Journal of Theoretical Biology 395 (2016) 194–203

Contents lists available at ScienceDirect

Journal of Theoretical Biology

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

Assortativity evolving from social dilemmas

Heinrich H. Nax [a,*], Alexandros Rigos [b]

[a] Department of Social Sciences, ETH Zürich, Clausiusstrasse 37-C3, 8092 Zurich, Switzerland
[b] Department of Economics, University of Leicester, Leicester LE1 7RH, UK

HIGHLIGHTS

- Assortative matching is studied in various social dilemmas.
- Assortativity is endogenized via democratic consensus.
- The matching process co-evolves with cooperation.
- The long-run levels of cooperation are evaluated.

ARTICLE INFO

Article history:
Received 7 August 2015
Received in revised form
13 December 2015
Accepted 20 January 2016
Available online 4 February 2016

Keywords:
Cooperation
(co-)evolution
Assortativity
Democratic consensus

ABSTRACT

Assortative mechanisms can overcome tragedies of the commons that otherwise result in dilemma situations. Assortativity criteria include various forms of kin selection, greenbeard genes, and reciprocal behaviors, usually presuming an exogenously fixed matching mechanism. Here, we endogenize the matching process with the aim of investigating how assortativity itself, jointly with cooperation, is driven by evolution. Our main finding is that full-or-null assortativities turn out to be long-run stable in most cases, independent of the relative speeds of both processes. The exact incentive structure of the underlying social dilemma matters crucially. The resulting social loss is evaluated for general classes of dilemma games, thus quantifying to what extent the tragedy of the commons may be endogenously overcome.

1. Introduction

What happens when a population would collectively benefit from cooperative behavior by all its individuals, while each individual has a private incentive to defect? In some such 'social dilemma' situations, collective action (Olson, 1965) may fail and the tragedy of the commons (Hardin, 1968) may result. However, many mechanisms in nature exist through which cooperative behaviors evolve (see Sachs et al., 2004; West et al., 2007, 2011, for reviews). Hence, the 'puzzle of cooperation' (Darwin, 1871) is that nature, involving humans and animals alike, provides us with many examples of social dilemma situations that are successfully resolved by suitable mechanisms, but also with many other examples that result in the tragedy of the commons.

Perhaps the best methodology to study the evolution of cooperation is provided by game theory (von Neumann and Morgenstern, 1944; Nash, 1951). Without suitable mechanisms, game theory predicts non-cooperative behavior in social dilemmas. The game-theoretic literature has addressed this issue at length (beginning with Hamilton, 1963, 1964a,b; Axelrod, 1984). It was shown that cooperation in social dilemma situations is not favored if interactions in the population are well-mixed/random (Nash, 1950; Lehmann and Keller, 2006; Young, 2011).

The class of mechanisms that we study in this paper function by assorting cooperators. The first accounts of assortative mechanisms date back to Wright (1921, 1922, 1965). Indeed, such mechanisms can lead to cooperative behavior in social dilemma situations; well-known examples include kin selection (Hamilton, 1964a,b; Domingue et al., 2014) via limited dispersal/locality ('spatial interactions'; Nowak and May, 1992; Eshel et al., 1998; Skyrms, 2004; Hauert, 2006; Abdellaoui et al., 2014), greenbeard genes (Dawkins, 1976; Frank, 2010; Jansen and Baalen, 2006; Sinervo et al., 2006; Brown and Fuckling, 2008; Fletcher and Doebeli, 2009, 2010; Gardner and West, 2010), preferences ('homophily'; Alger and Weibull, 2012, 2013; Xie et al., 2015), or are based on behavior ('reciprocal/meritocratic matching'; Clutton-Brock, 2010; Gunnthorsdottir et al., 2010; Rabanal and Rabanal, 2014; Nax et al., 2014, 2015). Importantly, assortment based on behavior is key for (but not restricted to) sustaining cooperation in humans as both theoretical models (Biernaskie

* Corresponding author. Tel.: +41 44 632 05 94; fax: +41 44 632 17 67.
E-mail addresses: hnax@ethz.ch (H.H. Nax), ar374@le.ac.uk (A. Rigos).

http://dx.doi.org/10.1016/j.jtbi.2016.01.032
0022-5193/

et al., 2011) and experiments (Wang et al., 2012) show. In this study we focus on this class of behavior-assortative mechanisms.

Under sufficiently assortative mechanisms, high levels of cooperation are predicted (e.g. Hamilton and Taborsky, 2005a,b; Bergstrom, 2003; Jensen and Rigos, 2014; Nax et al., 2014). It is unlikely, however, that assortativity fell from the sky. More likely, it evolved driven by evolutionary dynamics within the population and across populations. In this paper, we contribute to the assortativity literature by providing a model to endogenize the evolution of assortativity, in particular of behavior.

In our model, assortativity evolves by 'democratic consensus', a standard mechanism to reach consensus in humans. Democratic consensus therefore is a natural candidate for studying the evolution of behavior assortativity, which is particularly relevant for human interactions (Biernaskie et al., 2011). Known as range, average, cardinal, utility or score voting in voting theory, such decision-making rules are used by numerous (proto-)democratic human collectives (Staveley, 1972). Voting by clapping/shouting, financial lobbying, and other mechanisms resembling a tug-of-war-like competition in two opposite directions are examples. The basic feature of democratic consensus in our model is that the underlying mechanism of our interaction gets more or less assortative depending on which direction yields greater surplus. Democratic consensus is also similar to biological auctions, which are aggregation rules used by many animal species (Couzin et al., 2011; Chatterjee et al., 2012) such as bees selecting hive-locations (Seeley and Visscher, 2004) or ants choosing nest sites (Franks et al., 2002). To the best of our knowledge there exists no comparable prior study of evolving assortativity based on democratic consensus dynamics. In biology, other models have been proposed based on different factors such as invasion by mutants (Dieckmann and Doebeli, 1999; Jiang et al., 2013; Dyson-Hudson and Smith, 1978; Bearhop, 2005). Related is also Newton (2014) who studies evolving assortativity in the indirect evolutionary models by Alger and Weibull (2012, 2013, 2014, 2015). Other ways of endogenizing the matching rule such as dynamical networks may lead to different results, and these are avenues for further research we shall sketch in our concluding discussion.

In terms of underlying games, we focus on a class of symmetric two-player social dilemmas that nests the standard prisoners' dilemma (PD) (Rapoport and Chammah, 1965) but also includes other games. All agents are of the same kind, one whose strategy choices are driven by his own material self-interest alone. All social dilemmas we consider, not just the PD, are important situations that often occur in reality with potential detrimental consequences to cooperation.

The PD is the best-known example of social dilemmas, that is, of situations with the common characteristic that individuals have an incentive to defect when facing cooperators. The evolution of cooperation amongst humans and animals in social dilemma situations has received enormous attention, and the PD in particular has been studied widely in this context beginning with Trivers (1971) and Maynard Smith and Price (1973); Maynard Smith (1987) (see also Axelrod and Hamilton, 1981). Beyond the PD, there are related, less well-known social dilemmas of comparable practical importance. All our social dilemmas share the public goods character, but games differ with respect to which outcomes (i) are Nash equilibria and (ii) maximize total payoffs.

Our social dilemma situations include the prisoners' dilemma, the snowdrift game (also known as the hawk-dove game, the game of chicken, or the volunteer's dilemma, Maynard Smith and Price, 1973; Doebeli and Hauert, 2005; Diekmann, 1985; Myatt and Wallace, 2008; Raihani and Bshary, 2011), the missing hero dilemma (Schelling, 1971) and the underprovision dilemma. As a byproduct of our operationalization, we introduce the 'underprovision dilemma', a variant of the snowdrift game, which to the best of our knowledge has not previously been considered but certainly also represents an important class of games deserving investigation.

Our dynamical analyses rely on standard evolutionary replicator (Taylor and Jonker, 1978; Taylor, 1979). In the standard mathematical formulation of such a dynamic (e.g. Eshel, 1983; Helbing, 1992; Weibull, 1995; Eshel et al., 1997), we would assume a well-mixed population, that is, pairs would be drawn uniformly at random from the population. Here, we shall focus on action-assortative matching instead, using recently introduced methods (Bergstrom, 2003; Jensen and Rigos, 2014). In our dilemma games, such a rule is 'meritocratic' as it matches cooperators (defectors) with other cooperators (defectors). Assortativity itself evolves by democratic consensus. In the PD game, for example, cooperators prefer more assortativity in order to be matched less often with defectors, while defectors prefer less assortativity for the opposite reason. In which direction this struggle evolves depends on how many people stand on either side, and by how much they benefit from either change.

Our analysis proceeds in three steps. First, we study the stability of equilibria given an exogenous level of assortativity. Second, we endogenize the evolution of assortativity and investigate the stability of regimes under our voting dynamic. Finally, we evaluate which outcome is more stable in the long run.

2. The model

2.1. Social dilemmas

We start by laying out the general setup. Here, we have an infinite population taken to be the closed interval $[0, 1]$ that can follow one of the two strategies, either 'cooperate' (C) or 'defect' (D). (Alternative labels could be 'contribute' and 'free-ride'.) Denote by x the proportion of individuals playing C. Individuals in the population follow one of the two strategies, get matched to one other individual in the population, and then carry out their strategy in their pair. The exact process by which they get selected in pairs will be discussed in the next section.

Social dilemma: A social dilemma game in our setting is represented by a matrix of the form shown in Table 1.

Hence a social dilemma is defined by $G=(r, a)$. To ensure that C–C is not an equilibrium under random matching, we impose $0 < r < 1$ for all G, which defines the common 'public goods character': defection is always an individual best response against cooperation, but cooperation always increases the opponent's payoff. Moreover, we restrict $a \in (-1, r)$, so that C–D outcomes are associated with either higher or lower total payoffs than C–C, while D–D remains the outcome with lowest total payoffs in all cases. We therefore investigate the following four different types of (well-known) social dilemma games:

Prisoners' dilemma (Rapoport and Chammah, 1965): The PD game is obtained by setting $2r > 1+a$ and $a < 0$. Defection is a strictly dominant strategy, and total payoffs are highest in C–C. The unique Nash equilibrium is D–D.

Snowdrift game: The SD game is obtained by setting $2r < 1+a$ and $a > 0$. Cooperation is a best response against defection, and the

Table 1
The payoff matrix of a social dilemma.

	C	D
C	r, r	$a, 1$
D	$1, a$	0,0