Centre for Behavioural and Experimental Social Science

University of East Anglia

2020-01

THE SOPHISTICATION OF CONDITIONAL COOPERATORS: EVIDENCE FROM PUBLIC GOODS GAMES

Francesco Fallucchi

R. Andrew Luccansen III

Theodore L. Turocy

www.uea.ac.uk/cbess

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

Francesco Fallucchi
Luxembourg Institute of Socio-Economic Research (LISER)†

R. Andrew Luccasen III
Mississippi University for Women‡

Theodore L. Turocy
University of East Anglia§

February 5, 2020

Abstract

Experiments which elicit preferences for conditional cooperation in public goods games with linear payoffs find that about one-quarter of people approximately match the average contributions of others. To identify from among possible explanations proposed for this strong form of conditional cooperation, we extend the elicitation method of Fischbacher et al. (2001) and study voluntary contributions games with a broader range of economic and strategic incentives. We find that most strong conditional cooperators are sophisticated in responding to these incentives, by matching contributions only when doing so leads to an overall welfare improvement. Our data favour an account of conditional cooperation based on social norm compliance, and are not consistent with accounts in which these people are motivated by inequity aversion or warm-glow giving, or are confused about the economic incentives presented by the elicitation mechanism.

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

JEL Classifications: C72, D62, D71, H41.

*For many helpful comments, we thank seminar participants at University of Alabama, University of Arkansas, University of Nottingham, and Queen's University Belfast; conference participants at the European Meetings of the Economic Science Association and the Thurgau Experimental Economics Meeting; as well as Abigail Barr, Maxwell Burton-Chellew, Marco Faravelli, Simon Gächter, Werner Güth, Daniele Nosenzo and Simone Quercia. Jan Niederreiter provided excellent research assistance. 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.

†Address: 11 Porte des Sciences, 4366 Esch-sur-Alzette, Luxembourg

‡Address: Department of Business, Columbus MS 39701, United States

§Address: School of Economics and Centre for Behavioural and Experimental Social Science, Norwich Research Park, Norwich NR4 7TJ, United Kingdom. Corresponding author: T.Turocy@uea.ac.uk

1 Introduction

People are often willing to make voluntary contributions to public goods, and not infrequently do so in amounts which cannot easily be rationalised by accounting for the private benefits people might receive from their own use of the public good. The observation that many people exhibit some form of prosocial attitudes or behaviour is not novel or surprising. However, most people are not deep-pocketed enough to be able to respond, for example, to every opportunity to give to a cause they deem worthy. We need some criterion to determine which opportunities to respond to, and the appropriate level of contribution.

Fischbacher et al. (2001) and a series of subsequent studies, including Fischbacher and Gächter (2010) and Fischbacher et al. (2012), have proposed that at least some contributions to public goods are motivated by a preference for conditional cooperation. They build on an extensive literature of laboratory and lab-in-the-field experiments using a simple voluntary contribution mechanism (VCM) game. In the standard environment, a person's cash earnings are linear and additively separable in the amount of their own contribution, and the total contributions of others in the (laboratory) economy. In the standard structure, making a voluntary contribution incurs a private financial (opportunity) cost which results in a financial benefit to others in the interaction. Many people make positive contributions in this setting (Chaudhuri, 2011).

One reason a person might choose to take advantage of a given opportunity to make a contribution is because they know, or anticipate, that others will also make contributions to the same public good. Fischbacher et al. (2001) develop a protocol which takes advantage of the control provided by laboratory economy to explore this hypothesis. They ask participants to specify a contribution strategy, that is, how much they would contribute in response to different levels of contributions by the other members of their group. When applied to the environment with linear and additively separable payoffs, most people state they would increase their own contributions in response to larger contributions by others. The magnitude of this responsiveness differs widely across people. Many contribution strategies prescribe a weak response, contributing more than zero but less than others. Fischbacher and Gächter (2010) note this would be consistent with the tendency for contributions to decline over time in both experimental and naturally-occurring settings.

Looking only at the average response masks the existence of several distinct patterns of behaviour. Fallucchi et al. (2019) re-analysed the data from several studies which have used the Fischbacher et al. (2001) protocol, and point out that the distribution of responses is multi-modal. The single most common contribution strategy is to "free ride" by contributing nothing. The next most common strategy is to match the average contributions of others exactly one-for-one. Overall, about one-quarter of people can be characterised as being strongly conditionally cooperative, matching at a rate at or close to one-for-one.

Matching the contributions of others is a simple and salient rule of thumb, with a number of potential justifications. A person may hold the conviction that matching the generosity of others is the socially appropriate convention, generating one-for-one matching as a reasoned response to the economic and social environment. One-for-one matching may be a salient option for selecting from among the many possible contributions which are socially improving for the group. Alternatively, a person might not fully understand the potential consequences of their actions in the interaction; in this case one-for-one matching assumes (or hopes) that others are making informed decisions and their contributions are therefore a guide as to what is a good or appropriate amount to contribute.

Are strong conditional cooperators naïvely and unthinkingly importing a behavioural rule without reference to, or perhaps even understanding of, the details of a given economic environment, or are they making a reasoned response to the incentives they are presented with? This question lies under a more general interest of economists in studying how the distribution of heterogeneous agents affects aggregate outcomes (e.g. Fehr and Tyran, 2005; Bruhin et al., 2018). Since the presence of a relevant share of conditional cooperators may generate important complementarities, we aim at generalising this result by checking the consistency of their behaviour across settings with strategically different structures.

We investigate the nature of strong conditional cooperation by studying how people adapt their contribution strategies across a set of related VCM environments. We include environments in which one-for-one matching leads to socially inefficient outcomes in some situations. The responses of strong conditional cooperators across these different incentive schemes provide evidence to distinguish among possible explanations for their behaviour.

We find that the distinct behavioural types proposed in the analyses of Fischbacher et al. (2001) and Fallucchi et al. (2019) in the linear and additively separable setting are strongly predictive of contribution strategies used by those people in other environments. People who are classified as "free riders" continue to maximise their own earnings. Strong conditional cooperators remain well-distinguished from other types. They exhibit a sophisticated response that incorporates the economic incentives of the environment: they generally follow the one-for-one matching rule only when doing so is *not* Pareto-dominated. The avoidance of Pareto-dominated outcomes is evidence against an account that strong conditional cooperators are unaware of or misunderstand the economic incentives in our environment (Burton-Chellew et al., 2016).

Our experimental design distinguishes sophisticated responses by strong conditional cooperators because the contribution amount which maximises a participant's own earnings is always positive. A contribution strategy which specifies a lower contribution than the own-earnings-maximising amount in some contingency lowers the earnings of both the participant and the group. In two of our environments, the own-earnings-maximising contribution amount varies as a function of the contributions of others, offering a further complexity for people, especially strong conditional cooperators, to take into account. This requires that the consequences of one's own contribution are not separable from those of others, which occurs in real-world public goods settings. 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 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. 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. In this setting, actions are strategic substitutes. Strategic complementarity may arise, for example, on online review sites which 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 customised recommendations. The more reviews a person submits of their own experiences, the more it is possible for the 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 other people are contributing many recommendations, a purely self-interested person would have incentives to contribute many recommendations themselves, to help "train" the system to give them good recommendations for new products to try.

We use the additional economic complexity of our environments to assess the sophistication of people's responses, especially those of strong conditional cooperators. The method of Fischbacher et al. (2001) already adds a degree of complication in the experimental protocol, because it asks participants to specify a contribution strategy instead of a single contribution amount. In our additional environments we provide incentives which, further, are not linear and, in two of the three cases, not separable. To focus on the effects of the complexity of the economic environment, we designed our choice architecture to minimise the cognitive burden that our explanation and presentation of the incentives placed on our participants. We represent the participant's endowment by separately-numbered "tokens", represented virtually on the screen. Each token is labeled with its marginal earnings consequences if allocated to the public good, or if allocated to private consumption. These tokens are presented vertically on the screen, sorted by the difference between their value if retained for private consumption and if contributed to the public good. Participants indicate their preferred allocation by clicking on a token to indicate a division. In our design:

- We present the two options, private consumption and contributions to the public good, symmetrically, and pose the decision as an allocation between the two. Most protocols ask how many tokens to contribute to the project; others ask how many to allocate to private consumption. It is known that how this allocation question is framed can influence behaviour (Andreoni, 1995).

- We follow good interaction design practices by integrating the information about the economic incentives (by labeling each token with earnings amounts) and the decision interface (allocations are made by clicking on one of the tokens) into the same graphical device (the visual representation of the tokens themselves). This contrasts with more standard practices of asking participants to type numbers into an array of text boxes.

- Our presentation communicates the earnings consequences of the allocation for the participant herself and for the group. Previous experiments which used non-linear payoff structures, such as Andreoni (1993), Keser (1996), Chan et al. (2002), and Gronberg et al. (2012), used payoff tables or visualisation of payoff surfaces to explain how the participant's earnings depending on the decisions, while leaving the earnings consequences for the rest of the group implicit.

Our implementation of the choice architecture therefore differs from the one used in the experiments to which we will compare our results. We find, under the linear earnings environment, that the proportions of participants whose strategies reveal own-earnings-maximising, and of participants who match at or about one-to-one the contributions of others, are comparable to those reported previously. Our experiment therefore provides a robustness check that confirms the relative proportions of behavioural types previously claimed in this environment.

The remainder of the paper is structured as follows. In Section 2 we introduce the economic environments and mechanisms used in the experiment, and discuss two theories of behavioural types. In Section 3 we describe the experimental design and the choice architecture. In Section 4 we state the hypotheses which motivate the data analysis and results of Section 5. We conclude in Section 6 with a discussion.

2 Theory

2.1 A public goods environment with linear-quadratic earnings

There are N players, $i = 1, \ldots, N$, whom we refer to collectively as the *group*. Each player i has an endowment $\omega > 0$ of a resource, which we call *tokens*, which she can allocate between a *private account* x_i and a contribution g_i towards a public good, which we call the *project*. Player i's feasible actions are allocations $\mathcal{A}_{\mathbb{R}} \equiv \{(g_i, x_i) \in \mathbb{R}^2_+ : g_i + x_i = \omega\}$. The total amount contributed towards the project by the group is $G \equiv \sum_{j=1}^{N} g_j$, with $G_{-i} = \sum_{j \neq i} g_j$ denoting the total contributions of players other than i. Given private consumption x_i and contributions to the project by other players G_{-i}, the monetary payoff of player i is given by a function $\Pi_i(g_i, x_i, G_{-i})$.

In our experiment monetary payoffs are determined using functions of the form

$$\Pi_i(g_i, x_i, G_{-i}; \beta_1, \beta_2, \lambda) = (\beta_1 - \lambda G_{-i})\, x_i - \beta_2 x_i^2 + 0.4\,[G_{-i} + g_i]\,, \tag{1}$$

where $\beta_1 > 0$ and $\beta_2 \geq 0$.[1] We hold constant the *marginal per-capita return* (MPCR) $\frac{\partial \Pi_i}{\partial G} = 0.4$. By varying β_1, β_2, and λ, we can manipulate the location and slope of the reaction function, given G_{-i}, for a player who wants to maximise her monetary earnings.

When $\beta_2 = \lambda = 0$, we have

$$\Pi_i(g_i, x_i, G_{-i}; \beta_1, 0, 0) = \beta_1 x_i + 0.4 \left[G_{-i} + g_i\right], \tag{2}$$

which is the payoff function for a standard VCM game with linear and additively separable payoffs. The earnings-maximising reaction function for player i is to allocate all tokens to her private consumption,

$$(\tilde{g}_i(G_{-i}), \tilde{x}_i(G_{-i})) = (0, \omega). \tag{3}$$

Note that $\frac{\partial^2 \Pi_i(g_i, x_i; G_{-i})}{\partial x_i^2} - \frac{\partial^2 \Pi_i(g_i, x_i; G_{-i})}{\partial g_i^2} = -\beta_2$, and therefore when $\beta_2 > 0$, earnings are strictly concave in the number of tokens allocated to the private account. The reaction function for player i to maximise her own earnings is

$$(\tilde{g}_i(G_{-i}), \tilde{x}_i(G_{-i})) = \left(\omega - \frac{\beta_1 - \lambda G_{-i} - 0.4}{2\beta_2}, \frac{\beta_1 - \lambda G_{-i} - 0.4}{2\beta_2}\right), \tag{4}$$

when the $\tilde{g}_i(G_{-i})$ so defined is in $[0, \omega]$. The parameter λ captures the degree of complementary or substitutability of contributions, and therefore the slope of the reaction function. When $\lambda = 0$ in (4), the reaction function is constant, as in the specification used by Keser (1996) and Sefton and Steinberg (1996).[2] When $\lambda > 0$, player i wants to contribute more tokens to the project when others are making larger contributions, whereas when $\lambda < 0$ she wants to contribute fewer tokens to the project when others' contributions are higher.[3]

In our experiment tokens are discrete. In passing to the restricted action space $\mathcal{A}_{\mathbb{Z}} \equiv \{(g_i, x_i) \in \mathbb{Z}_+^2 : g_i + x_i = \omega\}$, we observe that the strict concavity (when $\beta_2 > 0$) of $\Pi_i(g_i, x_i, G_{-i})$ with respect to the allocation to the private account ensures that the earnings-maximising allocation is either the integer immediately above or below the reaction function given by (4). In what follows we focus on this discretised case.

Allocation decisions are made in two stages, using an extensive form game introduced by Fischbacher et al. (2001) as the *p-experiment game*. In Stage 1, players $i = 1, \ldots, N-1$ simultaneously and independently choose their allocations (g_i, x_i). Then in Stage 2, the remaining player

[1] Potters and Suetens (2009) used a similar quadratic specification in an experiment with repeated interaction between fixed pairs.

[2] Other studies using a quadratic specification with an interior earnings-maximising dominant strategy are Willinger and Ziegelmeyer (1999), and Gronberg et al. (2012).

[3] Other approaches have been used to generate interior equilibria. Andreoni (1993) used a Cobb-Douglas payoff specification; Cason and Gangadharan (2015) a piecewise-linear specification; and Chan et al. (2002) a quadratic specification with a different structure than ours.

$i = N$ learns the average contribution of those $N-1$ players, rounded to the nearest integer, which we refer to as $\overline{G}$; she then decides her allocation. Therefore, the strategy spaces for players $i = 1, \ldots, N-1$ are the same as the action space, $\mathcal{S}_i = \mathcal{A}_{\mathbb{Z}}$. For player N, the strategy space is $\mathcal{S}_N = \{s : \{0, \ldots, \omega\} \to \mathcal{A}_{\mathbb{Z}}\}$. The rounding involved in determining $\overline{G}$ makes this game formally a game of imperfect information; each level of $\overline{G}$ is an information set. We refer to the component of the action of players $i = 1, \ldots, N-1$ specifying the contribution to the project as the *unconditional contribution* u_i, and the strategy in $\mathcal{S}_N$ specifying the contribution to the project as the *contribution strategy* $c(\cdot)$.

For each game, we identify the set of rationalisable strategies (for players whose objective function is to maximise their own earnings), and the set of perfect Bayesian equilibria in pure strategies. We refer to an equilibrium as symmetric when the unconditional contributions u_i are the same for all $i = 1, \ldots, N-1$.

2.2 Experimental parameterisation

Groups in our experiments consisted of $N = 4$ players. Participants made decisions in four games. In each game participants had an endowment of $\omega = 20$ tokens. Earnings in our baseline game, LINEAR (Γ^L), were determined the same way as in Fischbacher et al. (2001), Fischbacher and Gächter (2010), and Fischbacher et al. (2012),[4]

$$\Pi_i^L(g_i, x_i, G_{-i}) \equiv \Pi_i(g_i, x_i, G_{-i}; 1, 0, 0) = x_i + 0.4\left[G_{-i} + g_i\right],$$

That is, the value of each token allocated by player i to her private account was £1.00, irrespective of how many tokens she allocated or the decisions of others in the group.

We compare participants' decisions in Γ^L with their decisions in three games in which $\beta_2 = .03$. The parameters for each game generate a reaction function for the Stage 2 player which is in the interior of the action space for all values of $u_1 + u_2 + u_3$, and, importantly, the earnings-maximising response for the Stage 2 player is the same for all values of $u_1 + u_2 + u_3$ consistent with each information set $\overline{G}$. Therefore the loss of precision in information about the play of others due to the rounding of average contributions is not strategically relevant for the Stage 2 player. In Appendix A we discuss our choice of parameters and provide a complete analysis of the equilibria of each game for own-earnings-maximising players.

In DOMINANT (Γ^D) earnings are determined by

$$\Pi_i^D(g_i, x_i, G_{-i}) \equiv \Pi_i(g_i, x_i, G_{-i}; 1.18, .03, 0) = 1.18x_i - .03x_i^2 + 0.4\left[G_{-i} + g_i\right].$$

[4] All earnings are expressed in GBP.

This game can be solved by iterated elimination of strictly dominant strategies. The Stage 2 player has a strictly dominant strategy $c^{\star D}(\overline{G}) = 7$ for all $\overline{G}$. Given this, a contribution of $u_i^{\star D} = 7$ is strictly dominant for each player i.

In SUBSTITUTES (Γ^S) earnings are determined by

$$\begin{aligned} \Pi_i^S(g_i, x_i, G_{-i}) &\equiv \\ \Pi_i\left(g_i, x_i, G_{-i}; 1.06, .03, -\frac{.02}{3}\right) &= \left(1.06 + \frac{.02}{3}G_{-i}\right)x_i - .03x_i^2 + 0.4\left[G_{-i} + g_i\right]. \end{aligned}$$

The Stage 2 player has a strictly dominant strategy, which is a nonincreasing strategy $c^{\star S}(\overline{G})$ with $c^{\star S}(0) = 9$ and $c^{\star S}(20) = 4$. The rationalisable unconditional contributions are $3 \leq u_i^S \leq 7$ for $i = 1, 2, 3$. There is a unique symmetric equilibrium with $u_1^{\star S} = u_2^{\star S} = u_3^{\star S} = 7$, with $c^{\star S}(7) = 7$ on the equilibrium path, and asymmetric equilibria with $u_1^{\star S} + u_2^{\star S} + u_3^{\star S} = 13$, with $c^{\star S}(4) = 8$ on the equilibrium path.

In COMPLEMENTS (Γ^C), earnings are determined by

$$\begin{aligned} \Pi_i^C(g_i, x_i, G_{-i}) &\equiv \\ \Pi_i\left(g_i, x_i, G_{-i}; 1.34, .03, +\frac{.02}{3}\right) &= \left(1.34 - \frac{.02}{3}G_{-i}\right)x_i - .03x_i^2 + 0.4\left[G_{-i} + g_i\right]. \end{aligned}$$

The stage 2 player has a strictly dominant strategy, which is a nondecreasing strategy $c^{\star C}(\overline{G})$ with $c^{\star C}(0) = 4$ and $c^{\star C}(20) = 11$. The rationalisable unconditional contributions are $7 \leq u_i^C \leq 10$ for $i = 1, 2, 3$. There is a unique symmetric equilibrium with $u_1^{\star C} = u_2^{\star C} = u_3^{\star C} = 7$, with $c^{\star C}(7) = 7$ on the equilibrium path, and asymmetric equilibria with $u_1^{\star C} + u_2^{\star C} + u_3^{\star C} = 29$, with $c^{\star C}(10) = 8$ on the equilibrium path.

For any fixed G_{-i}, the group's total earnings are always maximised when player i contributes all of her tokens to the project. Furthermore, for games $\gamma \in \{D, S, C\}$, any contribution $g < c^{\star\gamma}(\overline{G})$ by the Stage 2 player is Pareto-dominated by a contribution of $c^{\star\gamma}(\overline{G})$. There are no such *conditionally Pareto-dominated* contribution levels in LINEAR.

2.3 Theories of types in the linear p-experiment game

The p-experiment game is typically combined with the strategy method, which enables elicitation of the full strategy of the player making her allocation in Stage 2. The contribution components of these strategies have been used as the basis for identifying different types of behaviour across participants.

In previous studies using the p-experiment game with linear payoffs, there are only two contribution strategies which are followed exactly by more than a small number of participants: (1)

free-riding (FR), which corresponds to the contribution strategy $c(\overline{G}) = 0$ for all $\overline{G}$, and (2) exact one-for-one matching (OFO), which corresponds to $c(\overline{G}) = \overline{G}$ for all $\overline{G}$ (Fallucchi et al., 2019). Fewer than one-half of participants adopt one of these strategies exactly. Within an experiment, most participants adopt a strategy which is unique among participants in the experiment's sample, although many of these unique strategies differ only in the contributions in a small number of contingencies.

Because of the prevalence of similar-but-not-identical contribution strategies, methods for classifying different strategies into a small number of types have been proposed. There is inherently an element of judgement in dividing the heterogeneous contribution strategies into a small number of types. We therefore consider two type schemata, which we will use jointly to help summarise the contribution strategy data. Each contribution strategy will therefore have a "type" in each schema; which schema is being referenced will be clear from the context.

Fischbacher et al. (2001) proposed a schema (which we call FGF) which classifies strategies into four types. Free-riders (FR) contribute exactly zero in all contingencies, $c(\overline{G}) = 0$. Conditional cooperators (CC) increase their contributions based on higher contributions by others. Formally, participant i is a conditional cooperator if the Spearman's ρ correlation coefficient between the vector $[0, 1, \ldots, \omega]$ of possible average contributions and the participant's contribution strategy $[c(0), c(1), \ldots, c(\omega)]$ is significantly positive with p-value less than some threshold (typically 0.001, which is the value we use in this paper). Hump-shaped (HS) contributors are identified by visually classifying contribution strategies in which $c(0)$ and $c(\omega)$ are zero or small, but $c(\overline{G})$ is larger for some intermediate information sets $0 < \overline{G} < \omega$. Any contribution strategy not matching one of the above criteria is placed in a residual type.

Fallucchi et al. (2019) re-visit data from six p-experiment studies, and use cluster analysis to propose there are five "stereotypical" strategies: own-maximisers (OWN, $\hat{c}^{OWN}(\overline{G}) = 0$), weak conditional cooperators (WCC, $\hat{c}^{WCC}(\overline{G}) = \frac{1}{2}\overline{G}$), strong conditional cooperators (SCC, $\hat{c}^{SCC}(\overline{G}) = \overline{G}$), unconditional high contributors (UNH, $\hat{c}^{UNH}(\overline{G}) = \omega$), and mid-range contributors (MID, $\hat{c}^{MID}(\overline{G}) = \frac{1}{2}\omega$). We use this observation to construct a second schema (which we call FLT).[5] We define the distance between two contribution strategies c and c' using the Manhattan distance,

$$d(c, c') = \sum_{\overline{G}=0}^{\omega} \left| c(\overline{G}) - c'(\overline{G}) \right| . \tag{5}$$

Letting $c^{(j)}$ denote the contribution strategy of a given participant j, the type of the strategy in this

[5] Fallucchi et al. (2019) use the output of their cluster analysis to assign types. A feature of using cluster analysis is that types are defined endogenously based on the full dataset presented to the algorithm, and therefore the classification of strategies on the "border" between types may change as new data are included. The deterministic version presented here is based on their observation that these stereotypical behaviours, which have simple intuitive behaviours, emerge robustly as the centres of mass of types even when resampling the data.

schema, $T^{FLT}(c^{(j)})$, and by extension the participant, is determined by the stereotypical strategy to which $c^{(j)}$ is closest,

$$T^{FLT}(c^{(j)}) = \underset{t \in \{OWN, WCC, SCC, UNH, MID\}}{\arg\min} d(c^{(j)}, \hat{c}^t). \tag{6}$$

3 Experimental design

3.1 Payoff structure treatments

Participants were assigned at random into groups of four. The member identifiers of the group were the four suits of a standard deck of cards (clubs, diamonds, hearts, and spades). The standard 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 (♢♡♠)."[6]

Participants were asked to make their decisions in each of the four games without any feedback on the choices of others or outcomes of any of the games. The games were presented in one of four orderings, which differed across sessions. Games 1 and 3 were always LINEAR and DOMINANT, in either order, and Games 2 and 4 where always COMPLEMENTS and SUBSTITUTES, again in either order.

3.2 Timing of moves

The decisions in each game were elicited using the p-experiment protocol created by Fischbacher et al. (2001).

3.2.1 Step 1: Explain the earnings structure

For each game, the first screen explained to participants how their token allocation would affect their earnings and those of others in their group. This was explained both in a brief prose description, and using a table; the contents of the screens for each of the games are included in a separate Appendix.

We present the allocation task as allocating twenty individually-numbered tokens to either the project or the private account. The structure of the earnings function (1) allows us to express the earnings consequences of the allocation of each individual token. Because the MPCR is held constant at £0.40 for all tokens in all games, the consequence of allocating any token to the project

[6]Complete instructions are available as a separate Appendix.

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, with 6 tokens allocated to the project and 14 to the private account.

is shown as "40p each." The four games vary the consequence of allocating different tokens to the private account. By convention, token #1 was the token which generated the smallest return when allocated to the private account, and token #20 the token which generated the largest return.

3.2.2 Step 2: The Stage 1 allocation

We elicited decisions using the graphical device shown in Figure 1, which we referred to as the *allocation panel*. 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.[7] Participants were able to adjust their allocations as many times as they

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

wished before confirming. Colour-coding was used to indicate the currently-selected allocation; tokens allocated to the project were shown in yellow and those to the private account were shown in orange.

The allocation panel differs from most choice architectures in VCM experiments. With the use of individually-identifiable tokens, the decision is represented as the choice of an allocation, (g, x) in the notation used in Section 2, between the project and the private account. Our expression of the choice as division of tokens in a lab setting is novel.[8] Most experiments elicit decisions by asking participants to type numbers into text boxes. Many VCM experiments, including those of Fischbacher et al. (2001); Fischbacher and Gächter (2010); Fischbacher et al. (2012) to which we compare our results, ask participants to specify only the contribution to the group project; others ask for participants to state how many tokens are retained in the private account or removed from a common resource pool (Gächter et al., 2017). The effects of these choice architecture decisions on allocations have been studied (Brandts and Schwieren, 2009; Dufwenberg et al., 2011; Cubitt et al., 2011; Cox et al., 2013; Cox, 2015; Khadjavi and Lange, 2015; Kingsley, 2015; Cox et al., 2018). When the decision is expressed as a contribution to a public project from an endowment initially allocated to an individual account, contributions are generally higher.

The allocation panel also incorporates information about the consequences of an allocation directly into the graphical instrument used to express the choice.[9] Each token is individually labeled with the consequence of allocating that token to the project. The project column includes the identifiers of all four group members and each consequence in this column includes the word "each." The private account column includes only the identifier of the participant making the decision. This integration is likewise a novelty in LINEAR, but is essential for the other games in which the value of retaining different tokens in the private account is different.

3.2.3 Step 3: The Stage 2 allocation

Figure 2 displays the choice architecture for the Stage 2 allocation strategy, 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*. 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

[8] We acknowledge that this choice architecture is more common in field studies with primary school children (see e.g. Harbaugh and Krause, 2000; Hermes et al., 2019) to provide an easier understanding of the payoff consequences of choices in the linear public good game.

[9] Gronberg et al. (2012) also used a device for making earnings-maximising responses straightforward to discover, but their architecture did not directly represent the social benefits of contributing to the project.

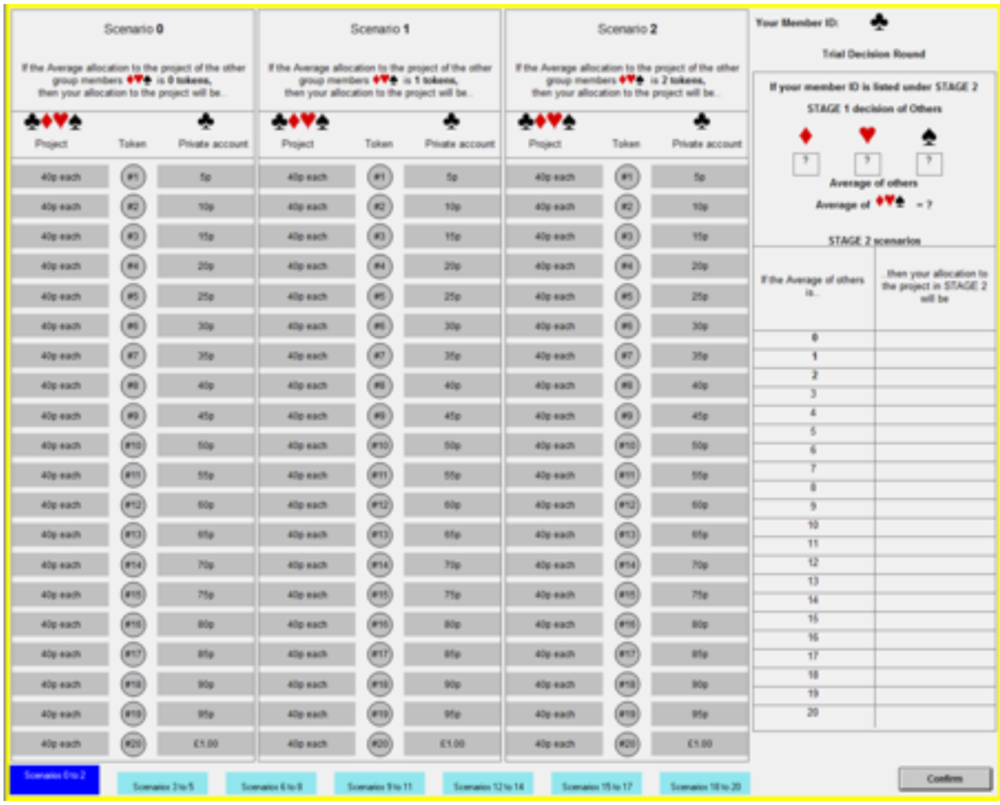


Figure 2: Indication of Stage 2 allocation strategy decisions. Three scenarios were available on the screen at a time. Navigation across scenarios was available using tabs at the bottom of the screen. A panel at the right summarised the allocation decisions made so far.

with the button at the bottom-right of the screen.[10]

3.3 Determination of earnings

One of the four games was selected at random to determine the earnings for the session. At the time participants made their decisions, they did not know which game would be selected, nor whether they would make their decisions in Stage 1 or Stage 2. Prior to any decisions being made, the experimenter placed four cards, one with each member identifier, into four sealed envelopes, and asked a participant to draw one of those envelopes. The sealed envelope was posted on the wall at the front of the lab, but not opened; participants were told the experimenter would open the envelope at the end of the experiment, and the member with the selected identifier would be the one who would play in Stage 2. Then, the experimenter placed four cards, with numbers 1 through 4, into four sealed envelopes, and asked another participant to select one envelope. The sealed envelope was likewise posted on the wall at the front of the lab, but not opened; participants were

[10]Fischbacher et al. (2001) elicit this using an array of 21 text boxes referred to as the "contribution table." In our instructions we simply refer to Stage 1 and Stage 2 choices.

told the experimenter would open the envelope at the end of the experiment, and the corresponding game would be the one which would determine earnings.[11] At the end of the experiment, the experimenter opened both envelopes to determine the game that would be played out, and which group members would make their decisions in Stage 1 and Stage 2, respectively. The software then computed the results of the corresponding game using the participants' decisions to determine earnings.

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. We recruited 148 participants from the standing participant pool, maintained using the hRoot system. (Bock et al., 2014) 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 interquartile range of £5.34.[12]

4 Hypotheses

We designed our choice architecture to organise and communicate the participants' financial incentives in each game in a standard way, accommodating the presence of a nonlinear payoff function without requiring participants to do extensive calculations. This new design allows us to test the robustness of behaviour in the linear VCM using the p-experiment, when using a different method of eliciting the decisions.

Hypothesis 1. *The proportions of types of contribution strategies in* LINEAR *will be the same in our experiment as in previously-reported experiments.*

Hypothesis 2. *The distribution of unconditional contributions in* LINEAR *will be the same in our experiment as in previously-reported experiments.*

In re-analysing previous linear VCM experiments using the p-experiment protocol, Fallucchi et al. (2019) identified the two most common types of Stage 2 strategies as own-maximisers (25.8%) and strong conditional cooperators (38.8%). Making decisions to maximise one's own expected earnings is commonly observed across experiments in economics, especially in one-shot

[11] Both the game and the roles are determined at random in this experiment and not revealed until the end of the session. With these two layers of random selection, we found it easier to write the instructions clearly and concretely by following this approach of pre-selecting the envelopes but keeping them visible but sealed, as we could then refer specifically to the contents of the selected but still unknown card in the posted envelope.

[12] For comparison, the living wage in the United Kingdom at the time of the experiments was £8.25 per hour.

interactions such as the games in our experiment. As such it is a reasonable presumption that own-maximisers in LINEAR do so because the understand the economic context in the game and are expressing an informed decision. Because the design of the choice architecture expresses the financial incentives of all games in a parallel way, we expect these participants to choose contribution strategies which maximise their own earnings across the other games as well.

Hypothesis 3. *Participants who are identified as own-maximisers in* LINEAR *will exhibit own-earnings-maximisation across all games; that is, they will use contribution strategies given by the version of (4) restricted to discrete integer choices.*

In contrast, different accounts have been given for the behaviour of strong conditional co-operators. Participants might match the contributions of others one-to-one because they do not understand the economics of the game as given by the financial incentives, or because they are ignoring those incentives in favour of a rule of thumb transferred from a different context. In each of DOMINANT, SUBSTITUTES, and COMPLEMENTS, following a one-to-one rule of thumb when others contribute below the Nash level would lead to conditionally Pareto-dominated contributions. If strong conditional cooperators were to match one-to-one for contributions lower than the Nash level, this would implicate confusion or heuristic transfer.

Hypothesis 4. *Participants who are identified as strong conditional cooperators in* LINEAR *will follow the same rule of thumb of (approximate) one-to-one matching across all games; this will result in conditionally Pareto-dominated choices in some contingencies.*

5 Results

5.1 Behaviour in the linear VCM

We benchmark our LINEAR data against the series of studies, which use the p-experimental protocol, by Fischbacher et al. (2001), Fischbacher and Gächter (2010), and Fischbacher et al. (2012) (which we refer to as the "Fischbacher sample").

To get a handle on whether our contribution strategies are similar to those in the Fischbacher sample, we classify behaviour according to the two type schemata, FGF and FLT, introduced in Section 2. In Table 1 we report classifications based on the two approaches.[13] Although matching exactly one-for-one is the second most common contribution strategy, it is not a type in its own

[13] We do not report proportions of "hump-shaped" (HS) contributors. Among the 37 participants who do not satisfy the criteria for FR or CC, none exhibit a clearly hump-shaped pattern. The presence of any clearly HS strategies would be evident in the heatmaps in Figure 3 and Figure 6. The absence of HS contribution strategies in our data is a notable difference from most previous studies.

Study	N	FGF		OFO	FLT				
		FR	CC		OWN	WCC	SCC	UNH	MID
Our data	148	22.3	52.7	7.4	32.4	22.3	27.0	2.7	15.5
Fischbacher et al. (2001)	44	29.5	50.0	9.1	43.2	20.5	27.3	2.3	6.8
Fischbacher et al. (2012)	136	14.7	70.6	11.0	35.3	21.3	38.2	2.2	2.9
Fischbacher and Gächter (2010)	140	22.9	52.1	9.3	42.1	20.0	32.9	2.1	2.9

Table 1: Type classifications based on contribution strategies in LINEAR.

right in either schema (it is a subset of CC in FGF and SCC in FLT). We therefore report the proportion of these separately as OFO.

Result 1. *The proportions of strategies which exactly or approximately maximise the participant's own earnings, and the proportions of strategies which exactly or approximately match the average contributions of the group, are similar in our data and the Fischbacher sample.*

Support. Contribution strategies which exactly or approximately maximise the participant's earnings appear with similar frequencies. For exact maximisation, we find a proportion of FR similar to the Fischbacher sample (22.3 compared to 20.3, $p = .62$ using the binomial test). Relaxing to approximate maximisation, our proportion of OWN is also similar (32.4 compared to 38.8; $p = .14$).

Exact one-for-one matching occurs at a similar rate (7.4 compared to 10.0, $p = .37$), as does the more relaxed criterion of strong conditional cooperation (27.0 compared to 34.1, $p = .13$).[14] □

Result 2. *The distribution of unconditional contributions in our data differs from the Fischbacher sample. In our experiment, fewer participants make a positive unconditional contribution, and the overall amounts contributed in Stage 1 are also lower. We observe a higher frequency of zero contributions, and typically lower contributions overall. These differences are consistent at type level.*

Support. Using the Mann-Whitney-Wilcoxon (MWW) test, the distribution of our unconditional contributions differs from the Fischbacher sample ($r = .385$, $p < .001$).[15] The proportion of participants contributing a positive amount is lower in our data ($z = 3.71$, $p < .001$). The increase

[14] If we consider all types, our type distribution is similar to the Fischbacher sample under the FGF classification (χ^2 test, $p = 0.329$), but differs under the FLT classification (χ^2 test, $p = 0.004$). The differences arise from contribution strategies which are HS in the Fischbacher sample, which are classified as OWN, WCC, or SCC in FLT, and the existence in our data of participants who divide tokens more or less equally between the private account and the project, who do not feature in the Fischbacher sample.

[15] For MWW tests we report the test statistic in terms of the effect size r, which is defined as the probability a randomly-selected observation in the first-named sample is greater than a randomly-selected observation in the second-named sample, with ties broken equiprobably.

Type	Sample	N	All participants		Positive contributions only			
			Median	Mean	% Positive	N	Median	Mean
FR	Our data	33	0.0	0.1	2%	2	1.5	1.5
	Fischbacher	65	0.0	1.8	22%	14	5.5	8.4
OFO	Our data	11	0.0	2.9	36%	4	8.5	8.0
	Fischbacher	32	7.5	8.6	81%	26	10.0	10.5
CC	Our data	78	5.0	4.8	59%	46	7.5	8.2
	Fischbacher	191	8.0	8.4	82%	157	9.0	9.8
OWN	Our data	48	0.0	0.5	15%	7	3.0	3.1
	Fischbacher	124	0.0	3.2	43%	53	5.0	7.5
WCC	Our data	33	5.0	4.1	67%	22	6.0	6.1
	Fischbacher	65	7.0	7.1	92%	60	8.0	7.8
SCC	Our data	40	5.5	6.1	65%	26	9.0	9.4
	Fischbacher	109	10.0	10.2	89%	97	10.0	11.4

Table 2: Stage 1 contributions by various Stage 2 types.

in zero contributions accounts for much although not all of the lower contributions in our data. Conditional on contributing a positive amount, contributions in our data are also somewhat lower (MWW, $p = .055$, $r = .429$).

Although we observe broadly similar contribution strategies in our experiment as in the Fischbacher sample, we found very different unconditional contributions. We ask whether the difference in our unconditional contribution data is driven by a certain type or types. Table 2 summarises the distribution of contributions of the main types in Stage 1 for these three studies and our experiment. Stage 1 contributions are lower in our data type-for-type. Of particular note are the Stage 1 contributions for exact free-riders; in our data only 6% (2 of 33) of these participants contribute a positive amount, in contrast to 22% in the Fischbacher sample. Strong conditional cooperators, and in particular the one-for-one subset of them, contribute about half as much in our study as in the Fischbacher sample. □

5.2 Prediction of contribution strategies

Do the types identified by the contribution strategies chosen in LINEAR predict the contribution strategies in games DOMINANT, SUBSTITUTES, and COMPLEMENTS? We focus first on OWN, WCC, and SCC, who together comprise 83.5% of our participants.

To give an initial visual summary of the data, we use the method developed in Fallucchi et al. (2019) and construct heatmaps for the contribution strategies of each type in each game. Let $T(j)$ denote the type classification of a given participant j. The heatmap for a type t is produced by taking the contribution strategies of all participants assigned to that type, and constructing the

	OWN vs. SCC	SCC vs. WCC	OWN vs. WCC
LINEAR	88.861 (<0.001)	70.889 (<0.001)	69.877 (<0.001)
DOMINANT	40.543 (0.006)	41.294 (0.005)	25.229 (0.237)
COMPLEMENTS	50.173 (0.015)	38.380 (0.012)	20.557 (0.486)
SUBSTITUTES	38.724 (0.012)	33.120 (0.045)	34.292 (0.034)

Table 3: Test statistics for Oja (2010) location test of the difference in the contribution vectors between types. p-values are reported in parentheses.

multiset $\{(\overline{G}, c^{(j)}(\overline{G}))\}_{j:T(j)=t,\overline{G}=0,\dots,20}$. The frequencies of these ordered pairs are used to generate the heatmap. Cells with darker shades correspond higher frequencies; the modal behaviour for any given information set $\overline{G}$ can therefore be identified by the darkest cells.

For each type t we also define the medoid strategy $\bar{c}(t)$ as the strategy with the smallest average distance from all the strategies in the type,

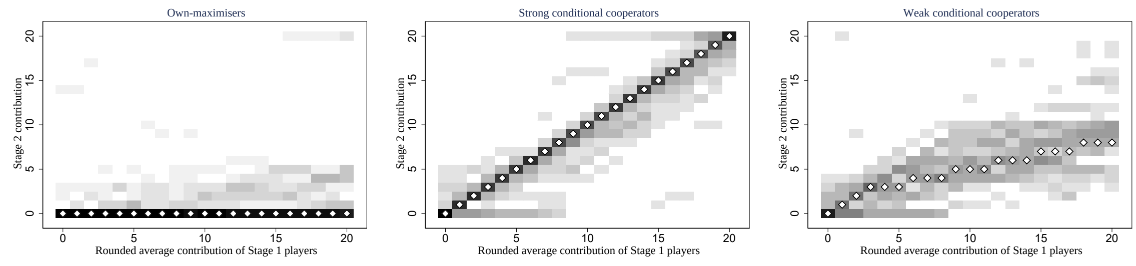
$$\bar{c}(t) = \underset{\{c^{(j)}:T(j)=t\}}{\arg\min} \frac{1}{|\{j : T(j) = t\}|} \sum_{k:T(k)=t} d\left(c^{(j)}, c^{(k)}\right). \tag{7}$$

The medoid for type t is always a strategy that was chosen by at least one participant classified as type t. It coincides with the more familiar centroid when the centroid is a member of the set $\{c^{(j)} : T(j) = t\}$. The medoid strategy is one way to express a "most typical" strategy for the type, and is plotted using small white diamonds in the heatmaps.

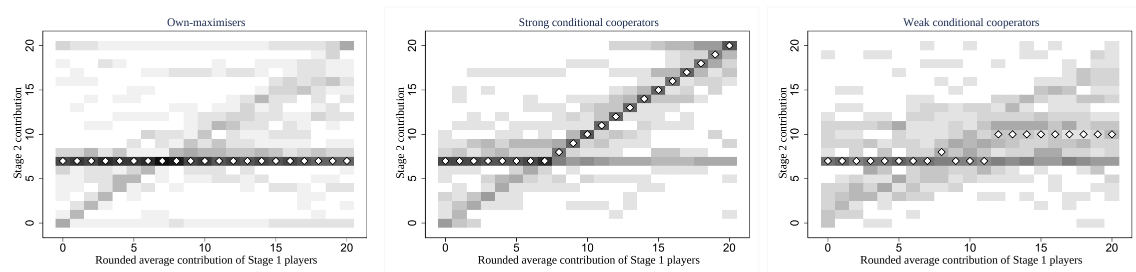
In Figure 3, we set each participant i's type $T(i)$ as their type determined by their contribution strategy in LINEAR. Then, for each type t and for each game Γ, we take the participants classified as type t and use their contribution strategies in Γ to construct the heatmap for type t in game Γ.

For participants classified as own-maximisers in LINEAR, the medoid contribution strategy in each of the three nonlinear games is to contribute exactly the own-maximising number of tokens in every contingency, except when other participants contribute 20 in SUBSTITUTES. The contributions of own-maximisers are typically at or close to the contribution strategy given by the reaction function (4).

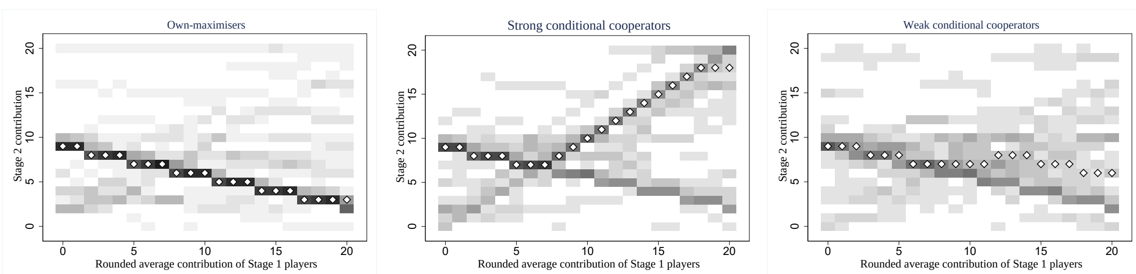
Participants identified as strong conditional cooperators in LINEAR also show a consistent pattern across the other games. The medoid contribution strategies in the nonlinear games match average contributions at or very near one-for-one, but - importantly - *only* when doing so is socially improving. When contemplating possible low levels of contributions by the rest of the group, the medoid contribution strategy of strong conditional cooperators selects the own-earnings-maximising contribution. This is particularly striking in SUBSTITUTES, as this results in a non-monotonic contribution strategy.



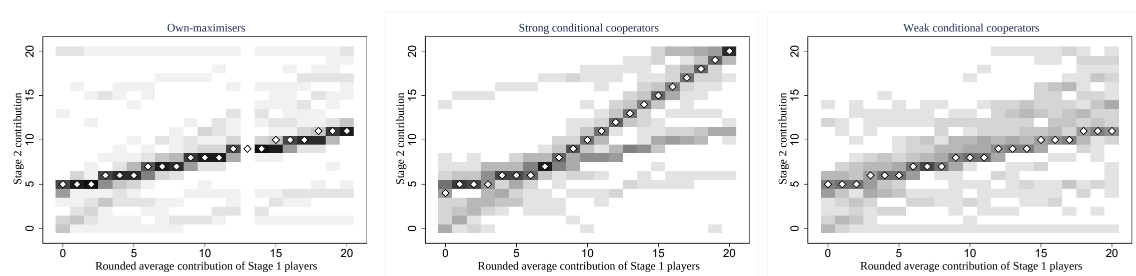
(a) LINEAR



(b) DOMINANT



(c) SUBSTITUTES



(d) COMPLEMENTS

Figure 3: Heatmaps of Stage 2 strategies for allocations to project, derived from Stage 2 strategies in LINEAR and grouped by FLT stereotypical strategies.

Result 3. *The type identifications in* LINEAR *are predictive of behaviour in other games.*

Support. The medoid contribution strategies of OWN and SCC are distinct across all four games, and contribution strategies for both types cluster primarily near the medoid. For OWN, the medoid contribution strategy is almost exactly the own-earnings-maximising reaction function in all games. For SCC, the medoid contribution strategy is almost exactly to follow one-for-one matching above the Nash equilibrium contribution level; below it, the medoid strategy chooses the own-earnings-maximising contribution level, and avoids Pareto-dominated contribution levels.[16] In contrast, the distribution of contribution strategies for WCC is much more dispersed in all games. The medoid contribution strategy for WCC in DOMINANT and SUBSTITUTES prescribes contributing a few more tokens than the own-earnings-maximising amount when responding to group contributions above the Nash level, while in COMPLEMENTS the medoid contribution strategy is exactly the own-earnings-maximising reaction function.

To formalise the discussion, we are testing a hypothesis about whether the distributions of contribution strategies are different among these three groups of participants. Distributions over strategies can be quantified in various ways; we therefore take two approaches to testing for differences in these distributions. Our first approach uses a non-parametric test proposed by Oja (2010) based on spatial signed ranks. The null hypothesis is that the treatment difference between samples is equal to the zero vector. Let $\mathcal{T}$ be the set of types being considered, and n_t be the number of participants in type $t \in \mathcal{T}$. Let R be the vector of centred rank scores, with elements $R_i = \sum_{c^{(j)}} \frac{1}{n} \left(\frac{c^{(i)} - c^{(j)}}{d(c^{(i)}, c^{(j)})} \right)$. The average centred rank score for type t is then $\overline{R}_t = \sum_{i:c^{(i)} \in t} \frac{1}{n_t} R_i$. $\hat{B}$ is the covariance matrix given by RR'. To test the null hypothesis that the true difference in ranks is the zero vector, form the test statistic

$$Q^2 = \sum_{t \in \mathcal{T}} n_t \overline{R}_t' \hat{B}^{-1} \overline{R}_t. \tag{8}$$

The limiting distribution of Q^2 is $\chi^2_{(|\mathcal{T}|-1)k}$, where k is the number of dimensions in the data.[17]

We report in Table 3 the results of pairwise comparisons among OWN, SCC, and WCC for each game. SCC are different in all games from OWN (all $p \leq 0.015$) and from WCC (all $p \leq 0.045$). WCC, however, are not well-distinguished from OWN in either DOMINANT or COMPLEMENTS; in the latter, as noted the upward-sloping reaction function is also the medoid contribution strategy for WCC. □

[16] We say "almost exactly" because both OWN and SCC vary slightly from these descriptions when $\overline{G} = 20$ in SUBSTITUTES, and SCC when $\overline{G} = 0$ in COMPLEMENTS. These are information sets which would be reached with very small probabilities given the empirical distributions of unconditional contributions (see Section 5.3).

[17] Unlike in the univariate case, in the multivariate analysis there are no natural orderings of the data points. See Oja (2010) for an overview of the different rules to rank observations using the Manhattan distance.

The advantage of the analysis above is that it takes the whole contribution strategy as the observation. It allows us to quantify whether strategies are different across types, but not at which information sets $\overline{G}$ they differ. Note, however, that the medoids in Figure 3 in all games below the Nash equilibrium outcome are similar for among OWN, SCC and WCC, while the stereotypical contribution patterns differ across the three types in information sets $\overline{G}$ above the equilibrium level.

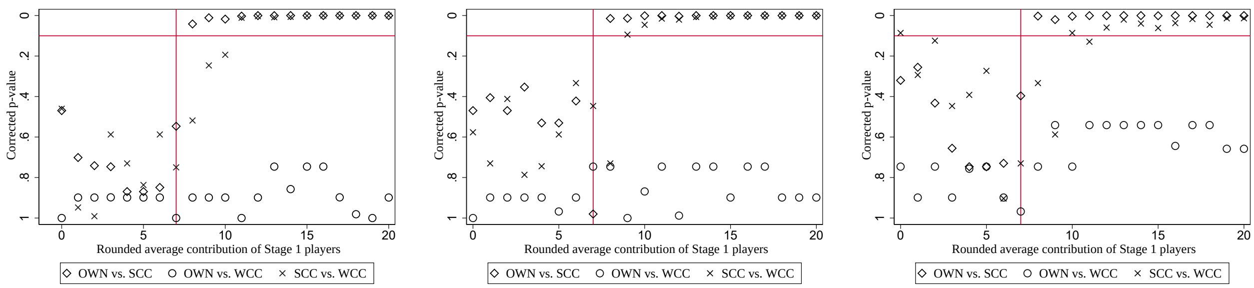
Result 4. *The majority of participants in all three main types do not choose conditionally Pareto-dominated contributions.*

Support. If we look at the heatmaps in Figure 3 we note that some of the choices, and in particular for information sets $\overline{G}$ below the equilibrium level, are inefficient. Overall the subjects making these choices range from 8 in COMPLEMENTS to 15 in DOMINANT and 16 in SUBSTITUTES. Of these subjects, half of them are classified as WCC, while the remaining half is equally distributed between OWN and SCC. However, most of these are just one step below the own-earnings maximisation contribution. We identify only four subjects (3% of the sample) that systematically make Pareto-dominated choices in all the three games, two SCC and one for each of the other two types. The contribution patterns for information sets $\overline{G}$ below the equilibrium level should be similar across types, as already hinted by looking at the heatmaps.

To test for this and for differences between types at any level of the information set, we follow Barr et al. (2018) and perform, for each information set $\overline{G}$, MWW tests comparing the distributions of contributions between each pair of types. We report in Figure 4 the p-values of these tests, corrected for the fact we are performing multiple tests.[18] OWN and SCC are distinguished at all information sets $\overline{G}$ above the Nash equilibrium level. SCC and WCC are distinguished when the contributions of the group exceed the Nash by more than a few tokens; the distinction is weaker in SUBSTITUTES, in which the reaction function is upward-sloping. □

We observe generally consistent behaviour among strong conditional cooperators across all games, as well as a clear distinction of strong conditional cooperators from weak conditional cooperators. Extensions to the standard model such as pure altruism or inequity aversion are not able to account for the patterns we observe in the data. A recently-proposed model, *social cooperation norm compliance*, does help to rationalise our results. Fehr and Schurtenberger (2018) propose a model which incorporates preferences for complying with a norm, following an idea of e.g. Elster (1989). In the case of the conditional contribution decision in the p-experiment, the rounded average contribution $\overline{G}$ of other players might establish such a norm. A player in Stage 2 who gives

[18]Specifically, we apply the Benjamini-Hochberg False Discovery Rate method (Simes, 1986; Benjamini and Hochberg, 1995). We sort the p-values in ascending rank, divide them by the rank and multiply for the number of multiple tests performed. We report the list of corrected p-values in Table 7 in Appendix B.



(a) DOMINANT (b) COMPLEMENTS (c) SUBSTITUTES

Figure 4: Results of pairwise comparisons of contribution strategies by information set $\overline{G}$. Each point is the p-value of a Mann-Whitney-Wilcoxon test; these are adjusted for multiple testing using the Benjamini-Hochberg False Discovery Rate method. The vertical reference line is at $\overline{G} = 7$, below which differences among the types are not expected. The horizontal reference line is set at $p = 0.10$.

consideration to social cooperation norm compliance would have utility

$$U_N(g_N, x_N, \overline{G}; \rho) = \begin{cases} \Pi_N(g_N, x_N, \overline{G}) - \rho\left(x_N - \overline{G}\right)^2 & \text{if } g_N < \overline{G} \\ \Pi_N(g_N, x_N, \overline{G}) & \text{if } g_N \geq \overline{G} \end{cases} \tag{9}$$

The parameter $\rho \geq 0$ captures the strength of any psychological costs that the player incurs by contributing *less* than the amount prescribed by the norm set by $\overline{G}$. Denote the best response contribution in game γ for a player with utility of the form (9) as $\tilde{c}^{\gamma}_{\rho}(\overline{G})$. Contributions in excess of the norm do not incur psychological costs or generate additional benefits; therefore, when $c^{\star\gamma}(\overline{G}) \geq \overline{G}$, it follows that $\tilde{c}^{\gamma}_{\rho}(\overline{G}) = c^{\star\gamma}(\overline{G})$. The stylised stereotypical behaviour of strong conditional cooperators, matching average contributions one-for-one when doing so is not conditionally Pareto-dominated, is generated by a sufficiently large value of ρ. In particular, for each game γ, there exists some threshold $\overline{\rho}^{\gamma}$ such that, when $\rho > \overline{\rho}^{\gamma}$, $\tilde{c}^{\gamma}_{\rho}(\overline{G}) = \overline{G}$ for all information sets $\overline{G}$ at which $c^{\star\gamma}(\overline{G}) < \overline{G}$. These threshold values are 0.61 for LINEAR, 0.51 for COMPLEMENTS, 0.76 for DOMINANT, and 0.88 for SUBSTITUTES. If $0 < \rho < \overline{\rho}^{\gamma}$, $\tilde{c}^{\gamma}_{\rho}(\overline{G}) \in \left(0, \overline{G}\right)$ for all information sets $\overline{G}$ at which $c^{\star\gamma}(\overline{G}) < \overline{G}$, corresponding broadly to the behaviour of weak conditional cooperators. The medoid contribution strategies for WCCs are best rationalised by (9) with values of ρ between 0.015 and 0.03. These parameters contrast sharply with those which rationalise SCC strategies.

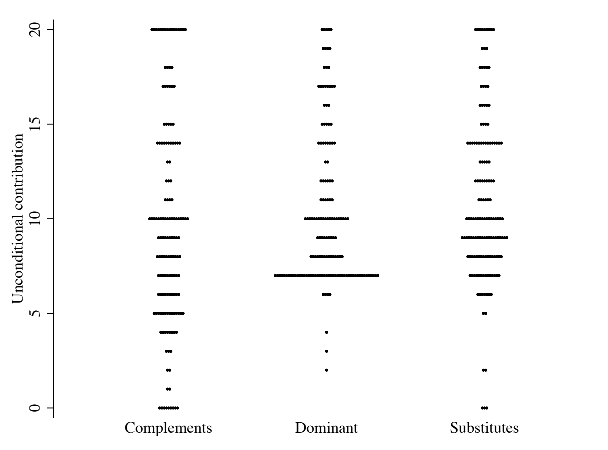


Figure 5: Distributions of unconditional contributions.

	LINEAR	DOMINANT	COMPLEMENTS	SUBSTITUTES
LINEAR	1.00			
DOMINANT	0.38 (<0.001)	1.00		
COMPLEMENTS	0.23 (0.030)	0.38 (<0.001)	1.00	
SUBSTITUTES	0.21 (0.068)	0.43 (<0.001)	0.40 (<0.001)	1.00

Table 4: Spearman rank-order correlation of unconditional contributions across games. Numbers in parentheses are Bonferroni-corrected significance levels.

5.3 Unconditional contributions across games

Our experimental design focuses on eliciting the contribution strategies. In Section 4 we did not specify hypotheses about the unconditional contributions. Nevertheless, under the assumption of own-earnings-maximisation, rationalisable unconditional contributions can be ranked, with $u_i^S \leq 7 = u_i^D \leq u_i^C$. This ranking arises because of the Stackelberg intuition that Stage 1 players can strategically influence the contribution response of the Stage 2 player; a participant might recognise this strategic opportunity qualitatively without reacting quantitatively as equilibrium predicts. We therefore use this prediction as a benchmark, while bearing in mind that unconditional contributions are a function both of preferences, which based on the evidence from type classifications vary substantially across participants, and beliefs about the behaviour of others, which we did not attempt to elicit.

Result 5. *In contrast to the ranking of games provided by rationalisability, unconditional contributions are highest in* SUBSTITUTES*, followed by* DOMINANT*, followed by* COMPLEMENTS*. The unconditional contributions of individual participants are positively correlated across games.*

***Support*.** We plot the distributions of unconditional contributions in Figure 5. The average unconditional contribution is 9.7 tokens in COMPLEMENTS, 10.3 in DOMINANT, and 11.1 in SUBSTITUTES. 80 (50) participants contribute more (fewer) tokens in SUBSTITUTES than COMPLEMENTS ($p = .003$, Wilcoxon matched-pairs test). Contributions in DOMINANT are in between: 75 (48) participants contribute more (fewer) tokens in SUBSTITUTES than DOMINANT ($p = .091$), and 75 (54) contribute more (fewer) tokens in DOMINANT than COMPLEMENTS ($p = .040$). Overall, 37.1% of participants contribute at least as many tokens in SUBSTITUTES than DOMINANT and at least as many tokens in DOMINANT than COMPLEMENTS, while 20.0% of participants exhibit the reverse order.

Individual participants are systematically more or less generous in contributing to the project across games. Table 4 reports the Spearman rank-order correlations of participants' unconditional contributions. The correlations between contributions in each pair of games are systematically positive, ranging from 0.21 between LINEAR and SUBSTITUTES to 0.43 between DOMINANT and SUBSTITUTES. We additionally aggregate unconditional contributions by type in Table 5, which shows that the pattern of contributing more tokens in SUBSTITUTES than in COMPLEMENTS is not driven solely by the behaviour of any one type.

□

Unconditional contributions exceed the predictions for DOMINANT, SUBSTITUTES, and COMPLEMENTS, even among own-maximisers, who almost all contribute zero tokens in LINEAR. Our baseline analysis assumed the Stage 2 player maximised her own earnings; however, it is certainly

	LINEAR	DOMINANT	COMPLEMENTS	SUBSTITUTES
OWN	0.4 (1.2)	8.8 (3.5)	8.8 (6.4)	10.1 (4.4)
WCC	4.2 (3.6)	9.8 (3.7)	8.5 (4.0)	11.1 (3.9)
SCC	6.1 (6.2)	11.1 (4.4)	10.9 (5.8)	11.9 (4.4)
UNH	12.5 (9.6)	14.5 (4.4)	17.3 (3.2)	13.5 (3.3)
MID	7.8 (4.1)	12.1 (4.0)	9.9 (5.2)	11.3 (4.9)

Table 5: Average and standard deviations (in parentheses) of Stage 1 contributions for each type in each game.

reasonable that at least some players anticipated at least some reciprocity by the Stage 2 players as exhibited by WCC and SCC. In Appendix A we extend our theoretical analysis to the case where Stage 1 players are own-maximisers but anticipate that the Stage 2 player follows a SCC strategy. Under this assumption, in SUBSTITUTES, $u_1^{\star S} + u_2^{\star S} + u_3^{\star S} \in \{13, 26\}$; in DOMINANT, $u_1^{\star D} + u_2^{\star D} + u_3^{\star D} \in \{26, 29\}$; and in COMPLEMENTS, $u_1^{\star D} + u_2^{\star D} + u_3^{\star D} \in \{26, 29, 32\}$.

We can offer some circumstantial evidence for anticipated reciprocity by comparing the unconditional contributions of conditionally-cooperative participants between COMPLEMENTS and SUBSTITUTES. Recall that the contribution strategies of strong conditional cooperators are particularly interesting in SUBSTITUTES because they contribute the fewest tokens in response to the information set $\overline{G} = 7$; either decreasing or increasing unconditional contributions from this level leads to an increase in their contributions. In contrast, their contribution strategies are increasing in COMPLEMENTS. We take the set of participants classified as WCC and SCC and consider their unconditional contributions in LINEAR. If we assume that these participants approach their decisions in Stage 1 and Stage 2 in similar ways, we can take their unconditional contribution in LINEAR as a rough proxy for their beliefs about the general expected contribution levels of other participants. A conditional cooperator who believes others will contribute few tokens should contribute more tokens in SUBSTITUTES than in COMPLEMENTS, because they anticipate contributions are likely to be in the region $\overline{G} < 7$ and therefore their preferences for conditional cooperation do not operate. In contrast, a conditional cooperator who believes contributions of others will be high should contribute roughly the same in SUBSTITUTES and COMPLEMENTS, because in that case their preferences for conditional cooperation would encourage them to contribute similar amounts in either game. We divide conditional cooperators into two groups based on whether their unconditional contribution in LINEAR is above or below the median of their type. We find that those who choose unconditional contributions below the median in LINEAR contribute significantly more tokens in SUBSTITUTES than in COMPLEMENTS (11.2 versus 8.5, Wilcoxon matched-pairs test $p = 0.005$). Those with above-median unconditional contributions in LINEAR contribute similar amounts in

Type	N	Mean	SE	Quartiles		
Own-maximisers	48	112.9	78.5	57	77	150
Strong conditional cooperators	42	115.0	70.8	62	97	153
Weak conditional cooperators	32	117.5	80.2	63	97	131
Mid-range	22	201.1	113.1	113	162	253
Unconditional high	4	150.3	100.7	79	117	222

Table 6: Time spent, in seconds, by participants answering control questions, by behavioural type.

SUBSTITUTES and COMPLEMENTS (11.9 versus 11.4, $p = 0.541$).

Anticipated reciprocity can therefore account for the unconditional contribution levels exceeding the baseline set by assuming an own-earnings-maximising Stage 2 player, but not the ordering of unconditional contributions we observe across games.

5.4 Response times to control questions

We have observed that strong conditional cooperators across games choose Stage 2 contribution strategies which appear to follow consistent principles incorporating some sophisticated consideration of the financial incentives of the game. This is suggestive that strong conditional cooperators are making a well-informed and conscious decision in forming their Stage 2 strategies.

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.[19] Table 6 reports descriptive statistics on the distribution of these times, by behavioural type.

Result 6. *Strong conditional cooperators are not different from own-maximisers or weak conditional cooperators in response time to control questions. Own-maximisers, weak and strong conditional cooperators take significantly less time to complete the control questions than mid-range.*

Support. Own-maximisers on average take 112.9 seconds to complete the control questions, strong conditional cooperators 115.0 seconds and weak conditional cooperators 117.5 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. (Kruskal-Wallis test, $p = 0.84$)

Mid-range contributors take notably longer to complete the questions, at 201.1 seconds. This differs from the response times of own-maximisers, weak conditional cooperators, and strong conditional cooperators, (MWW, $p = 0.001$; the probability the completion time of a randomly-chosen MID participant is longer than that of a randomly-chosen OWN/WCC/SCC is .76.) □

[19] Note however, that Bigoni et al. (2016) control for the task comprehension on the level of contribution in a repeated game, finding no correlation.

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.[20] 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. Our strong conditional cooperators appear to be as well-engaged and understand the task as well as our own-maximisers.[21]

5.5 Mid-range and unconditional high contributors

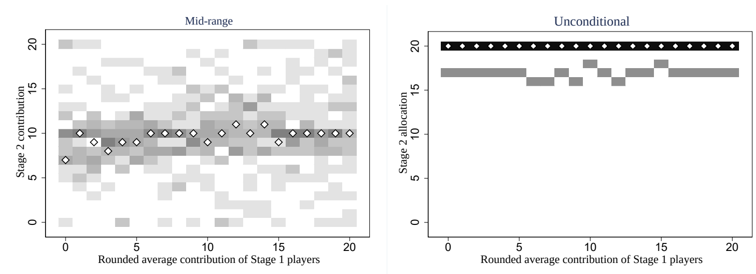
Only 22 participants are classified as mid-range contributors and 4 as unconditional high contributors. These participants do not show a systematic response to the anticipated contributions of the other members of their group in LINEAR. Recall that the stereotype strategy for mid-range contributors is a constant contribution of 10 tokens irrespective of $\overline{G}$ and the stereotype for unconditional high is full contribution of 20 tokens.

Figure 6 shows the heatmaps of Stage 2 strategies for these types. With only 4 participants classified as unconditional high contributors no meaningful conclusions can be drawn. The much longer control question response times for mid-range types reported in Table 6 suggest further qualitative comment on their behaviour across games. Their longer times to complete the control questions successfully suggests they had a harder time comprehending the experiment, were less engaged with the task, or both.

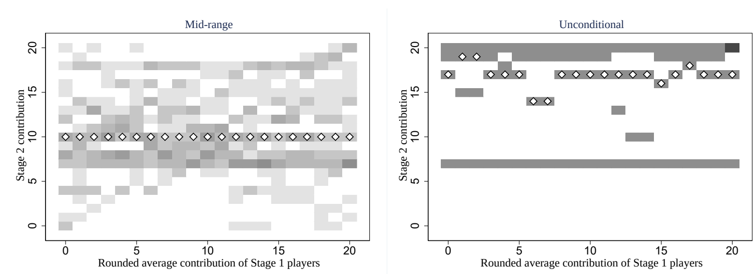
The medoid contribution strategy for mid-range contributors in LINEAR is exactly, or nearly, a contribution of 10 for all $\overline{G}$ for all four games. The contribution strategies do not shift systematically in response to the varying financial incentives across the games. The contributions in these contribution strategies are not uniformly random; contributions closer to 10 tokens are more frequent than those at either the high or low end of the strategy space. This contrasts with the clustering results in Fallucchi et al. (2019), in which the heatmap of the analogous "various" group is roughly uniform for all information sets $\overline{G}$. This may be a response to a lack of confidence in their comprehension of the game, in which a "choose the middle" heuristic may seem the safest

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

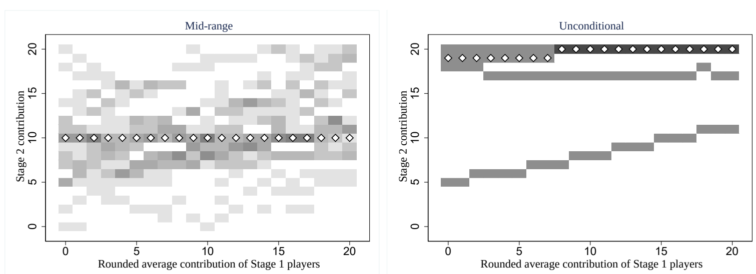
[21]We look at response times to the control questions rather than response times on choices because there are confounds in interpreting the latter. SCC generally take the longest to complete their Stage 2 decisions, while OWN complete Stage 2 more quickly. Fast decision times, however, are consistent both with clarity in one's own responses and with a lack of deliberation. There is a more prosaic reason why SCC take longer to complete Stage 2, which is simply that it takes more mouse movement to input the SCC strategies. It is interesting that some participants systematically adopt the SCC strategies even though it is more work for them to input it in the software.



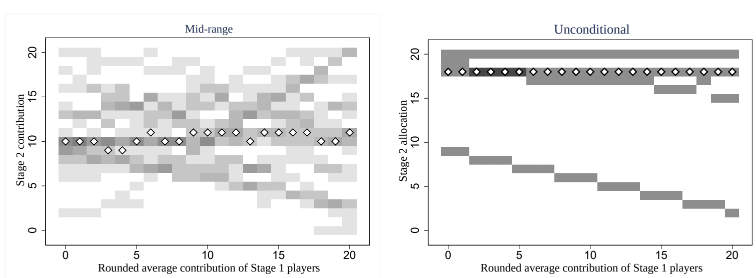
(a) LINEAR



(b) DOMINANT



(c) COMPLEMENTS



(d) SUBSTITUTES

Figure 6: Heatmap of Stage 2 strategies for allocation to the project, for mid-range and unconditional high types, across games

option. Alternatively, some may choose a contribution of 10 tokens on the principle of "share and share alike," where sharing is not done in terms of the financial incentives of the game, but instead based on the strategy space of tokens. Our experiment is not designed to identify these or other possible motivations for the mid-range contributor strategies, but we can note that as a group they are not systematically responsive to the economic environment given by the financial incentives. This suggests the instinctive response of participants who are unengaged, confused, uncertain, and/or importing heuristics from other settings is to split the token endowment more or less equally between the private account and the project; and not, in contrast to Hypothesis 4, reflexively to match the contributions of others one-for-one.

6 Discussion

We investigate the robustness of pro-social behaviour in VCM games by eliciting the behaviour of the same participants in games with different economic and strategic structures. Using a novel choice architecture to present the financial incentives and elicit the decisions, in the game with linear payoffs we find proportions of participants who adopt contribution strategies at or near the two most common stereotypes, exact free-riding and exact one-for-one matching, which are comparable to previous studies. We find contribution strategies in the linear game are strongly predictive of the contribution strategies adopted in games with other economic structures. Taken together, type classifications based on contribution strategies are fairly robust to how the decisions are elicited, and have predictive value for how participants will choose in other, related social dilemmas.

Our results support the distinction between strong and weak forms of conditional cooperation, as proposed in Fallucchi et al. (2019). In particular, strong conditional cooperators generally avoid choosing conditionally Pareto-dominated contribution levels, which implies an understanding of the economic structure of our environments and a sophistication in adapting the principle conditional cooperation to each. In contrast, weak conditional cooperators exhibit more heterogeneous behaviour, and are not well-distinguished from own-earnings-maximisers across the games we study.

The behaviour of strong conditional cooperators across this family of games allows us to rule out further candidate 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.

Andreoni (1990) proposed people contribute (more) to a public good because they receive "warm glow" utility from the mere act of contributing. In the linear VCM, 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 COMPLEMENTS, DOMINANT, and SUBSTITUTES, a warm glow theory of contribution would predict contributions in excess of the own-earnings-maximising amount for information sets $\overline{G}$ below the equilibrium level; 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 our other three games, 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) raises the possibility that the conditional contribution procedure creates an experimenter demand effect, by suggesting that contributions *should* depend on the actions of others. Under such a theory, participants would be more likely to specify a contribution strategy which is responsive to the information about the contributions $\overline{G}$ of others, even though the financial incentives of the game would indicate otherwise for participants seeking to maximise their own earnings. In our games SUBSTITUTES and COMPLEMENTS, own-maximising participants should indeed respond to $\overline{G}$, and we find that participants who are classified as own-maximisers in LINEAR predominantly also choose contribution strategies in SUBSTITUTES and COMPLEMENTS which maximise their own earnings as well. In contrast, strong conditional cooperators continue to adopt at or near one-for-one matching of contributions when it is efficiency-enhancing to do so.

The main difference in our data compared to previous experiments using the p-experiment protocol is that we find significantly lower unconditional contributions in LINEAR. Under the assumption that participants choose unconditional contributions according to their beliefs about the choices of others while using a strategy similar to their stated contribution strategy, lower unconditional contributions by conditional cooperators (of either type) would imply more pessimistic beliefs about the anticipated contributions of others (Kölle et al., 2014).[22] However, unconditional contributions by own-maximisers are also lower; indeed almost all participants who are exact free-riders in their contribution strategy also specify an unconditional contribution of zero. So beliefs alone cannot explain the lower unconditional contributions in our data. Equally, our lower levels of unconditional contributions makes the persistence of conditionally-cooperative contribution

[22]We did not attempt to measure beliefs in this experiment. The protocol is already complex for a participant to digest, even with our choice architecture and concrete phrasing of many parts of the instructions.

strategies in our data more remarkable.

The discrepancy between unconditional contributions and contribution strategies for free-riders might be understood in terms of Selten's original conception of the "strategy method." In the original sense of a "strategy method" experiment, a participant would formulate and write down a strategy for an extensive game. The participant would then play out the game, while being free to deviate from the prepared strategy if they wished. In Stage 2 of the p-experiment, participants are asked to think through the possible contingencies of G_{-i} that might arise, and how they would respond; such contingency-by-contingency reasoning might encourage own-earnings-maximising participants to realise that, in LINEAR, free riding always maximises earnings. In contrast, Stage 1 does not enforce such best-response reasoning.[23] Customarily, the Stage 2 elicitation always follows the Stage 1, so the learning from Stage 2 comes too late to inform Stage 1 decisions. Our choice architecture labels the tokens individually with their earnings consequences, which in its own way makes the allocation which maximises one's own earnings more transparent.

Based on the re-analysis of Fallucchi et al. (2019) and the data in this paper, about 30% of participants choose strongly conditionally cooperative contribution strategies in a linear VCM game. We show that the contribution strategies these participants adopt in other games points to them understanding the financial incentives they face in the games, and reacting to those in a sophisticated way. Whether for reasons of conformity to a norm, inequity aversion, or another motivation, these strong conditional cooperators match the average contributions of others when - and only when - doing so is efficiency enhancing. The "only when" in the previous statement allows us to rule out confusion, misunderstanding, or a lack of engagement with the experimental task as an explanation for this behaviour. Strong conditional cooperators are, in general, 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, 1995.

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

[23] Anyone who has taught introductory game theory will know from experience that the contingency-by-contingency reasoning to generate a reaction function does not come naturally to most students!

Abigail Barr, Tom Lane, and Daniele Nosenzo. On the social inappropriateness of discrimination. *Journal of Public Economics*, 164:153–164, 2018.

Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300, 1995.

Maria Bigoni, Stefania Bortolotti, Marco Casari, Diego Gambetta, and Francesca Pancotto. Amoral familism, social capital, or trust? the behavioural foundations of the italian north-south divide. *The Economic Journal*, 126(594):1318–1341, 2016.

Olaf Bock, Andreas Nicklisch, and Ingmar Baetge. hRoot: Hamburg registration and organization online tool. *European Economic Review*, 71:117–120, 2014.

Jordi Brandts and Christiane Schwieren. Frames and economic behavior. Working paper, 2009.

Adrian Bruhin, Ernst Fehr, and Daniel Schunk. The many faces of human sociality: Uncovering the distribution and stability of social preferences. *Journal of the European Economic Association*, 17(4):1025–1069, 2018.

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.

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.

Caleb Cox, Oleg Korenok, Edward Millner, and Laura Razzolini. Giving, taking, earned money, and cooperation in public good games. *Economics Letters*, 171:211–213, 2018.

Caleb A Cox. Decomposing the effects of negative framing in linear public goods games. *Economics Letters*, 126:63–65, 2015.

James C. Cox, Elinor Ostrom, Vjollca Sadiraj, and James M. Walker. Provision versus appropriation in symmetric and asymmetric social dilemmas. *Southern Economic Journal*, 79(3):496–512, 2013.

Robin P. Cubitt, Michalis Drouvelis, and Simon Gächter. Framing and free riding: Emotional responses and punishment in social dilemma games. *Experimental Economics*, 14(2):254–272, 2011.

Martin Dufwenberg, Simon Gächter, and Heike Hennig-Schmidt. The framing of games and the psychology of play. *Games and Economic Behavior*, 73(2):459–478, 2011.

Jon Elster. Social norms and economic theory. *Journal of Economic Perspectives*, 3(4):99–117, 1989.

Francesco Fallucchi, R. Andrew Luccasen, and Theodore L. Turocy. Identifying discrete behavioural types: a re-analysis of public goods game contributions by hierarchical clustering. *Journal of the Economic Science Association*, 2019.

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

Ernst Fehr and Ivo Schurtenberger. Normative foundations of human cooperation. *Nature Human Behaviour*, 2(7):458, 2018.

Ernst Fehr and Jean-Robert Tyran. Individual irrationality and aggregate outcomes. *Journal of Economic Perspectives*, 19(4):43–66, 2005.

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.

Urs Fischbacher, Simon Gächter, and Simone Quercia. The behavioral validity of the strategy method in public good experiments. *Journal of Economic Psychology*, 33(4):897–913, 2012.

Simon Gächter, Felix Kölle, and Simone Quercia. Reciprocity and the tragedies of maintaining and providing the commons. *Nature human behaviour*, 1(9):650, 2017.

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.

William T Harbaugh and Kate Krause. Children's altruism in public good and dictator experiments. *Economic Inquiry*, 38(1):95–109, 2000.

Henning Hermes, Florian Hett, Mario Mechtel, Felix Schmidt, Daniel Schunk, and Valentin Wagner. Do children cooperate conditionally? adapting the strategy method for first-graders. *Journal of Economic Behavior & Organization*, 2019.

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

Menusch Khadjavi and Andreas Lange. Doing good or doing harm: experimental evidence on giving and taking in public good games. *Experimental Economics*, 18(3):432–441, 2015.

David C Kingsley. Peer punishment across payoff equivalent public good and common pool resource experiments. *Journal of the Economic Science Association*, 1(2):197–204, 2015.

Felix Kölle, Simon Gächter, and Simone Quercia. The abc of cooperation in voluntary contribution and common pool extraction games. Verein für Socialpolitik/German Economic Association, 2014.

Hannu Oja. *Multivariate nonparametric methods with R: An approach based on spatial signs and ranks*. Springer Science & Business Media, 2010.

Jan Potters and Sigrid Suetens. Cooperation in experimental games of strategic complements and substitutes. *The Review of Economic Studies*, 76:1125–1147, 2009.

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

R. John Simes. An improved bonferroni procedure for multiple tests of significance. *Biometrika*, 73(3): 751–754, 1986.

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 games: Design and perfect Bayesian equilibrium analysis

In this section we analyse the four p-experiment games, under the assumption that all players maximise their own earnings. We show that for each game the Stage 2 player has a unique best response strategy. Given this, we can identify the set of unconditional contributions which are rationalisable for the Stage 1 players, and characterise the set of perfect Bayesian equilibria.

We number the three Stage 1 players as $i = 1, 2, 3$, and the Stage 2 player as $i = 4$. For LINEAR and DOMINANT the analysis is straightforward, as the allocation which maximises earnings does not depend on the allocations of other players. In LINEAR, the contribution which maximises earnings is 0, irrespective of what other players do, and therefore the Stage 2 player's equilibrium contribution strategy is $c^{\star L}(\overline{G}) = 0$ for all $\overline{G} \in \{0, \ldots, 20\}$, and for the Stage 1 players, the equilibrium unconditional contributions are $u_1^{\star L} = u_2^{\star L} = u_3^{\star L} = 0$. In DOMINANT, the contribution which maximises earnings is 7, and therefore the Stage 2 player's equilibrium contribution strategy is $c^{\star D}(\overline{G}) = 7$ for all $\overline{G} \in \{0, \ldots, 20\}$, and for the Stage 1 players, the equilibrium unconditional contributions are $u_1^{\star D} = u_2^{\star D} = u_3^{\star D} = 7$.

In turning to the analysis of SUBSTITUTES and COMPLEMENTS, our design of these games incorporated several considerations. As a starting point, we chose parameters such that the equilibria of the simultaneous-move VCM with payoff functions Π^S and Π^C would have a unique symmetric equilibrium with all players contributing 7, which coincides with the unique equilibrium in DOMINANT. The p-experiment game is a two-stage game, however. If our games were played with payoff functions Π^S and Π^C but with continuous action spaces and perfect information about the actions of the Stage 1 players, there would be a unique equilibrium in each game following the usual Stackelberg-type logic. In SUBSTITUTES, Stage 1 players would have an incentive to reduce contributions below 7, anticipating that the Stage 2 player would respond with a higher contribution; in COMPLEMENTS, the Stage 1 players would have an incentive to increase contributions above 7, anticipating that the Stage 2 player would respond with a higher contribution.

This intuition applies to our games, but is complicated by the discreteness of the action space and especially the imperfect information resulting from the rounding of the average contributions of the Stage 1 players. The latter creates a discontinuity in the reaction function for the Stage 2 player; it is this discontinuity that results in multiple equilibria in games with this structure (and not just for our chosen parameters). There are Stackelberg-type equilibria, which for our parameters involve asymmetric contributions among the Stage 1 players, where the asymmetry is a consequence of the discreteness of action spaces. Meanwhile, the unique symmetric equilibrium of the game corresponds to the equilibrium of the simultaneous-move version, and therefore to the dominant strategies in DOMINANT. This equilibrium survives because any of the Stage 1 player would have to decrease (in SUBSTITUTES) or increase (in COMPLEMENTS) their contribution by multiple tokens in order to change the response of the Stage 2 player, and such a unilateral deviation is unprofitable.

Another important design feature is that the best response for the Stage 2 player is constant for all G_{-i} consistent with each information set $\overline{G}$. This is a consequence of the discreteness of the action space, and is useful both theoretically and practically. Theoretically, it means that beliefs at each information set are not important for computing the reaction functions, and therefore the Stage 2 player has a dominant strategy response. This is important because the Stage 2 contribution strategies are the focus of our experiment. In LINEAR, the rounding of average Stage 1 contributions is not problematic for the baseline case of own-earnings-maximising players, but is important practically because it cuts down the number of choices participants need to specify in their contribution strategies. Our parameter choices allow us to extend the p-experiment design to our payoff structure, while not losing any information which is strategically relevant to an own-earnings-maximising player in Stage 2.

The multiplicity of equilibria is a consequence of having discrete action spaces, the rounding of the average contributions to report to the Stage 2 player, and having a systematic pattern for how the value of tokens allocated to the private account change. Multiplicity could in principle be eliminated by manipulating the formula for the value of tokens in the private account to destroy one of the classes of equilibria, but at the cost of not having an easily-explainable rule for how these values are determined. Our design retains the nice properties that the symmetric equilibrium contributions are the same across SUBSTITUTES, DOMINANT, and COMPLEMENTS, and further that the iteratively rationalisable unconditional contributions are weakly orderable, with contributions in SUBSTITUTES no more than in DOMINANT, and contributions in DOMINANT no more than in COMPLEMENTS.

We now turn to the detailed analysis of SUBSTITUTES and COMPLEMENTS.

A.1 Analysis of SUBSTITUTES

We begin with the Stage 2 player's decision. For any fixed G_{-4}, player 4's earnings are strictly concave in her allocation decision. Therefore, if two allocations (g_4, x_4) and $(g_4 - 1, x_4 + 1)$, which are adjacent in $\mathcal{A}_{\mathbb{Z}}$, result in the same earnings, they must jointly be the two (and only two) earnings-maximising allocations. We can therefore characterise the best response of player 4 by identifying the values of G_{-4} at which she is indifferent between adjacent allocations.

$$\begin{aligned}\Pi_4^S(g_4, x_4, G_{-4}) - \Pi_4^S(g_4 - 1, x_4 + 1, G_{-4}) &= -\left(1.06 + \frac{.02}{3}G_{-4}\right) + .03\,(2x_4 + 1) + 0.40 \\ &= -.63 + .06(20 - g_4) - \frac{.02}{3}G_{-4}. \qquad (10)\end{aligned}$$

Setting this equal to zero, we see that player 4 is indifferent between (g_4, x_4) and $(g_4 - 1, x_4 + 1)$ if and only if $\hat{G}_{-4} = 9(9.5 - g_4)$. If $G_{-4} < \hat{G}_{-4}$, she strictly prefers (g_4, x_4) to $(g_4 - 1, x_4 + 1)$, and if $G_{-4} > \hat{G}_{-4}$, she strictly prefers $(g_4 - 1, x_4 + 1)$ to (g_4, x_4). Note that there are no solutions where g_4 and $\hat{G}_{-4}$ are both integers. Therefore the best response is strict for all G_{-4}, and (g_4, x_4) is a best response to G_{-4} if and only if $25\frac{2}{3} - 3g_4 \leq \frac{G_{-4}}{3} \leq 28\frac{1}{3} - 3g_4$. If this inequality is satisfied for some G_{-4} such that $\frac{G_{-4}}{3}$ is an integer, it is also satisfied for $G_{-4} - 1$ and $G_{-4} + 1$ for the same g_4. Therefore, for all information sets $\overline{G}$, the best response is constant over all total contributions G_{-4} for which $\frac{G_{-4}}{3}$ rounds to $\overline{G}$. The beliefs that player 4 might have over the values G_{-4} which are consistent with information set $\overline{G}$ do not affect the best response. The unique rationalisable contribution strategy $c^{\star S}(\overline{G})$, and the contribution strategy in any perfect Bayesian equilibrium, is therefore

$c^{\star S}(\overline{G})$	2	3	4	5	6	7	8	9
G_{-4}	59-60	50-58	41-49	32-40	23-31	14-22	5-13	0-4
$\overline{G}$	20	17-19	14-16	11-13	8-10	5-7	2-4	0-1

Turning to the Stage 1 players, the discrete jumps in the contribution strategy $c^{\star S}(\overline{G})$ complicate the analysis. It is most straightforward to tabulate the best response function by direct calculation. Player 1's reaction function is given by

$u_2 + u_3$	$\tilde{u}_1^S$	$u_2 + u_3$	$\tilde{u}_1^S$	$u_2 + u_3$	$\tilde{u}_1^S$	$u_2 + u_3$	$\tilde{u}_1^S$
0	8						
1	8	11	7	21	6	31	5
2	8	12	7	22	6	32	5
3	8	13	7	23	6	33	5
4	8	14	7	24	6	34	5
5	8	15	7	25	6	35	5
6	7	16	6	26	5	36	4
7	6	17	5	27	4	37	3
8	5	18	4	28	3	38	2
9	4	19	3	29	2	39	1
10	3	20	2	30	1	40	0

By symmetry the best responses of the other players are identical up to the appropriate permutation of the indices. In many contingencies, the optimal response for a player is to reduce their contribution by exactly the number of tokens required to trigger a one-token increase by the Stage 2 player. For example, consider a situation in which player 1 believes that $u_2 + u_3 = 18$. Consider any $5 \leq u_1 \leq 13$; this results in $23 \leq G_{-4} \leq 31$ and the Stage 2 player responds with a contribution of 6 at these information sets. Among these, $u_1 = 6$ would result in the highest earnings for player 1. However, if instead player 1 contributes $u_1 = 4$, then $G_{-4} = 32$ and the Stage 2 player responds instead with a contribution of 7 tokens. Player 1's earnings from $u_1 = 4$ are higher than from $u_1 = 6$; there is a net loss of 0.40 from there being one fewer token contributed overall to the project, but player 1 is more than compensated by his private account tokens being more valuable due to the Stage 2 player's larger contribution.

To identify the rationalisable strategies for the Stage 1 players, without loss of generality assume $u_1^\star \leq u_2^\star \leq u_3^\star$. Because $u_3^\star \leq 8$, $u_2^\star + u_3^\star \leq 16$, and so $3 \leq u_1^\star$. This implies $u_1^\star + u_2^\star \geq 6$, and so $u_3^\star \leq 7$. Therefore, $6 \leq u_2^\star + u_3^\star \leq 14$, and the rationalisable strategies are $3 \leq u_i^\star \leq 7$.

The remaining pure strategy profiles $(u_1^\star, u_2^\star, u_3^\star)$ in which $u_1^\star \leq u_2^\star \leq u_3^\star$ and $u_1^\star$ is a best response to $u_2^\star + u_3^\star$ are $(7, 7, 7)$ and $\{(3, 3, 7), (3, 4, 6), (3, 5, 5), (4, 4, 5)\}$. By inspection, the first is the unique symmetric equilibrium, in which total unconditional contributions are $G_{-4} = 21$, and the Stage 2 player responds on the equilibrium path with a contribution of 7. The second set are asymmetric equilibria, in which total unconditional contributions are $G_{-4} = 13$, and the Stage 2 player responds on the equilibrium path with a contribution of 8.

A.2 Analysis of COMPLEMENTS

We begin with the Stage 2 player's decision. For any fixed G_{-4}, player 4's earnings are strictly concave in her allocation decision. Therefore, if two allocations (g_4, x_4) and $(g_i - 4, x_i + 4)$, which are adjacent in $\mathcal{A}_{\mathbb{Z}}$, result in the same earnings, they must jointly be the two (and only two) earnings-maximising allocations. We can therefore characterise the best response of player 4 by identifying the values of G_{-4} at which she is indifferent between adjacent allocations.

$$\begin{aligned} \Pi_i^C(g_4, x_4, G_{-4}) - \Pi_i^C(g_4 - 1, x_4 + 1, G_{-4}) &= -\left(1.34 - \frac{.02}{3} G_{-4}\right) + .03\,(2x_4 + 1) + 0.40 \\ &= -.91 + .06(20 - g_4) + \frac{.02}{3} G_{-4}. \end{aligned} \tag{11}$$

Setting this equal to zero, we see that player 4 is indifferent between (g_4, x_4) and $(g_4 - 1, x_4 + 1)$ if and only if $\hat{G}_{-4} = 9\left(g_4 - 4\frac{5}{6}\right)$. If $G_{-4} > \hat{G}_{-4}$, she strictly prefers (g_4, x_4) to $(g_4 - 1, x_4 + 1)$, and if $G_{-4} < \hat{G}_{-4}$, she strictly prefers $(g_4 - 1, x_4 + 1)$ to (g_4, x_4). Note that there are no solutions where g_4 and $\hat{G}_4$ are both integers. Therefore, the best response is strict for all G_{-4}, and (g_4, x_4) is a best response to G_{-4} if and only if $3g_4 - 14\frac{1}{3} \leq \frac{G_{-4}}{3} \leq 3g_4 - 11\frac{2}{3}$. If this inequality is satisfied for some G_{-4} such that $\frac{G_{-4}}{3}$ is an integer, it is also satisfied for $G_{-4} - 1$ and $G_{-4} + 1$ for the game g_4. Therefore, for all information sets $\overline{G}$, the best response is constant over all total contributions G_{-4} for which $\frac{G_{-4}}{3}$ rounds to $\overline{G}$. The beliefs that player 4 might have over the values G_{-4} which are consistent with information set $\overline{G}$ do not affect the best response. The unique rationalisable contribution strategy $c^{\star C}(\overline{G})$, and the contribution strategy in any perfect Bayesian equilibrium, is therefore

$c^{\star C}(\overline{G})$	11	10	9	8	7	6	5	4
G_{-4}	56-60	47-57	38-46	29-37	20-28	11-19	2-10	0-1
$\overline{G}$	20-19	16-18	13-15	10-12	7-9	4-6	1-3	0

Turning to the Stage 1 players, again it is most straightforward to tabulate the best response function by direct calculation. Player 1's reaction function is given by

u_2+u_3	$\tilde{u}_1^C$	u_2+u_3	$\tilde{u}_1^C$	u_2+u_3	$\tilde{u}_1^C$	u_2+u_3	$\tilde{u}_1^C$
0	5						
1	5	11	9	21	8	31	9
2	5	12	8	22	8	32	9
3	8	13	7	23	8	33	9
4	7	14	7	24	8	34	9
5	6	15	7	25	8	35	12
6	6	16	7	26	8	36	11
7	6	17	7	27	11	37	10
8	6	18	7	28	10	38	10
9	6	19	10	29	9	39	10
10	6	20	9	30	9	40	10

By symmetry the best responses of the other players are identical up to the appropriate permutation of the indices. The intuition for the jumps in the reaction function is parallel to SUBSTITUTES, except in COMPLEMENTS Stage 1 players may find it profitable to contribute just enough tokens to ensure an additional token contribution by the Stage 2 player.

To identify the rationalisable strategies for the Stage 1 players, without loss of generality assume $u_1^\star \leq u_2^\star \leq u_3^\star$. Because $u_2^\star \geq 5$, $u_2^\star + u_3^\star \geq 10$, and so $6 \leq u_1^\star$. This implies $u_2^\star + u_3^\star \geq 12$. Therefore, the rationalisable strategies are $7 \leq u_i^\star \leq 10$.

We consider remaining pure strategy profiles $(u_1^\star, u_2^\star, u_3^\star)$ in which $u_1^\star \leq u_2^\star \leq u_3^\star$ in increasing lexicographic order. The profile $(7, 7, 7)$ is a symmetric equilibrium, in which the total unconditional contributions are $G_{-4} = 21$, and the Stage 2 player responds on the equilibrium path with a contribution of 7. In order for $u_3^\star \geq 8$, $u_1^\star + u_2^\star \geq 19$, so the next profile to consider is $(9, 9, 10)$, at which player 3 is not best-responding. The profile $(9, 10, 10)$ is an asymmetric equilibrium, in which the total unconditional contributions are $G_{-4} = 29$, and the Stage 2 player responds on the equilibrium path with a contribution of 8.

A.3 Extension to the case where the Stage 2 player is SCC

As an extension, we characterise the set of rationalisable strategies and the pure-strategy perfect Bayesian equilibria under the assumption that Stage 1 players maximise their own earnings, while the Stage 2 player is a strong conditional cooperator.

First, observe that in LINEAR the equilibrium unconditional contributions remain $u_1^{\star L} = u_2^{\star L} = u_3^{\star L} = 0$ even if the Stage 2 player follows the strategy $\hat{c}^{SCC}(\overline{G}) = \overline{G}$. Without loss of generality, consider player 1; we claim that for any given u_2 and u_3, player 1's earnings are strictly decreasing in u_1. To see this, suppose player 1 increases their contribution from some level u_1 to u_1+1. Either this results in no change in the Stage 2 player's contribution, in which case player 1's earnings decrease by 0.60; or, it results in the Stage 2 player contributing an additional token, in which case player 1's earnings decrease by 0.20. Therefore, $u_i = 0$ remains a strictly dominant strategy for all Stage 1 players $i = 1, 2, 3$.

Based on our experimental results, for games $\gamma \in \{D, S, C\}$ we define the Stage 2 player's SCC strategy as

$$c^{SCC,\gamma}(\overline{G}) = \begin{cases} c^{\star\gamma}(\overline{G}) & \text{if } \overline{G} \leq 7 \\ \overline{G} & \text{if } \overline{G} > 7. \end{cases}$$

We then proceed to compute the reaction functions for the Stage 1 players under the assumption the Stage 2 player uses strategy $c^{SCC,\gamma}$. The method for enumerating the rationalisable strategies and equilibrium unconditional contribution profiles is the same as used in the previous subsections; we present an abbreviated summary for compactness.

For SUBSTITUTES, the reaction function for player 1 is

$u_2 + u_3$	$\tilde{u}_1^S$	$u_2 + u_3$	$\tilde{u}_1^S$	$u_2 + u_3$	$\tilde{u}_1^S$	$u_2 + u_3$	$\tilde{u}_1^S$
0	8						
1	8	11	7	21	8	31	7
2	8	12	7	22	7	32	6
3	8	13	10	23	9	33	8
4	8	14	9	24	8	34	7
5	8	15	8	25	7	35	6
6	7	16	10	26	9	36	5
7	6	17	9	27	8	37	8
8	5	18	8	28	7	38	7
9	4	19	7	29	6	39	6
10	3	20	9	30	8	40	7

The rationalisable strategies are $3 \leq u_i^\star \leq 10$. The equilibria with $G_{-4} = 13$ previously identified remain equilibria in this modified setting. However, the symmetric profile $(7,7,7)$ is no longer an equilibrium. Instead, there is a family of equilibria at profiles $(7,9,10)$, $(8,8,10)$, and $(8,9,9)$, with $G_{-4} = 26$, to which the Stage 2 player responds on the equilibrium path with a contribution of 9.

For COMPLEMENTS, the reaction function for player 1 is

$u_2 + u_3$	$\tilde{u}_1^C$	$u_2 + u_3$	$\tilde{u}_1^C$	$u_2 + u_3$	$\tilde{u}_1^C$	$u_2 + u_3$	$\tilde{u}_1^C$
0	5						
1	5	11	9	21	11	31	10
2	5	12	8	22	10	32	12
3	8	13	7	23	9	33	11
4	7	14	9	24	11	34	13
5	6	15	8	25	10	35	12
6	6	16	10	26	9	36	11
7	6	17	9	27	11	37	13
8	6	18	8	28	10	38	12
9	6	19	10	29	12	39	14
10	6	20	9	30	11	40	13

The rationalisable strategies are $8 \leq x_i^\star \leq 11$. The symmetric profile $(7,7,7)$ is no longer an equilibrium. There exist equilibria at $(8,8,10)$ and $(8,9,9)$, with $G_{-4} = 26$ and a Stage 2 contribution of 9; at $(9,10,10)$, with $G_{-4} = 29$ and a Stage 2 contribution of 10; and at $(10,11,11)$, with $G_{-4} = 32$ and a Stage 2 contribution of 11.

For DOMINANT, the reaction function for player 1 is

$u_2 + u_3$	$\tilde{u}_1^D$	$u_2 + u_3$	$\tilde{u}_1^D$	$u_2 + u_3$	$\tilde{u}_1^D$	$u_2 + u_3$	$\tilde{u}_1^D$
0	7						
1	7	11	7	21	8	31	10
2	7	12	7	22	10	32	9
3	7	13	10	23	9	33	8
4	7	14	9	24	8	34	10
5	7	15	8	25	10	35	9
6	7	16	10	26	9	36	8
7	7	17	9	27	8	37	10
8	7	18	8	28	10	38	9
9	7	19	10	29	9	39	8
10	7	20	9	30	8	40	10

The rationalisable strategies are $8 \leq u_i^\star \leq 10$. The symmetric profile $(7, 7, 7)$ is no longer an equilibrium. There exist equilibria at $(8, 8, 10)$ and $(8, 9, 9)$, with $G_{-4} = 26$ and a Stage 2 contribution of 9; and at $(9, 10, 10)$, with $G_{-4} = 29$ and a Stage 2 contribution of 10.

B Adjusted p-values of the multiple pairwise comparisons

	Dominant			Complements			Substitutes		
$\overline{G}$	OWN vs. SCC	SCC vs. WCC	OWN vs. WCC	OWN vs. SCC	SCC vs. WCC	OWN vs. WCC	OWN vs. SCC	SCC vs. WCC	OWN vs. WCC
≤ 7	>0.469	>0.460	>0.898	>0.405	>0.333	>0.746	>0.255	>0.086	>0.746
8	<0.001	0.518		0.014	0.730		0.003	0.333	
9	0.011	0.246		0.014	0.094		0.020	0.587	
10	0.018	0.194		0.001	0.046		0.004	0.086	
11	0.003	0.011		<0.001	0.015		<0.001	0.072	
12	<0.001	0.003		0.003	0.020		<0.001	0.059	
13	<0.001	0.009		<0.001	0.007		<0.001	0.020	
14	<0.001	0.006	>0.746	<0.001	0.002	>0.746	<0.001	0.039	>0.541
15	<0.001	0.001		<0.001	0.001		0.001	0.062	
16	<0.001	<0.001		<0.001	0.001		<0.001	0.037	
17	<0.001	<0.001		<0.001	0.001		<0.001	0.017	
18	<0.001	<0.001		<0.001	<0.001		<0.001	0.046	
19	<0.001	<0.001		<0.001	<0.001		0.001	0.014	
20	<0.001	<0.001		<0.001	<0.001		<0.001	0.014	

Table 7: Pairwise comparisons of contribution strategies by information set $\overline{G}$. Each cell is the p-value of a Mann-Whitney-Wilcoxon test; these are adjusted for multiple testing using the Benjamini-Hochberg False Discovery Rate method. $\overline{G}$ at or below the Bayes-Nash contribution level are grouped, as no difference is expected in these contingencies.