Journal of Mathematical Economics 63 (2016) 131–138

Contents lists available at ScienceDirect

Journal of Mathematical Economics

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

Repeated two-person zero-sum games with unequal discounting and private monitoring[☆]

Guilherme Carmona[a,1], Luís Carvalho[b,c,*]

[a] University of Surrey, School of Economics, Guildford, GU2 7XH, UK

[b] Instituto Universitário de Lisboa (ISCTE-IUL), Av das Forças Armadas, 1649-026, Lisboa, Portugal

[c] Business Research Unit (BRU-IUL), Av das Forças Armadas, 1649-026, Lisboa, Portugal

ARTICLE INFO

Article history:
Received 18 May 2015
Received in revised form
10 February 2016
Accepted 11 February 2016
Available online 23 February 2016

Keywords:
Repeated games
Two-person zero-sum games

ABSTRACT

We consider discounted repeated two-person zero-sum games with private monitoring. We show that even when players have different and time-varying discount factors, each player's payoff is equal to his stage-game minmax payoff in every sequential equilibrium. Furthermore, we show that: (a) in every history on the equilibrium path, the pair formed by each player's conjecture about his opponent's action must be a Nash equilibrium of the stage game, and (b) the distribution of action profiles in every period is a correlated equilibrium of the stage game. In the particular case of public strategies in public monitoring games, players must play a Nash equilibrium after any public history.

1. Introduction

In a two-person zero-sum game, one player's gain is the other's loss. Therefore, this class of games is regarded as the prototype example of a strictly competitive game. Von Neumann's minimax theorem (see, e.g., Myerson, 1991, Theorem 3.2, p. 123) shows that in all Nash equilibria of such games, every player receives his minmax payoff, i.e., the lowest payoff that he can guarantee to himself.

It is clear that the same conclusion applies to a discounted repeated two-person zero-sum game with perfect monitoring when both players have the same discount factor. However, as Lehrer and Pauzner (1999) have pointed out, when players have different discount factors, the repeated game is no longer a zero-sum one. Thus, one may conjecture that the equilibrium set will expand, in particular, by allowing players to obtain higher payoffs than the minmax one. Intuitively, we could think that the player with the smaller discount factor is willing to bear losses in the future if she is compensated with some gains in the present; and vice versa for the other player. Lehrer and Pauzner (1999) show that this intuition is misleading and that, despite the fact that the repeated game is no longer a zero-sum game, each player receives his stage-game minmax payoff in each subgame perfect equilibrium of the repeated game.

We extend Lehrer and Pauzner's (1999) above result to discounted repeated two-person zero-sum games with private monitoring (as in Mailath and Samuelson, 2006, Chapter 12) and with possibly different and time-varying discount factors for both players. Specifically, for each game in this class, we show that each player's payoff is equal to his stage-game minmax payoff in every sequential equilibrium.

Furthermore, we show that (a) in every history on the equilibrium path, the pair formed by each player's conjecture about his opponent's action must be a Nash equilibrium of the stage game, and (b) the distribution of action profiles in every period is a correlated equilibrium of the stage game. In the particular case of public strategies in public monitoring games, we show that players must play a Nash equilibrium after any public history. An analogous result may or may not hold in the case of private monitoring games as we show by two examples.

Our results differ from those in Lehrer and Pauzner (1999) as follows. First, we weaken their assumption of perfect monitoring to private monitoring. This extension is interesting in light of the results in Lehrer and Yariv (1999), where it is shown that conclusions of Lehrer and Pauzner (1999) fail in the context of repeated games with incomplete information on one side. Specifically, one-sided incomplete information and private

☆ We wish to thank Konrad Podczeck for helpful comments. Any errors, of course, are our own.

* Corresponding author at: Instituto Universitário de Lisboa (ISCTE-IUL), Av das Forças Armadas, 1649-026, Lisboa, Portugal. Tel.: +351 217903024.

E-mail addresses: g.carmona@surrey.ac.uk (G. Carmona), luis.carvalho@iscte.pt (L. Carvalho).

1 Tel.: +44 148368 3475.

http://dx.doi.org/10.1016/j.jmateco.2016.02.005
0304-4068/

monitoring are two ways of extending the perfect monitoring, complete information case by introducing some degree of incomplete information. In contrast to what happens in the one-sided incomplete information case of Lehrer and Yariv (1999), we show that the conclusions of Lehrer and Pauzner (1999) extend to the case of private monitoring.[2]

Second, we weaken the assumption that each player has a constant discount factor. In particular, our setting allows for the case in which there are two distinct discount factors, one player uses the lower one in odd periods and the higher one in even periods, and vice versa for the other player. In such a case, there is no longer a uniformly less patient player as in Lehrer and Pauzner (1999). Nevertheless, as we show, their conclusion for two-player zero-sum games remains valid.

Third, our results apply to repeated n-player zero-sum games provided that the sum of the players' minmax payoffs equals zero. This is unlike Lehrer and Pauzner (1999) who consider only two-player games. While some of Lehrer and Pauzner's (1999) results are hard to extend beyond the two-player case (see, e.g. footnote 2 in Sugaya, 2015), we show that their result regarding two-player zero-sum games does extend.

Our results contrast with recent folk theorems for repeated games with unequal discounting obtained by Chen and Takahashi (2012) (for perfect monitoring games) and Sugaya (2015) (for imperfect public monitoring games). The latter assumes that the set of feasible and individually rational payoffs of the stage game has full dimension, a condition that clearly fails in two-person zero-sum games. Our results show that the condition in Chen and Takahashi (2012), dynamically non-equivalent utilities, also fails in repeated two-person zero-sum games for any profile of (time-invariant and) player-specific discount factors. Many other papers have studied repeated games with unequal discounting, although none has considered the case of zero-sum games with private monitoring. These include Chen (2007) in the case of finitely repeated games, Takahashi (2005), Chen (2008), Salonen and Vartiainen (2008), Fong and Surti (2009), Guéron et al. (2011) and Dasgupta and Ghosh (2013) in the perfect monitoring case, Houba and Wen (2006), Haag and Lagunoff (2007), Acemoglu et al. (2008), Houba and Wen (2008), Houba and Wen (2011), Fainmesser (2012) and Opp (2012) in specific economic applications, and Lehrer and Scarsini (2013) in the case of dynamic cooperative games.

The paper is organized as follows. The setting and main definitions are presented in Section 2. Section 3.1 contains our main result for private monitoring games (Theorem 1), our result for public monitoring games (Theorem 2) and their extension for n-player zero-sum games (Remark 2). The proof of our results is in Section 3.2. Section 3.3 presents an example where a Nash equilibrium of the stage game is played in every period in any sequential equilibrium; Section 3.4 presents an example where this property does not hold.

2. Notation and definitions

The stage game: A two-player, zero-sum *private monitoring game* $G = (N, (A_i, Z_i, u_i^*)_{i\in N}, \rho)$ is defined as follows. The set of players is $N = \{1, 2\}$ and A_i is the finite set of player i's actions. Let $A = \prod_{i\in N} A_i$ be the set of action profiles. The set of mixed actions of player $i \in N$ is denoted by $\Delta(A_i)$, and we let $M(A) = \prod_{i\in N} \Delta(A_i)$ be the set of the mixed action profiles. Each player does not observe the action played by the other player but instead only a private signal. We let Z_i be the finite set of player i's private signals, $Z = \prod_{i\in N} Z_i$ and ρ a mapping assigning a probability distribution over Z to each action profile $a \in A$. Thus, for each $a \in A$, $\rho(\cdot|a) \in \Delta(Z)$ and, for each $z = (z_1, z_2) \in Z$, $\rho(z|a)$ is the probability that z_i is observed by player $i = 1, 2$ when a is played. Player i's (ex-post) payoff is $u_i^* : A_i \times Z_i \to \mathbb{R}$.

Ex-ante stage game payoffs are given by $u_i : A \to \mathbb{R}$ for each $i \in N$, where $u_i(a) = \sum_{z\in Z} \rho(z|a)u_i^*(a_i, z_i)$ for each $a \in A$ and $\sum_{i\in N} u_i(a) = 0$ for every $a \in A$.

The following notation will be useful. Let $\rho(z|\alpha) = \sum_{a\in A} \alpha(a)\rho(z|a)$ for each $\alpha \in \Delta(A)$. Furthermore, let A_{-i} and $\Delta(A_{-i})$ be, respectively, the action set and the mixed action set of player i's opponent. The mixed extension of player i's payoff function is also denoted by u_i and defined by $u_i(\sigma) = \sum_{a\in A} \sigma(a)u_i(a)$ for each $\sigma \in \Delta(A)$. Denote the *minmax payoff* for player i by $v_i = \min_{\sigma_{-i}\in\Delta(A_{-i})} \max_{a_i\in A_i} u_i(a_i, \sigma_{-i})$. We will consider, without loss of generality, the normalized game in which both players' minmax payoff is zero, i.e.

$$v_1 = v_2 = 0.$$

We say that G is a *public monitoring game* if $Z_1 = Z_2$ and $\rho(z|a) = 0$ for all $a \in A$ and $z = (z_1, z_2) \in Z$ with $z_1 \neq z_2$.

The repeated game: The infinitely repeated game G^∞ consists of an infinite sequence of repetitions of G. At any stage of the game, starting at date 1, each player takes an action which may depend on his previous private signals and his previous actions. He then receives a private signal according to $\rho \cdot i$.

For any $t \in \mathbb{N}$, the set of all t-stage histories is denoted by $H_t = (A \times Z)^t$, with the initial history containing only the empty set, $H_0 = \{\emptyset\}$. The set of all histories is defined by $H = \bigcup_{t\in\mathbb{N}_0} H_t$. For each $i \in N$, the set of player i's private histories is $H_i = \bigcup_{t\in\mathbb{N}_0}(A_i \times Z_i)^t$ and the set of player i's t-stage private histories is $H_{i,t} = (A_i \times Z_i)^t$. When G is a public monitoring game, a *t-stage public history* is $h_t = \left((a^1, z^1), \ldots, (a^t, z^t)\right) \in H_t$ such that $z_1^\tau = z_2^\tau$ for all $\tau = 1, \ldots, t$. Let H_t^p denote the set of all t-stage public histories.

A behavior strategy for player $i \in N$ is a function $f_i : H_i \to \Delta(A_i)$ mapping private histories into mixed actions, with $f_{i,a_i}(h_{i,t})$ denoting the probability of a_i being chosen by player i after history $h_{i,t} \in H_i$ has occurred. Given a strategy f_i and a history $h_{i,t} \in H_i$, the strategy induced by f_i at $h_{i,t}$ is denoted by $f_i|h_{i,t}$. Let F_i be the set of all strategies of player i, and $F = \Pi_{i\in N}F_i$ the set of all strategies profiles.

We say that a strategy profile $f \in F$ is *action-free* if $f_i(h_{i,t}) = f_i(\bar{h}_{i,t})$ for each $i \in N$, $t \in \mathbb{N}_0$, $h_{i,t} = \left((a_i^1, z_i^1)\ldots, (a_i^t, z_i^t)\right) \in H_{i,t}$ and $\bar{h}_{i,t} = \left((\bar{a}_i^1, \bar{z}_i^1)\ldots, (\bar{a}_i^t, \bar{z}_i^t)\right) \in H_{i,t}$ with $(z_i^1, \ldots, z_i^t) = (\bar{z}_i^1, \ldots, \bar{z}_i^t)$. Note that, in the case where G is a public monitoring game, action-free strategies are called *public strategies*.

A strategy profile f induces a probability distribution π_t^f over H_t for each $t \in \mathbb{N}_0$ and a probability distribution P^f over the infinite sequences of signal and action profiles, the set of such sequences being denoted by $Y = (A \times Z)^\infty$. Expectations with respect to P^f will be denoted by E^f. A strategy profile f also induces a probability distribution over the set of actions in each period $t \in \mathbb{N}$, which will play a role in our results. For each $t \in \mathbb{N}$, such distribution is denoted by $\alpha^t(f) \in \Delta(A)$ and is defined by

$$\alpha_a^t(f) = \sum_{h_{t-1}\in H_{t-1}} \sum_{z\in Z} \pi_t^f(h_{t-1} \cdot (a, z))$$

for all $a \in A$, where $h_{t-1} \cdot (a, z) = ((a^1, z^1), (a^2, z^2), \ldots, (a^{t-1}, z^{t-1}), (a, z))$ for each $h_{t-1} = \left((a^1, z^1), (a^2, z^2), \ldots, (a^{t-1}, z^{t-1})\right) \in H_{t-1}$.

We assume that each player $i \in N$ discounts the future with discount factors $\delta_t^i \in (0, 1)$ for $t \in \mathbb{N}$, satisfying $\overline{\lim}_t\, \delta_t^i < 1$. The

[2] It would be interesting to consider a general model of repeated zero-sum games having the one-sided incomplete information setting of Lehrer and Yariv (1999) as well as the private monitoring framework considered here as special cases, and to obtain conditions under which the results in Lehrer and Pauzner (1999) do, and do not, extend. We leave this for further research.