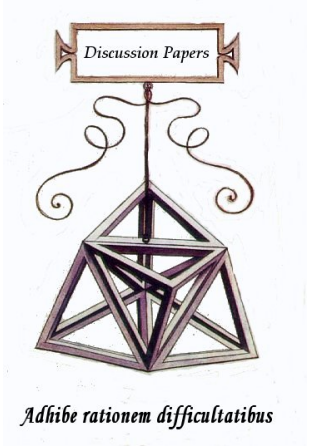
Discussion papers
E-papers of the Department of Economics and Management – University di Pisa



Andrej Angelovski, Irene Maria Buso,
Daniela Di Cagno, Werner Gueth,
Lorenzo Spadoni

How Behavioral Heterogeneity Shapes Institution Formation and Cooperation: The Case of Cost-Sharing Alliances

Discussion paper n. 333

2026

Discussion paper n. 333, presented: **November 2026**

Authors' address/Indirizzo degli autori:

Andrej Angelovski — Xi'an Jiaotong–Liverpool University (XJTLU), Suzhou, China. E-mail: Andrej.Angelovski@xjtlu.edu.cn

Irene Maria Buso — Department of Economics, University of Bologna, Piazza Scaravilli 2, 40126 Bologna, Italy. E-mail: irenemaria.buso@unibo.it

Daniela Di Cagno — Department of Economics and Finance, Luiss University, Rome, Italy. E-mail: ddicagno@luiss.it

Werner Gueth — Max Planck Institute for Behavioral Economics, Bonn, Germany. E-mail: gueth@coll.mpg.de

Lorenzo Spadoni — Department of Management and Economics, Pegaso University. E-mail: lorenzo.spadoni@unipegaso.it

© Andrej Angelovski, Irene Maria Buso, Daniela Di Cagno, Werner Gueth and Lorenzo Spadoni

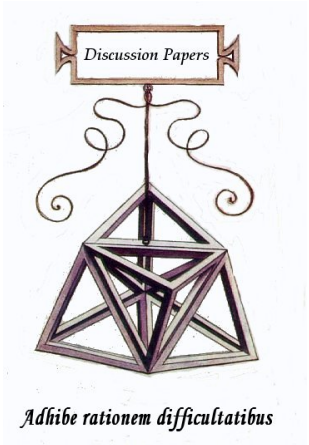
Please cite as:/Si prega di citare come:

Andrej Angelovski, Irene Maria Buso, Daniela Di Cagno, Werner Gueth, Lorenzo Spadoni (2026), "How Behavioral Heterogeneity Shapes Institution Formation and Cooperation: The Case of Cost-Sharing Alliances", Discussion Papers, Department of Economics and Management – University of Pisa, n. 333 (http://www.ec.unipi.it/ricerca/discussion-papers).

Discussion Papers Series contact: pietro.battiston@unipi.it

Discussion paper

n. 333



Andrej Angelovski, Irene Maria Buso, Daniela Di Cagno, Werner Gueth, Lorenzo Spadoni

How Behavioral Heterogeneity Shapes Institution Formation and Cooperation: The Case of Cost-Sharing Alliances

Abstract

While endogenous institution formation and behavioral heterogeneity in public-good provision have received substantial attention, they have largely been studied separately. Yet, differences in cooperative attitudes may shape not only contributions but also individuals' willingness to initiate or join cooperative institutions. We study behavioral heterogeneity in a two-stage public-good game, where participants sequentially decide whether to join a cost-sharing alliance whose members share contribution costs equally while outsiders remain unconstrained. A strategy-vector method elicits conditional joining decisions at all possible positions in the sequence and contributions for all compatible alliance outcomes, allowing us to identify individual profiles across both margins and explore how their interplay enables such alliances to sustain and stabilize cooperation. Joining behavior is heterogeneous but structured: roughly a quarter of choices reflect unconditional joining, about 40

Keywords: Cost-sharing alliance; Behavioural heterogeneity; Strategy method; Conditional cooperation; Public goods game

JEL CLassification: C92, H41, D71, C72

How Behavioral Heterogeneity Shapes Institution Formation and Cooperation: The Case of Cost-Sharing Alliances

Andrej Angelovski * Irene Maria Buso† Daniela Di Cagno‡
Werner Güth§ Lorenzo Spadoni ¶

November 11, 2026

Abstract

While endogenous institution formation and behavioral heterogeneity in public-good provision have received substantial attention, they have largely been studied separately. Yet, differences in cooperative attitudes may shape not only contributions but also individuals' willingness to initiate or join cooperative institutions. We study behavioral heterogeneity in a two-stage public-good game, where participants sequentially decide whether to join a cost-sharing alliance whose members share contribution costs equally while outsiders remain unconstrained. A strategy-vector method elicits conditional joining decisions at all possible positions in the sequence and contributions for all compatible alliance outcomes, allowing us to identify individual profiles across both margins and explore how their interplay enables such alliances to sustain and stabilize cooperation. Joining behavior is heterogeneous but structured: roughly a quarter of choices reflect unconditional joining, about 40% conditional joining, and only a small minority never join, with unconditional joining becoming more prevalent over time. Joining profiles predict contributions even under identical incentives, ruling out self-selection as the sole explanation. Overall, sustained cooperation results from the combination of behavioral heterogeneity and institutional incentives rather than convergence to the benchmark strategy.

JEL Codes: C92, H41, D71, C72

Keywords: Cost-sharing alliance; Behavioural heterogeneity; Strategy method; Conditional cooperation; Public goods game

*Xi'an Jiaotong–Liverpool University (XJTLU), Suzhou, China, e-mail: Andrej.Angelovski@xjtlu.edu.cn.

†Department of Economics, University of Bologna, Piazza Scaravilli 2, 40126 Bologna, 40126, e-mail: irene-maria.buso@unibo.it

‡Department of Economics and Finance, Luiss University, Rome, Italy, e-mail: ddicagno@luiss.it.

§Max Planck Institute for Behavioral Economics, Bonn, Germany, e-mail: gueth@coll.mpg.de.

¶Department of Management and Economics, Pegaso University, e-mail: lorenzo.spadoni@unipegaso.it.

1 Introduction

Solid evidence has accumulated that individuals differ systematically in their cooperative attitudes and that the resulting heterogeneity in behavior can play an important role in determining the success of cooperative attempts (e.g., Fischbacher et al., 2001; Gächter and Thöni, 2005; Chaudhuri et al., 2017; Thöni and Volk, 2018; Engel and Rockenbach, 2024). Furthermore, cooperation can be supported institutionally through alliances that emerge to address the free-riding problem. Several studies have examined, both theoretically and experimentally, the emergence and performance of such alliances, as well as how these depend on the rules governing participation and cooperative effort, particularly in the context of international environmental agreements (e.g., Carraro and Siniscalco, 1993a; Barrett, 1994a; Kosfeld et al., 2009; Gürerk, 2013; Dannenberg et al., 2014; Dannenberg and Gallier, 2020; Casari and Tavoni, 2024; Buchholz and Sandler, 2021). Yet, these two perspectives on the determinants of cooperation have largely been studied separately.

This leaves open an important question: whether the behavioral heterogeneity documented in contribution decisions also extends to the process through which cooperative institutions emerge endogenously. Differences in cooperative attitudes may shape both contribution and joining decisions, such that individuals may differ systematically not only in how much they contribute to a public good, but also in their willingness to initiate or join institutions designed to sustain cooperation. Moreover, heterogeneity in institutional participation may be systematically related to subsequent contribution behavior. Aggregate increases in cooperation following alliance formation may therefore mask substantially different individual-level mechanisms: cooperation may reflect a relatively uniform response to the incentives created by the institution, the self-selection of more cooperative individuals into the alliance, or systematic differences in how individuals condition both participation and contribution on others' willingness to join. Understanding why and when endogenous alliance formation sustains cooperation therefore requires looking beyond aggregate formation and contribution outcomes to individual behavior at both the formation and contribution stages.

Theoretical work on coalition formation suggests that social preferences as inequity aversion, altruism and reciprocity can affect individuals' willingness to join cooperative coalitions and, in turn, influence the outcomes of the coalition-formation process, including coalition size and public-good provision (e.g., Lange and Vogt, 2003; Vogt, 2016; van der Pol et al., 2012; Buchholz et al., 2018). Experimental evidence is also suggestive that social preferences matter in coalition formation: Kosfeld et al. (2009) find that individuals' willingness to form cooperative institutions is substantially reduced when these institutions leave outsiders with the opportunity to free ride, a pattern consistent with inequity aversion influencing decisions at the joining stage. Dannenberg et al. (2010) find evidence of inequity aversion among climate negotiators. Evidence on the

role of social preferences in alliance formation extends beyond cooperative settings: in a multilateral bargaining environment, Okada and Riedl (2005) show that reciprocity systematically affects coalition-formation decisions and, through them, bargaining outcomes. This literature further indicates that heterogeneity in social preferences may matter for coalition formation and its effects on cooperation. Ball et al. (2025) provide experimental evidence that, when social preferences are heterogeneous and privately known, willingness to join a coalition can signal underlying differences in these preferences, thereby facilitating cooperation in subsequent contribution decisions. Relatedly, the importance of heterogeneous group composition in terms of cooperative attitudes for cooperation outcomes has also been shown in the literature on sorting (e.g., Guido et al., 2019).[1] Importantly, social preferences need not affect both the joining and contribution margins or affect them in the same cooperative direction: depending on their nature and distribution within and outside the coalition, they may affect only one margin and not the other, or foster participation while reducing individual contributions (e.g., Lange and Vogt, 2003; Schopf, 2024).

Taken together, these studies suggest that it is important to deepen our understanding of how heterogeneity in cooperative attitudes influences not only the contribution stage but also the joining stage. To address this question, we study endogenous alliance formation under equal cost sharing. Cost-sharing alliances resemble real-world cooperative arrangements such as international climate agreements, large scientific collaborations, space programmes, and peace alliances.[2] Moreover, the equal cost-sharing rule provides a simple and transparent environment, limiting the possibility that behavioral differences reflect misunderstandings of the allocation mechanism. These features make cost-sharing alliances a useful setting for studying the behavioral heterogeneity underlying the formation of cooperative institutions and its relationship with subsequent cooperation. Importantly, in our setting, both participation and contribution decisions are endogenous: alliance members are not required to contribute at a prescribed level, nor are outsiders prevented from contributing, allowing heterogeneity along both margins to emerge without institutional constraints. Hence, while previous studies of coalition formation have typically started from prespecified heterogeneous social preferences to examine their effects on individuals' willingness to join a coalition, we focus on heterogeneity in joining decisions themselves.

Building on evidence that cost-sharing alliances can emerge endogenously and sustain cooperation (Buso et al., 2025), we employ a strategy-vector approach that links conditional joining

[1] In the traditional sorting literature, information or signals about individuals' cooperativeness allow cooperative and less cooperative individuals to sort into different interaction groups. In our setting, instead, the joining decision itself may signal cooperativeness, as in Ball et al. (2025), and alliance members and outsiders remain in the same public-good interaction group.

[2] They combine benefits that extend to the group as a whole with cost-sharing rules that apply only to alliance members, thus belonging to the class of institutions characterized by global benefits but participation-specific rules discussed by Dannenberg and Gallier (2020).

and contribution decisions. Participants first decide whether to join the alliance sequentially, with joining decisions elicited for every relevant position in the sequence, conditional on the number of previous group members who have joined. Before learning the realized alliance outcome, each participant then states their intended contribution for every alliance outcome compatible with their joining profile: no alliance, a partial alliance as a member, a partial alliance as an outsider, and the grand alliance. This conditional elicitation across both stages yields rich individual-level data on each participant's joining and contribution behavior, while reducing the extent to which contribution choices are anchored on a realized role. It thereby enables a classification of participant types based on the joint structure of their behavior along the two margins. We also vary the MPCR between $\alpha = 0.6$ and $\alpha = 0.8$ while keeping the benchmark solution unchanged across conditions, allowing us to examine whether the heterogeneous joining profiles and their associated contribution patterns persist as the return to cooperation changes.

We find that voluntary equal cost-sharing alliances form frequently, while contributions remain stable across rounds. Individual behavior, however, departs substantially from the benchmark. We find that voluntary equal cost-sharing alliances form frequently, while contributions remain stable across rounds, although individual behavior departs substantially from the benchmark. The theoretically predicted joining profile is rare, while a small set of unconditional and conditional joining profiles accounts for most observed choices and recurs across participants and rounds. At the same time, the prevalence of unconditional joining increases with experience, indicating that joining behavior is structured but not entirely fixed over time. Importantly, joining profiles predict contribution behavior: contributions are highest under Unconditional Join, lowest under Never Join, and intermediate under Conditional Join. These differences persist even when participants face the same no-alliance incentives, indicating that self-selection into alliance membership alone cannot account for the relationship between joining and contribution behavior, and the main patterns are broadly robust across MPCR conditions. Overall, the evidence suggests that joining profiles capture systematic differences in cooperative behavior that persist across incentive conditions but can evolve with experience, while contribution behavior reflects both these individual differences and the realized alliance context. The success of cost-sharing alliances in sustaining cooperation therefore appears to rest on the interplay between behavioral heterogeneity and the institutional incentives generated by alliance formation.

2 Experimental Design

2.1 Game Form

Our experiment implements a two-stage game. In the first stage, subjects choose whether or not to join an alliance that determines how members' contribution costs are shared in the second stage. The optimal decision of whether to join the alliance should therefore be based on anticipating the different contribution incentives implied by alliance membership by backward-induction.

Consider a group of n contributors endowed with e. Each contributor independently chooses an integer contribution ($c_i \in \{c_{min}, \ldots, c_{max}\}$) to a public good. Let ($C = \sum_{k=1}^{n} c_k$) denote the total group contribution and $\alpha \in (0, 1)$ the marginal per capita return (MPCR; Isaac et al., 1984), where $n\alpha > 1$. Thus, full contribution is socially efficient, even though $\alpha < 1$ makes free-riding individually rational.

An alliance is formed whenever at least $m \geq 2$ participants choose to join it. When $m < n$, we denote alliance members by i and outsiders by j. Alliance members equally share the total contribution cost incurred by the alliance: if the m members choose contributions $c_1, \ldots, c_m$, each member bears a cost of $C(m)/m$, where $C(m) = \sum_{i=1}^{m} c_i$. Consequently, all alliance members receive the same payoff:

$$\pi_i = e - \frac{C(m)}{m} + \alpha C. \tag{1}$$

An outsider j bears the full cost of their own contribution and earns:

$$\pi_j = e - c_j + \alpha C. \tag{2}$$

When no alliance is formed ($m = \emptyset$), all participants are effectively outsiders and equation (2) applies to everyone.

Assuming common and anticipated opportunism, outsiders free-ride (since $\alpha < 1$), i.e.,

$$c_j^*(m) = c_{min}.$$

Instead, for an alliance member, the optimal contribution depends on m: as the marginal cost of contributing one additional unit is $1/m$ (due to cost-sharing)[3], contributing maximally is optimal whenever $m\alpha > 1$:

$$c_i^*(m) = \begin{cases} c_{max} & \text{if } m\alpha > 1, \\ c_{min} & \text{if } m\alpha < 1. \end{cases}$$

[3]This is the externality that makes splitting a restaurant bill inefficient (Gneezy et al., 2004): each member bears only $1/m$ of what they order, so they order too much. Here the good being over-ordered is public rather than private, and since $n\alpha > 1$, the same externality raises welfare instead of destroying it. In this respect, the cost-sharing rule also operates like mechanisms that sustain contributions by altering effective marginal incentives, such as the tax–subsidy scheme of Falkinger (1996), which raised contributions to near-efficient levels experimentally (Falkinger et al., 2000). The difference is that here the change in marginal incentives is self-selected through the joining decision rather than imposed on the whole group.

In the first stage, the alliance is formed sequentially, with each participant observing how many preceding players have already joined. A participant will opportunistically join the alliance only when their decision is pivotal, i.e., when, anticipating opportunistic incentives in the contribution stage, their joining increases the alliance size from a level such that $m^*\alpha < 1$ to one such that $(m^* + 1)\alpha > 1$. Hence, alliances of size m^* are both externally stable (no outsider wishes to opt in) and internally stable (no member gains from unilaterally opting out) (Selten and Güth, 1982; d'Aspremont et al., 1983).[4] It follows that the first $(n - m^*)$ players in the sequence will never join and will instead enjoy the outsider payoff, anticipating that the remaining m^* players will form the alliance and contribute maximally.[5]

2.2 Experimental Procedure and Data Collection

In the experiment, we introduce groups of $n = 3$ contributors. In the first stage, alliance formation proceeds sequentially, but all decisions are elicited simultaneously using the strategy method. A random ordering of the $n = 3$ group members is drawn by the computer to determine who decides first, second, and third. Before this ordering is drawn, each participant makes a set of conditional binary choices, $\delta \in \{0, 1\}$, one for each possible combination of sequence position and number of prior joiners they could face, as illustrated in Figure 1.

[4] A similar stability logic underlies the theory of self-enforcing international environmental agreements (Carraro and Siniscalco, 1993b; Barrett, 1994b).

[5] A second equilibrium exists in which all three participants choose $\delta = 0$ everywhere, so that no alliance is ever formed. This equilibrium relies on weakly dominated strategies and is therefore ruled out by trembling-hand perfection (Selten, 1975): in any slightly perturbed game, a participant whose joining choice could be decisive strictly prefers to join, making the never-join strategy suboptimal.

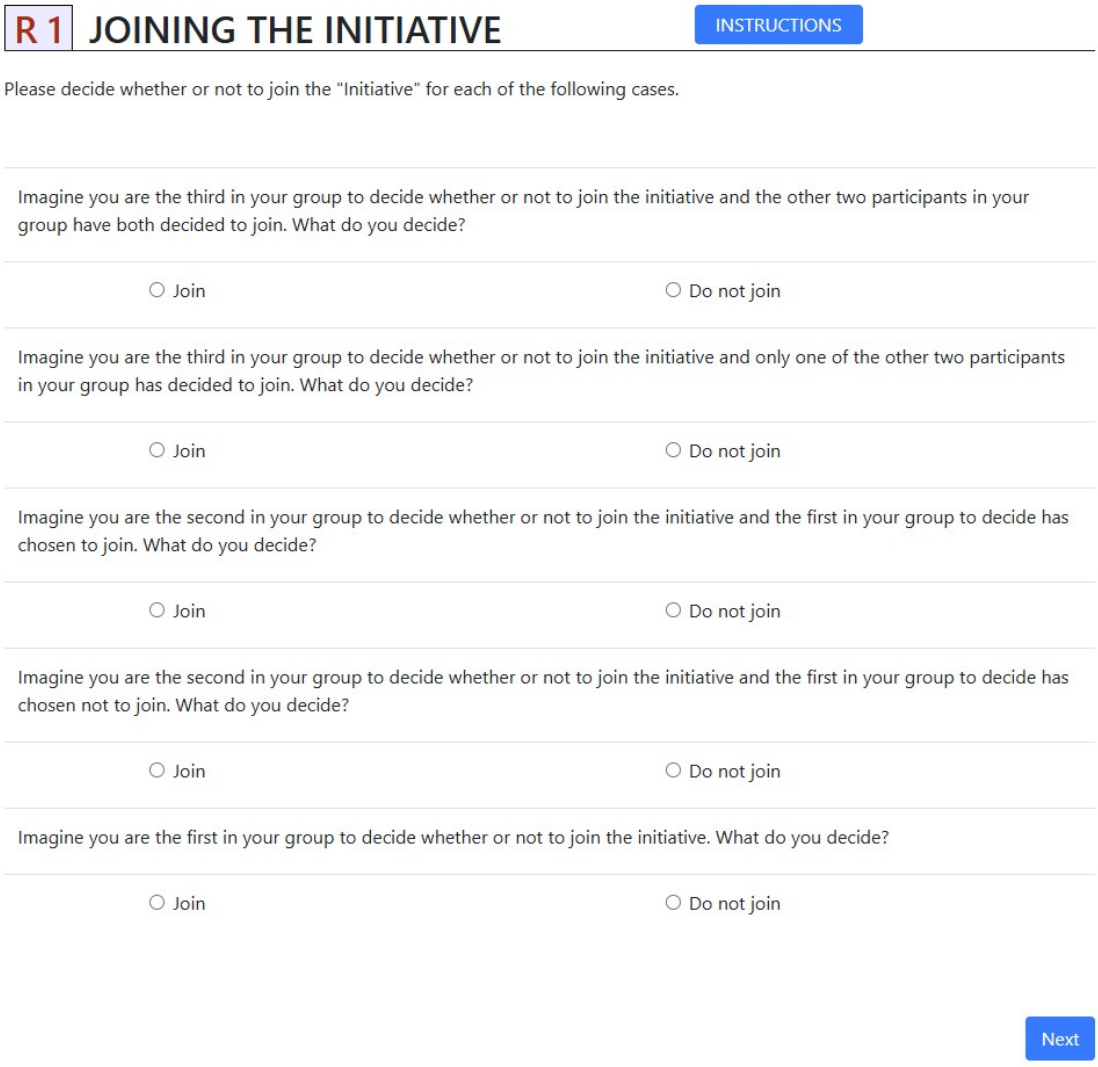


Figure 1: This is the oTree screen (translated into English) displayed to participants during the joining phase.

Each participant thus makes five binary decisions between joining ($\delta = 1$) and not joining ($\delta = 0$), which together constitute their *joining profile*,$(\delta_1\delta_2\delta_3\delta_4\delta_5)$. As it is illustrated in Table 1, this string corresponds, in order, to: 1st position with 0 joined before; 2nd position with 0 joined before; 2nd position with 1 joined before; 3rd position with 1 joined before; 3rd position with 2 joined before.

After all participants have submitted their profiles, a random sequence is drawn, and the realised joining decisions are determined accordingly. In the second stage each participant is endowed with $e = 25$ experimental currency units (ECU). Each contributor independently chooses an integer

Table 1: The five conditional joining choices

How many have joined before you	Your position in the sequence		
	3rd	2nd	1st
0	X	δ_2	δ_1
1	δ_4	δ_3	
2	δ_5		

Each cell requires a binary choice between joining ($\delta = 1$) and not joining ($\delta = 0$). When third in the sequence and no one has joined before, any choice is irrelevant since an alliance requires at least two members; this case is marked X.

contribution $c_i \in \{5, 6, \ldots, 19, 20\}$ to a public good.[6] Our experimental design distinguishes two between-subjects conditions that differ only in the MPCR: the low MPCR condition sets $\alpha = 0.6$ and the high MPCR condition sets $\alpha = 0.8$.

To illustrate, suppose $\alpha = 0.6$ and two participants form an alliance ($m = 2$) while the third remains an outsider. The two members choose contributions of 20 and 14, respectively, from their endowment of 25, and the outsider chooses 5. Total public good contribution is $C = 39$, of which 34 is contributed by alliance members. Each alliance member's cost is averaged: $C(m)/m = 34/2 = 17$, so both members earn $25 - 17 + 0.6 \times 39 = 31.4$ ECU, regardless of which one chose 20 and which chose 14. The outsider earns $25 - 5 + 0.6 \times 39 = 43.4$ ECU. One feature is worth noting. The cost-sharing rule implies that the member who contributes 14 bears the same cost as the member who contributes 20. Thus, joining an alliance commits a participant to paying an equal share of the alliance's total contribution cost. However, the design does not constrain individual contribution choices, nor does it impose the theoretically optimal contribution levels.

Choices in this stage are elicited using the strategy method: after submitting their joining profile, but before receiving any information about whether an alliance has formed, its size, or their role within it, each participant independently chooses how much to contribute for each alliance outcome that their joining profile has not excluded. The possible contribution decisions are:

(i) contribution when no alliance is formed ($c^{\emptyset}$);

(ii) contribution when a two-person alliance is formed and the participant is an outsider ($c^{2,o}$);

(iii) contribution when a two-person alliance is formed and the participant is a member ($c^{2,p}$);

[6] Restricting contributions to the interval $[5, 20]$ guarantees that, regardless of the interaction outcome, each participant earns both from what they keep for themselves and from the public good. Buso et al. (2025) implement a standard public goods game under the same contribution restriction and document the well-established pattern of contribution decay over time.

(iv) contribution when the grand alliance is formed (c^3).

A participant whose joining profile contains at least one $\delta = 0$ and at least one $\delta = 1$ must provide all four conditional contributions. A participant who never joins provides only (i) and (ii), since they have excluded the possibility of being an alliance member. A participant who always joins provides only (i), (iii), and (iv), since they have excluded the possibility of being an outsider to a formed alliance. The experimental software restricts the displayed contribution choices to those that are relevant given the participant's joining profile, as illustrated in Figure 2:

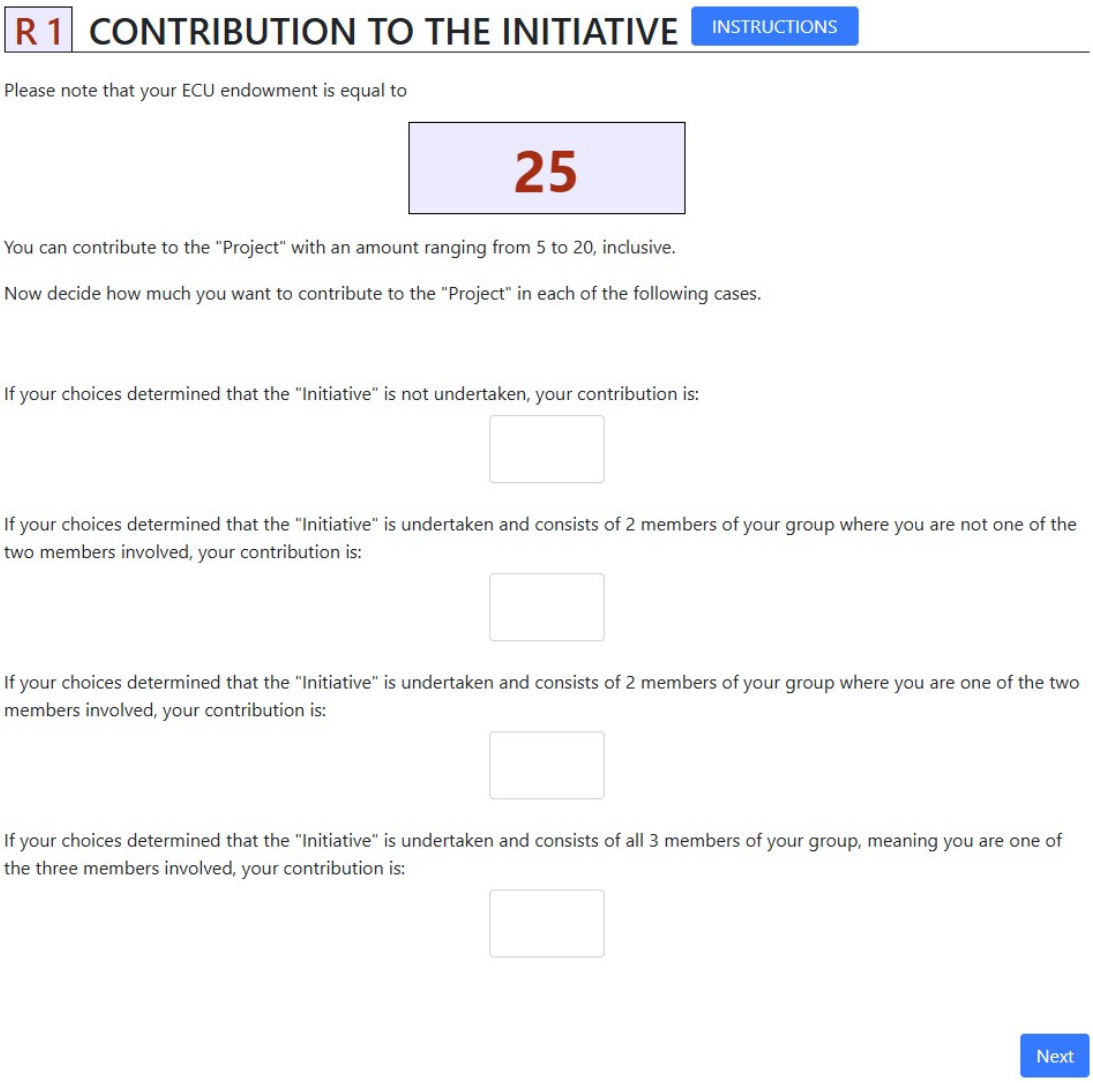


Figure 2: This is the oTree screen (translated into English) displayed to participants during the contribution phase whenever during the joining phase at least one joining decision is equal to 0 and at least one joining decision is equal to 1.

Both joining decisions and contribution decisions are elicited conditionally and simultaneously: participants commit to their full strategy before any uncertainty is resolved. This yields, for each participant in each round, a complete conditional schedule linking every possible alliance outcome to a contribution choice. The design maps behaviour across contingencies. It does not show how participants would reason in a realised one-shot environment.

After conditional contributions are elicited and before the contributions of the other group members are revealed, participants receive feedback on whether an alliance has been formed and, if so, on its size and their membership status. They are then asked to state their unincentivized beliefs about the contributions of the other group members. Finally, they are informed of the outcome of the round: their own contribution, the total group contribution, and their payoff for that round.

We collected data from 96 participants across four sessions conducted in January 2023.[7] We conducted four lab-like online sessions using the methodology developed in Buso et al. (2021), which recreates key laboratory conditions in an online environment through virtual cubicles on Cisco WebEx, ensuring participant monitoring and experimenter interaction while preserving the control and separation of a physical laboratory.

The experiment was programmed in oTree (Chen et al., 2016) and conducted with student participants from the LUISS Cesare Lab, recruited via ORSEE (Greiner, 2015) among students of Economics, Law, and Political Science. No participant took part in more than one session. The study was approved by the CESARE Lab, LUISS. All participants provided informed consent before participating in the experiment.

Within each session, participants were divided into matching groups of 6, from which groups of 3 were randomly formed in each round. Participants were aware that they would not be matched with the same group of 3 in two consecutive rounds, but were not informed of the matching group size.[8] No participant could identify the other members of their group during or after the session. After reading the instructions and answering control questions to verify comprehension, participants proceeded to the 12 rounds of play.[9] Each ECU was converted to euros at an exchange rate of 1 ECU = 0.50 euro. At the end of the experiment, one round was randomly selected for payment. Participants received their earnings from the selected round via PayPal, with average earnings of €19.90, in addition to a €6 show-up fee. Sessions lasted approximately 90 minutes.

[7] One matching group of 6 participants was removed from the analysis because a participant dropped during the session and had to be replaced by a laboratory assistant. The analysis therefore uses data from 90 participants.

[8] The MPCR condition varied across sessions; following convention in this literature, the matching group is the unit of independent observation throughout.

[9] The original Italian instructions and their English translation are reported in Appendix **??** and Appendix **??**, respectively.

3 Hypotheses

In a standard linear public goods game with $n = 3$ and $\alpha \in \{0.6, 0.8\}$, free-riding is the dominant strategy despite full contribution being socially efficient ($n\alpha > 1$). The typical experimental finding in such settings is contributions of around 40–60% of the endowment in early rounds, declining steadily toward minimal levels as the game is repeated (Zelmer, 2003; Ledyard, 1995). The cost-sharing alliance mechanism alters this environment by reducing the effective marginal cost of contributing for alliance members from 1 to $1/m$. When $m\alpha > 1$, which holds for all $m \geq 2$ in both α conditions, alliance members face no free-riding incentive and should contribute maximally. Indeed, under the parameterisation implemented in the experiment, the optimal prediction for both α conditions is that an alliance of size $m^* = 2$ forms, with members contributing maximally ($c = 20$) and outsiders contributing minimally ($c = 5$). If alliances form as predicted, contributions should therefore exceed those observed in standard Voluntary Contribution Mechanism experiments and should not exhibit the typical decay, since the institutional structure stabilises incentives across rounds.[10]

Hypothesis A (Alliance formation and cooperation)**.** *The equal cost-sharing mechanism generates persistent alliance formation; contributions are higher when an alliance exists than when none does; and contributions do not exhibit the decline over rounds that is typical of voluntary contribution mechanisms.*

We state these as a joint prediction because they are not independent. Alliance formation without sustained member contributions would indicate that participants join but then fail to exploit the cost-sharing rule's incentive to contribute, while sustained contributions without alliance formation would suggest that cooperation is driven by something other than the institutional structure. This hypothesis serves as a baseline confirmation that the mechanism operates as intended in our parametric environment ($n = 3$, $\alpha \in \{0.6, 0.8\}$).

The strategy-vector elicitation provides a complete conditional profile of joining and contribution decisions for each participant in every round, enabling us to investigate behavioural heterogeneity and shed light on the mechanisms through which cooperation is sustained. Because each participant makes five binary joining decisions per round, there are $2^5 = 32$ possible joining profiles. Only one is consistent with the benchmark,i.e., $(\delta_1 = 0, \delta_2 = 1, \delta_3 = 0, \delta_4 = 1, \delta_5 = 0)$. We do not expect participants to align with it. Rather, joining decisions may reflect behavioural considerations that depart from the backward-induction logic underlying the equilibrium prediction. Participants may, for instance, exhibit a form of conditional participation, joining only when others have already committed, or they may fail to engage in the backward-induction reasoning

[10] See Ledyard (1995); Zelmer (2003); Chaudhuri (2011), on the VCM decline, usually attributed to imperfectly conditional cooperators revising downwards (Fischbacher and Gächter, 2010)

required by the equilibrium strategy. Whether systematic joining patterns emerge and generate profiles that remain stable within individuals across rounds, and whether these profiles predict contribution behaviour in ways that reflect distinct cooperative attitudes, are empirical questions that the strategy-vector design is uniquely suited to address.

Hypothesis B (Behavioural heterogeneity)**.** *Joining behaviour is heterogeneous and structured: distinct joining profiles emerge as recurring patterns across participants and rounds, rather than as idiosyncratic or round-specific noise. Whether these profiles also predict contribution behaviour is examined as part of the same analysis.*

This hypothesis has a two-part structure. The first part, that distinct joining profiles emerge and persist within individuals, is a precondition for any meaningful behavioural typology. The second part, whether these profiles predict contribution behaviour, and if so with what pattern, is treated as an open empirical question rather than a directional prediction.

Finally, we examine the robustness of our findings across the two MPCR conditions. If the joining types and associated contribution patterns that may emerge from the analysis reflect stable individual dispositions toward participation in cost-sharing institutions, they should be similar across both MPCR levels.[11]

Hypothesis C (Robustness of behavioural heterogeneity)**.** *The distribution of joining types and the associated contribution patterns persist across the two MPCR conditions.*

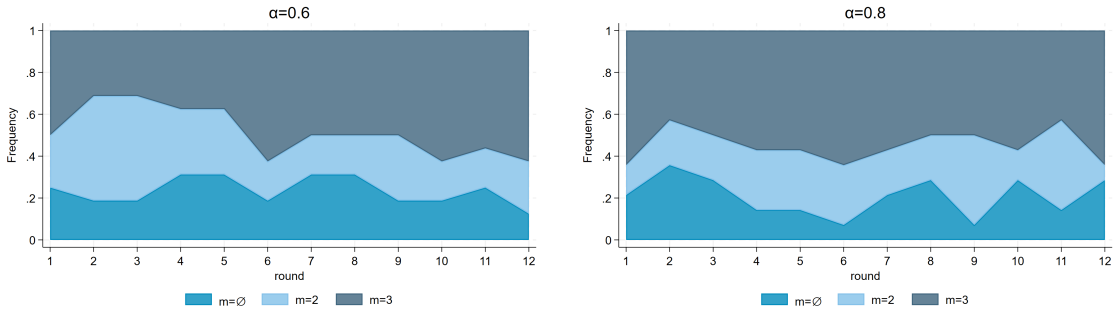
4 Results

Although the hypotheses in Section 3 are organized by topic—aggregate outcomes (A), behavioural heterogeneity (B), and robustness across MPCR conditions (C)—to provide a more streamlined presentation of the findings, the Results section is structured around the two key decisions in the experiment. We first analyse alliance formation at both the aggregate and individual levels, and then examine contribution behaviour using the same approach, while assessing the robustness of the findings across MPCR conditions throughout.

[11] Although the benchmark predicts identical behaviour across MPCR conditions, the weaker free-riding incentive under the higher MPCR condition (cost of contributing $1 - \alpha = 0.2$) may encourage contributions, joining, and grand alliance formation at the margin. In standard VCMs, higher MPCRs raise contributions (Isaac and Walker, 1988; Zelmer, 2003), although distributions of conditional-cooperation types appear insensitive to MPCR and group size, with beliefs accounting for differences in levels (Gächter and Marino Fages, 2023).

4.1 Alliance Formation

Alliances form in 78% of rounds across the full sample, with grand alliances ($m = 3$) accounting for roughly half of all alliances formed (51%). Formation is slightly more frequent when the MPCR is equal to 0.8 than when 0.6: 80% versus 77% overall, and 54% versus 48% for grand alliances. None of these differences reach statistical significance.[12] Figure 3 shows that the overall formation pattern is broadly stable across rounds. However, grand alliance formation in the high MPCR condition displays a gradual upward trend, starting from a slightly lower level than when MPCR is equal to 0.6 and converging to a similar level by the end of the experiment.[13]



$\alpha = 0.6$ $\alpha = 0.8$

Figure 3: Alliance size frequencies over rounds, by MPCR condition.

Result 1a *In both MPCR conditions, alliances consistently emerge across rounds. However, the theoretically predicted alliance size is observed in only about half of the cases, indicating substantial deviations from the benchmark prediction.*

Table 2 reports the empirical relative frequency of joining ($\delta = 1$) for each cell of the conditional joining table, by MPCR condition. The joining frequencies are far higher than the theoretical prediction. Indeed, the optimal profile, i.e., the one consistent with the theoretical benchmark developed in Section 2.1, prescribes $\delta = 1$ in only two of the five cells - joining as second-mover when no one has joined before, and joining as third-mover when one has joined. Yet participants join with frequency between 0.38 and 0.82 across all cells, with no cell exhibiting near-zero joining. Instead, joining seems conditional on others' prior decisions in a pattern more consistent with reciprocity than with the benchmark: when in second or third position, participants are substantially more

[12] Two-sample t-tests yield $p = 0.55$ for overall formation and $p = 0.23$ for grand alliance formation; Wilcoxon rank-sum tests yield identical conclusions.

[13] All alliance formation results reported in this section are based on the realized random sequences and the corresponding implemented alliance outcomes. Because the strategy-vector design elicits complete joining profiles, Appendix **??** complements this analysis with simulations based on all feasible sequence realizations, showing that the main patterns are robust when the full set of potential alliance outcomes is considered.

likely to join when one or more others have already joined than when none have. Paired tests at the individual level comparing the frequency of choices in the optimal and non-optimal cells in the scenario in which the participant is second in the sequence (0.38 vs. 0.82 in $\alpha = 0.6$; 0.41 vs. 0.79 in $\alpha = 0.8$) confirm that this within-subject difference is statistically significant in both conditions ($p < 0.001$). A similar, though smaller, gap appears when participants are third in the sequence (0.68 vs. 0.78 in $\alpha = 0.6$; 0.68 vs. 0.77 in $\alpha = 0.8$).[14]

Table 2: Relative frequency of joining ($\delta = 1$) by cell and MPCR condition.

	$\alpha = 0.6$			$\alpha = 0.8$		
How many joined	Position in sequence			Position in sequence		
before you	3rd	2nd	1st	3rd	2nd	1st
0	X	**0.38**	0.71	X	**0.41**	0.79
1	**0.68**	0.82		**0.68**	0.79	
2	0.78			0.77		

Cells in which $\delta = 1$ is the benchmark-prediction are in **bold**. The optimal profile prescribes $\delta = 1$ only in these cells; observed frequencies far exceed the prediction in the non-bold cells and are well below 1 in the bold cells.

The five binary joining choices generate $2^5 = 32$ possible joining profiles. If joining behaviour were unstructured, we would expect observations to be roughly uniformly distributed across these profiles. Instead, observations are concentrated in a small number of interpretable patterns. Table 3 reports the five most prevalent profiles - those accounting for at least 5% of all observations in at least one MPCR condition - together with the optimal profile.

Table 3: Most prevalent joining profiles and the benchmark profile.

Profile	Pattern ($\delta_1\delta_2\delta_3\delta_4\delta_5$)	$\alpha = 0.6$	$\alpha = 0.8$
Unconditional Join	11111	25%	27%
Initiator-Followers	10111	23%	23%
Followers	00111	7%	3%
Initiator-Selective Followers	10101	10%	13%
Never Join	00000	5%	7%
Optimal	01010	2.5%	3%

Cells of the profile string $\delta_1\delta_2\delta_3\delta_4\delta_5$ correspond, in order, to: 1st position with 0 joined before; 2nd position with 0 joined before; 2nd position with 1 joined before; 3rd position with 1 joined before; 3rd position with 2 joined before. The full distribution over all 32 profiles is reported in Appendix **??**.

The five most prevalent profiles together account for approximately 70% of all observations in

[14]The difference in the low MPCR condition is significant at the 5% level.

both conditions. The most common is the *Unconditional Join* profile ($\delta = 1$ in every cell), which represents around 26% of choices. The second most common is the *Initiator-Follower*, who joins as the first mover and joins whenever at least one other player has already joined, refusing only when in second position with no prior joiner. The *Follower* never initiates but, like the *Initiator-Follower*, joins whenever at least one other player has already joined. The *Initiator-Selective Follower* joins as the first mover and otherwise joins only when all preceding players have joined. As such, it can be viewed as a more selective form of conditional joiner than the other follower types: they only join if there is still a chance to complete a grand alliance. Under the *Never Join* profile, participants refuse under all conditions.

The optimal profile, which prescribes refusing as first mover, joining only when pivotal, accounts for fewer than 3% of observations in both conditions. Although the institutional mechanism generates the cooperation that theory predicts, the joining patterns through which it does so bear little resemblance to the sub-game perfect strategy. Participants converge on cooperation, but they do so primarily by being more willing to join than benchmark-prediction prescribes, not by following the optimal joining logic. In other words, the mechanism succeeds behaviourally through systematic "pro-entry" deviations from the benchmark, rather than through convergence to the benchmark itself.

Result 1b *A small number of systematic joining profiles account for most observed behaviour and sustain alliance formation across rounds, despite deviating substantially from the benchmark predictions.*

Profiles	$\alpha = 0.6$				$\alpha = 0.8$			
	% choices		% subjects	% subject round 1	% choices		% subjects	% subject round 1
	All rounds	Rounds 7–12			All rounds	Rounds 7–12		
Unconditional Join	25.2	31.6	27.1	12.5	27.4	34.1	33.3	4.8
Initiator-Followers	22.9	21.9	25.0	25.0	22.8	20.6	21.4	26.2
Followers	7.1	5.2	4.2	8.3	3.4	3.2	0.0	4.8
Initiator-Selective Followers	10.2	9.7	6.2	10.4	13.3	13.1	11.9	14.3
Never Join	5.0	4.9	4.2	0.0	6.9	6.7	4.8	4.8
Optimal	**2.4**	**2.1**	**2.1**	**0.0**	**3.0**	**2.8**	**2.4**	**2.4**

Table 4: This table reports the frequencies of profiles that account for at least 5% of choices across rounds in one of the two MPCR conditions. At the bottom of the table in bold it is reported also the profile of joining predicted by theory. For each profile, we report its frequency across all rounds and in the second half of the experiment. In addition, the third column for each condition reports the percentage of subjects who use the profile in at least 6 out of 12 rounds, and the fourth column reports the percentage of subjects who use the profile in round 1.

Table 3 shows that the most frequent joining profiles emerge in both MPCR conditions. Table 4 adds the time dimension and reveals two distinct patterns. The conditional profiles are present at close to their steady shares from the first round onwards: Initiator-Followers account for 25% and

26% of round-1 choices and 22–23% of choices overall, with similarly stable figures for Initiator-Selective Followers. Unconditional joining, by contrast, grows markedly with experience: it is used by only 12% ($\alpha = 0.6$) and 5% ($\alpha = 0.8$) of subjects in round 1 but accounts for 31% and 34% of choices in the second half of the experiment, while outright refusal, absent in round 1 in the low MPCR condition, also emerges only later. The frequency of the optimal profile does not increase over time, so the movement is not convergence towards the benchmark; it is a shift towards greater willingness to join. The subject-level classification (% subjects) shows that individuals nevertheless employ the same profile in at least half of the rounds, so the drift reflects gradual changes in which profile individuals settle on rather than round-to-round noise.[15]

Result 1c *The same small set of joining profiles recurs in both MPCR conditions from the first round onwards, and individuals employ their profile consistently across rounds. The composition is not static, however: the conditional profiles hold roughly constant shares throughout, while the share of unconditional joining rises from 12% and 5% of round-1 choices to 31% and 34% of choices in rounds 7–12. Heterogeneity in joining is thus stable in kind but drifts towards greater cooperativeness over the course of the experiment.*

Finally note that, on average, subjects earn more per round when they are members of a three-person alliance (21.70 euros) than in any other configuration, followed by outsiders (18.72 euros) and members of a partial alliance (18.09 euros); when no alliance exists subjects earn the least (17.11 euros). Regression estimates reported in Appendix **??** confirm that the earnings advantage of three-person alliances and of outsiders relative to subjects with no alliance is statistically significant, while the difference for partial-alliance members is not. Unconditional Join subjects earn slightly more on average (20.76 euros) than Conditional Join subjects (20.10 euros), and both outperform Never Join subjects (18.44 euros); here, only the gap between Unconditional Join and Never Join subjects is statistically significant.

4.2 Contribution Behaviour

Table 5 reports average conditional contributions for each alliance outcome.[16] Alliance membership substantially increases contributions: average conditional contributions rise from around 8.5 when no alliance is formed to 14.5 as a member of a partial alliance ($m = 2$) and 15.9 in the grand

[15] Table **??** in Appendix **??** shows that the only significant increase in joining profiles when comparing round 1 to rounds 7–12 is that of unconditional joining.

[16] The actual contribution implemented depends on the random sequence drawn. The conditional contributions record each participant's individual stated amount for each relevant alliance outcome before sequence resolution and before averaging. The actual contribution for Table 5 as well as the simulation for all possible random sequence draws are presented in the appendix as a robustness check for our finding in Appendices **??** and **??**. In addition, the dynamics of actual and simulated contributions are reported in these appendices.

alliance ($m = 3$), and this pattern holds in both MPCR conditions. Table 6 reports OLS regressions on conditional contributions, and confirms that alliance membership significantly increases contributions: relative to the no-alliance baseline, contributions increase by approximately 6 ECU when a partial alliance is formed and by 7.5 ECU in the grand alliance. Also, the MPCR condition does not have a statistically significant effect on conditional contributions.

Table 5: Average conditional contributions by alliance outcome and MPCR condition.

	$c^{\emptyset}$	$c^{2,o}$	$c^{2,p}$	c^{3}
$\alpha = 0.6$	8.08 (4.67)	8.50 (3.81)	14.32 (5.28)	15.55 (5.21)
$\alpha = 0.8$	8.87 (5.20)	8.31 (4.56)	14.70 (5.36)	16.27 (5.57)
Overall	8.45 (4.94)	8.41 (4.17)	14.50 (5.32)	15.88 (5.39)

Notes: Standard deviations in parentheses. Conditional contributions are the stated amounts for each scenario before the alliance outcome is realised. $c^{\emptyset}$ denotes contributions in the absence of an alliance, $c^{2,o}$ contributions by non-members of a partial alliance, $c^{2,p}$ contributions by members of a partial alliance, and c^{3} contributions by members of a grand alliance.

Figure 4 displays the time dynamics of conditional contributions. In contrast to the well-documented decline of contributions in standard VCM experiments, contributions in our setting are stable or moderately increasing over rounds, even with rematching between rounds (the partners–strangers comparisons of Andreoni, 1988). This pattern holds across both MPCR conditions and across alliance outcomes. Because the design uses rematching rather than fixed groups, these patterns should be read as round-level aggregate trends rather than stability within the same cooperating group. The positive and statistically significant time trend that emerges in Table 6 confirms that contributions do not decay over rounds. The aggregate pattern is therefore consistent with the mechanism working in this new parametric environment. Several additional features emerging from Table 5 merit discussion. In particular, contribution patterns do not appear to align with the theoretical solution. The benchmark predicts a sharp discontinuity: alliance members should contribute maximally ($c = 20$) regardless of alliance size, while outsiders and participants without an alliance should contribute minimally ($c = 5$). The data depart from this prediction in both directions. Members contribute substantially above the minimum but well below the maximum: around 14.5 in partial alliances and 15.9 in the grand alliance, against a benchmark of 20. Outsiders and participants without an alliance contribute well above the predicted minimum of 5, averaging around 8.4–8.5 in the conditional data. Neither alliance members nor outsiders behave as the benchmark predicts, though both move in the predicted direction: members contribute considerably more than non-members.

Furthermore, the benchmark predicts no difference between contributions as a member of a partial alliance ($m = 2$) and contributions in the grand alliance ($m = 3$), since both scenarios yield

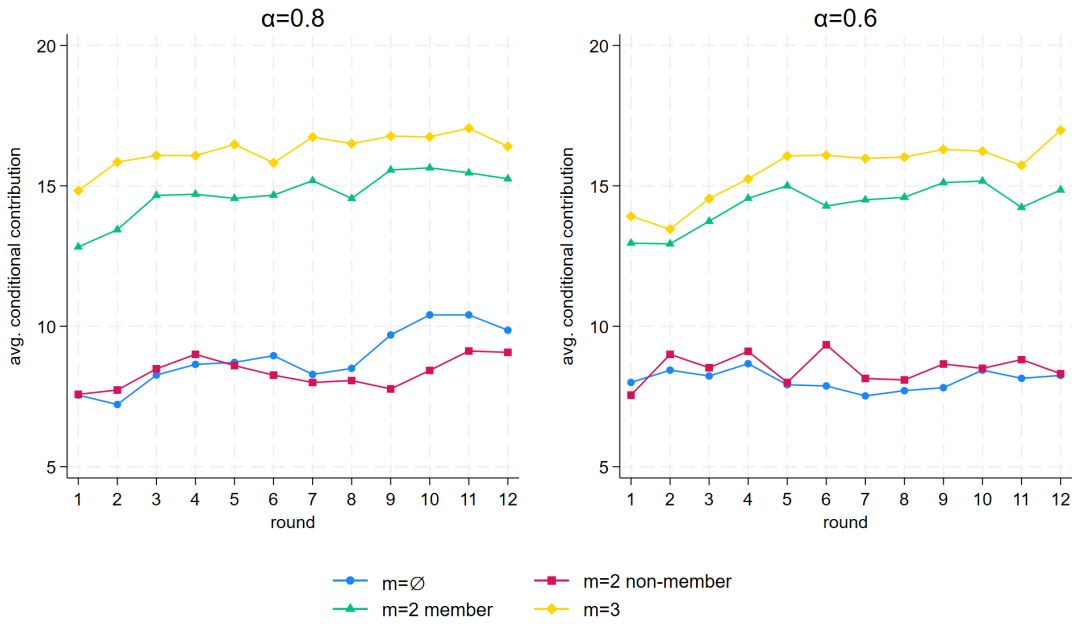


Figure 4: Average conditional contributions across rounds, by alliance outcome and MPCR condition.

$m\alpha > 1$ and should therefore elicit maximal contributions. Similarly, the benchmark predicts no difference between contributions when no alliance exists ($c^{\emptyset}$) and contributions as an outsider to a formed alliance ($c^{2,o}$), since in both cases the dominant strategy is to contribute minimally.

According to the test provided in Table 7, the conditional contribution data reject the first prediction but not the second. Average contributions as a member of the grand alliance (15.9) exceed contributions as a member of a partial alliance (14.5), a difference of approximately 1.4 ECU.

Average contributions as an outsider to a formed alliance (8.4) are essentially the same as contributions when no alliance exists (8.5), providing no evidence of conditional cooperation on the outsider margin. Outsiders do contribute above the benchmark minimum of 5, but they do not contribute more when an alliance exists than when it does not. The evidence, therefore, points to conditional cooperation on the member side but not on the outsider side, which suggests that alliance participation itself matters for how cooperative signals are received.

Result 2a *Alliance formation sustains and stabilises cooperation over rounds. At the same time, contribution behaviour deviates systematically from the benchmark predictions, with members contributing below the theoretical maximum and non-members above the theoretical minimum. Moreover, contributions are higher in the grand alliance than in a partial alliance, suggesting conditional cooperative responses among alliance members.*

Table 6: OLS regression on conditional contributions.

	Conditional Contribution
m=2, outsider	0.01
	(0.42)
m=2, member	6.05***
	(0.60)
m=3	7.43***
	(0.57)
MPCR = 0.6	-0.47
	(0.58)
Period	0.14***
	(0.03)
Constant	7.82***
	(0.44)
Observations	3909

Notes: OLS regression with robust standard errors clustered at the matching group level in parentheses. The baseline category for alliance condition is m = $\emptyset$. The baseline for MPCR is 0.8. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Contributions by Join Profiles

The evidence presented above establishes that the cost-sharing alliance mechanism generates cooperation and that contributions respond to alliance context in ways that go beyond payoff maximisation. We now ask what drives these aggregate patterns at the individual level. The strategy-vector design makes this possible: because each participant's joining profile and conditional contribution schedule are observed jointly, we can test whether the joining types identified in Section 4.1 predict contribution behaviour. Table 8 reports average conditional contributions separately for the Unconditional Join, Conditional Join (pooling Initiator-Followers, Followers and Initiator-Selective Followers), and Never Join profiles.

Our categories are reminiscent of the contribution types identified by Fischbacher et al. (2001), in that both separate participants who cooperate unconditionally from those whose cooperation is contingent on others and from those who never cooperate. The classifications are not interchangeable: ours is built on entry decisions, where joining coincides with the best response in two of the five conditional cells, whereas the contribution schedules used by Fischbacher et al. (2001) are costly in every state. What the two have in common is that a small number of interpretable types account for most behaviour, and that conditional cooperation is the most common of them.

Table 7: Wilcoxon signed-rank tests for conditional contributions.

Null hypothesis	N	p-value
$c^{\emptyset} = c^{2,o}$	15	0.5245
$c^{2,o} = c^{2,p}$	15	0.0001
$c^{2,p} = c^{3}$	15	0.0004

Notes: This table reports the p-values of two-sided Wilcoxon signed-rank tests for pairwise comparisons of conditional contributions across alliance structures. $c^{\emptyset}$ denotes contributions in the absence of an alliance, $c^{2,o}$ contributions by non-members of a partial alliance, $c^{2,p}$ contributions by members of a partial alliance, and c^{3} contributions by members of a grand alliance. N indicates the number of independent observations (matching groups).

Table 8: Average conditional contributions by joining type.

	Actual	$c^{\emptyset}$	$c^{2,o}$	$c^{2,p}$	c^{3}
Unconditional Join	16.03 (4.95)	9.93 (5.47)	—	15.32 (4.87)	16.73 (4.68)
Conditional Join	13.61 (6.32)	7.52 (4.28)	8.31 (4.00)	15.47 (4.84)	17.03 (4.52)
Never Join	7.03 (3.55)	7.34 (4.01)	7.64 (4.13)	—	—

Notes: Standard deviations in parentheses. Actual denotes the amount actually contributed; $c^{\emptyset}$, $c^{2,o}$, $c^{2,p}$, and c^{3} denote conditional contributions stated for each alliance scenario before the outcome is realised. Cells marked —- indicate contribution choices that are excluded by the joining profile: under Unconditional Join, participants are never outsiders to a formed alliance; under Never Join, participants are never alliance members. Conditional Join pools the Initiator-Follower, Initiator-Selective Follower, and Follower profiles.

Three patterns emerge. First, joining type predicts actual contributions: contributions are highest under the Unconditional Join profile (16.0), intermediate under Conditional Join (13.6), and lowest under Never Join (7.0). This evidence is confirmed by the regression analysis of actual contributions reported in Table 9: actual contributions under Unconditional Join are 2.1 ECU higher than under Conditional Join and 8.9 ECU higher than under Never Join. These differences in actual contributions may reflect both a direct effect—participants following Unconditional Join are more likely to be in an alliance (self-selection), where contributions are higher—and a behavioural pattern whereby joining willingness and contribution levels are positively correlated even within the same alliance outcome, reflecting stable differences in cooperativeness.

The analysis of conditional contributions sheds light on this issue. The between-type differences persist even when comparing contributions within the same alliance outcome. When no alliance is formed, average conditional contributions are 9.9 under Unconditional Join, compared to 7.5 under Conditional Join and 7.3 under Never Join. In this setting, all three types face identical incentives (no cost-sharing and standard VCM incentives), yet contributions under Unconditional Join are still higher. The differences between Unconditional Join and both Conditional Join and

Never Join in the no-alliance condition are statistically significant at the 5% level, as reported in Table 9. The same patterns emerge when considering the subject-level classification of profiles, which requires a profile to be used in at least 6 out of 12 rounds, as reported in Table **??** in the Appendix.

Within each joining profile, contributions increase with alliance size. Under Unconditional Join, average contributions rise from 15.3 as a partial alliance member to 16.7 in the grand alliance. Under Conditional Join, the increase is from 15.5 to 17.0.[17] Results are robust to classifying subject profiles at the individual level, requiring a profile to be used in at least 6 out of 12 rounds.[18] This confirms the conditional cooperation finding from Section 4.2 and shows that it is not driven by compositional differences across alliance sizes: even the same types of participants contribute more when the alliance is larger.[19]

Result 2b *Joining profiles predict actual contribution levels: contributions are highest under Unconditional Join, intermediate under Conditional Join, and lowest under Never Join. The higher contributions under Unconditional Join persist even when no alliance is formed and all participants face identical incentives, suggesting that joining profiles capture stable differences in cooperativeness rather than merely differences in alliance membership.*

It is worth noting that the MPCR coefficient is generally negative, although it does not reach statistical significance in any of the conditional-contribution specifications and is significant only at the 5% level in the actual-contribution regressions. The negative sign is nonetheless consistent with the weaker free-riding incentive under the higher MPCR condition ($\alpha = 0.8$) relative to the lower one ($\alpha = 0.6$), which may reduce the marginal impact of the cost-sharing mechanism on contribution behaviour. The overall weakness of MPCR effects suggests that, in this setting, the institutional structure matters more than variation in α. The mechanism reduces each member's marginal cost of contributing from 1 to $1/m$, which is a substantially larger change than the variation in α between our two conditions. This is consistent with the cost-sharing rule mattering more than the MPCR once participants are inside the alliance. Once they are outside, the entry decision rather than the marginal contribution incentive is the dominant strategic margin.

Result 2c *Differences in MPCR have at most a limited effect on contribution behaviour, suggesting*

[17] Wilcoxon signed-rank tests on matching-group averages (averaging first within subjects across rounds, then across subjects within matching group; $N = 15$ independent observations) confirm that the increase from $c^{2,p}$ to c^3 is statistically significant for both Unconditional Join ($p < 0.001$) and Conditional Join ($p < 0.01$).

[18] The difference between $c^{2,p}$ and c^3 remains statistically significant for both Unconditional Joiners ($p < 0.001$) and Conditional Joiners ($p < 0.001$).

[19] The beliefs by joining type reported in Table **??** in Appendix **??** consistently show that Unconditional Join profiles expect greater contributions from other group members when no partnership exists, while when a partnership exists, the Conditional Join profile shows slightly higher beliefs than Unconditional Join profiles, coherently with their slightly greater contributions.

that the institutional structure of the alliance plays a more important role than variation in the underlying return to the public good.

Table 9: OLS regressions on contributions by profile.

	(1)	(2)	(3)	(4)	(5)
	Actual	$c^{\emptyset}$	$c^{2,o}$	$c^{2,p}$	c^{3}
Conditional Join	-2.15**	-2.33*	0.00	0.41.	0.56.
	(0.67)	(0.96)	(0.00)	(0.74)	(0.50)
Never Join	-8.92***	-2.60*	-0.71.	–	–
	(0.77)	(0.89)	(0.72)		
MPCR = 0.6	-2.13*	-1.09.	-0.54.	-1.21.	-1.54.
	(0.89)	(0.75)	(0.46)	(0.99)	(0.89)
Period	0.16.	0.04.	-0.00.	0.17**	0.16*
	(0.09)	(0.06)	(0.05)	(0.05)	(0.06)
Constant	15.94***	10.21***	8.61***	14.68***	16.30***
	(1.26)	(0.87)	(0.38)	(0.90)	(0.93)
Observations	778	778	495	714	714
R-squared	0.1860	0.0708	0.0076	0.0290	0.0433

Notes: OLS regressions with robust standard errors clustered at the matching group level in parentheses. The baseline category for profile is *profile32* (Unconditional Join). In column (3), participants following Unconditional Join never state $c^{2,o}$, so the stated baseline is absent from that subsample: the Conditional Join coefficient is collinear with the constant and is omitted (reported as 0.00), the constant refers to Conditional Join, and the Never Join coefficient is measured relative to it. The baseline for MPCR is 0.8. Columns (1)–(5) refer respectively to actual contribution and to conditional contributions. Categories not present in the estimation subsample are indicated with –. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Conditional Join Profiles

The preceding analysis pools the three prominent conditional joining profiles: Initiator-Followers, Initiator-Selective Followers, and Followers, into a single category. While this pooling is appropriate for the main analysis, the three sub-types differ in their joining logic and may therefore differ in their contribution behaviour. Table 10 disaggregates the Conditional Join category.

Table 10: Average conditional contributions by conditional joining sub-type (pooled across MPCR conditions).

	Actual	$c^{\emptyset}$	$c^{2,o}$	$c^{2,p}$	c^{3}
Initiator-Followers	13.99 (6.32)	7.49 (4.27)	8.56 (4.28)	16.29 (4.55)	17.50 (4.31)
Initiator-Selective Followers	13.67 (6.55)	7.32 (4.34)	7.67 (3.52)	15.27 (4.81)	17.00 (4.77)
Followers	11.84 (5.57)	8.09 (4.24)	8.60 (3.56)	12.38 (4.92)	15.09 (4.37)

Standard deviations in parentheses. Averages computed over all observations in which the participant's joining profile in that round matches the indicated sub-type. *Initiator-Followers* (profile 10111) join as first mover and join whenever at least one other has joined, refusing only as second mover with no prior joiner. *Initiator-Selective Followers* (profile 10101) join as first mover and join only when their entry could still produce the grand alliance ($m = 3$), refusing to be the second member of a partial alliance. *Followers* (profile 00111) never initiate, joining only after at least one other has joined.

Initiator-Followers and Initiator-Selective Followers exhibit similar contribution patterns: both contribute substantially as alliance members (around 15–16 in partial alliances and 17 in the grand alliance) and modestly above the benchmark minimum as outsiders or without an alliance. Followers, by contrast, contribute noticeably less as alliance members: around 12.4 in partial alliances and 15.1 in the grand alliance. This is consistent with the behavioural logic of the Follower profile: participants who are reluctant to initiate alliance formation may also be more cautious in their contribution commitments. The gap between Followers and the other two conditional sub-types is most pronounced in the partial alliance member condition ($c^{2,p}$), where Followers average 12.4 compared to 16.3 for Initiator-Followers and 15.3 for Initiator-Selective Followers. Still, these differences are not statistically significant, as reported in Table 11.

All three conditional sub-types share the conditional cooperation pattern: contributions increase from $c^{2,p}$ to c^{3}. Initiator-Followers increase from 16.3 to 17.5, Initiator-Selective Followers from 15.3 to 17.0, and Followers from 12.4 to 15.1. The increase is significant for both Initiator-Selective Followers and Followers, whereas it is not significant for Initiator-Followers. The increase is largest for Followers (2.7 ECU), indicating that the contribution gap between Followers and the other conditional profiles narrows when moving from a partial to the grand alliance.[20]

[20] Within-profile differences between $c^{2,p}$ and c^{3} are tested in a pooled regression including interactions between the contribution condition and the conditional joining profiles, with standard errors clustered at the matching-group level. The estimated increases are 1.21 ECU for Initiator-Followers ($p = .204$), 1.73 ECU for Initiator-Selective

Table 11: OLS regressions on contributions by conditional profile.

	(1) Actual	(2) $c^{\emptyset}$	(3) $c^{2,o}$	(4) $c^{2,p}$	(5) c^{3}
Initiator-Selective Follower	-0.49.	-0.25.	-0.92.	-1.12.	-0.62.
	(0.92)	(0.63)	(0.85)	(1.12)	(1.17)
Followers	-1.53.	0.80.	0.09.	-3.60.	-1.99*
	(1.00)	(1.39)	(1.10)	(1.93)	(0.84)
MPCR = 0.6	-2.63*	-1.12.	-0.34.	-1.50.	-1.71.
	(0.92)	(1.06)	(0.43)	(1.12)	(0.81)
Period	0.18.	0.02.	-0.00.	0.06.	0.14.
	(0.14)	(0.08)	(0.06)	(0.07)	(0.08)
Constant	14.26***	7.98***	8.77***	16.74***	17.51***
	(1.30)	(0.84)	(0.56)	(0.94)	(0.89)
Observations	431	431	431	431	431
R-squared	0.0638	0.0199	0.0124	0.0969	0.0772

Notes: OLS regressions with robust standard errors clustered at the matching group level in parentheses. The baseline category is *Initiator-Follower* (profile24). The baseline for MPCR is 0.8. Columns (1)–(5) refer respectively to actual contribution and to conditional contributions. Categories not present in the estimation subsample are indicated with –. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Path Dependence and beliefs

We briefly examine whether contribution behaviour responds to past experience and beliefs on others' contribution. If participants adjust their contributions in response to previous-round outcomes, this would suggest that the cooperation dynamics observed in Section 4.2 are also shaped by learning or reciprocity across rounds. Rematching makes such carry-over less likely, but it does not rule it out: contribution behaviour has been shown to spill over across behaviourally distinct public good interactions, both within and between individuals (Angelovski et al., 2018).

Table 12 reports regressions of actual contributions on the participant's own contribution in the previous round, the contributions of other group members in the previous round, the MPCR condition, and round.

Across all types, a participant's own previous contribution is a strong predictor of current contribution, with coefficients in the range of 0.27 to 0.37. By contrast, the contributions of other group members in the previous round have no significant effect. This pattern is consistent with the conditional elicitation design: because participants commit to their contributions before

Followers ($p = .001$), and 2.71 ECU for Followers ($p = .027$).

Table 12: Path dependence in actual contributions.

	(1)	(2)	(3)	(4)
	Overall	Never Join	Unconditional Join	Conditional Join
Own contribution $(t-1)$	0.37^{***}	0.36^{**}	0.27^{**}	0.30^{*}
	(0.05)	(0.09)	(0.09)	(0.10)
Others' contribution $(t-1)$	$0.01^{\cdot}$	$0.05^{\cdot}$	$-0.02^{\cdot}$	$-0.03^{\cdot}$
	(0.03)	(0.05)	(0.04)	(0.04)
$\alpha = 0.6$	$-1.11^{\cdot}$	$-1.12^{\cdot}$	$-0.99^{\cdot}$	-2.09^{*}
	(0.71)	(0.71)	(1.02)	(0.94)
Round	0.15^{*}	$0.02^{\cdot}$	$0.08^{\cdot}$	$0.20^{\cdot}$
	(0.06)	(0.09)	(0.12)	(0.12)
Constant	7.78^{***}	3.70^{*}	12.46^{***}	10.02^{***}
	(1.15)	(1.25)	(1.88)	(1.74)
Observations	990	62	275	391

OLS regressions with standard errors clustered at the matching-group level. Columns (2)–(4) restrict to observations in which the participant's joining profile in that round matches the indicated type. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

learning the realised alliance outcome, they cannot directly condition on others' contemporaneous behaviour. The within-individual persistence in contributions, combined with the absence of response to others' past behaviour, is consistent with contributions being driven primarily by persistent individual dispositions rather than by round-to-round imitation of others' contribution levels. This is a departure from the dynamics of standard VCMs, where responsiveness to others' contributions, filtered through beliefs, drives the decay of cooperation (Fischbacher and Gächter, 2010); here the elicitation is conditional, and the response runs through the alliance structure rather than through others' amounts. Indeed, the previous round's alliance outcome affects contributions, and beliefs have a pattern that closely resembles contribution in each alliance condition: Table 13 shows that having experienced a grand alliance in the previous round is associated with significantly higher contributions overall (2.5 ECU above the no-alliance baseline), and the point estimate is largest for Never Join profiles (4.0 ECU), though this result should be interpreted with caution given the small number of observations in this category (N=62, corresponding to fewer than ten individuals). Having been a member of a partial alliance in the previous round also raises contributions overall (2.2 ECU), though this effect does not reach significance within any individual type. By contrast, having been an outsider to a partial alliance in the previous round is associated with significantly lower contributions overall (1.9 ECU below the no-alliance baseline), although this effect is not significant within any individual type.

Beliefs not only have a general positive correlation with contribution (see Table **??** in Appendix **??**), but, as shown in Table 14, they also follow a pattern that closely resembles contribution across each alliance condition: significantly higher contributions are expected from partners, both with respect to outsiders and to the no-alliance condition; also, although not significant, there is a slightly higher expectation of partners' contribution in grand alliances than in partial ones, and a similar pattern holds across MPCR conditions.[21]

Table 13: Effect of previous-round alliance type on actual contributions.

	(1)	(2)	(3)	(4)
	Overall	Never Join	Unconditional Join	Conditional Join
$m = 2$, non-member $(t-1)$	-1.90^{*}	$0.70^{\cdot}$	$-2.51^{\cdot}$	$2.01^{\cdot}$
	(0.85)	(0.50)	(2.59)	(1.66)
$m = 2$, member $(t-1)$	2.16^{**}	$1.34^{\cdot}$	$0.54^{\cdot}$	$-0.24^{\cdot}$
	(0.69)	(1.34)	(0.97)	(0.76)
$m = 3$ $(t-1)$	2.52^{***}	4.01^{*}	$0.66^{\cdot}$	$1.23^{\cdot}$
	(0.55)	(1.53)	(0.88)	(0.74)
$\alpha = 0.6$	$-1.54^{\cdot}$	-1.85^{**}	$-1.17^{\cdot}$	-2.78^{**}
	(0.94)	(0.54)	(1.26)	(0.93)
Round	0.22^{**}	$0.12^{\cdot}$	$0.10^{\cdot}$	$0.24^{\cdot}$
	(0.07)	(0.09)	(0.14)	(0.14)
Constant	11.17^{***}	6.30^{***}	15.41^{***}	12.65^{***}
	(1.14)	(0.81)	(1.60)	(1.29)
Observations	990	62	275	391

OLS regressions with standard errors clustered at the matching-group level. The reference category for previous-round alliance type is no alliance ($m = \emptyset$). $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

This pattern qualifies the "stable dispositions" interpretation: while joining profiles appear stable across rounds, contribution behaviour remains responsive to the cooperative history of the group. Even participants who consistently refuse to join contribute more after witnessing a successful grand alliance, suggesting that the stability of the type classification operates on the joining margin rather than the contribution margin.

[21] See regression in Table **??** in the Appendix **??** for significance levels.

Table 14: Mean beliefs by partnership condition and alpha treatment

	$m = \emptyset$	$m = 2$, outsider	$m = 2$, member		$m = 3$
			belief on partner	belief on outsider	
$\alpha = 0.6$	10.79	13.17	15.36	8.14	14.94
	(135)	(54)	(108)	(108)	(279)
$\alpha = 0.8$	11.27	13.10	14.09	8.95	16.13
	(105)	(41)	(82)	(82)	(276)
Overall	11.00	13.14	14.81	8.49	15.54
	(240)	(95)	(190)	(190)	(555)

Notes: $m = \emptyset$: belief on average others' contribution when no alliance. $m = 2$, outsider: belief of the subject excluded from a partial alliance on the average alliance members' contribution. $m = 2$, member, belief on partner: belief of a partial alliance member on the partner's contribution. $m = 2$, member, belief on outsider: belief of a partial alliance member on the outsider's contribution. $m = 3$: belief on average others' contribution when full alliance.

5 Discussion and Conclusion

This paper studies behavioral heterogeneity in the formation of a cooperative institution and its relationship with subsequent cooperation. We use a two-stage public-good game in which participants first decide whether to join a voluntary cost-sharing alliance and then choose how much to contribute to the public good. By eliciting complete conditional strategies at both stages, our design allows us to characterize heterogeneity not only in contribution behavior, as in the standard literature on heterogeneity, but also in how individuals condition their willingness to join a cooperative institution on the participation of others.

Our first finding is that joining behavior is heterogeneous but highly structured. Although the five conditional joining decisions allow for 32 possible profiles, a small number of interpretable patterns account for most observed behavior. Approximately one quarter of choices correspond to Unconditional Joining, around 40% to three forms of Conditional Joining, whereas only a small minority correspond to Never Joining. The benchmark profile is rare. Thus, the frequent formation of cost-sharing alliances does not arise because participants converge to the theoretically predicted joining strategy. Instead, it is sustained by systematic heterogeneity in willingness to initiate and respond to others' participation.

Importantly, this heterogeneity on the joining margin is systematically related to heterogeneity on the contribution margin. Unconditional Joiners contribute more than Conditional Joiners, who in turn contribute more than Never Joiners. Differences in actual contributions could partly arise mechanically because participants who are more willing to join are also more likely to become

alliance members and therefore face different contribution incentives. However, this cannot account for the entire relationship. When no alliance is formed, Unconditional, Conditional, and Never Joiners face identical institutional conditions and monetary incentives, yet Unconditional Joiners still contribute significantly more than the other two groups. Joining profiles therefore contain behavioral information about subsequent cooperation beyond the institutional position into which joining decisions select participants. This suggests that heterogeneity in institutional participation is behaviorally meaningful and systematically related to heterogeneity in cooperativeness.

These behavioral differences map onto well-known psychological distinctions: the unconditional engagement of Unconditional Joiners corresponds to a prosocial orientation, the contingent participation of Conditional Joiners to reciprocal motivation, and the persistent non-engagement of Never Joiners to an individualistic orientation. The correspondence extends beyond this broad three-way distinction. Among the conditional sub-types, the willingness to initiate versus the insistence on following reflects different thresholds of assurance before entering a cooperative arrangement. This within-category variation parallels evidence from standard public-good settings, where even among conditional cooperators, differences in beliefs about others' cooperativeness generate divergent contribution trajectories (Chaudhuri et al., 2017). That the ordering in cooperativeness across these profiles persists when no alliance forms and all participants face identical contribution incentives suggests that joining profiles capture stable cooperative dispositions rather than context-specific strategies, consistent with evidence that cooperation types in public-good games have stable psychological foundations (Volk et al., 2012).

At the same time, these behavioral profiles should not be interpreted as immutable types. The same small set of joining profiles is already present from the beginning of the experiment, and individuals tend to use the same profiles repeatedly across rounds. Yet their relative prevalence changes with experience. In particular, Unconditional Joining increases substantially over time, whereas the benchmark profile does not become more prevalent. The resulting evolution is therefore not convergence toward the equilibrium joining strategy, but a shift toward a greater willingness to participate in the cooperative institution. Moreover, contribution behavior displays substantial within-individual persistence while remaining responsive to previous alliance configurations. Taken together, these findings suggest that joining profiles may reflect underlying cooperative attitudes while also evolving with experience, rather than representing purely fixed behavioral traits. This interpretation is consistent with evidence that cooperative behavior reflects both relatively persistent individual dispositions and responses to the environment (Katuščák and Miklánek, 2023; Baader et al., 2026).

The relationship between institutional participation and cooperation is also visible within joining profiles. Participants contribute more as members of the grand alliance than as members of a partial alliance, even though the benchmark predicts maximal contributions in both cases. This

increase is observed within the same joining profiles and therefore cannot be explained solely by differences in the composition of participants across alliance sizes. The extent of others' participation thus matters for contribution behavior even after an individual has chosen to enter the institution. Joining and contribution decisions consequently appear to represent related but distinct margins of cooperative behavior: individuals differ in their willingness to participate, these differences predict how much they contribute, and contribution decisions remain responsive to the institutional environment they face.

These individual-level patterns help explain the aggregate performance of the cost-sharing institution. Alliances form in around 78% of rounds, contributions are substantially higher when participants are alliance members, and cooperation does not exhibit the decline commonly observed in repeated voluntary-contribution mechanisms. Yet these outcomes emerge despite substantial departures from the benchmark on both margins: participants join considerably more often than predicted, including in situations in which joining is not prescribed by the benchmark, while alliance members contribute less than the theoretical maximum. Sustained cooperation therefore does not require convergence toward a homogeneous benchmark strategy. Instead, the institution accommodates heterogeneous joining and contribution patterns while generating high and persistent levels of cooperation.

More broadly, our findings connect two strands of the literature that have largely examined behavioral heterogeneity and endogenous institution formation separately. Previous work has shown that heterogeneous social preferences can affect sorting and self-selection into cooperative groups and institutions. Our results highlight a complementary dimension: individuals also differ systematically in the structure of the joining decision itself—in whether they initiate cooperation, join conditionally on others' participation, or refrain from joining—and these participation profiles are systematically associated with subsequent contribution behavior. Behavioral heterogeneity therefore matters not only for behavior within cooperative institutions, but also for the process through which those institutions form.

Several limitations should qualify these conclusions. Our groups are small ($n = 3$), which restricts the number of possible alliance configurations and yields a single benchmark alliance size in both MPCR conditions. Larger groups would provide a richer environment in which to study conditional joining and would allow more complex participation profiles to emerge. Moreover, although our strategy-vector method is essential for observing complete conditional joining and contribution schedules, behavior elicited conditionally may differ from behavior following a single realized history. The similarity between our aggregate results and the direct-response results of Buso et al. (2025), together with evidence that strategy-method classifications in public-good games correspond closely to direct-response behavior (Fischbacher et al., 2012), provides some reassurance, but does not constitute a direct test of behavioral equivalence.

Future research could examine whether the joining profiles identified here generalize to environments with larger groups, alternative cost-sharing rules, or heterogeneity in endowments and returns to cooperation. An especially relevant question is whether heterogeneity in conditional participation is stable across institutions in the same way that cooperative dispositions have been shown to display persistence across contribution environments, and how changes in institutional incentives reshape the distribution of joining profiles. Distinguishing persistent cooperative dispositions from responses to particular institutional environments would help clarify when behavioral heterogeneity facilitates institution formation and when institutions themselves transform the behavioral composition of participation.

Replication files

The experimental screens, oTree package, and data and code for replicating the results of this paper are available at https://osf.io/b5dqh/. All files are licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Werner Güth would like to acknowledge the Max Planck Institute for Behavioral Economics for supporting this study.

References

Andreoni, J. (1988). Why free ride? strategies and learning in public goods experiments. Journal of Public Economics, 37(3):291–304.

Angelovski, A., Di Cagno, D., Güth, W., Marazzi, F., and Panaccione, L. (2018). Behavioral spillovers in local public good provision: An experimental study. Journal of Economic Psychology, 67:116–134.

Baader, M., Gächter, S., Lee, K., and Sefton, M. (2026). Social preferences and the variability of conditional cooperation. Economic Theory, 81(1):341–366.

Ball, S., Sarangi, S., and Upadhyay, U. (2025). Coalitions improve the coordination and provision of public goods. Journal of Public Economic Theory.

Barrett, S. (1994a). Self-enforcing international environmental agreements. Oxford Economic Papers, 46(Supplement 1):878–894.

Barrett, S. (1994b). Self-enforcing international environmental agreements. Oxford Economic Papers, 46:878–894.

Buchholz, W., Peters, W., and Ufert, A. (2018). International environmental agreements on climate protection: A binary choice model with heterogeneous agents. Journal of Economic Behavior & Organization, 154:191–205.

Buchholz, W. and Sandler, T. (2021). Global public goods: A survey. Journal of Economic Literature, 59(2):488–545.

Buso, I. M., Di Cagno, D., Ferrari, L., Larocca, V., Lorè, L., Marazzi, F., Panaccione, L., and Spadoni, L. (2021). Lab-like findings from online experiments. Journal of the Economic Science Association, 7(2):184–193.

Buso, I. M., Di Cagno, D., Güth, W., and Spadoni, L. (2025). Voluntary partnerships for equally sharing contribution costs: Theoretical aspects and experimental evidence. The B.E. Journal of Theoretical Economics, 25(1):1–41.

Carraro, C. and Siniscalco, D. (1993a). Strategies for the international protection of the environment. Journal of Public Economics, 52(3):309–328.

Carraro, C. and Siniscalco, D. (1993b). Strategies for the international protection of the environment. Journal of Public Economics, 52(3):309–328.

Casari, M. and Tavoni, A. (2024). Climate clubs in the laboratory. Journal of Behavioral and Experimental Economics, 110:102211.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. Experimental Economics, 14(1):47–83.

Chaudhuri, A., Paichayontvijit, T., and Smith, A. (2017). Belief heterogeneity and contributions decay among conditional cooperators in public goods games. Journal of Economic Psychology, 58:15–30.

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

Dannenberg, A. and Gallier, C. (2020). The choice of institutions to solve cooperation problems: A survey of experimental research. Experimental Economics, 23(3):716–749.

Dannenberg, A., Lange, A., and Sturm, B. (2014). Participation and commitment in voluntary coalitions to provide public goods. Economica, 81(322):257–275.

Dannenberg, A., Sturm, B., and Vogt, C. (2010). Do equity preferences matter for climate negotiators? an experimental investigation. Environmental and Resource Economics, 47(1):91–109.

d'Aspremont, C., Jacquemin, A., Gabszewicz, J. J., and Weymark, J. A. (1983). On the stability of collusive price leadership. Canadian Journal of Economics, 16(1):17–25.

Engel, C. and Rockenbach, B. (2024). What makes cooperation precarious? Journal of Economic Psychology, 101:102712.

Falkinger, J. (1996). Efficient private provision of public goods by rewarding deviations from average. Journal of Public Economics, 62(3):413–422.

Falkinger, J., Fehr, E., Gächter, S., and Winter-Ebmer, R. (2000). A simple mechanism for the efficient provision of public goods: Experimental evidence. American Economic Review, 90(1):247–264.

Fischbacher, U. and Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. American Economic Review, 100(1):541–556.

Fischbacher, U., Gächter, S., and Fehr, E. (2001). Are people conditionally cooperative? evidence from a public goods experiment. Economics Letters, 71(3):397–404.

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

Gächter, S. and Marino Fages, D. (2023). Using the strategy method and elicited beliefs to explain group size and mpcr effects in public good experiments. IZA Discussion Paper 16605, IZA Institute of Labor Economics.

Gächter, S. and Thöni, C. (2005). Social learning and voluntary cooperation among like-minded people. Journal of the European Economic Association, 3(2–3):303–314.

Gneezy, U., Haruvy, E., and Yafe, H. (2004). The inefficiency of splitting the bill. Economic Journal, 114(495):265–280.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. Journal of the Economic Science Association, 1(1):114–125.

Guido, A., Robbett, A., and Romaniuc, R. (2019). Group formation and cooperation in social dilemmas: A survey and meta-analytic evidence. Journal of Economic Behavior & Organization, 159:192–209.

Gürerk, Ö. (2013). Social learning increases the acceptance and the efficiency of punishment institutions in social dilemmas. Journal of Economic Psychology, 34:229–239.

Isaac, R. M. and Walker, J. M. (1988). Group size effects in public goods provision: The voluntary contributions mechanism. Quarterly Journal of Economics, 103(1):179–199.

Isaac, R. M., Walker, J. M., and Thomas, S. H. (1984). Divergent evidence on free riding: An experimental examination of possible explanations. Public Choice, 43(2):113–149.

Katuščák, P. and Miklánek, T. (2023). What drives conditional cooperation in public good games? Experimental Economics, 26(2):435–467.

Kosfeld, M., Okada, A., and Riedl, A. (2009). Institution formation in public goods games. American Economic Review, 99(4):1335–1355.

Lange, A. and Vogt, C. (2003). Cooperation in international environmental negotiations due to a preference for equity. Journal of Public Economics, 87(9–10):2049–2067.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In Kagel, J. H. and Roth, A. E., editors, The Handbook of Experimental Economics, pages 111–194. Princeton University Press, Princeton, NJ.

Okada, A. and Riedl, A. (2005). Inefficiency and social exclusion in a coalition formation game: Experimental evidence. Games and Economic Behavior, 50(2):278–311.

Schopf, M. (2024). Self-enforcing international environmental agreements and altruistic preferences. Environmental and Resource Economics, 87:2309–2359.

Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. International Journal of Game Theory, 4(1):25–55.

Selten, R. and Güth, W. (1982). Equilibrium point selection in a class of market entry games. In Deistler, M., Fürst, E., and Schwödiauer, G., editors, Games, Economic Dynamics, and Time Series Analysis, pages 101–116. Physica-Verlag.

Thöni, C. and Volk, S. (2018). Conditional cooperation: Review and refinement. Economics Letters, 171:37–40.

van der Pol, T., Weikard, H.-P., and van Ierland, E. (2012). Can altruism stabilise international climate agreements? Ecological Economics, 81:112–120.

Vogt, C. (2016). Climate coalition formation when players are heterogeneous and inequality averse. Environmental and Resource Economics, 65(1):33–59.

Volk, S., Thöni, C., and Ruigrok, W. (2012). Temporal stability and psychological foundations of cooperation preferences. Journal of Economic Behavior & Organization, 81(2):664–676.

Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. Experimental Economics, 6(3):299–310.