

Instrumental reciprocity as an error

Ernesto Reuben

Sigrid Suetens



Instrumental reciprocity as an error

Ernesto Reuben

IZA, LISER, and New York University Abu Dhabi, e-mail: ereuben@nyu.edu

Sigrid Suetens

Tilburg University, e-mail: s.suetens@uvt.nl

Abstract

We study the strategies used by experimental subjects in repeated sequential prisoners' dilemma games to identify the underlying motivations behind instrumental reciprocity, that is, reciprocation of cooperation only if there is future interaction. Importantly, we designed the games so that instrumental reciprocity is a mistake for payoff-maximizing individuals irrespective of their beliefs. We find that, despite the fact that instrumental reciprocity is suboptimal, it is one of the most frequently used cooperative strategies. Moreover, although the use of instrumental reciprocity is sensitive to the costs of deviating from the payoff-maximizing strategy, these costs alone cannot explain the high frequency with which subjects choose to reciprocate instrumentally.

This version: August 2018.

JEL codes: C72, C92, D00.

Keywords: reciprocity, reputation building, instrumental reciprocity, strategic cooperation, cooperation heuristics

Note: This is the author's version of a work that was accepted for publication in *Games*. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. A final version is published in <https://doi.org/10.3390/g9030066>.

1 Introduction

Experiments have shown that individuals often use reciprocal strategies in repeated games and seem well aware of the fact that reciprocity can be used instrumentally. Namely, when playing with the same partners, reciprocating cooperation in early repetitions of a finitely repeated game before switching to defection often leads to higher earnings than simply defecting from the beginning. For example, it is often observed that subjects cooperate more frequently if they know that they will play at least once more with each other than if they know that they are playing one last time, the so-called “end-game effect” (e.g., Selten and Stoecker, 1986; Sonnemans, Schram, and Offerman, 1999; Reuben and Suetens, 2012; Cabral, Ozbay, and Schotter, 2014). More recently, in their meta-analysis of finitely-repeated prisoners’ dilemma games, Embrey, Fr chet te, and Yuksel (2018) establish that most subjects converge to using strategies that reciprocate cooperation until a threshold repetition, after which they start defecting.¹

Why would subjects use reciprocal strategies instrumentally in games with a known end? The most common view about instrumental reciprocity is that it is used by players who want to maximize their own material payoff and who are sophisticated enough to understand that, in finitely-repeated games, the presence or believed presence of cooperative player types results in the existence of equilibria with high levels of cooperation, at least in early repetitions of the game (as shown by the seminal paper of Kreps, Milgrom, Roberts, and Wilson, 1982).² However, it is also possible that instrumental reciprocity instead reflects the use of general reputation-building heuristics individuals have learned to apply over the course of their lives (e.g., see Delton et al., 2011). Deciding whether it is optimal or not to reciprocate someone’s cooperation in situations where there is possible future interaction is not a trivial task, even for calculative individuals. In this view, subjects use instrumental reciprocity strategies even if they do not maximize material payoffs. Previous repeated-game experiments, including the ones cited above, cannot differentiate between these two views because in all of them it is possible to rationalize cooperation through a reputation-building framework   la Kreps, Milgrom, Roberts, and Wilson (1982). In the current paper, we design

¹Other often-cited evidence for instrumental reciprocity is the observation that cooperation rates are typically higher in games where subjects play repeatedly with the same partners than in games where partners change after each repetition (see Zelmer, 2003; Andreoni and Croson, 2008).

²A general reputation-building argument need not specify why cooperative types choose to cooperate. One of the most common explanations is that some players have social preferences and that is why they cooperate (e.g., as argued by Andreoni and Miller, 1993; Camerer and Fehr, 2006). However, the same logic applies if cooperative types are cooperating due to other reasons, such as inability to backward induct (e.g., Selten and Stoecker, 1986), having naive prior beliefs (Cox, Jones, Pflum, and Healy, 2015), or because they are prone to make mistakes (e.g., Palfrey and Prisbrey, 1997).

an experiment that allows us to observe the use of instrumental reciprocity among experienced subjects in a setting where we are certain that it is an error from the point of view of material-payoff maximization.

The key features of our design are as follows. First, we use a sequential prisoners' dilemma (SPD) game and allow for possible future interaction by repeating the game once with a known probability. Second, to detect instrumental reciprocity and avoid the confounding effects of beliefs, we elicit the strategies of second movers. Specifically, we allow second movers to condition their choice in each repetition of the SPD game on whether the first mover cooperates or defects. This design allows us to identify players who use instrumental reciprocity because they are willing to reciprocate cooperation by first movers in the first repetition of the SPD game but plan to defect if first movers cooperate in the second repetition of the SPD game (if played). We refer to this strategy as *reciprocate then defect*. Third, we choose payoffs in the game such that instrumental reciprocity is not a rational strategy for second movers who maximize material payoffs irrespective of their beliefs about the behavior of their matched first mover.

Our results can be summarized as follows. In line with the literature, we find that almost all of the second movers' cooperative behavior in the SPD games is accounted for by reciprocal strategies. Overall, about 80% of the cooperative strategies of experienced players can be attributed to two reciprocal strategies: *reciprocate then defect* and the strategy to *always reciprocate* (which is basically tit-for-tat, i.e. reciprocate cooperation in both the first and the second repetition of the game). Instrumental reciprocity is particularly prevalent, corresponding to 47% of the observed cooperative strategies, if the gains to mutual cooperation are relatively high. If the gains to mutual cooperation are low, so that instrumental reciprocity becomes a more costly error from the point of view of a material payoff maximizer, the strategy *reciprocate then defect* accounts for just 22% of the cooperative strategies. Finally, we find that the strategy *always reciprocate* is not sensitive to the gains to cooperation.

2 Experimental design and procedures

Experimental game

In the experiment, pairs of participants play a SPD. The SPD is played once with certainty (period 1), and it is played a second time by the same two players (period 2) with a known continuation probability equal to 0.5. When playing the game, first movers make their decisions using the direct-response method: at the beginning of period 1 and period 2 (if played), they choose to cooperate (*c*) or to defect (*d*). Second movers make their decisions using the contingent-response method:

they can condition their choice on the decision of the first mover and on the period being played.³ In particular, they are asked to choose to cooperate or to defect in four cases: (i) in period 1 when the first mover cooperates, (ii) in period 1 when the first mover defects, (iii) in period 2 when the first move cooperates, and (iv) in period 2 when the first mover defects.

The game develops as follows. After first movers make their initial decisions and second movers submit their strategies, players learn the outcome of the game in period 1. Subsequently, a coin toss is used to determine whether period 2 is played. If period 2 is played, first movers submit their second choice, which is then matched with the strategy that had already been entered by the second mover. Thereafter, both players learn the choice made by the partner (not the strategy) and the outcome of the game in period 2.

This method allows to directly observe the second movers' strategy choice. To be precise, we observe Markov strategies since second movers can condition their choice in period 2 on the first mover's decision in period 2 but not on decisions in period 1.⁴ Table 1 shows the sixteen possible strategies second movers may adopt, highlighting four strategies that are of interest. Two of these strategies refer to the above-mentioned reciprocal strategies: *reciprocate then defect* (instrumental reciprocity) and *always reciprocate*.

Treatments

We implement two treatments, SPD-High and SPD-Low, which differ solely in the gains from mutual cooperation. Both treatments have the same mutual defection payoff (equal to 25 points), the same temptation payoff (equal to 50 points), the same sucker payoff (equal to 9 points), but a different mutual cooperation payoff (equal to 37 points in SPD-High and to 30 points in SPD-Low).

Given that we are interested in whether individuals reciprocate when it is not optimal to do so from a monetary perspective, the payoffs in the two treatments were selected so that second movers can never maximize their material payoff in the two-period SPD by choosing either reciprocal strategy. Specifically, in both treatments, the expected material payoff of the strategy to *always defect* dominates *reciprocate then defect*, which in turn dominates *always reciprocate*.⁵ To see this, first note that in period 2 a second mover is never worse off from a monetary perspective if she defects,

³Previous experimental evidence suggests that decisions made with the contingent-response method are not different than those made with the direct-response method in prisoner's dilemma games and the like (Muller, Sefton, Steinberg, and Vesterlund, 2008; Brandts and Charness, 2011; Reuben and Suetens, 2012).

⁴We decided to elicit Markov strategies because it simplifies the instructions and previous experimental work has demonstrated that the vast majority of observed strategies are described by Markov strategies (see for example, Selten et al., 1997; Bruttel and Kamecke, 2012).

⁵In section A of the appendix we show that *always defect* dominates all the other strategies of second movers.

Table 1. Possible strategies of second movers

First mover's action in:	Period 1		Period 2	
	<i>c</i>	<i>d</i>	<i>c</i>	<i>d</i>
Second mover's strategies:				
<i>always defect</i>	<i>d</i>	<i>d</i>	<i>d</i>	<i>d</i>
<i>reciprocate then defect</i>	<i>c</i>	<i>d</i>	<i>d</i>	<i>d</i>
	<i>d</i>	<i>c</i>	<i>d</i>	<i>d</i>
	<i>d</i>	<i>d</i>	<i>c</i>	<i>d</i>
	<i>d</i>	<i>d</i>	<i>d</i>	<i>c</i>
<i>always reciprocate</i>	<i>c</i>	<i>c</i>	<i>d</i>	<i>d</i>
	<i>c</i>	<i>d</i>	<i>c</i>	<i>d</i>
	<i>c</i>	<i>d</i>	<i>d</i>	<i>c</i>
	<i>d</i>	<i>c</i>	<i>c</i>	<i>d</i>
	<i>d</i>	<i>c</i>	<i>d</i>	<i>c</i>
	<i>d</i>	<i>d</i>	<i>c</i>	<i>c</i>
	<i>c</i>	<i>c</i>	<i>c</i>	<i>d</i>
	<i>c</i>	<i>c</i>	<i>d</i>	<i>c</i>
	<i>c</i>	<i>d</i>	<i>c</i>	<i>c</i>
	<i>d</i>	<i>c</i>	<i>c</i>	<i>c</i>
<i>always cooperate</i>	<i>c</i>	<i>c</i>	<i>c</i>	<i>c</i>

Note: The table shows the possible strategies a second mover can use in the experiment. Strategies depend on the period (period 1 or period 2) and on the first mover's choice (*c* or *d*).

which implies that *reciprocate then defect* weakly dominates *always reciprocate* (both strategies imply the same actions in period 1). Second, to see that *always defect* dominates *reciprocate then defect*, define p_1 as the probability the first mover cooperates in period 1. For a first mover who cooperates in period 1, define p_2^{cc} as the probability that he cooperates in period 2 given that the second mover cooperates in period 1 and p_2^{cd} as the probability that he cooperates in period 2 given that the second mover defects in period 1. Similarly, for a first mover who defects in period 1, define p_2^{dc} as the probability that he cooperates in period 2 given that the second mover cooperates in period 1 and p_2^{dd} as the probability that he cooperates in period 2 given that the second mover defects in period 1. The expected payoff of *reciprocate then defect* for a second mover equals

$$(1 - p_1) \left[25 + \frac{1}{2} \left(50p_2^{dd} + 25 \left(1 - p_2^{dd} \right) \right) \right] + p_1 \left[X + \frac{1}{2} (50p_2^{cc} + 25 (1 - p_2^{cc})) \right], \quad (1)$$

where $X = 30$ in SPD-Low and $X = 37$ in SPD-High. The expected payoff of *always defect* equals

$$(1 - p_1) \left[25 + \frac{1}{2} \left(50p_2^{dd} + 25 \left(1 - p_2^{dd} \right) \right) \right] + p_1 \left[50 + \frac{1}{2} \left(50p_2^{cd} + 25 \left(1 - p_2^{cd} \right) \right) \right]. \quad (2)$$

It is easy to see that (1) is smaller than (2) as long as $X < 50 - 25\frac{1}{2}(p_2^{cc} - p_2^{cd})$, which is true for any $p_1, p_2^{dd}, p_2^{cc}, p_2^{cd} \in [0, 1]$ if $X < 37.5$.

The payoffs of SPD-Low and SPD-High were chosen to systematically vary the magnitude by which the expected payoff of *reciprocate then defect* and that of *always defect* differ. In the “best” case for *reciprocate then defect* — that is, when in period 1 the first mover cooperates and in period 2 he reciprocates period-1 cooperation by the second mover, so that $p_1 = 1$, $p_2^{cc} = 1$ and $p_2^{cd} = 0$ — the expected payoff of (1) in SPD-Low equals 55 points and that of (2) 62.5 points. By contrast, in SPD-High, the expected payoff of (1) equals 62 points and that of (2) 62.5 points.

Procedures

The experiment took one hour and was conducted in the laboratory of Northwestern University using z-Tree (Fischbacher, 2007). Participants were contacted through an online recruitment system. In total, 70 students participated in PD-Low and 72 in PD-High. Each student participated only once in a session of 10 to 12 people. After their arrival, participants drew a card to be randomly assigned to a seat in the laboratory and consequently to a treatment and role. Once everyone was seated, participants were given the instructions of the experiment. The instructions were written with neutral language. Participants were informed that the experiment consists of multiple parts and that the instructions for the subsequent parts would be provided after the first part had finished. After reading the instructions, they answered control questions to corroborate their understanding of the game. Thereafter, participants learned whether their role would be that of the first or the second mover. They kept the same role throughout the experiment. At the end of the session, participants were paid in private. Mean earnings were \$16.10 and ranged from \$11.70 to \$24.40.

In the first part of the experiment participants play the one-or-two-period SPD once. We refer to these data as coming from *inexperienced* participants. In the second part, participants play the game 15 times with randomly-matched opponents. After each repetition they are informed of the choice(s) of their partner and their own payoff in that repetition. The reason why we included the second part is that participants might need some experience to fully understand the use of reciprocal strategies (as argued by Selten and Stoecker, 1986; Cox, Jones, Pflum, and Healy, 2015; Embrey, Fréchette, and Yuksel, 2018). We refer to data based on the last five repetitions of the SPD in the second part as coming from *experienced* participants. See section C of the appendix for more detailed experimental procedures and a sample of the instructions.

Table 2. Cooperation rates for inexperienced and experienced second movers

Period	SPD-Low				SPD-High			
	Inexperienced		Experienced		Inexperienced		Experienced	
	1	2	1	2	1	2	1	2
If first mover cooperates	23%	23%	21%	19%	31%	19%	35%	19%
If first mover defects	6%	9%	2%	1%	11%	11%	3%	3%

Note: The table shows cooperation rates of inexperienced and experienced second movers in SPD-Low and SPD-High. The label “inexperienced” refers to behavior when participants play the experimental game for the first time. The label “experienced” refers to behavior in the last five repetitions of the experiment.

3 Results

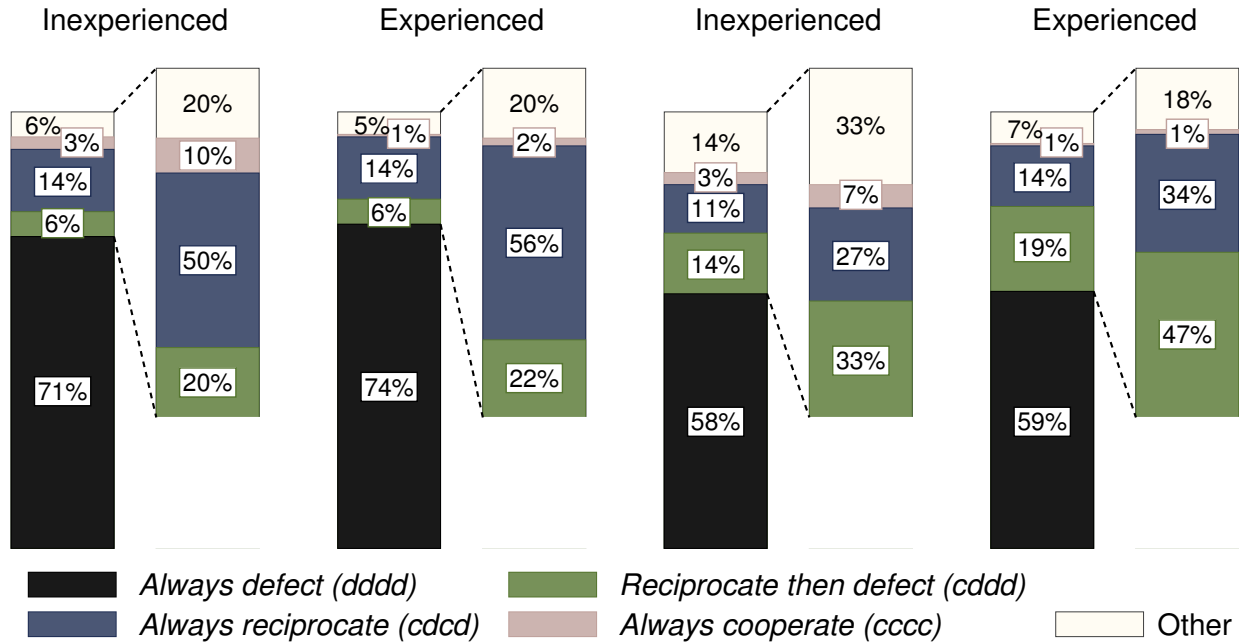
Cooperation and reciprocation rates

We distinguish between behavior when participants are “inexperienced” (in the first part of the experiment) and behavior once they have had time to learn (the last five repetitions of the second part), which we label as “experienced”. The observed cooperation rate of first movers is 14% in SPD-Low and 26% in SPD-High when participants are inexperienced and drops to respectively 5% and 9% once they have gained some experience. The realized cooperation rate of second movers is 10% in SPD-Low and 17% in SPD-High when inexperienced and drops to respectively 4% and 7% when experienced.

For our purposes, however, it is of more interest to look at how second movers condition their behavior on the period and action of the first mover. These conditional cooperation rates are provided in Table 2. From the table, it is clear that second movers reciprocate the first mover’s choice. A probit regression clustering standard errors on independent observations (sessions) confirms that second movers in both treatments cooperate significantly more when the first mover cooperates than when the first mover defects in period 1 ($p < 0.004$).⁶ In period 2, second movers significantly reciprocate in all cases ($p < 0.040$) except for inexperienced second movers in SPD-High ($p = 0.359$). Overall, these results suggest an important role for reciprocal strategies. We now turn to the paper’s main results.

⁶Throughout the results section, we report p -values from regressions used to test whether the frequency of various strategies and actions significantly differ. In all regressions, we cluster standard errors on sessions since errors may be correlated because participants are randomly re-matched within sessions. Section B of the appendix contains the output of all regressions and the precise description of each regression. Given that there is some concern about session effects and clustering in laboratory experiments (e.g., Fréchet, 2012), we checked whether our results hold if instead we cluster standard errors on subjects. We find that they do (see the appendix for details). Finally, the appendix also contains the results of the equivalent nonparametric tests.

Figure 1. Distribution of observed strategies of second movers
 SPD–Low SPD–High



Note: The figure shows distributions of all observed strategies (the bar on the left) and also zooms in on the distributions of observed strategies that involve some cooperation by second movers (the bar on the right). The label “inexperienced” refers to behavior when players play the experimental game for the first time. The label “experienced” refers to behavior in the last five repetitions of the experiment.

Strategies of second movers

The observed distribution of strategies of second movers by treatment is shown in Figure 1. The figure also shows the distribution of strategies conditional on them containing some cooperation. Although the most common strategy is *always defect*, which is the payoff-maximizing choice, a large share of the strategies involves some cooperation or reciprocation even after second movers have had plenty of opportunities to learn: 41% if strategies in SPD-High and 26% in SPD-Low. Next to *always defect*, *reciprocate then defect* and *always reciprocate* are the two most common strategies (the complete distribution of observed strategies is available in section B of the appendix). To illustrate, for experienced players, these two strategies account for 33% of all strategies in SPD-High and 20% in SPD-Low.

If we concentrate on strategies that include cooperative actions by second movers, we can see that *reciprocate then defect* and *always reciprocate* account for over 60% of the strategy choices for inexperienced second movers, and this percentage goes up to 78% once second movers have gained experience.⁷ A similar picture emerges if we concentrate on the second movers’ realized cooperation.

⁷This is well in line with evidence from infinitely repeated prisoner’s dilemmas showing that tit-for-tat is one of

In SPD-High, *reciprocate then defect* is the most common strategy behind the realized cooperation while in SPD-Low it is *always reciprocate*. In the same way, but not illustrated in the figure, these two strategies account for most of the second movers' reciprocity (i.e., the willingness in a given period to cooperate if the first mover cooperates and to defect if the first mover defects). Specifically, for both experienced and inexperienced second movers, they account for over 81% of the strategies that include reciprocity. In SPD-High, *reciprocate then defect* is the most common strategy behind reciprocation while in SPD-Low it is *always reciprocate*.

Looking at the change in the frequency at which strategies are used over time, we find that *reciprocate then defect* and *always reciprocate* are not used less frequently over time. Instead, the fraction of times second movers use these strategies increased slightly by the end of the experiment. In contrast, other strategies involving cooperation become less prevalent over time.

Finally, to test whether the distribution of strategies differs significantly depending on the treatment, we run a multinomial probit regression clustering standard errors on sessions (see section B of the appendix for details). We find a significantly lower frequency of *always defect* and a significantly higher frequency of *reciprocate then defect* in SPD-High compared to SPD-Low for experienced participants ($p = 0.044$ and $p = 0.025$ respectively, $p > 0.169$ for the other cases).

Is instrumental reciprocity a simple mistake?

Our results thus far are consistent with payoff-maximizing individuals mistakenly choosing strategies consistent with instrumental reciprocity as long as doing so is not too costly. In particular, the difference between SPD-High compared to SPD-Low in the frequency in which *reciprocate then defect* is used suggests that the costs of making mistakes plays a role in strategy choice. The fact that behavior is sensitive to errors but that individuals are less likely to make mistakes when doing so is more costly has been observed in many different games (McKelvey and Palfrey, 1995; Goeree et al., 2016). However, the high fraction of second movers choosing *reciprocate then defect* relative to other strategies suggests that the cost of deviating from the payoff-maximizing strategy is not sufficient to explain the observed distribution of strategies.

To evaluate whether the cost of deviation from profit maximization alone is enough to explain the high frequency of *reciprocate then defect*, we calculated the expected monetary payoff of each of strategy second movers could choose when we pit it against the distribution of choices made by first movers (see section B of the appendix for details). We find evidence that *reciprocate then defect* is chosen too often compared to the cost one incurs by choosing it. In particular, the second movers' strategy to defect in period 1 and reciprocate in period 2 (*ddcd*) has a higher expected payoff than

the most common cooperative strategies (Dal Bo and Fréchet, 2018).

reciprocate then defect but it is chosen considerably less often: it is chosen 3% of the time in both SPD-High and SPD-Low.⁸ This finding suggests that the popularity of *reciprocate then defect* is not simply the result of mere confusion or random mistakes due to low costs of deviating from the payoff-maximizing strategy.

4 Conclusion

We report the results of an experiment where individuals play a sequential prisoners' dilemma with possible future interaction where *always defect* is the only rational strategy for second movers who maximize material payoffs, even if they believe that first movers are reciprocators. We find that a large fraction of the strategies adopted by second movers in this game involve some cooperation or reciprocation: 41% of the strategies with high gains to mutual cooperation and 26% of strategies with low gains to mutual cooperation. Of the cooperative strategies, *reciprocate then defect* (i.e., instrumental reciprocity) and *always reciprocate* are clearly the most common ones.

In addition to finding that these types of reciprocal strategies are used, we also find that the use of instrumental reciprocity is sensitive to the gains from mutual cooperation. When the payoff of mutual cooperation is relatively high, *reciprocate then defect* accounts for 19% of all strategies of experienced second movers. In comparison, when the payoff of mutual cooperation is relatively low, *reciprocate then defect* accounts for just 6% of all strategies. By contrast, *always reciprocate* is not responsive to changes in the gains to cooperation in our setting: it accounts for 14% of all strategies in both cases.

The findings related to the strategy *always reciprocate* may not come as a surprise since the use of this strategy has been extensively discussed in the literature (it is often referred to as “strong reciprocity” Gintis, 2000). Individuals who strongly reciprocate can be modeled as being motivated by social preferences (for a review see Fehr and Schmidt, 2006), as following a social norm (Bicchieri, 2006; Boyd and Richerson, 2005), or as acting in accordance with a relatively hard-wired heuristic to conditionally cooperate (Axelrod, 1984; Rand, Peysakhovich, Kraft-Todd, Newman, Wurzbacher, Nowak, and Greene, 2014).

In contrast, for instrumental reciprocity we think that our findings are more surprising. *Reciprocate then defect* is typically thought of as a strategy adopted by sophisticated material-payoff maximizers who realize that cooperating in early periods to build a reputation and then defecting in later periods is in their best interest (Kreps, Milgrom, Roberts, and Wilson, 1982). Since in our

⁸The same is true for *always reciprocate*. It has a lower expected payoff than the strategy *ddcd*, but it is used considerably more often.

experiment *reciprocate then defect* is strictly dominated by *always defect* for any belief the second movers can hold, and subjects played the game multiple times, which gives them the opportunity to learn, our findings suggest that reputation-building strategies are at least partly chosen for reasons other than rational material-payoff maximization.

What are the reasons that a significant number of *experienced* second movers use instrumental reciprocity in our experiment? As mentioned previously, our findings suggest that choosing *reciprocate then defect* is not simply due to random mistakes by rational material-payoff maximizers, even if one takes into account the costs of deviating from the payoff-maximizing strategy (*always defect*). We believe that this leaves us with two broad explanations. The simplest explanation for the occurrence of instrumental reciprocity in our setting is that it is a pre-established heuristic to reputation-build that might be well-adapted for everyday life but happens to be ill-suited for this particular experiment (in line with the arguments found in Todd and Gigerenzer, 1999; Delton, Krasnow, Cosmides, and Tooby, 2011). A basic reputation-building heuristic explains why *reciprocate then defect* is chosen *per se*, but it is less straightforward to see why it would predict a difference between games with high and low gains to mutual cooperation. To predict this difference, one would need a model that explains when the heuristic is more or less likely to be used. Models of dual processes have been developed to explain the use of social heuristics for reciprocal strategies such as *always reciprocate* (e.g., see Rand et al., 2014). Hence, a fruitful line for future theoretical research would be to develop a similar model for reputation-building heuristics. In this respect, we think that an interesting extension to our work is to study how changes in the parameters of the game affect the popularity of *reciprocate then defect* when the cost of deviating from payoff maximization is kept constant. In particular, it would be interesting to vary the continuation probability in order to vary the salience of future interaction whilst changing the payoff of mutual cooperation such that the difference between *reciprocate then defect* and *always defect* is the same across treatments.⁹ If the continuation probability has an effect on the frequency of instrumental reciprocity in this setting, it would provide further evidence that the use of this strategy is driven by more than payoff maximization.

Another potential explanation for the use of *reciprocate then defect* is that it is chosen by utility-maximizing individuals with “weak” social preferences. For example, it can be shown that there exist beliefs for which *reciprocate then defect* is the utility-maximizing strategy of second

⁹For example, the payoff difference between *reciprocate then defect* and *always reciprocate* in a SPD with a continuation probability of 0.50 and a payoff of mutual cooperation of 37 points, as in SPD-High, can also be attained in a SPD with a continuation probability of 0.23 and a payoff of mutual cooperation of 44 points, or a SPD with a continuation probability of 0.77 and a payoff of mutual cooperation of 30 points.

movers who are averse to advantageous inequality but who are not adverse enough to prefer *always reciprocate* (see the inequity aversion model of Fehr and Schmidt, 1999).¹⁰ The advantage of this explanation is that, within one utility-maximizing framework, one can potentially explain the use of both *reciprocate then defect* and *always reciprocate* as well as changes in the popularity of these two strategies depending on the parameters of the game. The drawback of this explanation is that it is difficult to derive the precise equilibrium beliefs and strategies of players if one assumes individuals with selfish, “weak”, and “strong” social preferences coexist. Hence, it is not surprising that, to the best of our knowledge, there is no theoretical work that shows that the interaction between these three types can result in an equilibrium where all the three commonly-observed strategies (*always defect*, *reciprocate then defect*, and *always reciprocate*) are used.¹¹ Showing the conditions for such an equilibrium is an interesting question for future research.

References

- Andreoni, J. and R. Croson (2008). Partners versus Strangers: Random Rematching in Public Goods Experiments. In C. R. Plott and V. L. Smith (Eds.), *Handbook of Experimental Economics Results*, Chapter 82, pp. 776–783. Amsterdam: North-Holland.
- Andreoni, J. and J. H. Miller (1993). Rational cooperation in the finitely repeated prisoner’s dilemma: Experimental evidence. *The Economic Journal* 103, 570–585.
- Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.
- Bicchieri, C. (2006). *The Grammar of Society: The Nature and Dynamics of Social Norms*. New York: Cambridge University Press.
- Boyd, R. and P. J. Richerson (2005). Solving the Puzzle of Human Cooperation. In S. C. Levinson and P. Jaisson (Eds.), *Evolution and Culture*, pp. 105–132. Cambridge, MA: MIT Press.
- Brandts, J. and G. Charness (2011). The strategy versus the direct-response method: A first survey of experimental evidence. *Experimental Economics* 14, 375–398.

¹⁰Consider the example where second movers believe that their matched first mover is a reciprocator with certainty (i.e., $p_1 = 1$, $p_2^c = 1$ and $p_2^d = 0$). Second movers with this belief and with Fehr-Schmidt preference $\beta \in [0.03, 0.31]$ in SPD-High or $\beta \in [0.37, 0.48]$ in SPD-Low, derive a higher expected utility from *reciprocate then defect* (which gives second movers an expected utility of $EU = c + 0.5(t - \beta(t - s))$) than from both *always defect* (which gives $EU = t - \beta(t - s) + 0.5d$) and *always reciprocate* (which gives $EU = c + 0.5c$). Second movers with $\beta > 0.31$ in SPD-High or $\beta > 0.48$ in SPD-Low derive a higher expected utility from *always reciprocate*, and second movers with $\beta < 0.03$ in SPD-High or $\beta < 0.37$ in SPD-Low derive a higher expected utility from *always defect*.

¹¹Alternatively, one can always assume that second movers hold out-of-equilibrium beliefs, in which case it is not hard to find beliefs for which the three types choose different strategies (e.g., see footnote 10). However, making this assumption makes this explanation quite similar to simply assuming that different individuals use different social heuristics.

- Bruttel, L. V. and U. Kamecke (2012). Infinity in the lab: How do people play repeated games? *Theory and Decision* 72, 205–219.
- Cabral, L., E. Ozbay, and A. Schotter (2014). Intrinsic and instrumental reciprocity: An experimental study. *Games and Economic Behavior* 87, 100–121.
- Camerer, C. F. and E. Fehr (2006). When does ‘economic man’ dominate social behavior? *Science* 311, 47–52.
- Cox, C. A., M. T. Jones, K. E. Pflum, and P. J. Healy (2015). Revealed reputations in the finitely-repeated prisoners’ dilemma. *Economic Theory* 58, 441–484.
- Dal Bo, P. and G. R. Fréchet (2018). Strategy choice in the infinitely repeated prisoners’ dilemma. Working paper, New York University.
- Delton, A. W., M. M. Krasnow, L. Cosmides, and J. Tooby (2011, aug). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Sciences* 108(32), 13335–13340.
- Embrey, M., G. R. Fréchet, and S. Yuksel (2018). Cooperation in the finitely repeated prisoner’s dilemma. *Quarterly Journal of Economics* 133, 509–551.
- Fehr, E. and K. M. Schmidt (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114, 817–868.
- Fehr, E. and K. M. Schmidt (2006). The economics of fairness, reciprocity and altruism—Experimental evidence and new theories. In S. C. Kolm and J. M. Ythier (Eds.), *Handbook on the Economics of Giving, Reciprocity and Altruism*, pp. 615–691. Amsterdam: Elsevier.
- Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10, 171–178.
- Fréchet, G. R. (2012, sep). Session-effects in the laboratory. *Experimental Economics* 15(3), 485–498.
- Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology* 206, 169–179.
- Goeree, J. K., C. A. Holt, and T. R. Palfrey (2016). *Quantal Response Equilibrium A Stochastic Theory of Games*. Princeton, NJ: Princeton University Press.
- Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson (1982). Rational cooperation in the finitely repeated prisoners’ dilemma. *Journal of Economic Theory* 27, 245–252.
- McKelvey, R. D. and T. R. Palfrey (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior* 10, 6–38.
- Muller, L., M. Sefton, R. Steinberg, and L. Vesterlund (2008). Strategic behavior and learning in repeated voluntary-contribution experiments. *Journal of Economic Behavior and Organization* 67, 782–793.
- Palfrey, T. R. and J. E. Prisbrey (1997). Anomalous behavior in public goods experiments: How much and why? *American Economic Review* 87, 829–846.
- Rand, D. G., A. Peysakhovich, G. T. Kraft-Todd, G. E. Newman, O. Wurzbacher, M. A. Nowak, and J. D. Greene (2014, dec). Social heuristics shape intuitive cooperation. *Nature Communications* 5(1), 3677.
- Reuben, E. and S. Suetens (2012). Revisiting strategic versus non-strategic cooperation. *Experimental Economics* 15, 24–43.
- Selten, R., M. Mitzkewitz, and G. Uhlich (1997). Duopoly strategies programmed by experienced players.

- Econometrica* 65, 517–55.
- Selten, R. and R. Stoecker (1986). End behaviour in sequences of finite prisoner’s dilemma supergames. *Journal of Economic Behavior & Organisation* 3, 47–70.
- Sonnemans, J., A. Schram, and T. Offerman (1999). Strategic behavior in public good games: when partners drift apart. *Economics Letters* 62, 35–41.
- Todd, P. M. and G. Gigerenzer (1999). *Simple Heuristics that Make Us Smart*. Oxford: Oxford University Press.
- Zelmer, J. (2003). Linear public goods experiments:A meta-analysis. *Experimental Economics* 6, 299–310.