Discussion Paper No. 716

STRONG, PERFECT EQUILIBRIUM PAYOFFS OF INERTIA SUPERGAMES

by

Subir K. Chakrabarti*

January 1987

*Northwestern University Department of Managerial Economics and Decision Sciences, J.L. Kellogg Graduate School of Management, Evanston, Illinois 60201.

I remain grateful to Profs. Tatsuro Ichiishi and James Friedman for discussions on an earlier version. I thank Prof. Myerson for insisting on complete characterization of the equilibrium payoffs, and Prof. Ehud Kalai for some comments on an earlier draft.

STRONG, PERFECT EQUILIBRIUM PAYOFFS OF INERTIA SUPERGAMES

by

Subir K. Chakrabarti

Abstract

The set of payoffs of the strong perfect equilibrium points of the inertia supergames is examined. An inertia supergame is one in which changing strategies over time is not costless. A complete characterization of the set of payoffs of the equilibrium points mentioned above is given. It is shown that these payoffs are the same as the payoffs of the one-shot game which are called *sustainable* payoffs. Further, if one allows for the right correlating device, one finds that for games which have sustainable payoffs, the *correlated* α-core is the closure of these payoffs.

STRONG, PERFECT EQUILIBRIUM PAYOFFS OF INERTIA SUPERGAMES

by

Subir K. Chabrabarti

1. Introduction

In the literature on supergames, all the analysis has usually been carried out under assumptions which ignore any cost that may be incurred when strategies are changed over time. Under the assumption of zero cost of change, the two major lines of inquiry have either followed the "folk theorem" or the "Aumann proposition." These two major results provide characterizations of the payoffs of equilibria of the supergames in terms of the payoffs of the one shot games. The folk theorem establishes the connection between the individually rational payoffs of the one shot game, and the Nash equilibria of the supergame. The Aumann proposition establishes the connection between a stronger form of equilibrium concept, viz., the strong equilibrium of the supergame and the β-core of the one shot game.

The two groups of results are intriguing because each gives us a sense in which payoffs of one shot games could be implemented if there is sufficient repetition of these games. However, both strands in the literature deal with supergames arising out of repeating one shot games over time with no cost involved in changing strategies. In many situations changing strategies from one period to the next may not be costless as in the case of quantity adjustments by firms, which require changes in capacity and operating schedules, or price adjustments which require advertising the price changes. When such costs are incurred due to changing strategies over time, reporting a oneshot game gives rise to a supergame which has a slightly different payoff structure than the usual supergames. The penalty for changing strategies over time that is introduced with the assumption of change costs leads one to conjecture that a larger set of payoffs than the individually rational payoffs can be Nash equilibrium payoffs of these "inertia" supergames.

However, a more interesting conjecture is whether one can develop a connection between the α-core of the one shot game and strong equilibrium payoffs of the inertia supergame. This conjecture is of interest because the α-core is the weakest core concept that can be defined on the one shot games and has nice existence properties (see Scarf (1971)). The β-core is a far stronger concept than the α-core and no general existence result is known, and hence, a strengthenting of the Aumann proposition would be welcome.

Further, it would be of interest to characterize the set of strong equilibrium payoffs of the inertia supergames in terms of the payoffs of the one shot game and compare this to the existing core-like solution concepts like the α-core and the β-core.

In fact, we actually characterize the set of payoffs of the strong, perfect equilibrium points of the inertia supergames and show that these are exactly the same as the set of payoffs of the one shot game which is called "sustainable." These sustainable payoffs are always contained in the α-core but contains the β-core, the desired payoffs and the reactive payoffs. The desired payoffs are defined in Rubinstein (1980), in which it is shown that the desired payoffs coincides with the strong, perfect equilibrium payoffs of the supergames in which changing strategies are costless. The reactive payoffs are defined in Chakrabarti (1986) in which it is shown that every reactive payoff is the payoff of a strong, perfect equilibrium of the inertia supergame.

The correlated α-core is a concept that one may conceivably think should be close to the sustainable payoffs. However, what one finds is that the sustainable payoff is a slightly stronger core concept than the correlated α-core.

2. Model

Let $G = \{X^j, U^j\}_{j \in N}$ be a game in the normal form where

N is the set of players, and is finite.

X^j is the strategy set of player $j \in N$.

$X^S := \bigotimes_{j \in S} X^j$ is the cartesian product of X^j for $j \in S$.

Λ is the set of all nonempty subsets of N, or the set of all coalitions.

U^j: $X \to \mathbb{R}$ is the payoff function of player $j \in N$, where $X := \bigotimes_{j \in N} X^j$.

We will always retain the following two assumptions about the game G.

Assumption 1: The strategy set X^j is a nonempty, compact subset of a complete metric space for all $j \in N$.

Assumption 2: The payoff functions U^j are continuous in the product topology on X.

Assumption 3: The player set N is finite.

We now specify that G is played an infinite number of times starting at period 1. The strategies chosen in period t by the players specify the outcome in that period. If $x \in X$ is the outcome in period t than $U^j(x)$ is player j's payoff in period t. But if player j's strategy choice in period t is x^j_t and in period $t + 1$ is given by $x^j_{t+1} \neq x^j_t$, then the player incurs a change cost given by $C_j(x^j_t, x^j_{t+1})$ and the payoff to player j in period $t + 1$ is given by

$$U^j(x_{t+1}) - C_j(x^j_t, x^j_{t+1})$$

The cost function C_j: $X^j \otimes X^j \to \mathbb{R}$ takes only nonnegative values in its range and is bounded above. We will make the following assumption about these cost functions:

Assumption 4: For $S \in \Lambda$ define

$$\phi_j(S,y) = \sup\{[U^j(x^S,y^{N\setminus S}) - U^j(y)]/x^S \in X^S\}$$

We assume that

$$\phi_j(S,y) \leqslant C_j(y^j,x^j) \tag{2.1}$$

for all $j \in S$, $x^j \in X^j$, $y \in X$ and $S \in \Lambda$.

We now define the inertia supergame G^∞. A strategy for a player j in G^∞ is a sequence of functions $\{h_j(t)\}_{t=1}^{\infty}$, where $h_j(1) \in X^j$ is the strategy chosen in period 1 by player j, and

$$h_j(t)\colon X^{t-1} \to X^j, \quad t \geqslant 2$$

is the function defined on X^{t-1} where X^{t-1} is the cartesian product of X taken t - 1 times. Therefore a supergame strategy of a player is a complete plan of choices of strategies in each period, where the choices may be conditioned on past history.

We will denote the set of supergame strategies of player j by H_j. Then the set H of all n-tuple of strategies of the supergame is

$$H := H_1 \otimes H_2 \otimes \ldots \otimes H_n.$$

Note than an element of H is an n-tuple of sequences of functions $(\{h_1(t)\}_{t=1}^{\infty},\ldots,\{h_n(t)\}_{t=1}^{\infty})$. An <u>outcome path</u> of the supergame G^{∞} is a sequence $\underline{x} = (x_t)_{t=1}^{\infty}$ where $x_t \in X$ for all $t \geq 1$ and is the outcome in period t. With each outcome path is associated a sequence of vectors of payoffs; the payoff to player j in period t being defined to be $U^j(x_{t-1},x_t)$ for $t \geq 2$.

The payoff $U^j(x_{t-1},x_t)$ is given by

$$U^j(x_{t-1},x_t) := \begin{cases} U^j(x_t) \text{ if } x_{t-1}^j = x_t^j \\ \\ U^j(x_t) - C_j(x_{t-1}^j,x_t^j), \text{ if } t \geq 2 \text{ and } x_{t-1}^j \neq x_t^j. \end{cases}$$

for the outcome path $\underline{x} \in X^{\infty}$.

The payoff to player $j \in N$ in the inertia supergame G^{∞} when the outcome path $\underline{x} \in X^{\infty}$ is realized is then given by

$$(2.2) \qquad U_{\infty}^j(\underline{x}) := \liminf_{\nu\to\infty} \frac{1}{\nu} \sum_{t=1}^{\nu} U^j(x_{t-1},x_t).$$

Then $G^{\infty} := \{X^j,H_j,C_j,U_{\infty}^j\}_{j\in N}$ is an <u>inertia</u> supergame when Assumption 4 holds.

<u>Remark:</u> We should note at this point that given a strategy $h \in H$, there is a single outcome path $\underline{x} \in X^{\infty}$ that will be realized if all the players play according to plan h. Different outcome paths will result if one or more of the players deviates from h at some point.

<u>Notation:</u> For a strategy $h \in H$, we will write the outcome in period t as $h(t) \in X$, which, of course, depends on the past history. We will denote $h_j(t) \in X^j$ as player j's strategy choice.

3. Some Definitions

Before we prove the two main results, we need some definitions.

Definition 3.1: $v^* \in C_0\ [U(x)]$ is a sustainable payoff (or an s payoff) if for all $S \in \Lambda$ and $x_\ell^S \in X^S$ there exists an $x_\ell^{N\setminus S} \in X^{N\setminus S}$ such that for any finite index set L, and any $p \in \Delta(L)$, there exists a $j \in S$ such that

$$\sum_{\ell \in L} p_\ell U^j(x_\ell^S, x_\ell^{N\setminus S}) \leq v_j^*.$$

Note: Here $\Delta(L) = \{p: L \to (0,1]/ \sum_{\ell \in L} p_\ell = 1\}$, where L is some finite set.

Hence, v^* is a sustainable payoff if it is feasible when the grand coalition correlates strategies, and for any choice by a deviating coalition there is an appropriate response by the complement such that in whatever manner the deviating coalition may correlate deviations, someone in the coalition does not gain. We will later show that this is a slightly stronger notion than the correlated α-core.

Definition 3.2: A supergame strategy n-tuple $h^* \in H$ is a strong equilibrium of the inertia supergame G^∞ if for all $S \in \Lambda$ and $h^S \in H^S$, there exists a $j \in S$ such that

$$U_\infty^j(h^S,(h^*)^{N\setminus S}) \leq U_\infty^j(h^*)$$

where $H^S := \bigotimes_{j \in S} H_j$ and $U_\infty^j(h^S,(h^*)^{N\setminus S})$ is well-defined since the strategy n-tuple $(h^S,(h^*)^{N\setminus S})$ will give rise to an outcome path in X^∞. Similarly, $U_\infty^j(h^*)$ is well defined since the strategy n-tuple $h^* \in H$ gives rise to a unique outcome path in X^∞.

Therefore, a strong equilibrium is a Nash equilibrium that defines the condition that even deviations by coalitions are deterred. In other words, a strong equilibrium is a coalition proof Nash equilibrium.

Definition 3.3: A supergame strategy n-tuple $h^* \in H$ is a strong, perfect equilibrium of the inertia sueprgame G^∞ if for any history $(x_1, x_2, \ldots, x_T)$ up to period T, the strategy n-tuple $h^*|_{(x_1,x_2,\ldots,x_T)}$ that h^* induces on the subgame $G^\infty|_{(x_1,x_2,\ldots,x_T)}$ is a strong equilibrium of $G^\infty|_{(x_1,\ldots,x_T)}$.

Therefore, by a strong perfect equilibrium of the inertia supergame G^∞ we will mean a strong equilibrium which is subgame perfect.

We have a last definition:

Definition 3.4: A strategy n-tuple $h \in H$ is summable if the limit of the average of payoffs exists for all $j \in N$.

4. Strong, Perfect Equilibrium Payoffs of G^∞

It is shown here that the set of strong, perfect equilibrium payoffs coincides with the set of sustainable payoffs. The following lemma, a proof of which is to be found in Chakrabarti (1986) will be used. (See the Appendix for a proof.)

Lemma 4.1: Let v be a vector in the convex hull of the payoff vector of the game G. Then there is a summable n-tuple of strategies of the game G^∞ such that

$$\lim_{\nu\to\infty} \frac{1}{\nu} \sum_{t=2}^{\nu} U^j(x_{t-1}, x_t) = v_j \text{ for all } j \in N$$

where $\underline{x} \in X^\infty$ is the outcome path associated with h and x_t is the outcome in period t.

We first show that every sustainable or s payoff is the payoff of a strong, perfect equilibrium of G^∞.

Theorem 4.2: Let v^* be an s payoff. Then there is a supergame strategy n-tuple $h^* \in H$ which is a strong, perfect equilibrium of G^∞ such that the outcome path x^* associated with h^* satisfies

$$\lim_{\nu\to\infty} \frac{1}{\nu} \sum_{t=2}^{\nu} U^j(x^*_{t-1}, x^*_t) = v^*_j \text{ for all } j \in N.$$

Proof: From Lemma 4.1 we know that there is an assignment map $\psi\colon \mathbb{N} \to X$ and a class $H(\psi, v^*)$ of inertia supergame strategies such that for all $h \in H(\psi, v^*)$, we have

$$\lim_{\nu\to\infty} \frac{1}{\nu} \sum_{t=2}^{\nu} U^j(x_{t-1}, x_t) = v^*_j \text{ for all } j \in N$$

where $\underline{x} \in X^\infty$ is the outcome path associated with $h \in H$.

It will be shown that there is an $h^* \in H(\psi, v^*)$ such that h^* is a strong, perfect equilibrium of G^∞.

Consider any history $[x_1, \ldots, x_T]$ up to the time period T, and define the deviating coalition $S_{T+1}[x_1, \ldots, x_T]$ when $h^* \in H$ is agreed upon, as follows:

Let $S(\phi) = \phi$,

$$S_{T+1}[x_1, \ldots, x_T] := \begin{cases} \{j \in S_T / \sum_{\{t \leqslant T / x^j_{t-1} = x^j_t\}} U^j(x_{t-1}, x_t) \geqslant v^*_j + 1/\sqrt{T}\} \\ \cup \{j \notin S_T / h_j(T+1) \neq h^*_j(T+1) \text{ given the history } \\ x_1, \ldots, x_T\}. \\ \\ \phi, \text{ otherwise.} \end{cases}$$

Hence, given the plan to play $h^* \in H$, the deviating coalition at any time t is the collection of past deviators who are making profits from their deviations and those who have not conformed to the proposed play in time t.

The strategy $h^* \in H(\psi, v^*)$ is now defined as follows:

$h^*(1) = \psi(1)$

$$h_i^*(T+1) := \begin{cases} [x_\ell^{N\setminus S_T}(x_\ell^{S_T})]_i \text{ if } S_T[x_1,\ldots,x_{T-1}] \neq \phi \text{ and} \\ \quad i \notin S_T[x_1,\ldots,x_{T-1}] \text{ and } S_T \text{ played } x_\ell^{S_T} \text{ in period T.} \\ \\ \text{Arbitrary, if } S_T[x_1,\ldots,x_{T-1}] \neq \phi \text{ and} \\ \quad i \in S_T[x_1,\ldots,x_{T-1}]. \\ \\ \psi_i(T+1-k) \text{ if } S_T[x_1,\ldots,x_{T-1}] = \phi \text{ and} \\ \quad k := \left|\{t \leqslant T / S_t[x_1,\ldots,x_{t-1}] \neq \phi\}\right| \end{cases}$$

given the history $x_1,\ldots,x_{T-1}$ and the outcome in period T.

The plan is to play the response $x_\ell^{N\setminus S}$ if the deviators S played x_ℓ^S in the preceding period. The responses $x_\ell^{N\setminus S}$ exist because v^* is a sustainable payoff. If there are no deviators then the plan is to continue playing according to the assignment $\psi: \mathbb{N} \to X$.

It will be shown that $h^* \in H(\psi, v^*)$ is a strong, perfect equilibrium of the inertia supergame.

Consider $S \in \Lambda$ and $h^S \neq (h^*)^S$. Then either (i) there exists a T_0 such that for all $T \geqslant T_0$, $S_T = \phi$, or (ii) for every T_0, there exists a $T' \geqslant t_0$ such that $S_{T'} \neq \phi$, $S_{T'} \subseteq S$.

<u>Case (i)</u>: This is obvious since

$$U^j_\infty(h^S,(h^*)^{N\setminus S}) = \lim_{\nu\to\infty} \frac{1}{T_0 + \nu} \sum_{t=T_0}^{\nu} U^j(h^*) = v^*_j \text{ for all } j \in N.$$

Case (ii): Since for every T_0, there exists a $T' \geq T_0$ such that $S_{T'} \subseteq S$, and S is finite, there is a $B \subseteq S$ such that for every T_0, there exists a $T' \geq T_0$ for which $S_{T'} = B$. Hence, without any loss of generality we can assume $S = B$.

Then

$$(4.1) \qquad \frac{1}{T' + t_0} \sum_{t=1}^{T'+t_0} U^j[(h^S(t),(h^*)^{N\setminus S}(t)),(h^S(t-1),(h^*)^{N\setminus S}(t-1))]$$

$$\leq \frac{1}{T' + t_0} \sum_{\{t\geq(T'+t_0)/x^j_{t-1}=x^j_t\}} U^j[(h^S(t),(h^*)^{N\setminus S}(t)),$$

$$(h^S(t-1),(h^*)^{N\setminus S}(t-1))]$$

$$- \frac{1}{T' + t_0} \sum_{\{t\leq(T'+t_0)/x^j_{t-1}\neq x^j_t\}} [C_j(x^j_{t-1},x^j_t) - \phi_j(S_t,x_t)].$$

for all $j \in S$.

Now,

$$(4.2) \qquad \frac{1}{T' + t_0} \sum_{\{t\leq(T'+t_0)/x^j_{t-1}=x^j_t\}} U^j[(h^S(t),(h^*)^{N\setminus S}(t)),$$

$$h^S(t-1),(h^*)^{N\setminus S}(t-1))]$$

$$= \sum_{\{t\leq(T'+t_0)/x^j_{t-1}=x^j_t\}} U^j[(h^S(t),(h^*)^{N\setminus S}(t)),$$

$$(h^S(t - 1),(h^*)^{N\setminus S}(t - 1))]$$

$$\leqslant v_j^* + 1/\sqrt{T' + t_0}$$

for some $j \in S$ when it is sufficiently large. Inequality (4.2) holds because v^* is an s-payoff and $S_{T'} = S$.

Using inequality (4.2) in inequality (4.1) and recalling Assumption 4, we have

$$(4.3) \qquad \frac{1}{T' + t_0} \sum_{t=1}^{T'+t_0} U^j[(h^S(t),(h^*)^{N\setminus S}(t)),(h^S(t - 1),(h^*)^{N\setminus S}(t - 1))]$$

$$\leqslant v_j^* + 1/\sqrt{T' + t_0}$$

for some $j \in S$ if t_0 is sufficiently large.

Since S is finite and inequality (4.3) holds for infinitely many T', there exists a $j_0 \in S$ such that (4.3) holds for j_0 for infinitely many $T' \in N$. But this means that

$$(4.4) \qquad \underline{\lim}_{\nu\to\infty} \frac{1}{\nu} \sum_{t=1}^{\nu} U^{j_0}(h^S(t),(h^*)^{N\setminus S}(t)) \leqslant v_{j_0}^*$$

for some $j_0 \in S$.

Hence, since S and $h^S \in H^S$ is arbitrary, we have shown that $h^* \in H(\psi,v^*)$ is a strong equilibrium.

Given any history $(x_1,\ldots,x_T)$, the subgame $G^\infty|_{(x_1,\ldots,x_T)}$ has the same payoff structure as G^∞ since the payoffs from any finite history do not affect the limit of the averages. Hence, for the situation arising in case (i), $h^* \in H(\psi,v^*)$ clearly induces a strategy n-tuple on $G^\infty|_{(x_1,\ldots,x_T)}$ which gives a payoff of v_j^* to each $j \in N$.

For the situation in case (ii), take $T' \geqslant T$ and the same argument as in case (ii) will show that h^* induces a strategy n-tuple on $G^\infty|_{(x_1,\ldots,x_T)}$ which satisfies inequaltiy (4.4) for every S and for some $j \in S$. This concludes the proof. []

Theorem 4.2 has shown that every s-payoff is the payoff of some strategy n-tuple of the inertia supergame. The next result shows that the converse is also true--that is, the payoff vector generated by a strong, perfect equilibrium of the inertia supergame is an s-payoff of the one shot game.

Theorem 4.3: If $h^* \in H$ is a summable, strong, perfect equilibrium of the inertia supergame G^∞, then there is an s-payoff v^* such that

$$\lim_{\nu\to\infty} \frac{1}{\nu} \sum_{t=1}^{\nu} U^j(h^*(t)) = v_j^*$$

for all $j \in N$.

Proof: Suppose v^* is not an s-payoff. Then there exists an $S \in \Lambda$ and a finite index set $\bar{L}$ and $\{\bar{x}_\ell^S \in X^S / \ell \in \bar{L}\}$ and $\bar{p} \in \Delta(\bar{L})$ such that for all $\{x_\ell^{N\setminus S} \in X^{N\setminus S} / \ell \in \bar{L}\}$, and for all $j \in S$

$$\sum_{\ell\in\bar{L}} \bar{p}_\ell U^j(\bar{x}_\ell^S, x_\ell^{N\setminus S}) > v_j^*. \tag{4.5}$$

Define

$$\bar{L}(x^{N\setminus S}) = \{(x_\ell^{N\setminus S})_{\ell\in\bar{L}} / x_\ell^{N\setminus S} \in X^{N\setminus S}\}$$

Then

$$\min_{\bar{L}(X^{N\setminus S})} \sum_{\ell\in\bar{L}} p_\ell U^j(x_\ell^S, x_\ell^{N\setminus S})$$

is a continuous function on $\bar{L}(X^S) \otimes \Delta(\bar{L})$. Because of (4.5), for each $j \in S$ there is an open neighborhood U_j of $[\{\bar{x}_\ell^S \in X^S/\ell \in \bar{L}\}, \bar{p}]$ in $\bar{L}(X^S) \otimes \Delta(\bar{L})$ such that

$$\min_{\bar{L}(X^{N\setminus S})} \sum_{\ell\in\bar{L}} p_\ell U^j(x_\ell^S, x_\ell^{N\setminus S}) > v_j^* \tag{4.6a}$$

for all $[(x_\ell^S)_{\ell\in\bar{L}}, p] \in U_j$.

Hence, there is an open neighborhood $U = \bigcap_{j\in S} U_j$ of $[(\bar{x}_\ell^S)_{\ell\in\bar{L}}, \bar{p}]$ such that

$$\min_{\bar{L}(X^{N\setminus S})} \sum_{\ell\in\bar{L}} p_\ell U^j(x_\ell^S, x_\ell^{N\setminus S}) > v_j^* \tag{4.6b}$$

for all $[(x_\ell^S)_{\ell\in\bar{L}}, p]$ in U and for all $j \in S$.

Hence, for each $j \in S$ there is an $\varepsilon_j > 0$ such that

$$\begin{aligned} \sup_{\bar{L}(x^S)\otimes\Delta(\bar{L})} \min_{\bar{L}(X^{N\setminus S})} \sum_{\ell\in\bar{L}} p_\ell U^j(x_\ell^S, x_\ell^{N\setminus S}) \\ \geq \sup_{U} \min_{\bar{L}(X^{N\setminus S})} \sum_{\ell\in\bar{L}} p_\ell U^j(x_\ell^S, x_\ell^{N\setminus S}) > v_j^* - \varepsilon_j. \end{aligned} \tag{4.7}$$

Hence, if $\varepsilon = \min\{\varepsilon_j/j \in S\}$, then

$$\min_{j\in S} \sup_{\bar{L}(X^S)\otimes\Delta(\bar{L})} \sum_{\ell\in\bar{L}} p_\ell U^j(x_\ell^S, x_\ell^{N\setminus S}) > v_j^* - \varepsilon \tag{4.8}$$

for $\varepsilon > 0$.

Step 2: From (4.8) it follows that there exist $(\bar{x}_\ell^S)_{\ell\in\bar{L}}$ and $\bar{p} \in \Delta(\bar{L})$ such that

$$(4.9) \qquad \min_{\bar{L}(X^{N\setminus S})} \sum_{\ell\in\bar{L}} p_{\bar{\ell}} U^j(\bar{x}^S_\ell, x^{N\setminus S}_\ell) > v^*_j - \varepsilon/2$$

for all $j \in S$.

Hence we can define an assignment map ψ^S such that $\psi^S\colon \mathbb{N} \to X^S$ and

$$(4.10) \qquad \frac{1}{\nu}\sum_{t=1}^{\nu} U^j(\psi^S(t), h^{N\setminus S}(t) \geq \min_{\bar{L}(X^{N\setminus S})} \sum_{\ell\in\bar{L}} p_{\bar{\ell}} U^j(\bar{x}^S_\ell, x^{N\setminus S}_\ell) - \varepsilon/4$$

for all $j \in S$ and for all $\nu \geq T_0$, where T_0 is sufficiently large and for any $h^{N\setminus S} \in H^{N\setminus S}$.

The assignment map Ψ^S is defined by choosing $\bar{x}^S_\ell \in X^S$ in a proportion $p_{\bar{\ell}}$ out of the ν repetition of the game G for every $\ell \in \bar{L}$, taking care that changes of strategies occur a small enough number of times so that the cost from changing strategies goes to zero. A precise proof follows the same reasoning as used in the proof of Lemma 4.1.

For $S \in \Lambda$, define $h^S \in H^S$ as follows:

$h_j(t) := \psi_j(t)$, for all $t \in N$ and for all $j \in S$.

Then, form (4.10)

$$(4.11) \qquad \frac{1}{\nu}\sum_{t=1}^{\nu} U^j(h^S(t), h^{N\setminus S}(t)) \geq \min_{\bar{L}(X^{N\setminus S})} \sum_{\ell\in\bar{L}} p_{\bar{\ell}} U^j(\bar{x}^S_\ell, x^{N\setminus S}_\ell) - \varepsilon/4$$

for all $j \in S$ and $h^{N\setminus S} \in H^{N\setminus S}$, if $\nu \geq T_0$. Therefore,

$$(4.12) \qquad \frac{1}{\nu}\sum_{t=1}^{\nu} U^j(h^S(t), h^{N\setminus S}(t)) > v^*_j - \varepsilon/4$$

for $\nu \geq T_0$. This follows from (4.10).

In particular,

$$(4.13) \qquad \frac{1}{\nu} \sum_{t=1}^{\nu} U^j(h^S(t),(h^*)^{N\setminus S}(t)) > v_j^* - \varepsilon/4$$

for $\nu \geq T_0$ for all $j \in S$. Hence

$$(4.14) \qquad \varliminf_{\nu \to \infty} \frac{1}{\nu} \sum_{t=1}^{\nu} U^j(h^S(t),(h^*)^{N\setminus S}(t)) > v_j^* - \varepsilon/8$$

for all $j \in S$. This shows that $h^* \in H$ is not a strong equilibrium of G^∞. □

5. The s-payoffs of the Correlated α-core

Since the sustainable payoffs have the property that they coincide with the strong, perfect equilibrium payoffs of the inertia supergame G^∞, it is of interest to examine how different the set of sustainable is from any other core-like solution concept. The most natural concept to try would be the correlated α-core, since the s-payoffs seem to have a similar structure. We will show that the two concpets are very closely related when one makes certain assumptions about the correlating device used.

The structure that we will want on the strategy sets of the individual players is that they be compact subsets of complete, separable metric spaces with the Borel σ-algebra. Hence, we will make the following assumptions about the structure of the game G.

Assumption 5: The strategy sets X^j are compact metric spaces and β_{X^j} is the Borel σ-field of X^j, and $M(X^j)$ is the space of measures defined on X^j with that σ-field and is endowed with the weak topology.

The correlated α-core is now defined as follows:

Definition 5.1: v^* is in the correlated α-core of the game G if there is a measure $\mu^* \in M(X)$ such that

$$\int_X U^j(x)d(\mu^*) \geq v_j^*$$

for all $j \in S$, and if for all $S \in \Lambda$, and $\mu^S \in M(X^S)$ there exists a $\mu^{N\setminus S} \in M(X^{N\setminus S})$ such that

$$\int_S U^j(x)d(\mu^S \otimes \mu^{N\setminus S}) \leq v_j^*$$

for some $j \in S$.

Here $M(X^S)$, $M(X^{N\setminus S})$ is the space of probability measures which is the product of the measure spaces $M(X^j)$ for $j \in S$ and $j \in N\setminus S$, respectively.

Hence, an allocation in the correlated α-core is such that if a coalition deviates by playing some correlated strategy, then there is a correlated strategy that the complement can play such that at least one player of the deviating coalition is not made better off by the deviation.

Theorem 5.2: If the game G satisfies assumptions 1-5, then the correlated α-core is the closure of the set of s-payoffs of G, when the latter is not empty.

Proof. (Step 1): The set of s-payoffs is contained in the correlated α-core of G.

Suppose v^* does not belong to the correlated α-core of G. Then there exists an $S \in \Lambda$ and a $\bar{\mu}^S \in M(X^S)$ such that

$$\int_X U^j(x)d(\bar{u}^S \otimes \mu^{N\setminus S}) > v_j^* \tag{5.1}$$

for all $j \in S$, and for all $\mu^{N\setminus S} \in M(X^{N\setminus S})$.

Since $M(X^S)$ has the weak topology and X^S is a complete, separable metric space, the set of measures with finite support is dense in $M(X^S)$. Hence, there is a sequence of measures with finite supports $(\mu_n^S)_{n\in\mathbb{N}}$ which converges to $\bar{\mu}^S$. Now, since (5.1) holds there is an n_0, sufficiently large such that for every $n \geq n_0$ there is an open set U_n in $M(X^{N\setminus S})$ such that

$$(5.2) \qquad \int_X U^j(x)d(\mu_n^S \otimes z^{N\setminus S}) > v_j^*$$

for all $j \in S$ and $z^{N\setminus S} \in U_n$. Now, $\{U_n\}_{n\in\mathbb{N}}$ forms an open cover of $M(X^{N\setminus S})$. Since $X^{N\setminus S}$ is a compact metric space, $M(X^{N\setminus S})$ is a compact metric space. Hence, $\{U_n\}_{n\in\mathbb{N}}$ has a finite sub cover which covers $M(X^{N\setminus S})$. Since $(\mu_n^S)_{n\in\mathbb{N}}$ converges to $\bar{\mu}^S$, we can choose the open sets $\{U_n\}_{n\in\mathbb{N}}$ to be nested. Hence, there exists a $\mu_{n_0}^S \in M(X^S)$ such that

$$(5.3) \qquad \int_X U^j(x)d(\mu_{n_0}^S \otimes z^{N\setminus S}) > v_j^*$$

for all $j \in S$ and all $z^{N\setminus S} \in \mu(X^{N\setminus S})$. But this shows that v^* is not a sustainable payoff. This proves the claim.

(Step 2): The correlated α-core is a closed subset of $C_0 \cdot U(X)$ in the Euclidean space $\mathbb{R}^n$.

Let v^* be a limit point of the correlated α-core, and suppose v^* is not in the correlated α-core. Then there exists an $S \in \Lambda$, $\bar{\mu}^S \in M(X^S)$ such that for all $\mu^{N\setminus S} \in M(X^{N\setminus S})$,

$$(5.4) \qquad \int_X U^j(x)d(\bar{\mu}^S \otimes \mu^{N\setminus S}) > v_j^*$$

for all $j \in S$.

Now, the functions $f_j\colon M(X^{N\setminus S}) \to \mathbb{R}$ defined by

$$(5.5) \qquad f_j(\mu^{N\setminus S}) = \int_X U^j(x)d(\bar{\mu}^S \otimes \mu^{N\setminus S})$$

are continuous on $M(X^{N\setminus S})$, for all $j \in S$.

Since v^* is a limit point of the correlated α-core, for any $\varepsilon > 0$ there exists a v' in the correlated α-core such that

$$(5.6) \qquad \|v^* - v'\| < \varepsilon/2$$

But since (5.5) holds, we have

$$(5.7) \qquad f_j(\mu^{N\setminus S}) > v_j' - \varepsilon/2$$

for all $j \in S$ and $\mu^{N\setminus S} \in M(X^{N\setminus S})$. Hence, v' cannot be in the correlated α-core. We have a contradiction. This proves the claim.

(Step 3): If v^* is in the correlated α-core, then v^* is a limit point of the set of s payoffs.

Consider any v which is not a limit point of the set of s payoffs. Then there exists an $S \in \Lambda$, an index set $\bar{L}$,

$$\{(\bar{x}_\ell^S)_{\ell\in\bar{L}} / \bar{x}_\ell^S \in X^S\},$$

a $\bar{p} \in \Delta(\bar{L})$ and a $\delta > 0$ sufficiently small such that

$$(5.8) \qquad \sum_{\ell\in\bar{L}} \bar{p}_\ell U^j(\bar{x}_\ell^S, x_\ell^{N\setminus S}) > v_j^* - \delta$$

for all $j \in S$ and for all $\{(x_\ell^{N\setminus S})_{\ell\in\bar{L}} / x_\ell^{N\setminus S} \in X^{N\setminus S}\}$.

Now, $p^- \in \Delta(\bar{L})$ and $\{(\bar{x}_\ell^S)_{\ell\in\bar{L}} / x_\ell^S \in X^S\}$ induces a measure $\bar{\mu}^S \in M(X^S)$ which has finite support.

Since (5.8) holds for all $\{(x_\ell^{N\setminus S})_{\ell\in\bar{L}} / x_\ell^{N\setminus S} \in X^{N\setminus S}\}$ we have that

$$(5.9) \qquad \int_X U^j(x)d(\bar{\mu}^S \otimes \mu^{N\setminus S}) > v_j^* - \delta$$

for all $j \in S$ and for all $\mu^{N\setminus S} \in M(X^{N\setminus S})$ with finite support. Since the set of measures with finite support is dense in $M(X^{N\setminus S})$, we have

$$(5.10) \qquad \int_X u^j(x)d(\bar{\mu}^S \otimes \mu^{N\setminus S}) \geqslant v_j^* - \delta$$

for all $\mu^{N\setminus S} \in M(X^S)$ and for all $j \in S$. But this shows that there exists an $S \in \Lambda$ and a $\bar{\mu}^S \in M(X^S)$ such that for all $\mu^{N\setminus S} \in M(X^{N\setminus S})$ we have

$$(5.11) \qquad \int_X u^j(x)d(\bar{\mu}^S \otimes \mu^{N\setminus S}) > v_j^*$$

for all $j \in S$. Hence, v^* is not in the correlated α-core of G. []

6. Conclusion

We have shown that the s-payoffs completely describes the payoffs of the strong, perfect equilibrium points of the inertia supergame. The issue that is raised at this point is: How do the sustainable payoffs caompare with the other core-like concepts? Naturally, it is not difficult to show that it contains the β-core, the desired payoffs and the reactive oayoffs. However, the question of greatest interest is its relationship with the correlated α-core. In section 5 we showed that in a game G which has a nonempty set of s payoffs, if the correlating device gives rise to a Borel σ-algebra on the strategy sets of the players, then the correlated α-core is the closure of this set of s-payoffs. Hence, in this sense, the corelated α-core and the s-payoffs are quite closely related.

Throughout our analysis, we have only assumed that the strategy sets of the players be compact metric spaces. We have made no assumptions about the dimensionality of the strategy sets. However, an assumption that we have used in our proofs is that of the finiteness of the set of players. This assumption seems to be crucial to Theorems 4.2 and 5.2.

Appendix

Proof of Lemma 4.1: Since v is in the convex hull of the game G, there exists a finite index set F and $\{x_k/x_k \in X,\ k \in F\}$ such that v is a convex combinatiion of $\{U(x_k)/k \in F\}$. Let $v = \sum_{k \in F} \beta_k\ U(x_k)$, where

$$\beta_k > 0 \text{ for all } k \in F \text{ and } \sum_{k \in F} \beta_k = 1.$$

Case (i): The β_k's are all rational numbers. Since the coefficients β_k are all rational numbers there exists an integer T and an integer N_k for each $k \in F$ such that $N_k/T = \beta_k$ for all $k \in F$.

We define the map ψ_T: $\{1,\ldots,T\} \to X$ as follows:

$$\psi_T(t) = x_k \text{ for } \sum_{\ell<k} N_\ell < t < \sum_{\ell<k+1} N_\ell$$

where the elements in F are ordered in the order of integers determined according to some assignment. Then clearly,

$$\frac{1}{T} \sum_{t=1}^{T} U(\psi_T) = \sum_{k \in F} \beta_k U(x_k) = v.$$

Inductively, we define the map

$$\psi_{4T}\colon \{1,\ldots,T\} \to X$$

as follows

$$\psi_{4T}(t) := \begin{cases} \psi_T(t) \text{ for } t \leq T. \\ x_k \text{ for } T + 3 \sum_{\ell<k} N_\ell < t < T + 3\sum_{\ell<k+1} N_\ell \\ \text{for } k \geq 2 \\ x_1 \text{ for } T < t < 3N_1. \end{cases}$$

And $\psi_{n^2T}\colon \{1,\ldots,n^2T\} \to X$ as

$$\psi_{n^2T}(t) = \begin{cases} \psi_{(n-1)^2T}(t) \text{ for } t \leq T \\ x_k \text{ for } (n-1)^2T + [n^2 - (n-1)^2] \sum_{\ell<k} N_\ell \\ \quad < t < (n-1)^2T + [n^2 - (n-1)^2] \sum_{\ell<k+1} N_\ell \\ \text{for } k \geq 2 \\ x_1 \text{ for } (n-1)^2T < t < [n^2 - (n-1)^2]N_\ell. \end{cases}$$

We now define $\psi\colon \mathbb{N} \to X$ by extending the maps ψ_{n^2T} as follows:

$$\psi(t) = \psi_{n^2T}(t), \text{ if } (n-1)^2T < t \leq n^2T.$$

We now define the following strategy in the inertia supergame G^∞.

$$h_j(t) := \begin{cases} \psi_j(t) \text{ if the histroy up to } t-1 \text{ is given by} \\ \qquad \{h(s)/h(s) = \psi(s)\}_{s=1}^{p-1}. \\ \text{Arbitrary, otherwise.} \end{cases}$$

Then

$$\frac{1}{\nu}\sum_{t=1}^{\nu} U^j(h(t-1), h(t))$$

$$= \frac{1}{\nu}\sum_{t=1}^{\nu} [U^j(h(t) - C_j(h_j(t-1), h_j(t))]$$

$$\geq \frac{1}{\nu}\sum_{t=1}^{\nu} [U^j(h(t))] - \frac{\eta_\nu B}{\nu}.$$

where

$$\eta_\nu = \left|\{t/h_j(t) \neq h_j(t-1),\ t \geq \nu\}\right|,$$

$$B \geq C_j(x^j_{t-1}, x^j_t) \text{ for all } j \in N, \text{ all } x^j_t, x^j_{t-1} \in X^j.$$

Therefore, for all $\nu = n^2T$, we have

$$\frac{1}{\nu}\sum_{t=1}^{\nu} U^j(h(t-1), h(t))$$

$$= \frac{1}{n^2T}\sum_{k\in F} n^2N_kU^j(x_k) - \frac{1}{n^2T}\sum_{t=1}^{n^2T} C_j(h_j(t-1), h_j(t))]$$

$$\geq \sum_{k\in F}\frac{N_k}{T}U^j(x_k) - \frac{nM}{n^2T}$$

$$= \sum_{k\in F}\beta_kU^j(x_k) = \frac{n_kM}{n^2T}$$

Hence, for $\nu = n^2T$

$$\frac{1}{\nu}\sum_{t=1}^{\nu} U^j(h(t-1), h(t)) \geq v_j - \frac{kM}{nT} = v_j - C/n \text{ where } C = \frac{KM}{T}.$$

Hence, we have

(A.1) $$v_j = \sum_{k\in F} \beta_k U^j(x_k) \geqslant \frac{1}{\nu} \sum_{t=1}^{\nu} U^j(h(t-1), h(t)) \geqslant v_j - \frac{C}{n}.$$

for $\nu = n^2T$. (A.1) holds for all n and all $j \in N$. Now

$$\left|\frac{1}{n^2T+\ell} \sum_{t=n^2T+1}^{n^2T+\ell} U^j(h(t))\right|$$

where $1 \leqslant \ell \leqslant (2n+1)T$

(A.2) $$\leqslant \frac{1}{n^2T+\ell} \sum_{t=n^2T+1}^{n^2T+\ell} \left|U^j(h(t))\right|$$

$$\leqslant \frac{1}{n^2T+\ell}\, \ell S, \text{ where S is a bound for } \left|U^j(h(t))\right|$$

$$= \frac{S}{(n^2T)/\ell + 1}.$$

This is true for all n and all $j \in N$. This goes to zero as n goes to infinity for every $1 \leqslant \ell \leqslant (2n+1)T$.

Hence, we have that

$$\lim_{\nu\to\infty} \frac{1}{\nu} \sum_{t=1}^{\nu} U^j(h(t-1), h(t)) = v_j, \text{ for all } j \in N.$$

Case (ii): β_k's are not all rational numbers.

For $\varepsilon > 0$, we can find integer $N_k^{(1)}$ and $T^{(1)}$ such that

$$\left|\frac{N_k^{(1)}}{T^{(1)}} - \beta_k\right| < \varepsilon/M$$

for all $k \in F$, where $M = \sum_{k\in F} U(x_k)$. Then

$$v^{(1)} = \sum_{k \in F} \frac{N_k^{(1)}}{T^{(1)}} U(x_k) \text{ satisfies } \|v^{(1)} - v\| < \varepsilon.$$

We define the sequence $\{v^{(n)}\}_{n \in N}$ such that

$$v^{(n)} = \sum_{k \in F} \frac{N_k}{T^{(n)}} U(x_k)$$

where $N_k^{(n)}/T^{(n)}$ converges to β_k for each $k \in F$, and the sequence $T^{(n)}$ is increasing with $|T^{(n+1)} - T^{(n)}| \leq A(n)$, where the $A(n)$ satisfy:

(A.3) $$\frac{A(n)}{T^{(n)}} \to 0 \text{ as } n \to \infty.$$

Hence, v^n converges to v. We now define the map $\psi: \mathbb{N} \to X$, such that for the first $T^{(1)}$ period the blocks of time $N_k^{(1)}$ are assigned to x_k; the second $4T^{(2)} - T^{(1)}$ block is again broken into k blocks of length $4N_k^{(2)} - N_k^{(1)}$ each and assigns x_k to each block. The third $9T^{(3)} - 4T^{(2)}$ periods is broken into k blocks of length, $9N_k^{(3)} - 4N_k^{(2)}$ and each such block is assigned x_k.

The map $\psi: \mathbb{N} \to X$ is defined inductively as above.

We now define $h \in H$ as follows.

For $j \in N$:

$$h_j(t) := \begin{cases} \psi_j(t) \text{ if the history up to } t-1 \text{ is given by} \\ \qquad \{h(s)/h(s) = \psi(s)\}_{s=1}^{t-1} \\ \text{Arbitrary otherwise.} \end{cases}$$

Then

$$\frac{1}{\nu}\sum_{t=1}^{\nu} U^j(h(t-1),\ h(t))$$

$$= \frac{1}{\nu}\sum_{t=1}^{\nu} [U^j(h(t)) - C_j(h_j(t-1),\ h_j(t))]$$

$$\geqslant \frac{1}{\nu}\sum_{t=1}^{\nu} [U^j(h(t))] - \frac{\eta_\nu B}{B}$$

where

$$\eta_\nu = \left|\{t/h_j(t) \neq h_j(t-1),\ t \leqslant \nu\}\right|.$$

For $\nu = n^2 T^{(n)}$, we have $\eta_\nu = n$, so that

$$\frac{\eta_\nu B}{\nu} = \frac{B}{n^2 T^{(n)}},$$

and

$$\text{(A.4)} \qquad \frac{1}{\nu}\sum_{t=1}^{\nu} [U^j(h(t))] = v_j^{(n)},\ \text{for all } j \in N.$$

Now,

$$\left|\frac{1}{n^2 T^{(n)} + \ell}\sum_{t=n^2 T^{(n)}+1}^{n^2 T^{(n)}+\ell} U^j(h(t))\right|,\ \text{for } 1 \leqslant \ell < (n+1)^2 T^{(n+1)} - n^2 T^{(n)}$$

$$\leqslant \frac{1}{n^2 T^{(n)}}\sum_{t=n^2 T^{(n)}+1}^{n^2 T^{(n)}+\ell} \left|U^j(h(t))\right|$$

$$\leqslant \frac{1}{n^2 T^{(n)} + \ell}\,\ell S,\ \text{where } S \text{ is a bound for } \left|U^j(h(t))\right|$$

$$= \frac{S}{n^2 T^{(n)}/\ell + 1}$$

$$\leq \frac{S}{n^2 T^{(n)}/[(n + 1)^2 T^{(n+1)} - n^2 T^{(n)}] + 1}$$

and this goes to zero as $n \to \infty$ because of (A.3). Hence, by (A.4) and this fact, it follows that $1/\nu \sum_{t=1}^{\nu}[U^j(h(t))]$ converges to v_j for all $j \in N$. This concludes the proof. $\square$

References

Aummann, R. J., "Acceptable Points in General Cooperative n-person Games," in Contributions to the Theory of Games, vol. iv, edited by Tucker and Luce, Annals of Mathematical Studies, 40, 1959.

Aumann, R. J., "Acceptable Points in Games of Perfect Information," Pacific Journal of Mathematics, 10, 1960, 381-417.

Benoit, J. P. and V. Krishna, "Finitely Repated Games," Econometrica, 53, no. 4, July 1985, 905-922.

Chakrabarti, S. K., "A Set of Payoffs Which are Generated by Strong, Perfect Equilibria of the Inertia Supergames," working paper no. 8602, Committee for Mathematical Analysis in the Social Sciences, Graduate College, The University of Iowa, March 1986.

Friedman, J. W., "Cooperative Equilibria in Finite Horizon Noncooperative Supergames," Journal of Eocnomic Theory, 35, 1985, 390-398.

Marschak, T. and R. Selten, "Restabilizing Responses, Inertia Supergames, and Oligopolistic Equilibria," Quarterly Journal of Economics, February 1978, 71-91.

Parthasarathy, K. R., Probability Measures on Metric Spaces, Academic Press, 1967.

Rubinstein, A., "Strong Perfect Equilibrium in Supergames," International Journal of Game Theory, 9, no. 1, 1980, 1-12.

Selten, R., "Reexamination of the Perfectness Concepts for Equilibrium Points in Extensive Games," International Journal of Game Theory, 4, 1975, 25-55.