Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 157 (2015) 128–145

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Failure of gradualism under imperfect monitoring

Yves Guéron [1]

Sogang University, Republic of Korea

Received 6 June 2012; final version received 24 July 2014; accepted 14 December 2014

Available online 13 January 2015

Abstract

We consider a two-player Prisoner's Dilemma type game with continuous actions, where players choose how much to contribute to a public project. This game is played infinitely many times and actions are irreversible: players cannot decrease their actions over time. While it is strictly dominant for players not to contribute in the stage game, some strictly positive level of contribution is Pareto optimal. It is known that when players perfectly observe each other's actions, cooperation can be achieved through gradual increases in contribution levels. I show that introducing an arbitrarily small amount of smooth noise in the monitoring makes cooperation impossible and players play the static Nash equilibrium of the stage-game forever.

JEL classification: C72; C73

Keywords: Monotone games; Voluntary contributions; Gradualism; Irreversibility; Imperfect monitoring; Cooperation breakdown

1. Introduction

In many economic settings players often have incentives to free ride and benefit from the contributions of others without having to incur a private cost. As a result, many Pareto optimal outcomes cannot be sustained as equilibria of strategic interactions. This problem can usually be overcome through repeated interactions, in which dynamic incentives can counterbalance incentives to free ride. The folk theorem tells us that, in infinitely repeated games, any indi-

E-mail address: ygueron@sogang.ac.kr.

[1] School of Economics, Sogang University, 35 Baekbeom-ro, Mapo-gu, Seoul 121-742, Republic of Korea.

http://dx.doi.org/10.1016/j.jet.2015.01.006

vidually rational payoff profile can be obtained in equilibrium, provided players are sufficiently patient.[2] Dynamic interactions, however, are not necessarily stationary. For example competing firms could repeatedly interact while facing a declining market demand over time; environmental cooperation might require the installation of costly abatement technology which cannot be reversed; capital may be irreversibly destroyed to alleviate the tragedy of the commons. One may then worry that non-stationarity, and in particular irreversibility, might reduce the scope for cooperation.

The main insight from the literature on irreversibility (see for example Marx and Matthews [13] and Lockwood and Thomas [12]) is that cooperation has to occur gradually over time. This is because with irreversibility threats to return to a non-cooperative outcome are no longer feasible and instead punishments must consist in the withdrawal of future increases in cooperation, which implies gradualism.

In this paper we consider a repeated Prisoner's Dilemma with continuous actions in which actions are irreversible, as in Lockwood and Thomas [12], and introduce imperfect public monitoring in the model. Actions are interpreted as total contributions to a public project and irreversibility implies that past contributions are not refundable. Under perfect monitoring, we know that cooperation is feasible and asymptotically efficient as discount factors tend to one. We show that, under certain smoothness assumptions, the introduction of an arbitrarily small amount of noise in the monitoring technology reduces the set of public equilibrium outcomes to the infinite repetition of the unique and inefficient stage-game Nash equilibrium.[3] We also show that smoothness is necessary for our results by means of a counterexample in which the noise is no longer smooth and in which cooperation is again feasible.

To understand why cooperation breaks down under imperfect monitoring, consider a small deviation from a cooperative path in which a player contributes less than what he is expected to, and then returns to the cooperative outcome. Such a deviation yields a small instantaneous gain and a long-term cost which is the product of two effects: (*i*) the decrease in future continuation values; and (*ii*) the change in the distribution of signals. Under perfect monitoring, (*i*) occurs with probability one and is chosen so as to offset any instantaneous gains from a deviation.[4] When contributions are close to an upper bound then (*i*) is small, while if monitoring is smooth then (*ii*) is of similar order of magnitude than the instantaneous gain. The interaction between those two effects makes it such that eventually the cost from a small deviation cannot offset the instantaneous gain, and cooperation breaks down.

This paper finds itself at the intersection between two strands of literature: the literature on irreversibility in dynamic interactions and the literature on imperfect monitoring in repeated games. Unlike this paper, both strands of literature exhibit a series of positive results.

As mentioned previously, the main insight from the literature on irreversibility is that cooperation rises gradually over time and converges. Marx and Matthews [13] study a game of dynamic voluntary contribution to a public project where past contributions are not refundable and payoffs are linear in cumulative contributions, with a possible extra benefit when cumulative contributions are above a given threshold (the "completion point"). They construct an approx-

[2] For example, in an infinitely repeated prisoner's dilemma, cooperation can be achieved with a simple grim-trigger strategy, which prescribes a return to the stage-game Nash equilibrium if one player deviates from a given cooperative path.

[3] More specifically, we require that a small change in actions has an impact of similar order (at most) on the distribution of public signals. See Assumption 6.

[4] Provided initial contributions are sufficiently low this is always possible.