Equilibrium Outcomes of Repeated Two-Person, Zero-Sum Games[*]

Guilherme Carmona
Faculdade de Economia, Universidade Nova de Lisboa
Campus de Campolide, 1099-032 Lisboa, Portugal
email: gcarmona@fe.unl.pt
telephone: (351) 21 380 1671
fax: (351) 21 388 6073

January 7, 2003

---

[*] I wish to thank Andy McLennan for very helpful comments. All remaining errors are, of course, mine.

Abstract

We consider repeated two-person, zero-sum games in which the preferences in the repeated game depend on the stage-game preferences, although not necessarily in a time-consistent way. We assume that each player's repeated game payoff function at each period of time is strictly increasing on the stage game payoffs and that the repeated game is itself a zero-sum game in every period. Under these assumptions, we show that an outcome is a subgame perfect outcome if and only if all its components are Nash equilibria of the stage game.

Journal of Economic Literature Classification Numbers: C73.

Keywords: Repeated two-person, zero-sum games, time-inconsistent preferences.

Correspondent: Guilherme Carmona, Faculdade de Economia, Universidade Nova de Lisboa, Campus de Campolide, 1099-032 Lisboa, Portugal. Email: gcarmona@fe.unl.pt; telephone: (351) 21 380 1671; fax: (351) 21 388 6073.

Proposed Running Head: Repeated Two-Person, Zero-Sum Games

1 Introduction

When faced with a dynamic decision problem, individuals often display a desire to commit themselves to a particular plan of future actions. At first sight, a commitment to a plan of future actions may seem puzzling, since by doing so any individual is reducing the alternatives he will have in the future. However, commitment to a future plan of actions may be part of the optimal way to choose today when individuals' preferences change over time, i.e, when individual have time-inconsistent preferences.

Recently, there has been an increasing interest in time-inconsistent preferences, and its consequences for economic theory, and policy. Much of this interest was motivated by the fact that time-inconsistent preferences can change a model's implications for economic policy, as shown by Laibson [5], Jovanovic and Stolyarov [3], and Kocherlakota [4]. Is is the case that time-inconsistent preferences will always change a model's implications for economic policy? We show that in a two-person, zero-sum game time-inconsistent preferences have no effect over the equilibrium outcomes that can arise, and so have no effect over the model's implication for economic policy.

Our main result can be stated as follows: consider a repeated two-person, zero-sum games in which the preferences in the repeated game depend on the stage-game preferences, although not necessarily in a time-consistent way. Thus, players have a (possibly different) payoff function in every period. Assume that each player's repeated game payoff function at each period of time is strictly increasing on the stage game payoffs, and, if the repeated game is itself a zero-sum game in every period (i.e., the sum of the player's period t payoffs in the repeated game equals zero, for all t.) Under these assumptions, we will show that an outcome is a subgame perfect outcome if and only if all its components are Nash equilibria of the stage game.

Two-person, zero-sum (normal form) games are regarded as descriptions of highly competitive situations. This confirmed by the fact that in Nash equilibrium any player may assume that her opponent is choosing his action to minimize her payoff. Our result shows that this competitiveness still holds when the game is repeated countably many times.

It should be noted that the conclusion of our main result belongs to the oral tradition of game theory, at least when the repeated game payoffs are given by the discounted sum of stage game payoffs. A contribution of our work is to provide a simple proof of that result, and to show that it holds under quite general assumptions.

Given the recent interest on the economic effects of time-inconsistent preferences, it is interesting to know what game-theoretic results change by assuming time-inconsistent preferences. This question seems natural to us since, as Peleg and Yaari [6] and Goldman [2] pointed out, the appropriate way of modelling time-inconsistency in preferences is through the concept of subgame perfect equilibrium of a game between an agent and his future selves.[1] We see our work as a contribution towards answering this general question.

2 Notation and definitions

A *two-person, zero-sum game* G is defined by

$$G = (A_1, A_2, u_1, u_2)\,,$$

where for all $i = 1, 2$: (1) A_i is a finite set of player i' actions, and (2) $u_i : A \to \mathbb{R}$, where $A = A_1 \times A_2$, is player i' payoff function; the player's payoff functions satisfy

$$u_1(a) + u_2(a) = 0,$$

for all $a \in A$. Let $S_i = \Delta(A_i)$, $S = S_1 \times S_2$, and $u_i : S \to \mathbb{R}$ be the usual extension to mixed strategies.

Let, for $i = 1, 2$, $v_i = \min_{s_{-i}} \max_{s_i} u_i(s_i, s_{-i})$, and $NE = \{s \in S : \text{for all } i = 1, 2, u_i(s) \geq u_i(\tilde{s}_i, s_{-i}),\ \text{for all } \tilde{s}_i \in S_i\}$. The set NE is the set of

---

[1] The concept of time-inconsistent preferences was itself introduced by Strotz [7].

Nash equilibria of G, and v_i is the minmax level for player i.

The *supergame of* G consists of an infinite sequence of repetitions of G taking place in periods $t = 1, 2, 3, \ldots$. At period t the players make simultaneous moves denoted by $s_i^t \in S_i$ and then each player learn his opponent's move.

For $k \geq 1$, a $k-$*stage history* is a $k-$length sequence $h_k = (s^1, \ldots, s^k)$, where, for all $1 \leq t \leq k$, $s^t \in S$; the space of all $k-$stage histories is H_k, i.e., $H_k = S^k$ (the $k-$fold Cartesian product of S.)[2] The notation e stands for the unique 0–stage history — it is a 0–length history that represents the beginning of the supergame. The set of all histories is defined by $H = \bigcup_{n=0}^{\infty} H_n$.

For every $h \in H$, define $h^r \in S$ to be the projection of h onto its r^{th} coordinate. For every $h \in H$ we let $\ell(h)$ denote the *length of* h. For two positive length histories h and $\bar{h}$ in H we define the *concatenation of* h *and* $\bar{h}$, in that order, to be the history $(h \cdot \bar{h})$ of length $\ell(h) + \ell(\bar{h})$: $(h \cdot \bar{h}) = (h^1, h^2, \ldots, h^{\ell(h)}, \bar{h}^1, \bar{h}^2, \ldots, \bar{h}^{\ell(\bar{h})})$. We also make the convention that $e \cdot h = h \cdot e = h$ for every $h \in H$.

It is assumed that at stage k each player knows h_k, that is each player

---

[2] As in Aumann [1], we are assuming that players can observe the mixed strategies chosen. This assumption is not crucial to our work since, as Theorem 1 will show, equilibrium play is independent of the history.

knows the actions that were played in all previous stages. Regarding strategies, players chose behavioral strategies, that is, in each stage k, they choose a function from H_{k-1} to S_i denoted f_k^i, for player $i = 1, 2$. The set of player i's strategies is denoted by F_i, and $F = F_1 \times F_2$ is the joint strategy space. Finally, a strategy vector is $f = \left(\{f_k^i\}_{k=1}^{\infty}\right)_{i=1,2}$.

Given an individual strategy $f_i \in \Sigma_i$ and a history $h \in H$ we denote the *individual strategy induced by f_i at h* by $f_i|h$. This strategy is defined pointwise on H: $(f_i|h)(\bar{h}) = f_i(h \cdot \bar{h})$, for every $\bar{h} \in H$. We will use $(f|h)$ to denote $(f_1|h, \ldots, f_n|h)$ for every $f \in S$ and $h \in H$.

Any strategy $f \in F$ induces an outcome $\pi(f)$ as follows:

$$\pi^1(f) = f(e), \quad \pi^k(f) = f(\pi^1(f), \ldots, \pi^{k-1}(f)), \tag{1}$$

for $k \in \mathbb{N}$. Thus, we have define a function $\pi : F \to S^{\infty}$, where $S^{\infty} = S \times S \times \cdots$.

Let $M \geq 0$ be such that $|u_i(s)| \leq M$, for all $s \in S$, and $i \in N$. Then, any outcome $\pi \in S^{\infty}$ induces two elements in l^{∞}, one for each player, as follows

$$x_i^k(\pi) = u_i(\pi^k), \tag{2}$$

for all $k \in \mathbb{N}$. Thus, we have define a function $x_i : S^{\infty} \to l^{\infty}$, for all $i = 1, 2$.

For $x, y \in l^{\infty}$, $x \geqq y$, means $x_k \geq y_k$, for all $k \in \mathbb{N}$; $x \geq y$ means $x \neq y$ and $x \geqq y$.

Let for each $i = 1, 2$, and $k \in \mathbb{N}$, $U_i^k : l^\infty \to \mathbb{R}$ be given. The payoff for player i, $i = 1, 2$, from his point of view in period $k \in \mathbb{N}$ of a strategy $f \in F$ in the supergame of G is defined to be $U_i^k(x_i \circ \pi(f))$.

A strategy vector $f \in F$ is a *subgame perfect equilibrium* of the supergame of G if $U_i^k(x_i(h \cdot \pi(f|h))) \geq U_i^k(x_i(h \cdot \pi(g_i, f_{-i}|h)))$, for all $i = 1, 2$, $k \in \mathbb{N}$, $h \in H_{k-1}$ and $g_i \in F_i$. Let $E\mathit{\Pi}$ denote the set of subgame perfect equilibrium outcomes.

3 Equilibrium outcomes

In this section we state and prove our main result.

Theorem 1 *Suppose that for all $k \in \mathbb{N}$, and $i = 1, 2$,*

1. $U_1^k(x_1(\pi)) + U_2^k(x_2(\pi)) = 0$, *for all* $\pi \in S^\infty$,
2. U_i^k *is strictly increasing:* $x, y \in l^\infty$ *and* $x \geq y$ *implies* $U_i^k(x) > U_i^k(y)$.

Then, $E\mathit{\Pi} = NE^\infty$ *and* $u_i(\pi^k) = v_i$ *for all* $\pi \in E\mathit{\Pi}$, $i = 1, 2$, *and* $k \in \mathbb{N}$.

Proof. Clearly, we have that $NE^\infty \subseteq E\mathit{\Pi}$, and so it is enough to show that $E\mathit{\Pi} \subseteq NE^\infty$.

Let $\pi \in E\mathit{\Pi}$, $i = 1, 2$, and $k \in \mathbb{N}$. By 2,

$$\begin{aligned} U_i^k(x_i(\pi)) \geq U_i^k(x_i^1(\pi), \ldots, x_i^{k-1}(\pi), \max_{s_i} u_i(s_i, \pi_{-i}^k), v_i, v_i \ldots) \geq \\ \geq U_i^k(x_i^1(\pi), \ldots, x_i^{k-1}(\pi), v_i, v_i, \ldots) := \bar{v}_i^k. \end{aligned} \tag{3}$$

Let α be a Nash equilibrium of G; thus, in particular, $u_i(\alpha) = v_i$. By letting $\tilde{\pi} = (\pi^1, \ldots, \pi^{k-1}, \alpha, \alpha, \ldots)$, we conclude by 1 that $\bar{v}_1^k + \bar{v}_2^k = U_1^k(x_1(\tilde{\pi})) + U_2^k(x_2(\tilde{\pi})) = 0$. Also, by 1, $U_1^k(x_1(\pi)) + U_2^k(x_2(\pi)) = 0$. Hence, $U_k(x_i) = \bar{v}_i^k$.

We therefore conclude that,

$$\begin{aligned} U_k(x_i^1(\pi), \ldots, x_i^{k-1}(\pi), \max_{s_i} u_i(s_i, \pi_{-i}^k), v_i, v_i \ldots) = \\ = U_k(x_i^1(\pi), \ldots, x_i^{k-1}(\pi), v_i, v_i, \ldots), \end{aligned} \tag{4}$$

and so by 2, $\max_{s_i} u_i(s_i, \pi_{-i}^k) = v_i$.

Since $u_i(\pi^k) = x_i^k(\pi) \leq \max_{s_i} u_i(s_i, \pi_{-i}^k) = v_i$, for all $k \in \mathbb{N}$, and $U_i^1(x_i(\pi)) \geq U_i^1(v_i, v_i, \ldots)$, it follows that

$$u_i(\pi^k) = v_i = \max_{s_i} u_i(s_i, \pi_{-i}^k); \tag{5}$$

hence, π^k is a Nash equilibrium. ■

A interesting particular case is when both player use the same function to evaluate repeated game payoffs. In this case, then it is enough that this common function be additive, and strictly increasing in order for the conclusion of Theorem 1 to hold. A particular case of an additive function is a continuous linear functional, but we note that our result hold for functions outside this class.

Corollary 1 *Suppose that for all* $k \in \mathbb{N}$

1. $U_i^k = U_k$, *for* $i = 1, 2$,
2. U_k *is additive:* $U_k(x + y) = U_k(x) + U_k(y)$, *for all* $x, y \in l^\infty$,
3. U_k *is strictly increasing:* $x, y \in l^\infty$ *and* $x \geq y$ *implies* $U_k(x) > U_k(y)$.

Then, $E\Pi = NE^\infty$ *and* $u_i(\pi^k) = v_i$ *for all* $\pi \in E\Pi$, $i \in N$, *and* $k \in \mathbb{N}$.

The following example shows that we cannot dispense with additivity in the case when both players use the same function to evaluate repeated game payoffs. The stage game is the matching pennies:

$1\backslash 2$	H	T
H	$1, -1$	$-1, 1$
T	$-1, 1$	$1, -1$

Table 1: Payoff Function for the Matching Pennies

Assume time-consistency, and let $w = (-1, 1, -1, 1, ...)$.

The preferences are, for $\delta \in (0,1)$ and $M > 0$,

$$U(x) = \begin{cases} (1-\delta)\sum\limits_{k=1}^{\infty}\delta^{k-1}x_k + M & \text{if} \quad x \geqq w \\ (1-\delta)\sum\limits_{k=1}^{\infty}\delta^{k-1}x_k & \text{otherwise.} \end{cases} \tag{6}$$

This preferences are strictly increasing because $\geqq$ is transitive and the discounted sum is strictly increasing. Define π as follows $\pi^1 = (H,H)$, $\pi^2 = (T,H)$, $\pi^3 = (H,H)$, $\pi^4 = (T,H), \ldots$, and define

$$f_i(h) = \begin{cases} \pi_i^k \quad \text{if} \quad h = (\pi^1, ..., \pi^{k-1}) \\ \text{play H with 1/2 probability otherwise.} \end{cases} \tag{7}$$

Then f is a subgame perfect equilibrium and π is a SPE outcome (and it doesn't consist of Nash equilibria of the stage game). This is so because the payoff of the equilibrium path is $M - \frac{(1-\delta)^2}{1-\delta^2}$, and the payoff from deviating is 1. By choosing M big enough, we can deter deviations.

However, we can weaken to requirement of additivity by adding some extra conditions. One extra condition we need is independence from the past: Let for $h \in H$ and $r \in \mathbb{N}$, $\lambda^r h := (h^1, \ldots, h^r)$ and $\mu^r h := (h^r, h^{r+1}, \ldots)$. We say that U_k, $k \in \mathbb{N}$, is *independence of the past* if for all $x, y \in S^\infty$ satisfying $\lambda^{k-1}x = \lambda^{k-1}y$ then $U_k(x) \geq U_k(y)$ if and only if $U_k(\mu^k x) \geq U_k(\mu^k y)$. If, for all $k \in \mathbb{N}$, U_k is independent of the past, then $f \in F$ is a subgame perfect equilibrium if and only if $U_k(x_i \circ \pi(f|h)) \geq U_k(x_i \circ \pi(g_i, f_{-i}|h))$, for all $i = 1, 2$, $k \in \mathbb{N}$, $h \in H_{k-1}$ and $g_i \in F_i$.

We can now state:

Theorem 2 *Suppose that for all $k \in \mathbb{N}$*

1. *$U_i^k = U_k$, for $i = 1, 2$,*
2. *U_k is independent of the past,*
3. *$U_k(x + y) \geq U_k(x) + U_k(y)$, for all $x, y \in l^\infty$,*
4. *$U_k(\alpha, \alpha, \ldots) = \alpha$, for all $\alpha \in \mathbb{R}$,*
5. *U_k is strictly increasing: $x, y \in l^\infty$ and $x \geq y$ implies $U_k(x) > U_k(y)$.*

Then, $E\Pi = NE^\infty$ and $u_i(\pi^k) = v_i$ for all $\pi \in E\Pi$, $i = 1, 2$, and $k \in \mathbb{N}$.

Proof. Clearly, we have that $NE^\infty \subseteq E\Pi$, and so it is enough to show that $E\Pi \subseteq NE^\infty$.

Let $\pi \in E\Pi$, $i = 1, 2$, and $k \in \mathbb{N}$. For $t \geq k$, let $x_i^t = u_i(\pi^t)$, and $x_i = (x_i^k, x_i^{k+1}, \ldots)$. By 4, and 5,

$$U_k(x_i) \geq U_k(\max_{s_i} u_i(s_i, \pi_{-i}^k), v_i, v_i \ldots) \geq U_k(v_i, v_i, \ldots) = v_i. \tag{8}$$

By 3, and 4,

$$U_k(x_1) + U_k(x_2) \leq U_k(x_1 + x_2) = 0. \tag{9}$$

Because $v_1 + v_2 = 0$, it follows that

$$U_k(x_i) = v_i. \tag{10}$$

Thus, by 5, $\max_{s_i} u_i(s_i, \pi^k_{-i}) = v_i$. Since $u_i(\pi^k) = x^k_i \leq \max_{s_i} u_i(s_i, \pi^k_{-i}) = v_i$, then $u_i(\pi^k) = v_i$. Since this equality holds for all $k \in \mathbb{N}$, it follows that $x_i = (v_i, v_i, \ldots)$, and so by 5,

$$u_i(\pi^k) = v_i = \max_{s_i} u_i(s_i, \pi^k_{-i}); \tag{11}$$

hence, π^k is a Nash equilibrium. $\blacksquare$

4 Conclusion

We have shown that, under general conditions, equilibrium outcomes of repeated two-person, zero-sum games have the property that in every period a Nash equilibrium of the stage game is played. This result is interesting for at least two reasons: First, it shows that the strict competitiveness embodied in two-person, zero-sum (normal-form) games extends to the repeated version. Second, it shows that this is true, even if players have time-inconsistent preferences; in particular, in economic situations described by a repeated two-person, zero-sum game, the introduction of time-inconsistent preferences will not change the equilibrium outcomes that can arise, and so have no effect over the model's implication for economic policy.

References

[1] Aumann, R. (1964), “Mixed and Behavioural Strategies in Infinite Extensive Games”, *Advances in Game Theory,* ed. by M. Dresher, L. Shapley and A. Tucker, Princeton University Press.

[2] Goldman, S., (1980), “Consistent Plans”, *Review of Economic Studies*, **47**, 533-537.

[3] Jovanovic, B., and D. Stolyarov, (2000), “Ignorance is Bliss,” Manuscript, New York University.

[4] Kocherlakota, N., (2001), “Looking for Evidence of Time-Inconsistent Preferences in Asset Market Data,” *Federal Reserve Bank of Minneapolis Quarterly Review*, **v. 25, no. 3**, 13-24.

[5] Laibson, D., (1997), “Golden Eggs and Hyperbolic Discounting,” *Quarterly Journal of Economics*, **112**, 443-477.

[6] Peleg, B., and M. Yaari (1973), “On the Existence of a Consistent Course of Action when Tastes are Changing,” *Review of Economic Studies*, **40**, 1-24.

[7] Strotz, R. (1955), “Myopia and Inconsistency in Dynamic Utility Maximization,” *Review of Economic Studies*, **23**, 165-180.