Journal of Economic Behavior & Organization 116 (2015) 481–499

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

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

It's the thought that counts: The role of intentions in noisy repeated games

David G. Rand[a,b,*,1], Drew Fudenberg[c,1], Anna Dreber[d]

[a] Department of Psychology, Yale University, Box 208205, New Haven, CT 06520-8205, United States
[b] Department of Economics, Yale University, Box 208205, New Haven, CT 06520-8205, United States
[c] Department of Economics, Harvard University, United States
[d] Department of Economics, Stockholm School of Economics, Sweden

ARTICLE INFO

Article history:
Received 24 June 2014
Received in revised form 6 May 2015
Accepted 19 May 2015
Available online 28 May 2015

JEL classification:
C7
C9
D00

Keywords:
Cooperation
Prisoner's dilemma
Repeated games
Intentions

ABSTRACT

We examine cooperation in repeated interactions where intended actions are implemented with noise but intentions are perfectly observable. Observable intentions lead to more cooperation compared to control games where intentions are unobserved, allowing subjects to reach similar cooperation levels as in games without noise. Most subjects condition exclusively on intentions, and use simpler, lower-memory strategies compared to games where intentions are unobservable. When the returns to cooperation are high, some subjects are tolerant, using good outcomes to forgive attempted defections; when the returns to cooperation are low, some subjects are punitive, using bad outcomes to punish accidental defections.

1. Introduction

This paper studies cooperation in infinitely repeated games where the intended actions are implemented with error, so that the actions played are only a noisy or implicit signal of what was intended. The possibility of error is pervasive in social interactions, and many if not most of these interactions do not have a fixed and known termination date. The resulting imperfect public monitoring has received a large amount of attention in the theoretical literature on infinitely repeated games (e.g., Green and Porter, 1984; Radner et al., 1986; Abreu et al., 1990; Fudenberg et al., 1994), but only a handful of experimental studies have explored infinitely repeated games with errors (e.g. Aoyagi and Frechette, 2009; Bigoni et al., 2012; Fudenberg et al., 2012; Aoyagi et al., 2013).[2]

Our setup differs from that of these past studies in that we consider the effect of players directly observing the intended actions of their opponents, in addition to the realized ones. This sort of information is available in some real-world settings,

* Corresponding author at: Department of Psychology, Yale University, Box 208205, New Haven, CT 06520-8205, United States. Tel.: +1 2034324500.
E-mail address: david.rand@yale.edu (D.G. Rand).

[1] Rand and Fudenberg are co-lead authors.

[2] Van Lange et al. (2002) studied play in a repeated continuous-choice (rather than binary) PD with errors where subjects were matched against computer partners playing either tit-for-tat or tit-for-two-tats (but were told that the partners were actual people).

http://dx.doi.org/10.1016/j.jebo.2015.05.013
0167-2681/

for example compensation for hedge fund managers where both the positions taken and the actual outcomes are observable and thus explicit, or in a homicide when it is clear that the accused shot the victim but extenuating circumstances may exist – here the legal system pays attention to both intentions and outcomes, differentiating between manslaughter and various levels of murder.

From a theoretical standpoint, the impact of explicitly observing intentions is clear: the highest equilibrium payoff can be obtained with strategies that completely ignore the realized outcomes and condition only on intended play, and moreover this best equilibrium is the same as when actions are implemented without error. Note that this is very different from the situation in one-shot games, where maximizing money payoffs would lead subjects to ignore intentions entirely. Nonetheless, a substantial proportion of subjects in one-shot games do condition on intentions in addition to outcomes when both pieces of information are available.[3] One possible explanation for this apparent "preference for reciprocity" is that it reflects a heuristic that fosters cooperation in repeated interactions. If so, we might expect to see even more reliance on intentions in settings where conditioning on intentions leads to a cooperative equilibrium even in the absence of a preference for reciprocity. At an empirical level, the question of how extensively people condition their play on intentions in infinitely repeated games remains open, as does the extent to which they also condition on outcomes, and the effect of all this on the level of cooperation.

To begin to understand these issues, we study the experimental play of the repeated prisoner's dilemma when intended actions are implemented with error. Our main goals are to understand when and in what ways subjects use data on intentions and outcomes, and how cooperation when intentions are revealed compares to either a setting with error when intentions are not observed, or one in which error is not present (so the actions themselves reveal the intentions). Our experiment presents evidence from a set of infinitely repeated prisoner's dilemma games with a continuation probability of 7/8 and an error rate of 1/8. In our main treatments, intentions are explicit; as controls, we also consider the same games but where only actions are observable (thus leaving intentions implicit), as well as the same games without exogenously imposed error (where the observed action corresponds to the intended one). We explore two different payoff specifications for the stage game actions "Cooperate" ("C") and "Defect" ("D") (neutral language was used in the experiment itself). In the "high benefit" treatment, the benefit that playing "C" gives to the other player is high enough that there is a cooperative equilibrium in the game with errors whether or not intentions are observed. In the "low benefit" treatment, the benefit that C gives is low enough that the only equilibrium with errors and unobserved intentions is for both players to always defect, although cooperation remains an equilibrium outcome when intentions are observed.

1.1. Summary of results

We use two different methods to analyze the data: a structural estimation of the distribution of strategies using the "structural frequency estimation method" (SFEM) of Dal Bó and Frechette (2011), and a descriptive analysis that relates play in a given period of a supergame to the opponent's intention and action in the period before (which implicitly assumes subjects use strategies that mostly depend on that information). Both methods show that most subjects condition almost exclusively on intentions and thus play consistently with predictions based on maximizing money payoffs. In our descriptive analysis, the effect of opponent's intention is dramatically larger than that of the actual outcome. Similarly, in the strategy estimation, more than two thirds of subjects use strategies that do not condition on outcomes.

To the extent that subjects do condition on outcomes, interestingly, they do so in different ways depending on the payoff specification. In the treatment where there is less of an incentive to cooperate, some people (about 15%) are punitive in treating both accidental cooperation (partner meant to play D but played C) and accidental defection (partner meant to play C but played D) as defection; only when the partner both intended to play C and actually did so was this treated as cooperation. This behavior is not observed in the treatment with high returns to cooperation, where instead some people (about 19%) are tolerant in that they only retaliate against intentional defections – these subjects forgive both accidental defection (partner meant to play C but played D) as well as accidental cooperation (partner meant to played D but played C). Thus the "punitive" subjects in the low-benefit treatment use realized outcomes to punish cooperators that defect by accident, while in the high-benefit treatment "tolerant" subjects use the realized outcomes to forgive defectors that accidentally cooperated.

By conditioning largely on intentions, subjects are able to achieve high levels of cooperation in both treatments. Compared with the controls in which intentions are implicit, explicitly revealing intentions lead to significantly more cooperation. Interestingly, this increase in cooperation is not associated with more leniency (where the subject overlooks the partner's first defection) but instead with an increase in simple non-lenient strategies that conditioned on at most the previous period. This suggests that many of the longer memory lenient strategies seen in the noisy repeated games of Fudenberg et al. (2012) were the result of subjects trying to infer the intentions of their opponent, either because doing so leads to higher monetary payoffs or because preferences depend on the intentions of others.

[3] Past work on intentions in one-shot games is discussed in Section 2. Also, see the Bereby-Meyer and Roth (2006) and Kunreuther et al. (2009) studies of intentions in the finitely-repeated prisoner's dilemma.