IZA

DISCUSSION PAPER SERIES

IZA DP No. 4107

Revisiting Strategic versus Non-Strategic Cooperation

Ernesto Reuben
Sigrid Suetens

April 2009

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Revisiting Strategic versus Non-Strategic Cooperation

Ernesto Reuben
Northwestern University and IZA

Sigrid Suetens
Tilburg University

Discussion Paper No. 4107
April 2009

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 4107
April 2009

ABSTRACT

Revisiting Strategic versus Non-Strategic Cooperation*

We use a novel experimental design to disentangle strategically- and non-strategically-motivated cooperation. By using contingent responses in a repeated sequential prisoners' dilemma with a known probabilistic end, we differentiate end-game behavior from continuation behavior within individuals while controlling for expectations. This design allows us to determine the extent to which strategically-cooperating individuals are responsible for the so-called end-game effect. Experiments with two different subject pools indicate that the most common motive for cooperation in repeated games is strategic and that the extent to which end-game effects are driven by strategically-cooperating individuals depends on the profitability of cooperation.

JEL Classification: C91, D01, D74

Keywords: reputation building, strong reciprocity, conditional cooperation, strategic cooperation

Corresponding author:

Ernesto Reuben
Northwestern University
Kellogg School of Management
2001 Sheridan Rd.
Evanston, IL 60208
USA
E-mail: ereuben@northwestern.edu

* We thank Tore Ellingsen, Dirk Engelmann, Simon Gächter, Mark Isaac, Charles Noussair, Jan Potters, Arno Riedl, seminar participants at the University of Nottingham and Tilburg University, 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."

1 Introduction

Even though there is considerable of experimental evidence suggesting that individuals cooperate in social dilemma games with repeated interaction because of both strategic and non-strategic motivations, the share of cooperation that can be attributed to each of these motivations has not been precisely identified. To a large extent, this is because in repeated games one cannot tell from observing players' choices whether they are cooperating strategically or not. In this paper, we report results from experiments in which we solve this identification problem by eliciting the players' strategies.

That strategic motivations are in play in finitely repeated social dilemma experiments is alluded to by the so-called end-game effect: a sharp decline in cooperation in the last period(s) of a repeated game, particularly in two-player games (see Selten and Stoecker, 1986; Keser and van Winden, 2000). Clearly, cooperation can drop because strategically-motivated individuals, who reciprocate others' cooperation solely when there is future interaction, do not have an incentive to cooperate in the last period.[1] However, it can also drop because non-strategically-motivated individuals, who reciprocate others' cooperation even in the absence of future interaction, believe others will stop cooperating in the last period.[2] In other words, since both strategically- and non-strategically-motivated individuals can cause the decline in cooperation, it is difficult to know what the contribution of each type of motivation is.[3]

The same confounding effect is found in other experiments that offer evidence for strategic cooperation. For example, one could take the increase in cooperation between repeated games and (repetitions of) one-shot games as being caused by strategically-motivated individuals who now have a reason to cooperate.[4] However, this increase can also be driven by non-strategically-

[1]Kreps et al. (1982) propose a model that rationalizes this pattern by assuming incomplete information concerning preferences for cooperation.

[2]Fischbacher et al. (2001) and Fischbacher and Gächter (2009), for example, show that people often reciprocate others' contributions in public good experiments without future interaction.

[3]This problem cannot be easily resolved by eliciting beliefs since only on rare occasions will the differently-motivated individuals exhibit different belief-action combinations. 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.

[4]An increase in cooperation between partner-matching and random-matching treatments has 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., 2008), prisoners' 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,

motivated individuals who cooperate more because they expect that with repeated interaction others will be more cooperative. Similarly, the observation that cooperation is more frequent when it is more profitable can be attributed to strategic behavior and the existence of additional cooperative equilibria.[5] However, the increase in cooperation can also be due to an increase in non-strategically-motivated cooperation that results from intrinsically-motivated individuals who now find cooperation relatively more attractive or from boundedly-rational individuals who make relatively more mistakes.[6]

To distinguish strategic from non-strategic motivations for cooperation, we run experiments where subjects repeatedly play a prisoners' dilemma game—strategically equivalent to the sequential prisoners' dilemma—with a probability of continuation. Specifically, we use the contingent-response 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 playing *is* or *is not* the final 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 final period of the game, and (ii) whether the first mover *cooperates* or *defects*.

This design allows us to simultaneously observe choices in the continuation game and in the end game, and crucially, observe *counterfactual* behavior (of second movers) in the final period. That is, we know whether second movers cooperate or defect in the final period of the game when it is certain that the first mover will *cooperate*. Second movers who are willing to cooperate in the final period must be motivated, at least in part, by non-strategic reasons.[7] In contrast, second movers who always defect in the final period, but who are willing to cooperate in non-final periods, are clearly cooperating for strategic reasons.[8] In addition, this design clearly isolates

Andreoni (1988) finds more cooperation under random matching (for an explanation of these mixed results see Andreoni and Croson, 2003).

[5]For example, in settings described by Friedman (1971).

[6]The inequality aversion model of Fehr and Schmidt (1999), for example, predicts such an effect. So does the quantal-response equilibrium, which assumes that mistakes are less likely if they are more costly (McKelvey and Palfrey, 1995), and the logit model of Anderson et al. (1998), which allows for altruism and errors.

[7]It is not our intention to suggest that individuals who cooperate in the final period are failing to take into account strategic considerations. Instead, we suggest that players who exhibit this behavior are not *exclusively* motivated by strategic considerations.

[8]We can make this inference for second movers as they face no uncertainty with respect to the first mover's behavior. Expectations about the behavior of others do play a role for first movers. Consequently, we cannot fully identify their motivations for cooperation. For this reason, we concentrate most of our analysis on second movers.

the end-game effect. In games with a finite number of periods, differing abilities to perform backward induction (Selten and Stoecker, 1986; McKelvey and Palfrey, 1992; Katok et al., 2002) can distribute the end-game effect over the last few periods. In our experiment, cooperation in non-final and final periods is clearly differentiated.

Another feature of our design is that, just like in finitely-repeated games, mutual cooperation is *not* an equilibrium if it is common knowledge that all players are rational own-payoff maximizers. In this setting, the motivation to strategically cooperate is limited to individuals who believe that sufficiently many others reciprocate cooperative actions for non-strategic reasons (see Kreps et al., 1982).

There are other studies that investigate behavior in the sequential prisoners' dilemma (Brandts and Charness, 2000; Clark and Sefton, 2001) and/or examine motivations for cooperation by eliciting strategies (Fischbacher et al., 2001; Fischbacher and Gächter, 2009).[9] However, they use one-shot settings, which preclude strategic behavior. Other studies use the subjects' behavior to infer their repeated-game strategies (Engle-Warnick and Slonim, 2004, 2006). In these experiments, it is found that subjects use strategies that produce an end-game effect in finitely-repeated games and punishment of free riders that resemble grim-trigger strategies in infinitely-repeated games.[10] However, since both of these results can be due to both strategic and non-strategic motives, these papers cannot differentiate between the two motivations.

Another related study is Muller et al. (2008). It consists of 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. They conclude that strategic behavior has a more pronounced effect than learning in explaining the usually-observed decline in contributions in public good games. Our design improves on theirs for the purpose of identifying strategically- and non-strategically-motivated cooperation. In particular, in their design, choices within each period are simultaneous and therefore affected by expectations. For example, even if subjects are intrinsically motivated to reciprocate cooperation, they might still choose to defect in the second period—even for high first-period contributions—because they believe others cooperated strategically in the first period and will defect in the second.

[9]Muller et al. (2008) provide an extensive review of this literature.

[10]Dal Bó (2005), Dal Bó and Fréchette (2007), and Duffy and Ochs (2009) study cooperation in indefinitely-repeated prisoners' dilemma games. They find high cooperation rates in treatments where mutual cooperation is an equilibrium. Unlike these studies, mutual cooperation is *not* an equilibrium in our experiments (with common knowledge that all players are rational own-payoff maximizers).

On the basis of three treatments, with two different subject pools, we find that most of the observed cooperation is strategically motivated. However, the relative importance of non-strategic motivations increases with the profitability of cooperation. Lastly, with the use of two control treatments we find that using contingent responses does not alter the behavior of the second movers.

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

2 Description of the game and methodology

The game played in the experiment is the strategic equivalent of a repeated sequential prisoners' dilemma. In each period, with probability $(1-\delta)$ 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 design is that, in each period, players can condition their action on whether they are currently playing the last period of the game or whether the game will continue. 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, additionally, 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 the last period of the game or not, and they are informed about the corresponding action of the other player.[11]

As is well known, full cooperation in repeated games with an unknown end 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 profile of play can be sustained as part of a subgame perfect

[11]Players are informed only of the *realized* action of the other player and not of all their contingent choices.

equilibrium (Rubinstein, 1979; Fudenberg and Maskin, 1986). In our game, mutual cooperation by rational own-payoff maximizers is supported if the continuation probability falls above the threshold $\delta^* = (\pi^T - \pi^C)/(\pi^T - \pi^D)$.[12] This follows from the fact that both players always defect when they play the final period of the game, which makes the game in non-final periods equivalent to a game with an unknown end.[13] In this case, mutual cooperation can be sustained by a trigger strategy only if $\delta \geq \delta^*$, in which case the second mover gets a higher payoff by cooperating than by defecting.[14]

In order to match as closely as possible the conditions in a finitely-repeated game, we use parameter values such that mutual cooperation is not an equilibrium if it is common knowledge that all players are rational own-payoff maximizers. This reduces the motivation to strategically cooperate to individuals who believe others will reciprocate cooperation for non-strategic reasons. Specifically, we set the probability of continuation and the payoffs of the stage game such that $\delta < \delta^*$ in all our treatments.[15]

As previously mentioned, our design allows us to observe the stage-game strategies used by second movers.[16] There are two strategies that are of special interest. The first one consists of conditionally cooperating with the first mover irrespective of whether it is the last period or not. 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 conditionally

[12]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.

[13]A general worry with playing games with an unknown end 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 very small decrease in the frequency of strategically motivated cooperation (individuals who are not strategically motivated and those who are already unconditionally defecting are not affected by more discounting).

[14]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.

[15]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 in the laboratory. We report whether there is evidence for these type of strategies in footnote 21. Alternatively, mutual cooperation can be supported in equilibrium if there is a fraction of first and/or second movers who (are believed to) play according to a tit-for-tat strategy (see Kreps et al., 1982). We discuss this possibility in the conclusions.

[16]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.

Table 1: Strategies of second movers

	Not the last period		*Last period*	
Strategy	First mover Cooperates	First mover Defects	First mover Cooperates	First mover Defects
Reputation building	Cooperate	Defect	Defect	Defect
Strong reciprocity	Cooperate	Defect	Cooperate	Defect
Unconditional defection	Defect	Defect	Defect	Defect
Unconditional cooperation	Cooperate	Cooperate	Cooperate	Cooperate

cooperating as long as the last period is not being played and defecting if it is. We refer to this strategy as *reputation building* as 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 strategy of unconditional defection. We should note, however, that we cannot differentiate between second movers who are strategic but defect because they play a defection equilibrium and second movers who defect for non-strategic reasons.

3 Experimental design and procedures

We ran two experiments, each with a different subject pool. Both experiments were conducted in 2007 with z-Tree (Fischbacher, 2007) and lasted about 45 minutes. Subjects were recruited through online recruitment systems. In total, 312 subjects participated in the experiments. 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 giving us one independent observation per second mover. Once the game finished, subjects answered a debriefing questionnaire after which they were paid in cash and dismissed. See the Appendix for the specific treatment parameters of both experiments (A.1) and an example of the instructions (A.2).

Experiment I

The first experiment was run in CentERlab at Tilburg University. It consisted of one treatment where we implemented the game described above, which we refer to as *Tilburg*, and two control treatments. The purpose of *Tilburg* 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 contingent response method.

In principle, it is possible that the use of the contingent response 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 prisoners' 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 contingent responses than without it (Brosig et al., 2003; Falk et al., 2005). For this reason, 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 *Tilburg* except that they no longer submit a decision for both final and non-final periods. In this treatment, 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 separate choices depending on whether the first mover cooperates or not. By comparing choices between this control and *Tilburg*, 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 contingent responses (i.e., they play the game "normally"). Here, subjects are 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 the two control treatments, 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 $\delta^* = 0.61$. We used a computer to randomly determine when the game ends, and in order to make the three treatments more comparable, the random sequence of each pair in *Tilburg* was used for one pair in each control treatment. The average number of periods played equaled 2.73, and average earnings were €9.38.

Experiment II

The second experiment was run at Northwestern University. It consisted of two treatments in which we implemented the game with the full use of contingent responses and used a coin toss to determine whether the game continued or ended. In other words, the continuation probability was equal to $\delta = 0.50$. In the first of the two treatments, *Northwestern High*, we used a high payoff for mutual cooperation such that $\delta^* = 0.56$. In the second treatment, *Northwestern Low*, we used a low payoff of mutual cooperation (keeping everything else constant) such that $\delta^* = 0.72$. The average number of periods played in *Northwestern Low* and *High* was 2.32 and 4.06, respectively, and average earnings equaled $11.20.

These two treatments were run for the following reasons. First, they allow us to check whether the results from experiment I extend to cases where δ is not as close to δ^*. Second, it is interesting to see how the elicited strategies change as one varies the benefit of cooperation. It is well established that even if there is a dominant strategy to defect, subjects cooperate more if it is more profitable to do so (e.g., Isaac and Walker, 1988). With our design we can evaluate how much of this increase is due to strategic and how much to non-strategic cooperation. Third, although the evidence is mixed, there is some worry that the use of economics and business students biases results in experiments involving cooperative behavior (e.g., Marwell and Ames, 1981; Engelmann and Strobel, 2006; Fehr et al., 2006). Since Tilburg's subject pool consists mainly of such students, we ran this experiment in Northwestern University and excluded students who study economics or a related field.[17]

4 Results

In this section we present the experimental results. Throughout the section, we test differences in frequencies using likelihood-ratio χ^2 tests. Since this is our most common test, in order to avoid unnecessary repetition, we simply report the resulting p-values when it is used. Moreover, since different second movers played the game a different number of periods, all reported values that are based on aggregate data across periods are adjusted by the inverse number of periods played. This way, each pair receives an equal weight. Note that none of the qualitative results change if we concentrate on the first period, which was played only once by all second movers.

[17] In Northwestern, the areas of study were: journalism/communication (21%), engineering (21%), biology/chemistry/physics (16%), anthropology/political science/sociology (15%), history/languages/philosophy, (12%), arts (8%), and others (8%).

Table 2: Realized cooperation rates

Note: The table shows realized mutual cooperation rates and realized cooperation rates by role and action of the first mover in the three main treatments. Statistics are weighted by the inverse of the number of periods played by each pair.

	Tilburg	*Northwestern High*	*Northwestern Low*
Mutual cooperation rate	19%	19%	3%
Cooperation rate	33%	28%	18%
of first movers	41%	33%	31%
of second movers	25%	23%	4%
if first mover cooperates	46%	60%	10%
if first mover defects	10%	5%	2%

Summary statistics for the *realized* outcomes are presented in Table 2. As can be seen, subjects cooperate at similar rates in *Tilburg* and *Northwestern High* and somewhat less in *Northwestern Low*. Comparing the two Northwestern treatments, we find mutual cooperation to be significantly more frequent in *Northwestern High* vis-à-vis *Northwestern Low* ($p = 0.025$). The same is true, albeit weakly, for the overall cooperation rate ($p = 0.054$). This finding confirms our expectation that the occurrence of cooperation is sensitive to its profitability. We can also observe that first movers cooperate more often than second movers (as in Clark and Sefton, 2001). Moreover, consistent with existent literature, cooperation by second movers is strongly conditioned on the action of the first mover.[18]

Next, we proceed to the paper's main results. In subsection 4.1, we provide an overview of the subjects' strategies in order to observe their motivations for cooperation. In subsection 4.2, we turn to the causes of the end-game effect. We would also like to stress that behavior in *Tilburg* does not significantly differ from that in the two control treatments. In other words, in this game, we do not see that the use of contingent responses induces different behavior compared to "hot" decision-making. Given that this analysis is not central to our paper, we put it in the Appendix (A.3).

[18] Average cooperation rates between first and second movers are significantly different in all treatments with Wilcoxon signed-ranks tests ($p < 0.009$). Similarly, average cooperation rates by second movers are significantly higher if the first mover cooperates than if the first mover defects (Wilcoxon signed-ranks tests, $p < 0.004$).

4.1 First and second movers' strategies

We begin with the first movers' strategies. A large percentage of the first movers' stage-game strategies consist of always defecting: 39% in *Tilburg*, 53% in *Northwestern High*, and 56% in *Northwestern Low*. For first movers who cooperate, we find that most do so only in non-final periods. In other words, their stage-game strategy consists of cooperating if it is not the last period and defecting otherwise. In *Tilburg*, 75% of first movers who cooperate submit this strategy; in *Northwestern High* it is 59% and in *Northwestern Low* 46%.[19] This 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.

Now, we turn to the second movers' strategies in order to identify their motivation for cooperation. Figure 1 presents the distribution of second movers' strategies in *Tilburg* and the two Northwestern treatments using the classification of Table 1. Overall, unconditional defection is the most common strategy. It is chosen from 28% to 60% of the time. However, there is still considerable space for strategies that involve some cooperation.

In all treatments, the most frequent strategy that includes some cooperation is reputation building. It accounts for around 30% of all strategies.[20] This is even the case in *Northwestern Low* where the continuation probability is well below δ^*. The third most common strategy is strong reciprocity, whose frequency varies between 6% and 23%. Unconditional cooperation is used less than 5% of the time and other strategies between 2% and 14%.[21]

[19]Cooperating in both non-final and final periods is the next most common strategy. It accounts for 18% in *Tilburg*, 26% in *Northwestern High*, and 35% in *Northwestern Low* of the strategies that involve some cooperation.

[20]Reputation building is also the most common reason for second-movers' *realized* cooperation. The fraction of cooperative actions of second movers that are due to reputation building is 57% in *Tilburg*, 32% in *Northwestern High*, and 67% in *Northwestern Low*. Strong reciprocity accounts for 14% of the second movers' cooperation in *Tilburg*, 29% in *Northwestern High*, and 0% in *Northwestern Low*. Other relatively important strategies are unconditional cooperation, which accounts for 17% of cooperative actions in *Tilburg* and 14% in *Northwestern High* (0% in *Northwestern Low*), and the strategy that consists of always cooperating if it is not the last period and conditionally cooperating if it is, which accounts for 14% of cooperative actions in *Northwestern High* (0% in the other two treatments).

[21]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 conditionally cooperating if it is. The second is always cooperating if it is not the last period and conditionally cooperating 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 non-final periods to cooperate if the first mover defects and defect if the first mover cooperates (and always defect in final periods). This strategy was chosen only once by one subject.

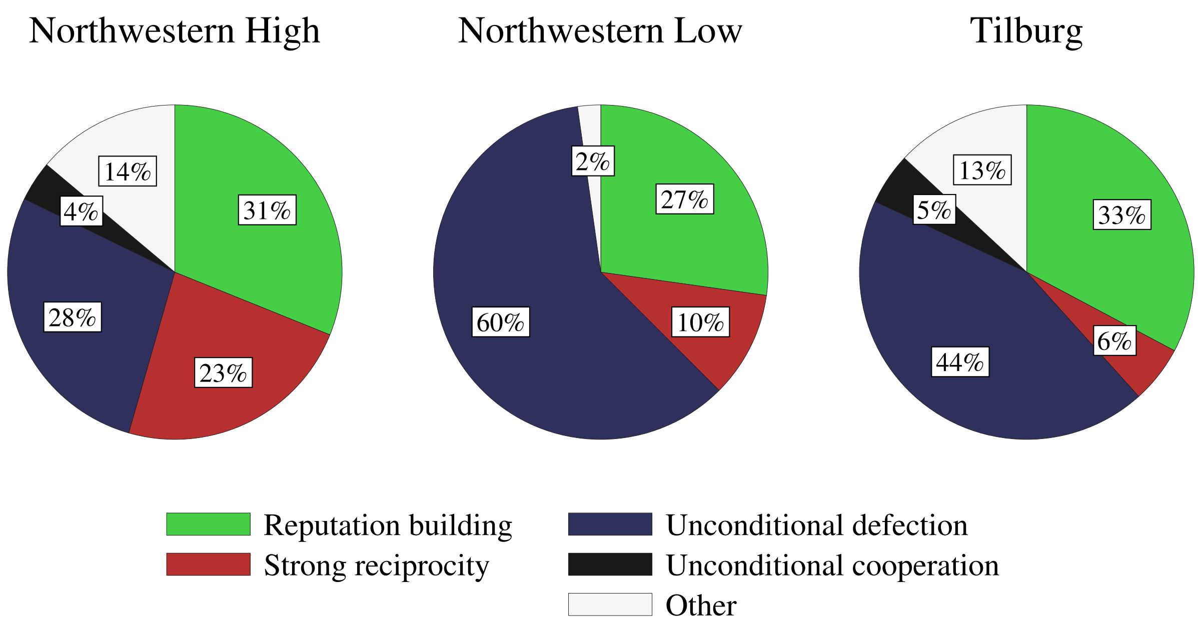

Figure 1: Frequency of strategies

Note: The pie charts show, for each treatment and across all periods, the frequency of strategies used by second movers classified according to Table 1. Strategies are weighted by the inverse number of periods played by each subject.

In Figure 1 one can also observe which strategies are responsible for the higher cooperation rate in *Northwestern High* compared to *Northwestern Low* (the distributions are significantly different, $p = 0.043$). The increase in cooperation rates is driven by a sharp decrease in the frequency of unconditional defection and an increase in strong reciprocity, unconditional cooperation, and strategies under "other." Interestingly, the frequency of reputation building is almost identical.

Next 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 their strategies to calculate the stage-game strategy we would have observed if second movers could condition their choice on the action of the first mover but could not condition on whether it was the last period or not. If this had been the case, we would have observed that most cooperative outcomes are the result of second movers who conditionally cooperate: 76% in *Tilburg*, 86% in *Northwestern High*, and 92% in *Northwestern Low*. Now, if we look at the strategies that are behind the conditional-cooperation pattern, we find that reputation building is the most common. In *Tilburg*, reputation building produces 64% of all conditional cooperation; in *Northwestern High* it produces 48% and

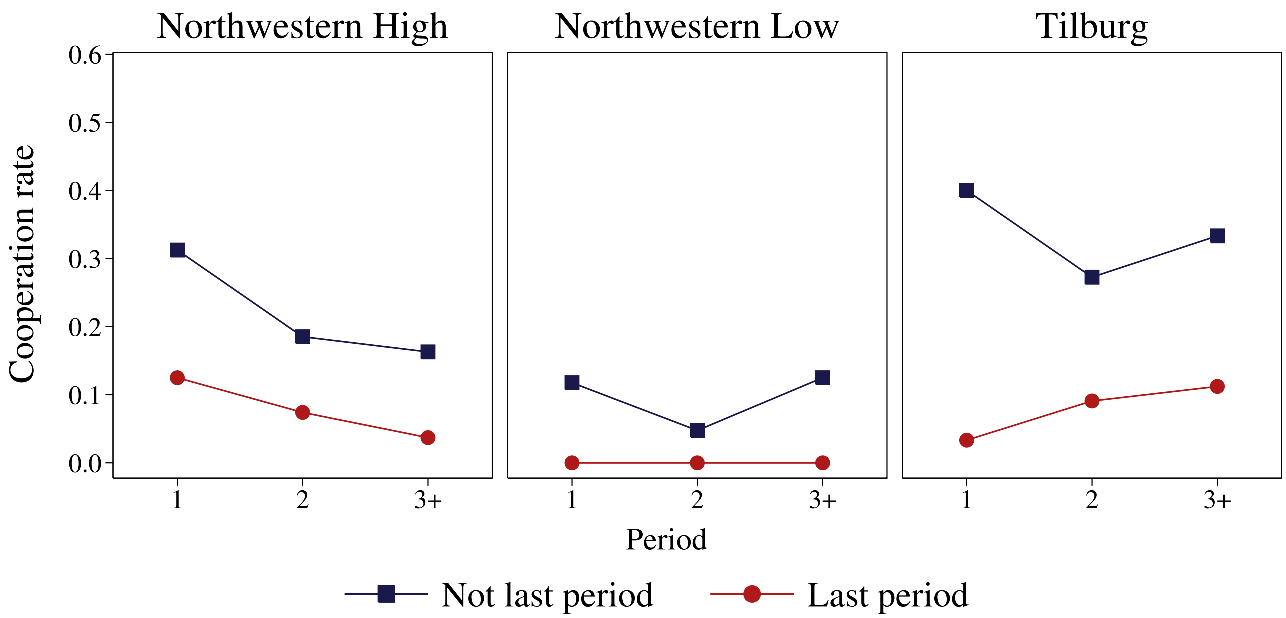

Figure 2: Mutual Cooperation Rates

Note: The figure shows, for each treatment, the rate of mutual cooperation per period calculated using both the subjects' non-final-period action and their final-period action. Rates for period three and up are pooled as by then the number of observations has decreased considerably in some treatments.

in *Northwestern Low* 64%. The respective percentages for strong reciprocity are 12%, 44%, and 36%.

Finally, we find that the strategies listed in Table 1 are fairly stable both across periods and within subjects, whereas other strategies are less stable. This analysis is provided in the Appendix (A.4).

4.2 Disentangling the end-game effect

On average, subjects cooperate less often in final periods compared to non-final periods. The mutual cooperation rate calculated with the subjects' continuation strategies (i.e., their contingent choices in non-final periods) is significantly higher in all treatments than the mutual cooperation rate calculated with their end-game strategies (i.e., the subjects' contingent choices in final periods): in *Tilburg*, it drops from 33% to 4%, in *Northwestern High* it drops from 25% to 10%, and in *Northwestern Low* it drops from 10% to 0% (Wilcoxon signed-ranks tests, $p < 0.026$).[22] This drop can be observed in Figure 2 for each period $t \in \{1, 2, 3+\}$. From the figure, one can see that already in the first period there is an obvious difference between behavior in non-final and final periods,[23] which indicates that, even before they have had the opportunity to interact, subjects make a clear distinction between the two situations.[24]

Generally, the end-game effect occurs due to two distinct causes. First, individuals that cooperate solely because of the existence of future interaction switch to defection in the hope of getting the temptation payoff. Second, individuals that are willing to conditionally cooperate even in the final period of a game switch to defection because they expect others will now defect. In our design, each role (first and second mover) is affected by only one of these causes. On the one hand, since second movers can condition their choice on that of the first mover, their strategies are independent of their expectation of first-mover cooperation. On the other hand, given that cooperation is almost exclusively conditional, the first movers' willingness to cooperate should be driven exclusively by their expectation of second-mover cooperation.

To observe the impact of each of the two causes, we calculate the effect on the mutual cooperation rate of varying one cause while keeping the other constant. In other words, to observe the impact on the end-game effect of the strategic behavior of second movers we: (i) compare the cooperation rate calculated with the continuation strategy of both first and second movers to the cooperation rate calculated with the continuation strategy of first movers and the end-game strategy of second movers, and (ii) compare the cooperation rate calculated with the end-game strategy of both first and second movers to the cooperation rate calculated with the end-game strategy of first movers and the continuation strategy of second movers. To observe the effect of expectations, we do the equivalent comparisons. That is, we: (i) compare the cooperation rate calculated with the continuation strategy of both first and second movers to

[22]We concentrate on mutual cooperation rates calculated using the subjects' strategies. However, the strong end-game effect is also evident with *realized* rates of mutual cooperation: depending on whether it is the last period or not, mutual cooperation drops from 36% to 3% in *Tilburg*, from 19% to 9% in *Northwestern High*, and from 6% to 0% in *Northwestern Low* ($p < 0.001$).

[23]Wilcoxon signed-ranks tests of differences in the mutual cooperation rate in $t = 1$ depending on whether it is calculated with continuation or with end-game strategies yield $p = 0.001$ in *Tilburg*, $p = 0.083$ in *Northwestern High*, and $p = 0.046$ in *Northwestern Low*.

[24]We do not find evidence in any of the treatments for a time trend in the mutual cooperation rate (calculated either with continuation or end-game strategies). In all treatments, Spearman's rank correlation coefficients between mutual cooperation rates and periods are not significantly different from zero ($p > 0.204$).

Table 3: Causes of the end-game effect

Note: The table shows the mutual cooperation rate calculated using either the continuation or the end-game strategy of first movers in combination with either the continuation or end-game strategy of second movers. ≫ and > indicate, at 5% and 10% respectively, statistically significant differences using Wilcoxon signed-ranks tests based on independent observations.

		First mover's strategy								
		Tilburg			*Northwestern High*			*Northwestern Low*		
		Cont.		End	Cont.		End	Cont.		End
Second mover's strategy	Cont.	33%	≫	13%	25%	≫	15%	10%		10%
		⩔	⇘	⩔		⇘		∨	⇘	⩔
	End	10%	>	4%	20%	≫	10%	1%		0%

the cooperation rate calculated with the end-game strategy of first movers and the continuation strategy of second movers, and (ii) compare the cooperation rate calculated with the end-game strategy of both first and second movers to the cooperation rate calculated with the continuation strategy of first movers and the end-game strategy of second movers.

These comparisons can be seen in Table 3 where we present the mutual cooperation rates calculated with the different combinations of continuation and end-game strategies. In the table, we also indicate statistically significant differences based on Wilcoxon signed-ranks tests. In *Tilburg*, both strategic behavior and expectations cause a significant drop in the mutual cooperation rate. In *Northwestern High*, we find that whereas expectations cause a significant drop in cooperation, strategic behavior does not. Conversely, in *Northwestern Low* strategic behavior causes a significant drop in cooperation but expectations do not.

All in all, this analysis demonstrates that the end-game effect is caused both by the strategic behavior of second movers and by the expectation of such behavior by first movers. However, the relative impact of each of these causes changes with the profitability of cooperation. In *Northwestern Low*, where the ratio of non-strategic to strategic cooperation is low, the end-game effect is mostly due to the disappearance of strategic incentives to cooperate. In *Northwestern High*, where this ratio is high, the end-game effect is mostly due to overly pessimistic expectations by first movers.[25]

[25] This difference might disappear if first movers have opportunities to update beliefs about the distribution of second movers, for example, by playing the indefinitely repeated prisoners' dilemma game repeatedly and receiving

5 Conclusions

In this paper we provide evidence of the relative importance of strategic versus non-strategic motivations for cooperation in social dilemmas. We report the results of a laboratory experiment where subjects play the strategic equivalent of a repeated sequential prisoners' dilemma with a probabilistic end. The novelty of our design is that choices can be conditioned on whether the period of play is the final period or not, which allows us to identify the frequency of strategically-motivated and non-strategically-motivated cooperation.[26]

We find that the behavior of second movers is basically the result of three strategies: (i) unconditional defection, which accounts for 28% to 60% of all strategies, (ii) reputation building, which accounts for 27% to 33% of all strategies, and (iii) strong reciprocity, which accounts for between 6% and 23% of all strategies. As reported in the Appendix (A.4), we also find that this distribution of strategies is fairly stable in time and within subjects.

If we concentrate on second movers who cooperated, we find that the strategy with the highest weight is reputation building—it accounts for between 32% and 67% of the realized cooperative actions—followed by strong reciprocity, which accounts for between 0% and 29% of cooperative actions.

Although in our experiment mutual cooperation is not supported in equilibrium when it is common knowledge that all players are rational own-payoff maximizers, the observed share of reputation building could be explained with the model of Kreps et al. (1982). In this model, reputation building pays off due to the existence of tit-for-tat players, which in our analysis could be seen as subjects who use the strong reciprocity strategy. Since the 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 whether the observed amount of reputation building is consistent with the model's predictions.[27]

We also find that an increase in the payoff of mutual cooperation increases the frequency of strong reciprocity. This finding is consistent with the idea that individuals possess social preferences—which induce them to act as strong reciprocators—but nevertheless react to changes

information about outcomes of other pairs.

[26] In the Appendix (A.3) we show that using contingent responses does not affect behavior compared to the direct-response method.

[27] 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).

in financial incentives (e.g., Fehr and Schmidt, 2006). The finding is also consistent with a quantal-response equilibrium that treats strong reciprocity as an error which is more likely to be made the lower its cost (i.e., the higher the cooperation payoff relative to the temptation payoff).[28]

Interestingly, the increase in the payoff of mutual cooperation also increases the ratio of reputation builders to unconditional defectors (from 45% to 75%). This change is roughly consistent with Kreps et al. (1982) since an increase in the share of strong reciprocators in a population makes reputation building more profitable for the remaining "rational" players, and therefore, they ought to be more likely to reputation-build as opposed to unconditionally defect.

Lastly, we decompose the causes of the end-game effect and demonstrate that cooperation drops due to both strategic behavior and the anticipation of such behavior. Interestingly, depending on the profitability of cooperation (and possibly the subject pool),[29] the cause of the end-game effect changes.

[28]Goeree et al. (2002) integrate the two interpretations (social preferences and errors) in a one-shot public good game.

[29]For example, Bohnet et al. (2008) and Gächter et al. (2008) are recent studies that find differences in cooperation and trust across different cultures.

A Appendix

A.1 Overview of experiments and treatments

Table A1: Overview of experiments and treatments

	Experiment I			Experiment II	
	Tilburg	*Control I*	*Control II*	*Northwestern High*	*Northwestern Low*
Parameters					
Mutual defection π^D	15	15	15	15	15
Unilateral defection π^T	33	33	33	33	33
Unilateral cooperation π^S	10	10	10	6	6
Mutual cooperation π^C	22	22	22	23	20
Probability of continuation δ	0.60	0.60	0.60	0.50	0.50
Threshold for cooperation δ^*	0.61	0.61	0.61	0.56	0.72
Characteristics					
Location	Tilburg	Tilburg	Tilburg	Northwestern	Northwestern
Field of study	Economics	Economics	Economics	Not economics	Not economics
Number of second movers	30	30	30	32	34

Note: 10 points equaled €1.50 in experiment I and \$2.00 in experiment II; amounts exclude a show-up fee of \$6 in experiment II.

A.2 Instructions

These are the instructions for the *Northwestern High* treatment. The instructions for other treatments are very similar and are available from the authors upon request.

General

You are participating in an experiment on economic decision making and will be asked to make a number of decisions. If you follow the instructions carefully, you can earn money. At the end of the experiment, you will be paid your earnings in private and in cash.

You are not allowed to communicate with other participants. If you have a question, raise your hand and one of us will help you.

During the experiment your earnings will be expressed in points. Points will be converted to US dollars at the following rate: *10 points = $2.00*.

The experiment is *strictly anonymous*: that is, your identity will not be revealed to others and the identity of others will not be revealed to you.

In the experiment, participants will be randomly divided into groups of 2 participants. You will therefore be in a group with one other participant. The composition of the groups will *remain the same during the entire experiment*.

In each group, one participant will be randomly assigned to the *first mover* position. The other participant in the group will be in the *second mover* position. Your position as first or second mover will remain the same during the entire experiment.

Your decision in each period

The experiment is divided into periods. In each period, both the first and the second mover make a choice between option A and option B. The first mover makes his/her decision first. Thereafter the second mover makes his/her decision. The following table shows what the first and second movers earn (in points) depending on their choices:

	first mover's earnings	second mover's earnings
both choose A	23	23
first mover chooses A and the second mover chooses B	6	33
first mover chooses B and the second mover chooses A	33	6
both choose B	15	15

Number of Periods

For each group, the number of periods of in the experiment is determined *randomly*. At the end of each period, we will throw a coin to determine whether that period was the last period of the experiment or whether the experiment continues (heads means the experiment continues and tails means the experiment ends). Thus, in every period the probability that the experiment continues is 50% and the probability that the experiment ends is 50%. Your total earnings in the experiment will equal the sum of earnings across all periods.

After each period, you will receive feedback concerning the decision of the other participant in your group and on your earnings.

The decision of the first mover

In each period, the first mover makes his/her decision in each of the two following situations:

- Do you choose A or B if the current period *is not* the final period (in other words the experiment proceeds to a next period)?
- Do you choose A or B if the current period *is* the final period (in other words the experiment does not proceed)?

If the result of the coin toss is that the experiment continues (heads), then earnings in that period will depend on the answer to the first question. If the result of the coin toss is that the experiment ends (tails), then earnings depend on the answer to the second question.

The decision of the second mover

In each period, the second mover makes his/her decision in each of the four following situations:

- *If the first mover chooses A:* Do you choose A or B if the current period *is not* the final period (in other words the experiment proceeds to a next period)?
- *If the first mover chooses A:* Do you choose A or B if the current period *is* the final period (in other words the experiment does not proceed)?
- *If the first mover chooses B:* Do you choose A or B if the current period *is not* the final period (in other words the experiment proceeds to a next period)?
- *If the first mover chooses B:* Do you choose A or B if the current period *is* the final period (in other words the experiment does not proceed)?

If the result of the coin toss is that the experiment continues (heads) and the first mover chooses A, earnings will depend on the answer to the first question. If the result of the coin toss is that the experiment ends (tails) and the first mover chooses A, earnings will depend on the answer to the second question. If the result of the coin toss is that the experiment continues (heads) and the first mover chooses B, earnings will depend on the answer to the third question. If the result of the coin toss is that the experiment ends (tails) and the first mover chooses B, earnings will depend on the answer to the fourth question.

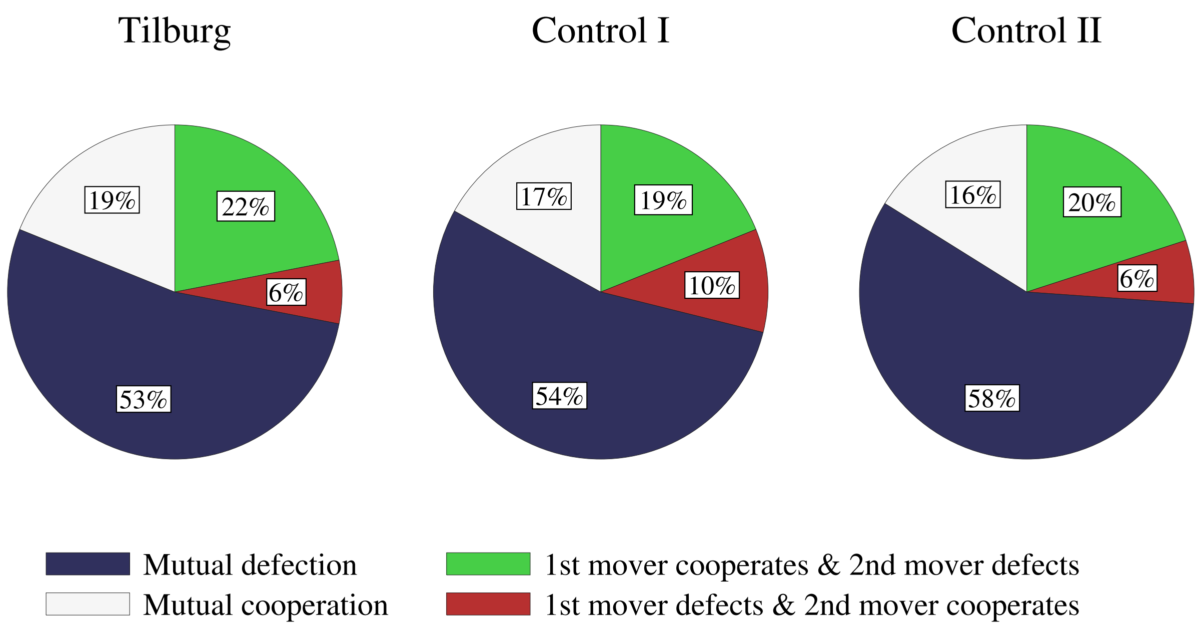

Figure A1: Comparing outcomes with control treatments

Note: The pie charts show the frequency of each of the four possible outcomes in *Tilburg* and the two control treatments. Outcomes are weighted by the inverse of the number of periods played by each pair.

A.3 Control treatments

To ensure that the use of the contingent-response method does not result in different behavior than the direct-response method, we use this subsection to compare behavior in *Tilburg* and the two control treatments.

Figure A1 gives an overview of the distribution of *realized* outcomes in the three treatments. It is clear that outcomes are highly similar. This is corroborated if we test for equality of distributions across the three treatments ($p = 0.992$).[30]

Next, we compare *Tilburg* with *Control I* to determine whether the elicited stage-game strategies change when they are conditioned on whether it is the last period or not. One could worry that conditioning on the final period might trigger more strategic thinking than otherwise, and therefore, if contingent responses are elicited, there could be less conditional cooperation in the last period. However, we do not find this to be the case. As is shown in the top part of Table A2, the frequencies of stage-game strategies for both treatments for non-final and final

[30] We do not get statistical significance either if we do pairwise comparisons between treatments ($p > 0.910$), or if we compare separately across treatments the frequency of each strategy ($p > 0.745$).

Table A2: Comparing strategies with control treatments

Note: The top half of the table shows, for *Tilburg* and the control treatments, the distribution of the second movers' stage-game strategies depending on whether it is the last period or not. The bottom half shows the second movers' cooperation rates depending on the first mover's choice (actual cooperation rates for *Control II* and cooperation rates implied by the stage-game strategies in the other two treatments). Strategies and rates are weighted by the inverse of the number of periods played by each pair.

	if not last period			if last period		
	Tilburg	*Control I*	*Control II*	*Tilburg*	*Control I*	*Control II*
Comparing stage-game strategies						
Always cooperate	9%	2%		5%	10%	
Cooperate if first mover cooperates	39%	50%		16%	17%	
Cooperate if first mover defects	3%	5%		1%	7%	
Always defect	49%	44%		78%	67%	
Comparing cooperation rates						
When the first mover cooperates	48%	52%	54%	21%	27%	13%
When the first mover defects	12%	7%	8%	6%	17%	14%

periods are very similar (distributions are not significantly different: $p = 0.642$ for non-final periods and $p = 0.581$ for final periods). Furthermore, in *Control I* we do not see a higher frequency of conditional cooperation in final periods.[31]

Lastly, we compare cooperation rates between *Control II* and the other two treatments to test whether second movers' choices are affected by conditioning them on the first movers' action. The actual cooperation rates in *Control II* and the ones implied by the strategies in the other treatments are seen in the bottom part of Table A2. We do not find statistically significant differences between the frequencies of the three treatments when running tests that compare separately non-final periods and final periods depending on whether the first mover cooperates or defects ($p > 0.428$). This also holds if we do pairwise tests between treatments ($p = 0.202$). Thus, we conclude that, for our purpose, the contingent response or strategy method is a valid technique since it does not induce different behavior compared to "hot" decision-making.

[31] There are no significant differences if we compare separately the frequency of each stage-game strategy across treatments ($p > 0.268$).

A.4 Stability of strategies

We briefly analyze the stability over time of the distribution of strategies. In all treatments, the frequencies of strategies do not change considerably across periods. If we test in each treatment for equality of distributions between the first three periods, we find no significant differences ($p > 0.570$). This finding tells us that the relative influences of reputation building and strong reciprocity vary little with repetition.[32] Next, we check whether this relative stability is hiding substantial changes at the individual level.

In order to analyze the stability of strategies within each subject, we take a look at how often subjects choose the same strategy. Specifically, we calculate the probability that a second mover picks in period t the same strategy that she picked in period $t-1$. Overall, second movers pick the same strategy for consecutive periods 64% of the time in *Tilburg*, 72% in *Northwestern High*, and 86% in *Northwestern Low*. The stability of individual strategies can be seen in Table A3 where this probability is calculated separately for each strategy and treatment. From the table, one can see that the three main strategies (reputation building, strong reciprocity, and unconditional defection) are quite stable. A second mover who chooses one of these strategies has around an 80% chance of choosing the same strategy in the next period. In comparison, the strategies that fall within "other" are considerably less robust. In most cases, these strategies are chosen for only one period at a time.[33] With respect to the motivation of second movers to switch strategies, besides choosing a strategy under "other," we do not find that either the previously chosen strategy or the outcome in the previous period has a significant effect.[34] In summary, strategies are fairly stable both across periods and within subjects. A majority of subjects consistently

[32]We do not find significant differences even if we do, for each treatment and strategy, pairwise comparisons between each of the three periods ($p > 0.110$).

[33]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 (weakly) reject it in all treatments for reputation building ($p < 0.001$), unconditional defection ($p < 0.001$), and strong reciprocity ($p < 0.056$). For unconditional cooperation it is rejected in *Northwestern High* ($p = 0.032$). Treating strategies under 'other' as a group, we cannot reject the null in any treatment ($p > 0.998$).

[34]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. We used the following independent variables: dummy variables for the strategy chosen in t, dummy variables for the realized outcome in t, treatment indicator variables, and the period number. We find that choosing a strategy from "other" in period t is associated with a 32% higher probability of choosing a different strategy in $t+1$ ($p = 0.001$, using White's heteroskedasticity consistent covariance matrix estimator to cluster on each subject). However, we find no other significant effect. We get the same result if we run a separate regression for each strategy or for each treatment.

Table A3: Stability of strategies

Note: The table shows, for each strategy and treatment, the probability that a second mover picks in period t the same strategy that she picked in period $t-1$. Probabilities are weighted by the inverse of the number of periods played by each subject.

	Reputation building	Strong reciprocity	Unconditional defection	Unconditional cooperation	Other
Tilburg	62%	100%	77%	–	0%
Northwestern High	84%	88%	68%	80%	35%
Northwestern Low	88%	80%	88%	–	0%
All treatments	77%	89%	78%	80%	21%

chose one of the strategies in Table 1, while other strategies are chosen less consistently.

References

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

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

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

Andreoni, J. and Croson, R. (2003). Partners versus strangers: Random rematching in public goods experiments. In Plott, C. R. and Smith, V. L., editors, *Handbook of Experimental Economics Results*, volume forthcoming.

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

Bohnet, I., Greig, F., Herrmann, B., and Zeckhauser, R. (2008). Betrayal aversion: Evidence from Brazil, China, Switzerland, Turkey, the United Arab Emirates and the United States. *American Economic Review*, 98:294–310.

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

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

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

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

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

Cochard, F. and 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. *Economic 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. *American Economic Review*, 95:1591–1604.

Dal Bó, P. and Fréchette, G. R. (2007). The evolution of cooperation in infinitely repeated games: Experimental evidence. Working paper 2007-7, Brown University.

Duffy, J. and Ochs, J. (2009). Cooperative behavior and the frequency of social interaction. *Games and Economic Behavior*, forthcoming.

Engelmann, D. and Strobel, M. (2006). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments: Reply. *American Economic Review*, 96:1918–1923.

Engle-Warnick, J. and Slonim, R. L. (2004). The evolution of strategies in a repeated trust game. *Journal of Economic Behavior & Organisation*, 55:553–573.

Engle-Warnick, J. and Slonim, R. L. (2006). Inferring repeated-game strategies from actions: Evidence from repeated trust game experiments. *Economic Theory*, 28:603–632.

Falk, A., Fehr, E., and Fischbacher, U. (2005). Driving forces behind informal sanctions. *Econometrica*, 73:2017–2030.

Fehr, E., Fischbacher, U., and Gächter, S. (2002). Strong reciprocity, human cooperation and the enforcement of social norms. *Human Nature*, 13:1–25.

Fehr, E., Naef, M., and Schmidt, K. M. (2006). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments: Comment. *American Economic Review*, 96:1912–1917.

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114:817–868.

Fehr, E. and Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism–experimental evidence and new theories. In Kolm, S. C. and Ythier, J. M., editors, *Handbook on the Economics of Giving, Reciprocity and Altruism*, pages 615–691. Elsevier, Amsterdam.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10:171–178.

Fischbacher, U. and Gächter, S. (2009). Social preferences, beliefs, and the dynamics of free riding in public good experiments. *American Economic Review*, forthcoming.

Fischbacher, U., Gächter, S., and 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. and Maskin, E. (1986). The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54:533–554.

Gächter, S. and Falk, A. (2002). Reputation and reciprocity: Consequences for the labour relation. *Scandinavian Journal of Economics*, 104:1–26.

Gächter, S., Herrmann, B., and Thöni, C. (2008). Antisocial punishment across societies. *Science*, 319:1362–1367.

Gächter, S. and 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.

Goeree, J. K., Holt, C. A., and Laury, S. K. (2002). Private costs and public benefits: Unraveling the effects of altruism and noisy behavior. *Journal of Public Economics*, 83:255–276.

Huck, S., Ruchala, G., and Tyran, J.-R. (2006). Competition fosters trust. Working paper 06-22, University of Copenhagen.

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

Katok, E., Sefton, M., and Yavas, A. (2002). Implementation by iterative dominance and backward induction: An experimental comparison. *Journal of Economic Theory*, 104:89–103.

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

Kreps, D. M., Milgrom, P., Roberts, J., and 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., and van Winden, F. (2008). After the war: An experimental study of post-conflict behavior. Working paper, University of Amsterdam.

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

McKelvey, R. D. and Palfrey, T. R. (1992). An experimental study of the centipede game. *Econometrica*, 60:803–836.

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

Muller, L., Sefton, M., Steinberg, R., and Vesterlund, L. (2008). Strategic behavior and learning in repeated voluntary-contribution experiments. *Journal of Economic Behavior and Organization*, 67:782–793.

Palacios-Huerta, I. and 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 Sauermann, H., editor, *Beiträge zur Experimentellen Wirtschaftsforschung*, pages 136–168. J.C.B. Mohr (Paul Siebeck), Tübingen.

Selten, R. and Stoecker, R. (1986). End behaviour in sequences of finite prisoner's dilemma supergames. *Journal of Economic Behavior & Organisation*, 3:47–70.

Stahl, D. O. I. (1991). The graph of prisoners dilemma supergame payoffs as a function of the discount factor. *Games and Economic Behavior*, 3:368–84.