural type of the participant's Stage 2 strategy.

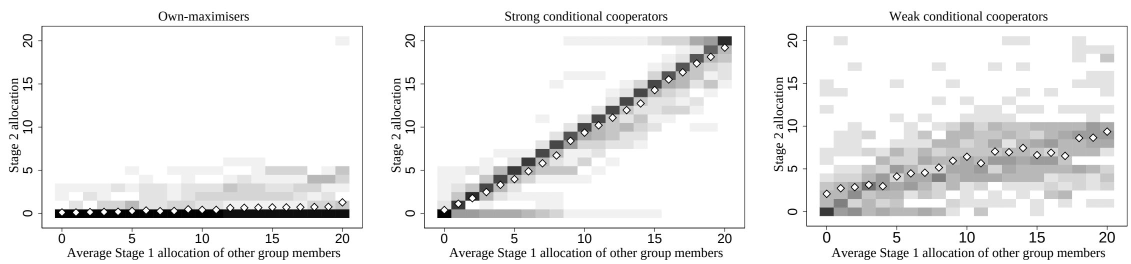
after the Bonferroni correction for multiple comparisons). The contributions of weak and strong conditional cooperators are not significantly different ($p = 0.69$). □

4.2 Coherence of types across games

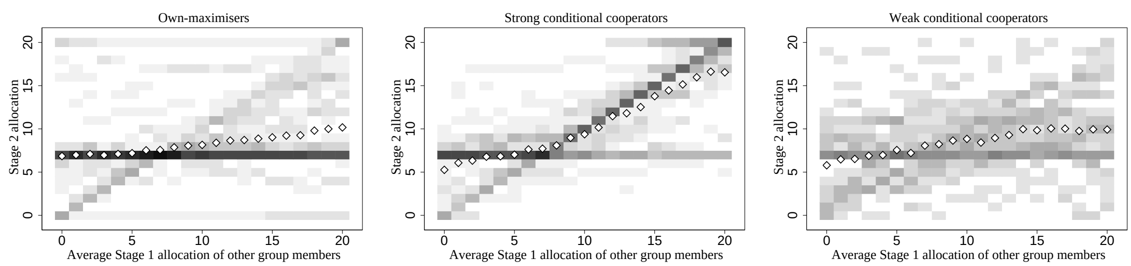
We now investigate how informative are the types identified by clustering on Stage 2 strategies in LINEAR in predicting behaviour in the other three games DOMINANT, SUBSTITUTES, and COMPLEMENTS. We focus first on own-maximisers, weak conditional cooperators, and strong conditional cooperators, who together comprise 83.5% of our participants. In Figure 6, we group participants based on their LINEAR type classification, and produce heatmaps of their Stage 2 behaviour in each of the other three games.

The heatmaps for the group identified as own-maximisers in LINEAR all feature a dark line corresponding to the contributions which maximise individual earnings in each game. That is, by far the modal behaviour for these participants is to choose the contributions which maximise their earnings throughout the experiment. The behaviour of these own-maximisers is therefore consistent across games.

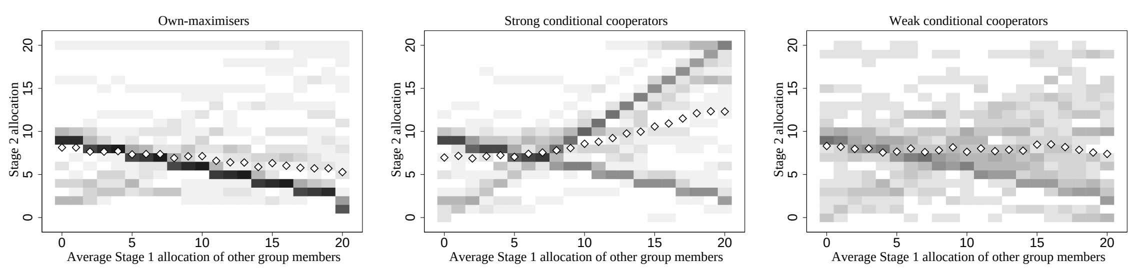
A parallel picture emerges among participants identified as strong conditional cooperators in LINEAR, albeit one which demonstrates an interesting level of sophistication. The majority of strong conditional cooperators match the average contributions of other group members one-for-one, but *only* when doing so results in a contribution above the own-earnings-maximising level. When contemplating low contribution levels by the rest of the group, strong conditional cooperators tend to choose the own-earnings-maximising contribution.



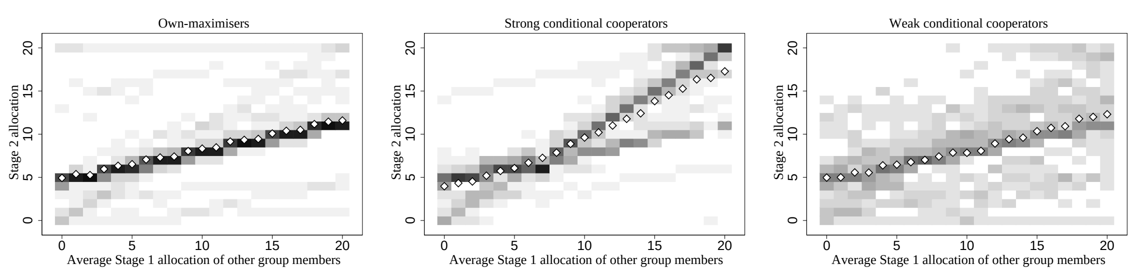
(a) LINEAR



(b) DOMINANT

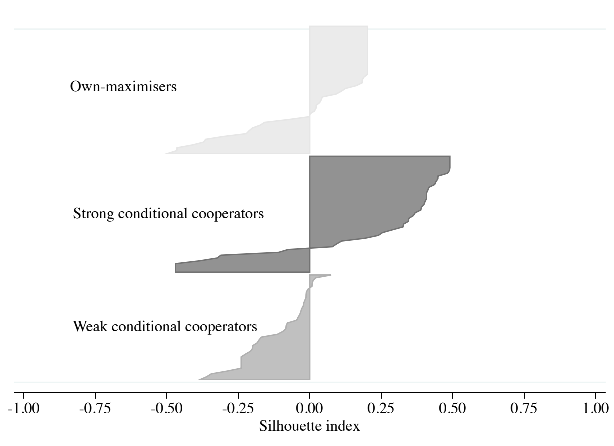


(c) SUBSTITUTES

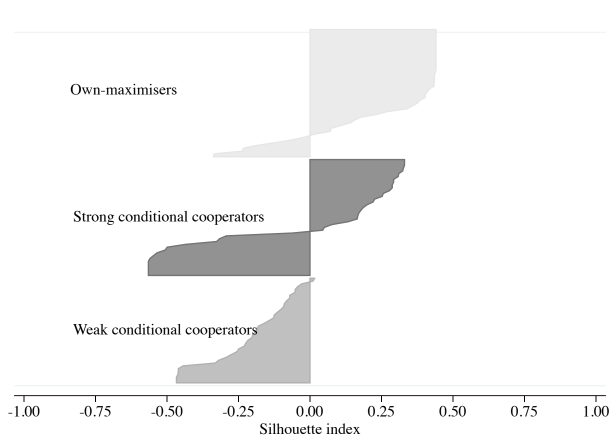


(d) COMPLEMENTS

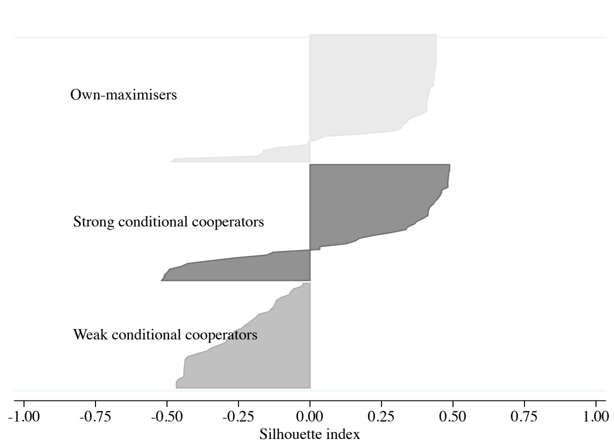
Figure 6: Heatmaps of Stage 2 strategies for allocations to project, grouped by clustering derived from Stage 2 strategies in LINEAR.



(a) DOMINANT



(b) SUBSTITUTES



(c) COMPLEMENTS

Figure 7: Silhouette plots for nonlinear games, using the type clusters derived from Stage 2 strategies in LINEAR.

The within-group consistency and striking sophistication of the strong conditional cooperators contrasts with the behaviour of the group identified as weak conditional cooperators in LINEAR. Weak conditional cooperators in LINEAR continue, on average, to be weakly conditionally cooperative in other games, insofar as the average contribution of the group lies at or above the own-earnings-maximising response. However, the heterogeneity in behaviour of the group of weak conditional cooperators in other games distinguishes them clearly from the groups of own-maximisers and of strong conditional cooperators. Although a fraction of participants become essentially own-maximisers in other games, as evidenced by the visibility of the own-earnings-maximising line in the heatmaps, the heatmaps convey the heterogeneity of this group. Notably, weak conditional cooperators make contributions below the own-earnings-maximising level far more often than those who are classified as own-maximisers or strong conditional cooperators.

In Figure 7 we evaluate the coherence of the clusters generated from the Stage 2 contributions in LINEAR, as applied to the nonlinear games, using silhouette plots. These are generated by taking the LINEAR clustering, but then computing the silhouette indices using the Stage 2 contributions in the respective nonlinear game. The salient feature of these plots is that almost all weak conditional cooperators have negative silhouette indices in the nonlinear games. This means that the Stage 2 contributions of almost all players considered WCCs in LINEAR are closer either to the average behaviour of own-maximisers or the average behaviour of strong conditional cooperators, than to the average behaviour of weak conditional cooperators. A majority of own-maximisers and strong conditional cooperators have positive silhouette indices across the nonlinear games. This out-of-sample application of the LINEAR typology to the nonlinear games provides strong support that own-maximisers and strong conditional cooperators are distinct, coherent, and robust types.

The analysis of Figures 6 and 7 preferentially define types based on the Stage 2 contribution strategies in LINEAR. This is useful in view of the extensive literature which has used the linear payoff specification in public goods games. Nevertheless, the clustering algorithm can be applied separately to the Stage 2 contributions of each the three games with nonlinear payoffs, producing another classification of behaviour based on different data.

We take the 126 participants classified in LINEAR as being own-maximisers, weak conditional cooperators, or strong conditional cooperators. For each nonlinear game, we cluster those participants into three clusters, based only on their Stage 2 strategies in each respective game. If behavioural types were entirely dependent on the details of the game, there would be no carryover of behaviour from one game to the next, and we would expect the assignment of participants to clusters in LINEAR to be independent of the assignment in each other game.

In Table 3 we compare the clustering in LINEAR with the clustering of each nonlinear game. Let $\mathcal{T}(G)$ denote the clusters generated for game G, where formally $t \in \mathcal{T}(G)$ is a subset of the participants. Each table on the left side of Table 3 compares the clusters $\mathcal{T}(G_1)$ and $\mathcal{T}(G_2)$ for a

	D_1	D_2	D_3	Total
OWN	33	11	2	46
	(25.2)	(19.3)	(1.5)	
SCC	8	33	1	42
	(23.0)	(17.7)	(1.3)	
WCC	28	9	1	38
	(20.8)	(16.0)	(1.2)	
Total	69	53	4	
Fisher's exact test $p < .001$				

	OWN	SCC	WCC
OWN	.564***	.674***	.426***
	(.470)	(.529)	(.529)
SCC		.646***	.657***
		(.471)	(.528)
WCC			.563***
			(.472)
Overall: .588*** (random baseline .508)			

(a) LINEAR and DOMINANT

	C_1	C_2	C_3	Total
OWN	37	3	6	46
	(25.2)	(4.7)	(16.1)	
SCC	14	1	27	42
	(23.0)	(4.3)	(14.7)	
WCC	18	9	11	38
	(20.8)	(3.9)	(13.3)	
Total	69	13	44	
Fisher's exact test $p < .001$				

	OWN	SCC	WCC
OWN	.661***	.646***	.566
	(.429)	(.573)	(.572)
SCC		.513**	.650***
		(.429)	(.572)
WCC			.347***
			(.429)
Overall: .590*** (random baseline .524)			

(b) LINEAR and COMPLEMENTS

	S_1	S_2	S_3	Total
OWN	12	28	6	46
	(12.8)	(18.3)	(15.0)	
SCC	6	12	24	42
	(11.7)	(16.7)	(13.7)	
WCC	17	10	11	38
	(10.6)	(15.1)	(12.4)	
Total	35	50	41	
Fisher's exact test $p < .001$				

	OWN	SCC	WCC
OWN	.443***	.714***	.685**
	(.334)	(.664)	(.664)
SCC		.415***	.695***
		(.335)	(.664)
WCC			.336
			(.335)
Overall: .602*** (random baseline .554)			

(c) LINEAR and SUBSTITUTES

Table 3: Consistency of clustering between LINEAR and each other game. In the contingency tables at left, each cell is the number of participants classified in LINEAR to the cluster given by the row label, and in the other game to the cluster given by the column label. Given in parentheses is the expected number of participants under the null hypothesis that the clustering is independent across the games. The tables at the right give the Rand (1971) index for each pair of types. Given in parentheses is the expected index under independent clustering. * denotes a significant difference at 10%, ** at 5%, *** at 1%.

pair of games G_1 and G_2. For a given pair of types $t_1 \in \mathcal{T}(G_1)$ and $t_2 \in \mathcal{T}(G_2)$, each cell in a table gives the number of participants in $t_1 \cap t_2$, that is, the number of participants classified as type t_1 in G_1 and type t_2 in G_2.

Result 3. *There is significant correlation of types between* LINEAR *and each of the three games with nonlinear payoffs.*

Proof. If knowing the type t_1 of a participant based on their behaviour in G_1 gave no information about their behaviour in G_2 (or vice versa), then the distribution of participants to types in G_1 would be independent of that in G_2. To test this null hypothesis, we conduct Fisher's exact test on the joint distribution of type assignments in LINEAR and each nonlinear game. We reject the null hypothesis of independence for each pair at the .001 significance level. □

To quantify the consistency of classification across games relative to the baseline of random assignment, we use the agreement measure proposed by Rand (1971). This measure is constructed by considering each pair of participants i and j. Let $t_i(G_1)$ and $t_j(G_1)$ denote the types of these participants in game G_1, and $t_i(G_2)$ and $t_j(G_2)$ the types of the participants in game G_2. Two partitions are said to agree on the classification of i and j if either (a) $t_i(G_1) = t_j(G_1)$ and $t_i(G_2) = t_j(G_2)$, or (b) $t_i(G_1) \neq t_j(G_1)$ and $t_i(G_2) \neq t_j(G_2)$. Rand's index of agreement is given by the proportion of pairs i and j for which the two clusterings agree. On the right side of Table 3 we report this index comparing LINEAR against each nonlinear game. We also break down this measure conditional on the types of the participants i and j in LINEAR.

The Rand index is useful in understanding the structure of the correlation in the clustering across the games. Strong conditional cooperators in LINEAR remain the core of a cluster in each nonlinear game, forming the majority of clusters D_2 in DOMINANT, C_3 in COMPLEMENTS, and S_3 in SUBSTITUTES. Conditional on two participants being SCC in LINEAR, there is a 64.6% chance the participants are classified as the same type in DOMINANT, 51.3% they are the same type in COMPLEMENTS, and 41.5% they are the same type in SUBSTITUTES. The tables give, in parentheses, the agreement rate that would be obtained if the types in the two games were independently distributed, conditional on the observed number of participants in each cluster.[8] These exceed the random-assignment benchmark by substantial margins. Likewise, own-maximisers in LINEAR tend to classify into the same cluster in each of the nonlinear games.

The weak conditional cooperators in LINEAR are, in contrast, a less coherent group. In DOMINANT weak conditional cooperators are mainly grouped with own-maximisers. Given a participant i from OWN and a participant j from WCC, there is only a 42.6% chance they are classified as different in DOMINANT, which is significantly less than the expected 52.9% under random assignment. In this game the own-maximisers contribute about 7 tokens, irrespective of the behaviour

[8]That is, under the null hypothesis of Fisher's exact test.

of others. Weak conditional cooperators do not distinguish themselves substantially from this behaviour. Although LINEAR and DOMINANT have by definition different payoff structures and therefore caution must be exercised in comparing choices in the strategy spaces of the two games, the presence of the interior own-earnings-maximising contribution may crowd out some of the pro-social instinct of the WCCs.

The clustering in SUBSTITUTES assigns two WCCs from LINEAR to the same cluster just about as often as expected by chance (33.6% in the data versus 33.5% under the random benchmark). In COMPLEMENTS WCCs from LINEAR are actually considered different more often than random chance would predict; WCCs are considered similar in COMPLEMENTS only 34.7% of the time against the benchmark 42.9%.

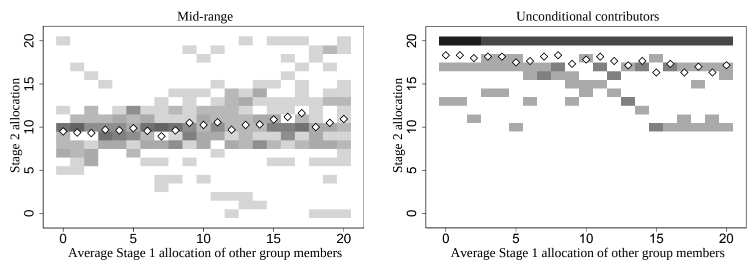
4.3 Mid-range and high contributors

Only 22 participants are classified in the mid-range or high clusters. These participants do not show a systematic response to the anticipated contributions of the other members of their group in LINEAR. The mid-range cluster would include the stereotype strategy of contributing exactly one-half of one's endowment irrespective of the actions of others, while the high contributor cluster would include the strategy of unconditional full contribution, which is the decision that maximises the group earnings from the player's endowment.

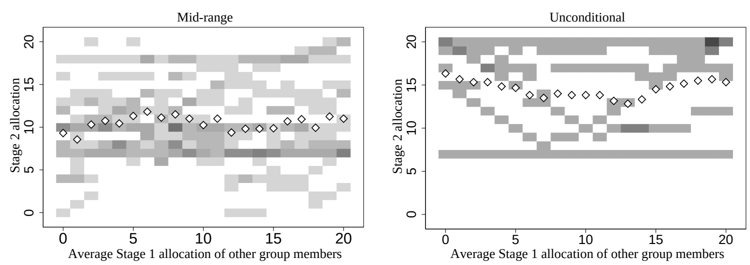
Figure 8 shows the heatmaps of Stage 2 strategies for the mid-range and high clusters. As these are based on small numbers of participants, we offer primarily qualitative judgments. The pattern of Stage 2 contributions among mid-range contributors does not differ noticeably between LINEAR and DOMINANT. These participants do not respond to the increase in the own-earnings maximising strategy from 0 tokens to 7 tokens, which is broadly consistent with an account where mid-range contributors anchor on a 50-50 split of tokens in the strategy space as an appropriate response. There is likewise only a weak suggestion of a response to the financial incentives in games COMPLEMENTS and SUBSTITUTES. The few high contributors in LINEAR likewise contribute generally high amounts in the other games as well. Notable in the heatmaps for both of these types is that the number of contributions below the own-earnings response is relatively small.

4.4 Response times to control questions

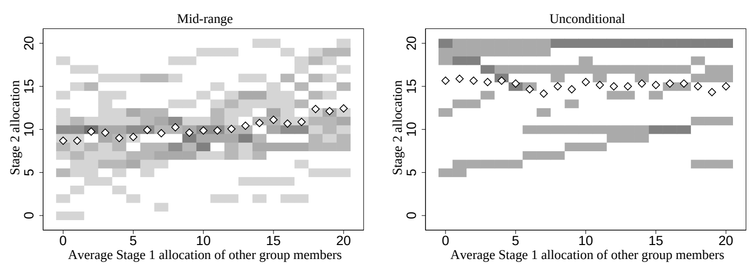
We have observed that strong conditional cooperators, as a group, adopt behaviour which is consistent across games and displays some sophisticated consideration of the financial incentives of the game, especially compared to weak conditional cooperators. This is suggestive that strong conditional cooperators are making a well-informed and conscious decision in forming their Stage 2 strategies.



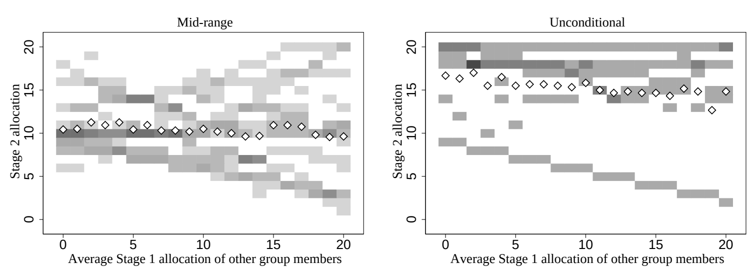
(a) LINEAR



(b) DOMINANT



(c) COMPLEMENTS



(d) SUBSTITUTES

Figure 8: Heatmap of Stage 2 strategies for allocation to the project, for mid-range and high clusters, across games

Type	N	Mean	SE	Quartiles		
Own-maximisers	46	106.9	9.9	56	73	150
Strong conditional cooperators	42	110.9	10.6	60	96	151
Weak conditional cooperators	38	142.9	16.8	70	105	165
Mid-range	16	191.9	28.1	94	162	253
High	6	160.7	36.1	81	136	242

Table 4: Time spent by participants answering control questions, by behavioral type. All times in seconds.

To look for further evidence, we look at the time participants spent reviewing and answering the battery of comprehension control questions at the end of the instructions. Table 4 reports descriptive statistics on the distribution of these times, by behavioural type.

Result 4. *Strong conditional cooperators are not different from own-maximisers in response time to control questions. Strong conditional cooperators and own-maximisers take significantly less time to complete the control questions than other types.*

Support. Own-maximisers on average take 106.9 seconds to complete the control questions, and strong conditional cooperators 110.9 seconds. The lower and upper quartiles of the distribution of response times are likewise similar between the groups. We cannot reject the null hypothesis of these distributions being the same. (Mann-Whitney-Wilcoxon test, p-value 0.69; the probability a randomly chosen SCC's response time is longer than a randomly chosen OWN's response time is .525.)

Comparing the response times of own-maximisers and strong conditional cooperators against the response times of all other participants, we can reject the null hypothesis these distributions being the same. (Mann-Whitney-Wilcoxon test, p-value 0.002; the probability a randomly chosen SCC/OWN's response time is longer than a randomly chosen response time of another type is .351.) □

There are many factors which might feed into how long it takes a participant to complete the control questions. A participant could spend a longer time on the control questions because of one or more incorrect answers, as participants could only continue once they gave a correct response. Participants of different cognitive abilities might need more or less time to process and respond to a question. Some participants with long response times may simply be less engaged with the experimental task.[9] However, in order to complete the control questions in a relatively small length of time, a participant would need to be engaged with the task and provide the correct responses to questions quickly. So, in contrast to the claim of Burton-Chellew et al. (2016), our

[9]The distributions of the completion times for all groups have long right tails.

strong conditional cooperators appear to be as well-engaged and understand the task as well as our own-maximisers.

5 Discussion

We investigate the robustness of pro-social behaviour in social dilemma games by eliciting the behaviour of the same participants across public goods environments with different payoff structures. Following the literature which seeks to classify participants into discrete behavioural types, we first show that our protocol replicates the distribution of types found in other studies using the standard linear VCM payoff scheme. Retaining the same choice architecture but moving to nonlinear payoff structures, the behavioural type classification from the linear game is a strong predictor of the type classification in other games.

The consistency of types across related games is interesting on its own, as it suggests that classifying participants based on their Stage 2 strategies in the conditional choice protocol we use does capture information about individual differences. More significant, however, is how participants of each of the various types react to different incentive structures. It has previously been suggested that conditional cooperators, originally identified as a single behavioural type by e.g. Fischbacher et al. (2001), actually consists of two distinct groups, which we label strong and weak conditional cooperators, respectively. Our data extend this distinction; strong conditional cooperators emerge as a coherent group of sophisticated participants. Weak conditional cooperators, in contrast, are much more heterogeneous, and also more likely to behave similarly to own-earnings maximisers in nonlinear games.

Looking at the behaviour of strong conditional cooperators across this family of games rules out some possible explanations for their behaviour. If strong conditional cooperation were due to conformity (Bardsley and Sausgruber, 2005), we would expect these players to match the average contributions of other group members in all games, irrespective of the financial incentives of the game. However, most strong conditional cooperators respond to the financial incentives, insofar as they do not choose contributions below the own-earnings-maximising amount. For these players, if conformity is a consideration, it is not undertaken blindly or naively, as the one-to-one matching behaviour is primarily observed only when the individual sacrifice is beneficial to the group.

People might be motivated by a "warm glow" (Andreoni, 1990) from the mere act of contributing to the public good. In the linear game, as other group members increase contributions to the public good, the resulting income effect could result in a warm glow-motivated player to increase her own contribution in response. However, in DOMINANT and SUBSTITUTES, we would expect players who valued the act of contribution in its own right to contribute in excess of the own-earnings-maximising amount even when the contributions of others were low; the majority of strong conditional cooperators do not do so.

Strong conditional cooperation in the linear VCM would also be consistent with inequity aversion. (Fehr and Schmidt, 1999) As other players increase contributions from zero, a sufficiently inequity averse player would experience disutility from contributing zero herself, so she would contribute to reduce her advantageous inequity. However, in the nonlinear game, contributing the own-earnings-maximising amount to the project in response to low contributions by others would accentuate, rather than reduce, advantageous inequity.

Zizzo (2010) cites the two-stage conditional contribution procedure we use as a possible example of an experimenter demand effect. The use of the conditional contribution table could suggest to participants that conditionality is an important consideration in playing the game, and therefore make participants more likely to choose Stage 2 strategies of the sort we classify as conditionally cooperative. However, because strong conditional cooperators do not, in general, choose contributions below the own-earnings-maximising amount, strong conditional cooperators are not following such a suggestion dogmatically.

Based on the re-analysis of Fallucchi et al. (2017) and the data in this paper, about 30% of participants in experiments report Stage 2 strategies which are classified as strongly conditionally cooperative. The evidence from our experiment is that a substantial majority of these understand the financial consequences of their actions, and therefore can be presumed to be making an informed decision to behave pro-socially. When the contingency calls for contributions to the public good in excess of the amount that would maximise their own earnings, these participants may be influenced by considerations of conformity or inequity aversion to identify one-for-one matching as the appropriate strategy, from among the large set of possible ways to be express a pro-social attitude. However, our data are not consistent with an account that this mode of pro-social behaviour is due to confusion, misunderstanding, or a lack of engagement with the experimental task. Rather, many strong conditional cooperators are expressing a sophisticated response to the social dilemma of the VCM.

References

James Andreoni. Impure altruism and donations to public goods: A theory of warm-glow giving. *Economic Journal*, 100:464–477, 1990.

James Andreoni. An experimental test of the public-goods crowding-out hypothesis. *American Economic Review*, 83:1317–1327, 1993.

James Andreoni. Warm-glow versus cold-prickle: The effects of positive and negative framing on cooperation in experiments. *Quarterly Journal of Economics*, 60:1–14, 1995a.

James Andreoni. Cooperation in public-goods experiments: Kindness or confusion? *American Economic Review*, 85:891–904, 1995b.

Nicholas Bardsley and Rupert Sausgruber. Conformity and reciprocity in public good provision. *Journal of Economic Psychology*, 26:664–681, 2005.

Olaf Bock, Andreas Nicklisch, and Ingmar Baetge. hRoot: Hamburg registration and organization online tool. WiSo-HH Working Paper Series, Number 1, 2012.

Maxwell N. Burton-Chellew, Claire El Mouden, and Stuart A. West. Conditional cooperation and confusion in public-goods experiments. *Proceedings of the National Academy of Sciences*, 113(5):1291–1296, 2016.

Tadeusz Caliński and Joachim Harabasz. A dendrite method for cluster analysis. *Communications in Statistics*, 3:1–27, 1974.

Timothy N. Cason and Lata Gangadharan. Promoting cooperation in nonlinear social dilemmas through peer punishment. *Experimental Economics*, 18:66–88, 2015.

Kenneth S. Chan, Rob Godby, Stuart Mestelman, and R. Andrew Muller. Crowding-out voluntary contributions to public goods. *Journal of Economic Behavior & Organization*, 48:305–317, 2002.

Ananish Chaudhuri. Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Experimental Economics*, 14:47–83, 2011.

Ananish Chaudhuri, Tirnud Paichayontvijit, et al. Conditional cooperation and voluntary contributions to a public good. *Economics Bulletin*, 3(8):1–14, 2006.

Richard O. Duda and Peter E. Hart. *Pattern Classification and Scene Analysis*, volume 3. John Wiley & Sons, New York, 1973.

Francesco Fallucchi, R. Andrew Luccasen, and Theodore L. Turocy. Identifying discrete behavioural types: A re-analysis of public goods game contributions by hierarchical clustering. CBESS Discussion Paper No 17-01, 2017.

Ernst Fehr and Klaus M. Schmidt. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114:817–868, 1999.

Urs Fischbacher. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10:171–178, 2007.

Urs Fischbacher and Simon Gächter. Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review*, 100:541–546, 2010.

Urs Fischbacher, Simon Gächter, and Ernst Fehr. Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, 71:397–404, 2001.

Simon Gächter, Daniele Nosenzo, Elke Renner, and Martin Sefton. Who makes a good leader? Cooperativeness, optimism, and leading-by-example. *Economic Inquiry*, 50(4):953–967, 2012.

Timothy Gronberg, R. Andrew Luccasen, Theodore L. Turocy, and John Van Huyck. Are tax-financed contributions to a public good completely crowded out? Experimental evidence. *Journal of Public Economics*, 96:596–603, 2012.

Claudia Keser. Voluntary contributions to a public good when partial contribution is a dominant strategy. *Economics Letters*, 50:359–366, 1996.

Stuart Lloyd. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2):129–137, 1982.

William M. Rand. Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*, 66(336):846–850, 1971.

Peter J. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65, 1987.

Martin Sefton and Richard Steinberg. Reward structures in public good experiments. *Journal of Public Economics*, 61(2):263–287, 1996.

Joe H. Ward. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301):236–244, 1963.

Marc Willinger and Anthony Ziegelmeyer. Framing and cooperation in public good games: An experiment with an interior solution. *Economics Letters*, 65:323–328, 1999.

Daniel John Zizzo. Experimenter demand effects in economic experiments. *Experimental Economics*, 13: 75–98, 2010.

A Experimental instructions

Welcome!

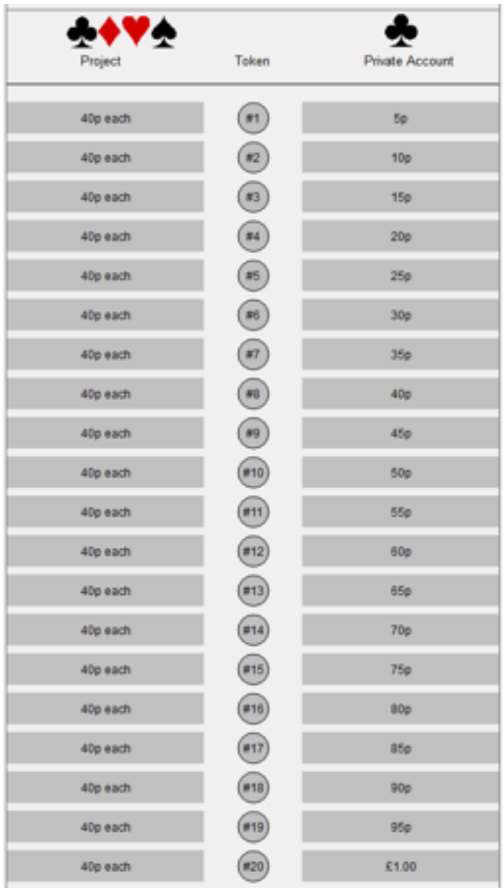
You are about to participate in an experiment in the economics of decision-making. Various agencies have provided funds for this research. If you follow the instructions and make appropriate decisions, you can earn an appreciable amount of money. At the end of today's session you will be paid in private and in cash. These instructions are solely for your private information. You are not allowed to communicate during the experiment. If you have any questions, please raise your hand and a member of the experimental team will come to you.

Groups and decisions

The participants will be divided into groups of four members. You will not know who is in each group, nor will other participants know which group you are in. The four members of each group will be identified by one of the four suits of a standard deck of cards. One member will be identified by clubs (♣), one by diamonds (♢), one by hearts (♡), and one by spades (♠). Your member ID (♣) will not change during the experiment. The experiment will consist of 4 decision rounds. All decisions will be made anonymously.

The decision round

We now describe how each decision round will be conducted. In each decision round, there is a project, which benefits all members of the group equally. In addition, each member has their own private account, which benefits that member alone. Each member receives 20 tokens. Each member will allocate their tokens between the project and their own private account. The tokens will be numbered from #1 through #20. Each token has two possible values, depending on whether the token is allocated to the project or to the private account. Information about these values will be organised for you in a display like this, which we will call an **allocation panel**.



For each token, the value in the Project column (on the left) tells you the value of the token if it is allocated to the project. The value in the Private account column (on the right) tells you the value of the token if it is allocated to your private account.

Earnings from the project

Each group member will profit equally from the tokens you or any other group member allocate to the project. In each decision round, each token you or the other group members allocate to the project will generate an income of 40p to each member of the group. This is indicated in the allocation panel by the text 40p each in the Project column next to each token. The header on the Project side includes all four member IDs (♣♢♡♠), indicating that each member of the group will earn 40p from the token, if that token is allocated to the project.

Your earnings from your private account

You also earn from each token you allocate to your private account. No one except you earns anything from tokens you allocate to your private account. The header on the Private account side includes only your member ID (♣), indicating that only you will receive earnings from the token, if the token is allocated to the private account. The value of each token may be different. In the example, token #20 is worth £1.00 if allocated to the private account; token #19 is worth 95p, and so on. The tokens will be sorted in the allocation panel according to their value in the Private account column, from lower to higher.

Your total earnings

Your total earnings for a decision round will be the sum of your earnings from your private account and from the project:

Your total earnings = Earnings from your private account + Earnings from the project
= Earnings from the tokens you allocated to your private account +
40p $\times$ total number of tokens allocated to the project by
all members of the group

Example 1. Suppose you allocated 18 tokens (#1 through #18) to the project, and 2 tokens (#19 through #20) to your private account. Then, your earnings from your private account would be £1.95, computed by adding up the values in the Private account column for tokens #19 through #20: £1.00 + 95p = £1.95. Each of the 18 tokens allocated to the project would result in earnings from the project of 40p $\times$ 18 = £6.40 for each member of the group, including yourself.

Example 2. Suppose you allocated 14 tokens (#1 through #14) to the project, and 6 tokens (#15 through #20) to your private account. Then, your earnings from your private account would be £5.25, computed as £1.00 + 95p + 90p + 85p + 80p + 75p. Each of the 14 tokens allocated to the project would result in earnings from the project of 40p $\times$ 14 = £5.60 for each member of the group, including yourself.

Summary

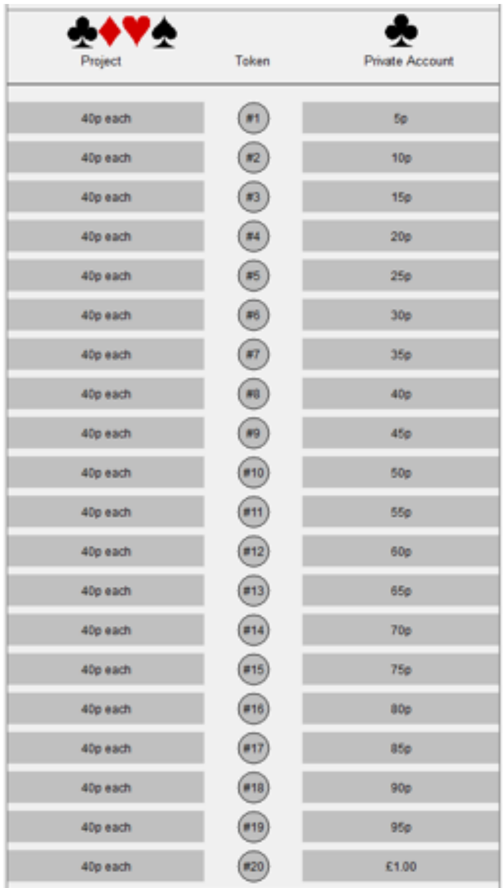
In each decision round, you can derive earnings both from your private account and from the project. Each token you allocate to the project will result in 40p of earnings to each member of your group, including yourself. Tokens allocated to your private account result in earnings for you alone. Each token's value to you, if it is allocated to your private account, will be indicated on your allocation panel in each decision round.

Practice questions

Please answer the questions that will shortly appear on your screen. These will help you to gain an understanding of the calculation of your earnings.

The decision round

There will be four decision rounds. In each round, you will make your allocation decisions on the screen of the computer in front of you. Recall that you will have 20 tokens at your disposal in each round. Each of them can be allocated either to the project or to your private account.



Indicating an allocation

You indicate your allocation decisions using allocation panels. On an allocation panel, you can indicate your decision in one of two ways:

- If you click on a value in the Project column to the left of a token, then all tokens from #1 up through that token are coloured yellow. This indicates these tokens are allocated to the project. Any remaining tokens will be coloured orange; this indicates these tokens are allocated to your private account.

- If you click on a value in the Private account column to the right of a token, then all tokens from #20 down through that token are coloured orange. This indicates these tokens are allocated to your private account. Any remaining tokens will be coloured yellow; this indicates these tokens are allocated to the project.

For example, suppose you click on the value in the Project column to the left of token #6. This allocates the 6 tokens labeled #1 through #6 to the project; the remaining 14 tokens labeled #7 through #20 are allocated to your private account.

This is how the allocation panel would indicate this allocation. Tokens #1 through #6 are

♣♦♥♠ Project	Token	♣ Private Account
40p each	#1	5p
40p each	#2	10p
40p each	#3	15p
40p each	#4	20p
40p each	#5	25p
40p each	#6	30p
40p each	#7	35p
40p each	#8	40p
40p each	#9	45p
40p each	#10	50p
40p each	#11	55p
40p each	#12	60p
40p each	#13	65p
40p each	#14	70p
40p each	#15	75p
40p each	#16	80p
40p each	#17	85p
40p each	#18	90p
40p each	#19	95p
40p each	#20	£1.00

coloured yellow, indicating they are allocated to the project; tokens #7 through #20 are coloured orange, indicating they are allocated to the private account. Each member of the group, including you, would receive 40p from each of the 6 tokens you allocated to the project. You, and you alone, will also receive earnings from the 14 tokens allocated to your private account: 35p from token #7, 40p from token #8 etc. You can also set allocations by clicking on the values in the Private account column. If you click on the value in the Private account column to the right of token #7, then the 14 tokens labeled #7 to #20 are allocated to the private account, and the 6 tokens labeled #1 through #6 are allocated to the project. You can revise your decision on an allocation panel any number of times. You confirm the decisions you indicate on allocation panels by clicking the Confirm button, which will be located at the bottom-right of the screen.

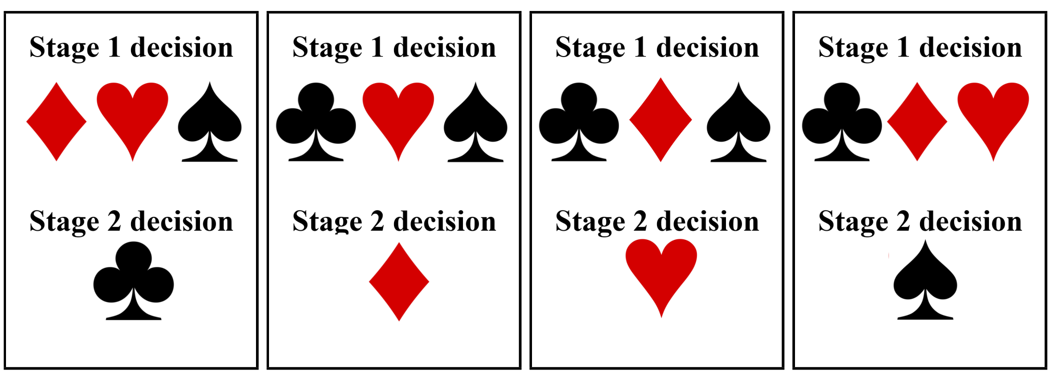
Practice questions

You will now have an opportunity to practice indicating decisions on the allocation panel. There will be three screens. Each screen will ask you to specify a given allocation of tokens to the project and the private account.

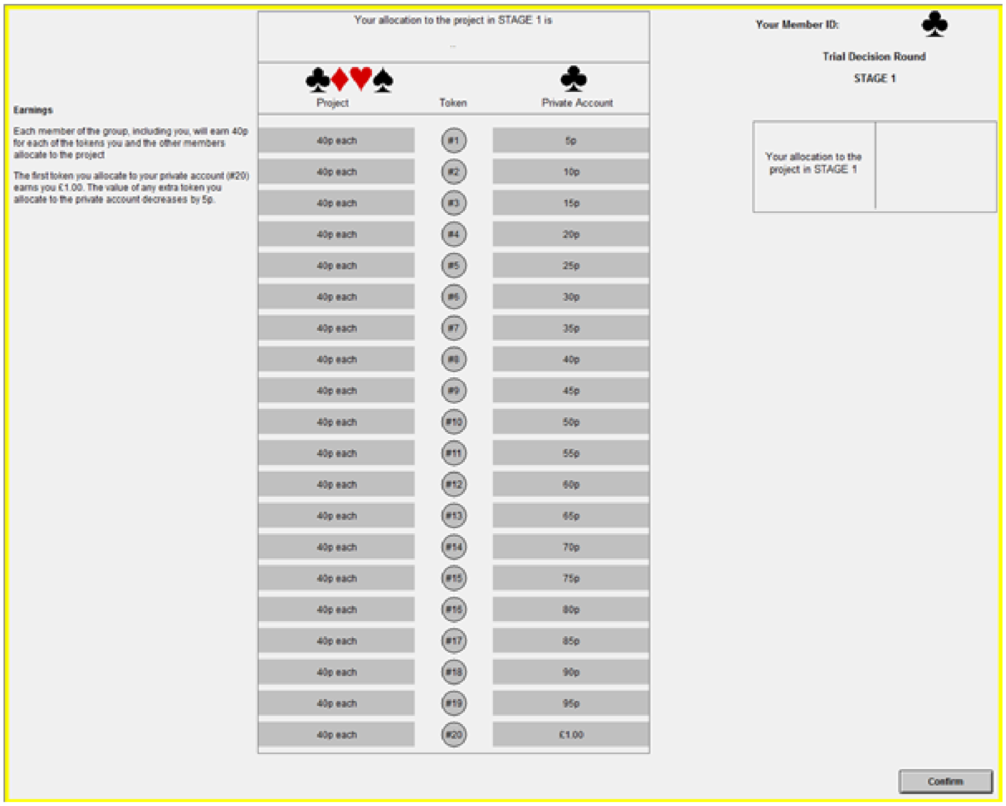
Determining the result of the decision round

The allocations of each group will be determined in two stages, Stage 1 and Stage 2. The allocations of three group members will be determined in Stage 1, and the allocation of the fourth group member will be determined in Stage 2.

Prior to today's experiment, the following four cards were placed, one each, into four unmarked white envelopes, which were sealed.



The participant seated at station 4 will now select one of these four envelopes. This white envelope will only be opened at the end of the experiment, after all decisions have been made. The card in the selected envelope will indicate which group members will have their allocations determined in Stage 1, and which member will have their allocation determined in Stage 2. Because the envelope will not be opened until the end of the experiment, we will ask you to make allocations for both the case in which your ID (♣) is listed under Stage 1, and in which your ID (♣) is listed under Stage 2. Once we open the envelope and reveal which IDs are listed under Stage 1 and Stage

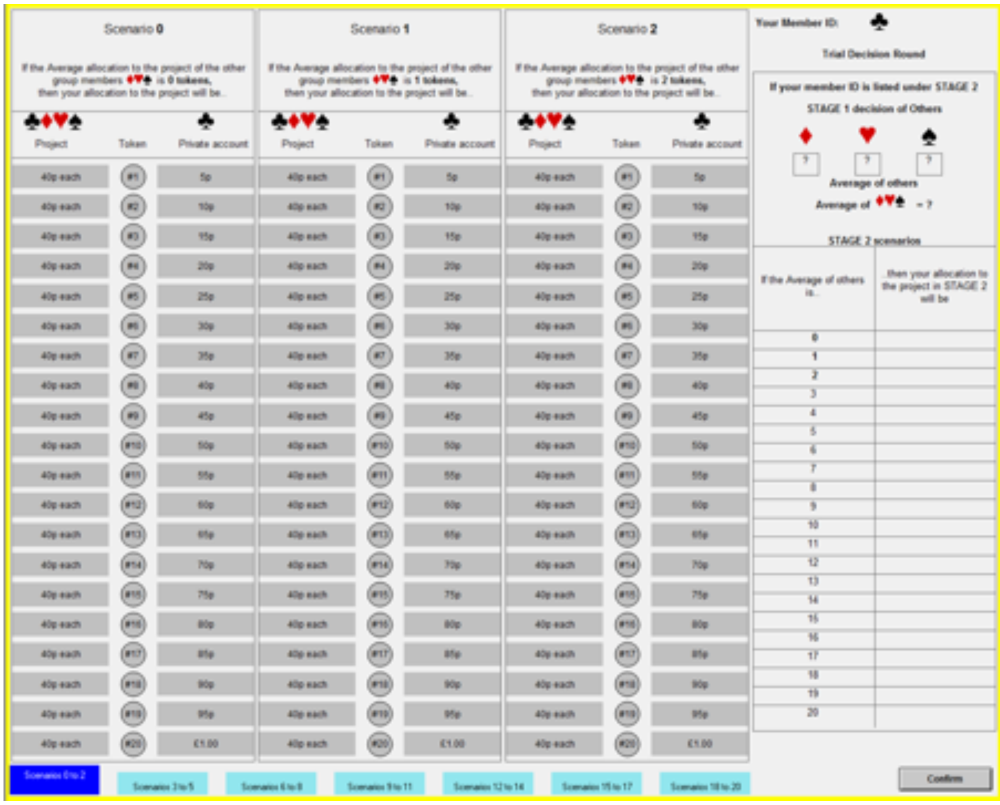


2, the computer will use the decisions that you and the other members of your group indicated to determine the outcome of the decision round.

If your member ID (♣) is listed under Stage 1 on the selected card, then your Stage 1 decision will be the one used to determine your allocation. The Stage 1 decision consists simply of deciding which tokens you allocate to the project, and which to the private account. Here is a sample of what your screen will show when you make your Stage 1 decision in a typical decision round.

Once all participants have indicated and confirmed their Stage 1 decisions for a decision round, we will move to Stage 2.

If your member ID (♣) is listed under Stage 2 on the selected card, then your Stage 2 decision will be the one used to determine your allocation. In the Stage 2 decision, we will ask for your allocation separately for a number of scenarios. There are 21 scenarios, based on the average number of tokens allocated to the project by the other three members of your group () in Stage 1: one scenario in which the average is 0 tokens, one in which it is 1 token, and so on up to 20 tokens. If the average is not a whole number, it will be rounded to the nearest whole number to determine the scenario. Your Stage 2 decision consists of making an allocation for each of these possible scenarios.



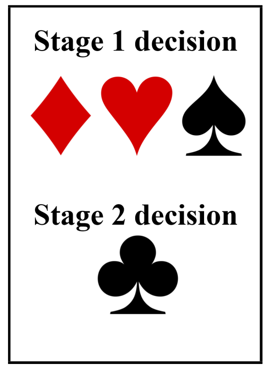
A typical Stage 2 decision screen is shown above. Each allocation panel corresponds to one scenario. The allocation panels for three scenarios are displayed on screen at one time; this screen shows the allocation panels for the scenarios in which the average allocation of the other members (♢♡♠) to the project is 0, 1, or 2 tokens, respectively. To navigate to other scenarios, click the corresponding scenario numbers in the row of buttons at the bottom of the screen. The table on the right side of the screen summarises the allocations you have indicated so far for various scenarios. When you have indicated your allocation for each of the scenarios, click the Confirm button at the bottom-right of the screen to finalise your decision.

After all participants have made and confirmed their Stage 2 allocations, we will proceed to the next decision round. At the end of the experiment, we will open the white envelope to reveal which group members are listed under Stage 1 and which one is under Stage 2. The computer will take the three allocations to the project of the group members listed under Stage 1, compute the average, and round it to the nearest whole number. This will determine the scenario which is relevant for the group member listed under Stage 2. The computer will consult the Stage 2 allocation for that member for that scenario to determine that member's allocation.

Examples

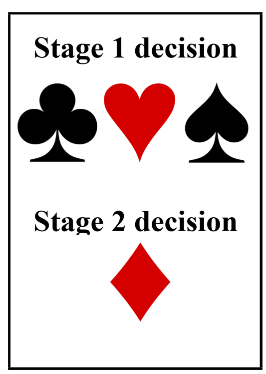
We will now go through two hypothetical examples to illustrate how the Stage 1 and Stage 2 decisions made by you and the other members of your group will be used to determine the allocations in each decision round.

Example 1. Suppose the card below is the card in the selected white envelope:



This card lists your ID (♣) under Stage 2. The computer will then use the Stage 1 allocations to the project of the other three group members (♢♡♠); suppose these are 0, 2, and 4 tokens, respectively. The average allocation of these three group members (♢♡♠) is 2 tokens: 0 + 2 + 4 = 6, which divided by 3 equals 2. This determines the relevant scenario, in which the average Stage 1 allocation of the others is 2 tokens. If you indicated in your Stage 2 decision that in this scenario you would allocate 1 token to the project, then the total allocation to the project is 0 + 2 + 4 + 1 = 7 tokens. Each group member would therefore earn 40p $\times$ 7 = £2.80 from the project. Each member additionally would receive earnings from tokens allocated to their own private account.

Example 2. Suppose the card below is the card in the selected white envelope:



This card lists your ID (♣) under Stage 1. The computer will then use your Stage 1 allocation; suppose it was 16. It will also use the Stage 1 allocations of the other two group members (♡♠) also listed under Stage 1; suppose these were 18 and 20 tokens respectively. The average allocation of these three group members (♣♡♠) would be 18 tokens: 16 + 18 + 20 = 54, divided by 3 equals 18. This determines the relevant scenario, in which the average Stage 1 allocation is 18 tokens. Then, the computer would consult the Stage 2 decision of the member whose ID (♢) is listed under Stage 2 on the card for this scenario. Suppose that in this scenario, this member (♢) indicated that they would allocate 3 tokens. Then the total allocation of the group to the project would be 16 + 18 + 20 + 3 = 57 tokens. Each group member would therefore earn 40p $\times$ 57 = £22.80 from the project. Each member additionally would receive earnings from tokens allocated to their own private account.

How earnings will be determined for the experiment

One of the four decision rounds will be selected at random to determine the earnings you will receive from your decisions. Prior to today's experiment, the numbers 1 through 4 were printed on separate slips of paper, and placed, one each, into four unmarked brown envelopes, which were sealed. We will now ask the participant seated at station 3 to select one of these four envelopes. The number in the selected envelope will determine the decision round used to determine your earnings for the session. This brown envelope will only be opened at the end of the experiment, after all decisions have been made.

Beginning the experiment

If you have any questions please raise your hand and an experimenter will come to your desk to answer.

A.1 Instructions on Screen for LINEAR

Decision Round #

Project

Each member of the group, including you, will earn 40p for each of the tokens you and the other members allocate to the project.

Your Private Account

Each token you allocate to your private account earns you £1.00.

Token	Your Private Account
#1	£1.00
#2	£1.00
#3	£1.00
#4	£1.00
#5	£1.00
#6	£1.00
#7	£1.00
#8	£1.00
#9	£1.00
#10	£1.00
#11	£1.00
#12	£1.00
#13	£1.00
#14	£1.00
#15	£1.00
#16	£1.00
#17	£1.00
#18	£1.00
#19	£1.00
#20	£1.00

A.2 Instructions on Screen for DOMINANT

Decision Round #

Project

Each member of the group, including you, will earn 40p for each of the tokens you and the other members allocate to the project.

Your Private Account

The first token you allocate to your private account earns you £1.15.

The value of each additional token you allocate to your private account decreases by 6p.

Token	Your Private Account
#1	£1.15
#2	£1.09
#3	£1.03
#4	97p
#5	91p
#6	85p
#7	79p
#8	73p
#9	67p
#10	61p
#11	55p
#12	49p
#13	43p
#14	37p
#15	31p
#16	25p
#17	19p
#18	13p
#19	7p
#20	1p

A.3 Instructions on Screen for COMPLEMENTS

Decision Round #

Project

Each member of the group, including you, will earn 40p for each of the tokens you and the other members allocate to the project.

Your Private Account

Your earnings from the first token you allocate to your private account depends on the average allocation of other group members (AO) to the project.

If AO is zero, the first token you allocate to your private account earns you £1.03. This amount increases by 2p for each one-token increase in AO.

The value of each additional token you allocate to your private account decreases by 6p.

Token	Your Private Account
#1	£1.30 - 2p x AO
#2	£1.24 - 2p x AO
#3	£1.18 - 2p x AO
#4	£1.12 - 2p x AO
#5	£1.06 - 2p x AO
#6	£1.00 - 2p x AO
#7	94p - 2p x AO
#8	88p - 2p x AO
#9	82p - 2p x AO
#10	76p - 2p x AO
#11	70p - 2p x AO
#12	64p - 2p x AO
#13	58p - 2p x AO
#14	52p - 2p x AO
#15	46p - 2p x AO
#16	40p - 2p x AO
#17	34p - 2p x AO
#18	28p - 2p x AO
#19	22p - 2p x AO
#20	16p - 2p x AO

A.4 Instructions on Screen for SUBSTITUTES

Decision Round #

Project

Each member of the group, including you, will earn 40p for each of the tokens you and the other members allocate to the project.

Your Private Account

Your earnings from the first token you allocate to your private account depends on the average allocation of other group members (AO) to the project.

If AO is zero, the first token you allocate to your private account earns you £1.30. This amount decreases by 2p for each one-token increase in AO.

The value of each additional token you allocate to your private account decreases by 6p.

Token	Your Private Account
#1	£1.03 + 2p x AO
#2	97p + 2p x AO
#3	91p + 2p x AO
#4	85p + 2p x AO
#5	79p + 2p x AO
#6	73p + 2p x AO
#7	67p + 2p x AO
#8	61p + 2p x AO
#9	55p + 2p x AO
#10	49p + 2p x AO
#11	43p + 2p x AO
#12	37p + 2p x AO
#13	31p + 2p x AO
#14	25p + 2p x AO
#15	19p + 2p x AO
#16	13p + 2p x AO
#17	7p + 2p x AO
#18	1p + 2p x AO
#19	-5p + 2p x AO
#20	-11p + 2p x AO