1-13-2024

How Personalized Networks Can Limit Free Riding: A Multi-Group Version of the Public Goods Game

Aaron S. Berman
Chapman University, aaberman@chapman.edu

Laurence R. Iannaccone
Chapman University, iannacco@chapman.edu

Mouli Modak
Chapman University, modak@chapman.edu

How personalized networks can limit free-riding: A multi-group version of the public goods game*

Aaron S. Berman† Laurence R. Iannaccone‡ Mouli Modak§

January 13, 2024

Abstract

People belong to many different groups, and few belong to the same network of groups. Moreover, people routinely reduce their involvement in dysfunctional groups while increasing involvement in those they find more attractive. The net effect can be an increase in overall cooperation and the partial isolation of free-riders, even if free-riders are never punished, excluded, or recognized. We formalize and test this conjecture with an agent-based social simulation and a multi-good extension of the standard repeated public goods game. Our initial results from three treatments suggest that the multi-group setting indeed raises overall cooperation and dampens the impact of free-riders. We extend our understanding of this setting by imposing greater heterogeneity between groups through interweaving automated bot players amongst human subjects; whereby initial sessions of this amplify the aforementioned effects.
Keywords: cooperation, PGG, lab experiment, multi-group
JEL Codes: C72, C73, C91, C92, H41

*We would like to thank David Porter, Stephen Rassenti, Jared Rubin, Bart Wilson, Jason Shachat, Kevin James, and Tristan Tran for their helpful comments and suggestions. We would also like to thank the Economic Science Institute, ESI, and the IFREE Summer Scholars of 2023, as well as Megan Luetje for her assistance in our laboratory sessions. Any errors are entirely our own.

†Economic Science Institute, Argyros College of Business and Economics, Chapman University, 1 University Drive, Orange, CA 92866, USA, email: aaberman@chapman.edu

‡Institute for the Study of Religion, Economics and Society, Argyros College of Business and Economics, Chapman University, 1 University Drive, Orange, CA 92866, USA, email: iannacco@chapman.edu

§Smith Institute for Political Economy and Philosophy, Chapman University, 1 University Drive, Orange, CA 92866, USA, email: modak@chapman.edu

1 Introduction

In the laboratory, experimenters strive for realism in their design to ensure external validity but are simultaneously bound by a need to maintain analytical control. This realism-tractability trade-off is evident in the social dilemmas literature. The public goods game (PGG), a subset of the social dilemmas space, explores the factors that are omnipresent in social dilemmas – the individual incentive to free-ride. Numerous papers implement institutional remedies to free-riding that are costly and require bountiful information to successfully circumvent this behavior. Although often successful in reducing the proclivity of free-riding, these well-studied mechanisms are only applicable to a subset of *all* public goods. For instance, punishment, as implemented by Fehr and Gächter (2000), is appropriate in large scale public goods such as tax collection that have the capacities to credibly threaten and utilize coercion. In contrast, smaller scale public goods, a group project, congregation, family unit etc., may be unfit or incapable of retaliating against free-riding behavior via fines, excommunication, or other coercive tactics. Despite their prevalence in the laboratory setting, these institutional mechanisms are insufficient to explain the persistence of smaller scale public goods: club goods.

There *must* exist another factor that is currently omitted by the standard public goods literature which enables the survival of these groups. We propose the multi-group dimension as such a factor with the potential to innately remedy free-riding behavior. The multi-group setting proves to be a natural inspection as outside of the lab, individuals are simultaneously involved in a plethora of social groups and must allocate their resources across these groups. Therefore, in their decision-making process it is not merely a decision of cooperate or defect, but rather *where* and *when* to cooperate or defect at any moment.

The multi-group setting does not require one to possess a wild imagination and conjure up a convoluted thought experiment to visualize the pulls that an individual faces from various social groups in which they must constantly decide how to allocate their time, effort, and money to. Picture Noah, a long-standing Boy Scout, devoted to his current chemistry group project, and a member of a synagogue that he attends with his family each Saturday. With the three-group subset of Noah's social groups, Noah enters a constrained optimization problem in which he must decide, consciously or unconsciously, how to invest his various endowments across his groups whereby he faces a trade-off with each allocation. If Noah attends synagogue with his family, that is inherently time that he cannot be working on his chemistry project, unless Noah plans to attend services, but not actively participate by working on his phone: not *fully* contributing to either.

With Noah's situation in mind, we explore the multi-group extension of the PGG in the lab by placing individuals in *two* groups under *two* endowment schemes with varying degrees of realism: the multi-group public goods game. Having subjects participate in multiple games simultaneously, often two, has received attention in various 2-player games Bednar et al. (2012), in coordination games Cason, Savikhin and Sheremeta (2012), and in contests Savikhin and Sheremeta (2013), but has received minimal attention in the realm of social dilemmas – with notable exceptions Falk, Fischbacher and Gächter (2013), McCarter, Samek and Sheremeta (2014). A throughline of the standing experimental multi-group literature has shown that the multi-group structure alters individual behavior, and this deviation is attributed to the increased strategic uncertainty and number of social interactions.

In the multi-group setting, we consider two cases for handling the endowment of subjects; (1) each subject receives a separate endowment for each group they are a member of, Multi-Split, and (2) each agent receives one endowment that is shared across all their groups, Multi-Shared. We include the Multi-Split institution as it directly analogs the standard PGG and serves as a baseline for the endowment handling within the multi-group setting. The Multi-Split ensures that the two simultaneously played games remain structurally independent. The Multi-Shared handling maintains independence of non-overlapping group membership but relaxes the economic independence of the public goods; the Multi-Shared treatment makes the trade-off inherent in social group investment salient through its one-endowment, two-group structure. In the existing multi-group PGG literature, Falk et al. utilize the Multi-Split handling, but lacks consideration of the Multi-Shared design, whereas McCarter et al. treat the Multi-Shared design as given, but provide no comparison to the Multi-Split nor the single-group PGG.

Our primary includes both multi-group endowment handlings, and the single-group PGG as a baseline. We find that the multi-group treatments produce higher average investment in the first period, innate proclivity to invest, but only the Multi-Shared treatment produces increased investment that persists over the repeated game horizon when compared against both the Multi-Split and single-group treatments. We attribute the investment dominance of the Multi-Shared treatment to the heightened capacity for players to shift their investments here that is weaker or absent in our other treatments.

We opted to use the well-studied player typings of the PGG Fischbacher, Gächter and Fehr (2001) as a means to understand the impacts of the multi-group setting on average investment to the public good. This analysis proved less than fruitful as these typings did not characterize subject behavior in either multi-group PGG setting, and the random sample of our primary study lacked heterogeneity of the player types. This absence of diversity in typings dampens the potential gains from the inclusion of the multi-group setting (i.e., if a player is in two groups which are nearly composed of identical players, any shift in investment would (a) not be incentivized to occur, or (b) be reduced in magnitude). Therefore, we also produce an extension study (7) in which we correct for the homogeneity in player typings by imposing heterogeneity by bringing bot players into the mix. These bot players are programmed to behave as classic free-riders, and cooperators, and are positioned on the grouping lattice to ensure each subject has a more and less favorable group, in expectation. We find the amount that subjects invest to their group accounts in any period to be similar regardless of the presence of bots, but subjects better allocate with the presence bots. We note this to do to the decrease in strategic uncertainty in the extension's modified multi-group setting. We also find that the average payoffs of free-rider bots are lower than that of all other-typings, human and bot subjects, when in the Multi-Shared setting. This hints that free-riders face isolation as many subjects practice a 'leave with their money' strategy.

The remainder of the paper comprises of the multi-group Literature Review (2), the Experimental Design of the primary study (3), the Simulation of the multi-group setting (4), the Hypotheses (5), the primary experiment Results (6), the Extension study (7), and ends with the Conclusion (8).

2 Literature Review

We direct you to Ledyard (1995) for an excellent synopsis and entry point into the PGG literature and to Zelmer (2003) for a meta-analysis of the linear voluntary contribution mechanism, VCM, form of the single-group PGG that we are extending to the multi-group setting. With ample access to the breadth of the standard PGG literature in which each paper provides an adequate foundation for what a public good is and how they are tested in the laboratory, we focus our literature review on the less often curated, multi-group PGG studies.

Falk et al. (2013) extend the linear-VCM to the multi-group using a grid-grouping mechanism to ensure non-overlapping group membership and compare this setting against the single-group standard. In their experiment, subjects receive two identical but separate endowments for the two groups that they are members of. Their design nearly mirrors our Multi-Split treatment, but we account for the total endowment size differently. Subjects in Falk et al. each receive 20 points in each period for each group they are a member of, regardless of whether they are in the single-group baseline or 'Multi-Split' treatment. Subjects in our Multi-Split treatment each receive 10 points for each group they are a member of, and subjects in our single-group baseline receive 20 points for their one group. Our design controls for endowment size and aids in in cross-treatment comparisons with the baseline and with the Multi-Split treatment. Their paper uncovers 'social interaction' effects in which individual behavior is swayed in different directions from the behavior of their respective groups – individuals may act differently in their first versus their second group based on the individuals that they are interacting with.

McCarter et al. (2014) also produce a multi-group linear-VCM. Their subjects simultaneously interact in two groups for 20 periods and seek to determine if individuals continue to behave as conditional cooperators in the multi-group setting, as has been readily shown in its single-group counterpart. To answer their hypothesis, they have two handlings of group membership in the two-group, multi-group space; the *same* treatment in which the subject's groups are composed of identical members, and the *different* treatment in which the subject's groups have no membership in common. Both treatments feature a 'Multi-Shared' endowment of 160 points and they find that investments, at the treatment level across rounds, are higher in the *different* treatment as opposed to the *same* treatment. They attribute the treatment difference to individuals being better able to find other conditional cooperators when interacting with a greater number of individuals. Through their incorporation of a shared endowment across both treatments, McCarter et al.; lose access to the comparison between the Multi-Split and Multi-Shared endowment, to evaluate the behavioral change, if any, of making the transition to further realism; confound their research agenda of analyzing how group composition impacts free-riding with an untested experimental treatment – the incorporation of the shared endowment; and through the exclusion of the single-group PGG as a baseline, they are unequipped to determine whether subject behave differently in the multi-group setting. Prior to their publication, we have been unable to uncover multi-group literature with a Multi-Shared endowment scheme, and ex-post, any behavioral differences from handling the endowment from split to shared has been unstudied or at least unpublished. Experimental studies that have been launched in the wake of McCarter et al. take the shared endowment as a prior in their experimental design such as with Kreig and Samek

(2016) where they seek to determine best practice for charities when in amidst competition for donations.

Our study builds upon the existing multi-group PGG literature by connecting the Multi-Split treatment of Falk et al. and the Multi-Shared treatment of McCarter et al. Our Multi-Split design maintains the non-overlap conditions of both, and controls for the Multi-Split total endowment size for tractable comparisons. We improve upon the Multi-Shared design by using a more conventional endowment size, as well as conducting an experiment with two-levels of baselines to evaluate the gains from the Multi-Shared. These changes along with our designs access to player typings allow us to quantify gains, if any, from the multi-group dimension, and from the further shift from the Multi-Split to Multi-Shared endowment.

3 Experimental Design

3.1 Treatments

The experiment has a baseline, the single-group PGG, and two multi-group treatments, Multi-Split and Multi-Shared, which vary the nature of the endowment that subjects receive. We utilized between subject design so each subject participated in at most one of the three treatment groups. All sessions of the primary experiment were conducted in the following order; repeated PGG or repeated multi-group PGG treatment, risk elicitation task Laury and Holt (2002), linear-VCM strategy method Fischbacher, Gächter and Fehr (2001), and end with a post-play questionnaire. The experimental materials are available in full via Appendix A.

We opt for the multi-group treatment to begin all experimental sessions as we want to take advantage of the inexperienced first-period investment decision of subjects. In the public goods literature, this metric is interpreted as being representative of subject's innate propensity to invest to the group account(s) given the PGG setting, instruction framing, etc. Since the strategy method is not the core purpose of our study, we are content with and acknowledge that the typings produced will be from experienced subjects. The experienced typings were found to be representative of subject's PGG type by Fischbacher, Gächter and Quercia (2012) who utilize the within-subject design in combination with the strategy method and repeated play, and for the Single Group PGG find that the order in which subjects experience experienced each, has no significant effect on typings and repeated play decisions.

We place the risk elicitation task in between the repeated PGG and the strategy method as a buffer to break up the history of the two: an attempt at mitigating sequential behavior spillover via a form of a cognitive reset task. Below, we describe each experimental component and treatment.

3.1.1 Experimental Treatments

We utilize bank-oriented diction rather than the connotatively burdensome words of cooperation or contribution. In the single-group baseline, we pose the decision as an investment to your Group Account, and in both of the multi-group treatments we pose the decision as an investment to your Blue Group Account and your Green Group Account, with the Personal Account presentation held constant. We follow McCarter et al's convention of using colors to distinguish between the two groups rather than the use of numbered groups, Group 1 and Group 2, of Falk et al.[1] We opted for the color delineator as it does not reinforce an arbitrary hierarchy of the groups nor pre-impose an order for subjects to consider them. In building the multi-group presentation, we follow the precedent set by both McCarter et al. and Falk et al. by presenting the two group decisions horizonallty, on one page, with one group on the left and another on the right.[2] We chose the bulk of our parameters from standard choices

[1]The use of colored groups is prevalent in many papers that study individual behavior in multiple groups Savikhin and Sheremeta (2013), McCarter, Samek and Sheremeta (2014), Cason, Savikhin and Sheremeta (2012) where Blue and Green being common selections as to not elicit color-dependent emotional responses Valdez and Mehrabian (1994).

[2]This is potentially problematic as English is read left-to-right and a subject *may* unconsciously make their decision in the left group before they consider the right group – despite the intended simultaneous

within the linear-VCM literature with deviations when necessary.

For our multi-group sessions, we enforce that their is no overlap in group membership for a subject's two groups besides himself, which is known to subjects. We achieve this feat by placing each subject on a node on a torus and grouping them accordingly. For the group size of four and participating in two groups simultaneously, the smallest session size possible is 16, and thus we use these parameters for the primary experiment sessions.

Single

In each period a subject receives an endowment of 20 points where any integer amount, inclusive, between 0 and 20 can be invested to their Group Account with the remainder entering their Personal Account. For ease of explanation to subjects, the multiplier, marginal per capita returns (MPCR), is set to 0.5 for all treatments. Each group is composed of 4 members with known continuation of 20 periods. In any period, a subject's payoff is

$$\pi_{Single} = e - c_i + \frac{r}{N}\left(c_i + \sum_{j \neq i} c_j\right), \tag{1}$$

where e is the endowment, c_i is a subject's investment to their Group Account, r is the scale factor of 2[3], N is the group size of 4, and $c_i + \sum c_{-i}$ is the group's total investment to the Group Account.

Multi-Split

For comparability between the single and multi-group treatments, each treatment's total endowment(s) sum to 20 points. In the Multi-Split treatment, each subject receives two endowments of 10 points, one for their Blue group and one for their Green group. Both

nature of the decision.

[3]To maintain our environment as a social dilemma our $\frac{r}{N}$, MPCR, needs to satisfy the following inequality $\frac{1}{n} < \frac{r}{N} < 1$.

groups have the same MPCR of 0.5 In any period, a subject's payoff is

$$\pi_{Split} = (e_{\mathrm{B}} - c_{i,\mathrm{B}}) + \frac{r}{N}\left(c_{i,\mathrm{B}} + \sum_{j \neq i} c_{j,\mathrm{B}}\right) + (e_{\mathrm{G}} - c_{i,\mathrm{G}}) + \frac{r}{N}\left(c_{i,\mathrm{G}} + \sum_{j \neq i} c_{j,\mathrm{G}}\right), \tag{2}$$

with added B and B subscript to represent the Blue and the Green choice variables, respectively.

Multi-Shared

The Multi-Shared treatment only deviates from the Multi-Split in the handling of the endowment. In a Multi-Shared session, a subject receives an endowment of 20 points which they can invest however they please. In the Multi-Shared treatment, a subject can invest 20 points to one group and 0 points to the other which is an infeasible allocation under the Multi-Split regime. In any period, a subject's payoff is

$$\pi_{Shared} = (e - c_{i,\mathrm{B}} - c_{i,\mathrm{G}}) + \frac{r}{N}\left(c_{i,\mathrm{B}} + \sum_{j \neq i} c_{j,\mathrm{B}}\right) + \frac{r}{N}\left(c_{i,\mathrm{G}} + \sum_{j \neq i} c_{j,\mathrm{G}}\right) \tag{3}$$

which captures the Multi-Group dimension with B and G subscript, and recombines e to be 20 points.

3.1.2 Risk Elicitation

Our method for risk elicitation comes from Laury and Holt (2002). A subject is presented with a series of 10 pairs of lotteries. Lottery A has constant 50-50 chance of a high prize of \$3 and a low prize of \$1 whereas Lottery B has an increasing probability of receiving the high prize of \$5 and a decreasing chance of receiving the low prize of \$0.1. The number of times a subject opts for Lottery B over Lottery A provides a measure of their risk aversion. Hajikhameneh and L. (2023) found risk aversion, with this mechanism, to be insignificant in their Single Group PGG with additive shocks. Despite this, there is potential that risk may play an explanatory role in a subject's decision-making process in the multi-group version.

With additional avenues for investment, as well as a third avenue for investment being their Personal Account, risk may influence a subject to invest to one group over another based on prior period stability or perception thereof.

3.1.3 Strategy Method

We utilize the strategy method implementation of the PGG as presented by Fischbacher, Gächter and Quercia (2012). Like Fischbachet et al., we maintain the saliency of payments between the repeated PGG and strategy method by weighting the strategy method payment to the repeated play. We opt for the weighting of 1:10 where the strategy method payment is equivalent to 10 rounds of the repeated PGG as our repeated design has 20 periods and theirs has 10. From the strategy method, we type players in familiar archetypes, free-rider, hump-shaped, conditional cooperator, etc., and use the typings to simulate behavioral expectations in the analogous multi-group setting.

3.1.4 Questionnaire

The questionnaire contains standard demographic questions regarding age, gender, major, etc., but also includes a strategy elicitation portion. The strategy elicitation questions consist of open-ended questions that attempt to uncover the decision-making process of individuals in the experiment.

3.2 Logistics

The experiment was implemented via oTree (Chen, Schonger and Wickens, 2016). Each treatment had 5 sessions with 16 subjects per session. We conducted a pilot session in summer 2023 with IFREE Summer Scholars which was used to estimate our sample size, shown in Appendix B. All subsequent sessions were conducted at Chapman University in fall 2023 with ESI's recruitment platform and subject pool. We recruited subjects with no prior experience in the public goods environment, and no subject participated in more than one session. Subjects received a $7 show-up fee, their cumulative earnings from the 20 periods, their payment from their risk elicitation task, and strategy method choice. Total payments, including the show-up fee, averaged $27.38.

4 Simulation

We use a simulation to bolster our intuition regarding how involvement in multiple groups under various endowment schemes effects the behavior of individuals. With these predictions and results from the previous multi-group literature, we construct our hypotheses. A simulation of this type appears absent from the multi-group literature, and the model we utilize can readily explain the phenomenon in this setting. The learning model in question is the Experience-Weighted Attraction, EWA, developed in Camerer and Hua Ho (1999) along with the algorithm used in Bühren et al. (2023).

Our simulation environment mirrors our experimental design with the following deviations. First, in the experiment, players participate in 20 periods; however, in the model, the simulated players learn over hundreds of periods. Second, as for the nature of repetition, players in the lab know the terminal round, whereas, in the simulation, the simulated players have bounded rationality – they do not know when the game is supposed to end and how that should influence their choices. Simulated players choose the action that should give them a high payoff, given their experience. Third, players in experimental sessions are randomly grouped and we elicit their types in the strategy method block. As for the simulation, we run multiple iterations where we vary the distribution of the two utility types we consider – free-riders and conditional cooperators – and evaluate whether this distribution alters any treatment effect. To note, in both the experiment and the simulation, players know *their* utility function, but not the utility function of others. The simplified environment of our model allows for any observed differences between the treatments to be precisely attributed to the multi-group structure and the endowment schemes.

4.1 Model Details

In the EWA learning model, there are two key variables – $\hat{N}(t)$ (observation equivalent of past experience), $A_i^j(t)$ (player i's attraction for strategy $c_i^j(t)$ in period t). These two variables update using the experience in prior periods to optimize the given problem stochastically. Each player finds a solution that provides a high payoff under the PGG framework subject to the stochastic nature of the environment.

For the Single Group baseline, a strategy is the investment in one public good $c_i(t)$ where $c_i(t) \in \{0, 1, \cdots, e\}$, whereas for the multi-group treatments, Multi-Split or Multi-Shared, a strategy is the vector of investments in two public goods – $c_i(t) = (c_{i,B}(t), c_{i,G}(t))$, where $c_{i,k}(t) \in \{0, 1, \cdots, e_k\},\ k \in \{B, G\}$ for Multi-Split treatment and for the Multi-Shared treatment $c_{i,k}(t) \in \{0, 1, \cdots, e\},\ k \in \{B, G\}$ with $c_{i,B}(t) + c_{i,G}(t) \leq e$.

Parameter	Description	Value
r	Scale Factor	2
N	Number of players in group	4
	Number of players in lattice	16
e	Common endowment (Single, Multi-Shared Treatment)	10
e_B, e_G	Endowment for Blue and Green Groups (Multi-Split Treatment)	5

Table 1: Simulation Parameters

In each period t, the action $c_i(t)$ is chosen randomly with probabilities defined by

$$P_i^j(t) = \frac{e^{A_i^j(t-1)/\lambda_i(t-1)}}{\sum_{k=1}^{J} e^{A_i^k(t-1)/\lambda_i(t-1)}},$$

where $\lambda(t) \in (0, 1]$ measures the sensitivity of players to attractions. Given the choices $c_i(t)$

and $c_{-i}(t)$, we can find the payoffs for each player i. The payoffs depend on the utility type of the players. Table 2 lists the different utility functions we use in the simulation.

Given the player's choice, payoff, and others' choices, we can update A_i and λ_i. The updating of $\hat{N}$ is independent of these variables. The initial values $\hat{N}(0)$, $A_i^j(0)$, and $\lambda_i(0)$ are parameters that can be influenced by pregame experience. We initialize $A^j(0)$ for every player with a randomly drawn integer less than equal to the endowment available. The following equations show the updating for each of these variables.

- For $\hat{N}$

$$\hat{N}(t) = \rho\hat{N}(t-1) + 1, \quad t \geq 1 \tag{4}$$

  where ρ is a depreciation rate that measures the fractional impact of previous experience compared to one new period. Note that, $\hat{N}(t) \geq 1$.

- For A_i

$$A_i^j(t) = \frac{\phi\hat{N}(t-1)A_i^j(t-1) + \left(\delta + (1-\delta)I(c_i^j, c_i(t))\right)\pi\left(c_i^j, c_{-i}(t)\right)}{\hat{N}(t)} \tag{5}$$

  Here, ϕ is a discount factor to depreciate previous attractions. The parameter $\delta \in [0,1)$ is the weight provided to the strategies not chosen by player i. $I(c_i^j, c_i(t)$ is the indicator that strategy choice at period t is strategy j and $\pi\left(c_i^j, c_{-i}(t)\right)$ is the payoff given player i's strategy j and others' choice at period t.

- For λ_i

$$\lambda_i(t) = \begin{cases} \max\left(\underline{\lambda}, (1-\lambda_\Delta)\times\lambda(t-1)\right), & c_t^* = c_{t-1}^* \text{ \& } c^*(t), c^*(t-1) \text{ are not null} \\ \min\left((1+\lambda_\Delta)\times\lambda(t-1), \bar{\lambda}\right), & c_t^* \neq c_{t-1}^* \text{ \& } c^*(t), c^*(t-1) \text{ are not null} \\ \lambda_i(t-1), & \text{otherwise} \end{cases} \tag{6}$$

  where c^* is the strategy that has maximum attraction only if unique. We dynamically adapt players' sensitivity to attraction in lines of Bühren et al. (2023). High values of λ implied an exploratory nature of the learning process. We want the players to explore their strategy in the early period and gradually reduce λ over time. λ increases if the algorithm finds that the optimally attractive strategy in a period differs from the last period. This allows exploring other strategies by increasing their probability of being chosen. When the optimal strategy (maximum attraction) does not change over time, λ goes down to increase the probability of choosing that strategy.

Each player type in our simulation, altruists, conditional cooperators, and free-riders, have utility functions defined in Table 2 and the selected values for the aforementioned parameters are presented in Table 3.

Type	Function			Parameter Details
	Single	Multi-Split	Multi-Shared	
Altruist	$\pi_i = c_i$	$\pi_i = c_{i,B} + c_{i,G}$	$\pi_i = c_{i,B} + c_{i,G}$	-
Conditional Cooperator (CC)	$\pi_i = \pi_i^{FR} + c_i^{\alpha} c_{-i}^{1-\alpha}$	$\pi_i = \pi_i^{FR}$ $+c_{i,B}^{\alpha} c_{-i,B}^{1-\alpha}$ $+c_{i,G}^{\alpha} c_{-i,G}^{1-\alpha}$	$\pi_i = \pi_i^{FR}$ $+c_{i,B}^{\alpha} c_{-i,B}^{1-\alpha}$ $+c_{i,G}^{\alpha} c_{-i,G}^{1-\alpha}$	$\alpha \in (0,1)$ $\alpha = 0.25$
Free-Rider (FR)	Equation 1	Equation 2	Equation 3	-

Table 2: Utility Details

Parameters	Description	Value
$\lambda(0)$	Initial Value of $\lambda(t)$	0.7
$\underline{\lambda}$	Lower bound of $\lambda(t)$	0.05
$\bar{\lambda}$	Upper bound of $\lambda(t)$	1
λ_Δ	Change in $\lambda(t)$	0.05
$\hat{N}(0)$	Initial value of $\hat{N}(t)$	1
ρ	Depreciation Rate of previous observation	0
ϕ	Discount factor of previous attraction	0.9
δ	Weight of not chosen strategies	0.6

Table 3: EWA Parameters

Finally, for all our simulation exercises, we run 30 or 50 iterations, and the results presented are the averages of the iterations. The pseudo-code for this simulation can be found in Appendix C.

4.2 Model Verification

We first verify whether our learning model operates appropriately. A natural test is to run iterations with a sole player-type in the population for which we have intuition regarding what investments should resemble; we simulate this setting where altruists, free-riders, and conditional cooperators constitute 100% of the population. Altruists should invest any allocation that expends their entire endowment, free-riders should invest nothing, and conditional cooperators will respond to one another based on their utility functions. Figure 1 shows the results of this exercise.

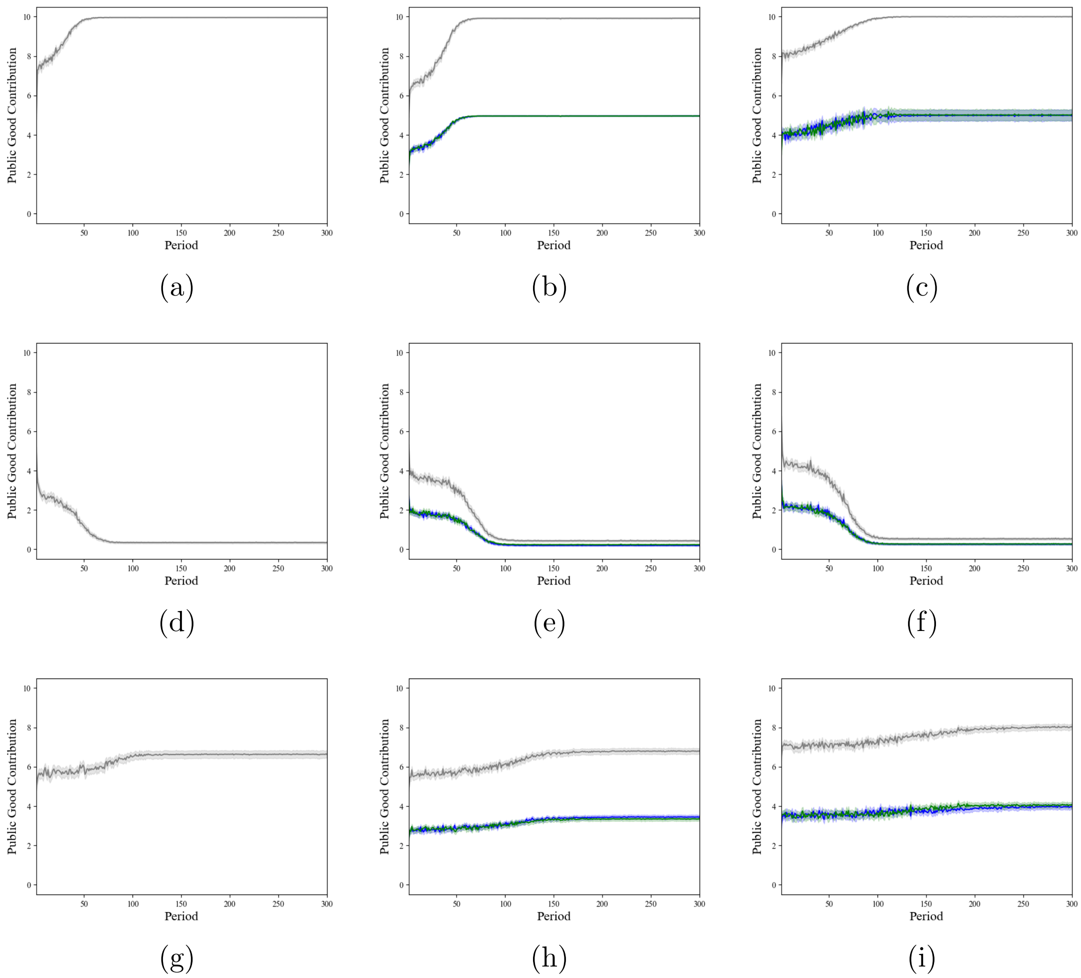
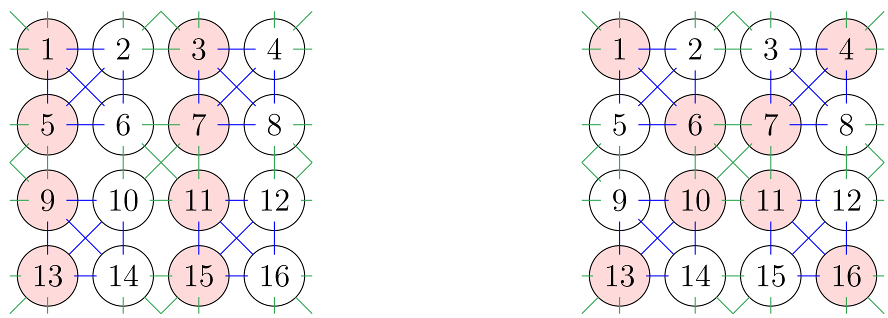


Figure 1: Simulation Verification

Notes: The columns depict the treatments, single, Multi-Split, and Multi-Shared respectively, Whereas the rows are player-types, altruist, free-rider, and CC. Each graph is composed of 30 iterations for 300 periods.

4.3 Simulation Results

With our simulation explained, and evaluated, we now utilize it to discover the impacts of group composition on the investment behavior of its members. For this inquiry, we produce two patterns of group composition on our lattice: Alternating and Segregating (see Figure 2). In the Alternating pattern, each group (Blue or Green) has two free-riders and two conditional cooperators. With this pattern, there is no ex-ante difference between the two groups any player is a member of. Therefore, we expect a player's investment to be the same to both groups. In the Segregating pattern, each Blue group has two free-riders and two cooperators, and there are two Green groups with four cooperators and two Green groups of four free-riders. In this pattern, players 1, 4, 6, 7, 10, 11, 13, and 16 are in Blue groups which have equal distribution of free-riders and cooperators and Green groups which only have other conditional cooperator players. We expect that these players will shift their investments towards their Green group as it is composed of more cooperators than their Blue groups. Players 2, 3, 5, 8, 9, 12, 14, and 15 are free-riders and will not invest to either group. In the aggregate, we should find that the Green groups have higher investment than the Blue groups in the segregating pattern.



Alternating Pattern

Segregating Pattern

Figure 2: Distribution of Utility Types

Notes: The red nodes represent Free-Riders and the white nodes represent any utility type that favors cooperation. The blue line represents the connection between nodes in a Blue Group. The green line represents the connection between nodes in a Green Group. The Green Groups are {1, 4, 16, 13}, {2, 3, 15, 14}, {5, 8, 12, 9}, {6, 7, 11, 10}. The Blue groups are {1, 2, 6, 5}, {3, 4, 8, 7}, {9, 10, 14, 13}, {11, 12, 16, 15}.

We produce Figure 3 and 4 which provide these simulated results for both the Multi-Split, Figure 3, and Multi-Shared, Figure 4, treatments. For both, as we anticipated, under the Alternating pattern there is no differentiation between investments to one's Blue and Green group, see the bottom-left most graph in both Figures. Whereas we do see the anticipated distinction between a player's investment to their groups in the Segregating pattern – the bottom right most graph in both figures. Interestingly enough, this difference under the Segregating pattern is heightened under the Multi-Shared, the gap between the investment between the Green and Blue group is far larger under the Multi-Shared than the Multi-Split treatment. From this we bolster our intuition with our simulation which showcases that the Multi-Shared treatment is responsive to the composition of groups, and players can be more receptive of discrepancies between their groups due to the *shared* nature of the endowment that is absent in the standard public goods literature.

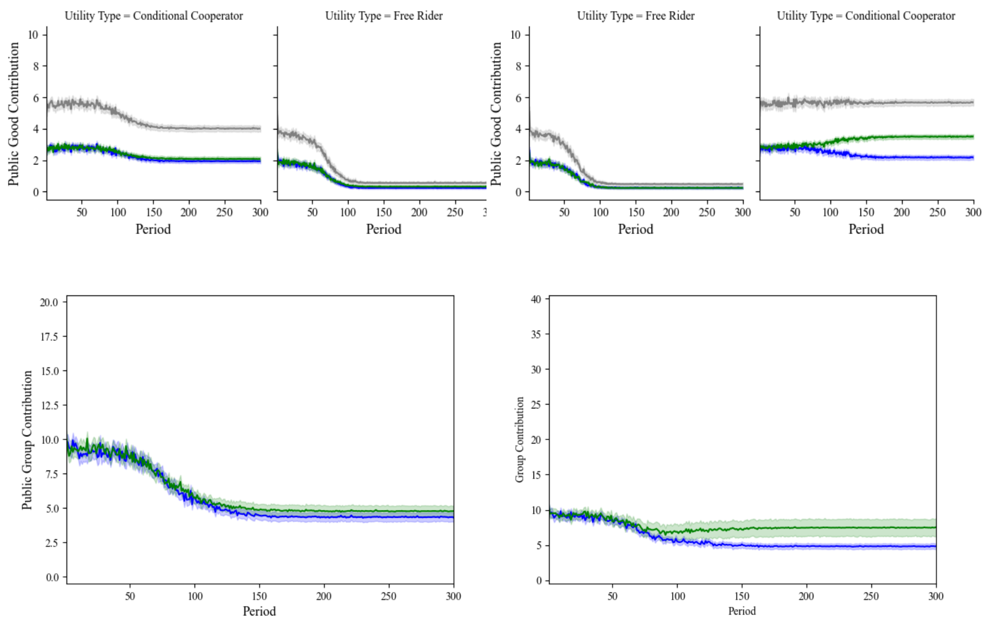


Figure 3: Multi-Split Treatment

Notes: Left Panel is for Alternating Pattern, Right Panel for Segregating Pattern.

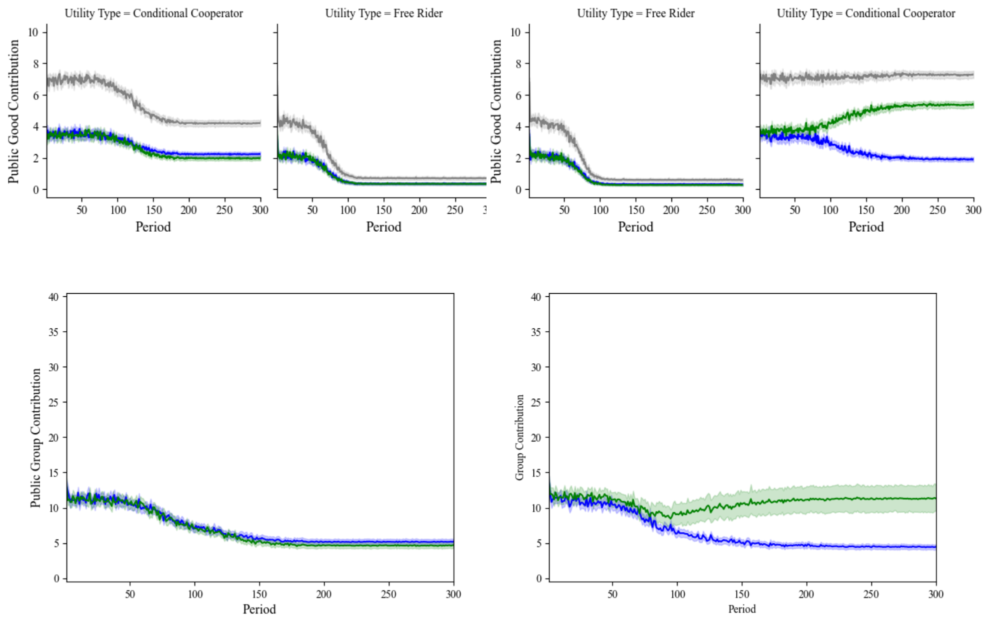


Figure 4: Multi-Shared Treatment

Notes: Left Panel is for Alternating Pattern, Right Panel for Segregating Pattern.

5 Hypothesis

Previous multi-group literature has emphasized social interaction effects in affecting the decision-making behavior of subjects in simultaneously played games. These social interaction effects manifest as behavioral spillovers where one's experience in one game affects their behavior in the other. For the simultaneous multi-group public goods game, this enables subjects to compare their groups based on the period history both *within* and *across* groups. To this end, we focus on our three treatments and generate falsifiable hypotheses.

Hypothesis 1

(Single vs. Multi-Split) On average, we anticipate more investment in the group's joint account under the Multi-Split than under the single group baseline.

(Single vs. Multi-Shared) On average, we anticipate more investment in the group's joint account under the Multi-Shared than under the single group baseline.

(Multi-Split vs. Multi-Shared) On average, we anticipate more investment in the group's joint account under the Multi-Shared than under the Multi-Split treatment.

The hypothesized directions are bolstered by the player-typing data we collect from each subject via the strategy method. We are particularly interested in the behavior of free-riders and conditional cooperators in the multi-group setting. With access to each subject's typing, we hypothesize the following treatment and type-specific directions.

Hypothesis 2

(Free-riders) Free-riders should behave consistently in all treatments.

(Conditional Cooperators) Conditional cooperators attribute for the difference between multi-group and the single group baseline, and are responsible for the furthered difference between the multi-group treatments.

6 Results

We first inspect our treatment results (6.1) in isolation, and bolster our analysis with the player typings (6.2).

6.1 Treatment Results

6.1.1 First Period Investments

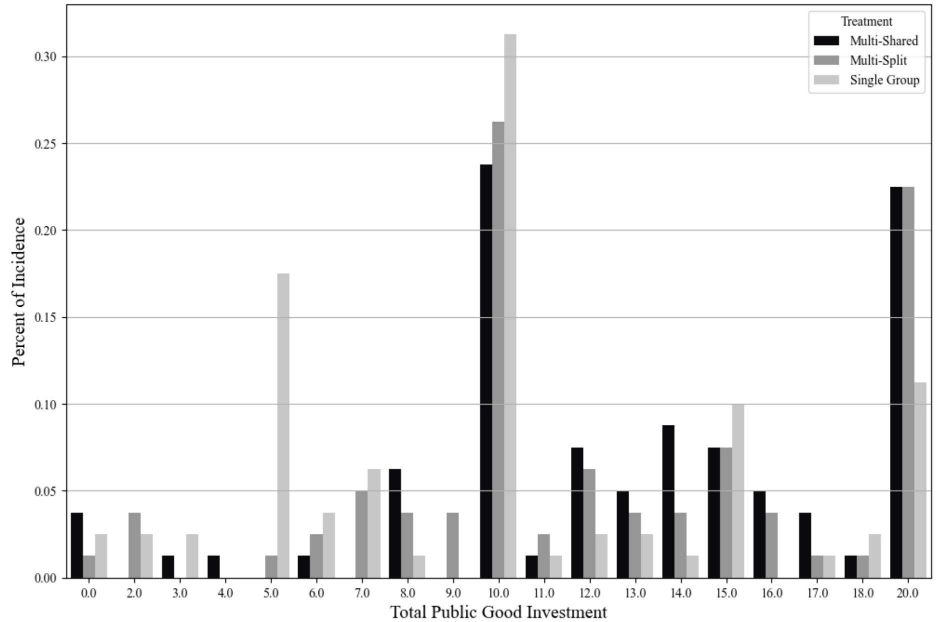
As noted in our Experimental Design (3), a subject's first-period investment is often used as a measure of innate proclivity to cooperate within this literature. Therefore, the choices can be expected to be unaffected by the history of the environment. The average first-period investment for the Single Group baseline is 10.24 points, for the Multi-Split treatment, 6.34 points in the Blue group and 6.26 points in the Green group, a combined investment of 12.60 and for the Multi-Shared treatment, 6.89 points and 6.36 points in the Blue and Green groups respectively, a combined investment of 13.25. Treating each individual's first-period investment as an observation, Tukey's multiple comparison test (Tukey, 1949) reveals that there is not a significant difference between Multi-Split and Multi-Shared first-period behavior (p-value of 0.717) in terms of the total investment, but the investment of both multi-group treatments are significantly different from the investment in the single group baseline (p-values of 0.012 and 0.001 respectively). This suggests that novice subjects, who have only read depictions of their respective environments via instructions and answered subsequent comprehension questions, respond differently to being in a single versus multi-group setting. The multi-group setting incurs a greater average investment to group accounts, but the nature of the endowment lacks significant explanatory power at the games inception.

We also look at the empirical distribution of the first period investments. In Figure 5, the top plot presents the frequency of total investment in public goods and the bottom panel provides the frequency of investment for the multi-group sessions split by group color. For the Single Group baseline, since there is only one public good, total public good investment is the investment in that public good. In multi-group treatments, the total public good is the sum of investments in Blue and Green public goods. We find that in the Multi-Split treatment, more than 20% of choices are 10 points (all of their endowment is invested in that account) in each account, but not in both accounts simultaneously. The Multi-Shared treatment has a similar pattern, the difference being there are some incidences where subjects invested more than 10 points in either of the accounts. In all treatments, we find that a proportion of subjects have invested all their endowment in the public good (see the top panel), but the percentage is almost double in the multi-group treatments than in the Single Group treatment.[4] Another intriguing takeaway is that less than 5% of the subjects invested nothing in each treatment.[5]

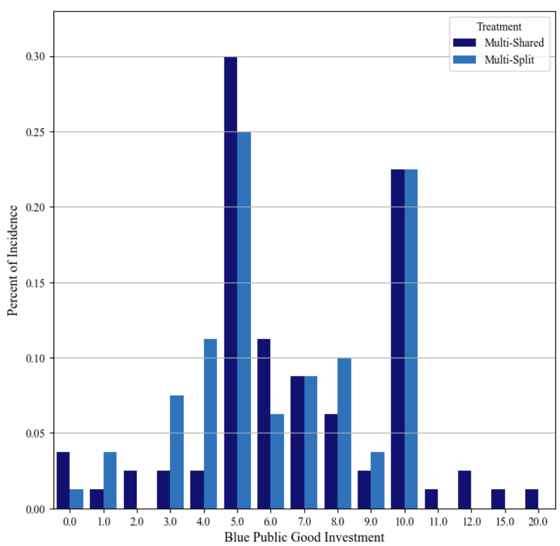
Given the empirical distribution of the choices, we compare the distributions using the

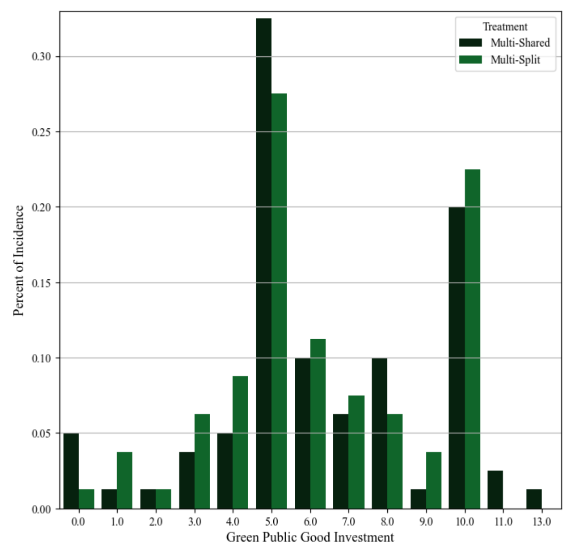
[4]The percentage is 22.5% for multi-group treatments and 11.25% for Single Group treatment. However, the difference is not statistically significant.

[5]The percentage is 1.25% and 3.75% for Multi-Split and Multi-Shared treatments, respectively, and 2.50% for Single Group baseline. However, the difference is not statistically significant.



All Treatments – Total Investment





Blue Account

Green Account

Figure 5: First Period Investment Frequencies by Treatment and Group

Krusal-Wallis test. We find that the distribution of choices in Single Group treatment is significantly different from those in Multi-Split (p-value of 0.004) and Multi-Shared (for α-level less than 0.001) treatments. However, we can not reject the null hypothesis (p-value of 0.289) that the first-period choices in the two multi-group treatments come from the same distribution. This supports our result from the Tukey test that we presented earlier

	Total PG Investment (1)	Blue Account PG Investment (2)	Green Account PG Investment (3)
Constant	10.485*** (0.822)	6.323*** (0.219)	6.248*** (0.274)
Treatment Effect			
Single Group	(BASE)	-	-
Multi-Split	2.822** (1.038)	(BASE)	(BASE)
Multi-Shared	3.432** (1.215)	0.751 (0.472)	0.369 (0.653)
$\chi^2_{(1)}$ $H_0 : \beta_{SG} = \beta_{MSp}$	7.97*	-	-
$\chi^2_{(1)}$ $H_0 : \beta_{SG} = \beta_{MSh}$	7.39*	-	-
$\chi^2_{(1)}$ or $F_{(1,159)}$ $H_0 : \beta_{MSp} = \beta_{MSh}$	0.29	2.54	0.32
$\chi^2_{(2)}$ $H_0 : \beta_{SG} = \beta_{MSp} = \beta_{MSh}$	10.19**	-	-
# Observation	240 [6, 189, 45]	160 [4, 133, 23]	160 [5, 136, 19]

Table 4: Coefficients from Tobit Regression on First-Period Choices

Notes: This table shows the coefficients and test statistics from the Tobit regression of total public goods investments on treatment dummies in column (1) and account-specific public good games in columns (2) and (3). We report the chi-squared statistic for the test of equality between coefficients of different treatments in column (1) and F-statistic in columns (2) and (3). The robust standard errors are listed in parentheses below the coefficients. The array in the row "# Obsversation" shows the breakdown of censored data, the first showing the observation at the lower bound at 0, the second showing the unbounded data, and the third, the number of observations at the upper bound at 20.

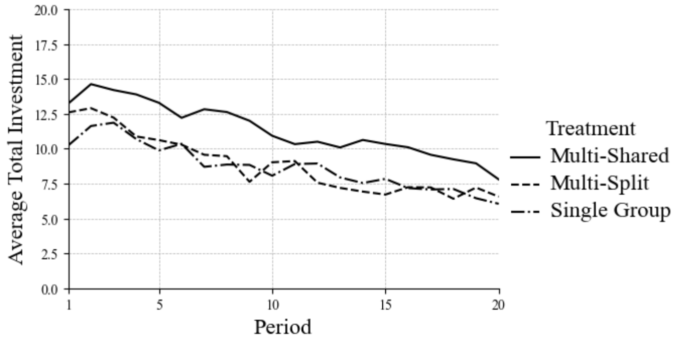
Symbols: p-values *** < 0.001, ** < 0.01, * < 0.05, $^+ < 0.1$

(see column (1) in Table 4). Another robustness check we perform is a Tobit regression of the first-period choices on the treatments. This again confirms that even novice subjects perceive a difference between the Single Group baseline and the multi-group treatments and invest differently. However, the finding that there is no statistical difference between Multi-Split and Multi-Shared treatments supports our theory that the difference between the two treatments should stem from the beliefs regarding the types of group members formed by looking at the history of the game. There is no ex-ante cause to differentiate between the two treatments. To further elucidate this fact, we look at the Blue and Green account investments. We compare the distributions of investments in each account of the two treatments and can not reject the null hypothesis that the data is generated from the same distribution (p-value of 0.25 for Blue Account and p-value of 0.715 for Green account). The result is also supported by a Tobit regression of choices of the treatments (see columns (2) and (3) in Table 4). Furthermore, the subjects should not distinguish between the two accounts in the multi-group treatments, which is held true in the data (p-value of 0.943 for Multi-Split treatment and p-value of 0.374 for Multi-Shared treatment from Krusal-Wallis test).

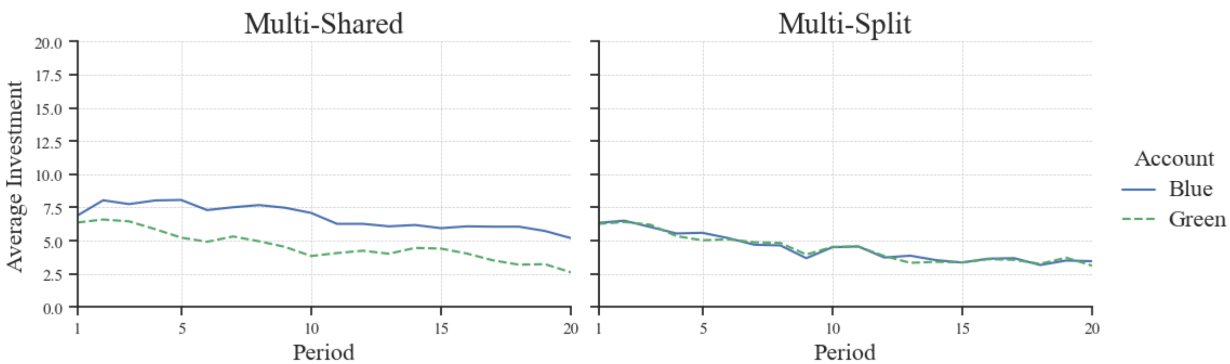
Result 1 *In Period 1, without prior experience of the environment, subjects invest more in public goods in multi-group environments than in a single-group environment.*

6.1.2 All Periods Investment

Now, we address the investments of subjects in the full duration of the experiment. Figure 6 displays the path of the average total public good investment. Visually, the investment path of the Multi-Shared treatment dominates its Multi-Split and Single Group treatment counterparts, while the paths for the Single Group and Multi-Split treatments helix about one another. We also see a distinct downward trend in the average investment in public goods in all treatments. Noteworthy in Figure 6, is the well-documented decline in investment to group accounts in the multi-group treatments over time and particularly at the end of the game (Zelmer, 2003; Ledyard, 1995); interestingly enough, our results in both Single Group and multi-group formulations, appear less plagued by investment decline and common end-game effects. However, this representation does not show the heterogeneity of the choices. Three important actions in the PGG are 0 (no public good investment), 10 (entire endowment invested in the public good(s) in Multi-Split treatment), and 20 (entire endowment invested in the public good(s) in Single Group and Multi-Shared treatments). 7 produces a bar plot of the frequency of total investment in all periods. Note that these frequencies of choices deviate from the first period plot 5. The frequency of investing all of their endowment in public goods increases for Multi-Shared treatment slightly but is the same for Multi-Split, dropping to around 10%. Similarly, the proportion of those investing 0 points increases drastically.



Average Total Public Good Investment by Periods



Average Account-Based Public Good Investment by Periods

Figure 6: Average Investment by Periods

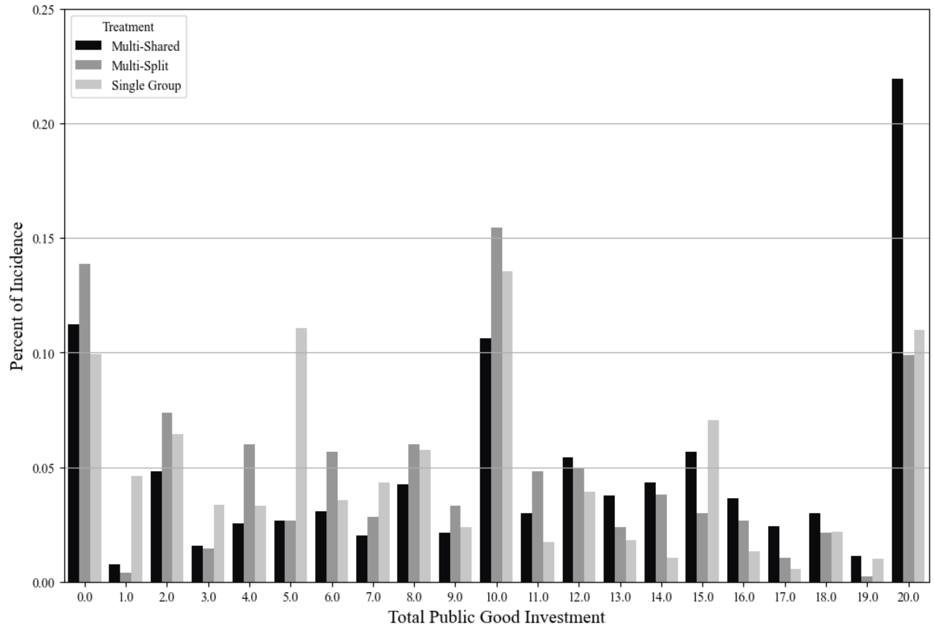


Figure 7: All Period Investment Frequencies by Treatment

We first use Tukey's multiple comparison test, where for each treatment, each subject's total public good investment averaged across all periods is treated as an observation. We find that the total public good investment in the Multi-Shared treatment is significantly higher than both the Multi-Split and Single Group treatments (p-values of 0.001 and 0.0003, respectively), whereas those of the Multi-Split and Single Group treatments are not statistically different (p-value of 0.969).[6] This non-parametric test does not account for our panel data, time trends, and censoring at the upper and lower bounds. Therefore, we complement the non-parametric test with a Panel Tobit Regression. The results of the regressions are listed in Table 5. These results confirm the results from the non-parametric test. We can reject the null hypothesis that the coefficients for Multi-Shared are the same as those of Multi-Split and Single Group. We can also reject the composite test that all the coefficients are the same. These hold even after considering treatment level time trends. There is a significant negative

[6]For robustness, we perform the non-parametric Kruskal-Wallis and post-hoc Connover Iman test (Conover and Iman, 1979). The resulting Kruskal-Wallis H-test statistic is 17.74, producing a p-value of 00014. The post-hoc component validates this finding with a shared-single p-value of 0.0002, a shared-split p-value of 0.002, and single-split of 1.

time trend in Single Group treatment as is expected with finite repetition. Moreover, this is extended to the multi-group treatments as well.

	Total PG Investment (1)	Total PG Investment (2)
Constant	8.735***	12.129***
	(0.719)	(0.764)
Treatment Effect		
Multi-Split	-0.181	0.756
	(1.017)	(1.08)
Multi-Shared	3.408**	4.481***
	(1.02)	(1.088)
Time Trend		
Single Group	-	-0.326***
		(0.026)
Multi-Split	-	-0.413***
		(0.026)
Multi-Shared	-	-0.428***
		(0.027)
ρ	0.505	0.545
$\chi^2_{(1)}$ $H_0 : \beta_{SG} = \beta_{MSp}$	0.03	0.49
$\chi^2_{(1)}$ $H_0 : \beta_{SG} = \beta_{MSh}$	11.16**	16.94***
$\chi^2_{(1)}$ $H_0 : \beta_{MSp} = \beta_{MSh}$	12.37**	11.17**
$\chi^2_{(2)}$ $H_0 : \beta_{SG} = \beta_{MSp} = \beta_{MSh}$	15.67***	19.33***
# Observation	4800	4800

Table 5: Coefficients from Panel Tobit Regression on All-Periods Choices

Notes: This table shows the coefficients and test statistics from the Panel Tobit regression of total public goods investments on treatment dummies (with Single Group as the control treatment) in column (1) and separate time trends in column (2). The data is censored at 0 at the lower end (561 observations) and 20 (685 observations) at the upper end. The robust standard errors are listed in parentheses below the coefficients. Symbols: p-values $^{***} < 0.001$, $^{**} < 0.01$, $^{*} < 0.05$, $^{+} < 0.1$

Result 2 *Considering all periods, subjects invest more in public goods in Multi-Shared treatment compared to Multi-Split and Single Group treatments.*

Our results partially satisfy our hypotheses. We find that subjects consistently invest more in the Multi-Shared treatment than the Single Group baseline. However, the increased investment in the Multi-Split treatment dissipates over time. This warrants investigation pertaining to the cause of these effects. To better understand how subject's alter their investments across their group accounts and personal accounts, we use a Panel Tobit regression to evaluate how the game's history impacts investment behavior. This history is composed of the individual's last period investment, as well as the total investment of the group(s) in the last period. These past actions of others influence a subject's beliefs regarding others' choices (Fischbacher and Gächter, 2010) and hence her own choices. We also produce a Panel Vector Autoregressive Model (VAR) (Abrigo and Love, 2016) with other exogenous co-variates for the multi-group treatments to incorporate both accounts separately.

Table 6 summarizes two regressions – Panel Tobit and Panel VAR models. We find that the subject's current investment is highly positively dependent on their past investment for both the total PG investment and by account (significant with p-value < 0.001 for all regressions). In the case of the Single Group, the current investment is also positively dependent on the others group members' investments. This implies that the choices of conditional cooperators can sway group investment patterns. We argue that a history of positive or increasing investment from others influences the subject's beliefs regarding the composition of player types in theri group, which influences the subject to be more cooperative themselves, if he is a conditional cooperator. In the case of multi-group treatments, the influence of others' investments varies. When we look at the multi-group treatments, for the total PG investment, others' investment in the last period increases a subject's total investment in the current period, but for Multi-Shared treatment, the effect is significant only at α-level of 0.1. However, we do not put much stalk in these results as we use aggregated investments from the two accounts, and the mechanisms are not on aggregate investment. The coefficients of concern are found in the Panel VAR model to evaluate how the decisions of others' affect one's decisions. Our theory suggests that others' investments in Blue(Green) account should positively impact a subject's investment in that account, and others' investments in the

	Single	Multi-Split			Multi-Shared		
	(1)	(2)	(3)		(4)	(5)	
	$c_{i,t}$	$c_{i,t}$	$c_{i,B,t}$	$c_{i,G,t}$	$c_{i,t}$	$c_{i,B,t}$	$c_{i,G,t}$
$c_{i,t-1}$	0.327***	0.418***	-	-	0.433***	-	-
	(0.033)	(0.034)			(0.037)		
$c_{i,B,t-1}$	-	-	0.272***	0.009	-	0.375***	-0.078
			(0.057)	(0.063)		(0.079)	(0.081)
$c_{i,G,t-1}$	-	-	0.080+	0.360***	-	0.005	0.250***
			(0.043)	(0.049)		(0.076)	(0.068)
$c_{-i,t-1}$	0.096***	0.095***	-	-	0.032+	-	-
	(0.017)	(0.022)			(0.017)		
$c_{-i,B,t-1}$	-	-	0.062+	0.039	-	0.074	-0.165*
			(0.035)	(0.038)		(0.070)	(0.066)
$c_{-i,G,t-1}$	-	-	0.059*	0.088**	-	-0.015	0.100**
			(0.023)	(0.026)		(0.049)	(0.038)
Period	-0.186***	-0.153***	-0.020	-0.033	-0.288***	-0.085	-0.169*
	(0.029)	(0.038)	(0.038)	(0.039)	(0.035)	(0.081)	(0.069)
Constant	5.225***	3.713***	-	-	9.014***	-	-
	(0.859)	(1.057)			(1.219)		
ρ	0.355	0.177	-	-	0.543	-	-
# Observations	1520	1520	1440		1520	1440	
	[157, 1196,167]	[221, 1159,140]			[177, 1010, 333]		

Table 6: Coefficient from Regressing Current Investment on the Past Investment by Self and Others

Notes: This table shows the coefficients from the Panel Tobit regression of total public goods investments on last period's investments in columns (1), (2), and (4) and from the VAR model (with X variables) in columns (3) and (5). The data for the Tobit Regression is censored at 0 at the lower end (left number in the array under "# Observations") and 20 (right number in the array under "# Observations") at the upper end. The robust standard errors are listed in parentheses below the coefficients.

Symbols: p-values $^{***} < 0.001$, $^{**} < 0.01$, $^{*} < 0.05$, $^{+} < 0.1$

Blue(Green) account should negatively impact a subject's investment in that account. We find evidence of this in the Green account of the Multi-Shared treatment (others' last period investment in Green account positively and that in Blue account negatively impacts current investment in Green account) but not the Multi-Split treatment where we find the others' investment own and cross-lag to be positive and significant in Blue account, and others' investment own-lag to be positive and significant in Green account.

However, this regression does not involve the entire mechanism behind our model. Here, we do not know which group/account the subject believes to be housing more cooperators. The theory stipulates that there is a comparison between the two accounts, which the current regression lacks. We, therefore, remove the arbitrary Blue-Green delimiters of groups and renames the groups as minimum and maximum groups depending on the investments, under the assumption that a subject invests more in the group she thinks has more cooperators. We employ the panel VAR regression on the investments in the Maximum and Minimum groups and others' investments in those corresponding accounts in the last period. To elaborate, we assume that looking at the investment in each account by others, a subject updates or retains their beliefs regarding the account that is more cooperative.

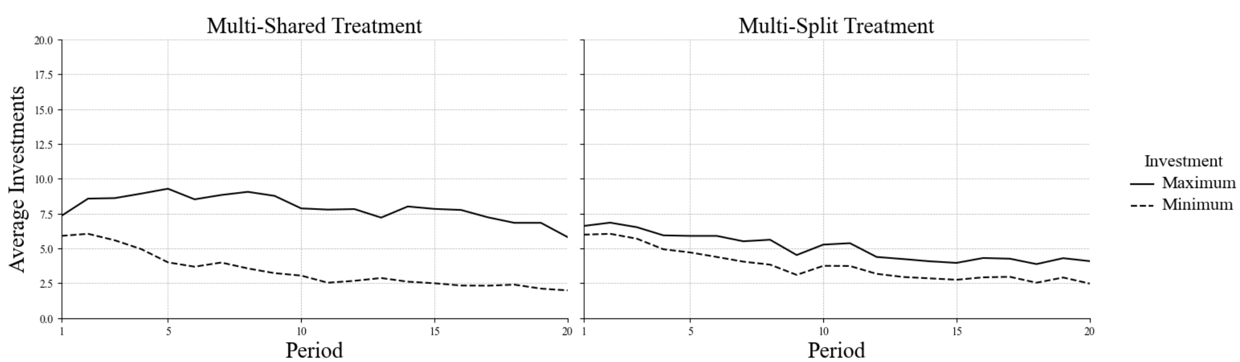


Figure 8: Average Maximum and Minimum Investments per Period

In Table 7, lags for both minimum and maximum investments are positive and highly significant with their auto- and cross-lags in the Multi-Split treatment. These positive coefficients reveal that the system, a session of networked public goods games, bolsters investment in both the maximum and minimum groups. Higher investments in the maximum group produce an interaction effect that raises investment in the minimum group. This heightened investment in the minimum group, in turn, raises subsequent investment for the maximum in a similar mechanism to that of a positive feedback loop. However, we do not find the expected effect of others' investment in the Multi-Shared treatment. Lags for both minimum and maximum investments are still positive and highly significant with their auto-lags in the Multi-Shared treatment. But the own cross-lags differ between the Maximum and Minimum investments. The others' investment cross-lag coefficient is only weakly significant in the case of the Maximum investment, but that, for Minimum, is positive and weakly significant.

	Multi-Split (3)		Multi-Shared (5)	
	$c_{i,Max,t}$	$c_{i,Min,t}$	$c_{i,Max,t}$	$c_{i,Min,t}$
$c_{i,Max,t-1}$	0.292***	0.037	0.248**	0.089+
	(0.071)	(0.070)	(0.072)	(0.046)
$c_{i,Min,t-1}$	0.068	0.331***	-0.181*	0.463***
	(0.058)	(0.061)	(0.091)	(0.059)
$c_{-i,Max,t-1}$	0.073*	0.033	-0.064	0.068+
	(0.028)	(0.028)	(0.051)	(0.033)
$c_{-i,Min,t-1}$	0.054*	0.096***	-0.089+	0.099**
	(0.032)	(0.025)	(0.051)	(0.031)
Period	-0.018	-0.028	-0.261**	-0.018
	(0.037)	(0.036)	(0.081)	(0.05)
# Observations	1440		1440	

Table 7: Coefficient from Regressing Current Maximum and Minimum Investment on the Past Investment by Self and Others

Notes: This table shows the coefficients from the VAR model (with X variables) corresponding to the columns (3) and (5) in Table 6. The data for the Tobit Regression is censored at 0 at the lower end (left number in the array under "# Observations") and 20 (right number in the array under "# Observations") at the upper end. The robust standard errors are listed in parentheses below the coefficients.
Symbols: p-values $^{***} < 0.001$, $^{**} < 0.01$, $^{*} < 0.05$, $^{+} < 0.1$

With this, we have thoroughly illustrated that the Multi-Shared treatment, when subjects are involved in two groups simultaneously with no overlap in group membership and receive one endowment, is significantly, statistically and economically, different from the Multi-Split treatment as well as the baseline PGG. In terms of Hypothesis 1, we discover our expected result for the difference between the Single Group versus the Multi-Shared, and the Multi-Split and the Multi-Shared, but fail to discover any significant difference between the Single Group and the Multi-Split. Moreover, we can not find evidence to support the notion of the importance of group composition.

6.2 Strategy Method

We now look at the data we collected from our Strategy part of the experiment. We follow the procedure introduced in Fischbacher, Gächter and Quercia (2012). Subjects are asked what their unconditional investment to the group account is, followed by 21 questions regarding what their investment to the group account would be if the average investment of the rest of their group is some integer between 0 and the size of the endowment. With this modified strategy method, we can assign types to the subjects in a single PGG. We use the definitions of the different types from Fischbacher, Gächter and Quercia (2012). Note that our Single Group treatment can be treated as the replication of the C-P treatment in FGQ. FGQ shows that the order of the direct response and strategy method does not matter statistically. They could not reject the null hypothesis the data from the C-P and P-C treatments come from different distributions. Therefore, our analysis is under the assumption that eliciting the conditional investment schedules (investment in PG as a function of average investment by the rest of the group) by strategy method after the main treatment (Single Group, Multi-Shared, Multi-Split) is administered does not affect the choices in the strategy method.

Table 8 shows the distribution of player types across our three treatments. Figure 16 provides the average conditional investments by a subject's type. *Conditional Cooperators*, CC, are classified by the standard criteria in which their investment schedules are either monotonic with at least one increase or have a positive and significant (at α-level of 1%) Spearman coefficient. *Free-riders*, FR, are rather strictly defined, necessitating that all of their conditional values are 0. *Hump-shaped*, HS, depict subjects who increase their investment to the group account to some satiation point and decrease their investment thereafter. This type is partitioned by a positive and significant (at α-level of 1%) Spearman coefficient up to some maximum conditional value and then a negative and significant (at α-level of 1%) Spearman coefficient from that maximum to the endowment. *Others* are those that are unclassifiable by these three aforementioned typings.

	CC	FR	HS	Other	Total
Multi-Shared	51	7	11	11	80
Multi-Split	47	4	12	17	80
Single-Group	47	3	14	16	80

Table 8: Distribution of Types

Table 8 reveals that in each treatment, over half of all subjects can be typed as conditional cooperators.[7] Given the skewed distribution of types in favor of CC, we can not test our hypotheses regarding the heterogeneity of player types driving the differences in subject behavior between our treatments. In figures 9 and 10, we show the average investment levels in each public good in each of our treatments (1 in Single Group, 2 in Multi-Split and Multi-Shared treatments) decomposed by subject's type. Figure 10 illustrates the lack of predictive power that these typings hold in the multi-group context.[8] This can possibly be attributed to the following reasons: (1) the experienced typing misclassifies individuals, (2) the lack of heterogeneity in player types in our sample, and (3) these player types from the Single Group strategy method are simply not accurate predictors of subject behavior in

[7]One drawback of the definition we use for CC is that we can not further categorize them except as perfect CC (perfectly matching others' average investment) or CC with self-serving bias (contribute less than others but still have an overall positively sloped schedule). Among the subjects who can be typed as CC with self-serving bias, the bias can be of differing degrees, which can impact the investment decision. A finer typing algorithm could help us tease out this heterogeneity, which could, in turn, help explain the behavior in our experiment.

[8]When performing similar regressions to those in (6.1), the player types, when compared against one another, maintain their differences. Free-riders invest less than any other player type, conditional cooperators the most, and hump-shaped somewhere in between.

trivial extensions of the single linear-VCM to multiple linear-VCMs.

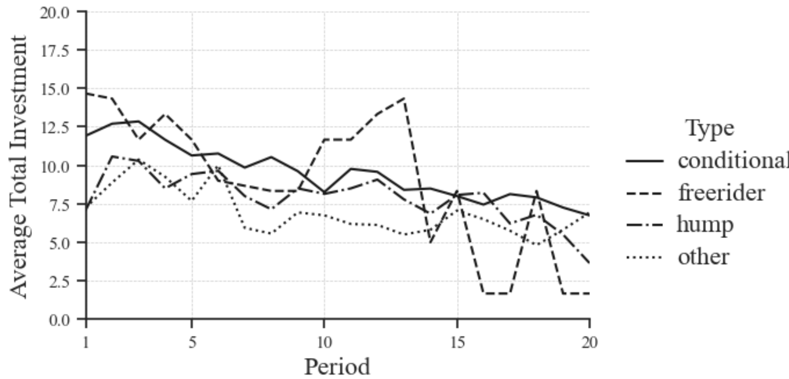


Figure 9: Average Investment of Different Types in Single Group Treatment by Periods

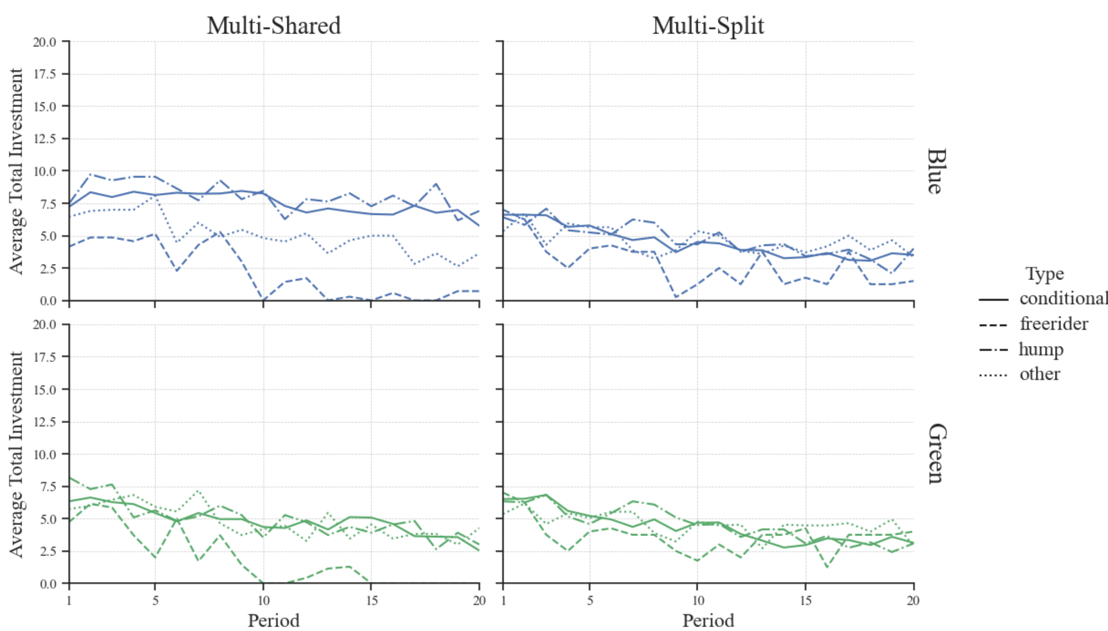


Figure 10: Average Investment of Different Types in each Account (Multiple Groups) by Periods

In terms of Hypothesis 2, we are not adequately able to accept or reject its claim regarding which type of agents are driving the treatment difference as the typings, as collected via the strategy method, hold little merit in the multi-group extension of that same game.

7 Extension

The second panel of Figure 6 plots the average investment across all sessions of a given multi-group treatment, Multi-Shared on the left and Multi-Split on the right, by periods from the original experimental session. One would expect the averages of each group, Blue and Green, should be the same as one another as, by the experimental design, both groups are defined by identical parameter sets (group size, scale factor, decision, etc.). This expectation is realized with the Multi-Split treatment, but visually, it is not the case for the Multi-Shared treatment. This difference between the Blue and Green group, on average, is puzzling as it does not appear for *both* multi-group treatments, rather only the Multi-Shared. The decision screen, showcased in Appendix A, is nearly identical to the decision for the Blue group on the left and the Green group on the right. The only difference between the screens is that the Multi-Split notes that the subject receives *two* endowments in each period rather than one.

Moreover, despite the homogeneity of our subjects, in regard to typing, in our primary experiment, we uncovered a statistically significant difference between the Multi-Shared, Multi-Split, and single-group sessions. Mainly, the Multi-Shared elicits increased investment to group accounts compared to both its Multi-Split counterpart and the baseline PGG. However, it remains an open question if this feat will be exacerbated when the sample is more heterogeneous. Our simulation attests this to be the case. Thus, with our primary experiment results, we present this extension to investigate these issues.

7.1 Experimental Design – Bots

Rather than relying on a random sample of subjects to create heterogeneity, we opt to impose heterogeneity in each session by involving automated players, bots, in addition to recruiting subjects. These bots are placed at various nodes of our lattice as shown in Figure 11. The numbers represent human subjects, B_F represents a bot who will behave as a free-rider, and

B_C is another type of bot who will act as a cooperator. The rows form a subject's Blue group, whereas the columns are a subject's Green group. This structure ensures that each subject has a more (less) cooperative group. We anticipate subjects to shift their investment toward (away from) their more (less) cooperative group. For instance, player 1's Blue group is composed of himself, player 2, and two free-rider bots, whereas his Green group is composed of himself, player 3, and two cooperator bots. Since we do not ex-ante know the typing of player 1's group mates, and assuming player 1 maintains his typing across groups, we would anticipate his Green group to be more cooperative than his Blue group. With this in mind, player 1, if some semblance of a conditional cooperator himself, should, over time, shift his investments from the Blue group to the Green group. For subjects 1, 2, 3, and 4, their Blue group is "less cooperative" than their Green group: with the order reversed for subjects 5, 6, 7, and 8.

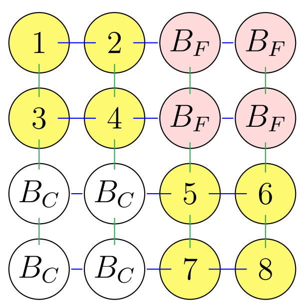


Figure 11: Bot Session Multi-Shared Grouping

The free-rider bots invest zero and have a small probability of deviating from this strategy and increasing its investment to the group account by drawing from a Poisson distribution. Cooperator bots prefer to invest their whole endowment and have a preference for fairness, 10,10, with probability deviations of reducing their investment via draws from an exponential distribution.

With the addition of bots, we rephrase our instructions by informing subjects that there are bots in both of their groups and these bots follow programmed algorithms to determine their investments in each period, which matches the standard of previous automated player experiments March (2021). We also incorporate additional survey questions to elicit subjects' opinions regarding the bots that appeared in their groups; these questions included considerations of how many bots were in each group, if the bots invested more or less than human players, etc.

Our extension is primarily concerned with the multi-group setting and, more specifically, the Multi-Shared treatment. To isolate this, the extension sessions contain two treatments within the Multi-Shared treatment: the screen order of Blue-Green and Green-Blue. Instructions for both bot extensions remain the same, with the only alteration being the description of where the two groups appear on the screen. Our motivation behind introducing the two orders is to control for the bizarre deviation between average Blue and Green investment found in our primary experiment for the Multi-Shared treatment. All additional aspects of the Multi-Shared bots extension are inherited from the Multi-Shared treatment as described in Section 3.1.1.

7.1.1 Logistics – Bots

The experiment was implemented via oTree Chen, Schonger and Wickens (2016). We ran 4 sessions of each Multi-Shared screen ordering. Each session included 8 human participants and 8 bot participants. All extension sessions were conducted in the fall of 2023 at Chapman University from ESI's subject pool with recruiting criteria of no prior experience in the public goods environment, single- or multi-group, and no subject participating in more than one session. Subjects received a $7 show-up fee and their cumulative earnings from the 20 periods. Total payments, including the show-up fee, averaged $18.89.

7.2 Results – Bots

A preliminary task of this extension was to check if the screen ordering affect that manifested for the priamry experiments Multi-Shared experiment would carry over to its extension counterpart. Figure 12 reveals the average investment paths of the 4 sessions on the left with the Blue-Green ordering and the 4 sessions on the right with the Green-Blue ordering. In the extension, the screen ordering effect vanishes. This figure showcases that subjects, initially, have no mechanism to discriminate their two groups; however, given enough periods, they shift their investments to their group with more cooperator bots and decrease their investment from the group with more free-rider bots.

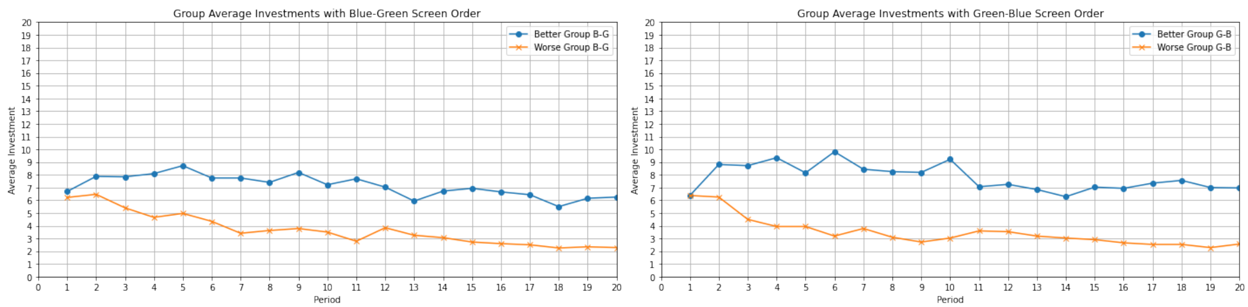


Figure 12: Group Average Investment for Multi-Shared Bots with Varying Screen Orders

In the absence of this effect in our extension data we are confident in combining the session data as that screen ordering was the sole difference in treatments and produce the averaged plot below in Figure 13. This figure also reports the investments of the comparable Multi-Shared treatment from our primary study. Also to note, our extension data follows more closely with the stylized fact of investment decay in finitely repeated social dilemmas. In addition to the alignment with conventional public goods results, our extension data has a larger fanning effect than in our primary experiment.

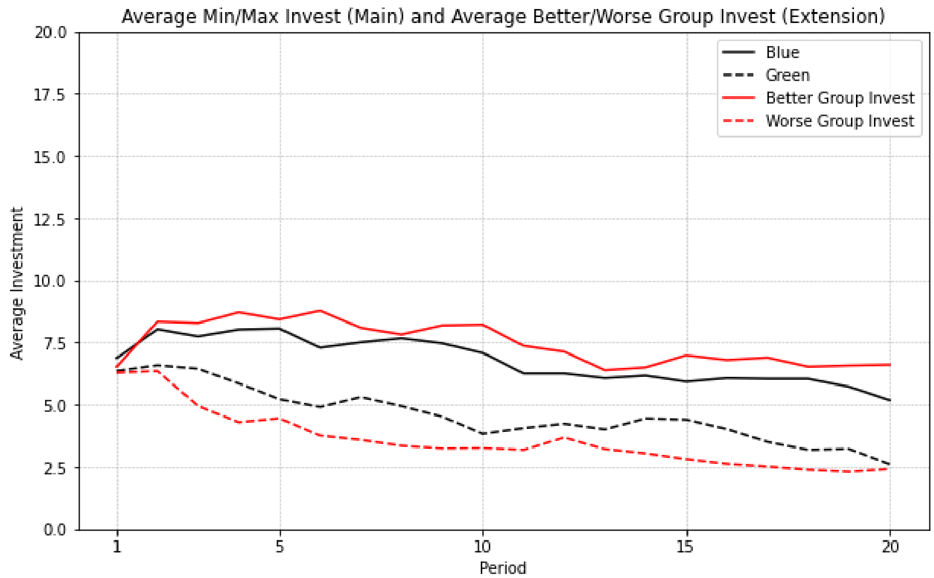


Figure 13: Average Investment for Multi-Shared with and without Bots

There remains an important distinction to be made: are subjects better at allocating the investment that they would make to their 'better' group, and/or does this added heterogeneity invoke subjects to contribute more than they otherwise would have? We produce Figure 14, which plots the average number of points players leave in the personal account in each period for the Multi-Shared primary and extension. As is readily seen, the amount subjects leave in their personal accounts, on average, between the two handlings of Multi-Shared, helix about one another, and are rather similar. This attests that subjects are better at allocating the investments that they would make, already bolstered by the multi-group setting, not necessarily contributing more just because they have a group with exceptionally favorable returns.

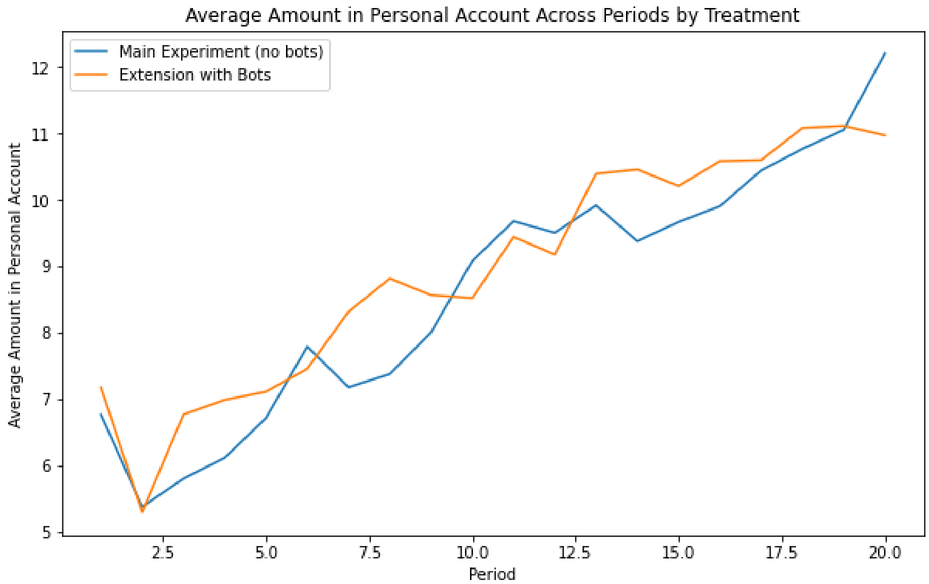


Figure 14: Average Personal Accounts Multi-Shared and Multi-Shared with Bots

A final consideration for this extension is thinking of the evolution of social groups and player types within these groups. If looked at as an evolutionary game, would free-riders be able to survive, or out perform cooperators in this multi-group setting, given that most human players behave as conditional cooperators – indicated by our own results and (McCarter, Samek and Sheremeta, 2014). Figure 15 showcases the average payoffs for the free-rider and cooperator bots, the human players in the primary and extension experiments. The figure indicates that the cooperator bots perform the best, where human players fall in the middle, with free-riders being outperformed. Intuitively, this would attest to the fact that free-riders, in the multi-group setting, when the population is sufficiently mixed, can be 'left behind' and isolated from the rewards of other public goods in the system, even in a system with no inherent mechanism of excommunication, punishment, nor sorting.

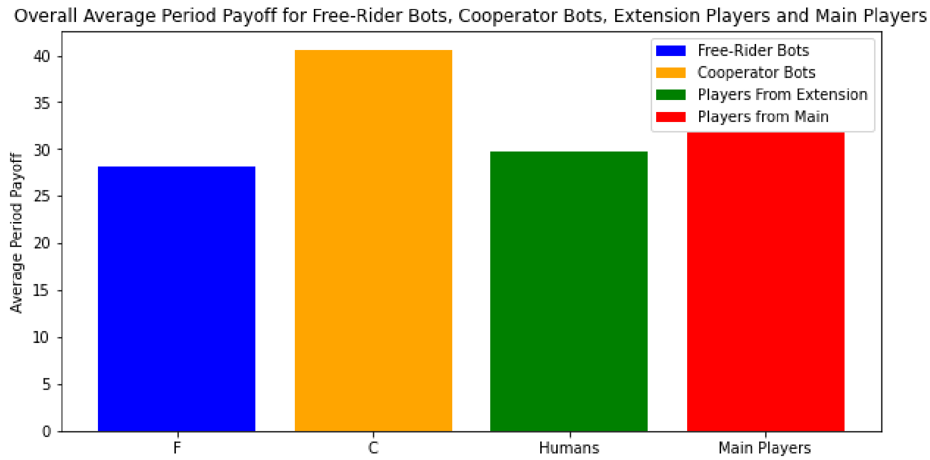


Figure 15: Average Payoffs for Bot Types and Players across Multi-Group Experiments

8 Conclusion

People are involved in a plethora of social groups in their lives; these groups vary in what they successfully provide to the lives of their members, which is predominantly a function of group composition. The multi-group aspect of public and club goods has been untapped by the overarching social dilemmas literature. Although ever present in laboratory studies, free-riding behavior cannot be ubiquitous in the real world as many social groups persist, grow, and are cornerstones of human life.

Our primary study evaluates the gains from analyzing this multi-group extension in the controlled lab setting. We find a significant treatment difference between the average investment of the Multi-Shared treatment against both the Multi-Split treatment and single-group baseline. From our standard regression techniques and conventional player typings, which do not hold explanatory power in the multi-group environment, we are unable to pinpoint an exact motivation as to why the Multi-Shared treatment bolsters cooperation as much as it does. With this discovery and lack of explanatory power in tow, we continue our exploration with an extension study that accounts for screen ordering, a homogeneous sample, and weak notions of multi-group player types in an effort to uncover a *why*.

The extension reduces the complexity of the environment and increases the heterogeneity between a subject's group. From our sessions of this extension, we find that the amount that subjects invest in any period is similar regardless of the presence of bots, but human subjects better allocate their investments when bots are present. Subjects are able to recognize their better group and invest more there whilst leaving behind their worse group by investing less. This would hint that groups with rampant free-riding would not persist in an evolutionary game with free-riders being isolated without the need for any institutional mechanism to sanction this behavior.

We conclude that multi-group considerations are important to incorporate and consider in the lab. Although we gain control from the inspection of a Single Group interaction over a period, we lose the notion of the economic trade-offs that are inherent in every decision we make. Especially that is most salient being where we invest our time, energy, and effort. The monolithic notion of free-riding in the public goods and social dilemmas literature at large may be a by-product of the design of the standard linear-VCM and other related frameworks, namely heralded by their single-group design.

References

Abrigo, M. and I. Love. 2016. "Estimation of panel vector autoregression in Stata." The Stata Journal 16(3):778–804.

Bednar, J., Y. Chen, T. Xiao Liu and S. Page. 2012. "Behavioral spillovers and cognitive load in multiple games: An experimental study." Games and Economic Behavior 74:12–31.

Bühren, Christoph, Jan Haarde, Christian Hirschmann and Janis Kesten-Kühne. 2023. "Social preferences in the public goods game–An Agent-Based simulation with EconSim." Plos one 18(3):e0282112.

Camerer, Colin and Teck Hua Ho. 1999. "Experience-weighted attraction learning in normal form games." Econometrica 67(4):827–874.

Cason, T., A. Savikhin and R. Sheremeta. 2012. "Behavioral spillovers in coordination games." European Economic Review 56:233–245.

Chen, D.L., M. Schonger and C. Wickens. 2016. "An open-source platform for laboratory, online and field experiments." Journal of Behavioral and Experimental Finance 9:88–97.

Conover, W.J. and R. Iman. 1979. "On-Multiple Comparisons Procedures." Technical Report LA-7677-MS, Los Alamos Scientific Laboratory.

Falk, A., U. Fischbacher and S. Gächter. 2013. "Living in Two Neighborhoods – Social Interaction Effects in the Laboratory." Economic Inquiry 51:563–578.

Fehr, E. and S. Gächter. 2000. "Cooperation and Punishment in Public Goods Experiments." The American Economic Review 90(4):980–994.

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

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

Fischbacher, Urs and Simon Gächter. 2010. "Social preferences, beliefs, and the dynamics of free riding in public goods experiments." American economic review 100(1):541–556.

Hajikhameneh, A. and Iannaccone L. 2023. "God games: An experimental study of uncertainty, superstition, and cooperation." Games and Economic Behavior 139:88–116.

Kreig, J. and A. Samek. 2016. "When charities compete: A laboratory experiment with simultaneous public Goods." Journal of Behavioral and Experimental Economics .

Laury, S. and C Holt. 2002. "Risk aversion and incentive effects." American Economic Review 92(5):1644–1655.

Ledyard, J.O. 1995. Public Goods: A Survey of Experimental Research In: The Handbook of Experimental Economics. 1 ed. Princeton: Princeton University Press.

March, C. 2021. "Strategic interactions between humans and artificial intelligence: Lessons from experiments with computer players." Journal of Economic Psychology 87.

McCarter, M.W., A. Samek and R.M. Sheremeta. 2014. "Divided Loyalists or Conditional Cooperators? Creating Consensus About Cooperation in Multiple Simultaneous Social Dilemmas." Group & Organization Management 39(6):744–771.

Savikhin, A. and R. Sheremeta. 2013. "Simultaneous Decision-Making in Competitive and Cooperative Environments." Economic Inquiry 51(2):1311–1323.

Tukey, John W. 1949. "Comparing Individual Means in the Analysis of Variance." Biometrics 5(2):99–114.

Valdez, P. and A. Mehrabian. 1994. "Effects of Color on Emotions." Journal of Experimental Psychology 123(4):394–409.

Zelmer, J. 2003. "Linear Public Goods Experiment: A Meta-Analysis." Experimental Economics 6:299–310.

Appendix A. Experiment, Data, & Data Analysis

Our full experiment and subsequent data analysis can be found on the following two GitHub repositories.

Experiment Implementation (oTree): Multi-Group Experiment on GitHub

Data and Data Analysis: Data Analysis for Multi-Group on GitHub

Additional access is available on this public Heroku web-page found at the following link. Demos of all experiment content, instructions, comprehension questions, decision screens, and surveys can be found on this website.

Heroku Web-Page with Experiment Content

8.1 Results

8.1.1 Strategy Method

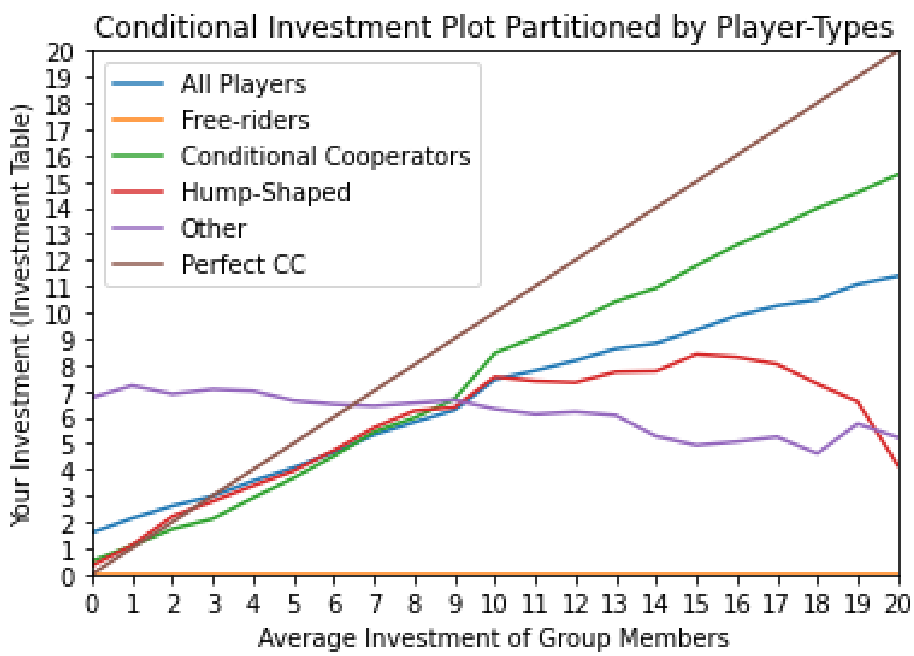


Figure 16: Average Conditional Investments for Each Typing

Appendix B. Power Analysis

We conduct an initial power analysis with data from our Multi-Shared pilot session and metrics reported in Falk, Fischbacher and Gächter (2013) that resembles our Multi-Split treatment.

The general metrics they provide are rather limited, so we utilized the total average investment that they report for the first 5 rounds of play, out of 20 rounds, which is 11.3 tokens out of 40 tokens they receive each period – 20 for their group one and 20 for their group two. Our pilot data differs in two significant ways, in addition to the treatment, (1) all subjects received an endowment of 20 points that can be invested in their Blue and Green group accounts and (2) our pilot ran for 10 periods. To correct for (1), we convert the averages in both datasets to the average percent of the endowment subjects give in the first 5 periods. For Falk et al. $\frac{11.3}{40} = 0.2825$. (2) cannot be adequately addressed by normalizing the data in anyway, so this is kept in mind as a setback of the subsequent power analysis.

From our pilot data, our subjects' average investment to their group accounts in the first 5 periods is 7.05625 points which is 0.3528% their endowment.

In lieu of access to the complete dataset of Falk, Fischbacher and Gächter (2013), we impose equal standard deviations between their and our samples. With this in mind, we calculate our standard deviation to be 3.086 points or 0.154% of the endowment.

We opt for standard choices for the remainder of the parameters, $\alpha = 0.05$ and power $= 0.80$. With the respective averages from both studies, and our standard deviation, we find the effect size to be 0.456. Via the 'TTestIndPower' function from the 'statsmodels' package in Python, we estimate the required sample size for both the Multi-Shared and Multi-Split treatment to be 77 subjects. With our torus grouping procedure and group size 4 parameter, our minimum session size is 16. With this we will round 77 up to our nearest session size of 80 subjects. Our power analysis informs us to run 5 sessions of each multi-group treatment.

All code for this power analysis was conducted in Python and is available on GitHub via the following link or upon request for a .zip file: Power Analysis for Multi-Group PGG on GitHub.

Appendix C. Simulation Pseudo-Code

Pseudo-code

S0: Initiate simulation parameters – $nsim$

S1: For each simulation **do**

S10: Initiate Public Good Game (PGG) Environment → Initiate Lattice of the PGG → Initiate Agents/Players on the Lattice

S11: Run One Simulation – One simulation consists of T periods → For period $t, \forall t \in [1, T]$ **do**

S111: $c_t \leftarrow$ Randomly (given the probabilities P_t) pick investment given A_{t-1}

$$A^*_{t-1} = \max_j A^j_{t-1}$$

$$P^j_t = \frac{e^{\frac{A^j_{t-1}}{\lambda_{t-1} A^*_{t-1}}}}{\sum_k e^{\frac{A^k_{t-1}}{\lambda_{t-1} A^*_{t-1}}}}$$

S112: $\hat{N}_t \leftarrow$ Update $\hat{N}$ given $\hat{N}_{t-1}$

$$\hat{N}_t = \rho \hat{N}_{t-1} + 1$$

S113: $A_{i,t} \leftarrow$ Update A given $c_t, A_{i,t-1}$

$$A^j_{i,t} = \frac{\phi \hat{N}_{t-1} A^j_{i,t-1} + (\delta + (1-\delta) I(c^j_i, c_{i,t}))\pi(c^j_i, c_{-i,t})}{\hat{N}_t}$$

S114: $\lambda_{i,t} \leftarrow$ Update λ given $\lambda_{i,t-1}, A_{i,t-1}, A_{i,t}$

S1141: $A^*_t \leftarrow \max_j A^j_t$

S1142:

$$c^*_t = \begin{cases} \text{argmax}_j A^j_t & \text{if unique maximum} \\ null & \text{otherwise} \end{cases}$$

S1143:

$$\lambda_t = \begin{cases} \max\left(\underline{\lambda}, \lambda_{t-1}(1 - \lambda_\Delta)\right), & c^*_t = c^*_{t-1} \text{ \& } c^*_t, c^*_{t-1} \text{are not null} \\ \min\left(\lambda_{t-1}(1 + \lambda_\Delta), \bar{\lambda}\right), & c^*_t \neq c^*_{t-1} \text{ \& } c^*_t, c^*_{t-1} \text{are not null} \\ \lambda_{t-1}, & \text{otherwise} \end{cases}$$

S12: Store Variables from Simulation

S2: Generate Plots