Exp Econ (2012) 15:24–43
DOI 10.1007/s10683-011-9286-4

Revisiting strategic versus non-strategic cooperation

Ernesto Reuben · Sigrid Suetens

Received: 1 September 2010 / Accepted: 4 May 2011 / Published online: 17 May 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract We propose a novel experimental method that disentangles strategically- and non-strategically-motivated behavior. We apply it to an indefinitely-repeated prisoner's dilemma game to observe simultaneously how the *same* individual behaves in situations with future interaction and in situations with no future interaction, while controlling for expectations. This method allows us to determine the extent to which strategically-cooperating individuals are responsible for the observed pattern of cooperation in experiments with repeated interaction, including the so-called endgame effect. Our results indicate that the most common motive for cooperation in repeated games is strategic.

Keywords Repeated games · Cooperation · Reputation building · Strong reciprocity

JEL Classification C92 · D01 · D70

1 Introduction

In order to design policies and institutions that promote cooperation in social dilemmas, it is important to understand and measure the pervasiveness of the various motivations behind cooperation. Broadly speaking, in social dilemma games with re-

Electronic supplementary material The online version of this article (doi:10.1007/s10683-011-9286-4) contains supplementary material, which is available to authorized users.

E. Reuben (✉)
Columbia University and IZA, New York, USA
e-mail: ereuben@columbia.edu

S. Suetens
CentER, TILEC, and TIBER, Tilburg University, Tilburg, The Netherlands
e-mail: s.suetens@uvt.nl

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

peated interaction, cooperation can be the result of strategic and non-strategic motivations. Strategic cooperation is cooperation motivated solely by the prospect of future interaction and supported by the future behavior of others. By contrast, non-strategic cooperation refers to cooperation that is not forward looking. It refers to individuals motivated to cooperate (at least under certain circumstances) in situations where there is no future interaction. Whereas both strategic and non-strategic cooperation are supported by experimental evidence (see Sobel 2005, for a review of the literature), it is not known which share of cooperation can be attributed to each of these motivations.[1] To a large extent, this is because in repeated games it is not sufficient to observe individual choices to tell whether cooperation is strategic or not. Does an individual cooperate solely because there is future interaction and stops cooperating when this possibility disappears? Or is he a conditional cooperator who expects others to cooperate when there is future interaction and stop cooperating when there is no future interaction?[2] In this paper, we propose a novel experimental method that solves this identification problem, and we use it to measure the amount of cooperation in a repeated prisoner's dilemma game that can be attributed to strategic and to non-strategic motivations.

To distinguish strategic from non-strategic motivations for cooperation, we run an experiment where subjects play an indefinitely-repeated prisoner's dilemma game where they know whether the period they are currently playing is the last period of the game or not. However, since we are interested in motivations for cooperation that come up in finitely-repeated games, we use parameters such that cooperation in our game cannot be rationalized as being part of an equilibrium of the indefinitely-repeated game assuming common knowledge of rationality and own-payoff maximization. We use the strategy method developed by Selten (1967) to allow a first type of players, henceforth *first movers*, to condition their decision on whether the period they are currently playing *is the last period* of the game or *is not the last period* of the game. Moreover, there is a second type of players, henceforth *second movers*, who are allowed to condition their decision on: (i) whether the period they are playing *is* or *is not* the last period of the game, and (ii) whether the first mover *cooperates* or *defects*. We validate our use of the strategy method by comparing behavior in this setting to behavior in control treatments that do not use the strategy method.

[1] On the one hand, cooperation in experiments without future interaction suggests there is scope for non-strategic cooperation because, in these settings, cooperation cannot be strategic (and rational) by definition. On the other hand, the existence of strategic cooperation is suggested by differences in cooperation between random-matching and fixed-matching treatments, which have been found in bribery games (Abbink 2004), principal-agent games (Cochard and Willinger 2005), trust games (Huck et al. 2006), conflict games (Lacomba et al. 2010), prisoner's dilemma games (Duffy and Ochs 2009), gift-exchange games (Gächter and Falk 2002), and public good games (Croson 1996; Keser and van Winden 2000). In contrast, Andreoni (1988) finds more cooperation under random matching (for an explanation of these mixed results see Andreoni and Croson 2008). Moreover, studies with indefinitely-repeated games (see e.g., Dal Bó 2005; Aoyagi and Fréchette 2009; Dal Bó and Fréchette 2011) convincingly show that the subjects' behavior is consistent with theories that predict strategic cooperation by rational own-payoff maximizing players.

[2] The same confounding effect is also found when comparing fixed-matching and random-matching experiments (where a one-shot game is played repeatedly). The lower rate of cooperation that is typically observed under random matching compared to fixed matching can be caused by strategically-motivated individuals who do not have a reason to cooperate under random matching, or by conditional cooperators who believe others will cooperate less under random matching.

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

The novelty of this experimental method is that it allows us to simultaneously observe choices in the continuation game and in the end game *within* subjects, and crucially, observe *counterfactual* behavior (of second movers) in the last period. That is, we know whether second movers cooperate or defect in the last period of the game *when it is certain that the first mover will cooperate*. Strictly speaking, second movers who are willing to cooperate in the last period must be motivated by non-strategic reasons.[3] In contrast, second movers who always defect in the last period, but who are willing to cooperate in non-last periods, are clearly cooperating solely for strategic reasons. We can make this inference for second movers as they face no uncertainty with respect to the first mover's behavior.[4]

We find that most of the cooperation in the repeated prisoner's dilemma game is strategically motivated. We also find that the individuals' motivation to cooperate is relatively stable over time. In other words, individuals consistently choose either a strategically-motivated strategy or a non-strategically-motivated strategy across periods. Moreover, eliciting strategies does not produce behavior that is different from behavior in the control treatments where strategies were not elicited.

The method proposed in this paper can also be used to disentangle the so-called endgame effect: a sharp decline in cooperation in the last periods of a repeated game (see Selten and Stoecker 1986; Selten et al. 1997; Keser and van Winden 2000). If cooperation is entirely strategic, it drops in the last periods because strategically-motivated individuals no longer have an incentive to cooperate. However, cooperation can also drop when it is non-strategic if individuals believe others will stop cooperating in the last periods and this induces them to stop cooperating themselves.[5] In existing experiments, one cannot tell how much of the drop in cooperation is due to the presence of strategic cooperation and how much is due to expectations of strategic cooperation, whereas in our experiment one can.

Our paper can be seen as building on the work of Fischbacher et al. (2001), who introduce a method to identify different types of motivations for cooperation in a one-shot game, which precludes strategic behavior by nature.[6] By eliciting strategies in a repeated setting, we can identify different types of both non-strategic and strategic motivations for cooperation.

A number of studies estimate repeated-game strategies on the basis of experimental data (e.g. Engle-Warnick and Slonim 2004, 2006; Aoyagi and Fréchette 2009;

[3] We do not suggest that individuals who cooperate in the last period are failing to take into account the (future) behavior of others. Instead, we suggest that players who exhibit this behavior are not *exclusively* motivated to cooperate by strategic considerations.

[4] We cannot fully identify the motivations for cooperation of first movers because their expectation of the behavior of others does affect their choice. In fact, as we argue in the body of the paper, the behavior of first movers is likely to be driven exclusively by their expectations, which allows us to measure the impact that expectations have on cooperation.

[5] The experiments of Croson (2000) and Gächter and Renner (2006), for example, show that the observed pattern of beliefs closely resembles the observed pattern of actions. Kreps et al. (1982) propose a model that rationalizes this pattern by assuming incomplete information concerning preferences for cooperation.

[6] Duffy and Ochs (2009) and Fischbacher and Gächter (2010) use this method to show that the elicited strategies predict subsequent behavior in a repeated version of the game where subjects are randomly rematched in every period.

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

Dal Bó and Fréchette 2011; Fudenberg et al. 2011). Overall, individuals use strategies that produce an endgame effect in finitely-repeated games and punishment of free riders that resemble (lenient) grim-trigger or tit-for-tat strategies in indefinitely-repeated games.[7] The strategies observed in these studies can be the result of both strategic and non-strategic cooperators, and therefore, these papers cannot differentiate between the two motivations.

Muller et al. (2008), Bruttel and Kamecke (2011), Dreber et al. (2011), and Cabral et al. (2011) are the closest related studies. Muller et al. (2008) report an experiment where subjects play a two-period linear public good game in which they can condition their contribution in the second period on the total contribution of others in the first period. The result is that strategic behavior has a more pronounced effect than learning in explaining the endgame effect. Bruttel and Kamecke (2011) have participants play a prisoner's dilemma game three times in a row and then, for the remaining periods of the repeated game, submit conditional choices depending on the four potential outcomes in the previous period. They find that such short-memory strategies account for a large share of actual decisions in the indefinitely-repeated game. Our method adds another component, which is the identification of strategically- and non-strategically-motivated cooperation. In particular, in Muller et al. (2008) and Bruttel and Kamecke (2011), choices within each period are simultaneous and therefore affected by expectations. For example, even when contributions in the previous period are high, subjects who are conditional cooperators would still defect in the current period if they believe others will defect in the current period (after cooperating strategically in the previous period). Dreber et al. (2011) take a different approach and identify strategic from non-strategic cooperation by studying the correlation between cooperative strategies in indefinitely-repeated games and behavior in a one-shot dictator game. They find no correlation between the two. In our paper, we identify the motivation for cooperation within the same game, and therefore, our method is not affected by behavioral changes induced by differences in the games played. Finally, Cabral et al. (2011) also disentangle strategic and non-strategic motivations by informing subjects if the period they are playing is the last period of an indefinitely-repeated game. It differs from our study in that it does not use the strategy method, and focuses on motivations to cooperate in infinitely repeated games where cooperation can be supported in equilibrium.

The paper is organized as follows. In Sect. 2 we describe the game and methodology used in the experiments. In Sect. 3 we describe the experimental design and procedures. We present the results in Sect. 4 and conclude in Sect. 5.

2 Description of the game and methodology

The game played in the experiment is the strategic equivalent of an indefinitely-repeated sequential prisoner's dilemma game. In each period, with probability $(1-\delta)$

[7] Engle-Warnick and Slonim (2004) find that the cooperation rate of inexperienced subjects does not differ between finitely and indefinitely-repeated games (as do Normann and Wallace 2006), but does differ with experienced subjects (see also Dal Bó 2005).

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

the game ends after the period is played and with probability δ the game continues. In the stage game, each of the two players can either *cooperate* or *defect*. If both players cooperate they each get π^C, if both defect they each get π^D, and if one defects and the other cooperates the defector gets the temptation payoff π^T and the cooperator gets the sucker payoff π^S. Payoffs are such that defecting is the dominant strategy ($\pi^T > \pi^C > \pi^D > \pi^S$), and mutual cooperation is the efficient outcome ($2\pi^C > \pi^T + \pi^S$).

The main feature of our experimental method is that, in each period, players can condition their action on whether they are currently playing the final period of the game, to which we refer to as the *last* period, or whether they are not playing the last period of the game, to which we refer to as *continuation* periods. In addition, one of the two players is designated to be the *first mover* and the other to be the *second mover*. The second mover can condition her action on whether the first mover cooperates or defects. In other words, the first mover submits an action for two cases: (i) the game continues and (ii) the game ends; and the second mover submits an action for four cases: (i) the first mover cooperates and the game continues, (ii) the first mover defects and the game continues, (iii) the first mover cooperates and the game ends, and (iv) the first mover defects and the game ends. After both players make a decision, they learn whether they were playing a continuation period or the last period of the game, and they are informed about the corresponding action of the other player. Note that players are informed only of the *realized* action of the other player and not of the other player's contingent choices.

As previously mentioned, our method allows us to observe the stage-game strategies used by second movers.[8] There are two strategies that are of special interest. The first one consists of reciprocating the first mover's choice irrespective of whether it is a continuation period or the last period. We refer to this strategy as *strong reciprocity*, that is, reciprocity irrespective of potential future interaction (Gintis 2000; Fehr et al. 2002). The second one corresponds to reciprocating only if a continuation period is being played and defecting if the last period is being played. We refer to this strategy as *reputation building* since it is a clear example of strategically-motivated cooperation. These and other important strategies are described in Table 1. For example, it is also informative to know the prevalence of second movers who choose the strategies unconditional defection and unconditional cooperation.

As is well known, cooperation in indefinitely-repeated games can be achieved by rational own-payoff maximizing individuals with the use of trigger strategies (Friedman 1971). In fact, for a sufficiently high δ, any payoff between the cooperative and the defection payoff can be supported by some equilibrium strategy (Rubinstein 1979; Fudenberg and Maskin 1986). In our game, mutual cooperation by rational own-payoff maximizers is supported in continuation periods if δ is larger than the threshold $\delta^* = (\pi^T - \pi^C)/(\pi^T - \pi^D)$.[9] This follows from the fact that both players always

[8] To facilitate reading, we will often refer to stage-game strategies simply as strategies. We do not elicit the actual strategies for the whole game as doing so could require an infinite number of questions.

[9] We are assuming no time discounting in the experiment, which we think is reasonable since the time interval between periods is very short and subjects are not paid until the end.

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

Table 1 Strategies of second movers

Strategy	Continuation period		Last period	
	First mover Cooperates	First mover Defects	First mover Cooperates	First mover Defects
Unconditional defection	Defect	Defect	Defect	Defect
Unconditional cooperation	Cooperate	Cooperate	Cooperate	Cooperate
Reputation building	Cooperate	Defect	Defect	Defect
Strong reciprocity	Cooperate	Defect	Cooperate	Defect

defect when they play the last period of the game, which makes the game in continuation periods equivalent to an indefinitely-repeated game.[10] Therefore, mutual cooperation can be sustained in continuation periods by a trigger strategy if $\delta \geq \delta^*$, in which case the second mover can get a higher payoff by cooperating than by defecting.[11]

The pervasiveness of strategic versus non-strategic cooperation and the study of endgame effects has been done mostly in finitely-repeated games (e.g., see the references in footnote 1). To reproduce the motivations for cooperation present in these studies, we use parameter values such that mutual cooperation is *not* part of an equilibrium under common knowledge that all players are rational own-payoff maximizers.[12] We do so by setting the probability of continuation and the payoffs of the stage game such that $\delta < \delta^*$ in all our treatments.[13] Under this condition, individuals can be motivated to cooperate by two distinct reasons. First, individuals might cooperate for non-strategic reasons (e.g., due to other-regarding preferences). Second, individuals might cooperate strategically because they believe there are people in the population who reciprocate cooperation for non-strategic reasons (see Kreps et al. 1982). Specifically, first (second) movers might rationally cooperate in not-last periods (i.e., reputation-build) if they believe there are sufficiently many second (first) movers who are willing to cooperate in the last period as long as they have not experienced defection in not-last periods, or put differently, if the probability of being paired with a *strong reciprocator* is sufficiently high.[14]

[10] A general worry with playing indefinitely-repeated games is that subjects know the experiment cannot last for an extremely long time. Thus, they might discount future interactions at a rate that is lower than δ. However, for the purpose of our experiment, this can at most induce a small decrease in the frequency of strategically motivated cooperation (individuals who are cooperating for non-strategic reasons and those who are already defecting are not affected by more discounting).

[11] Given that the second mover has no reason to cooperate if the first mover defects, the first mover does not have an incentive to deviate from an equilibrium with mutual cooperation.

[12] One can expect some degree of unilateral cooperation if subjects play a correlated equilibrium and $\delta \geq (\pi^D - \pi^S)/(\pi^T - \pi^D)$ (see Stahl 1991). However, these equilibria require a high degree of coordination that is hard to achieve. We report whether there is evidence for these type of strategies in footnote 20.

[13] As previously mentioned, we choose a low continuation probability so that cooperation by solely rational own-payoff maximizers is not an equilibrium. However, it would also be interesting to have a longer time horizon where $\delta \geq \delta^*$ in order to study equilibrium selection and conditional cooperation in these conditions.

[14] Since the Kreps et al. (1982) model utilizes mixed strategies—which are hard to observe by eliciting only stage-game strategies (see the discussion in Palacios-Huerta and Volij 2008)—we are unable to test

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

3 Experimental design and procedures

The experiment was conducted in 2007 in CentERlab at Tilburg University. It was programmed using z-Tree (Fischbacher 2007) and lasted about 45 minutes. Subjects were recruited through online recruitment systems. In total, 180 subjects participated and each subject played only once. After their arrival, subjects drew a card to be randomly assigned to a seat in the laboratory, and consequently to a role and a treatment. Once everyone was seated, subjects were given the instructions for the experiment. The instructions are written with neutral language. Thereafter, roles were revealed and subjects had to answer control questions to corroborate their understanding of the game. Next, they played the game until the random draw indicated that it ended. Roles and pairs were kept constant throughout the experiment so it is plausible to assume that there is no correlation across the decisions of the different pairs.[15] Once the game finished, subjects answered a debriefing questionnaire after which they were paid in cash and dismissed. See the online supplementary materials for an example of the instructions.

The experiment consists of one treatment where we implemented the game described in the previous section, which we refer to as *Main*, and two control treatments. The purpose of *Main* is to identify the various strategies used by second movers, while the purpose of the control treatments is to identify any behavioral changes induced by the strategy method.

In principle, it is possible that the use of the strategy method induces a change in behavior. In the experimental literature there is yet no consensus if this is indeed the case. Various authors report no significant differences in, for example, sequential dictator games (Cason and Mui 1998) and, closest to our study, chicken and prisoner's dilemma games (Brandts and Charness 2000). However, there are also studies that do find differences in behavior. For instance, some authors have found less punishment with the use of the strategy method than without it (Brosig et al. 2003; Falk et al. 2005). For a comprehensive survey of the effects of the strategy method see Brandts and Charness (forthcoming).

We use two control treatments to test whether the method affects behavior in our setting. In the first control treatment, *Control I*, subjects play the same game as in *Main* except that they do not submit a decision for both last and not-last periods. In *Control I*, subjects are told whether the game ends or continues before the start of each period, and then they make their decision. Note that second movers still submit contingent choices depending on whether the first mover cooperates or not. By comparing choices between *Control I* and *Main*, we can test whether behavior is affected by conditioning decisions on whether it is the last period or not. In the second control, *Control II*, we again implement the same game but this time without the use of the strategy method (i.e., subjects play the game "normally"). Here, subjects are

whether the observed amount of reputation building is consistent with that model's predictions. In addition, the uncertain duration of the game used in this paper makes it unclear how to calculate the subjects' optimal behavior once cooperation starts to unravel. For experimental tests of the Kreps et al. (1982) model, see Camerer and Weigelt (1988) and Andreoni and Miller (1993).

[15]The multiple observations of a given pair will be treated as correlated throughout the data analysis in Sect. 4.

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

told whether the game ends or continues before the start of each period, and second movers learn what the first mover did before they make their choice. By comparing behavior between *Control I* and *Control II*, we can test whether the decisions of second movers are affected by the possibility to condition their choice on the action of the first mover.

In all three treatments, we chose a continuation probability of $\delta = 0.60$ and the payoffs of the stage game were selected so that $\delta^* > \delta$. In particular, payoffs in the prisoner's dilemma were the following: $\pi^D = 15, \pi^C = 22, \pi^T = 33$ and $\pi^S = 10$. Subjects played the repeated game once. Therefore, by not having more iterations of the repeated game, we potentially identify a lower bound for strategic cooperation. For example, Dal Bó and Fréchette (2011) show that in indefinitely-repeated games where cooperation is part of an equilibrium, cooperation increases as subjects gain experience, which suggests that subjects are learning to strategically cooperate. We used a computer to randomly determine when the game ends, and in order to make the three treatments easily comparable, the same random sequences generated for pairs in *Main* were used for pairs in the control treatments. The average number of periods played equaled 2.73, and average earnings were €9.38.

4 Results

In this section we present the experimental results. Since pairs of subjects played a different number of periods, when we give descriptive statistics or perform hypothesis tests that are based on data aggregated across periods, we adjust the statistic/test by the inverse number of periods played. This way, each pair receives an equal weight. Importantly, none of the qualitative results change if we concentrate on data from the first period, which was played only once by all pairs.

We proceed as follows. In Sect. 4.1 we provide summary statistics of the realized cooperation rates in the three treatments and test whether there are differences between *Main* and the two control treatments. In Sect. 4.2, we provide an overview of the subjects' strategies in *Main* in order to identify their motivations to cooperate. In Sect. 4.3, we analyze the stability over time of the subjects' strategies. Lastly, in Sect. 4.4, we turn to the causes of the endgame effect.

4.1 Cooperation rates

Figure 1 gives an overview of the distribution of *realized* outcomes in all periods in the three treatments. The mutual cooperation rates lie between 16% and 19% and outcomes are clearly similar across the three treatments. This is corroborated if we run a multinomial probit regression with the four outcomes as the dependent variable, treatment dummies as independent variables, and using White's heteroskedasticity consistent estimator to cluster on pairs of subjects. We find that we cannot reject the hypothesis that all outcomes occur with the same frequency in all treatments ($p = 0.690$). We also do not observe statistically significant differences across treatments if we use probit regressions to compare separately the frequency of each outcome ($p > 0.147$). Hence, we can safely conclude that our proposed method of eliciting

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

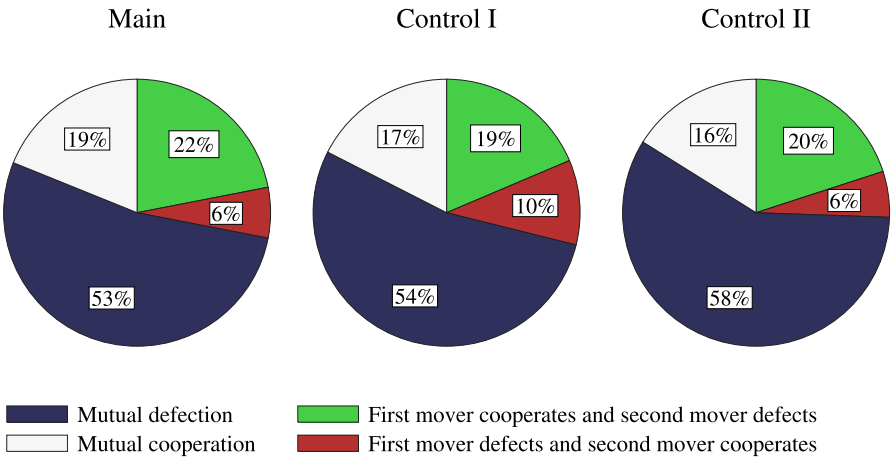


Fig. 1 Realized outcomes. *Note:* The pie charts show the frequency of each of the four possible outcomes in *Main*, *Control I*, and *Control II*. To give each pair of subjects the same weight, frequencies are weighted by the inverse of the number of periods played

Table 2 Cooperation rates

	All periods			Cont. periods			Last periods		
	M	C-I	C-II	M	C-I	C-II	M	C-I	C-II
Both players	33%	32%	29%	48%	39%	42%	15%	25%	20%
First movers	41%	36%	36%	56%	47%	48%	19%	27%	27%
Second movers	25%	28%	22%	39%	32%	36%	11%	23%	13%
Obs.	30	30	30	30	22	22	30	30	30
Second movers if the first mover:									
Cooperates	37%	35%	34%	48%	52%	54%	21%	27%	13%
Obs.	30	30	18	30	22	13	30	30	8
Defects	8%	13%	10%	12%	7%	8%	6%	17%	14%
Obs.	30	30	25	30	22	13	30	30	22

Note: Cooperation rates and the number of independent observations in *Main* (M), *Control I* (C-I), and *Control II* (C-II). Cooperation rates of first movers are based on the observed behavior in C-I and C-II and on stage-game strategies in M. Cooperation rates of second movers are based on the observed behavior in C-II and on stage-game strategies in C-I and M. To give each pair of subjects the same weight, frequencies are weighted by the inverse of the number of periods played

choices within subjects—that is depending on whether the period of play is the last one or not—does not translate into outcomes different from "hot" decision-making.

Table 2 provides summary statistics of the cooperation rates of both players in all periods and depending on whether they are playing a continuation period or the last period. The cooperation rates of second movers depending on the choice of the first mover are also included. The table shows that cooperation rates are largely similar in the three treatments. In fact, for none of the cooperation rates in Table 2 do we

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

find significant differences across the three treatments (Kruskal-Wallis tests based on independent observations, $p > 0.515$).[16]

The table further shows that first movers cooperate more often than second movers (as reported by Clark and Sefton 2001). Average cooperation rates of first movers are 36–41% and those of second movers are 22–28%. This difference is driven by behavior in continuation periods. In the last periods the difference largely disappears. Wilcoxon signed-ranks tests (WSR tests) confirm that the difference between first and second movers in continuation periods is significant in *Main* ($p = 0.004$) and *Control I* ($p = 0.009$), and is weakly significant in *Control II* ($p = 0.072$)

Also consistent with existent literature, is that cooperation is mostly conditional (see e.g. Keser and van Winden 2000; Fischbacher et al. 2001). In particular, cooperation by second movers is strongly conditioned on the action of the first mover. Average cooperation rates by second movers given that the first mover cooperates are 34–37% and those given that the first mover defects are 13% or lower. Again, the effect is mostly due to continuation periods where the difference in cooperation rates is statistically significant in *Main* (WSR test, $p = 0.001$) and *Control I* (WSR test, $p = 0.002$).[17]

Finally, we can see in Table 2 that endgame effects are prevalent: cooperation rates in continuation periods are higher than those in the last period. The differences are at least weakly significant (WSR tests, $p < 0.001$ in *Main*, $p = 0.048$ in *Control I*, and $p = 0.089$ in *Control II*).

To summarize, we recover results that are common in the cooperation literature, and we find that the use of the strategy method to separate continuation-game from endgame behavior within subjects does not induce different behavior compared to "hot" decision-making.[18] Hence, our proposed method is well-suited to disentangle the different motivations that underly the observed cooperation.

4.2 Strategies

In this subsection we focus on the strategies observed in *Main*. We begin with the first movers' strategies. Always defecting accounts for 39% of the first movers' stage-game strategies. For first movers who cooperate, we find that most do so as long as there is future interaction. In other words, their stage-game strategy consists of cooperating in continuation periods and defecting in the last period: first movers submit this strategy 42% of the time.[19] This finding is interesting as first movers have an incentive to defect in the last period only if they anticipate that a large fraction of second movers cooperate strategically.

[16] If we use Mann-Whitney U tests to do pairwise comparisons between the three treatments, we do not find a statistically significant difference for any of the cooperation rates ($p > 0.320$).

[17] There are too few observations in *Control II* to do a within-subjects WSR test. However, a between-subjects Mann-Whitney U test does result in a statistically significant difference ($p = 0.006$).

[18] In Appendix A.1 we compare *Main* to *Control I* to further show that using the strategy method to condition on whether it is the last period or not, does not affect how subjects condition their choice on the action of the first mover.

[19] Cooperating in continuation periods and also in the last period is the next most-common strategy. It accounts for 14% of the first movers' stage-game strategies.

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

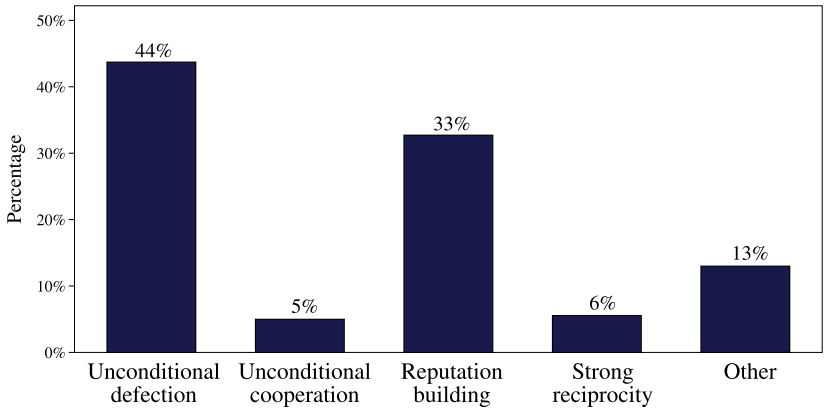


Fig. 2 Distribution of second movers' strategies. *Note:* The bar chart shows the frequency of the strategies used by second movers in *Main*, classified according to Table 1. To give each subject the same weight, frequencies are weighted by the inverse of the number of periods played

Next, we identify the various motivations for cooperation by looking at the elicited strategies of second movers. Figure 2 presents the distribution of second movers' strategies in *Main* using the classification of Table 1. The complete distribution of strategies is available in the Appendix (A.2). Overall, unconditional defection is the most common strategy. It is chosen 44% of the time. However, this still leaves considerable space for strategies that involve some cooperation.

The most frequent strategy that includes some cooperation is reputation building. It accounts for 33% of all strategies. Strong reciprocity is observed 6% of the time. Unconditional cooperation is used less than 5% of the time and all other strategies 13%.[20]

Next, we use the elicited strategies of second movers to describe how the different motivations to cooperate account for two commonly reported behaviors in social dilemma experiments: namely, positive levels of cooperation and the fact that cooperation is mostly conditional on the cooperation of others (i.e., subjects reciprocate). Both of these behavioral patterns are present among second movers in our experiment (see Table 2).

Reputation building is the most common reason for second-movers' *realized* cooperation. Of all the cooperative actions of second movers, 57% are due to reputation building. The other relatively important strategies are unconditional cooperation and strong reciprocity, which respectively account for 17% and 14% of the second movers' cooperative actions.

[20]Two strategies account for around 70% of those in the "Other" category. The first is always defecting if it is not the last period and reciprocating if it is. The second is always cooperating if it is not the last period and reciprocating if it is. Note that we don't find support for cooperation due to correlated equilibria (Stahl 1991)—perhaps due to the lack of a suitable coordination device. In these equilibria, we ought to observe some second movers choosing in continuation periods to cooperate if the first mover defects and defect if the first mover cooperates (and always defect in the last period). This strategy was chosen only once.

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

We turn to the relative importance of the various strategies in explaining the prevalence of conditional cooperation among second movers. To do so, we use the strategies to calculate the contingent choice second movers would have made if they could condition their choice on the action of the first mover but not on whether it was the last period or not. If this had been the case, we would observe that "always defect" and "cooperate only if the first mover cooperates" (i.e., conditional cooperation) respectively account for 61% and 31% of the second movers' contingent choices. Accordingly, conditional cooperation accounts for 79% of the contingent choices that involve some cooperation. Now, if we look at the strategies that are behind conditional cooperation, we find that reputation building is the most common: it produces 64% of all conditional cooperation. In contrast, strong reciprocity produces 12% of all conditional cooperation.

Finally, we evaluate whether it makes sense for second movers to use a reputation building strategy. First, we calculate the expected payoff of using either a reputation building or an unconditional defection strategy given the observed reaction of first movers to the previous behavior of second movers. We find that the expected payoffs of the two strategies are essentially equal: unconditional defection pays 48.1 points and reputation building pays 48.0 points. Second, we show that under simplifying assumptions, own-payoff maximizing second movers prefer reputation building over unconditional defection if more than 2.8% of first movers are strong reciprocators (i.e., they cooperate in period t as long as the second mover cooperated in period $t-1$, see the Appendix for the calculations). Guiding ourselves by the observed percentage of strong reciprocity among second movers, we conclude that reputation building is indeed a plausible strategy.

4.3 Dynamics

We start looking at how subjects change their strategies over time by first checking whether the choices of second movers have an effect on the future behavior of first movers. In particular, we see whether first movers reciprocate the behavior of second movers in the previous period. On average, the cooperation rate of first movers is 49% if the second mover cooperated in the previous period and 25% if the second mover defected. If we concentrate on continuation periods, we find that first movers cooperate at a rate of 79% if the second mover cooperated and 28% if the second mover defected. For last periods, cooperation rates are lower but the same qualitative pattern remains: a cooperation rate of 31% if the second mover cooperated and 17% if the second mover defected.[21] This indicates that cooperation by second movers with strategies such as reputation building does indeed elicit cooperation by first movers.[22]

[21] A probit regression with the decision of the first mover in continuation periods as the dependent variable and the lagged action of the second mover as an independent variable results in a positive and significant coefficient ($\beta = 1.821$, $p < 0.001$). For decisions corresponding to last periods, the same regression results in a positive but not significant coefficient ($\beta = 0.686$, $p = 0.165$). If in additional regressions we control for the previous action of the first mover, we find that these results are driven by first movers who cooperated in the previous period (i.e., first movers who defected do not react to the action of the second mover). Regressions were run using White's heteroskedasticity consistent estimator to cluster on first movers.

[22] The fact that not all first movers cooperate after meeting a cooperating second mover fits within the model of Kreps et al. (1982), which predicts that cooperation partially unravels in a finitely-repeated game.

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

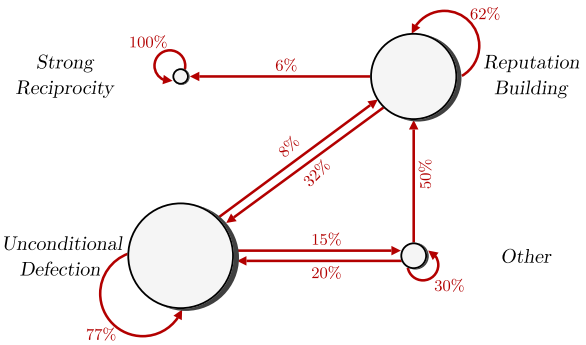


Fig. 3 Stability of strategies. *Note*: The figure shows how second movers change their strategy over time. The size of the circles is proportional to the mean frequency of each strategy in period t. For each strategy, curved arrows indicate the mean fraction of subjects in period t who choose that same strategy in period $t+1$. For each strategy, *straight arrows* indicate the direction and mean fraction of subjects in period t who switch to another strategy in period $t+1$ (missing arrows imply no switching to that particular strategy). To give each subject the same weight, sizes and percentages are weighted by the inverse of the number of periods played

A full picture of the stability of strategies and the way in which second movers switch between strategies is seen in Fig. 3. For simplicity we concentrate on the three most common strategies: unconditional defection, reputation building, and strong reciprocity. All other strategies are placed under "Other." Moreover, given that we are looking at changes over time, the data used includes only second movers who played more than one period. The size of the circles is proportional to the frequency in which each strategy is chosen in period t. For each strategy, the arrows indicate the percentage of subjects that keep the same strategy in the next period, and if not, the percentage of subjects that switch to each other strategy. Specifically, for each strategy, curved arrows indicate the mean fraction of subjects in period t who choose that same strategy in period $t+1$. Similarly, for each strategy, straight arrows indicate the direction and mean fraction of subjects in period t who switch to each other strategy in period $t+1$ (missing arrows imply no switching between to a particular strategy). Hence, for example, 100% of second movers who chose strong reciprocity make the same choice in the next period. As throughout the paper, all calculations are weighted by the inverse of the number of periods played.[23]

From Fig. 3 one can see that all of the three main strategies are quite stable. A second mover who chooses one of these strategies has between a 62% and a 100% chance of choosing the same strategy in the next period. In comparison, the strategies

Specifically, they show that, if the fraction of cooperation that is due to strategically-motivated second movers is high-enough then players play a mixed strategy that implies some first movers will defect even after meeting a cooperating second mover. That some first movers cooperate after meeting a defecting second mover is consistent with subjects using strategies that are more forgiving than the well-known grim-trigger and tit-for-tat strategies. Such strategies can induce higher levels of cooperation if players are prone to making mistakes (for further evidence see Bruttel and Kamecke 2011; Fudenberg et al. 2011).

[23]The reason unconditional cooperation appears in Fig. 2 but not it Fig. 3 is that a majority of the occurrences of this strategy happened for players that played only one period, and hence we do not have enough observations to treat it as a separate category in Fig. 3.

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

that fall within "Other" are considerably less robust: second movers are more likely to switch than to choose one of the strategies in this group for two consecutive periods. Using binomial probability tests and the null hypothesis that the probability of choosing the same strategy in period t and $t+1$ is less than 50% (i.e., a subject is more likely to switch than to choose the same strategy), we can reject it for reputation building, unconditional defection, and strong reciprocity ($p < 0.048$). We do reject the null for strategies under "Other" ($p = 0.813$).[24]

With respect to the motivation of second movers to switch strategies *per se*, besides choosing a strategy under "Other," we do find that neither the previously chosen strategy nor the outcome in the previous period predicts whether a second mover changes strategy or not.[25]

4.4 Disentangling the endgame effect

A nice feature of our experimental method is that it isolates the endgame effect into one (the last) period, which makes studying it much easier. If subjects differ in their ability to perform backward induction, in games with a finite number of periods, it is difficult to know whether a subject defects in periods other than the last period because he managed to backward-induct or due to other reasons such as reciprocating the defection of others (Selten and Stoecker 1986; McKelvey and Palfrey 1992; Katok et al. 2002).

As seen in Table 2, there is a pronounced endgame effect in our experiment. In *Main*, the cooperation rate calculated with the subjects' strategies is considerably higher in continuation periods than in last periods: it drops from 48% to 15% (WSR test, $p < 0.001$). Moreover, there is already an obvious difference in the first period (WSR test, $p < 0.001$), which indicates that, even before they have had the opportunity to interact, subjects make a clear distinction between the two situations.

Generally, endgame effects can be thought as being caused by three distinct reasons. First, individuals who are willing to conditionally cooperate in the last periods of a game switch to defection because they *expect* others will start to defect. Second, individuals who cooperate solely because of the existence of future interaction—what

[24] An alternative to looking at the stability of specific strategies is to think there are two types of individuals: those who are willing to cooperate in the last period (i.e., for non-strategic reasons) and those who are not. This gives us a group of "non-strategically-motivated" strategies and a group of "strategically-motivated" strategies. As a group, strategically motivated strategies are remarkably stable: 89% of second movers who chose one of these strategies in period t do so again in $t+1$. Non-strategically motivated strategies are only slightly less stable: 75% of second movers who select a non-strategically motivated strategy select one again in the next period.

[25] We ran a probit regression with a binary variable indicating whether a subject changes strategy from period t to $t+1$ as the dependent variable. As independent variables we use dummy variables indicating which was the strategy chosen in t and the period number. We find that, compared to choosing unconditional defection, choosing a strategy from "Other" in period t is associated with a 38% higher probability of choosing a different strategy in $t+1$ ($p = 0.009$). A similar regression using dummy variables to indicate the realized outcomes in t (instead of strategies) does not result in significant coefficients. We also do not get significant coefficients for realized outcomes if we run a separate regression for unconditional defection and reputation building (we cannot run a regression for strong reciprocity since it does not vary or for strategies under "Other" since we have too few observations). Regressions were run using White's heteroskedasticity consistent estimator to cluster on second movers.

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

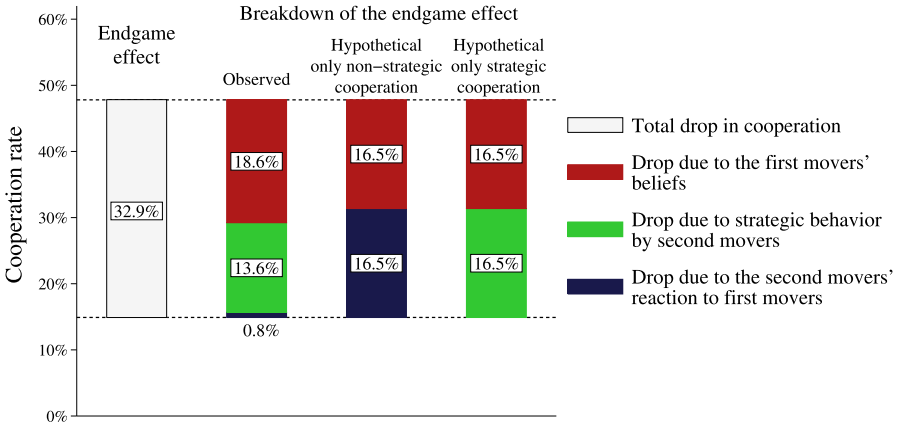


Fig. 4 Causes of the endgame effect. *Note:* The *leftmost bar* shows the total endgame effect: the difference between the cooperation rate in continuation and last periods. The bar labeled "observed" shows the breakdown of the actual endgame effect. To give each subject the same weight, percentages are weighted by the inverse of the number of periods played. The *rightmost bars* show how the breakdown would look like in two hypothetical cases: a case where second movers reciprocate in the last period but first movers defect because they expect defection (bar labeled "only non-strategic cooperation"), and a case where second movers are cooperating strategically and first movers defect because they anticipate their defection (bar labeled "only strategic cooperation")

we call *strategic* cooperation—switch to defection as the prospects of future interaction disappear. Third, individuals who switch to defection because they *experienced* defection due to the previous two causes and as a consequence start defecting themselves.

With our experimental method, we can observe the impact of each of these three reasons. On the one hand, since second movers can condition their choice on the choice of the first mover, their strategies are independent of their expectations of first-mover behavior. Hence, their contribution to the endgame effect is restricted to the second and third reasons, which we can further differentiate because we observe how second movers condition their choices: conditioning between continuation periods and last periods allows us to see how much cooperation drops due to strategic reasons, and conditioning on the action of the first mover (in the last period) allows us to see how much cooperation drops due to experienced defection. On the other hand, given that cooperation by second movers is almost exclusively conditional (i.e., first movers cannot reasonably expect to defect and get the temptation payoff), first movers who switch from cooperation in continuation periods to defection in the last period should be doing so exclusively because they expect second movers will defect more.

Figure 4 depicts the breakdown of the endgame effect in our experiment. The leftmost bar equals the total endgame effect (a drop in the cooperation rate of 32.9 percentage points). The bar labeled "observed" shows the breakdown of the endgame effect into the three reasons discussed above. It shows that slightly more than half of the total endgame effect (18.6 percentage points) can be attributed to the first movers' anticipation of second-mover defection. About 40% of the endgame effect (13.6 percentage points) is due to second movers who were cooperating for strategic reasons, and less then 3% of the endgame effect (0.8 percentage points) is due to

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

second movers reacting to the increase in first-mover defection.[26] WSR tests indicate that the first two reasons are significantly different from zero ($p < 0.002$) whereas the third reason is not ($p = 0.537$). Lastly, to provide a comparison, the rightmost bars depict how the breakdown of the endgame effect would look like in two focal examples (keeping the size of the endgame effect constant). The bar labeled "only non-strategic cooperation" shows the breakdown in a hypothetical world where second movers are willing to reciprocate in the last period but first movers defect because they (wrongly) expect second movers will defect. By contrast, the bar labeled "only strategic cooperation" shows the breakdown in a world where second movers are cooperating strategically and first movers defect because they (correctly) anticipate second movers will defect. As we can see, the actual breakdown closely resembles the case where the endgame effect is solely caused by correctly anticipated strategic cooperation.

5 Conclusions

In this paper, we propose a novel experimental method that disentangles, *within subjects*, strategic from non-strategic behavior in repeated games. We use it to distinguish strategic from non-strategic motivations for cooperation in prisoner's dilemma games. In the experiment, subjects play the strategic equivalent of an indefinitely-repeated sequential prisoner's dilemma game. The novelty of our method is that choices can be conditioned on whether the period of play is the last period or not. This allows us to identify whether a subject is using a strategically-motivated strategy, which conditions behavior on whether it is the last period or not, or a non-strategically-motivated strategy, which does not condition behavior on the endgame. Hence, we extend the method introduced by Fischbacher et al. (2001) and Fischbacher and Gächter (2010) to identify not only different types of non-strategically-motivated cooperators but also types of strategically-motivated cooperators. We validate the use of our method by comparing behavior in treatments with and without the strategy method. We show that individuals' "hot" decisions are not different from the decisions derived by eliciting strategies.

We find that the three most common strategies of second movers are unconditional defection, reputation building, and strong reciprocity. Not only are these the most common strategies, our data suggests that they are also the most stable strategies within subjects. If we concentrate on the strategies that account for the observed cooperation, we find that reputation building is responsible for a majority of all cooperative actions whereas strong reciprocity accounts for much less of cooperative actions. These results can be seen as good news for inducing infinitely-repeated games

[26] This breakdown of the endgame effect is specific to the matching of first and second movers that occurred in the experiment. One can also make the calculations using the cooperation rates in Table 2 (i.e., for any matching). In this case, the total endgame effect is a drop of 30.1 percentage points, the contribution of first mover beliefs is 18.6 percentage points, the contribution of strategic cooperation by second movers is 8.9 percentage points, and the contribution due to the reaction of second movers to more first-mover defection is 2.6 percentage points.

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

in the lab by random termination since they imply that subjects understand the difference between continuation and termination periods and change their behavior accordingly. These results are in line with the idea put forth by Kreps et al. (1982) that a small number of non-strategic cooperators can change the incentives of more strategic individuals and induce a lot of cooperation (see Camerer and Fehr 2006).

Acknowledgements We thank John Duffy, Tore Ellingsen, Dirk Engelmann, Guillaume R. Fréchette, Simon Gächter, Mark Isaac, Charles Noussair, Jan Potters, Arno Riedl, seminar participants at the University of Nottingham, Tilburg University, and WZB Berlin, and participants of the 2007 International ESA Meeting and the 2008 Nordic Conference on Behavioral and Experimental Economics for useful suggestions. We also thank the University of Antwerp for financial support. A previous version of this paper was distributed under the title "Conditional cooperation: Disentangling strategic from non-strategic motivations."

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Additional analysis and data

A.1 Additional analysis of the strategy method

In this subsection, we test whether allowing subjects to condition their choice on whether they are playing the last period or not affects how second movers condition their choice on the action of the first mover. Table 3 presents the frequencies of the four ways in which second movers can condition their choice on the action of the first mover.

As we can see, the distributions are very similar in *Main* and *Control I*. In fact, we cannot reject the hypothesis that choices in all treatments are derived from the same distribution using multinomial probit regressions with treatment dummies as independent variables and using White's heteroskedasticity consistent estimator to cluster on pairs of subjects ($p = 0.825$ in continuation periods and $p = 0.335$ in last periods).

Table 3 Conditional cooperation in *Main* and *Control I*

	Cont. periods		Last periods	
	Main	*Control I*	*Main*	*Control I*
Unconditional cooperation	9%	2%	5%	10%
Cooperate only if first mover cooperates	39%	50%	16%	17%
Cooperate only if first mover defects	3%	5%	1%	7%
Unconditional defection	49%	44%	78%	67%

Note: The table shows, for *Main* and *Control I*, the frequency of each of the four ways in which second movers can condition their choice on the action of the first mover. Frequencies are shown separately for continuation periods and last periods. To give each pair of subjects the same weight, frequencies are weighted by the inverse of the number of periods played

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

Table 4 Distribution of strategies of second movers

Strategy				Frequency
Continuation period		*Last period*		
First mover Cooperates	First mover Defects	First mover Cooperates	First mover Defects	
Defect	Defect	Defect	Defect	43.7%
Defect	Defect	Cooperate	Defect	5.6%
Defect	Defect	Defect	Cooperate	0.0%
Defect	Defect	Cooperate	Cooperate	0.0%
Cooperate	Defect	Defect	Defect	32.7%
Cooperate	Defect	Cooperate	Defect	5.6%
Cooperate	Defect	Defect	Cooperate	0.7%
Cooperate	Defect	Cooperate	Cooperate	0.0%
Defect	Cooperate	Defect	Defect	1.7%
Defect	Cooperate	Cooperate	Defect	1.1%
Defect	Cooperate	Defect	Cooperate	0.0%
Defect	Cooperate	Cooperate	Cooperate	0.0%
Cooperate	Cooperate	Defect	Defect	0.0%
Cooperate	Cooperate	Cooperate	Defect	3.3%
Cooperate	Cooperate	Defect	Cooperate	0.7%
Cooperate	Cooperate	Cooperate	Cooperate	5.0%

Note: The table shows the frequency of the strategies used by second movers in *Main*. To give each subject the same weight, frequencies are weighted by the inverse of the number of periods played

A.2 Additional data of second mover strategies

Table 4 presents the complete distribution of the elicited strategies of second movers in *Main*. Note that the strategies in the first, fifth, sixth, and sixteenth rows correspond respectively to the strategies in Table 1: unconditional defection, reputation building, strong reciprocity, and unconditional cooperation.

A.3 Additional analysis of rational reputation building

Here, we derive the condition under which rational own-payoff maximizing second movers choose reputation building over unconditional defection. Intuitively, second movers reputation build if the fraction of "strong reciprocators" among first movers is sufficiently high. If this is the case, a second mover has a high-enough chance that the last period of the game is reached without a deviation from mutual cooperation and that the first mover then cooperates, which allows the second mover to obtain some gains from cooperation and the temptation payoff.

To simplify, we assume that there are two types of first movers: "strong reciprocators" who cooperate in period t as long as the second mover cooperated in period $t-1$ and irrespective of whether the period played is the last period or not, and "own-payoff maximizers" who do defect in the last period and strategically cooperate in continuation periods (if doing so is profitable). Denote α as the fraction of strongly reciprocating first movers and $1-\alpha$ as the fraction of own-payoff maximizing first movers. A rational own-payoff maximizing second mover prefers reputation

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

building instead of unconditional defection in period $t = 0$ if the expected payoff from reputation building is higher than the one from unconditional defection:

$$\alpha\pi^T + (1-\alpha)\pi^D + \sum_{t=1}^{\infty}\delta^t\pi^C \geq (1-\delta)(\alpha\pi^T + (1-\alpha)\pi^D) + \delta\pi^T + \sum_{t=1}^{\infty}\delta^t\pi^D.$$

Taking $\sum_{t=1}^{\infty}\delta^t = \frac{\delta}{1-\delta}$ and solving for α, gives the condition

$$\alpha \geq \frac{\pi^T - \pi^C - \delta(\pi^T - \pi^D)}{\pi^T - \pi^D - \delta(\pi^T - \pi^D)}.$$

For our choice of parameters, this condition translates to $\alpha \geq 0.028$.

References

Abbink, K. (2004). Staff rotation as an anti-corruption policy: An experimental study. *European Journal of Political Economy*, *20*, 887–906.

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

Andreoni, J., & Croson, R. (2008). Partners versus strangers: Random rematching in public goods experiments. In C. R. Plott & V. L. Smith (Eds.), *Handbook of experimental economics results* (pp. 776–783). Amsterdam: North-Holland.

Andreoni, J., & Miller, J. H. (1993). Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence. *Economic Journal*, *103*, 570–585.

Aoyagi, M., & Fréchette, G. (2009). Collusion as public monitoring becomes noisy: Experimental evidence. *Journal of Economic Theory*, *144*, 1135–1165.

Brandts, J., & Charness, G. (2000). Hot vs. cold: sequential responses and preference stability in experimental games. *Experimental Economics*, *2*, 227–238.

Brandts, J., & Charness, G. (forthcoming). The strategy versus the direct-response method: a first survey of experimental comparisons. *Experimental Economics*.

Brosig, J., Weimann, J., & Yang, C.-L. (2003). The hot versus cold effect in a simple bargaining experiment. *Experimental Economics*, *6*, 79–90.

Bruttel, L. V., & Kamecke, U. (2011, forthcoming). Infinity in the lab: How do people play repeated games? *Theory and Decision*. doi:10.1007/s11238-011-9247-6

Cabral, L., Ozbay, E. Y., & Schotter, A. (2011). *Intrinsic and instrumental reciprocity: An experimental study* (Working paper). New York University.

Camerer, C. F., & Fehr, E. (2006). When does 'economic man' dominate social behavior? *Science*, *311*, 47–52.

Camerer, C., & Weigelt, K. (1988). Experimental tests of a sequential equilibrium reputation model. *Econometrica*, *56*, 1–36.

Cason, T. N., & Mui, V.-L. (1998). Social influence in the sequential dictator game. *Journal of Mathematical Psychology*, *42*, 248–265.

Clark, K., & Sefton, M. (2001). The sequential prisoner's dilemma: evidence on reciprocation. *Economic Journal*, *111*, 51–68.

Cochard, F., & Willinger, M. (2005). Fair offers in a repeated principal-agent relationship with hidden actions. *Economica*, *72*, 225–240.

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

Croson, R. (2000). Thinking like a game theorist: factors affecting the frequency of equilibrium play. *Journal of Economic Behavior & Organization*, *41*, 299–314.

Dal Bó, P. (2005). Cooperation under the shadow of the future: Experimental evidence from infinitely repeated games. *The American Economic Review*, *95*, 1591–1604.

Dal Bó, P., & Fréchette, G. R. (2011). The evolution of cooperation in infinitely repeated games: Experimental evidence. *The American Economic Review*, *101*, 411–429.

Dreber, A., Fudenberg, D., & Rand, D. G. (2011). *Who cooperates in repeated games?* (Working paper). Harvard University.

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press

Duffy, J., & Ochs, J. (2009). Cooperative behavior and the frequency of social interaction. *Games and Economic Behavior*, *66*, 785–812.
Engle-Warnick, J., & Slonim, R. L. (2004). The evolution of strategies in a repeated trust game. *Journal of Economic Behavior & Organization*, *55*, 553–573.
Engle-Warnick, J., & Slonim, R. L. (2006). Inferring repeated-game strategies from actions: Evidence from repeated trust game experiments. *Journal of Economic Theory*, *28*, 603–632.
Falk, A., Fehr, E., & Fischbacher, U. (2005). Driving forces behind informal sanctions. *Econometrica*, *73*, 2017–2030.
Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation and the enforcement of social norms. *Human Nature*, *13*, 1–25.
Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*, 171–178.
Fischbacher, U., & Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public good experiments. *The American Economic Review*, *16*, 541–556.
Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, *71*, 397–404.
Friedman, J. W. (1971). A non-cooperative equilibrium for supergames. *Review of Economic Studies*, *38*, 1–12.
Fudenberg, D., & Maskin, E. (1986). The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, *54*, 533–554.
Fudenberg, D., Rand, D. G., & Dreber, A. (2011, forthcoming). Slow to anger and fast to forget: Cooperation in an uncertain world. *American Economic Review*. Available online at http://www.aeaweb.org/forthcoming/output/accepted_AER.php
Gächter, S., & Falk, A. (2002). Reputation and reciprocity: Consequences for the labour relation. *The Scandinavian Journal of Economics*, *104*, 1–26.
Gächter, S., & Renner, E. (2006). *The effects of (incentivized) belief elicitation in public good experiments* (Discussion paper 2006-16). CeDeX.
Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology*, *206*, 169–179.
Huck, S., Ruchala, G., & Tyran, J.-R. (2006). *Competition fosters trust* (Working paper 06-22). University of Copenhagen.
Katok, E., Sefton, M., & Yavas, A. (2002). Implementation by iterative dominance and backward induction: An experimental comparison. *Journal of Economic Theory*, *104*, 89–103.
Keser, C., & van Winden, F. (2000). Conditional cooperation and voluntary contributions to public goods. *The Scandinavian Journal of Economics*, *102*, 23–39.
Kreps, D. M., Milgrom, P., Roberts, J., & Wilson, R. (1982). Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory*, *27*, 245–252.
Lacomba, J. A., Lagos, F. M., Reuben, E., & van Winden, F. (2010). *On the escalation and de-escalation of conflict* (Working paper). University of Amsterdam.
McKelvey, R. D., & Palfrey, T. R. (1992). An experimental study of the centipede game. *Econometrica*, *60*, 803–836.
Muller, L., Sefton, M., Steinberg, R., & Vesterlund, L. (2008). Strategic behavior and learning in repeated voluntary-contribution experiments. *Journal of Economic Behavior & Organization*, *67*, 782–793.
Normann, H.-T., & Wallace, B. (2006). *The impact of the termination rule on cooperation in a prisoner's dilemma experiment* (Working paper). University College London.
Palacios-Huerta, I., & Volij, O. (2008). Experientia docet: Professionals play minimax in laboratory experiments. *Econometrica*, *76*, 71–115.
Rubinstein, A. (1979). Equilibrium in supergames with the overtaking criterion. *Journal of Economic Theory*, *21*, 1–9.
Selten, R. (1967). Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments. In H. Sauermann (Ed.), *Beiträge zur Experimentellen Wirtschaftsforschung* (pp. 136–168). Tübingen: Mohr (Paul Siebeck).
Selten, R., Mitzkewitz, M., & Uhlich, G. (1997). Duopoly strategies programmed by experienced players. *Econometrica*, *65*, 517–555.
Selten, R., & Stoecker, R. (1986). End behaviour in sequences of finite prisoner's dilemma supergames. *Journal of Economic Behavior & Organization*, *3*, 47–70.
Sobel, J. (2005). Interdependent preferences and reciprocity. *Journal of Economic Literature*, *43*(2), 392–436.
Stahl, D. O. I. (1991). The graph of prisoner's dilemma supergame payoffs as a function of the discount factor. *Games and Economic Behavior*, *3*, 368–384.

https://doi.org/10.1007/s10683-011-9286-4 Published online by Cambridge University Press