Collective Action: Experimental Evidence[1]

María Victoria Anauati[2]
Universidad de San Andrés

Brian Feld[3]
University of Illinois at Urbana-Champaign

Sebastian Galiani[4]
University of Maryland

Gustavo Torrens[5]
Indiana University

November 2015

Abstract. We conducted a laboratory experiment to test the comparative statics predictions of a new approach to collective action games based on the method of stability sets. We find robust support for the main theoretical predictions. As we increase the payoff of a successful collective action (accruing to all players and only to those who contribute), the share of cooperators increases. The experiment also points to new avenues for refining the theory. We find that, as the payoff of a successful collective action increases, subjects tend to upgrade their prior beliefs as to the expected share of cooperators. Although this does not have a qualitative effect on comparative static predictions, using the reported distribution of beliefs rather than an ad hoc uniform distribution reduces the gap between theoretical predictions and observed outcomes. This finding also allows us to decompose the mechanism that leads to more cooperation into a "belief effect" and a "range of cooperation effect".

JEL classification codes: D72, C92, H41

Key words: Collective action, multiple equilibria, laboratory experiment.

[1] We thank the University of Maryland for its financial support, the Universidad de San Andrés and the Universidad Nacional de La Plata in Argentina for providing a laboratory in which to conduct the experiment, and the Department of Economics of the University of Zurich for allowing us to use Z-tree. We would especially like to thank Lucia Yanguas for helping us with the code and logistics at the Universidad de San Andres and the CEDLAS at Universidad de La Plata. We are grateful for the very insightful comments provided by two anonymous referees.
[2] Email address: victoria.anauati@gmail.com
[3] Email address: bhfeld2@illinois.edu
[4] Email address: galiani@eecon.umd.edu
[5] Email address: gtorrens@indiana.edu

1. Introduction

The rational-choice theory of collective action comprises two main paradigms. Olson's model regards collective action as a prisoners' dilemma with only one equilibrium (Olson, 1965), while Schelling's model depicts it as a tipping game with multiple equilibria (Schelling, 1978). Medina (2007) develops a unifying framework that covers both paradigms and produces novel comparative statics predictions about the effects of the parameters of the game on the probability of a successful collective action. In this paper we use a simple laboratory experiment to test some of these implications.

The unifying framework relies on the notion of stability sets to deal with multiple equilibria. The method of stability sets originally proposed by Harsanyi and Selten (1988) and further developed and applied to collective action problems by Medina (2007) is a very useful theoretical tool for studying large collective action games with multiple equilibria. The crucial advantage of the stability-sets method is that it provides an assessment of the likelihood of different equilibria as a function of the payoffs of the game and the distribution of prior beliefs. Thus, the method can be used to generate clear predictions on the comparative statics of the probability of a successful collective action with respect to any variable that affects the payoffs of the collective action game. The focus of this paper is to test these comparative statics predictions using a controlled, randomized laboratory experiment. In particular, we concentrate on testing a key theoretical prediction of the new framework. The probability of a successful collective action should increase in line with the benefit accrued to all players involved, including those who do not contribute if the collective action is successful, as well as in line with the extra benefit obtained by those who do contribute.

In order to test these predictions, we conducted a laboratory experiment at the Universidad de San Andrés and the Universidad Nacional de La Plata in the Province of Buenos Aires, Argentina. We recruited undergraduate and graduate students from any field of study and regardless of their knowledge of game theory and economics. We conducted 16 sessions (7 at the Universidad de San Andrés and 9 at the Universidad Nacional de La Plata) with 20 subjects each, totaling 320 participants. In each round of each session, subjects were randomly allocated into groups of 10 and asked to play a simple game. At the beginning, each subject has 1 point and must decide whether to invest it or not. The probability that the investment is successful depends on the share of subjects who contribute their point. If the investment is successful, all players obtain B points and those who contributed obtain s extra points. Depending on the values of B and s, the game has one Nash equilibrium in which nobody contributes, or three Nash equilibria, one in which nobody contributes, another in which all contribute and a third one in which each player contributes with positive probability (the same for all players). We consider 4 possible treatments. Treatment 1 is the baseline free-rider Olsonian model with one Nash equilibrium in which nobody contributes. In treatments 2 to 4, we gradually increase B and/or s, inducing multiple equilibria. Furthermore, the probability of a successful collective action predicted by the stability-sets method is 0 in treatment 1 and increases to 0.25, 0.50, and 0.75 in treatments 2, 3, and 4, respectively (assuming initial beliefs about the expected share of cooperators are uniformly distributed).

In general, we find robust support for the main theoretical predictions of the stability-sets method applied to collective action. As B and/or s are increased, the share of cooperators and, hence, the probability of a successful collective action increases. Analogous results are obtained for the payoffs. The effects are statistically significant whether or not we include controls for individual characteristics, level of understanding of the game as measured by performance on a quiz before playing the rounds, fixed effects by session, whether or not subjects are asked to report their prior beliefs about the expected share of cooperators, whether or not the collective action was successful in the previous round, and the number of players in the same group who decided to invest in the previous round.

We also find that, on average, there is more cooperation than predicted by the theory when theoretical predictions are obtained under different assumptions regarding the distribution of expected cooperators. As a benchmark, we first assume that subjects' prior beliefs about the share of cooperators have a uniform distribution over the interval $[0,1]$ for all treatments. This can be considered to be a Laplacian assumption when no information on prior beliefs is available. Second, in some randomly selected sessions, before subjects started playing, we asked them to report their prior beliefs as to the share of cooperators in each treatment. We find that subjects' prior beliefs are not uniformly distributed and vary among treatments. Specifically, as the benefit of cooperation increases, subjects upgrade their assessments concerning the expected share of cooperators. Using reported prior beliefs to compute the theoretical prediction regarding the probability of successful collective action reduces the gap between the model predictions and observed behavior. Even so, the data point to the existence of more cooperation than expected.

Finally, taking into account the fact that prior beliefs vary among treatments, we decompose the total effect on the probability of a successful collective action into two analytically different effects. In particular, as the benefit of cooperation increases, subjects upgrade their assessments of the expected share of cooperators. We illustrate how to compute the change in the probability of a successful collective action attributed to belief upgrading (belief effect) and to an increase in the range of prior beliefs that induce cooperation (range of cooperation effect).

Experiments on Collective Action and Multiple Equilibria. There are three branches of experimental literature connected with this work. First, there is a vast body of literature on laboratory experiments with public good games. Second, a large number of experiments that employ games with multiple equilibria have been conducted to study equilibrium selection. Third, there is the literature on belief elicitation, which is directly related to our decomposition.[6]

Experiments with Public Good Games.[7] Many laboratory experiments have been conducted with *public good games with only one Nash equilibrium* (see, among others, Marwell and Ames, 1981; Isaac, Walker and Williams, 1994; Ostrom, 1998; Cherry et al., 2005; and Hichri, 2005). Most of these studies have concentrated on the dynamics of behavior in finitely repeated interactions and have found levels of cooperation that are significantly greater than those indicated by theoretical predictions.[8] Although we also find more cooperation than predicted by the theory in most of our treatments and, in particular, in treatment 1, which has only one Nash equilibrium, the focus of our work is on testing the comparative static predictions of the stability-set method in the context of multiple equilibria.

---

[6] In addition to laboratory experiments, field experiments with collective action games have also been conducted. See, for example, Schmitt (2000), Cardenas (2003) and Barr et al. (2012). However, none of them has tested the comparative statics predictions derived from the stability-set method.

[7] An extensive number of studies using variations of the design of public good experiments have been synthesized in Davis and Holt (1993), Ledyard (1995), Offerman (1997) and Chaudhuri (2011).

[8] Several mechanisms have been proposed to explain this phenomenon, including kindness, altruism, conditional cooperation, reciprocity and repetition (see, for example, Anderson, Goeree and Holt, 1998, and Fischbacher, Gätcher and Fehr, 2001).

Experiments with the *probabilistic provision of public goods* have a particular bearing on our work. In particular, Dickinson (1998) introduced uncertainty in the provision of a public good in a standard voluntary contribution mechanism. In his first treatment, there is an ex post and exogenous probability of provision while, in his second treatment, the probability of provision is endogenous and increasing in aggregate contribution levels. Likewise, in our experiment the success of the collective action is probabilistic and depends positively on the share of contributors. Beyond this similarity, the aim and approach of our work are very different, however. While Dickinson (1998) studies how uncertainty affects contributions, we explore how changes in the payoffs interact with prior beliefs to induce changes in the probability of a successful collective action. Another minor difference is that in our experiment the probability of a successful collective action reaches 1 when all players cooperate. The method of stability sets, however, can also be applied to a collective action model where unanimous cooperation does not imply a probability of successful collective action equal to 1.

Our work is also related to the large body of literature on *threshold public good games*. In contrast to standard linear public good games, threshold public good games have multiple Nash equilibria. Furthermore, for a deterministic threshold, Pareto-efficient outcomes are sustainable as Nash equilibria (see, for example, Palfrey and Rosenthal, 1984; and Bagnoli and Lipman, 1989). Uncertain thresholds, on the other hand, can restore the free-riding incentives and lead to a Pareto-inefficient set of equilibria (see, for example, Nitzan and Romano, 1990; Suleiman, 1997; McBride, 2006; and Barrett, 2013). Our setting can also produce multiple Nash equilibria. Indeed, for treatments 2 to 4, the model has three Nash equilibria: a Pareto-efficient equilibrium in which all players cooperate and two Pareto-inefficient equilibria, one in which all players defect and another in which all defect with the same positive probability.

Several authors have experimented with deterministic thresholds. Isaac et al. (1989) conducted an experiment with deterministic thresholds in an otherwise standard public good game. They found that a rise in the thresholds typically increases contribution levels and decreases the likelihood that the threshold will be reached. Cadsby and Maynes (1998) also experimented with different thresholds in a variety of scenarios and obtained similar results. Croson and Marks (2000) considered a public good game with a deterministic threshold. They proved that the game has a set of efficient Nash equilibria in which the threshold is exactly met and a set of inefficient Nash equilibria in which the public good is not provided. They conducted an experiment and found that the higher the step return,[9] the higher the aggregate contributions and the greater the probability that the public project will be successfully funded. In our model, there is no exogenous threshold that we can randomize, but a variation in the payoffs induces a change in the tipping point beyond which players cooperate. In particular, an increase in the benefit from a successful collective action lowers the tipping point, which induces a higher probability of cooperation. Since our experiment supports this theoretical prediction, our results are consistent with the findings of Isaac et al. (1989), Cadsby and Maynes (1998) and Croson and Marks (2000).

Many researchers have also experimented with probabilistic thresholds. Wit and Wilke (1998) studied the effects of environmental uncertainty in the provision threshold on contributions in a public good game. They found lower contributions under high uncertainty. Gustafsson et al. (2000) compared voluntary contributions in public good games with the same expected provision threshold but with different variances. They found that the average contribution is smaller in the high-variance group. Au (2004) also compared voluntary contributions in a public good experiment with fixed and uncertain provision thresholds. He found that the provision rate for the public good was significantly higher when the provision point was known precisely.[10] Overall, these experiments suggest that uncertainty in the provision threshold makes cooperation harder. In our model, it is possible to induce environmental uncertainty about the location of the tipping point, introducing uncertainty in the payoffs. Although we do not have treatments that deal with this situation, it is worth noting that the theoretical predictions of the model are consistent with the findings reported in the experimental literature on probabilistic thresholds.

Up to this point, we have stressed the commonalities between our work and the literature on threshold public good games. There is, however, a fundamental difference between the stability-set approach and threshold public good games. While in threshold

[9] The concept of the Step Return (SR) is analogous to the concept of the Marginal Per-Capita Return adapted to threshold public goods games. The SR captures the private value of moving one unit of resources from an individual's private consumption to the public good. Specifically, the SR is expressed as: $SR = \frac{aggrgate\ group\ payoff\ from\ the\ public\ good}{total\ contribution\ threshold}$.

[10] While Wit and Wilke (1998) and Gustafsson et al. (2000) applied a simultaneous design, Au (2004) obtained similar results using a sequential decision-making protocol.

public good games the provision threshold is exogenous, the stability-sets approach produces an endogenous threshold or tipping point that separates the space of prior beliefs in two regions, the stability set of cooperation and the stability set of no cooperation. Then, the likelihood that prior beliefs belong to the cooperation region is the theoretical prediction of the probability of a successful collective action. Unlike threshold public good games, the tipping point in our experiment depends on the payoff structure and, hence, changes in the payoffs lead to changes in the probability of cooperation. Our main focus is on testing those comparative static effects.

Experiments with Multiple Equilibria, Selection and Learning. Our work is also related to a large body of literature on experiments with multiple equilibria games and equilibrium selection. For example, Van Huyck et al. (1990) experimented with coordination games; Van Huyck et al. (1991) with average opinion games; Battalio et al. (2001) and Golman and Page (2010) with stag-hunt games; Cason et al. (2004), Neugebauer et al. (2008) and Oprea et al. (2011) with hawk-dove games; and Haruvy and Stahl (2000) with symmetric normal-form games with multiple Nash equilibria. There is also a large body of literature on tests of equilibrium selection theories in multiple equilibrium games with repeated interactions. See, for example, Van Huyck et al. (1990, 1991) and Iwasaki et al. (2003). To the best of our knowledge, none of these authors has tested the predictions provided by the stability-sets method. Only Golman and Page (2010) use a related approach, but for a very different purpose, namely, to compare cultural learning with belief-based learning. They consider a class of generalized stag-hunt games in which agents can choose from among multiple potentially cooperative actions or can opt for a secure, self-interested action. Though the set of stable equilibria is identical under the two learning rules, the basins of attraction for the efficient equilibria are much larger for cultural learning. Moreover, as the stakes grow arbitrarily, cultural learning always locates an efficient equilibrium while belief-based learning never does. In some sense, we are adopting and testing a different approach to multiple equilibria. Instead of focusing on identifying different criteria for equilibrium selection, we use the stability-sets method to obtain theoretical predictions of the probability of occurrence of each of the Nash equilibria of the collective action game.

Many experimental efforts have found it useful to employ the notion of quantal-response equilibrium (QRE), which in many cases increases the accuracy of theoretical predictions (McKelvey and Palfrey, 1995). The learning model that underlies the notion of QRE is one in which players take into account the possibility that they can make mistakes in their choices, that other players can also make mistakes and that, hence, they can be mistaken about their perceptions of other players. According to Medina (2013), the tracing procedure can be used to represent more general processes of belief formation, above and beyond the one implicitly stipulated by QRE. In particular, we can study how players react if they are highly confident or highly skeptical about other players' cooperation levels. Indeed, we find evidence that taking into account those prior beliefs is an important factor in narrowing the gap between the model predictions and observed behavior.

Beliefs Elicitation. Our decomposition of the mechanism that leads to more cooperation owing to a belief effect and a range-of-cooperation effect is related to the experimental work on belief elicitation. Eliciting beliefs has long been of interest to researchers because it contributes to an understanding of subjects' motives.[11] In the context of public good games, Offerman (1997) and Offerman, Sonnemans and Schram (1996, 2001) elicited beliefs about the behavior of other agents in order to gain a deeper understanding of the observed results in step-level public good games. Fischbacher, Gächter and Fehr (2001) conducted a one-shot standard linear public good game experiment that directly elicited subjects' willingness to engage in conditional cooperation. In one treatment, subjects were asked to make a single decision as to how many tokens to invest in a common fund while, in a second treatment, they were asked to indicate, for each average contribution level of other group members, how much they wanted to contribute to the common fund. The researchers found that roughly 50% of the subjects showed conditional behavior, i.e., their own contribution increased in step with the other group members' average contribution.[12] Croson (2007) experimented with finitely repeated, simultaneous linear public good games. Before each period, in some treatments subjects were asked to reveal their priors about the share of

[11] See Schotter and Trevino (2014) for a recent survey on belief elicitation in laboratory experimental economics.

[12] In a similar vein, Neugebauer, Perote, Schmidt and Loos (2005) elicited beliefs regarding the reasons for the declining pattern of cooperation in finitely repeated games and Fischbacher and Gächter (2006) did so in order to measure how beliefs and contributions are correlated with public good games with random matching.

cooperators in their group. She found no significant difference in levels of contributions between situations in which participants revealed their priors or did not do so, but contributions were positively related to prior beliefs about others contributions.

Our findings are consistent with these works. For example, we also find that, as subjects upgrade their assessments of the expected share of cooperators, they are more likely to cooperate. More importantly, a novel contribution of our study is that we propose a method for separating the individual contribution of changes in prior beliefs from the observed behavior. Thus, we show that higher benefits from a successful collective action lead to a higher probability of cooperation, not only because the tipping point for cooperation is lower, but also because subjects update their beliefs as to the share of cooperators.

Advances in the Theory of Collective Action. Our work is also related to several advances in the formal theory of collective action. In particular, many authors have recently modeled collective action as a global game (see, among others, Edmond, 2008; Boix and Svolik, 2009; Egorov et al., 2009; Persson and Tabellini, 2009; and Egorov and Sonin, 2014). To the best of our knowledge, the precise relationships between the global game and stability-sets methods have not yet been plotted out. There are, however, many similarities between them. First, both approaches deal with multiple equilibria. Global games can be seen as a particular instance of equilibrium selection. While complete information games often have multiple equilibria, introducing a natural perturbation leads to a unique rationalizable action for each player. If players do not share common knowledge about the payoffs of the game, and instead rely upon privately observed signals with a small level of noise, the perturbation selects a unique equilibrium (for 2x2 games, see Carlsson and van Damme, 1993; and, for various generalizations, see Morris and Shin, 1998; Morris and Shin, 2001; Frankel, Morris, and Pauzner, 2003; and Morris, Shin and Yildiz, 2015). The problem of multiple equilibria is also at the core of the stability-sets method. This method partitions the space of prior belief profiles into regions assigned to each equilibrium and then employs the distribution of prior beliefs to produce an estimate of the probability of the occurrence of each equilibrium.

Second, in applications of the global game approach to collective action, it is usually assumed that there are global strategic complementarities, i.e., a setting in which a player's incentive to cooperate (defect) increases with the likelihood that others cooperate (defect).[13] These kinds of strategic complementarities are present in our model because cooperating is more attractive when the collective action is more likely to succeed, and the collective action is more likely to succeed when other players are more likely to cooperate. Finally, both approaches make it possible to perform comparative static analyses with respect to the underlying parameters of the model, which is the crucial challenge to be met by any useful collective action theory (see Medina, 2013).

The rest of this paper is organized as follows: Section 2 presents the theoretical framework. Section 3 describes the laboratory experiment. Section 4 explains how it was determined that subjects understood the game that they were playing and that the randomization was balanced. Section 5 presents descriptive statistics for the main variables. Section 6 presents the main results of the paper. Section 7 shows a decomposition of a change in the predicted share of cooperators in a "belief effect" that captures the change in prior beliefs and a "range of cooperation effect" that captures the change in the range of prior beliefs that induced cooperation. Finally, Section 8 concludes.

2. Theoretical Framework

In this section, we present a collective action model based on Medina (2007). Then, we adapt the model for use in a laboratory experiment. We focus on the comparative static results of the model under two different assumptions regarding the distribution of prior beliefs concerning the expected share of cooperators. First, we assume that the distribution of prior beliefs is fixed for the whole set of parameters of the collective action game. Second, we relax this assumption and assume that a change in parameters that increases the set of beliefs that induce cooperation leads to a new distribution of prior beliefs that first-order stochastically dominates the prior one.

2.1. A collective action model (based on Medina 2007). Consider a set of players $\boldsymbol{N > 2}$. For each player, the set of pure strategies is $\boldsymbol{A_i = \{C_i, D_i\}}$, where $\boldsymbol{C_i = 1}$ and $\boldsymbol{D_i = 0}$ denote "cooperate" and "defect", respectively. Let $\boldsymbol{a_i}$ indicate a generic

[13] An exception is Shadmehr and Dan Bernhardt (2011).

element of $\boldsymbol{A_i}$. The set of mixed strategies is $\Delta(\boldsymbol{A_i})$, and $\boldsymbol{\alpha_i}$ indicates a generic element of $\Delta(\boldsymbol{A_i})$, where $\boldsymbol{\alpha_i} = \mathbf{Pr}(\boldsymbol{C_i})$ and $(\mathbf{1} - \boldsymbol{\alpha_i}) = \mathbf{Pr}(\boldsymbol{D_i})$. Let $\boldsymbol{A} = \times_{\boldsymbol{i=1}}^{\boldsymbol{N}} \boldsymbol{A_i}$, and $\boldsymbol{a}$ indicates a generic element of $\boldsymbol{A}$. There are two possible outcomes: either the collective action is a success or it is a failure, indicated by $\boldsymbol{S}$ and $\boldsymbol{F}$, respectively. The probability that the collective action is successful is a function $\boldsymbol{G}$ of the proportion of players who cooperate. Formally, $\mathbf{Pr}(\boldsymbol{S}) = \boldsymbol{G}(\boldsymbol{\gamma}(\boldsymbol{a}))$, where $\boldsymbol{\gamma}(\boldsymbol{a}) = \frac{\mathbf{1}}{\boldsymbol{N}} \#\{\boldsymbol{i}: \boldsymbol{a_i} = \boldsymbol{C_i}\}$. Logically, $\mathbf{Pr}\ (\boldsymbol{F}) = \mathbf{1} - \mathbf{Pr}(\boldsymbol{S})$. $\boldsymbol{G}$ is assumed to be continuous, monotonically increasing (as the proportion of cooperators rises, the probability of success also increases) and $\boldsymbol{G}(\mathbf{0}) = \mathbf{0}$. The payoff for each player $\boldsymbol{u_i}$ depends only on the player's action and the outcome of the collective action. Thus, $\boldsymbol{u_i}$ can be fully described with just four numbers: $\boldsymbol{u_i}(\boldsymbol{C_i}, \boldsymbol{S})$ (the payoff when i cooperates and the collective action is successful), $\boldsymbol{u_i}(\boldsymbol{C_i}, \boldsymbol{F})$ (the payoff when i cooperates and the collective action does not prosper), $\boldsymbol{u_i}(\boldsymbol{D_i}, \boldsymbol{S})$ (the payoff when i defects and the collective actions is successful) and $\boldsymbol{u_i}(\boldsymbol{D_i}, \boldsymbol{F})$ (the payoff when i defects and the collective action does not prosper). Moreover, we will assume that, for all $\boldsymbol{i}$, it is always the case that $\mathbf{min}\,\{\boldsymbol{u_i}\,(\boldsymbol{C_i}, \boldsymbol{S}), \boldsymbol{u_i}(\boldsymbol{D_i}, \boldsymbol{S})\} > \boldsymbol{u_i}(\boldsymbol{D_i}, \boldsymbol{F}) > \boldsymbol{u_i}(\boldsymbol{C_i}, \boldsymbol{F})$.

Medina (2007) studies this game when $N \rightarrow \infty$, i.e., he focuses on a large game of collective action. We briefly summarize his results when all players have identical payoff functions. Any correlated equilibrium of a large game of collective action can be represented by an aggregate share γ_μ. Define $W = \frac{u_i(D_i,F) - u_i(C_i,F)}{u_i(D_i,F) - u_i(C_i,F) + u_i(C_i,S) - u_i(D_i,S)}$. If $W > 1$, then there exists a unique equilibrium where nobody cooperates $(\gamma_\mu = 0)$. However, if $0 < W < 1$, the large game of collective action has three correlated equilibria: One equilibrium in which all players cooperate $(\gamma_\mu = 1)$, another in which nobody cooperates $(\gamma_\mu = 0)$ and a third one in which there is an intermediate level of cooperation given by $G(\gamma_\mu) = W$. In order to deal with the multiplicity of equilibria, Medina (2007) extends the notion of stability sets originated by Harsanyi and Selten (1988). He uses a methodology known as the "tracing procedure" to assign a set of initial belief conditions to each equilibrium. These conditions can be represented as a share of expected cooperators γ_η. Then, the stability set of an equilibrium is defined as the set γ_η assigned to it. The key result states that $\gamma_\eta < W$ belongs to the stability set of $\gamma_\mu = 0$, while $\gamma_\eta > W$ belongs to the stability set of $\gamma_\mu = 1$. As Medina (2007) emphasizes, the threshold value of γ_η that separates the stability set of $\gamma_\mu = 0$ from the set of $\gamma_\mu = 1$ is associated with the mixed-strategy equilibrium implicitly given by $G(\gamma_\mu) = W$. In order to see this more clearly, assume that $G(\gamma) = \gamma$. Then, the stability set of $\gamma_\mu = 0$ is the set of all shares of expected cooperators lower than the mixed-strategy equilibrium share of cooperators $\gamma_\mu = W$, while the stability set of $\gamma_\mu = 1$ is the set of all shares of expected cooperators higher than the mixed-strategy equilibrium share of cooperators $\gamma_\mu = W$.

Finally, Medina (2007) shows how to use stability sets to compute the probability of cooperation. In order to do so, assume that the initial belief conditions γ_η are distributed with the CDF H. Then:

$$\Pr(\gamma_\mu = 1) = \Pr\left(\gamma_\eta > W\right) = 1 - \Pr\left(\gamma_\eta < W\right) = 1 - H(W)$$

This expression is very useful for deducing comparative static results. In particular, note that, as W increases, the probability of cooperation decreases.

Olson's Model (single Nash equilibrium for large N): The standard Olson's public good model of collective action is a special case of the above model when the payoffs are given by:

$$u_i(C_i, S) = B - c,\ u_i(D_i, S) = B,\ u_i(C_i, F) = -c,\ u_i(D_i, F) = 0 \tag{1}$$

where c >0. For this model, $W = \infty$. Hence, when $N \to \infty$, the unique equilibrium is $\gamma_\mu = 0$. More intuitively, in a large group ($N \to \infty$) there is a free rider problem (it is a dominant strategy for every player to defect) that impedes the members of the group from furthering their common interests.

Schelling's model (multiple Nash equilibria for large N): Consider a simple modification of Olson's model:

$$u_i(C_i, S) = B - c + s,\ u_i(D_i, S) = B,\ u_i(C_i, F) = -c,\ u_i(D_i, F) = 0 \tag{2}$$

where $s > c > 0$. For this model $W = \frac{c}{s} < 1$. Hence, when $N \to \infty$, there are three Nash equilibria $\gamma_\mu = 0$, $\gamma_\mu = 1$, and γ_μ such that $G(\gamma_\mu) = \frac{c}{s}$. The stability set of $\gamma_\mu = 0$ is $\{\gamma_\eta : 0 \le \gamma_\eta < W\}$, while the stability set of $\gamma_\mu = 1$ is $\{\gamma_\eta : W < \gamma_\eta \le 1\}$. More intuitively, introducing an extra payoff $s > c$ that is obtained only by those who cooperate when the collective action is successful transforms Olson's game into a coordination game with multiple equilibria. If everybody defect, the best strategy is to defect, but if everybody cooperates, the best strategy is to cooperate. Moreover, there is a threshold for the share of expected cooperators ($W = \frac{c}{s}$) such that players cooperate if and only if they expect there to be more cooperators than this threshold value. Finally, if the expected share of cooperators γ_η is distributed with the cumulative distribution function *H,* we have:

$$\Pr(\gamma_\mu = 1) = 1 - H\left(\frac{c}{s}\right)$$

Hence, as c decreases and/or s increases, the probability of cooperation increases. In Schelling's model, according to the stability-sets method, as c decreases and/or s increases, it is more likely that players will coordinate in the efficient equilibrium.

2.2. Laboratory adaptation. In order to test the predictions derived by Medina (2007) using a laboratory experiment, we need to make some adjustments to the model presented in the previous section. The most important change is that we must consider the case when N is finite. This implies that we need to compute a threshold for the number of players such that the game with finite N has the same set of equilibria as the large game. To do so, we focus on simple cases. In particular, we will assume $\boldsymbol{G}(\boldsymbol{\gamma}) = \boldsymbol{\gamma}$ when studying the Olson and Schelling models.

We begin by defining a Nash equilibrium for the game of collective action when N is finite. Let $S(k) = \{a: \sum_j a_j = k\}$ and $S(k,i) = \{a_{-1}: \sum_{j \neq i} a_j = k\}$. $S(k)$ is the set of pure strategy profiles in which k players cooperate, while $S(k,i)$ is the set of pure strategy profiles of all players except i in which k players cooperate. Given a strategy profile $\alpha = (\alpha_i, \alpha_{-i})$, we can compute the probability that k players cooperate given that player i does not cooperate. This is given by $P(k,i) = \sum_{a_{-i} \in S(k,i)} \prod_{j \neq i} \alpha_j^{a_j} (1 - \alpha_j)^{1-\alpha_j}$.

Therefore, the payoff for player i associated with the strategy profile $\alpha = (\alpha_i, \alpha_{-i})$ is given by:

$$U_i(\alpha_i, \alpha_{-i}) = \alpha_i \sum_{k=0}^{N-1} \left[\left(u_i(C_i, S) - u_i(C_i, F)\right) G\left(\frac{k+1}{N}\right) + u_i(C_i, F) \right] P(k, i)$$

$$+ (1 - \alpha_{-i}) \sum_{k=0}^{N-1} \left[\left(u_i(D_i, S) - u_i(D_i, F)\right) G\left(\frac{k}{N}\right) + u_i(D_i, F) \right] P(k, i)$$

Definition 1. *A Nash equilibrium for the collective action game with finite N is a strategy profile α such that, for each I, one of the following conditions holds:*

$U_i(1, \alpha_{-i}) \geq U_i(0, \alpha_{-i})$ *and* $\alpha_i = 1$

$U_i(1, \alpha_{-i}) \leq U_i(0, \alpha_{-i})$ *and* $\alpha_i = 0$

$U_i(1, \alpha_{-i}) = U_i(0, \alpha_{-i})$ *and* $\alpha_i \in (0,1)$

The following proposition characterizes the set of Nash equilibria for the Olson and Schelling collective action games when N is finite and $G(\gamma) = \gamma$.

Proposition 1. *Suppose that N is finite and* $G(\gamma) = \gamma$. *Then:*

1. ***Olson's model****: Assume that* u_i *is given by (1). Then, if* $N < \frac{B}{c}$, *the unique Nash equilibrium is* C_i *for all i, while, if* $N > \frac{B}{c}$, *the unique Nash equilibrium is* D_i *for all i.*

2. ***Schelling's model****: Assume that* u_i *is given by (2). Then, if* $N < \frac{B+s}{c}$, *then* C_i *for all i is the unique Nash equilibrium, while, if* $N > \frac{B+s}{c}$, *there are three Nash equilibria:* C_i *for all i,* D_i *for all i, and* $\alpha_i = \hat{\alpha} = \frac{cN-B-s}{s(N-1)}$ *for all i. Moreover, in the third Nash equilibrium, the expected share of cooperators is* $\mathbf{E}\left[\frac{k}{N}\right] = \hat{\alpha}$.

Proof: See Online Appendix 1. ■

Summing up, when $N > \frac{B}{c}$, the unique Nash equilibrium for Olson's model with finite N is D_i for all i, which coincides with the equilibrium for Olson's model when $N \to \infty$.

Similarly, when $N > \frac{B+s}{c}$, the set of Nash equilibria for Schelling's model with finite N is C_i for all i, D_i for all i, and $\alpha_i = \hat{\alpha} = \frac{cN-B-s}{s(N-1)}$ for all i, which is analogous to the set of Nash equilibria for Schelling's model when $N \to \infty$. Note, in particular, that $\lim_{N\to\infty} \hat{\alpha} = \frac{c}{s}$, which is the share of cooperators in the mixed strategy equilibrium in the large game.

Suppose that, as in the large collective action game, we use the mixed strategy equilibrium for Schelling's model with finite N to compute the probability of occurrence of the two pure strategy equilibria. In particular, assume that the share of expected cooperators γ_η is distributed according to the cumulative distribution function H. Then:

$$\Pr\ (C_i \text{ for all } i\ ) = 1 - H(\hat{\alpha})$$

Changes in $\hat{\alpha}$. The probability of cooperation increases with B and s and decreases with c. Formally:

$$\frac{\partial \Pr\ (C_i \text{ for all } i)}{\partial B} = -H'(\hat{\alpha})\frac{\partial \hat{\alpha}}{\partial B} > 0$$

$$\frac{\partial \Pr\ (C_i \text{ for all } i)}{\partial s} = -H'(\hat{\alpha})\frac{\partial \hat{\alpha}}{\partial s} > 0$$

$$\frac{\partial \Pr\ (C_i \text{ for all } i)}{\partial c} = -H'(\hat{\alpha})\frac{\partial \hat{\alpha}}{\partial c} < 0$$

because $\frac{\partial \hat{\alpha}}{\partial B} < 0$, $\frac{\partial \hat{\alpha}}{\partial s} < 0$, and $\frac{\partial \hat{\alpha}}{\partial c} > 0$. For example, if H is the uniform distribution, then we have $\Pr\ (C_i \text{ for all } i\ ) = \frac{(s-c)\ N+B}{s\ (N-1)}$ and, hence, $\frac{\partial \Pr\ (C_i \text{ for all } i)}{\partial B} = \frac{1}{s(N-1)} > 0$, $\frac{\partial \Pr\ (C_i \text{ for all } i)}{\partial s} = \frac{(N-1)cN+B}{[s(N-1)]^2} > 0$ and $\frac{\partial \Pr\ (C_i \text{ for all } i)}{\partial c} = \frac{-N}{s(N-1)} < 0$. Intuitively, as B and/or s increases or c decreases, the threshold for the share of expected cooperators that makes players indifferent to the choice of cooperating or defecting decreases. Cooperation becomes more attractive and, hence, players require a lower share of expected cooperators in order to cooperate. Thus, given the distribution of the share of expected cooperators, the probability of cooperation increases.

Changes in $\boldsymbol{H}$ induced by changes in $\widehat{\boldsymbol{\alpha}}$. Suppose that $\boldsymbol{H}$ is not independent of $\widehat{\boldsymbol{\alpha}}$. In particular, assume that we have a family of distributions indexed by $\widehat{\boldsymbol{\alpha}}$. We write $\boldsymbol{H}(\boldsymbol{\gamma}_{\boldsymbol{\eta}}, \widehat{\boldsymbol{\alpha}})$ to indicate the probability that the expected share of cooperators is less than or equal to $\boldsymbol{\gamma}_{\boldsymbol{\eta}}$ when $\boldsymbol{\alpha}_{\boldsymbol{i}} = \widehat{\boldsymbol{\alpha}} = \frac{\boldsymbol{cN-B-s}}{\boldsymbol{s(N-1)}}$ for all $\boldsymbol{i}$ is the mixed strategy Nash equilibrium of the collective action game. Furthermore, assume that $\boldsymbol{H}(\boldsymbol{\gamma}_{\boldsymbol{\eta}}, \widehat{\boldsymbol{\alpha}}') \leq \boldsymbol{H}(\boldsymbol{\gamma}_{\boldsymbol{\eta}}, \widehat{\boldsymbol{\alpha}})$ for all $\boldsymbol{\gamma}_{\boldsymbol{\eta}}$ whenever $\widehat{\boldsymbol{\alpha}}' < \widehat{\boldsymbol{\alpha}}$, i.e., $\boldsymbol{H}(.,\widehat{\boldsymbol{\alpha}}')$ first-order stochastically dominates $\boldsymbol{H}(.,\widehat{\boldsymbol{\alpha}})$ when $\widehat{\boldsymbol{\alpha}}'$ is lower than $\widehat{\boldsymbol{\alpha}}$. Then, we have:

$$\frac{\partial \Pr\ (C_i \text{ for all } i)}{\partial B} = -[H_1(\hat{\alpha},\hat{\alpha}) + H_2(\hat{\alpha},\hat{\alpha})]\frac{\partial \hat{\alpha}}{\partial B} > 0$$

$$\frac{\partial \Pr\ (C_i \text{ for all } i)}{\partial s} = -[H_1(\hat{\alpha},\hat{\alpha}) + H_2(\hat{\alpha},\hat{\alpha})]\frac{\partial \hat{\alpha}}{\partial s} > 0$$

$$\frac{\partial \Pr\ (C_i \text{ for all } i)}{\partial c} = -[H_1(\hat{\alpha},\hat{\alpha}) + H_2(\hat{\alpha},\hat{\alpha},)]\frac{\partial \hat{\alpha}}{\partial c} < 0$$

H_1 (H_2) is the partial derivative with respect to the first (second) argument, $H_2 > 0$, because $H(.,\hat{\alpha}')$ first-order stochastically dominates $H(.,\hat{\alpha})$ whenever $\hat{\alpha}' < \hat{\alpha}$, $\frac{\partial \hat{\alpha}}{\partial B} < 0$, $\frac{\partial \hat{\alpha}}{\partial s} < 0$, and $\frac{\partial \hat{\alpha}}{\partial c} > 0$. Note that if $H(.,\hat{\alpha}')$ first-order stochastically dominates $H(.,\hat{\alpha})$ whenever $\hat{\alpha}' < \hat{\alpha}$, then the dependence of H on $\hat{\alpha}$ magnifies all the comparative statics derivatives without affecting their signs. Intuitively, as B and/or s increases or c decreases, the probability of cooperation increases for two reasons. First, the threshold for the share of expected cooperators that makes players indifferent to the choice of cooperating or defecting decreases. Second, initial beliefs about the expected share of cooperators are updated in the sense that the new distribution gives at least as high a probability of an initial belief being at least γ_η as the prior distribution does.

3. The Laboratory Experiment

In this section we describe our laboratory experiment. First, we provide a general description of the experiment, including its monetary payoffs, number of sessions and rounds, matching procedure and the instructions received by the subjects. Second, we give a detailed description of the game that the subjects played. Finally, we summarize the treatments and compute the corresponding theoretically predicted outcomes.

3.1. General description of the experiment. The experiment was conducted in May - October 2014 at the Universidad de San Andrés and the Universidad Nacional de La Plata, both of which are located in the Province of Buenos Aires, Argentina. We recruited undergraduate and graduate students from any field of study and regardless of how familiar they were with game theory and economic theory. We conducted 16 sessions with 20 subjects each, totaling 320 participants. Subjects were allowed to participate in only one session. Every session included four treatments, which made it possible to avoid any treatment selection problem. In each treatment, subjects were asked to play a collective action game. The experiment was programmed and conducted using z-Tree software (Fischbacher, 2007). Each session lasted approximately 50 minutes. The experiment proceeded as follows:

Allocation to computer terminals. Before each session began, subjects were randomly assigned to computer terminals.

Instructions. After subjects were at their terminals, they received the instructions, which were also explained by the organizers. Subjects then had time to read the instructions on their own and ask questions. Online Appendix 2.1 contains an English translation (from Spanish) of the script that we employed to provide instructions while Online Appendix 2.2 contains the printed version of the instructions. This was the last opportunity that subjects had to ask questions.

Prior beliefs. At the beginning of randomly selected sessions, subjects were asked to report their assessments of how the game would unfold.[14] In particular, for each treatment, we asked each subject how many subjects from a group of 10 they thought would contribute their point. This allowed us to obtain an empirical distribution of individuals' prior beliefs regarding the expected share of cooperators for each treatment. The questions we asked can be found in Online Appendix 2.3.

Quiz. In order to check whether participants understood the rules of the game, we asked them to take a five-question quiz. The quiz was administered after we had given the instructions, but before the rounds began. Subjects were paid approximately US$ 0.25 per correct answer, but we never informed them which ones they had answered correctly. The quiz questions can be found in Online Appendix 2.4.

[14] For each session, all participants were asked to report their prior beliefs with probability 1/2. Thus, on average, subjects reported their prior beliefs in half of the sessions.

Rounds. After the subjects had finished the quiz, they began playing rounds, during which they interacted solely through a computer network using z-Tree software. Subjects played 16 rounds of the collective action game. The first 4 rounds were for practice, and the last 12 rounds were for pay. At the end of each round, subjects received a summary of the decisions taken by both themselves and their partners, including payoffs per round, their own accumulated payoffs for paid rounds and nature's decision. Online Appendix 2.5 provides some samples of the screens that the subjects saw.

Matching. In odd-numbered rounds, 10 players were randomly matched and played treatment 1 (T_1), while the other 10 players played treatment 3 (T_3). In even-numbered rounds, 10 players were randomly matched and played treatment 2 (T_2), while the other 10 players played treatment 1 (T_4). See below for a detailed explanation of the treatments.

Questionnaire. Finally, just before leaving the laboratory, all the subjects were asked to complete a questionnaire which was designed to enable us to test the balance across experimental groups and to control for their characteristics in the econometric analysis. This questionnaire is presented in Online Appendix 2.6.

Payments. All subjects were paid privately, in cash. After the experiment was completed, a password appeared on each subject's screen. The subjects then had to present this password to the person who was running the experiment in order to receive their payoffs. Subjects earned, on average, US$ 11.80, which included a US$ 2 show-up fee, US$ 0.25 per correct answer on the quiz and US$ 0.25 for each point they received during the paid rounds of the experiment. All payments were made in Argentine currency; at the time, US$ 1 was equivalent to AR$ 8.[15]

3.2. Treatments and predicted outcomes. Once they had finished the quiz, subjects directed their attention to their computers and proceeded to play the first round of the session. In each round, subjects were randomly assigned to one of two groups, each consisting of 10 participants. At the beginning of the round they received one point and then decided whether to keep it for themselves or to invest it in a common fund. The

[15] Since Argentina's rate of inflation was very high, we adjusted the conversion rate in order to maintain the purchasing power of the payments. Thus, from May to July, the conversion rate was 2 pesos per point, while, from August to October, it was 2.4 pesos per point.

probability that the investment in the common fund would be successful equals the share of subjects who contribute their point out of the group of 10. If the investment was successful, all players obtained $\boldsymbol{B}$ points and those who contributed obtained $\boldsymbol{s}$ extra points.

The experiment consisted of four different treatments. The first treatment represents a scenario of no cooperation opportunities ($\boldsymbol{B} = \mathbf{1.25}$ and $\boldsymbol{s} = \mathbf{0}$); in other words, this is the free-rider Olsonian model with one Nash equilibrium, in which nobody contributes. In treatments 2 to 4, we gradually increased $\boldsymbol{B}$ and/or $\boldsymbol{s}$, inducing multiple equilibria. Specifically, the second treatment represents a scenario of low cooperation opportunities ($\boldsymbol{B} = \mathbf{1.25}$ and $\boldsymbol{s} = \mathbf{1.25}$); the third treatment is associated with a scenario of high (but not full) cooperation opportunities ($\boldsymbol{B} = \mathbf{3}$ and $\boldsymbol{s} = \mathbf{1.25}$); and the fourth, a scenario where the incentives to cooperate are the highest ($\boldsymbol{B} = \mathbf{3}$ and $\boldsymbol{s} = \mathbf{1.75}$).

Table 1 summarizes the relevant parameters for each treatment and indicates the predicted share of contributors and the predicted profit if there are 10 players in each group and assuming that prior beliefs are uniformly distributed in the interval $[\mathbf{0}, \mathbf{10}]$.

Table 1: Treatments and Predicted Share of Cooperators with a Uniform Prior Belief Distribution

Treatment	N	B	C	s	Predicted share of cooperators (Priors uniformly distributed) Pr (C_i for all i)	Predicted payoff (Priors uniformly distributed)
	(1)	(2)	(3)	(4)	(5)	(6)
T_1	10	1.25	1	0.00	0.000	1.000
T_2	10	1.25	1	1.25	0.333	1.500
T_3	10	3.00	1	1.25	0.490	2.625
T_4	10	3.00	1	1.75	0.667	3.500

4. Understanding of the Game and Randomization Balance

In this section we show that subjects understood the game and that the randomization was balanced. Table 2 shows that, on average, subjects understood the rules of the game. Indeed, 81% of them answered question 1 correctly, 95% answered question 2 correctly, 78% answered question 3 correctly and 89% answered question 4 correctly. It seems that subjects found question 5 to be more difficult, since only 70% of them answered it correctly.

Table 3 shows the randomization balance across treatments. Note that the same group of 20 subjects was randomly matched to play T_1 or T_3 in odd-numbered rounds and T_2 or T_4 in even-numbered rounds. This enabled us to determine whether subjects with given characteristics were more frequently allocated to one treatment over another. In the comparisons among the four treatments, all characteristics and levels of understanding of the game were perfectly balanced between T_1 and T_2 and between T_3 and T_4. In some cases, there were slight imbalances in terms of undergraduate/graduate students and nationality, mostly at a 5% significance level. Nevertheless, we had to reject the null hypothesis at the 10% and 5% levels of statistical significance in less than 10% of the tests. Moreover, the imbalances in nationality and undergraduate/graduate students were due to the fact that there were very few foreigners (95% of the subjects were Argentines) and very few graduates in the sample (94% of the subjects were undergraduates).

Table 2: Balance across Treatments (I)

	All subjects			T_1		T_2		T_3		T_4	
	Number of subjects	Mean	S.d.	Mean	S.d.	Mean	S.d.	Mean	S.d.	Mean	S.d.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Subject characteristics											
Age	320	21.76	3.34	21.62	3.27	21.78	3.36	21.91	3.41	21.74	3.32
Nationality (Argentine=1)	320	0.95	0.21	0.95	0.22	0.96	0.19	0.96	0.20	0.94	0.23
Studied game theory (=1)	320	0.47	0.50	0.45	0.50	0.47	0.50	0.50	0.50	0.48	0.50
Gender (male=1)	320	0.51	0.50	0.51	0.50	0.51	0.50	0.51	0.50	0.51	0.50
Graduate studies (=1)	320	0.06	0.23	0.05	0.21	0.07	0.25	0.07	0.25	0.04	0.20
Spanish language (=1)	320	0.97	0.16	0.97	0.16	0.98	0.14	0.98	0.15	0.97	0.17
Understanding the experiment											
Answered correctly: question 1	320	0.81	0.39	0.82	0.39	0.82	0.39	0.80	0.40	0.80	0.40
Answered correctly: question 2	320	0.95	0.22	0.95	0.22	0.95	0.22	0.95	0.22	0.95	0.22
Answered correctly: question 3	320	0.78	0.42	0.77	0.42	0.77	0.42	0.78	0.41	0.78	0.41
Answered correctly: question 4	320	0.89	0.31	0.89	0.31	0.88	0.32	0.89	0.31	0.90	0.30
Answered correctly: question 5	320	0.70	0.46	0.71	0.46	0.71	0.46	0.69	0.46	0.70	0.46

Note: The mean is the sample mean and S.d. is the standard deviation for the corresponding variable in each line. Entries in columns (1)-(3) indicate the values for the complete sample; those in columns (4)-(5) represent the subjects who played treatment 1; columns (6)-(7) show those who played treatment 2; columns (8)-(9) show those who played treatment 3; and those who played treatment 4 are recorded in columns (10)-(11).

Table 3: Balance across Treatments (II)

	T_1/T_2 (1)	T_1/T_3 (2)	T_1/T_4 (3)	T_2/T_3 (4)	T_2/T_4 (5)	T_3/T_4 (6)
Characteristics of subjects						
Age	-0.164	-0.287*	-0.123	-0.123	0.041	0.164
Nationality (Argentine=1)	-0.017*	-0.012	0.004	0.005	0.021**	0.016*
Studied game theory (=1)	-0.021	-0.054**	-0.033	-0.033	-0.012	0.021
Gender (male=1)	0.004	0.003	-0.001	-0.001	-0.005	-0.004
Graduate studies (=1)	-0.023**	-0.021**	0.002	0.002	0.025**	0.023**
Spanish language (=1)	-0.007	-0.006	0.001	0.001	0.008	0.007
Understanding of the experiment						
Answered correctly: question 1	0.001	0.017	0.016	0.016	0.015	-0.001
Answered correctly: question 2	0.002	0.002	0.000	0.000	-0.002	-0.002
Answered correctly: question 3	-0.002	-0.014	-0.012	-0.012	-0.010	0.002
Answered correctly: question 4	0.008	0.002	-0.005	-0.006	-0.013	-0.007
Answered correctly: question 5	0.002	0.014	0.012	0.012	0.010	-0.002

Note: Each entry indicates the mean difference between the two treatments in the column for the corresponding variable in each line. * indicates that the difference of means test is significant at 10%; ** significant at 5%; *** significant at 1%.

5. Descriptive Analysis

In this section we first present descriptive statistics for the decisions taken by the subjects (share of cooperators and payoffs by treatment). Then, we study the distribution of the subjects' prior beliefs regarding the share of expected cooperators. Finally, we show that the average share of cooperators and average payoffs do not differ between the sessions at which subjects were asked to report their prior beliefs and those in which they were not asked to do so.

5.1. Cooperation decision. Table 4 provides descriptive statistics on the share of cooperators for all subjects (first panel), the subset of subjects who were asked to report their prior beliefs (second panel) and the subset of subjects who were not asked to report their prior beliefs (third panel). For each treatment, Table 4 indicates the total number of observations, sample means and standard deviations for the share of cooperators, computed as the proportion of players out of the 10 participants who decided to invest their point in each round, treatment and session. In order to facilitate comparisons with theoretical predictions, we also report the model prediction for the share of cooperators, assuming that the expected share of cooperators is uniformly distributed and assuming the model prediction for the share of cooperators when the empirical distribution of the expected share of cooperators is employed. As predicted by the model, the average share of cooperators increases from T_j to T_{j+1} for $j = 1,2,3$. However, for all treatments, it exceeds the share predicted by the model either when prior beliefs are assumed to be uniformly distributed or when the empirical distribution of prior beliefs is employed. Note, however, that in the former case, the gap between the observed average share of cooperators and theoretical predictions significantly decreases for T_2, T_3, and T_4 (by definition, the distribution of prior beliefs does not affect theoretical predictions for T_1).

Table 4: Share of Cooperators (Descriptive Statistics)

	Number of observations	Model prediction (Prior beliefs uniformly distributed)	Model prediction (Prior beliefs empirically distributed)	Mean	S.d.
All subjects					
T_1	96	0.000	0.000	0.072	0.085
T_2	96	0.333	0.340	0.590	0.238
T_3	96	0.490	0.599	0.811	0.181
T_4	96	0.667	0.910	0.927	0.103
Subjects who reported priors					
T_1	48	0.000	0.000	0.058	0.084
T_2	48	0.333	0.340	0.604	0.273
T_3	48	0.490	0.599	0.806	0.184
T_4	48	0.667	0.910	0.925	0.101
Subjects who did not report priors					
T_1	48	0.000	0.000	0.085	0.084
T_2	48	0.333	0.340	0.575	0.197
T_3	48	0.490	0.599	0.817	0.179
T_4	48	0.667	0.910	0.929	0.104

Note: For each treatment, there are 6 observations per session of the share of cooperators. Because we conducted 16 sessions, the total number of observations per treatment is 96.

5.2. Payoffs. Table 5 shows the sample mean and standard deviation of payoffs per treatment. As predicted by the model, the payoff is, on average, higher in T_{j+1} than in

T_j for $j = 1,2,3$ (4.327 points in T_4, 3.294 points in T_3, 1.710 points in T_2 and 1.019 points in T_1), but in all treatments the average payoff exceeds the one predicted by the model when prior beliefs are assumed to be uniformly distributed. Specifically, all players earned, on average, 1.9% more than what the model predicted in T_1, 14% more in T_2, 25.5% more in T_3 and 23.6% more in T_4. The average payoff is very close to theoretical predictions when the empirical distribution of prior beliefs is employed. Specifically, all players earned, on average, 1.9% more than what the model predicted in T_1, 13.3% more in T_2, 11.7% more in T_3 and 2% less than predicted by the model in T_4.

Table 5: Payoffs (Descriptive Statistics)

	Number of observations	Model prediction (Prior beliefs uniformly distributed)	Model prediction (Prior beliefs empirically distributed)	Mean	S.d.
All subjects					
T_1	960	1.000	1.000	1.019	0.396
T_2	960	1.500	1.509	1.710	0.966
T_3	960	2.625	2.948	3.294	1.663
T_4	960	3.500	4.414	4.327	1.266
Subjects who reported priors					
T_1	480	1.000	1.000	1.046	0.401
T_2	480	1.500	1.509	1.888	0.879
T_3	480	2.625	2.948	3.342	1.625
T_4	480	3.500	4.414	4.426	1.095
Subjects who did not report priors					
T_1	480	1.000	1.000	0.993	0.390
T_2	480	1.500	1.509	1.532	1.017
T_3	480	2.625	2.948	3.246	1.701
T_4	480	3.500	4.414	4.228	1.410

Note: For the payoffs in each treatment, there are 60 observations per session. Because we conducted 16 sessions, the total number of observations per treatment is 960.

5.3. Prior beliefs. For each treatment, Table 6 gives descriptive statistics for the subjects' prior beliefs regarding the expected share of cooperators.[16] Note that prior beliefs differ across treatments.[17] In particular, the expected share of cooperators is higher in T_{j+1} than in T_j for $j = 1,2,3$.

Table 6: Prior Belief (Descriptive Statistics)

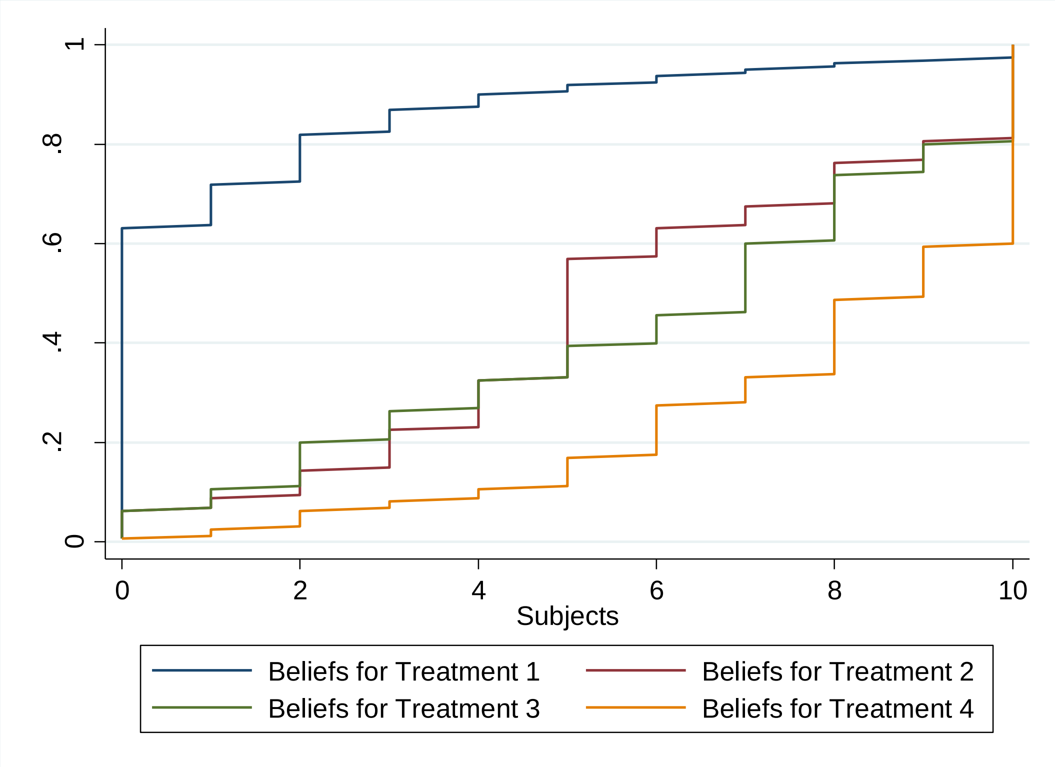
	Mean	Median	S.d.	Min	Max
Prior beliefs for T_1	1.325	0	2.423	0	10
Prior beliefs for T_2	5.713	5	3.000	0	10
Prior beliefs for T_3	6.056	7	3.200	0	10
Prior beliefs for T_4	7.863	9	2.517	0	10

Figure 1 shows the cumulative distribution function of prior beliefs regarding the expected share of cooperators across treatments. The horizontal axis measures how many participants out of a group of 10 subjects believed would contribute their point in each treatment. Let H^j denote the cumulative distribution function of prior beliefs for treatment $j = 1, 2, 3, 4$. Note that H^4 first order stochastically dominates H^3 and H^2 and that H^3 and H^2 first-order stochastically dominate H^1.

Figure 1: Cumulative Distribution Function of Prior Beliefs

[16] The reader will recall that, in randomly selected sessions, subjects were asked to estimate the expected number of cooperators before they started playing.

[17] In line with this finding, Palfrey and Rosenthal (1991) show that subjects' prior beliefs of the probability that a subject contributes is biased up with respect to an unbiased Bayes-Nash equilibrium. In the same vein, Orbell and Dawes (1991) argue that cooperators expect significantly more cooperation than do defectors.



In order to formally compare the distribution functions of prior beliefs, we conduct the nonparametric Wilcoxon matched-pair sing-rank test. The null hypothesis of this test is that the distribution of the prior belief for T_j (denoted as H^j) is equal to the distribution of the prior belief for T_k (denoted as H^k) and the alternative hypothesis is that H^j is shifted to the left of H^k. Table 7 shows the results of this test. Note that in all cases except T_2 vs T_3, the null hypothesis of equal distributions can be rejected at 0.17% of significance.[18] Therefore, there is evidence that the distribution of prior beliefs for T_1 is shifted to the left of T_2, T_3, T_4, while the distribution of prior beliefs for T_2 is shifted to the left of T_4 and the distribution of prior beliefs for T_3 is shifted to the left of T_4.

Table 7: Comparison of Distribution of Priors Beliefs

	Statistic	$p-value$
$H_0: H^1 = H^2$	385.0	0.000
$H_0: H^1 = H^3$	239.5	0.000
$H_0: H^1 = H^4$	173.5	0.000
$H_0: H^2 = H^3$	3247.0	0.098

[18] We use the Bonferroni correction to counteract the problem of d multiple simultaneous comparisons. The Bonferroni correction tests each individual hypothesis at a significance level of α/d. Therefore, if we test six hypotheses with an intended $\alpha = 0.01$, then the Bonferroni correction would test each individual hypothesis with $\alpha = 0.05/6 = 0.0017$.

$H_0: H^2 = H^4$	1447.5	0.000
$H_0: H^3 = H^4$	648.0	0.000

5.4. Reporting prior beliefs. The first panel (second panel) in Table 8 shows the results of the difference in means test of the share of cooperators (payoffs) between the sample composed by participants who reported their prior beliefs and those who did not report them. Standard errors are clustered by session. Note that it is not possible to reject the null hypothesis of equal means for the share of cooperators (payoff) in all treatments.

Table 8: Difference in Means Test (Share of Cooperators and Payoffs)

	$t - value$	$Pr(\lvert T\rvert > \lvert t\rvert)$
Share of cooperators		
All subjects	-0.06	0.950
T_1	-1.08	0.299
T_2	0.29	0.778
T_3	-0.13	0.896
T_4	-0.1	0.921
Payoff		
All subjects	0.80	0.436
T_1	1.19	0.251
T_2	1.48	0.159
T_3	0.20	0.848
T_4	0.53	0.604

Summing up, the descriptive analysis shows that: (i) The average share of cooperators increases from T_j to T_{j+1} for $j = 1,2,3$; (ii) The average share of cooperators in all treatments exceeds the share predicted by the model, but the gap is smaller once we compute theoretical predictions using reported prior beliefs rather than the uniform distribution; (iii) Payoffs are, on average, higher in T_{j+1} than in T_j for $j = 1,2,3$; (iv) Average payoffs exceed model predictions when prior beliefs are assumed to be uniformly distributed, but they are closer to theoretical predictions when the empirical distribution of prior beliefs is employed; (v) Prior beliefs differ across treatments. The average expected share of cooperators is higher in T_{j+1} than in T_j for $j = 1,2,3$. H^4 first-order stochastically dominates H^3 and H^2, and H^3 and H^2 first-order stochastically dominate H^1; (vi) The average share of cooperators and the average payoffs are not statistically different for the sessions in which subjects were asked to report their prior beliefs than they are for the sessions in which they were not asked to do so.

6. Results

In this section we formally test the main comparative static results using regression analysis. Note that in the context of perfect experimental data, where no controls are needed for identification of the causal effects of interest, the analysis is completely nonparametric, as all that it entails is a comparison of the mean outcome differences across treatment groups, and the inference can also be made nonparametric. In all cases, robust and clustered standard errors are computed by session.

6.1. Cooperation decision. In order to formally test the hypothesis that the probability of a successful collective action increases with B and s, we use the following regression model:

$$Coop_{ips} = \alpha + \beta_1 DT + \beta_2 X_{ips} + \sum_{s=1}^{16} \beta_3 D\theta_s + \varepsilon_{ips}$$

where i indexes subjects, $p = 1, 2, 3, \dots, 12$ indexes experimental rounds, and $s = 1, 2, 3, \dots, 16$ indexes experimental sessions. $Coop_{ips}$ is the dependent variable and indicates whether player i decided to invest his/her point in each session, round and treatment ($Coop_{ips} = 1$ if s/he contributes and $Coop_{ips} = 0$ if s/he does not). The explanatory variable of interest is DT, a dummy variable indicating treatment status (T_j for $j = 2, 3, 4$). In some specifications, we also include control variables. We control for individual characteristics X_{ips} (gender, age, nationality, university, whether or not the subjects have ever taken a course in game theory, whether the subjects are undergraduate or graduate students and the subjects' level of understanding of the game as measured by their answers to the quiz questions) and for fixed effects by session ($D\theta_s$). According to our theoretical predictions, we should expect $\hat{\beta}_1$ to be positive when comparing T_{j+1} with T_j for $j = 1, 2, 3$.

Columns (1), (3) and (5) in Table 9 summarize the results of regressing $Coop_{ipr}$ in each of the treatments separately without controls for all the subjects in the sample, those subjects who reported their beliefs and those who did not report them, respectively. Robust standard errors are reported in parentheses and standard errors clustered by session are shown in square brackets. In keeping with the model's prediction, the probability of cooperators in each treatment is significantly different (at a confidence level of 99% in most cases), and the coefficient associated with each treatment is positive in all cases. Indeed, note that when we compare the probability of cooperation in T_4 vs. T_1, the associated coefficient is the highest.[19] Thus, as predicted by the model, a higher value of B and/or s leads to a larger share of cooperators and, hence, to a higher probability of cooperation. Columns (2), (4) and (6) in Table 9 report the results once the entire set of controls is included. As the table shows, the results do not change in any meaningful way.

Table 9: Cooperation Decision (Regression Analysis)

	All subjects		Subjects who reported prior beliefs		Subjects who did not report prior beliefs	
	(1)	(2)	(3)	(4)	(5)	(6)
$T_1 = 0$ vs $T_2 = 1$						
$\hat{\beta}_1$	0.518***	0.518***	0.546***	0.548***	0.490***	0.489***
	(0.018)	(0.018)	(0.025)	(0.024)	(0.026)	(0.026)
	[0.044]	[0.044]	[0.080]	[0.079]	[0.040]	[0.042]
R-squared	0.303	0.315	0.336	0.362	0.271	0.287
$T_1 = 0$ vs $T_3 = 1$						
$\hat{\beta}_1$	0.740***	0.742***	0.748***	0.743***	0.731***	0.736***
	(0.015)	(0.015)	(0.021)	(0.021)	(0.022)	(0.021)
	[0.039]	[0.039]	[0.053]	[0.052]	[0.061]	[0.063]
R-squared	0.555	0.562	0.570	0.596	0.540	0.557
$T_1 = 0$ vs $T_4 = 1$						
$\hat{\beta}_1$	0.855***	0.855***	0.867***	0.864***	0.844***	0.846***
	(0.012)	(0.012)	(0.016)	(0.016)	(0.017)	(0.017)
	[0.025]	[0.025]	[0.037]	[0.037]	[0.037]	[0.037]
R-squared	0.731	0.734	0.751	0.762	0.712	0.715
$T_2 = 0$ vs $T_3 = 1$						
$\hat{\beta}_1$	0.222***	0.221***	0.202**	0.197**	0.242***	0.243***
	(0.020)	(0.020)	(0.029)	(0.028)	(0.029)	(0.028)
	[0.045]	[0.046]	[0.065]	[0.064]	[0.065]	[0.066]
R-squared	0.059	0.076	0.049	0.115	0.069	0.099

[19] The reader will recall that $\boldsymbol{T_1}$ represents a scenario of no cooperation opportunities ($\boldsymbol{B = 1.25}$ and $\boldsymbol{s = 0}$); in other words, this is the free-rider Olsonian model with one Nash equilibrium in which nobody contributes. $\boldsymbol{T_4}$ represents a scenario where the incentives to cooperate are the highest ($\boldsymbol{B = 3}$ and $\boldsymbol{s = 1.75}$).

$T_2 = 0$ vs $T_4 = 1$						
$\hat{\beta}_1$	0.338***	0.336***	0.321***	0.321***	0.354***	0.351***
	(0.018)	(0.018)	(0.025)	(0.025)	(0.025)	(0.025)
	[0.050]	[0.050]	[0.087]	[0.084]	[0.054]	[0.053]
R-squared	0.155	0.168	0.143	0.193	0.168	0.181
$T_3 = 0$ vs $T_4 = 1$						
$\hat{\beta}_1$	0.116***	0.115***	0.119**	0.124**	0.112**	0.111**
	(0.015)	(0.015)	(0.022)	(0.021)	(0.021)	(0.021)
	[0.025]	[0.025]	[0.034]	[0.037]	[0.037]	[0.036]
R-squared	0.029	0.043	0.030	0.108	0.029	0.093
Controls	No	Yes	No	Yes	No	Yes
Number of observations	1920	1920	960	960	960	960

Note: * significant at 10%; ** significant at 5%; *** significant at 1% (using standard errors clustered by session). Robust standard errors are shown in parentheses and standard errors clustered by session are shown in square brackets. Controls: (i) Individual characteristics X_{ipr}: gender, age, nationality, university, whether the subjects have ever taken a course in game theory and whether they are undergraduate or graduate students; (ii) Level of understanding of the game measured by the subjects' correct answers to the quiz questions; and (iii) Fixed effects by session $D\theta_r$.

6.2. Payoffs. In order to formally test the hypothesis that the average payoff of a player increases with B and/or s, we use the following regression model:

$$Payoff_{ips} = \gamma + \delta_1 DT + \delta_2 X_{ips} + \sum_{s=1}^{11} \delta_3 D\theta_s + \varepsilon_{ips}$$

The dependent variable $Payoff_{ips}$ is the payoff denominated in points obtained by subject i in round p and session s. The regressors are the same as in the model for the share of cooperators. The explanatory variable of interest is DT, a dummy variable indicating treatment status (T_j for $j = 2, 3, 4$). According to our theoretical predictions, we should expect $\hat{\delta}_1$ to be positive when comparing T_{j+1} with T_j for $j = 1,2,3$.

Table 10 summarizes the results. The corresponding clustered standard errors are shown in square brackets. As predicted by our model, the payoff in each treatment is significantly different, and the coefficient associated with each treatment is positive. Hence, operating under the parameters in T_{j+1} rather than in T_j for $j = 1, 2, 3$ induces a positive and statistically significant effect on the payoff.

Table 10: Payoffs (Regression Analysis)

	All subjects		Subjects who reported prior beliefs		Subjects who did not report prior beliefs	
	(1)	(2)	(3)	(4)	(5)	(6)
$T_1 = 0$ vs $T_2 = 1$						
$\hat{\delta}_1$	0.691***	0.691***	0.842***	0.847***	0.539**	0.539**

	(0.034)	(0.033)	(0.044)	(0.044)	(0.050)	(0.049)
	[0.120]	[0.120]	[0.155]	[0.153]	[0.177]	[0.176]
R-squared	0.180	0.206	0.276	0.297	0.109	0.163
$T_1 = 0$ vs $T_3 = 1$						
$\hat{\delta}_1$	2.275***	2.267***	2.296***	2.282***	2.253***	2.245***
	(0.055)	(0.055)	(0.076)	(0.074)	(0.080)	(0.080)
	[0.251]	[0.255]	[0.419]	[0.422]	[0.306]	[0.315]
R-squared	0.470	0.482	0.485	0.534	0.455	0.475
$T_1 = 0$ vs $T_4 = 1$						
$\hat{\delta}_1$	3.308***	3.307***	3.380***	3.379***	3.235***	3.238***
	(0.043)	(0.042)	(0.053)	(0.051)	(0.067)	(0.065)
	[0.187]	[0.187]	[0.303]	[0.304]	[0.237]	[0.236]
R-squared	0.757	0.771	0.808	0.826	0.710	0.727
$T_2 = 0$ vs $T_3 = 1$						
$\hat{\delta}_1$	1.584***	1.581***	1.454***	1.445***	1.714***	1.713***
	(0.062)	(0.061)	(0.084)	(0.080)	(0.090)	(0.091)
	[0.240]	[0.241]	[0.345]	[0.348]	[0.352]	[0.351]
R-squared	0.253	0.284	0.237	0.333	0.273	0.286
$T_2 = 0$ vs $T_4 = 1$						
$\hat{\delta}_1$	2.617***	2.616***	2.538***	2.542***	2.696***	2.706***
	(0.051)	(0.050)	(0.064)	(0.060)	(0.079)	(0.080)
	[0.204]	[0.204]	[0.271]	[0.264]	[0.322]	[0.323]
R-squared	0.575	0.607	0.621	0.674	0.546	0.557
$T_3 = 0$ vs $T_4 = 1$						
$\hat{\delta}_1$	1.033***	1.036***	1.084***	1.099***	0.982**	0.979**
	(0.067)	(0.066)	(0.089)	(0.081)	(0.101)	(0.099)
	[0.203]	[0.207]	[0.257]	[0.257]	[0.331]	[0.332]
R-squared	0.109	0.174	0.133	0.307	0.090	0.147
Controls	No	Yes	No	Yes	No	Yes
Number of Observations	1920	1920	960	960	960	960

Note: * significant at 10%; ** significant at 5%; *** significant at 1% (using standard errors clustered by session). Robust standard errors are shown in parentheses and standard errors clustered by sessions are shown in square brackets. Controls: see the note for Table 9.

As a robustness check, we repeated the estimations in Tables 9 and 10 while introducing two new explanatory variables, namely DR and I. DR is a dummy variable that indicates whether the collective action in the previous round was successful or not, and I is the number of players in the same group who decided to invest in the previous round. These variables capture the possibility that subjects decide to cooperate in a treatment just because either the collective action was successful in the previous round or the number of investors was relatively high. The results do not change in any meaningful way. The coefficients associated with each treatment are still significantly different and positive.

6.3. Prior beliefs. In order to determine if asking subjects to reveal their prior beliefs biased their decisions during the game, we performed a test of equality of the regression coefficients. Panel 1 of Table 11 summarizes the results of a test whose null hypothesis is that the effects of each treatment on the share of cooperators are the same for the subjects who reported their prior beliefs (β_1) and those who did not report them (β_1^*). Standard errors are clustered by sessions. In all cases the null hypothesis of equal coefficients cannot be rejected. Analogously, Panel 2 of Table 11 summarizes the results of a test whose null hypothesis is that the effects of each treatment on the payoffs are identical for subjects who reported their prior beliefs (δ_1) and subjects who did not report them (δ_1^*). Standard errors are clustered by sessions. In all cases the null hypothesis of equal coefficients cannot be rejected[20]. Thus, it is possible to confirm that asking subjects to reveal their prior beliefs before the game started did not introduce any bias in their decisions during the game.

Table 11: Reporting Versus Not Reporting Prior Beliefs
(Difference in Average Treatment Effects)

	F (1,7)	$Pr > F$
Share of cooperators (H_0: $\beta_1 = \beta_1^*$)		
$T_1 = 0$ vs $T_2 = 1$	1.99	0.201
$T_1 = 0$ vs $T_3 = 1$	0.08	0.791
$T_1 = 0$ vs $T_4 = 1$	0.39	0.551
$T_2 = 0$ vs $T_3 = 1$	0.36	0.568
$T_2 = 0$ vs $T_4 = 1$	0.38	0.558
$T_3 = 0$ vs $T_4 = 1$	0.03	0.871
Profit (H_0: $\delta_1 = \delta_1^*$)		
$T_1 = 0$ vs $T_2 = 1$	2.92	0.131
$T_1 = 0$ vs $T_3 = 1$	0.02	0.892
$T_1 = 0$ vs $T_4 = 1$	0.37	0.56
$T_2 = 0$ vs $T_3 = 1$	0.55	0.484
$T_2 = 0$ vs $T_4 = 1$	0.24	0.639
$T_3 = 0$ vs $T_4 = 1$	0.09	0.767

Note: $F(1,7)$ indicates the F statistic with 1 degree of freedom in the numerator and 7 degrees of freedom in the denominator. $Pr > F$ indicates the significance level of each test.

[20] The results of the tests hold when we add controls in the regression of the share of cooperators (and payoffs) in each of the treatments. We do not report the corresponding F statistics for the sake of simplicity.

Summing up, the regression analysis produces robust support for the main theoretical comparative statics predictions. Increases in B and/or s have a significant positive effect on the share of cooperators and, hence, on the probability of a successful collective action, as well as on the players' payoffs.[21] The effects are statistically significant whether or not we include controls for individual characteristics, level of understanding of the game or fixed effects by session. Asking subjects to report their prior beliefs before the game started did not have any significant effect on their decisions. Introducing control variables for whether the collective action in the previous round was successful or not and for the number of players in the same group who decided to invest in the previous round does not change the results in any meaningful way either.

7. Exploring a Decomposition of Changes in $\hat{\alpha}$

In this section we decompose a change in $\hat{\boldsymbol{\alpha}}$ in a 'belief effect' and a 'range of cooperation effect'. The idea is to learn about the mechanism that induce more cooperation when $\hat{\boldsymbol{\alpha}}$ decreases.

Table 12 shows the average share of cooperators as well as the predicted share of cooperators for each treatment, both using a uniform distribution for all treatments, and the empirical distribution of prior beliefs for each treatment.

Table 12: Model Prediction of the Share of Cooperators

	Empirical probability of cooperation	Model prediction (Prior beliefs, uniformly distributed)	Model prediction (Prior beliefs empirically, distributed for T1)	Model prediction (Prior beliefs, empirically distributed for T2)	Model prediction (Prior beliefs, empirically distributed for T3)	Model prediction (Prior beliefs, empirically distributed for T4)
T_1	0.072	0.000	**0.000**	0.000	0.000	0.000
T_2	0.590	0.333	0.054	**0.340**	0.448	0.687
T_3	0.811	0.490	0.079	0.424	**0.599**	0.819
T_4	0.927	0.667	0.121	0.742	0.717	**0.910**

[21] More cooperative prior beliefs with the same B and/or s (i.e., within a treatment) do not, however, induce more cooperation.

As discussed in section 5.3., the distribution of prior beliefs is not the same in every treatment. As the benefit of cooperation increases, subjects tend to raise their assessments of the share of cooperators. The effect of these changes in theoretical predictions can be observed in Table 12. Except for $\boldsymbol{T_1}$, for which theoretical predictions do not change with the distribution of prior beliefs, the predicted share of cooperators increases for all of the rest of the treatments, as we employ the prior beliefs associated with a treatment with a lower $\hat{\boldsymbol{\alpha}}$.[22] For example, for $\boldsymbol{T_2}$, if we use the priors of $\boldsymbol{T_1}$, the predicted share of cooperators is 0.054, while it is 0.340 with the priors of $\boldsymbol{T_2}$, 0.448 with the priors of $\boldsymbol{T_3}$ and 0.687 with the priors of $\boldsymbol{T_4}$. This suggests that we can decompose a change in the predicted share of cooperators into two analytically different effects: a "belief effect" that captures the change in prior beliefs, and a "range of cooperation effect" that captures the change in the range of prior beliefs that induced greater cooperation.

More technically, the distribution of the expected share of cooperators $\boldsymbol{H}$ is not independent of $\hat{\boldsymbol{\alpha}}$. Although this does not affect the sign of the comparative statics of the model, it is interesting to explore what fraction of the change in the predicted share of cooperators can be attributed to a change in prior beliefs and what fraction can be attributed to a change in the range of prior beliefs that induce greater cooperation. Thus, we are now interested in detecting the mechanism by which a decrease in $\hat{\boldsymbol{\alpha}}$ leads to a higher probability of a successful collective action.

Let H^j denote the cumulative distribution function of the expected share of cooperators for treatment T_j, and Pr^j the probability of a successful collective action in treatment T_j. Then:

$$Pr(T_j) - Pr(T_k) = H^k(\hat{\alpha}_k) - H^j(\hat{\alpha}_j) = \left[H^k(\hat{\alpha}_k) - H^j(\hat{\alpha}_k)\right] + \left[H^j(\hat{\alpha}_k) - H^j(\hat{\alpha}_j)\right]$$

Define $\Delta^I(T_k \rightarrow T_j) = \frac{H^k(\hat{\alpha}_k) - H^j(\hat{\alpha}_k)}{H^k(\hat{\alpha}_k) - H^j(\hat{\alpha}_j)}$. $\Delta^I(T_k \rightarrow T_j)$ as the proportion of the change attributed to a change in the distribution of expected cooperators. Naturally, $1 - \Delta^I(T_k \rightarrow T_j)$ is the proportion of the change in the probability of a successful collective action due to a change in the range of prior beliefs that induce greater

[22] The reader will recall from section 2.2 that a lower $\hat{\alpha}$ is associated with a higher B and/or s; in other words: $\frac{\partial \hat{\alpha}}{\partial B} < 0$, $\frac{\partial \hat{\alpha}}{\partial s} < 0$.

cooperation. Table 13 shows the decomposition of a change in the predicted share of cooperators into the belief and range of cooperation effects.

Table 13: Decomposition of Changes in $\hat{\alpha}$: I

	$H^k(\hat{\alpha}_k) - H^j(\hat{\alpha}_j)$	$\Delta^I(T_k \to T_j)$	$1 - \Delta^I(T_k \to T_j)$
$T_1 \to T_2$	0.340	0.00	1.00
$T_1 \to T_3$	0.599	0.00	1.00
$T_1 \to T_4$	0.910	0.00	1.00
$T_2 \to T_3$	0.259	0.42	0.58
$T_2 \to T_4$	0.570	0.61	0.49
$T_3 \to T_4$	0.311	0.71	0.29

To some extent, this decomposition is arbitrary, in the sense that we can first vary H and then $\hat{\alpha}$ or the other way around. Formally, we can also decompose $Pr(T_j) - Pr(T_k)$ as follows:

$$Pr(T_j) - Pr(T_k) = H^k(\hat{\alpha}_k) - H^j(\hat{\alpha}_j) = [H^k(\hat{\alpha}_k) - H^k(\hat{\alpha}_j)] + [H^k(\hat{\alpha}_j) - H^j(\hat{\alpha}_j)]$$

and define the proportion of the change attributed to a change in the distribution of expected cooperators by $\Delta^{II}(T_k \to T_j) = \frac{H^k(\hat{\alpha}_j) - H^j(\hat{\alpha}_j)}{H^k(\hat{\alpha}_k) - H^j(\hat{\alpha}_j)}$. Table 13 shows this decomposition.

Table 14: Decomposition of Changes in $\hat{\alpha}$: II

	$H^k(\hat{\alpha}_k) - H^j(\hat{\alpha}_j)$	$\Delta^{II}(T_k \to T_j)$	$1 - \Delta^{II}(T_k \to T_j)$
$T_1 \to T_2$	0.340	0.159	0.841
$T_1 \to T_3$	0.599	0.132	0.868
$T_1 \to T_4$	0.910	0.133	0.867
$T_2 \to T_3$	0.259	0.324	0.676
$T_2 \to T_4$	0.570	0.705	0.295
$T_3 \to T_4$	0.311	0.379	0.621

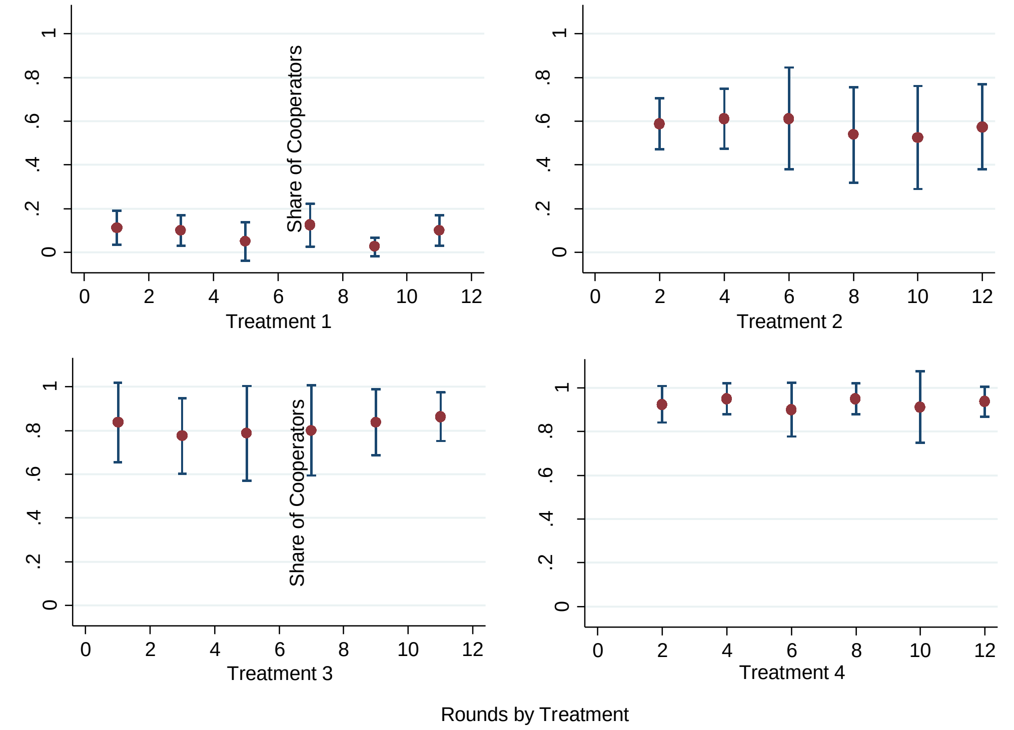
Except when we move from T_3 to T_4, both decompositions assign similar proportions to both effects. When the starting point is T_1, both decompositions assign a very high proportion of the change to the range of cooperation effect (at least 84%). When the starting point is T_2 and we move to T_3 (T_4), the first and second decompositions attribute 42% and 32.4% (61% and 70%) of the change to a switch in beliefs, respectively.

A potential concern about these decompositions is that they rely on the empirical distribution of prior beliefs that were reported by the subjects before the rounds began. It is possible that these prior beliefs evolve as the experiment proceeds and that subjects learn from previous rounds. However, we do not observe any temporal pattern in the data. For example, Figure 2 shows the mean share of cooperators per round across treatments for all the subjects in the sample (first panel), the subjects who reported their beliefs (second panel) and the subjects who were not required to report their beliefs (third panel). The mean share of cooperators fluctuates without forming any clear pattern.[23]

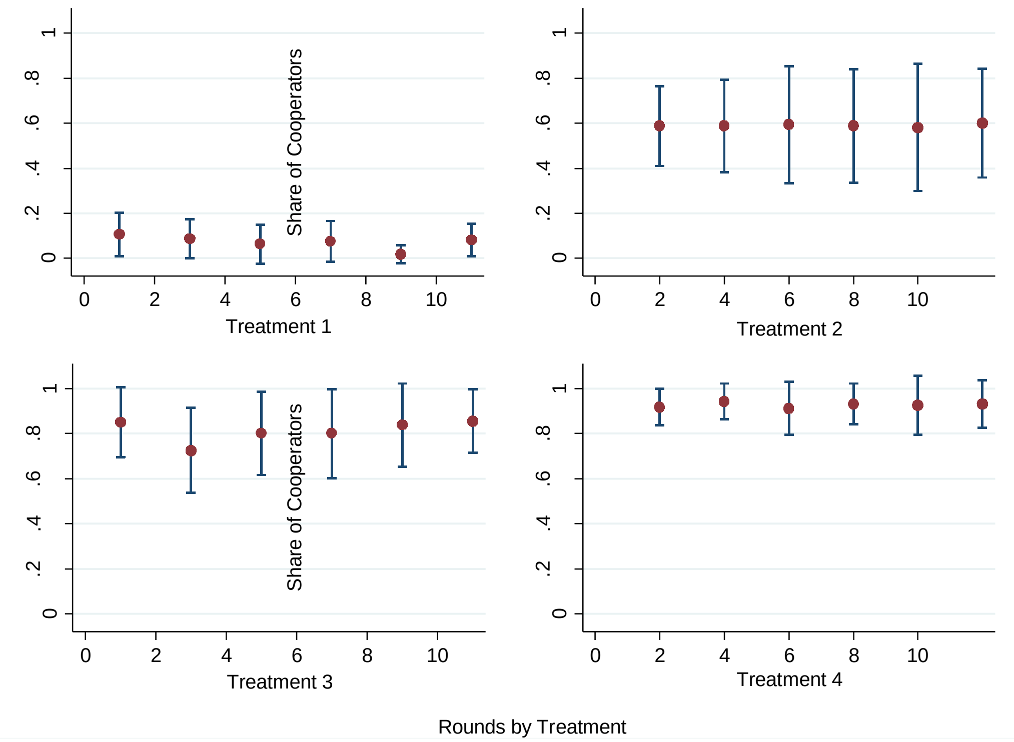
Figure 2: Share of Cooperators

All subjects

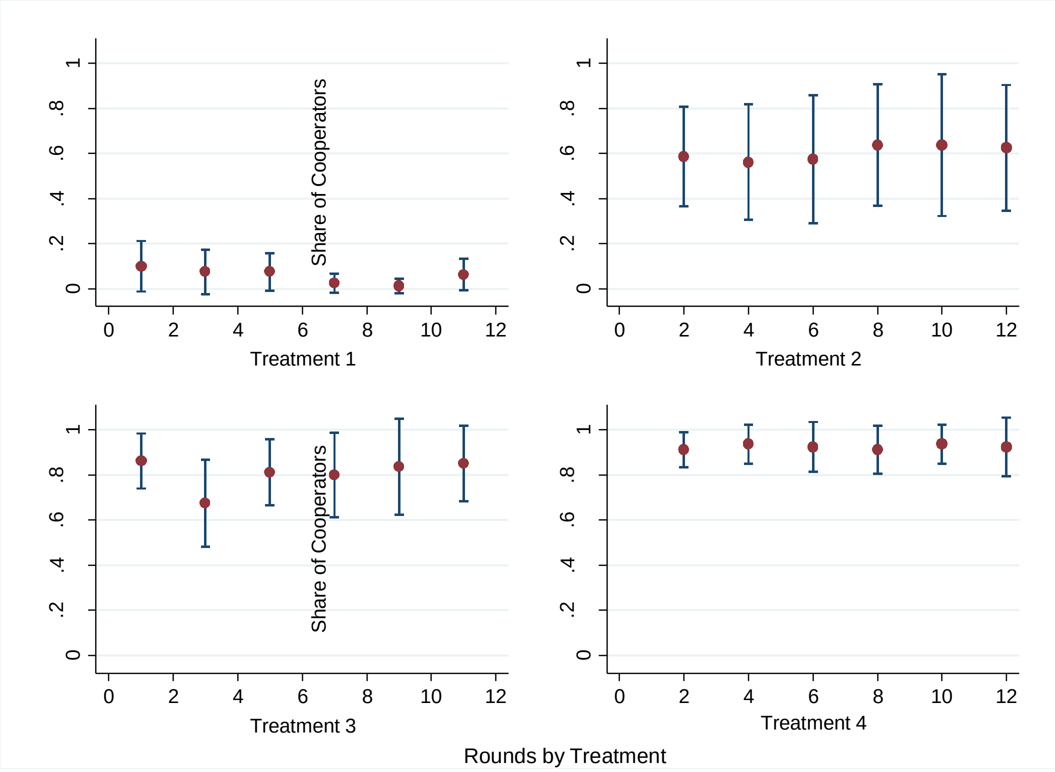
[23] Temporal patterns have been observed in public good games even in a stranger situation, i.e., when subjects face different group members in each repetition of the game. Usually, these patterns are interpreted as evidence of the existence of conditional cooperators (see, for example, Keser and van Winden 2000). In treatment 1, which has only one Nash equilibrium, the existence of a proportion of conditional cooperators would imply a decay in the share of cooperators in later rounds. However, as Figure 2 shows, we do not observe such behavior. In treatments 2 to 4, which have multiple Nash equilibria, it is not clear the temporal pattern implied by conditional cooperators. Indeed, in these treatments, players face a coordination problem, which mean that the best response function is to cooperate only if the share of cooperators is above some tipping point. It might be the case that initially conditional cooperators cooperate even when the expected share of cooperators is below the tipping point. This would induce higher cooperation than predicted by the model and no temporal pattern. Indeed, as Table 4 shows, there is a higher share of cooperators than predicted by the model (especially when we use a uniform distribution for prior beliefs). As Figure 2 indicates, for treatments 2 to 4, there is no temporal patterns in the mean share of cooperators.



Subjects who reported prior beliefs



Subjects who did not report prior beliefs



Note: Red diamonds denote the average values of the variable per treatment/round within treatment. Blue bars indicate one standard deviation from the mean, calculated in standard form.

Summing up, there are two mechanisms operating simultaneously that induce a higher predicted share of cooperators. First, as $\hat{\alpha}$ decreases, subjects raise their assessments of the expected share of cooperators (the belief effect). Second, given any distribution of the assessments, a lower $\hat{\alpha}$ raises the assessments that induce subjects to contribute (the range of cooperation effect). Except when we move from T_3 to T_4, both decompositions lead to similar results.

8. Conclusions

We have conducted a laboratory experiment in order to test the main implications of the stability-sets methods as applied to collective action games. We have found strong support for the key comparative static predictions of the theory. As we increase the payoff of a successful collective action accruing to all players (B) and only to those who contribute (s), the share of cooperators and payoffs both increase. As in many other laboratory experiments, we found that subjects behave more cooperatively than is predicted by the theory. But we have also shown that the gap between theoretical predictions and observed behavior narrows significantly when we refine the theory by allowing for a distribution of prior beliefs that varies with the parameters of the model. Overall, the experiment indicates that the stability-sets method could be a very useful tool for studying games with multiple equilibria.

The experiment also suggests a refinement of the theory. We found that, as the range of cooperation increases, subjects upgrade their prior beliefs relating to the expected share of cooperators. We have shown that if the new distribution of prior beliefs first-order stochastically dominates the preceding one, then the signs of the comparative static derivatives are not affected, but all effects are magnified. For practical purposes, this refinement improves the power of the theory to predict the observed behavior. Analytically, it allows us to decompose the mechanism that produces cooperation into a ”belief effect” and a ”range of cooperation effect”. Using our experiment, we have computed these decompositions and have found evidence of the presence of both effects. This may have interesting political economy implications. For example, a policy change that affects the payoffs of a collective action game can produce a bigger change in the likelihood of cooperation than what we would expect if we did not take the fact that agents update the distribution of prior beliefs into account.

Understanding the logic of collective action is crucial in terms of political economy. Explicitly or implicitly, collective action is a core component of many models of political influence, political representation and coalition formation. A new approach to collective action can produce significant impacts in terms of the way that we approach those topics. To illustrate this point, consider the following examples. In the standard common agency model of lobbying (Dixit, Grossman and Helpman, 1997 and Grossman and Helpman, 2000), it is assumed that groups are either organized (meaning that the group has solved the collective action problem and can lobby to advance its members’ common interests) or unorganized. The stability-sets approach can serve as the basis for an assessment of the likelihood that a group is organized as a function of structural parameters that characterize the collective action problem of group organization. Thus, by combining the common agency model of lobbying with the stability-sets approach to collective action, we can build a more accurate theory of political influence. Another interesting example is provided by Acemoglu's and Robinson's model of political regime determination (Acemoglu and Robinson, 2006). This is a dynamic model in which, with some exogenous probability, in every period a group with no de jure political power can organize and obtain de facto political power. Again, combining this model with the stability-sets approach to collective action can help us to refine the theory of political transitions.

References

Acemoglu, Daron, and James A. Robinson (2006). *Economic Origins of Dictatorship and Democracy*, Cambridge, Cambridge University Press.

Anderson, S., Goeree, J., and Holt, C. (1998). A theoretical analysis of altruism and decision error in public goods games. *Journal of Public Economics*, vol. 70, pp. 297–323.

Au W (2004). Criticality and environmental uncertainty in step-level public goods dilemmas. *Group Dynamics*, vol. 8(1), pp. 40–61.

Bagnoli, M, and Lipman, B.L. (1989). Provision of public goods: fully implementing the core through private contributions. *Review of Economic Studies*, vol. 56(4), pp. 583–601.

Banerjee, P., Chakravarty, S., and Mohanty, R. (2012). Equilibrium play and learning in a stochastic public goods experiment. *Journal of Quantitative Economics*, vol. 10(2), pp. 60-77.

Barr, A., Mugisha, F., Serneels, P., and Zeitlin, A. (2012). Information and collective action in the community monitoring of schools: Field and lab experimental evidence from Uganda (unpublished manuscript).

Barrett S (2013). Climate treaties and approaching catastrophes. *Journal of Environmental Economic Management*, vol. 66(2), pp. 235–250.

Battalio, R., Samuelson, L., and Van Huyck, J. (2001). Optimization incentives and coordination failure in laboratory stag hunt games. *Econometrica*, vol. 69(3), pp. 749-764.

Boix, C., and Milan Svolik. (2009). The Foundations of Limited Authoritarian Government: Institutions and Power-Sharing in Dictatorships. Mimeo.

Cadsby, C. B., and Maynes, E. (1999). Voluntary provision of threshold public goods with continuous contributions: Experimental evidence. *Journal of Public Economics*, vol. 71(1), pp. 53-73.

Cardenas, J. C. (2009). Experiments in environment and development. Working paper, Bogotá, Universidad de los Andes, School of Economics.

Carlsson, H., and Van Damme, E. (1993). Global games and equilibrium selection. *Econometrica: Journal of the Econometric Society*, vol. 61, pp. 989-1018.

Cason, T. N., Saijo, T., Yamato, T., and Yokotani, K. (2004). Non-excludable public good experiments. *Games and Economic Behavior*, vol. 49(1), pp. 81-102.

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

Cherry, T. L., Kroll, S., and Shogren, J. F. (2005). The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab. *Journal of Economic Behavior & Organization*, vol. 57(3), pp. 357-365.

Croson, R. (1996). Partners and strangers revisited. *Economics Letters,* vol. pp. 53, 25-32.

Croson, R. (2007). Theories of commitment, altruism and reciprocity: Evidence from linear public goods games. *Economic Inquiry*, vol. 45, pp. 199–216.

Croson, R.T.A. and Marks, M.B. (2000). Step Returns in Threshold Public Goods: A Meta- and Experimental Analysis. *Experimental Economics,* vol. 2, pp. 239-259.

Davis, D. D., and Holt, C. A. (1993). *Experimental Economics*. Princeton, New Jersey: Princeton University Press.

Dickinson, David L. (1998). The Voluntary Contributions Mechanism with Uncertain Group Payoffs. *Journal of Economic Behavior and Organization*, vol. 35, pp. 517-533.

Dixit, A., Grossman, G., and Helpman, E. (1997). Common agency and coordination: General theory and application to government policy making. *Journal of Political Economy*, vol. 105, pp. 752-769.

Edmond, Chris. (2008). Information Manipulation, Coordination and Regime Change. Mimeo.

Egorov, Georgy, and Konstantin Sonin. (2014). Incumbency Advantage in Non-Democracies. National Bureau of Economic Research. Working Paper, No. 20519, p. 1-43.

Egorov, Georgy, Sergei Guriev, and Konstantin Sonin (2009). Why Resource-poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data. *American Political Science Review*, vol. 103, pp. 645-68.

Fischbacher, U. (2007). Z-tree 2.1: Zurich Toolbox for Readymade Economic Experiment. *Experimental Economics*, vol. 10, pp. 171-178.

Fischbacher, U., and Gächter, S. (2006). Heterogeneous social preferences and the dynamics of free riding in public goods. CeDEx Working Paper No. 2006-01, University of Nottingham.

Fischbacher, U., Gächter, S., and Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, vol. 71(3), pp. 397-404.

Frankel, D. M., Morris, S., and Pauzner, A. (2003). Equilibrium selection in global games with strategic complementarities. *Journal of Economic Theory*, vol. 108(1), pp. 1-44.

Golman, R., and Page, S. E. (2010). Individual and cultural learning in stag hunt games with multiple actions. *Journal of Economic Behavior & Organization*, vol. 73(3), pp. 359-376.

Grossman, G., and Helpman, E. (2000). *Special Interest Politics*. Cambridge, Massachusetts and London, UK: MIT Press.

Gustafsson, M., A. Biel, and T. Garling. (2000). Egoism Bias in Social Dilemmas with Resource Uncertainty. *Group Processes and Intergroup Relations*, vol. 3, pp. 351–65.

Harsanyi, J., and Selten, R. (1988). *A General Theory of Equilibrium Selection in Games*, Cambridge, Massachusetts: MIT Press.

Haruvy, E., and Stahl, D. (2000). Initial play and equilibrium selection in symmetric normal-form games. Working paper.

Hichri, W. (2005). The individual behavior in a public goods game. *International Review on Public and Nonprofit Marketing*, vol. 2(1), pp. 59-71.

Isaac, M., Schmidtz, D., and Walker, J. (1989). The Assurance Problem in a Laboratory Market. *Public Choice*, vol. 62, 217–236.

Isaac, R. M., Walker, J. M., and Williams, A. W. (1994). Group size and the voluntary provision of public goods: Experimental evidence utilizing large groups. *Journal of Public Economics*, vol. 54(1), pp. 1-36.

Iwasaki, A., Imura, S., Oda, S. H., and Ueda, K. (2003). Accidental initial outcomes and learning in experimental games with multiple equilibria. International ESA meeting, Pittsburgh, Pennsylvnia, United States, 19-22 June 2003.

Keser, C. and van Winden, F. (2000). Conditional cooperation and voluntary contributions to public goods. *Scandinavian Journal of Economics*, vol. 102, pp. 23-39.

Ledyard, J. (1995). Public goods: A survey of experimental research, *Handbook of Experimental Economics* (eds. John Kagel and Alvin Roth), Princeton, New Jersey: Princeton University Press.

Marwell, G., and R. Ames (1981). Economists free ride, does anyone else? Experiments on the provision of public goods, IV. *Journal of Public Economics*, vol. 15, pp. 295–310.

McBride, M (2006). Discrete public goods under threshold uncertainty. *Journal of Public Economics*, vol. 90, pp. 1181–1199.

McKelvey, R.D., and Palfrey, T.R. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior*, vol. (10), pp. 6–38.

Medina, L. F. (2007). *A Unified Theory of Collective Action and Social Change.* Ann Arbor: University of Michigan Press.

Medina, L. F. (2013). The Analytical Foundations of Collective Action Theory: A Survey of Some Recent Developments. *Annual Review of Political Science*, vol. 16, pp. 259-283.

Morris, S. and Shin, H. S. (2001). Global Games: Theory and Applications. Working paper, Yale University.

Morris, S., and Shin, H. S. (1998). Unique equilibrium in a model of self-fulfilling currency attacks. *The American Economic Review*, vol. 88, pp. 587-597.

Morris, S., Shin, H. S. and Yildiz, M. (2015). Common belief foundations of global games. Working Paper. Levine's Bibliography No. 122247000000001638, UCLA Department of Economics. Available at: http://goo.gl/yWWFKA

Neugebauer, T., Perote, J., Schmidt, U. and Loos, M. (2005). Self-biased conditional cooperation: On the decline of cooperation in repeated public goods experiments. Mimeo, University of Hanover.

Neugebauer, T., Poulsen, A., and Schram, A. (2008). Fairness and reciprocity in the hawk–dove game. *Journal of Economic Behavior & Organization*, vol. 66(2), pp. 243-250.

Nitzan S, Romano RE (1990). Private provision of a discrete public good with uncertain cost. *Journal of Public Economics*, vol. 42(3), pp. 357–370.

Offerman, T. (1997). *Beliefs and Decision Rules in Public Goods Games: Theory and Experiments*, Dordrecht, the Netherlands: Kluwer Academic Publishers.

Offerman, T., Sonnemans, J., and Schram, A. (1996). Value Orientations, Expectations and Voluntary Contributions to Public Goods. *Economic Journal*, vol. 106, pp. 817-845.

Offerman, T., Sonnemans, J., and Schram, A. (2001). Expectation Formation in Step-Level Public Good Games. *Economic Inquiry*, vol. 39, pp. 250-269.

Olson, M. (1965). *The Logic of Collective Action: Public Goods and the Theory of Groups, Cambridge*, Massachusetts: Harvard University Press.

Oprea, R., Henwood, K., and Friedman, D. (2011). Separating the hawks from the doves: Evidence from continuous time laboratory games. *Journal of Economic Theory*, vol. 146(6), pp. 2206-2225.

Orbell, J., and Dawes, R. M. (1991). A ‘cognitive miser’ theory of cooperators’ advantage. *American Political Science Review*, vol. 85, pp. 515–528.

Ostrom, E. (1998). A behavioral approach to the rational-choice theory of collective action. *American Political Science Review*, vol. 92(1), pp. 1-22.

Palfrey, T. R., and Rosenthal H (1984). Participation and the provision of public goods: a strategic analysis. *Journal of Public Economics*, vol. 24, pp. 171–193.

Palfrey, T., and Rosenthal, H. (1991). Testing game-theoretic models of free riding: New evidence on probability bias and learning. In T. Palfrey (ed.), *Laboratory Research in Political Economy.* Ann Arbor, Michigan: University of Michigan Press.

Persson, T., and Tabellini, G. (2009). Democratic Capital: The Nexus of Political and Economic Change. *American Economic Journal: Macroeconomics*, vol. 1, pp. 88-126.

Saunders, D. K. (2012). Coordinating contributions to a threshold public good in a noisy best response framework. Working Paper, University of California, Santa Barbara.

Schelling, T. C. (1978). *Micromotives and Macrobehavior*. New York: W. W. Norton.

Schmitt, P., Swope, K., and Walker, J. (2000). Collective action with incomplete commitment: Experimental evidence. *Southern Economic Journal*, vol. 66, pp. 829-854.

Schotter, A., and Trevino, I. (2014). Belief elicitation in the laboratory. *Annual Review of Economics*, vol. 6(1), 103-128.

Shadmehr, M., and Bernhardt, D. (2011). Collective action with uncertain payoffs: coordination, public signals, and punishment dilemmas. *American Political Science Review*, vol. 105(04), pp. 829-851.

Suleiman R (1997). Provision of step-level public goods under uncertainty a theoretical analysis. *Rationality and Society,* vol. 9(2), pp. 163–187.

Van Huyck, J. B., Battalio, R. C., and Beil, R. O. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *The American Economic Review*, vol. 80, pp. 234-248.

Van Huyck, J. B., Battalio, R. C., and Beil, R. O. (1991). Strategic uncertainty, equilibrium selection, and coordination failure in average opinion games. *The Quarterly Journal of Economics*, vol. 106, pp. 885-910.

Wit A, Wilke A (1998). Public good provision under environmental and social uncertainty. *European Journal of Social Psychology*, vol. 28(2), pp. 249–256.

Online Appendix 1: Proposition 1

In this appendix we present the proof for proposition 1.

Olson's model: The expected payoff for players *i* is given by:

$$U_i(\alpha_i, \alpha_{-i}) = \alpha_i \sum_{k=0}^{N-1} \left[B\left(\frac{k+1}{N}\right) - c\right] P(k,i) + (1-\alpha_{-i}) \sum_{k=0}^{N-1} B\left(\frac{k}{N}\right) - P(k,i)$$

$$= \sum_{k=0}^{N-1} \left[\alpha_i B\left(\frac{k+1}{N}\right) - \alpha_i c + (1-\alpha_{-i}) B\left(\frac{k}{N}\right)\right] P(k,i)$$

$$= \alpha_i \left(\frac{B}{N} - c\right) \sum_{k=0}^{N-1} P(k,i) + \sum_{k=0}^{N-1} B\left(\frac{k}{N}\right) P(k,i)$$

The second term does not depend on α_i. When $\frac{B}{N} > c$ $\left(\frac{B}{N} < c\right)$, the first term adopts a maximum of $\alpha_i = 1$ $(\alpha_i = 0)$. Therefore, if $N < \frac{B}{c}$, the unique Nash equilibrium is C_i for all *i*, while, if $N > \frac{B}{c}$, the unique Nash equilibrium is D_i for all *i*. ■

Schelling's model: A Nash equilibrium is a profile α such that, for all $i = 1, \ldots, N$, one of the following conditions must hold:

$$\sum_{k=0}^{N-1} \left[\frac{B + s(k+1)}{N} - c\right] P(k,i) \geq 0 \quad \text{and} \quad \alpha_i = 1 \tag{3}$$

$$\sum_{k=0}^{N-1} \left[\frac{B + s(k+1)}{N} - c\right] P(k,i) \leq 0 \quad \text{and} \quad \alpha_i = 0 \tag{4}$$

$$\sum_{k=0}^{N-1} \left[\frac{B + s(k+1)}{N} - c\right] P(k,i) = 0 \quad \text{and} \quad \alpha_i \in (0,1) \tag{5}$$

Where $P(k,i) = \sum_{a_{-i} \in S(k,i)} \prod_{j \neq i} \alpha_j^{a_j} (1-\alpha_j)^{1-a_j}$ and $S(k,i) = \{a_{-i} : \sum_{j \neq i} a_j = k\}$.

Lemma 1: If $\alpha_i = 1$ and $\alpha_h = 0$, then $P(k,i) = P(k+1,h)$. **Proof**: Since players' strategies are not correlated, the probability that $k+1$ players cooperate when we exclude h is equal to the probability that k players cooperate when we exclude i and h times the probability that *i* cooperates plus the probability that $k+1$ players cooperate excluding i and h times the probability that *i* does not cooperate. Formally,

$$P(k+1,h) = \Pr\left(\sum\nolimits_{j\neq h} a_j = k+1\right)$$

$$= \Pr\left(\sum\nolimits_{j\neq i,h} a_j = k\right)\alpha_i + \Pr\left(\sum\nolimits_{j\neq i,h} a_j = k+1\right)(1-\alpha_i)$$

$$= \Pr\left(\sum\nolimits_{j\neq i,h} a_j = k\right)$$

$$= \Pr\left(\sum\nolimits_{j\neq i} a_j = k-1\right)\alpha_h + \Pr\left(\sum\nolimits_{j\neq i} a_j = k\right)(1-\alpha_h)$$

$$= \Pr\left(\sum\nolimits_{j\neq i} a_j = k\right)$$

The third line uses $\alpha_i = 1$. Again, since strategies are not correlated, the probability that k players cooperate when we exclude i and h is equal to the probability that $k-1$ players cooperate when we exclude i times the probability that h cooperates plus the probability that k players cooperate excluding i and h times the probability that h does not cooperate. This justifies the fourth line. Finally, the last line is due to $\alpha_h = 0$. ■

Lemma 2: If $\alpha_i > \alpha_h$ and $k \geq 1$, then $P(k,h) \geq P(k,i)$. Moreover, if there exist $k-1$ players different from i,h for which $\alpha_j > 0$, then $P(k,h) > P(k,i)$. **Proof**: Using the same argument we employed in Lemma 1 we have:

$$P(k,i) = \Pr\left(\sum\nolimits_{j\neq i} a_j = k\right)$$

$$= \Pr\left(\sum\nolimits_{j\neq i,h} a_j = k-1\right)\alpha_h + \Pr\left(\sum\nolimits_{j\neq i,h} a_j = k\right)(1-\alpha_h)$$

Analogously,

$$P(k,h) = \Pr\left(\sum\nolimits_{j\neq h} a_j = k\right)$$

$$= \Pr\left(\sum\nolimits_{j\neq i,h} a_j = k-1\right)\alpha_i + \Pr\left(\sum\nolimits_{j\neq i,h} a_j = k\right)(1-\alpha_i)$$

Therefore,

$$P(k,h) - P(k,i) = (\alpha_i - \alpha_h)\left[\Pr\left(\sum\nolimits_{j\neq i,h} a_j = k-1\right) - \Pr\left(\sum\nolimits_{j\neq i,h} a_j = k\right)\right]$$

$$= (\alpha_i - \alpha_h) \Pr\left(\sum\nolimits_{j \neq i,h} a_j = k-1\right) [1 - \alpha_i(1-\alpha_h) - \alpha_h(1-\alpha_i)]$$

$$= (\alpha_i - \alpha_h) \Pr\left(\sum\nolimits_{j \neq i,h} a_j = k-1\right) [(1-\alpha_i)(1-\alpha_h) + \alpha_i \alpha_h]$$

The second line uses the fact that $\Pr\left(\sum_{j \neq i,h} a_j = k\right) = \Pr\left(\sum_{j \neq i,h} a_j = k-1\right)[\alpha_i(1-\alpha_h) + \alpha_h(1-\alpha_i)]$. By assumption $(\alpha_i - \alpha_h) > 0, \Pr\left(\sum_{j \neq i,h} a_j = k-1\right) \geq 0$, and $1 - (1-\alpha_i)(1-\alpha_h) + \alpha_i \alpha_h > 0$. Moreover, if there exist $k-1$ players different from i, h for which $\alpha_j > 0$, then $\Pr\left(\sum_{j \neq i,h} a_j = k-1\right) > 0$ and, hence, $P(k,h) > P(k,i)$. ■

Case 1 (all cooperate): Suppose that $\alpha_i = 1$ for $i = 1, \ldots, N$. Then, $P(k,i) \neq 0$ if and only if $k = N-1$ and, hence, the Nash conditions become:

$$\left[\frac{B + sN}{N} - c\right] P(N-1, i) \geq 0$$

Since $s > c$, these conditions always hold. Therefore, $\alpha_i = 1$ for $i = 1, \ldots, N$ is always a Nash equilibrium.

Case 2 (nobody cooperates): Suppose that $\alpha_i = 0$ for $i = 1, \ldots, N$. Then, $P(k,i) \neq 0$ if and only if $k = 0$ and, hence, the Nash conditions become:

$$\left[\frac{B + s}{N} - c\right] P(0, i) \leq 0$$

These conditions hold if and only if $N \geq \frac{B+s}{c}$. Thus, if $N \geq \frac{B+s}{c}$, $\alpha_i = 0$ for all i is a Nash equilibrium.

Case 3 (some cooperate, some do not cooperate and some play a mixed strategy): Suppose that there is a Nash equilibrium in which n_1 players are cooperating, n_2 are playing a complete mixed strategy, and $N - n_1 - n_2$ are not cooperating. Without loss of generality, assume that $\alpha_i = 1$ for $i = 1, \ldots, n_1$, $\alpha_i \epsilon (0,1)$ for $i = n_1 + 1, \ldots, n_2$, and $\alpha_i = 0$ for $i = n_2 + 1, \ldots, N$. Then, for $i = 1, \ldots, n_1$, we have $P(k,i) \neq 0$ if and only if $n_1 - 1 \leq k \leq n_2 - 1$. Thus, the Nash conditions become:

$$\sum_{k=n_1-1}^{n_2-1} \left[\frac{B + s\,(k+1)}{N} - c\right] P(k,i) \geq 0$$

For $i = n_1 + 1, \dots, n_2$, we have $P(k,i) \neq 0$ if and only if $n_1 \leq k \leq n_2 - 1$. Thus, the Nash conditions become:

$$\sum_{k=n_1}^{n_2-1} \left[\frac{B + s\,(k+1)}{N} - c\right] P(k,i) = 0$$

Finally, for $i = n_2 + 1, \dots, N$, we have $P(k,i) \neq 0$ if and only if $n_1 \leq k \leq n_2$. Thus, the Nash conditions become:

$$\sum_{k=n_1}^{n_2} \left[\frac{B + s\,(k+1)}{N} - c\right] P(k,i) \leq 0$$

Arbitrarily select $i < n_1$ and $n_1 + 1 \leq h \leq n_2$. Then, $\alpha_i = 1$ and $\alpha_h = 0$, and Lemma 1 implies that $P(k,i) = P(k+1,h)$. Therefore:

$$\sum_{k=n_1-1}^{n_2-1} \left[\frac{B + s\,(k+1)}{N} - c\right] P(k,i) = \sum_{k=n_{1-1}}^{n_2-1} \left[\frac{B + s\,(k+1)}{N} - c\right] P(k+1,h)$$

$$= \sum_{k=n_1}^{n_2} \left[\frac{B + sk}{N} - c\right] P(k,h)$$

But, this leads to a contradiction, because the Nash condition for i implies that $\sum_{k=n_1}^{n_2} \left[\frac{B+sk}{N} - c\right] P(k,h) \geq 0$, while the Nash condition for h implies $\sum_{k=n_1}^{n_2} \left[\frac{B+s(k+1)}{N} - c\right] P(k,h) \leq 0$. Note that the argument does not depend on the existence of a group of players who are playing a complete mixed strategy. In other words, if $n_2 = n_1$, the same argument holds. Hence, there cannot be a Nash equilibrium in which some players cooperate with probability 1 and other players do not cooperate at all.

Case 4 (all play a mixed strategy): Suppose that $\alpha_i \epsilon (0,1)$ for $i = 1, \dots, N$. Then $P(k,i) \neq 0$ for all k. Thus, the Nash conditions become:

$$\sum_{k=0}^{N-1} \left[\frac{B + s\,(k+1)}{N} - c\right] P(k,i) = 0$$

Since $\sum_{k=0}^{N-1} P(k,i) = 1$, these conditions are equivalent to:

$$\sum_{k=0}^{N-1} k\,P(k,i) = \sum_{k=1}^{N-1} k\,P(k,i) = \frac{cN - B - s}{s}$$

Arbitrarily select i and h and, without loss of generality, assume that $\alpha_i > \alpha_h$. Then, from Lemma 2, it must be the case that $P(k,h) > P(k,i)$ for all $k \geq 1$. But this leads to a contradiction, because $\sum_{k=1}^{N-1} k\, P(k,i) = \frac{cN-B-s}{s}$ and $\sum_{k=1}^{N-1} k\, P(k,h) = \frac{cN-B-s}{s}$ cannot simultaneously hold. Thus, in a Nash equilibrium in which all players are playing a complete mixed strategy, it must be the case that $\alpha_i = \hat{\alpha}\, \epsilon\, (0,1)$ for $i = 1, \dots, N$. In this case, $P(k,i) = \binom{N-1}{k} \hat{\alpha}^k (1-\hat{\alpha})^{N-1-k}$, i.e. $k \sim binomial\,(N-1, \hat{\alpha}\,)$. Therefore,

$$\sum_{k=0}^{N-1} \left[\frac{B+s(k+1)}{N}\right] \binom{N-1}{k} \hat{\alpha}^k (1-\hat{\alpha})^{N-1-k} = c$$

$$\sum_{k=0}^{N-1} k \binom{N-1}{k} \hat{\alpha}^k (1-\hat{\alpha})^{N-1-k} = \frac{cN-B-s}{s}$$

$$\hat{\alpha}(N-1) = \frac{cN-B-s}{s}$$

The last line uses the fact that the expected value of $k \sim binomial\,(N-1, \hat{\alpha}\,)$ is $\hat{\alpha}(N-1)$. Therefore, $\hat{\alpha} = \frac{cN-B-s}{s(N-1)}$. Note that $s > c$ implies that $\hat{\alpha} < 1$, while $\hat{\alpha} > 0$ if and only if $N > \frac{B+s}{c}$. Thus, $\alpha_i = \hat{\alpha} = \frac{cN-B-s}{s(N-1)}$ for $i = 1, \dots, N$ is a Nash equilibrium if and only if $N > \frac{B+s}{c}$.

Case 5 (some cooperate and some play a mixed strategy): Suppose that there is a Nash equilibrium in which n_1 players are cooperating and $N - n_1$ are playing a complete mixed strategy. Without loss of generality, assume that $\alpha_i = 1$ for $i = 1, \dots, n_1$ and $\alpha_i\, \epsilon\, (0,1)$ for $i = n_1 + 1, \dots, N$. Then, for $i = 1, \dots, n_1$, the Nash conditions become $\sum_{k=n_1-1}^{N-1} \left[\frac{B+s(k+1)}{N} - c\right] P(k,i) \geq 0$ or, which is equivalent:

$$\sum_{k=n_1-1}^{N-1} k\, P(k,i) \geq \frac{cN-B-s}{s}$$

For $i = n_1 + 1, \dots, N$, the conditions become $\sum_{k=n_1}^{N-1} \left[\frac{B+s(k+1)}{N} - c\right] P(k,h) = 0$ or, which is equivalent:

$$\sum_{k=n_1}^{N-1} k\, P(k,h) = \frac{cN-B-s}{s}$$

Using the same argument that we employed to prove that if, in a Nash equilibrium, all players are playing a mixed strategy, they must play the same strategy, we can prove that in a Nash equilibrium $\alpha_i = \tilde{\alpha}$ for all $i = 1, \dots, n_1$. As a consequence, $P(k,i) = \binom{N - n_1 - 1}{k - n_1} \tilde{\alpha}^{k-n_1}(1-\tilde{\alpha})^{N-1-k}$ and, therefore:

$$\sum_{k=n_1}^{N-1} \binom{N - n_1 - 1}{k - n_1} \tilde{\alpha}^{k-n_1}(1-\tilde{\alpha})^{N-1-k} = \frac{cN - B - s}{s}$$

This implies that:

$$\tilde{\alpha} = \frac{cN - B - (1 + n_1)s}{s(N - 1 - n_1)}$$

For $i = n_1 + 1, \dots, N$ we have $P(k,i) = \binom{N - n_1}{k - n_1 + 1} \tilde{\alpha}^{k-n_1+1}(1-\tilde{\alpha})^{N-k-1}$. Therefore:

$$\sum_{k=n_1-1}^{N-1} k \binom{N - n_1}{k - n_1 + 1} \hat{\alpha}^{k-n_1+1}(1-\hat{\alpha})^{N-k-1} \geq \frac{cN - B - s}{s}$$

This implies:

$$\tilde{\alpha} \geq \frac{cN - B - n_1 s}{s(N - n_1)}$$

Since $s > c$, $\tilde{\alpha} = \frac{cN-B-(1+n_1)s}{s(N-1-n_1)}$ and $\tilde{\alpha} \geq \frac{cN-B-n_1 s}{s(N-n_1)}$ never hold simultaneously.

Case 6 (some do not cooperate and some play a mixed strategy): Suppose that there is a Nash equilibrium in which n_2 are playing a complete mixed strategy and $N - n_2$ are not cooperating. Without loss of generality assume $\alpha_i \epsilon (0,1)$ for $i = 1, \dots, n_2$, and $\alpha_i = 0$ for $i = n_2 + 1, \dots, N$. Then, for $i = 1, \dots, n_2$ we have $P(k,i) \neq 0$ for $0 \leq k \leq n_2 - 1$. Thus, the Nash conditions become:

$$\sum_{k=0}^{n_2-1} \left[\frac{B + s\,(k+1)}{N} - c\right] P(k,i) = 0$$

Since $\sum_{k=0}^{N-1} P(k,i) = 1$, these conditions are equivalent to:

$$\sum_{k=0}^{n_2-1} kP(k,i) = \sum_{k=1}^{n_2-1} kP(k,i) = \frac{cN - B - s}{s}$$

For $i = n_2 + 1, \dots, N$ we have $P(k, i) \neq 0$ for $0 \leq k \leq n_2$. Thus, the Nash conditions become:

$$\sum_{k=0}^{n_2} \left[\frac{B + s\,(k+1)}{N} - c \right] P(k, i) \leq 0$$

Since $\sum_{k=0}^{N-1} P(k, i) = 1$, these conditions are equivalent to:

$$\sum_{k=0}^{n_2} kP(k, i) = \sum_{k=1}^{n_2} kP(k, i) \leq \frac{cN - B - s}{s}$$

Arbitrarily select $i \leq n_2$ and $h > n_2$. Then, from Lemma 2, we have $P(k, h) > P(k, i)$ for all $k \geq 1$, which implies $\sum_{k=1}^{n_2-1} kP(k, h) \geq \frac{cN-B-s}{s}$. However, this leads to a contradiction because $\sum_{k=1}^{n_2-1} kP(k, h) \geq \frac{cN-B-s}{s}$ and $\sum_{k=1}^{n_2} kP(k, h) \leq \frac{cN-B-s}{s}$ cannot hold simultaneously. ■

Online Appendix 2. Description of the Experiment

In this appendix we present the script for the general instructions, the instructions given to the participants, the quiz and the questionnaire.

Appendix 2.1. Script for General Instructions

We would like to welcome everyone to this experiment. This is an experiment in decision-making, and you will be paid for your participation, in cash, at the end of the experiment. Different subjects may earn different amounts. What you earn depends partly on your decisions, partly on the decisions of others and partly on chance.

The entire experiment will be conducted through computer terminals, and all interaction between participants will take place through the computers. It is very important for you not to talk or to try to communicate with other subjects during the experiment in any way.

At your workstation, you will find a pencil, a paper with instructions and scratch paper. During the experiment you can use the scratch paper to make calculations.

We will now start with a brief instruction period. During this period, you will be given a complete description of the experiment. If you have any questions during the instruction period, please raise your hand and your question will be answered so that everyone can hear the response. If any difficulties arise after the experiment has begun, raise your hand, and one of the persons conducting the experiment will come to assist you.

You are one of 20 students who have been asked to participate in this experiment. In each round you will be randomly assigned to one of two groups, consisting of 10 persons each. Then, you will play a computer game, which will appear on the screen, with the other members of that same group. At the beginning of each round, the parameters of the game will appear on the screen, as will the timing. At the end of the round, you will be informed of the result of the game, the points you have earned and the points that you have accumulated so far. In the next round, all players will again be randomly assigned to one of the two new groups of 10 people each.

The experiment you are participating in is broken down into four unpaid practice rounds and twelve separate paid rounds. At the end of the last round, you will be paid the total amount you have accumulated during the course of the last twelve rounds. Your profit is denominated in POINTS. Your PESO profit is determined by multiplying your earnings in points by a conversion rate. In this experiment, the conversion rate is 2 pesos to 1 point.[24] Everyone will be paid in private and you are under no obligation to tell others how much you earned.

Please read the instructions that you will find on your desktop carefully. You have 10 minutes. Please, remember that if you have any questions, you should ask them aloud.

Appendix 2.2. Instructions

1. In each round you receive ONE point. You can keep it for yourself or invest it in a common fund. You have 90 seconds to make your decision. When you select an option, please press the "Next" button. If, after 90 seconds, you have not selected an option, the computer will randomly do it for you.

2. Once all players have taken a decision, the outcome of the game will appear on the screen: If the investment is successful, each of the ten players will receive B points, and those who have decided to invest their point will receive s additional points. If the investment fails, nobody gets a profit, and those who have decided to invest their point will lose the point that they initially invested.

Therefore:

- If you have decided to invest your point in the common fund and the investment is successful, you will accumulate $B + s$ points;
- If you have decided to keep your point for yourself and the investment is successful, you will accumulate $B + 1$ points;
- If you have decided to invest your point in the common fund and the investment fails, you will earn 0 points;
- If you have decided to keep your point for yourself and the investment fails, you will earn 1 point.

[24] The conversion rate was adjusted by inflation (20% since August). Hence, starting in August, the rate was adjusted to 2.4 pesos for 1 point. At that stage, 2 and 2.4 Argentine pesos were equivalent to approximately 0.25 and 0.28 dollars, respectively.

3. The probability of success of the investment depends on the proportion of players in your group who have decided to invest:

$$\text{Successful investment probability} = \frac{\text{(Number of players who invested their point)}}{10}$$

Thus, the greater the number of players who have decided to invest their point, the greater is the probability that the investment will be successful.

For example, if 6 out of 10 participants choose to invest their point in the common fund, the chances of success are 60%. If the investment is successful, those six participants will get $B + s$, and the remaining four will obtain $B + 1$. However, if the investment fails, the six participants who decided to invest their point get 0 units, while the remaining four get 1 point.

Let us suppose that, in another case, only 2 out of 10 players decide to invest their points in the common fund. Therefore, the chances of success are 20%. If the investment is successful, those two participants will get $B + s$, and the remaining eight will obtain $B + 1$. However, if the investment fails, the two participants who decided to invest their point will get 0 units, while the remaining eight will get 1 point.

At the end of each round, you will be told how many players have decided to invest their point in the common fund, whether the investment was successful or not, the gain in the round, and the total amount of points accumulated from the fifth round onward. To end the round, you will need to press the "Next" button.

At the beginning of the next round, you will be randomly assigned to a new group. Pay attention because the parameters of the game may have changed. That is, in each round, B and / or s may vary.

After the sixteenth, round you will be asked to answer a few questions about you. Finally, when you click "Finish", the screen will display a WORD. It is IMPORTANT to remember this word because you have to present this password to the person who was running the experiment in order to receive your payoff.

Appendix 2.3. Belief Questions

The following script provides a sample of the questions that subjects were asked about their beliefs before starting the game.

Screen: Before you begin to play, we would like to ask you some questions about the experiment. These questions are for information purposes only, and there is no right or wrong answer. You will not be paid for answering them.

1. Suppose that $B = 1.25$ and $s = 0$: How many players out of a group of 10 persons do you think will invest their point in the common fund? [11 options].

2. Suppose that $B = 1.25$ and $s = 1.25$: How many players out of a group of 10 persons do you think will invest their point in the common fund? [11 options].

3. Suppose that $B = 3$ and $s = 1.25$: How many players out of a group of 10 persons do you think will invest their point in the common fund? [11 options].

4. Suppose that $B = 3$ and $s = 1.75$: How many players out of a group of 10 persons do you think will invest their point in the common fund? [11 options].

Appendix 2.4. The Quiz

After a general explanation of the rules of the game, subjects took the following quiz:

1. Suppose the following parameters of the game: $B = 2$ and $s = 0$. If all players, including you, decide NOT to invest their point in the common fund and the investment fails, how many points do you obtain at the end of this round? [5 options]

2. Suppose the following parameters of the game: $B = 2$ and $s = 1$. If all players, including you, decide to invest their point in the common fund and the investment is successful, how many points do you get at the end of this round? [5 options]

3. Consider the following two possible games:

- First game: $B = 3$ and $s = 1$;
- Second game: $B = 4$ and $s = 1$;

If you decide NOT to invest your point and the investment fails, in which of the two games do you accumulate more points? [3 options]

4. If there are 10 players and 8 of them decide to invest their point, what is your best option if the parameters of the game are: $B = 0.5$ and $s = 2$? [3 options]

5. If there are 10 players and 4 of them decide to invest their point, what is your best option if the parameters of the game are: $B = 1$ and $s = 1$? [3 options]

Appendix 2.5. Sample Screen

At the end of each round, subjects were shown a summary of the decisions taken in the round and were told whether the investment was successful or not, what the payoff obtained in that round was and what their own accumulated payoffs for paid rounds was.

- **Screen*:***

  You have decided (not) to invest your point.

  (1, 2, 3, 4, 5, 6, 7, 8, 9 or all) subjects in your group have decided to invest their point.

  The investment was (not) successful.

  Your earning in this round was ____ points.

  You have accumulated ____ points since the start of the game.

Appendix 2.6. The Questionnaire

Thank you for participating in this experiment! Please complete the following questionnaire before leaving.

Question 1: Gender (male/female)

Question 2: Age (in years)

Question 3: Nationality

Question 4: Whether or not you are fluent in English (Yes/No)

Question 5: Whether you have ever taken a course in game theory (Yes/No)

Question 6: Current studies (Graduate/Undergraduate)

Question 7: Degree in: (a) Economics; (b) Business Administration or Accountancy; (c) Finance; (d) Political Science, International Affairs, Humanities or Law; (e) Marketing or Human Resources; (f) Other (please specify).

Question 8: Number of courses out of the total courses in your degree program that you have completed successfully.