Journal of Economic Behavior and Organization 233 (2025) 106943

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Institution formation in weakest-link games

Alejandro Caparrós [a,b,*], Esther Blanco [c,d], Michael Finus [e,f]

[a] Department of Economics, University of Durham, Mill Hill Lane, Durham, UK
[b] Spanish National Research Council (CSIC), Spain
[c] Department of Public Finance, University of Innsbruck, Austria
[d] The Ostrom Workshop, Indiana University, USA
[e] Department of Economics, University of Graz, Austria
[f] Department of Economics, University of Bath, UK

ARTICLE INFO

JEL classifications:
C72
C91
C92
H41

Keywords:
Weakest-link game
Minimum-effort game
Coalition formation
Experimental economics
Quantal response equilibrium
Agent quantal response equilibrium

ABSTRACT

We study the role of endogenous formation of institutions in overcoming coordination failures in weakest-link games with fixed neighborhoods. In our setting, institutions are weak and only form and make decisions by unanimity. Experimental results show that such institutions are formed and mitigate the coordination problem, raising equilibrium provision levels, but falling short of providing Pareto-optimal contributions. Given the multiplicity of Nash equilibria in weakest-link games, we consider several equilibrium refinements that allow for (small) errors by individuals. Without institutions, risk dominance and the Quantal Response Equilibrium (QRE) with (almost) perfectly rational agents select the worst equilibrium, while all equilibria are trembling-hand perfect and proper. With the possibility of forming an institution, all these concepts predict the Pareto-optimal equilibrium as the unique outcome. As we do not observe this outcome in our experimental results, only the Agent QRE model with bounded rationality can explain our data.

1. Introduction

The theoretical and experimental literature on weakest-link games (also known as minimum-effort games) is abundant, and there are many real-world examples where the benefits obtained from a good or service depend on the lowest contribution among all agents (see for example Vicary (1990), Sandler (2004), Caparrós and Finus (2020)). Hirshleifer (1983) introduced the weakest-link concept illustrating it with an island, Anarchia, which is perfectly circular and where each citizen owns a wedge-shaped slice from the center to the sea. Dikes protect Anarchia from storms that threaten to flood the land. Since Anarchia has no government, everyone decides on the height of the dike built on his piece of land. The benefit of protection against flooding will depend on the lowest section of the dike. Vicary (1990), Sandler (2004), Caparrós and Finus (2020) and others provide many real-life examples that share key features of Anarchia. These examples range from vaccination and other measures against the spread of diseases and pandemics, security measures at airports, and cyber security standards to the performance of teams in firms that may depend on the lowest effort.

As in other coordination games, there are multiple Nash equilibria that can be Pareto-ranked in weakest-link games. Thus, one should expect coordination to be trivial. However, many experiments have confirmed that this is not the case. The common pattern observed is a downward-sloping trend of average and minimum contribution levels as the game proceeds, resulting in low contribution levels (Van Huyck et al. 1990; Cooper et al. 1990; Cachon and Camerer, 1996; Feri et al. 2010; see Devetag and Ortmann (2007) for a

* Corresponding author at: Department of Economics, University of Durham, Mill Hill Lane, Durham, UK.
E-mail address: alejandro.caparros@durham.ac.uk (A. Caparrós).

https://doi.org/10.1016/j.jebo.2025.106943
Received 30 June 2024; Received in revised form 15 January 2025; Accepted 12 February 2025
Available online 9 March 2025
0167-2681/

review).

As discussed in our section "Related Literature" below, solutions to this coordination problem suggested in the literature include outside money or are coercive in nature. However, Anarchia has no government, which excludes most of these options. We consider what is probably the simplest *voluntary* institution for which theoretical results allow us to conjecture that it will solve the coordination problem (details below). Individuals announce whether they want to join or remain outside of the institution. Then, all individuals must confirm their membership after observing how many other players are willing to participate, otherwise the institution will not form. Members of the institution choose their effort levels by unanimity − implementing the minimum proposed by any member of the institution − whereby no member can be forced to contribute more than she is prepared to do. Payoffs derive from the minimum effort of the whole group. That is, payoffs depend on the minimum effort chosen by the members of the institution and the effort levels of all non-members, provided that not all individuals in a group are members of the institution. No group member can be excluded from affecting the benefits of all players, which we refer to as a setting with a fixed neighborhood.

In a nutshell, we analyze the endogenous formation of institutions in the context of the weakest-link game, excluding sanctions as an enforcement measure. Participation is voluntary and decisions within the institution are taken by unanimity. Hence, on the one hand, this is a weak institution in terms of enforcement power, on the other hand, it is an attractive institution, as players do not have to give up their sovereignty. The term "institution" that we use is a synonym for a (sub)-group of individuals who voluntarily accept a set of rules that facilitate coordination (it is also a synonym for the term "coalition" frequently used in the literature). We could also consider an institution that imposes exogenously a rule whereby the minimum proposed by all players (or by a subset of players) is implemented, as this would probably also increase efficiency. However, only an entity with enforcement power could impose such a rule, whereas we are interested in voluntary and endogenous solutions to the coordination problem faced by the citizens of Anarchia.

Loosely speaking, the institution that we analyze reduces or eliminates the strategic risk inherent in weakest-link games. In an institution with full membership, the risk is entirely eliminated. Among all members, the lowest effort proposal is implemented. Thus, individuals do not have to fear that other players choose lower contribution levels than they do, which would be costly to them. In an institution with partial membership, at least the strategic risk of a costly choice is reduced, as fewer "independent players" outside the institution can choose a Pareto-inferior and lower contribution level than the institution.

The specific weakest-link games which we consider are the Van Huyck et al. (1990) weakest-link game, henceforth abbreviated VHBB, and the "riskier" weakest-link game, first introduced by Feri et al. (2010), henceforth abbreviated FIS. Although we follow their abstract wording, their games are particularly close to Hirshleifer's scenario, simply substituting the "number you choose" with the "height of your dike". In VHBB, agents have incentives to coordinate on high contribution levels, which result in high individual and group welfare, but also face strong strategic uncertainty, as one single player suffices to cause substantial losses to all other group members if her contributions are below those of others. In FIS, the failure to coordinate entails even higher losses.[1] Throughout the exposition, we focus mainly on the results of the VHBB game, as our goal is to discuss the role of institution formation in this well-known weakest-link game. We present the results using the payoffs in FIS as a robustness check.

Our experimental results show that institutions are voluntarily formed and enhance coordination and welfare. In particular, institutions are capable of (i) stopping the sequence of downward adjustment towards the worst equilibria over time, and (ii) increasing coordination levels compared to situations where institutions are not possible, (iii) even though provision levels do not reach Pareto-optimal values. Thus, we show that institution formation can alleviate coordination failures, but also that it does not solve the coordination problem completely.

As mentioned above, in weakest-link games, there are many Nash equilibria. Pareto-dominance selects the best equilibrium, which is not in line with the experimental evidence. In the setting with institution formation, the number of Nash equilibria is even larger, and Pareto-dominance as a simple selection device still leaves many equilibria. Therefore, we consider several equilibrium refinements that allow for errors by individuals (mistakes), including risk-dominance (Harsanyi and Selten, 1988), trembling-hand perfect equilibrium (Selten, 1975), proper equilibrium (Myerson, 1978) and the Quantal Response Equilibrium (QRE), respectively Agent QRE (AQRE) (McKelvey and Palfrey, 1995, 1998). Without institution formation, either all Nash equilibria survive, or the worst equilibrium is selected. In contrast, with institution formation, only the grand institution where all individuals play the Pareto-dominant equilibrium survives.

These results lead us to conjecture that institution formation significantly improves coordination. However, as we do not observe perfect coordination in our experiment, only the (A)QRE, when defined for agents with bounded rationality, is compatible with our empirical result that institution formation alleviates but does not completely solve the problem of coordination failure.

2. Related literature

Experimental results have established that weakest-link games follow the common pattern observed in coordination experiments with Pareto-ranked equilibria: a downward-sloping trend of average and minimum contribution levels as the game proceeds resulting in low contribution levels (Van Huyck et al. 1990; Cooper et al. 1990; Cachon and Camerer, 1996; Feri et al. 2010; see Devetag and Ortmann (2007) for a review). A broad body of literature has explored interventions to enhance effort levels, entailing higher efficiency. A first set of studies have analyzed the role of changes in marginal incentives, either through introducing financial incentives (Brandts and Cooper, 2006; Hamman et al., 2007) or reducing effort costs (Goeree and Holt, 2005; Brandts et al., 2007). Other studies

[1] FIS introduce their payoff table to construct an environment which reinforces the attraction of choosing the worst outcome because of high risk.

have investigated inexpensive but coercive policies. They show that efficiency increases substantially when people can be excluded from the group, so that they no longer determine the group's success (Croson et al., 2015; Riedl et al., 2016). Alternative policies include imposing costs to enter the game (Cachon and Camerer, 1996), allowing for advice by players who have played in previous rounds (Chaudhuri et al., 2009), time pressure in decision-making (Feri et al., 2010), costly pre-play communication and auctioning the right to participate (Van Huyck et al., 1993; Devetag, 2005), costly communication combined with subsidies (Kriss et al., 2016), costly communication for groups that merge (Fehr, 2017), costless pre-play communication (Blume and Ortmann, 2007), real-time monitoring (Deck and Nikiforakis, 2012) and communication with incremental commitments (Avoyan and Ramos, 2023).

Our focus in this paper is on *voluntary* institution formation. To the best of our knowledge, no research is available on the role of institution formation, neither in weakest-link games nor in coordination games generally. Voluntary institution formation has advantages over economic policy instruments and coercive policies, and can even be the only realistically feasible option in some settings. Policies changing incentives are costly and may take a toll on scarce public funding or displace private entities' expenditures. Mandatory and coercive policies may not be feasible or politically palatable, like excluding players from the game and the auctioning of rights to participate. Sanctions may be lacking, not legally enforceable, expensive and/or not credible. For instance, firms may want to refrain from top-down policies imposed on workers, as this may demotivate their performance (Gneezy et al., 2011). Similarly, the social costs of ostracism by peers may be prohibitively high in some groups and societies, e.g., leading to a high employee turnover rate in firms or the breakdown of family businesses in which social ties are traditionally strong. Also, enforcement of international law and sanctions often do not work in international relations because they are not credible.

A closely related body of literature has studied institution formation in voluntary contribution mechanism (VCM) public good games, or in "summation" public good games using the terminology introduced by Hirschleifer (1983).[2] See Kosfeld et al. (2009) for a prominent example and Dannenberg and Gallier (2020) for a survey of this literature. However, as highlighted by the numerous theoretical studies that have analyzed weakest-link games (Hirshleifer, 1983; Cornes, 1993; Caparros and Finus, 2020), the incentives are completely different from those of the VCM (summation) public good game. In fact, the games that we study have much more in common with other coordination games (Devetag and Ortmann, 2007).

The literature on communication and signaling games has shown that the opportunity to communicate before decisions are taken may increase contributions to the public good (e.g., Isaac and Walker, 1988; Palfrey and Rosenthal, 1991; Bochet et al., 2006; Bochet and Putterman, 2009; Koukoumelis et al., 2012; Oprea et al., 2014). In our institution, members propose a number, and the minimum is implemented as the contribution of all the members of the institution. However, as the minimum number is implemented automatically, no decisions are taken by the players after they are informed about the minimum number implemented by the institution. In addition, and unlike in the literature just mentioned, players are only informed about the minimum number suggested by all members of the institution, with no individual feedback on the suggestions by each institution member.

Precedents to our use of the QRE to analyze weakest-link games can be found in Goeree and Holt (2005), who used the QRE analysis developed in Anderson et al. (2001) to study experimental data of a weakest-link game with a continuous action space and a maximum of three players. Croson et al. (2015) also used payoff dominance, risk dominance, and the QRE as refinement tools when analyzing different types of weakest-link games with four players. To our knowledge, no previous paper has employed the AQRE to analyze a multi-stage weakest-link game.

The rest of the paper is organized as follows. Section 3 describes the two weakest-link public good games and their associated payoffs, as well as the setting without and with institution formation. Section 4 outlines the details of our experiment and Section 5 presents and discusses our results. Section 6 offers a theoretical analysis of the weakest-link game, with and without institution formation, using the different equilibrium concepts mentioned above, and presents our estimations for the (A)QRE. Section 7 concludes.

3. Weakest-link games

We consider first decision settings in which individuals make decisions without the opportunity to form an institution, and then settings where there is the opportunity to join an institution in which decisions are coordinated. In this section, we focus on Nash and Subgame Perfect Equilibria in pure strategies and consider extensions later.

3.1. No institution formation

We follow the mainstream of the experimental literature and consider a linear payoff function for individual i in the weakest-link game:

$$\Pi_i^V(e_i, e_{-i}) = be^{\min} - ce_i + K \text{ with } e^{\min} = \min\{e_1,\ e_2,\ \dots,\ e_n\}, \tag{1}$$

where n is the number of players; K is a constant scale parameter; b is a benefit parameter and c a cost parameter, with $b > 0$ and $c > 0$; e_i is the contribution level of individual i and e_{-i} denotes the vector of contribution levels of all other players except player i; $e^{\min}$ is the minimum contribution over all players. For positive Nash equilibrium provision levels, we need to assume $b > c$. Following VHBB, we

[2] Hirschleifer (1983) introduced the weakest-link game as a particular type of public good game, proposing a classification into "best-shot", "weakest-link" and "summation" public goods (the latter "aggregation technology" in Hirschleifer's terminology is the standard public good game where the benefit is a function of the total provision level).

assume a discrete action space with 7 provision levels, $e_i = \{1, 2, 3, 4, 5, 6, 7\}$ and let the number of players be $n = 8$. Moreover, as VHBB, we assume $b = 20$, $c = 10$ and $K = 60$. This results in the canonical payoff matrix introduced by VHBB which is presented in Table 1. In their terminology, which we henceforth use, choosing a contribution level means choosing a number.

In Table 1, a player can choose any number between 1 and 7, as listed in the first column. The smallest number chosen by all players is listed in the first row. The payoff a player obtains from choosing some number depends on the smallest number chosen among all players, i.e., the minimum number. For instance, if a player chooses the highest number 7, she will earn 130 if all players choose 7. However, if the lowest number chosen by others is 6, she will earn only 110. Other combinations have a similar interpretation.

As detailed in VHBB, all entries in the main diagonal are Nash equilibria, and all players choosing 7 is the Pareto-dominant Nash equilibrium in this game.

In addition to the classical VHBB payoffs, to check the robustness of our results, we investigate the weakest-link game introduced by FIS, which is displayed in Table 2. The payoff for all $e_{-i}^{\min} < e_i$ is zero and therefore much lower than in Table 1. The corresponding payoff function is a piecewise function:

$$\Pi_i^F(e_i, e_{-i}) = be^{\min} - ce_i + K \text{ if } e_i = e^{\min} \text{ and } \Pi_i^F(e_i, e_{-i}) = 0 \textit{ otherwise} \tag{2}$$

Compared to Table 1, coordination failure entails a larger loss in Table 2. For this reason, one could view Table 2 as a "riskier weakest-link game" than Table 1. Except for this difference, the parameter values and the action space are the same as in Table 1. The predictions for Table 2 are the same as in Table 1: there are 7 Nash equilibria along the diagonal and all players choosing number 7 is the Pareto-dominant Nash equilibrium.

In both tables, applying the maximin criterion by choosing the strategy that maximizes the minimum payoff would induce players to choose number 1, as it guarantees the largest payoff in the worst possible case. However, this yields the least efficient equilibrium. FIS point out that Table 2 keeps the property of Pareto-ranked equilibria, but reinforces the attraction of the maximin criterion as a selection device: any number greater than 1 can lead to a payoff of zero. We will discuss alternative equilibrium selection criteria in Section 6.

3.2. *Endogenous institution formation*

We consider a simple model of endogenous institution formation in a three-stage game.[3] In stage 1, all individuals in a group decide whether to join an institution S, $S \subseteq N$. In stage 2, the decision of stage 1 is revealed to all individuals. All individuals who announced joining the institution are asked to confirm their membership. If and only if all players in S confirm their membership, the institution will be established. Otherwise, the institution is dissolved. In stage 3, all s members of institution S ($s \leq n$) simultaneously make a proposal e_i^P for a number. Among all s proposals, the minimum e_i^P is adopted for the institution, $e_S = \min\{e_1^P, \ldots, e_s^P\}$. Note that players inside the institution do not see the individual proposals made by other members; they are only informed about the smallest proposal, which is automatically implemented. This simplifies the analysis: players cannot use their proposal as a communication device. Also at stage 3, and simultaneously, all outsiders $j \notin S$ freely choose their individual number e_j without knowing the choice of the institution. The outcome of stage 3 is $e^{\min} = \min\{e_S, e_j, e_k\}, j, k \notin S$. Thus, neither members nor non-members have an informational advantage in stage 3.

In our institution formation game, joining the institution entails accepting the rules that govern it. In addition, these rules imply a high degree of consensus. In terms of membership, this is immediately evident. In terms of the number that is implemented by the institution, no player can be overruled: all players announce a number and the smallest proposal is implemented. Thus, no player can be forced to implement a higher number that is not in accordance with his or her preferences.[4]

We expect that these rules will make it attractive to join the institution, addressing the concern of individuals of being overruled by other members, and allowing us to conjecture that this institution will alleviate the coordination problem. Nevertheless, as long as an institution does not comprise full membership, the institution cannot control the provision level by non-members. The institution's members face the risk that outsiders may not match any provision level e_S above 1. However, as long as $s \geq 2$, the risk for the members of the institution caused by any outsider $j \notin S$ choosing $e_j < e_s$ is lower than without the institution, as there are fewer "independent players". By symmetry, also the risk for every outsider j that any other player chooses a lower provision level $e_{-j} < e_j$ is lower than without the institution. If $s = n$, there is no risk left.

Solving the game via backward induction, we start analyzing stage 3. All numbers between 1 and 7 are Nash equilibria in the final third stage. In stage 2, institutions of any size s may be confirmed in equilibrium. And at stage 1, institutions of any size s may be initiated in equilibrium. For instance, in the grand institution, if all players choose number 3, the best response of player i is to match this number. Alternatively, consider an institution that comprises 3 members where all players (members and non-members) choose number 7, which is also a Nash equilibrium. Technically, even Pareto-dominance does not single out one equilibrium with institution

[3] For an overview of institution formation games, commonly called coalition formation games, see for instance Bloch (1997) and Yi (2003). For a theoretical discussion of coalition formation in the weakest-link game, see Caparrós and Finus (2020).

[4] In terms of membership, this resembles features of the exclusive membership Γ-Game of Hart and Kurz (1983) or the sequential move unanimity game by Bloch (1995), as discussed in Finus and Rundshagen (2009). In terms of the choice of the team number, this resembles the smallest common denominator proposal in bargaining games. Moulin (1994) has shown that this "conservative mechanism" is strategy-proof and leads to unbiased proposals.

Table 1
VHBB payoffs.

		Smallest number chosen by any individual in your group (including yourself)						
		7	**6**	**5**	**4**	**3**	**2**	**1**
Number you choose	**7**	130	110	90	70	50	30	10
	6		120	100	80	60	40	20
	5			110	90	70	50	30
	4				100	80	60	40
	3					90	70	50
	2						80	60
	1							70

Source: Van Huyck et al. (1990), Payoff Table A, p. 232, all entries multiplied by 100.

Table 2
FIS payoffs.

		Smallest number chosen by any individual in your group (including yourself)						
		7	**6**	**5**	**4**	**3**	**2**	**1**
Number you choose	**7**	130	0	0	0	0	0	0
	6		120	0	0	0	0	0
	5			110	0	0	0	0
	4				100	0	0	0
	3					90	0	0
	2						80	0
	1							70

Source: Feri et al. (2010), Table 1, Panel B, p. 1895.

formation, as number 7 may be played by all players in equilibrium regardless of the institution's size.[5] This starkly contrasts with the results in the literature on institution formation and VCM (summation) public good games surveyed in Dannenberg and Gallier (2020), where once the institution is introduced, the equilibrium (or equilibria in some games) implies a Pareto-improvement.

4. Experimental design

Subjects participated in one of four alternative treatments in a 2 × 2 between-subjects design, varying the possibility of forming institutions. This is summarized in Table 3. For the VHBB payoffs (Table 1), this gives rise to the treatments *Baseline* and *Baseline-Institution* and for FIS payoffs (Table 2) this gives rise to the treatments *Risk* and *Risk-Institution*. The experiment comprised 20 rounds, which were divided into two parts (Part 1: rounds 1–10 and Part 2: rounds 11–20). Subjects learned the details of Part 1 at the beginning, and, after the completion of Part 1, they were introduced to Part 2. In the treatments *Baseline* and *Risk*, all subjects played the weakest-link game with the payoffs in Tables 1 and 2, respectively, throughout Part 1 and Part 2.[6] In contrast, in the treatments *Baseline-Institution* and *Risk-Institution*, subjects started by playing 10 rounds without the possibility of institution formation, followed by 10 additional rounds where subjects had the possibility to form an institution. Including Part 1 without institution formation was important because we are interested in how institutional changes affect behavior in a cohort, given that there is a history where subjects have experienced the coordination problem before having the option to form an institution. In addition, Part 1 provides an initial measure of coordination failure in the absence of institution formation.

Instructions were read aloud to ensure that the details of the experiment were public information. The language used in the experiment was neutral. Subjects were randomly and anonymously assigned to groups of 8 people. Groups remained fixed for the duration of the experiment, and this was common information. Before making decisions in Part 1 and Part 2 in all treatments, subjects answered quizzes to check their understanding of the game. See Section II in the Supplementary Material for details.

In each round without institutions, subjects were confronted with the sole decision to simultaneously choose a number out of 1, 2, 3, 4, 5, 6, or 7. At the end of each round, each subject saw an information screen showing the number she had chosen, the minimum number chosen in her group, her earnings for that given round, and her cumulative earnings.

In rounds with institution formation, instructions were described in 4 phases for clarity. In phase 1, players had to choose their

[5] The maximin criterium is also of limited use to reduce the number of equilibria, unlike in the case without institution formation. Under the maximin criterium, playing 1 (with a payoff of 70) is the unique equilibrium as long as there is one player outside the institution. In the grand coalition, there are multiple equilibria, as the worst payoff is 70, irrespective of the number chosen. There are also multiple equilibria at stages 1 and 2. The worst expected payoff of joining the institution and confirming it is 70, and the same payoff is the worst possible outcome of not joining or not confirming.

[6] For consistency of treatments with institution formation, the instructions for treatments without institution formation also included a Part 1 and a Part 2. Subjects were told at the end of Part 1 that the game would continue for 10 more rounds in Part 2.

Table 3
Summary of experimental sessions.

Treatment	Payoff Function	Institution, Part 1 (Round 1–10)	Institution Part 2 (Round 11–20)	Number of subjects	Number of groups
Baseline	Van Huyck et al. (1990)	No	No	64	8
Baseline-Institution	Van Huyck et al. (1990)	No	Yes	72	9
Risk	Feri et al. (2010)	No	No	72	9
Risk-Institution	Feri et al. (2010)	No	Yes	64	8
Total				272	17

membership. At the beginning of phase 2, all individuals were informed about how many individuals had decided to join the institution, and members were asked to confirm their membership at phase 2. Members and outsiders where then informed about whether the institution had been formed (confirmed) and its size. Then stage 3 of the game was presented as one phase where the institution determined its minimum number (phase 3) and one phase where outsiders decided their number between 1 and 7 independently (phase 4). As no information was provided to the subjects between these phases and decisions were taken simultaneously, phases 3 and 4 constitute one single stage in modeling terms. As discussed in Section 3.2, outsiders do not learn the effort level adopted by the institution before making their decision, and members do not learn the effort level adopted by outsiders.

At the end of each round, all subjects saw an information screen summarizing key decisions taken and outcomes of the three stages: their own chosen (if a non-member) or proposed (if a member) number; the minimum number chosen in her group (minimum of the number implemented by the institution and the numbers chosen by non-members); her earnings for this round and her cumulative earnings; and whether an institution had formed and, if it had formed, the size of the institution and the number chosen by it.

All sessions were conducted at the University of Innsbruck EconLab. The experiments used z-Tree (Fischbacher, 2007) for programming and ORSEE (Greiner, 2015) for subject recruitment.[7] Sessions lasted for about an hour, and participants earned on average 13.28 Euros.

5. Experimental results

We first analyze whether institutions are initiated and established. We then analyze whether institution formation increases provision levels, by comparing average numbers and minimum numbers with and without institution formation. Finally, we analyze whether institution formation impacts the evolution of provision levels over time by comparing the trend of average and minimum numbers.

5.1. Institution formation

This section analyzes the sizes of the institutions initiated and eventually established by confirmation of all members. We first focus on those instances for which at least two subjects selected membership in an institution in phase 1 (*initiated* institution) and analyze whether those subjects who choose to become a member of an institution unanimously confirm their membership in phase 2 (*confirmed* institutions). Tables 4a and Table 4b provide an overview of the results, and Fig. 2 in Section 5.2 provides further details on the provision levels depending on the institution size.

In the treatment *Baseline-Institution,* in 90 cases membership in an institution was chosen, which implies that an institution was initiated in every round and in every group (100 percent). Of those 90 initiated institutions, in 80 membership was confirmed, which gives a confirmation rate of 88.89 percent. In the treatment *Risk-Institution*, in 78 out of 80 possible cases an institution was initiated (97.5 percent) of which 48 were confirmed, corresponding to a confirmation rate of 61.54 percent. While the rate of institutions initiated is not significantly different between treatments (Mann-Whitney U Test, $p = 0.1325$), the confirmation rate is statistically higher in *Baseline-Institution* than in *Risk-Institution* (Mann-Whitney U Test, $p = 0.0000$).

For both, *Baseline-Institution* and *Risk-Institution*, the majority of confirmed institutions contain all eight subjects in a group. Confirmed institutions with full membership constitute 51.25 percent in *Baseline-Institution* of the total confirmed institutions (i.e., 41 out of 80) and 93.75 percent in *Risk-Institution* (i.e., 45 out of 48). This is because in *Baseline-Institution* the size of confirmed institutions is diverse, while in *Risk-Institution* there are only three instances of confirmed institutions without full membership. For the institutions with full membership, confirmation rates are similar in *Baseline-Institution* (93.18 percent) and *Risk-Institution* (97.83 percent). However, the treatments differ in the confirmation rate for institutions with less than eight members. In *Baseline-Institution* the confirmation rate of initiated institutions without full membership is above 65 percent in all cases, while the corresponding confirmation rate in *Risk-Institution* is below 20 percent. Thus, members in *Baseline-Institution* are more likely to accept the risk that not all individuals in a group join the institution, whereas members in *Risk-Institution* prefer to withdraw their membership if they realize that

[7] We obtained ethical clearance from the Board of Ethical Questions in Science of the University of Innsbruck. Certificate of good standing 44/2020.

Table 4a
Initiated and confirmed institutions across the treatment "Baseline-Institution".

	Initiated institutions	Confirmed institutions	Confirmation rate (%)	Percentage of confirmed institutions by size
Total	90	80	88.89	100
Two Members	1	1	100	1.25
Three Members	8	8	100	10
Four Members	3	2	66.67	2.5
Five Members	5	4	80	5
Six Members	17	14	82.35	17.5
Seven Members	12	10	83.33	12.5
Eight Members	44	41	93.18	51.25

Table 4b
Initiated and confirmed institutions across the treatment "Risk-Institution".

	Initiated institutions	Confirmed institutions	Confirmation rate (%)	Percentage of confirmed institutions by size
Total	78	48	61.54	100
Two Members	3	0	0	0
Three Members	2	0	0	0
Four Members	2	0	0	0
Five Members	8	1	12.5	2.08
Six Members	6	0	0	0
Seven Members	11	2	18.18	4.17
Eight Members	46	45	97.83	93.75

Note: The table shows absolute and relative numbers of initiated and confirmed institutions broken down by institution size. In the treatment *Baseline-Institution* there are 9 groups observed in 10 rounds in Part 2 such that there is a maximum of 90 instances where institutions can be initiated and confirmed. In the treatment *Risk-Institution* there are 8 groups observed in 10 periods, making a maximum of 80 instances where institutions can be initiated and confirmed. The confirmation rate for each institution size is the percentage of successfully confirmed institutions relative to the number of initiated institutions within each class. The percentage of confirmed institutions by size measures the number of confirmed institutions for a given size relative to the total number of confirmed institutions.

not all subjects in their group have chosen to become members of the institution. Indeed, the difference between the average size of confirmed institutions in *Baseline-Institution* (6.7 members (SD=1.72)) and *Risk-Institution* (7.9 (SD=0.47)) is statistically significant (Mann-Whitney U Test, $p = 0.0000$). Our observations are summarized in Result 1 below.

Result 1. *Individuals use the opportunity to form institutions: they initiate and confirm institutions. The confirmation rate of larger institutions is higher than that of smaller institutions. For the treatment Risk-Institution, individuals almost exclusively confirm only the institution with full membership, whereas smaller institutions are confirmed for the treatment Baseline-Institution.*

Result 1 is further substantiated by an analysis of the evolution of institution formation over rounds, illustrated in Fig. 2 in Section 5.2. A Spearman's rank order correlation shows a significant negative correlation between the number of confirmed institutions and rounds in the treatment *Baseline-Institution* (ρ=−0.69, $p = 0.0283$) and no significant correlation in *Risk-Institution* (ρ=0.36, $p = 0.3059$). The negative correlation in *Baseline-Institution* is even stronger for the number of institutions with full membership over rounds (Spearman rank order correlation: ρ=−0.76, $p = 0.0102$), whereas it is significantly positive in *Risk-Institution* (Spearman rank order correlation: ρ=0.87, $p = 0.0009$). The number of confirmed institutions with fewer than eight members over rounds is constant in *Baseline-Institution* (Spearman rank order correlation: ρ=−0.08, $p = 0.8307$) and negative in *Risk-Institution* (Spearman rank order correlation: ρ=−0.68, $p = 0.0294$). Thus, the decreasing number of confirmed institutions over time in *Baseline-Institution* originates from a downward trend in the number of confirmed institutions with full membership. In *Risk-Institution*, the constant relationship between the number of confirmed institutions and rounds masks two opposing trends: an increasing number of confirmed institutions with full membership over rounds and a decreasing number of confirmed institutions with fewer than eight members. Hence, it appears that in the riskier weakest-link game, individuals understand over time that the confirmation of the grand institution is a reasonable strategy to reduce their risk, but shy away from the confirmation of smaller institutions, as the risk of a loss is perceived to be too high.

5.2. Provision levels

In order to address the impact of the opportunity to form institutions, in terms of average contributions and in terms of equilibrium provision levels, we investigate differences in two main output variables between treatments with institution formation, namely *average numbers* and *minimum numbers*. All analyses are conducted at the group level, where each group of 8 subjects is treated as an independent unit of observation. The minimum number in a round measures the group output level and influences the payoff that subjects receive in a particular round. The average number in a round is a measure of the numbers chosen by all subjects in this round. By definition, the minimum number is weakly smaller than the average group number, only for groups with perfect coordination do the two numbers coincide. The average and minimum numbers across all groups in each round are shown in Fig. 1 and summarized in Table 5; the details for each group are displayed in Figs. A1–A8 in Section I in the Supplementary Material.[8]

As shown in Table 5, average and minimum numbers in Part 2 are significantly higher in the treatment *Baseline-Institution* than in *Baseline*.[9] Although both average numbers and minimum numbers are higher in *Risk-Institution* than in *Risk,* these differences are not significant.[10] Thus, there is some evidence of the positive impact of institution formation on provision levels. However, even with institution formation, minimum numbers fall short of the Pareto-optimal level 7. It is also interesting to observe that institutions with full membership perform equally well in both treatments, exhibiting a minimum number of 5.63 (SD 1.59) in *Baseline-Institution* and 5.51 (SD 2.53) in *Risk-Institution* (Mann-Whitney U test $p = 0.0656$). This is not the case for institutions with smaller membership (see Fig. 2). In *Baseline-Institution*, institutions of different sizes are implemented throughout rounds, with generally high average and minimum numbers. In contrast, in *Risk-Institution*, after round 13, only the institution with full membership is sustained, which in round 14 supports a maximum effort by all and an average minimum number of 6 between rounds 15 and 20.

Result 2. *Average numbers and minimum numbers tend to be higher when institution formation is possible. Even with institution formation, provision levels do not reach the Pareto-optimal number.*

That is, institution formation alleviates the coordination problem but does not trivially solve it.

5.3. Trends

This section analyzes the effect of institution formation on the trends of our output variables average and minimum number throughout rounds within Part 1 and Part 2, using the Jonckheere-Terpstra test for ordered alternatives.[11] The impression of a decreasing trend of average numbers throughout rounds 1–10, as observed in Fig. 1– Panel A, is supported by this test. There is a statistically significant trend of lower average numbers in Part 1 over rounds with no institution formation in all treatments ($p = 0.0000$ in *Baseline; p* = 0.0000 in *Baseline-Institution; p* = 0.0000 in *Risk* and $p = 0.0000$ in *Risk-Institution*). This aligns with the downward-sloping trend commonly observed in coordination experiments without institution formation (Van Huyck et al., 1990; Feri et al., 2010; Riedl et al., 2016). Moreover, the trend of average numbers across rounds 11–20 in Part 2 is different between treatments with and without institution formation. While a negative trend is observed in Part 2 in the treatments without institution formation ($p = 0.0000$ in *Baseline* and $p = 0.0000$ in *Risk*), there is no statistically significant downward trend in the treatments with institution formation ($p = 0.6797$ in *Baseline-Institution* and $p = 0.5000$ in *Risk-Institution*). Thus, institution formation has a stabilizing effect on average numbers.[12]

Similar comments apply to the development of minimum numbers. In Part 1, a significant descending trend of minimum numbers throughout rounds 1–10 is observed in the Baseline, Baseline-Institution and Risk-Institution treatments. A Jonckheere-Terpstra test for a descending ordered alternative hypothesis yields $p = 0.0003$ in *Baseline,* $p = 0.0311$ in *Baseline-Institution,* $p = 0.1956$ in *Risk* and $p = 0.0066$ in *Risk-Institution*. Note that the initial minimum number in round 1 in the treatment *Risk* (1.7) is already close to the lowest possible number, making the possibility for a descending trend across rounds in Part 1 unfeasible.[13] In Part 2, Fig. 1–Panel B shows that treatments with and without institution formation exhibit different patterns of minimum numbers. Minimum numbers decrease throughout rounds in the treatment *Baseline*, show no trend in the treatments *Baseline-Institution* and *Risk* and increase across rounds in the treatment *Risk-Institution*. In the treatment *Risk*, minimum numbers are again close to the lowest possible number in all rounds of Part 2 such that a negative trend is unlikely to emerge. A Jonckheere-Terpstra test for a descending ordered alternative hypothesis for

[8] The average and minimum numbers in all twenty rounds separated by groups and by treatments are shown in Figures A1-A4 (average numbers) and A5-A8 (minimum numbers) in Section I in the Supplementary Material. A test of significance of differences between Parts 1 and 2 for the different treatments, based on a Wilcoxon Signed-Rank Test, can be found in Table A.1, in Section I in the Supplementary Material.

[9] In the treatments *Baseline* and *Risk*, there is an increase in the effort level at the beginning of Part 2. Even in closed groups without rematching, like our setting, it is not uncommon to observe a “restart effect”, with an increase in provision levels after stopping part 1 of a game which is followed by a part 2, even if the instructions remain the same (see for example Burton-Chellew (2022)).

[10] The lack of significant differences is partly driven by an outlier group in the treatment *Risk* that played the payoff-dominant equilibrium in all 20 rounds. Excluding this group, minimum numbers in part 2 in the treatment *Risk-Institution* are significantly higher than in *Risk* (p = 0.027), while there is no significant difference in average numbers (p = 0.875).

[11] The Jonckheere-Terpstra test for a decreasing trend tests the null hypothesis that average (minimum) numbers are equal in all rounds against the alternative hypothesis that average (minimum) numbers are equal or decreasing over rounds. In particular, for Part 1, if μ_i denotes the average (minimum) number in round *i*, then H0: $\mu_1 = \mu_2 = \ = \mu_{10}$ and HA: $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_{10}$ with at least one strict inequality (Lunneborg, 2005). Alternatively, the test can also be specified for an ascending trend.

[12] A Friedman test was conducted as a robustness check for the results of the Jonckheere-Terpstra trend test, providing the same results. There is a significant difference among average numbers in Part 1 in all treatments (X^2=61.8136, p = 0.0000 in *Baseline*; X^2=70.5818, p = 0.0000 in *Baseline-Institution*; X^2=41.2303, p = 0.0000 in *Risk* and X^2=43.5818, p = 0.0000 in *Risk-Institution*). Moreover, there is a significant difference among average numbers in Part 2 in the treatments *Baseline* and *Risk* (X^2=43.0977, p = 0.0000 in *Baseline*; X^2=26.6061, p = 0.0016 in *Risk*) while the null hypothesis of equal average numbers across rounds in Part 2 is not rejected in the treatments *Baseline-Institution* and *Risk-Institution* (X^2=13.7758, p = 0.1305 in *Baseline-Institution* and X^2=4.6364, p = 0.8648 in *Risk-Institution*). Similar qualitative conclusions are obtained for minimum numbers.

[13] There was one group in the treatment *Risk* which coordinated on the highest possible number in all rounds 1-20 (See Figure A3, Group *Risk_5* in the Supplementary Material I). Excluding this outlier group from the analysis does not change qualitative results on trend behavior but shifts average numbers and minimum numbers downwards.

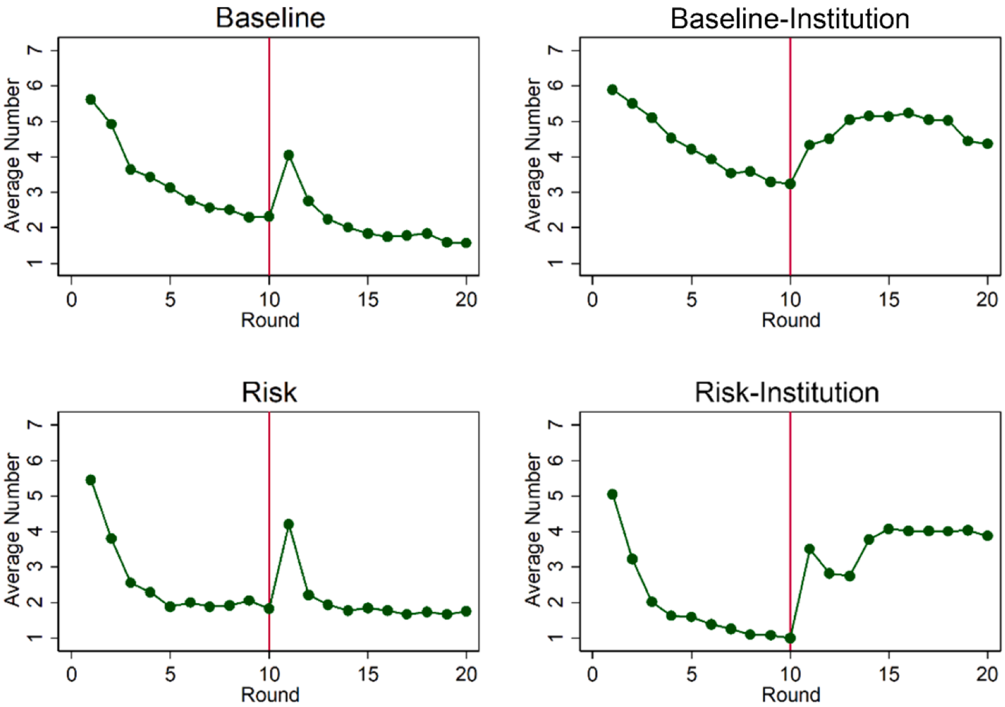


Fig. 1. Panel A: average numbers across rounds,
Note: The straight vertical line at round 10 separates rounds 1–10 of Part 1 from rounds 11–20 of Part 2.

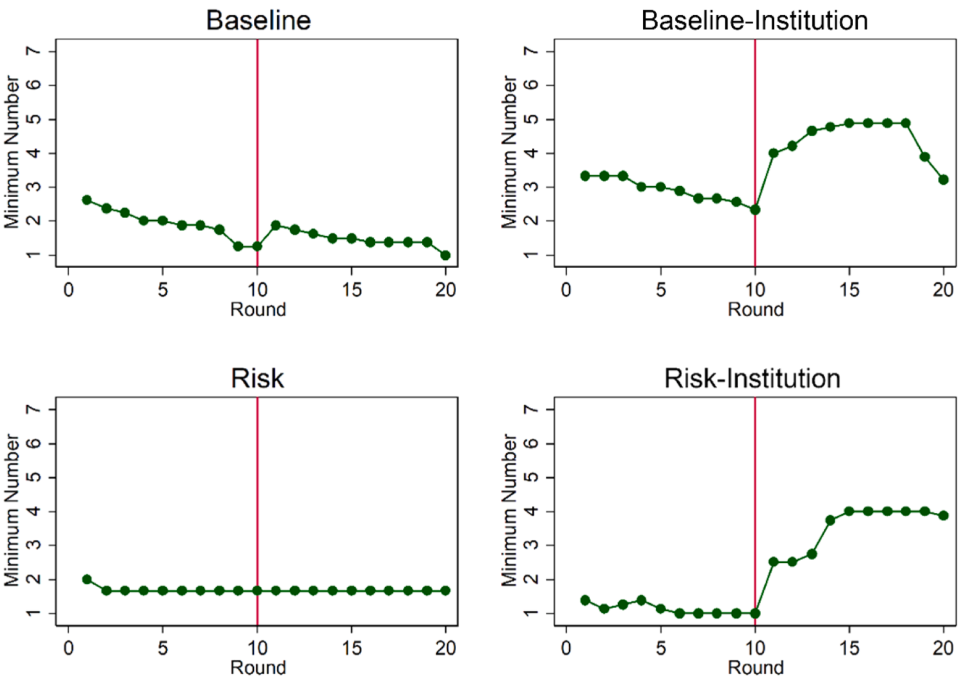


Fig. 1. Panel B: Minimum numbers across rounds.
Note: The straight vertical line at round 10 separates rounds 1–10 of Part 1 from rounds 11–20 of Part 2.

minimum numbers throughout all rounds in Part 2 yields $p = 0.0053$ in *Baseline*, $p = 0.5258$ in *Baseline-Institution*, $p = 0.5000$ in *Risk* and $p = 0.9313$ in *Risk-Institution*. In *Risk-Institution*, the alternative hypothesis of an ascending order of minimum numbers is not rejected ($p = 0.0687$). Result 3 summarizes our observations.

Table 5
Comparison between treatments of average numbers and minimum numbers in part 2.

	Baseline	*Baseline-Institution*	Mann-Whitney U Test[(3)]	*Risk*	*Risk-Institution*	Mann-Whitney U Test
Average[(1)] Number in Part 2	2.145	4.832	0.0209	2.056	3.688	0.8471
Minimum[(2)] Number in Part 2	1.475	4.433	0.0113	1.667	3.538	0.1348

Notes: (1) Average number: arithmetic mean of average numbers over all rounds in part 2 over all groups in the same treatment. (2) Minimum number: arithmetic mean of minimum numbers over all rounds in part 2 over all groups in the same treatment. (3) Significance of differences between parts is based on a Mann-Whitney U Test, displaying p-values.

Result 3. *Institution formation stops the downward trend of average and minimum numbers over rounds.*

Comparing the average and minimum numbers in the final round of Part 1 and Part 2 yields similar results. In *Baseline-Institution*, average and minimum numbers are not significantly different in rounds 10 and 20 ($p = 0.4156$ for average numbers and $p = 0.3221$ for minimum numbers), stopping the downward trend observed in *Baseline,* where average numbers are significantly lower in round 20 compared to round 10 ($p = 0.0350$). In *Risk-Institution*, both average and minimum numbers, are significantly higher in round 20 compared to round 10 ($p = 0.0474$), whereas in *Risk* there are no significant differences in average ($p = 0.2587$) or minimum numbers between rounds 10 and 20.

6. Equilibria analysis

6.1. Alternative equilibrium concepts (refinements)

From the discussion in Section 3, we know that there are multiple Nash equilibria in our game without institution formation and with institution formation the number of multiple Nash equilibria is even larger. Based on our experimental results, but also those of others on weakest-link games, it is clear that the concept of Pareto-dominance of Nash equilibria is not a good predictor. Given the nature of weakest-link games, it seems that the avoidance of risk plays some role for individuals' behavior, which suggests that equilibrium refinement concepts which allow for mistakes could perform better. Obvious candidates are different versions of risk-dominant equilibrium (Goeree et al., 2005; Harsanyi, 1995; Harsanyi and Selten, 1988), trembling-hand perfect equilibrium (Selten, 1975) and proper equilibrium (Myerson, 1978). In Section III, in the Supplementary Material, all concepts which are briefly introduced in this subsection are explained in more detail and applied to our weakest-link games without and with institution formation.

The concept of risk-dominant equilibrium is clearly defined for 2×2 coordination games (see Harsanyi and Selten, 1988): the risk-dominant action is a best response to the conjecture that assigns equal probability to each of the opponent's actions. There is less consensus on how the concept generalizes if the number of players and actions is larger (see, e.g., Harsanyi (1995) or Goeree et al. (2016)). As pointed out by Peski (2010), a straightforward option for coordination games is to extend this idea to multi-player and multi-action games, assigning equal probabilities to each of the opponent's actions for all opponents. We follow this approach.

The trembling-hand perfect equilibrium (Selten, 1975) and the proper equilibrium (Myerson, 1978) assume that all players make "small" mistakes, whereas only the latter assumes that individuals are better-responders, in the sense that players are less likely to make mistakes that are more costly to them. More precisely, Myerson's (1978) proper equilibrium attaches increasing and proportional errors to the probability that the second-best, third-best, and so on action is played by mistake.

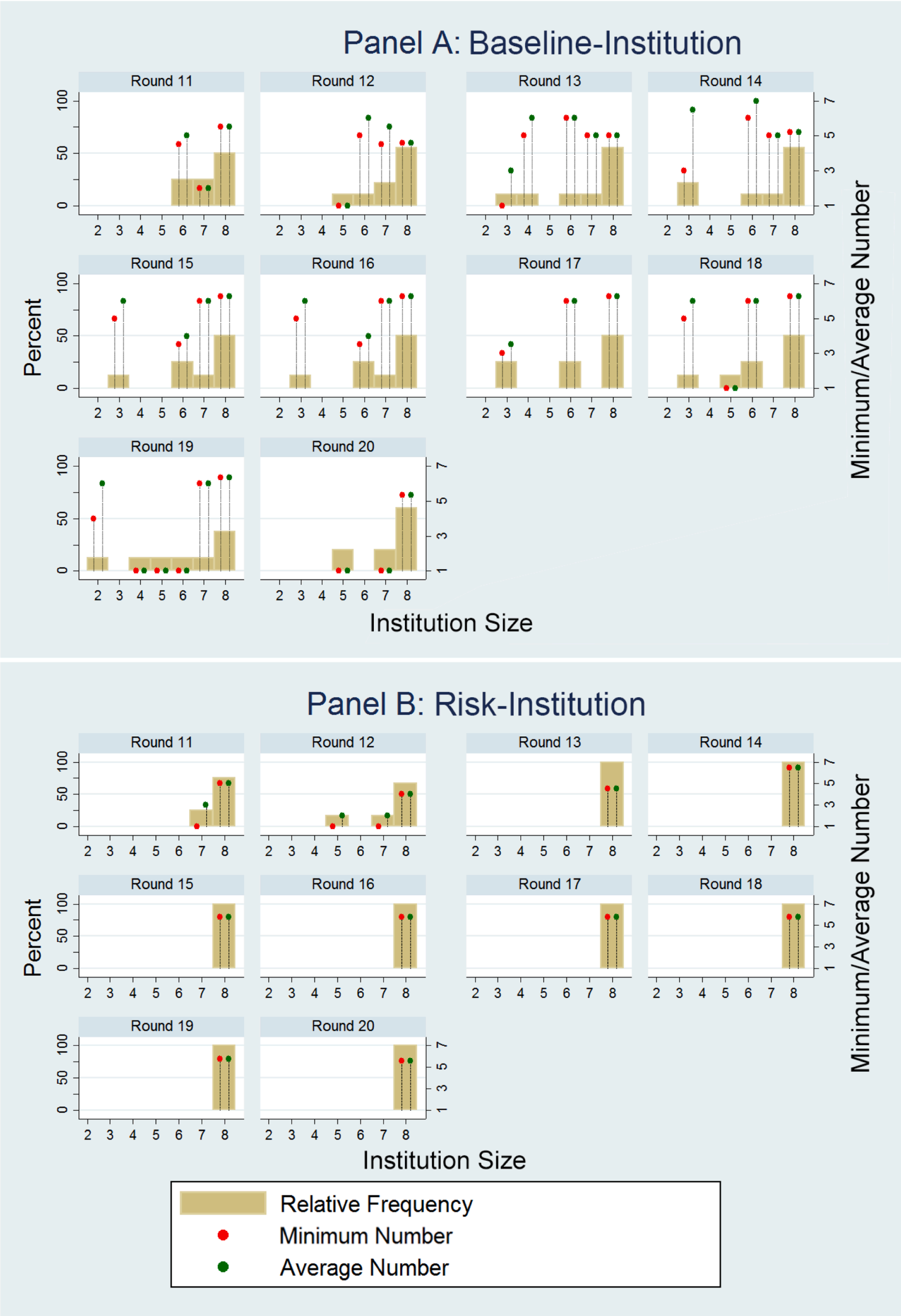
Finally, we consider the Quantal Response Equilibrium (QRE) by McKelvey and Palfrey's (1995), and its Agent version (AQRE) for games with more than one stage (McKelvey and Palfrey, 1998). These concepts assume that players are better-responders and make mistakes. When players are assumed to be almost perfectly rational, mistakes become less and less likely, and the QRE can be seen as an equilibrium refinement. It is closely related to risk dominance and the potential function approach proposed by Monderer and Shapley (1996), see Goeree and Holt (2005) for a general discussion on this relationship (and footnotes 14 and 15 for our case). However, the (A)QRE also accommodates less-than-perfect rational behavior, which proves helpful in explaining our data.

As shown in Table 6, in the context without institution formation, risk-dominance and QRE with almost perfectly rational agents select one unique equilibrium,[14] namely all individuals playing number 1, while trembling-hand and proper equilibrium do not reduce the number of Nash-equilibria. In the context with institution formation, all four equilibrium refinement concepts predict that all individuals join the institution, the grand institution is confirmed, and implements number 7.[15]

The intuition for the latter result is similar for all concepts considered, as they are all based on errors. For any intermediate institution structure, essentially nothing changes at stage 3 compared to the situation without institution formation, either all equilibria survive, or only the worst one survives. However, if the grand institution forms, equilibria that are not Pareto-optimal do not

[14] The same equilibrium emerges with the potential function proposed by Monderer and Shapley (1996). Using our notation, the authors show that for the VHBB weakest link game the selected equilibrium is to play 1, as long as $b<nc$ (or 20<8*10 with our parameter values). The same result holds for the FIS payoffs.

[15] Monderer and Shapley's (1996) potential function also predicts that the grand institution would implement 7. In this case, the potential function can be defined as $P(e_1, e_2, \ldots e_8) = (b - c)\min(e_1, e_2, \ldots e_8)$, which is maximized for e_i=7 for all i as long as $b > c$, which is the case for our parameter values.



(caption on next page)

Fig. 2. Distribution of sizes of implemented institutions over rounds.
Note: The figures present the distributions of different sizes of institutions over rounds in the treatments *Baseline-Institution* (Panel A) and *Risk-Institution* (Panel B) together with the *minimum number* and the *average number* for each size of institution. The relative frequency of each size of institution in a round (left-hand side of the vertical scale) is calculated as the percentage of this size of institution relative to the total number of institutions within this round. Relative frequencies sum up to 100 percent in each round. The minimum number for each size of institution (right-hand side of the vertical scale) is the arithmetic mean of minimum numbers over all groups with implemented institutions of that size within the respective round. The average number for each size of institution (right-hand side of the vertical scale) is the arithmetic mean of the number chosen by the institution over all groups with implemented institutions of that size within the respective round

Table 6
Equilibria with alternative concepts (refinements).

Concepts	Without institution formation	With institution formation
Nash equilibrium	Any number	Any institution size; any number
Pareto dominance	Play 7	Any institution size; play 7
Error-based refinements:		
Trembling hand, small errors	Any number	Grand institution; play 7
Proper equilibrium, small payoff-sensitive errors	Any number	Grand institution; play 7
(A)QRE with (almost) perfect rationality, payoff sensitive errors	Play 1	Grand institution; play 7
Risk dominance, large errors	Play 1	Grand institution; play 7
(A)QRE with bounded rationality, payoff sensitive errors	Probability of playing 1 increases with "rationality parameter"	Stage 1: probability of joining the institution increases with the degree of rationality Stage 2: probability of confirming the institution decreases with the degree of rationality for intermediate institutions and increases for the grand institution Stage 3: probability of playing 1 increases with the degree of rationality for intermediate institutions; and probability of playing 7 increases with the degree of rationality for the grand institution

survive. If I am in the grand institution, I can always profitably deviate from any equilibrium that is not Pareto-optimal, say playing number 3, by playing the Pareto-optimal number 7. If all the other members play number 3 as intended, I do not lose anything because I will get the benefit of 3 but will also pay the cost associated with number 3 (recall that all members implement the minimum suggested). However, if *all* other members play 7 by mistake, I will obtain a higher payoff if I play 7 (or place as much probability as possible on number 7). For similar reasons, all equilibria where not all members join (confirm) at stage 1 (respectively 2) do not survive the introduction of mistakes. See Section III.2 in the Supplementary Material for details.

In any case, this extreme prediction, moving to the best equilibrium with institution formation, cannot explain our experimental result that coordination is far from perfect. With a "smaller degree of rationality", the AQRE predicts that equilibrium numbers move up with the possibility of institution formation (but fail to reach 7) compared to when this option is not available. As this comes closer to what we have observed in our experimental setting, we devote the following two sections to the (A)QRE.

6.2. (A)QRE: concept and logit correspondence

The QRE is a Nash equilibrium based on a probabilistic choice function to model decision-making with payoff-sensitive errors (McKelvey and Palfrey, 1995), and the AQRE is an extension of the QRE to games with more than one stage (McKelvey and Palfrey, 1998). The basic idea on which the QRE and the AQRE are built is that decisions are stochastic, and all actions have a non-zero probability of being selected (Goeree et al., 2016). However, the probability of choosing non-optimal actions is inversely related to the loss induced. As in Myerson's (1978) proper equilibrium, players are better-response agents, not best-response agents. Unlike in Myerson's concept, where errors attached to the second-best, third-best, and so on, are increasing and proportional, in the "structural" formulation of the QRE, the assumption is that the expected payoff associated with any action includes an error term drawn from a commonly known distribution. That is, the disturbed expected payoff (denoted $\overline{V_{ij}}$) of action j for player i is:

$$\overline{V_{ij}}(\sigma) = V_{ij}(\sigma) + \varepsilon_{ij}$$

where V_{ij} is the expected payoff of agent i when she selects number j, ε_{ij} is a privately observed payoff disturbance associated to action j and player i (the unobservable component of utility), and σ is the set of all probabilities over all actions and all players (i.e., the set of completely mixed strategies over all players).

Note that $V_{ij}(\sigma)$ is not directly the payoff function shown in Eq. (1), as the expected payoff depends on the probabilities that all players in the group select each possible action (i.e., σ). Choosing 1 gives a certain payoff of 70 to a player, but the payoff associated with choosing 2 depends on the probability of any other player choosing 1, as a player only defines the minimum if no other player chooses 1. The same logic is applied to calculate the remaining expected payoffs. See Section III.1 in the Supplementary Material for details.

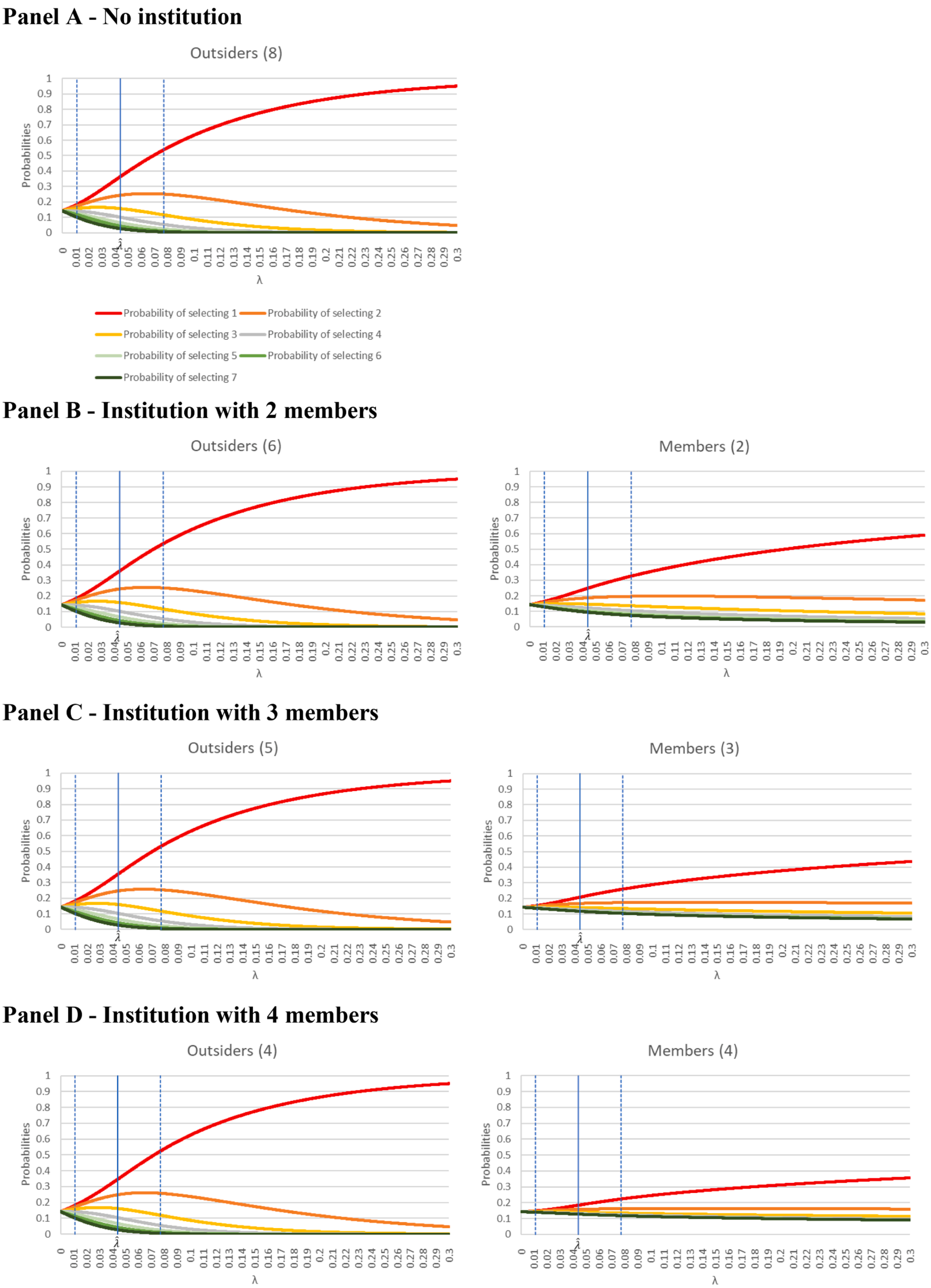
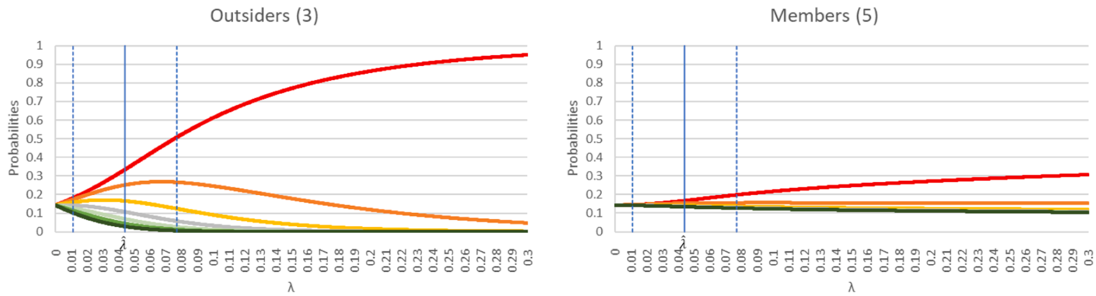


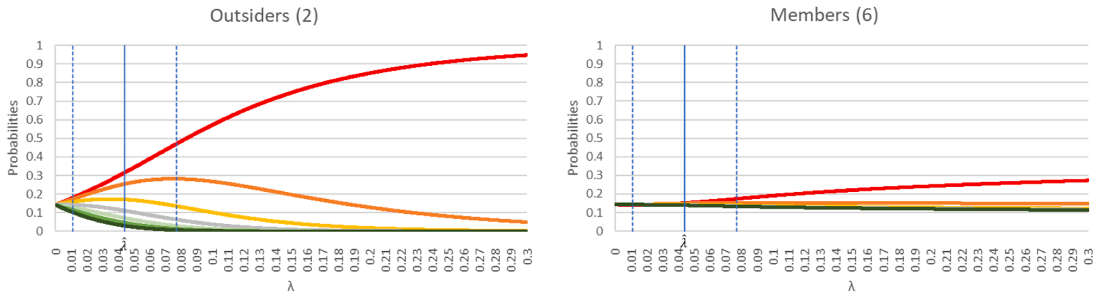
Fig. 3. AQRE correspondence for the Institution treatment and stage 3. Probabilities of selecting numbers between 1 and 7 for different institution structures and different values of λ.

Notes: Series 1 (red) shows the probability of selecting 1, Series 2 (brown) the probability of selecting 2, and so on (see Panel A for colors). Estimated values for $\widehat{\lambda}$ are shown using a solid blue line, and confidence intervals using a dashed blue line (see Table 7).

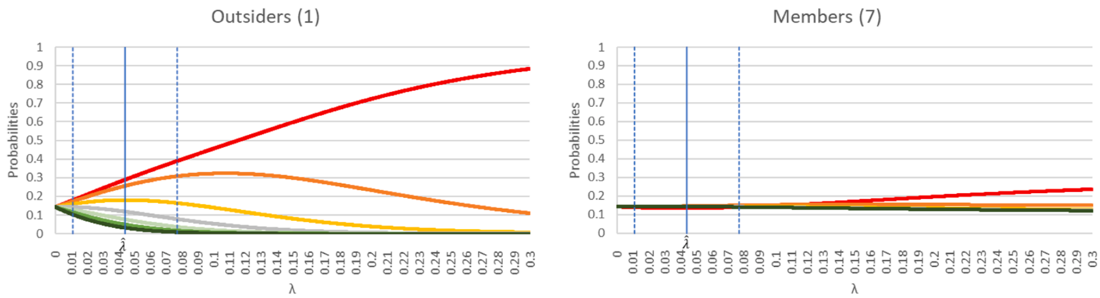
Panel E - Institution with 5 members



Panel F - Institution with 6 members



Panel G - Institution with 7 members



Panel H - Institution with 8 members

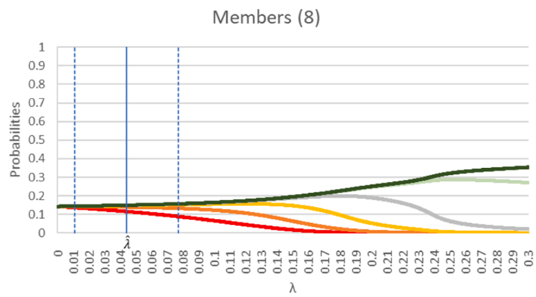


Fig. 3. (*continued*).

If one adds the assumption that the additive shock ε_{ij} follows an extreme value distribution, one can derive the logit quantal response function (McFadden, 1976; Goeree et al., 2016). Thus, we assume that ε_{ij} is independently and identically extreme value distributed with cumulative distribution $G(\varepsilon_{ij}) = \exp(-\exp(-\lambda\varepsilon_{ij}))$, where λ is a parameter. In our game, without institution formation, the logit QRE is the solution to the following system of equations:

$$\sigma_{ij}(\lambda) = \frac{e^{\lambda V_{ij}(\sigma)}}{\sum_{k=1}^{7} e^{\lambda V_{ik}(\sigma)}} \text{ for all } i \in \{1, 2, \ldots, 8\}, j \in \{1, 2, \ldots, 7\} \tag{3}$$

where σ_{ij} is the probability that player i selects number j.

The impact of the unobservable component of utility is summarized in the "error parameter"λ. Larger values of λ imply a smaller propensity to errors and make non-optimal choices less likely. For $\lambda = 0$, choices are purely random and for $\lambda \to \infty$ choices are (almost) perfectly rational, in the sense that options with larger expected payoffs are selected with a probability close to one. Therefore, λ can be viewed as a measure of the degree of "rationality" (Goeree et al., 2016), with all the limitations associated with using a single parameter for this purpose. Parameter λ can be estimated for our experimental data, as detailed in Section 6.3.

The logit equilibrium correspondence includes the solutions for $\lambda \in [0, \infty)$ and is defined as

$$\sigma^* : \lambda \to \left\{ \sigma : \ \sigma_{ij}(\lambda) = \frac{e^{\lambda V_{ij}(\sigma)}}{\sum_{i=1}^{7} e^{\lambda V_{ik}(\sigma)}} \ \forall\, i, j \right\}. \tag{4}$$

We find the logit correspondence in (4) numerically by focusing on symmetric equilibria (Turocy, 2010). As in all games, for $\lambda = 0$, there is a unique logit equilibrium close to the "centroid" of the game, where all strategies are adopted with equal probability. This implies that all numbers between 1 and 7 are played with equal probability. Then, we trace a principal branch of the logit correspondence for progressively higher values of λ until $\lambda \to \infty$. The limit point of the principal branch as λ approaches infinity is called the logit solution of the game. When a unique branch connects the centroid to exactly one Nash equilibrium, as in most simple games and in our weakest-link games, this can be seen as a strong refinement (McKelvey and Palfrey, 1995; Goeree et al., 2016). As the QRE is similar to the Bayesian equilibrium in Harsanyi's disturbed game, it is not surprising that it shares the feature that it selects a single equilibrium in "almost" all games (Harsanyi, 1973; McKelvey and Palfrey, 1995). In our weakest-link games without institution formation, the logit solution selects the worst equilibrium, all playing 1.

For our institution formation game, we apply the AQRE (McKelvey and Palfrey, 1998). A behavioral strategy is an AQRE if each player i is choosing her best response at every information set, taking σ_{-ij} and the error structure associated with the logit equilibrium as given. Furthermore, any limit point of a sequence of logit AQRE when $\lambda \to \infty$ corresponds to a sequential equilibrium of the game (Goeree et al., 2016). Basically, the solution is obtained via backward induction. More precisely, to calculate the AQRE, one needs to calculate the probabilities of all actions available at each information set, conditional on arriving at a particular information set and based on continuation values. The logit AQRE correspondence σ^* is then defined as the mapping from $\lambda \in [0, \infty)$ to the set of logit AQRE behavioral strategies:

$$\sigma^* : \lambda \to \left\{ \sigma : \ \sigma_{ij}(\lambda) = \frac{e^{\lambda V_{ijk}(\sigma)}}{\sum_{a_j \in A\left(h_i^k\right)} e^{\lambda V_{ijk}(\sigma)}} \ \forall\, i, j, k \right\} \tag{5}$$

where $A\left(h_i^k\right)$ is the finite set of actions a_j available to agent i at information set h_i^k. Details on how the payoffs and associated probabilities are calculated, focusing on symmetric equilibria, can be found in Section III.2 in the Supplementary Material.

As discussed in Section 3.2, the institution formation game has three stages. Let us consider first stage 3. At this stage, payoff functions are different for members and non-members. For members, benefits are given by the lowest number selected by the entire group, but the minimum selected by the institution determines costs. For non-members, benefits are also given by the lowest number selected by the entire group, but the number selected by the individual determines costs. If an institution with 2 up to 7 players is formed, at the third stage ($k = 3$), the solution to (5) implies solving a system of 14 equations for each value of λ, as we consider two types of agents in our symmetric equilibria (members and outsiders) and each type of agent can choose from 7 possible numbers ($j = 1, \ldots, 7$). If no institution has formed (or it has only 1 member), we have only one type of agent (outsiders). Hence, the system has 7 equations at this stage. If the institution has full membership, we have again only one type of agent (members), and a system of 7 equations.

At stage 2, the solution to (5) implies solving a system of two equations ($j = 0$, 1, with 1 indicating that membership is confirmed and 0 that membership is withdrawn). Finally, at stage 1, we also solve a system of two equations ($j = 0$, 1, with 1 indicating membership and with 0 indicating remaining a non-member). In both cases, payoffs are given by the continuation payoffs obtained by backward induction. For details, see Section III.2 in the Supplementary Material.

Let us now explore in more detail the correspondences for our weakest-link games, starting with the last stage of the game with institution formation (or the only stage of the game without institution formation). For stage three, Fig. 3 shows the correspondence for outsiders and members for all possible institution structures. The figure focuses on the game with the VHBB payoffs, but the correspondences for the FIS payoffs are qualitatively similar.

Panel A in Fig. 3 shows the case when no institution has been formed. This is relevant for the treatments *Baseline* but also *Baseline-Institution* provided at most one player joined the institution at stage 1, or when at least one player did not confirm her membership in stage 2. The graph focuses on the part of the distribution of λ that is relevant for our estimation (see Section 6.3). For the VHBB- and FIS-payoffs the logit equilibrium when $\lambda \to \infty$ implies that 1 is played with a probability close to one. In fact, for values of λ above 0.9211 for the VVHS-payoffs and above 0.1573 for the FIS-payoffs, the logit equilibrium yields a probability larger than 0.9999 of playing 1. That is, convergence to the worst possible outcome occurs faster under the FIS- than under the VVHS-payoffs, confirming the intuition that FIS-payoffs (Table 2) favor the risk-dominant action of choosing 1 even more than the VVHS-payoffs (Table 1).

When an institution with 2 to 7 members is formed, the correspondences differ for outsiders and members, as shown in Panels B) to G) in Fig. 3. Outsiders play 1 with a probability close to 1 for large values of λ for all institution structures. Furthermore, outsiders converge faster (i.e., for smaller values of λ) to the worst equilibrium (playing 1) the smaller the number of members joining the institution (and thus the larger the number of outsiders). The reason is that institutions tend to play larger numbers than outsiders. Thus, the smaller the number of outsiders in the game, the smaller the risk for an outsider of playing a low number (unless $\lambda \rightarrow \infty$).

The payoffs of members of the institution depend on the proposal made by their fellow members and the choice of outsiders. However, the impact is different. If a player is a member of the institution and an outsider plays a smaller number than the number implemented by the institution, the outsider defines the minimum, but not the effort (and hence cost) of this player. If a fellow member suggests a smaller number, she defines all members' minimum and effort. Hence, having an outsider play a smaller number is worse for a member than having a fellow member propose a smaller number. Thus, the larger the number of members in the institution, the smaller the risk for a member of playing a high number. This explains why convergence to the worst outcome occurs slower the higher the number of players in the institution. However, remarkably, even if there is only one outsider (Fig. 3, Panel G), the probability of playing the worst outcome is the only one that increases with λ, implying that ultimately, even inside the institution, all players will play 1 with probability close 1 for sufficiently large values of λ.

Only when there are no outsiders, the probability of large numbers being selected increases with λ (Fig. 3, Panel H). That is, only when there is absolutely no risk because all players have joined the institution, the AQRE predicts that players choose large numbers with a high probability with increasing values of λ.

At the second stage, the correspondence implies that the probability of confirming an institution decreases for all institutions, except for those with full membership, although very slowly. For full membership, the probability of confirming the institution increases with λ and is larger than 0.9999 for λ equal or larger than 0.2441. At the first stage, the probability of joining the institution also increases with λ, with a probability of joining equal or larger than 0.9999 for λ above 0.2585.

Taking the three stages together, for values of λ larger than 0.2585, the probability of joining the institution is basically 1; the probability of confirming it is essentially 1; but, as shown in Fig. 3, Panel H, players continue to play low numbers with relevant probabilities. Only for λ above 0.9216, the probability of playing 7 becomes larger than 0.9999. That is, the AQRE predicts that for large values of λ, eventually all players will join and confirm their membership and this institution with full membership plays the Pareto-dominant number 7. Note that this is the same equilibrium as the one predicted by the three alternative equilibrium concepts discussed in Section 6.1. However, this occurs for values of λ which are significantly larger than those that best explain our experimental results, as we will show below.

To sum up, when $\lambda \rightarrow \infty$, the A(QRE) provides opposite predictions with and without institution formation: (Almost) perfect rationality is a curse for the basic coordination game, while it is a blessing once institution formation is possible. For intermediate levels of λ, the prediction is that endogenous institution formation alleviates the coordination problem but does not achieve first-best outcomes.

6.3. (A)QRE: empirical estimation

To estimate the value of λ that best describes our data in the (A)QRE, we focus on the behavior observed in the last five rounds of the game (rounds 16 to 20), as we are interested in equilibrium behavior. For the case without institution formation, the log-likelihood function to be estimated is given by

$$\log L(\lambda, f) = \sum_{i=1}^{n} \sum_{j=1}^{7} f_{ij} \log\left(\sigma_{ij}^{*}(\lambda)\right) \tag{6}$$

where the observed empirical frequencies of strategy choices are denoted by f, and f_{ij} represents the number of observations of player i choosing number j. The maximum-likelihood estimates are $\widehat{\lambda} = \mathrm{argmax}_{\lambda} \log L(\lambda, f)$. For the case with institution formation, the estimation strategy is similar (see Section III.3 in the Supplementary Material in Section III.3 for details). Standard errors are estimated in both cases using the bootstrapping method (Efron, 1979; Train, 2009).

The first two columns of Table 7 show the results for the logit QRE model estimated for the two treatments without institution formation. The last two columns report the results for the logit AQRE model estimated for the two treatments with institution formation. The estimations of the parameter $\widehat{\lambda}$ are significant at the 1 % or 5 % level in all cases, except for the treatment *Risk-Institution* which has a p-value of 0.06.

For the third stage, Table 8 shows the average and minimum numbers observed in rounds 16 to 20 as well as the estimation-based predictions obtained with the different models, which are obtained by using the estimated $\widehat{\lambda}$-values shown in Table 7.

For the treatment *Baseline*, and as one should expect given the high significance of the estimated parameter λ, the model explains well the average behavior of the players in our sample, yielding an estimated average number almost identical to the observed one. In the case of the treatment *Risk*, the central value for the average number is within the confidence interval. However, the observed minimum numbers are above the predicted minimum numbers. The reason is that the numbers chosen by the players have less dispersion around the average than those predicted by the logit correspondence for $\widehat{\lambda}$ (the logit correspondence is a completely mixed equilibrium, where a positive probability is attached to all numbers).

For the case with institution formation, observed behavior implies an average number above the confidence interval for the treatment *Baseline-Institution* and within the confidence interval for the treatment *Institution-Risk*. The observed minimum numbers are above the numbers predicted by the model. Given that our estimates are significant, this is relatively surprising. Two effects explain

Table 7
Estimation results for the logit QRE and AQRE models for rounds 16–20.

	No institution treatments (QRE)		Institution treatments (AQRE)	
	Baseline	Risk	Baseline-Institution	Risk-Institution
Full sample				
$\hat{\lambda}$	0.0870*** (0.0242)	0.0527*** (0.0126)	0.0435** (0.0173)	0.0627*(0.0338)
Log-L	369.6219	223.6916	737.2310	360.4983
M	320	360	360	320

Note: Standard errors are shown in parentheses; $\hat{\lambda}$: estimated λ; M: number of observations. Significance at the 10 % (*), 5 % (**) and 1 % (***) level.

Table 8
Observed and estimation-based predicted average and minimum numbers for rounds 16 to 20.

	Observed Data		(A)QRE predictions[(1)]	
	Average Number	Minimum Number	Average Number	Minimum Number
Baseline	1.7063	1.3000	1.7069 [1.3523–2.6241]	1.0009 [1.0000–1.0380]
Risk	1.7167	1.6667	1.4565 [1.0907–2.6029]	1.0000 [1.0000–1.0025]
Baseline-Institution	4.8222	4.3556	2.5435 [1.8667–3.6699]	1.0314 [1.0029–1.2487]
Risk-Institution	3.9906	3.9750	1.2735 [1.0142–4.0000]	1.0000 [1.0000–1.3417]

[(1)] (A)QRE predictions use the $\hat{\lambda}$ shown in Table 7. 95 % confidence intervals are shown in square brackets. Average and minimum numbers for a given treatment are the arithmetic means over all groups and all five rounds.

this. The first effect is the lower dispersion around average numbers described in the previous paragraph. The second effect at play is that our model predicts that the probability of joining the institution is below the observed ones (see Table 4). For the values of λ that are relevant for our estimation (i.e., around the values that maximize the Log-likelihood function, roughly below 0.1000, see Table 7 and the blue lines in Fig. 3), choice probabilities vary significantly at stage 3 (Fig. 3), but hardly vary for stages 1 and 2, implying that the probabilities of joining and confirming an institution are still close to 0.50, the starting point of the logit correspondence at these stage for λ=0. We now highlight the main results discussed above.

Result 4. *Observed average numbers are within the confidence intervals predicted by the A(QRE) for the treatments Baseline, Risk and Risk-Institution, and above for the treatment Baseline-Institution. Observed minimum numbers are above the confidence intervals in all treatments.*

The main take-home message from this section is that our experimental results are best explained by levels of λ which clearly do not tend to infinity (Table 7).[16] As discussed in Section 6.2, for these values, the (A)QRE predicts that endogenous institution formation within a group alleviates the coordination problem but does not achieve first-best outcomes. This is the outcome that we observe, as highlighted in Result 2.

7. Conclusions

This study provides a first experimental analysis of the potential of endogenous institution formation in order to increase equilibrium provision levels in weakest-link games. Institutions facilitate subjects to coordinate on their preferred contribution levels in a setting without the possibility of excluding group members, i.e., fixed neighborhood, and without coercion. Institutions also reduce the potential risk for all players of costly high contributions which are undermined by other players who contribute less. The design of our institution requires little commitment by players and addresses the concern that individual members may not join an institution because of the possibility of being outvoted: an institution is only established if membership is confirmed unanimously, and also the provision level within the institution is agreed upon by unanimity, implementing the minimum proposed by any member of the institution. Our results show that even without the possibility of excluding individuals in a group, nor the punishment of group members, establishing a weak institution within a group can alleviate (despite not trivially solving) the coordination problem. This is a remarkable result in terms of testing theory, which also has interesting policy implications.

From a policy perspective, this result suggests that in settings where coordination failure is prevalent, establishing an institution by a group, even if governed by unanimity voting, can be an alternative to harsh interventions, entailing ostracism to group members or

[16] To improve predictions, one could introduce more flexibility in the AQRE model. As shown by Haile et al. (2008), introducing sufficient flexibility, a structural QRE and AQRE model can rationalize almost any behavior. This can be done by relaxing the Independence of Irrelevant Alternatives assumption in the (A)QRE and using a heterogeneous (A)QRE, where each individual follows a different quantal response function (Rogers et al., 2009). The main drawback of this approach is that it becomes more sample-specific. In any case, we have left this option for future research.

punishment. This conclusion is important as there are many situations in which exclusion is costly or simply impossible because it creates tensions among the members in a group or because of legal or physical constraints. Milder interventions, i.e., supporting the formation of institutions, come at the cost of not reaching first-best outcomes, but this might be a reasonable price to pay in exchange for contained conflicts.

We confirm the previous literature showing that the concept of Pareto-dominance of Nash equilibria is not a good predictor to explain experimental results in weakest-link coordination games. In fact, the problem of multiplicity of Nash equilibria was exacerbated in our analysis with institution formation. To shed some light, we have considered several mistake-based equilibrium refinements which all support the conjecture that our institution should solve the coordination problem.

The (Agent) Quantal Response Equilibrium was the most appropriate for explaining our experimental data. This equilibrium refinement, allowing for small deviations from perfectly rational agents, provided predictions for opposite behaviors with and without institution formation: (almost) perfect rationality is a curse for the basic coordination game, while it is a blessing once institution formation is possible. When considering agents with bounded rationality, the model captures the observed behavior that institution formation alleviates the coordination problem but does not solve it.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Esther Blanco reports financial support was provided by Austrian Science Fund. Alejandro Caparros reports financial support was provided by the Spanish Ministry of Economy and Competitiveness. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We are grateful to the participants to the Workshop on Experimental Economics for the Environment in Hamburg, Workshop on Behavioural and Experimental Methods in Manchester, International Meeting on Experimental and Behavioral Social Sciences in Lisbon, AERNA Conference, EAERE Conference, and the Stony Brook International Game Theory Conference. Seminar participants at the Universities of Columbia, Innsbruck and Durham also provided helpful comments. Philipp Buchenauer provided excellent research assistance. Financial support was provided by the Austrian Science Fund (Grant Numbers P 25973-G11 and P 32859) and the Spanish Ministry of Economy and Competitiveness (TrenGood, ECO2017-84461-R).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jebo.2025.106943.

Data availability

Data will be made available on request.

References

Anderson, S.P., Goeree, J.K., Holt, C.A., 2001. Minimum-effort coordination games: stochastic potential and logit equilibrium. Games Econ. Behav. 34, 177–199.

Avoyan, A., Ramos, J., 2023. A road to efficiency through communication and commitment. Am. Econ. Rev. 113 (9), 2355–2381.

Bloch, F., 1995. Endogenous structures of association in oligopolies. RAND J. Econ. 26, 537–556.

Bloch, F., 1997. Non-cooperative models of coalition formation in games with spillovers. In: Carraro, C and D Siniscalco (eds.). In: New Directions in the Economic Theory of the Environment, 10. Cambridge University Press, Cambridge, pp. 311–352. ch.

Blume, A., Ortmann, A., 2007. The effects of costless pre-play communication: experimental evidence from games with Pareto-ranked equilibria. J. Econ. Theory 132 (1), 274–290.

Bochet, O., Putterman, L., 2009. Not just babble: opening the black box of communication in a voluntary contribution experiment. Eur. Econ. Rev. 53, 309–326.

Bochet, O., Talbot, P., Louis, P., 2006. Communication and punishment in voluntary contribution experiments. J. Econ. Behav. Organ. 60, 11–26.

Brandts, J., Cooper, D.J., 2006. A change would do you good ... An experimental study on how to overcome coordination failure in organizations. Am. Econ. Rev. 96, 669–693.

Brandts, J., Cooper, D.J., Fatas, E., 2007. Leadership and overcoming coordination failure with asymmetric costs. Exp. Econ. 10, 269–284.

Burton-Chellew, M.N., 2022. The restart effect in social dilemmas shows humans are self-interested not altruistic. Proc. Natl. Acad. Sci. USA 119 (49), e2210082119, 1-7.

Cachon, G.P., Camerer, C.F., 1996. Loss-avoidance and forward induction in experimental coordination games. Q. J. Econ. 111 (1), 165–194.

Caparrós, A., Finus, M., 2020. Public good agreements under the weakest-link technology. J. Public Econ. Theory 22 (3), 555–582.

Chaudhuri, A., Schotter, A., Sopher, B., 2009. Talking ourselves to efficiency: coordination in inter-generational minimum effort games with private, almost common and common knowledge of advice. Econ. J. 119, 91–122.

Cooper, R.W., DeJong, D.V., Forsythe, R., Ross, T.W., 1990. Selection criteria in coordination games: some experimental results. Am. Econ. Rev. 80 (1), 218–233.

Cornes, R., 1993. Dyke maintenance and other stories: some neglected types of public goods. Q. J. Econ. 108 (1), 259–271.

Croson, R., Enrique, F., Neugebauer, T., Morales, A.J., 2015. Excludability: a laboratory study on forced ranking in team production. J. Econ. Behav. Organ. 114, 13–26.

Dannenberg, A., Gallier, C., 2020. The choice of institutions to solve cooperation problems: a survey of experimental research. Exp. Econ. 23, 716–749.

Deck, C., Nikiforakis, N., 2012. Perfect and imperfect real-time monitoring in a minimum-effort game. Exp. Econ. 15 (1), 71–88.
Devetag, G., 2005. Precedent transfer in coordination games: an experiment. Econ. Lett. 89 (2), 227–232.
Devetag, G., Ortmann, A., 2007. When and why? A critical survey on coordination failure in the laboratory. Exp. Econ. 10 (3), 331–344.
Efron, B., 1979. Bootstrapping methods: another look at the jackknife. Ann. Stat. 7, 1–26.
Fehr, D., 2017. Costly communication and learning from failure in organizational coordination. Eur. Econ. Rev. 93, 106–122.
Feri, F., Irlenbusch, B., Sutter, M., 2010. Efficiency gains from team-based coordination-large-scale experimental evidence. Am. Econ. Rev. 100 (4), 1892–1912.
Finus, M., Rundshagen, B., 2009. Membership rules and stability of coalition structures in positive externality games. Soc. Choice. Welf. 32, 389–406.
Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10 (2), 171–178.
Gneezy, U., Stephan, M., Pedro, R.B., 2011. When and why incentives (don't) work to modify behavior. J. Econ. Perspect. 25 (4), 191–209.
Goeree, J.K., Holt, C.A., 2005. An experimental study of costly coordination. Games Econ. Behav. 51, 349–364.
Goeree, J.K., Holt, C.A., Palfrey, T.R., 2016. Quantal Response Equilibrium: A Stochastic Theory of Games. Princeton University Press, Princeton.
Greiner, B., 2015. Subject pool recruitment procedures: organizing experiments with ORSEE. J. Econ. Sci. Assoc. 1 (1), 114–125.
Haile, P.A., Ali, H., Grigory, K., 2008. On the empirical content of quantal response equilibrium. Am. Econ. Rev. 98, 180–200.
Hamman, J., Rick, S., Weber, R.A., 2007. Solving coordination failure with "all-or-none" group-level incentives. Exp. Econ. 10, 285–303.
Harsanyi, J.C., 1973. Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points. Int. J. Game Theory 2, 1–23.
Harsanyi, J.C., 1995. A new theory of equilibrium selection for games with complete information. Games Econ. Behav. 8, 91–122.
Harsanyi, J.C., Selten, R., 1988. A General Theory of Equilibrium Selection in Games. MIT Press, Cambridge.
Hart, S., Kurz, M., 1983. Endogenous formation of coalitions. Econometrica 51, 1047–1064.
Hirshleifer, J., 1983. From weakest-link to best-shot: the voluntary provision of public goods. Public Choice 41, 371–386.
Isaac, R., Walker, M.J.M., 1988. Communication and free-riding behavior: the voluntary contribution mechanism. Econ. Inq. 26, 585–608.
Kosfeld, M., Akira, O., Arno, R., 2009. Institution formation in public goods games. Am. Econ. Rev. 99 (4), 1335–1355.
Koukoumelis, A., Vittoria Levati, M., Weisser, J., 2012. Leading by words: a voluntary contribution experiment with one-way communication. J. Econ. Behav. Organ. 8 (379), 390.
Kriss, P.H., Blume, A., Weber, R.A., 2016. Coordination with decentralized costly communication. J. Econ. Behav. Organ. 130, 225–241.
Lunneborg, C.E., 2005. Jonckheere–Terpstra test. Wiley StatsRef: Statistics Reference Online. Wiley.
McFadden, D., 1976. Quantal choice analysis: a survey. Ann. Econ. Soc. Meas. 5, 363–390.
McKelvey, R., Palfrey, T.R., 1995. Quantal response equilibrium for normal form games. Games Econ. Behav. 10, 6–38.
McKelvey, R., Palfrey, T.R., 1998. Quantal response equilibria for extensive form games. Exp. Econ. 1, 9–41.
Monderer, D., Shapley, L.S., 1996. Potential games. Games Econ. Behav. 14, 124–143.
Moulin, H., 1994. Serial cost sharing of excludable public goods. Rev. Econ. Stud. 55, 936–976.
Myerson, R.B., 1978. Refinements of the Nash equilibrium concept. Int. J. Game Theory 7 (2), 73–80.
Oprea, R., Gary, C., Daniel, F., 2014. Continuous time and communication in a public goods experiment. J. Econ. Behav. Organ. 108, 212–223.
Palfrey, T.R., Rosenthal, H., 1991. Testing for effects of cheap talk in a public goods game with private information. Games Econ. Behav. 3, 183–220.
Peski, M., 2010. Generalized risk-dominance and asymmetric dynamics. J. Econ. Theory 145, 216–248.
Riedl, A., Rohde, I.M.T., Strobel, M., 2016. Efficient coordination in weakest-link games. Rev. Econ. Stud. 83 (2), 737–767.
Rogers, B.W., Palfrey, T.R., Camerer, C.F., 2009. Heterogeneous quantal response equilibrium. J. Econ. Theory 144, 1440–1467.
Sandler, T., 2004. Global Collective Action. Cambridge University Press, New York.
Selten, R., 1975. Reexamination of the perfectness concept for equilibrium points in extensive games. Int. J. Game Theory 4 (1), 25–55.
Train, K.E., 2009. Discrete Choice Methods with Simulation: Second Edition. Cambridge University Press, New York.
Turocy, T.L., 2010. Computing sequential equilibria using agent quantal response equilibria. Econ. Theory 42, 255–269.
Vicary, S., 1990. Transfers and the weakest link: an extension of Hirshleifer's analysis. J. Public Econ. 43, 375–394.
Van Huyck, J.B., Battalio, R.C., Beil, R.O., 1993. Asset markets as an equilibrium selection mechanism: coordination failure, game form auctions, and tacit communication. Games Econ. Behav. 5 (3), 485–504.
Van Huyck, J.B., Battalio, R.C., Beil, R.O., 1990. Tacit coordination games, strategic uncertainty, and coordination failure. Am. Econ. Rev. 80 (1), 234–248.
Yi, S.S., 2003. Endogenous formation of economic coalitions: a survey of the partition function approach. In: Carraro, C. (ed.). In: The Endogenous Formation of Economic Coalitions, 3. Edward Elgar, Cheltenham, UK, pp. 80–127. et al., ch.