Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 161 (2016) 161–189

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

In and out of equilibrium I: Evolution of strategies in repeated games with discounting

Julián García [a], Matthijs van Veelen [b,*]

[a] *Faculty of Information Technology, Monash University, Clayton, 3800 VIC, Australia*
[b] *Department of Economics, Universiteit van Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands*

Received 7 November 2014; final version received 28 October 2015; accepted 21 November 2015

Available online 2 December 2015

Abstract

In the repeated prisoner's dilemma there is no strategy that is evolutionarily stable, and a profusion of neutrally stable ones. But how stable is neutrally stable? We show that in repeated games with large enough continuation probabilities, where the stage game is characterized by a conflict between individual and collective interests, there is always a neutral mutant that can drift into a population that is playing an equilibrium, and create a selective advantage for a second mutant. The existence of stepping stone paths out of *any* equilibrium determines the dynamics in finite populations playing the repeated prisoner's dilemma.

JEL classification: C73

Keywords: Repeated games; Evolutionary game theory

"Everything changes, nothing remains the same"
Buddha

1. Introduction

Repeated games typically have many equilibria. But how stable are these equilibria? And are some equilibria perhaps more stable than others? In this paper we use refinements from

* Corresponding author.
E-mail addresses: julian.garcia@monash.edu (J. García), c.m.vanveelen@uva.nl (M. van Veelen).

http://dx.doi.org/10.1016/j.jet.2015.11.007
0022-0531/

evolutionary game theory to determine how stable equilibria are, and to help understand the evolutionary dynamics in populations of individuals that are playing such games.

For the repeated prisoner's dilemma we know that there is no pure strategy that is evolutionarily stable (Selten and Hammerstein, 1984), and it is straightforward to extend that argument to non-trivial repeated games in general, and to mixed strategies with finite support. There are, on the other hand, very many neutrally stable strategies in the repeated prisoner's dilemma (Bendor and Swistak, 1995, 1997, 1998). These NSS'es range from fully defecting to fully cooperative. Neutral stability is a weaker version of evolutionary stability, that does not imply asymptotic stability in the replicator dynamics. That, however, does not rule out that there might be a *set* of NSS'es that is evolutionarily stable.

In order to investigate whether or not there are such evolutionarily stable sets, and, more in general, to determine how stable these NSS'es are, we will use the concept of robustness against indirect invasions (RAII, van Veelen, 2012). RAII is less strict than ESS, because it allows for neutral mutants. It is also more strict than NSS, because it does not allow for neutral mutants that serve as a stepping stone for other mutants that have an actual selective advantage, once the first mutant has gained enough of a foothold in the population, for instance through neutral drift. Robustness against indirect invasions preserves a tight link with the replicator dynamics for infinite populations, as well as with stochastic, finite population dynamics of which the replicator dynamics are the large population limit (see Traulsen et al., 2005, 2006). If a strategy is RAII, then it is an element of an ES set (van Veelen, 2012; Balkenborg and Schlag, 2001), which is, as a set, asymptotically stable in the replicator dynamics (Thomas, 1985). Vice versa, if there is an ES set, then all of its elements are RAII. Moreover, the way robustness against indirect invasions deals with neutral mutants implies that it matches the qualitative equilibrium analysis typically applied to stochastic, finite population dynamics, in which neutral mutants play a pivotal role (Nowak, 2006).

It turns out that for repeated games in which the stage game shows a conflict between individual and collective interests – like the prisoner's dilemma – no equilibrium is RAII, provided that the continuation probability is sufficiently high. In other words, any equilibrium can be upset by an at first harmless mutant, which serves as a stepping stone, or a springboard, for the invasion of a second mutant. Stepping stone paths with decreasing cooperation exist for all equilibria, unless there is no cooperation in equilibrium to begin with. Stepping stone paths with increasing cooperation exist for all equilibria that fall sufficiently short of full cooperation. What "sufficiently short" is, depends on the continuation probability.

Simulations show that not only do stepping stone paths out of any equilibrium exist for the repeated prisoner's dilemma, evolution also finds them. With a mutation procedure that is not biased, and that allows for all finite state automata to be reached as mutants, we find that indirect invasions are indeed the driver of the dynamics. The population finds itself in equilibrium most of the time, with regular transitions from equilibrium to equilibrium that do indeed follow these stepping stone paths, both with rising and with declining levels of cooperation. The implications for the dynamics are further illustrated by comparing the repeated prisoner's dilemma – which has no NSS'es that are RAII – to a repeated coordination game which does. We find that if the population size increases, the number of transitions out of equilibrium in the repeated coordination game quickly goes to zero, while in the repeated prisoner's dilemma the population keeps moving from equilibrium to equilibrium regularly. Whether or not there are equilibria that are RAII therefore makes a huge difference for the evolutionary dynamics in repeated games.

Under very reasonable dynamics, equilibria of the repeated games we tend to look at when we study cooperation in repeated interactions, are therefore relatively *un*stable when we compare