Journal of Mathematical Economics 18 (1989) 281–290. North-Holland

MORE ON THE 'ANTI-FOLK THEOREM'*

Jordi MASSO

University of Pittsburgh, Pittsburgh, PA 15260, USA

Robert W. ROSENTHAL

Boston University, Boston, MA 02215, USA

Submitted September 1987, accepted October 1988

For discounted repeated games with unobservable individual deviations, Kaneko's 'anti-folk theorem' states that the set of Nash-equilibrium plays coincides with the set of sequences of one-shot Nash-equilibrium plays. When the payoff criterion is long-run average, however, Kaneko's characterization is of a different sort. Here we show that with some additional topological assumptions a version of the anti-folk theorem is available under the long-run average criterion which is parallel to the characterization under the discounting criterion.

1. Introduction

The phrase 'folk theorem' in game theory refers to any of a collection of results having the rough form: The set of Nash-equilibrium payoffs in a repeated game equals the set of feasible, individually-rational payoffs in the one-shot (stage) game.[1] The proofs of such theorems invariably involve strategies which identify and punish individual deviators from the long-run plan. When restrictions are imposed on the information pattern of the repeated game such that individual deviators cannot be identified (for example, a game with a large number of individually insignificant players), it is therefore natural to expect the Nash equilibria of the repeated game to form a much smaller set. Results that characterize equilibria in this setting are termed collectively the 'anti-folk theorem' by Kaneko (1982) and Dubey and Kaneko (1984). For the case in which the payoff sequences in the repeated game are evaluated by discounting or by the overtaking criterion [e.g. Rubinstein (1979)], the anti-folk theorem of Kaneko takes its most extreme form: The set of Nash-equilibrium plays of the repeated game equals

*Research supported by NSF grant #SES 8603550 at SUNY, Stony Brook and by Bell Communications Research. We are grateful to A. Mas-Colell for helpful comments

[1]See, for example, Aumann (1981).

the set of sequences composed of one-shot Nash equilibria. For the case in which the payoff sequences are evaluated by the long-run average criterion, however, Kaneko's characterization (see Theorem 1 below) is both less extreme and less useful.

Our purpose here is to give another 'anti-folk' type characterization of Nash equilibria when individual deviators cannot be identified, under the long-run average criterion. We show that under certain additional assumptions a version of the anti-folk theorem can be obtained that is close in spirit to the discounted version; roughly: The set of equilibrium plays of the repeated game equals the set of sequences of one-shot outcomes having the property that every accumulation point of such a sequence that is the limit of a 'non-negligibly occurring subsequence' is itself an equilibrium outcome of the one-shot game.

In the next section of the paper, the model is described and the main result and a corollary are presented. Section 3 is devoted to a discussion of the main results, the differences between the results in the discounted versus long-run-average versions of the anti-folk theorem, and the need for the additional assumptions. The proof of the main result is deferred to section 4.

2. Model and results

Let (I, S, λ) be a measure space, where I is the set of players (possibly finite), S is a σ-algebra of subsets of I, and λ is a measure on (I, S).

In the one-shot (stage) game denoted G, A_i denotes the set of feasible actions for each $i \in I$. The elements $a_i \in A_i$ may be interpreted as either pure or randomized; we do not allow for additional randomizations over the elements of A_i. Assume that some σ-algebra is associated with $\bigcup_{i\in I} A_i$, and denote by $\hat{A}$ the set of feasible joint actions, defined by

$$\hat{A} \equiv \left\{ \hat{a}: I \to \bigcup_{i \in I} A_i \colon \hat{a}(i) \in A_i \ \forall\, i \in I, \text{ and } \hat{a} \text{ is measurable} \right\}.$$

For $\hat{a}, \hat{b} \in \hat{A}$ we say that $\hat{a}$ and $\hat{b}$ are equivalent if they differ only on a set of zero λ-measure. Denote by A the set of equivalence classes of $\hat{A}$. [Of course, if $\lambda(\{i\}) > 0\ \forall\, i \in I$, each equivalence class contains only one element; in this case, A and $\hat{A}$ may be viewed as identical.] For each $i \in I$, the payoff function is $h_i: A_i \times A \to R$. We assume that $(h_i)_{i \in I}$ is a collection of bounded functions such that for every $a \in A$ the function $h_i(a(i), a)$ is measurable as a function of i.[2] [For the case $\lambda(\{i\}) > 0$, we assume that if $a(j) = b(j)$ for λ-a.e. $j \in (I - \{i\})$, then $h_i(a_i', a) = h_i(a_i', b)\ \forall\, a_i' \in A_i$.]

[2]We adopt the usual convention that if a statement holds for all elements of an equivalence class, then we say that it holds for the class itself; in particular, $h_i(a(i), a)$ measurable means that for all $\hat{a}$ in the class a, $h_i(\hat{a}(i), a)$ is measurable.

A Nash equilibrium of G is an $a \in A$ such that λ-a.e.:

$$h_i(a(i), a) \geqq h_i(a'_i, a) \quad \forall\, a'_i \in A_i,$$

Let A^* denote the set of Nash equilibria of the stage game.

In the (undiscounted) repeated game G^∞, time is indexed by t taking values in N, the set of natural numbers. For each $t \in N$, the history of play through t is described by A^t, the t-fold Cartesian product of A with itself, with typical element $(a^1, \ldots, a^t)$.[3] A strategy for player i is a sequence of functions $f_i = (f_i^1, f_i^2, \ldots)$ satisfying:

(i) $f_i^1 \in A_i$;

and $\forall\, t \in N$,

(ii) $f_i^{t+1}: A^t \to A_i$.

The informational restriction which insures that individual deviations cannot be detected is $\forall\, t \in N$:

(iii) if $(a^1, \ldots, a^t), (b^1, \ldots, b^t) \in A^t$ and if for each $\tau = 1, \ldots, t$ the cardinality of $\{j \in I - \{i\}: a^\tau(j) \neq b^\tau(j)\}$ does not exceed one, then $f_i^{t+1}(a^1, \ldots, a^t) = f_i^{t+1}(b^1, \ldots, b^t)$.

[Note that (iii) is redundant when (I, S, λ) is atomless.] Let F_i denote the set of f_i satisfying (i), (ii), and (iii) (i's strategy set), and let

$$F = \{(f_i)_{i \in I}: f_i \in F_i \ \forall\, i \in I;\ f_i^1 \text{ is measurable as a function of } i; \text{ and } \forall t \geqq 1 \text{ and } \forall (a^1, \ldots, a^t) \in A^t,\ f_i^{t+1}(a^1, \ldots, a^t) \text{ is measurable as a function of } i\}$$

denote the set of feasible joint strategies. Given $f \in F$, the play it produces is identified as follows: let $a^1(f) \in A$ be such that $a^1(f)(i) = f_i^1$, λ-a.e., and, recursively, $a^{t+1}(f) \in A$ be such that $a^{t+1}(f)(i) = f_i^{t+1}(a^1(f), \ldots, a^t(f))$, λ-a.e.

For player $i \in I$, let

$$h_i^T(f) \equiv T^{-1} \sum_{t=1}^{T} h_i(f_i^t(a^1(f), \ldots, a^{t-1}(f)), a^t(f)).$$

To define payoffs generally in undiscounted repeated games, one must choose some function between the extremes

$$\bar{H}_i(f) \equiv \limsup_{T \to \infty} h_i^T(f) \quad \text{and} \quad \underline{H}_i(f) \equiv \liminf_{T \to \infty} h_i^T(f).$$

[3]Superscripts that do not designate footnotes are time indices throughout this paper. They are, in addition, exponents only when it is obvious from the context.

We need to retain the linear structure of finite averages, so the natural choice is as follows. First, associate with each $f \in F$ the sequence $\{h_i^t(f)\}_{t=1}^{\infty} \in \ell_{\infty}$. Next, take any Banach limit[4] on ℓ_{∞} and define the payoff function $H_i(f)$ to be the Banach limit evaluated at the sequence $\{h_i^t(f)\}$. Given $f \in F$ and $g_i \in F_i$, let $(f|g_i) \in F$ denote the joint strategy f with player i switching to g_i, defined by

$$(f|g_i)_j = \begin{cases} g_i & \text{if} \quad j=i \\ f_j & \text{if} \quad j \neq i. \end{cases}$$

A Nash equilibrium of the repeated game G^{∞} is an $f \in F$ such that λ-a.e.:

$$H_i(f) \geqq H_i(f|g_i) \quad \forall g_i \in F_i.$$

Let F^* denote the set of Nash equilibria of G^{∞}.

In Kaneko's anti-folk theorem the lim inf criterion is used. Hence, let $\underline{F}^*$ be the set of Nash equilibria relative to the payoff function $\underline{H}_i$.

Theorem 1 (Kaneko). *The strategy combination $f \in \underline{F}^*$ if and only if, λ-a.e.*:

$$\underline{H}_i(f) \geqq \liminf_{T \to \infty} T^{-1} \sum_{t=1}^{T} h_i(a_i^t, a^t(f)) \quad \forall (a_i^1, a_i^2, \ldots) \in \mathop{\times}_{t=1}^{\infty} A_i.$$

This result does not relate equilibria of G^{∞} to equilibria of G, however.

Before stating our main result, we need some additional notation. First, endow A with a topology. Next, for each $f \in F$, let

$$X(f) \equiv \{a \in A : a \text{ is the limit point of some subsequence of } \{a^t(f)\}\}.$$

For any infinite, (directed) subset B of N and any $T \in N$, let $B(T)$ denote the set $(B \cap \{1, \ldots, T\})$ and $\# B(T)$ its cardinality. Now, for $f \in F$, let

$$\underline{Z}(f) \equiv \left\{ a \in X(f) : \text{if, for any } B, \{a^t(f)\}_{t \in B} \to a, \right.$$

$$\left. \text{then } \liminf_{T \to \infty} \# B(T)/T = 0 \right\},$$

[4]A Banach limit on ℓ_{∞} is a linear function LIM satisfying,

$$\liminf_{t \to \infty} x^t \leqq \mathrm{LIM}(x) \leqq \limsup_{t \to \infty} x^t, \qquad \mathrm{LIM}(x^1, x^2, \ldots) = \mathrm{LIM}(x^2, x^3, \ldots).$$

The existence of such functions is a well-known consequence of the Hahn–Banach theorem.

and

$$\bar{Z}(f) \equiv \left\{ a \in X(f)\text{: if, for any } B,\ \{a^t(f)\}_{t\in B} \to a, \right.$$

$$\left. \text{then } \limsup_{T\to\infty} \# B(T)/T = 0 \right\}.$$

In words, $\underline{Z}(f)$ and $\bar{Z}(f)$ are two ways to describe the set of limit points of negligible subsequences of $\{a^t(f)\}$ only.

Theorem 2. If A is sequentially compact;[5] *h_i is continuous in its second argument λ-a.e.; $\{a^t\}_{t\in N} \to a$ implies $\{h_i(a^t(i), a^t)\}_{t\in N} \to h_i(a(i), a)$ λ-a.e.; and $f \in F$; then*

(1) $f \in F^*$ implies $X(f) \subset (A^* \cup \underline{Z}(f))$, and
(2) $X(f) \subset (A^* \cup \bar{Z}(f))$ implies $f \in F^*$.

Proof. See section 4.

Note that the continuity assumptions in Theorem 2 are weaker than joint continuity.

Finally, an easy consequence of Theorem 2.

Corollary. If I is finite or countably infinite; for every $i \in I$, A_i is a sequentially compact topological space; and each h_i is continuous in its second argument A, endowed with the product topology; then the conclusion of Theorem 2 holds.

Proof. Follows from continuity of the projection operator and the fact that the countable product of sequentially compact spaces is sequentially compact in the product topology.

3. Discussion

The logic behind the discounted version of the anti-folk theorem is straightforward: If at some stage of the repeated game an outcome that is not a one-shot equilibrium could arise, then some player (or non-null set of players in the continuum-player version) could deviate profitably at that stage with no other ramifications, since his action is by hypothesis not discernible to others; conversely, any repeated-game strategy combination

[5]The set A is sequentially compact if every sequence in A has a converging subsequence with limit in A. If A is compact and first countable, then A is sequentially compact.

generating a sequence of one-shot equilibria does not admit a profitable deviation. Note that this argument makes no use of topological assumptions.

Under the long-run-average payoff criterion, what occurs at any finite set of stages has no direct effect on average payoffs; one therefore expects a larger set of equilibrium plays in the repeated game. Indeed, it can be the case that no player ever uses his part of any one-shot equilibrium in an equilibrium play of the repeated game, as is made obvious by the trivial instance of a one-player stage game in which the typical action (a) is taken from the interval $[0, 1]$ and produces the payoff (a). Since

$$\lim_{t\to\infty} a^t = 1 \quad \text{implies} \quad \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} a^t = 1,$$

any such $\{a^t\}$ sequence constitutes an equilibrium of the repeated game.

Clearly it is only limits of subsequences that matter, and the same example illustrates the need to distinguish between $\bar{Z}(f)$ and $\underline{Z}(f)$ in Theorem 2: it is not hard to produce a strategy f for this example which generates a sequence $\{a^t\}$ of zeroes and ones s.t.

$$\liminf_{T\to\infty} T^{-1} \sum_{t=1}^{T} a^t < 1 = \limsup_{T\to\infty} T^{-1} \sum_{t=1}^{T} a^t$$

and such that if B is the index set for the sequence of zeroes, then

$$\liminf_{T\to\infty} \# B(T)/T = 0 < \limsup_{T\to\infty} \# B(T)/T.$$

Now, if the Banach limit agrees with lim sup on this sequence (which is possible using the standard constructive proof of the Hahn–Banach Theorem), the sequence constitutes a Nash equilibrium, but $0 \notin \bar{Z}(f)$. On the other hand, if the Banach limit agrees with lim inf on this sequence (similarly possible), the sequence does not constitute a Nash equilibrium, but $0 \in \underline{Z}(f)$.

The need for a little care with the hypotheses is illustrated by the following examples.

Example 1. In the stage game, the player set I is $(1, 2, \ldots)$; the action set A_i is $\{0, 1\}$, the same for each $i \in I$; and h_i for each i resulting from any action combination $\{a_j : j \in I\}$ is:

$$a_i \text{ if } \sum_{j\in I} a_j < \infty, \quad \text{and} \quad (-a_i) \text{ otherwise.}$$

It is known [Peleg (1969)] that this one-shot game has no equilibria – even when the natural set of randomized actions is allowed. Consider, however, the strategy combination for the repeated game given by: At each time $t\in\{1,2,\dots\}$ the players named $1,\dots,t$ play 1 while the rest play 0, independently of the history. Since every player's long-run average payoff from this strategy combination is 1, the maximum possible, the strategy combination must constitute an equilibrium of the repeated game (and it obviously does not rely on discerning the past actions of single players). The sequence of joint actions has the limit of $(1,1,\dots)$ (in the product topology), which is not a one-shot Nash equilibrium.

Example 1 illustrates the need for continuity assumptions on payoffs in the stage game. The next example, adapted from one in Schmeidler (1973), illustrates the need for a compactness assumption on the set of joint actions in the stage game.

Example 2. In the stage game, I is $[0,1]$; A_i is again $\{0,1\}$ for each $i\in I$; the joint action set A is the set of (equivalence classes of Lebesgue-) measurable functions from $[0,1]$ to $\{0,1\}$; and

$$h_i(a_i,a)=\begin{cases}\left|a_i-(1/i)\int_{[0,i]}a\,\mathrm{d}\lambda\right| & \text{if } i\neq 0\\ 0 & \text{if } i=0,\end{cases}$$

where λ is Lebesgue measure. As Schmeidler (1973) shows, this game has no equilibria. Consider the following strategy combination in the repeated game. At each time t, partition the player set into 2^t half-open subintervals of equal size. Almost all players in the odd subintervals play 0 while almost all those in the even subintervals play 1, all independently of the history. For almost all i, the sequence of stage-game payoffs converges to 1/2; hence no non-null set of players can deviate profitably, and the strategy combination must be an equilibrium of the repeated game. The sequence of plays converges (L_1-weak*) to the (equivalence class of the) constant function 1/2, which is not in the (pure) joint-action space for the one-shot game. Recalling Theorem 2 there can be no topology on A such that both A is sequentially compact and the h_i satisfy the required continuity assumptions. On the other hand, if A_i is extended to all of $[0,1]$ and h_i is extended linearly, then h_i satisfies the continuity hypotheses and, by Alaoglu's Theorem, A is L_1-weak* compact. As the L_1-weak* topology is first countable, A is sequentially compact; hence the hypotheses of Theorem 2 are satisfied.

The final example illustrates that the hypotheses of Theorem 2 admit situations in which both the stage game and the repeated game do not possess equilibria.

Example 3. The stage game is two-player 'matching-pennies' with only pure actions allowed (hence no equilibria). In the pure-strategy repeated game in which neither player is allowed to condition his action at any time on any of his opponent's past actions, the only pure strategies are simply sequences of pure actions, to any of which there is a response sequence which holds the player to his minimum payoff at each stage. It follows that the repeated game possesses no equilibria.

4. Proof of Theorem 2

(1) $f \in F^*$ *implies* $X(f) \subset (A^* \cup \underline{Z}(f))$.

Suppose $f \in F^*$ and $\exists B \subset N$ s.t. $\{a^t(f)\}_{t\in B} \to a \notin A^*$. Then $\exists \bar{I} \in S$ with $\lambda(\bar{I}) > 0$ s.t. for a.e. $i \in \bar{I}: \exists b_i \in A_i$ with $h_i(b_i, a) - h_i(a(i), a) = \delta_i > 0$. By the continuity hypothesis, for a.e. $i \in \bar{I}$, $\exists T_i$ s.t. $\forall t \geqq T_i$, $t \in B$,

$$|h_i(a^t(f)(i), a^t(f)) - h_i(a(i), a)| < \delta_i/3 \tag{1}$$

and

$$|h_i(b_i, a) - h_i(b_i, a^t(f))| < \delta_i/3. \tag{2}$$

Define $g_i \in F_i$ (a.e. $i \in \bar{I}$) as follows:

$$g_i^t(\cdot) \equiv \begin{cases} a^t(f)(i) & \text{if} \quad t \notin B \text{ or } t < T_i \\ b_i & \text{if} \quad t \in B \text{ and } t \geqq T_i. \end{cases}$$

Then,

$$H_i(f) - H_i(f|g_i)$$

$$\leqq \limsup_{T\to\infty} T^{-1} \sum_{t=1}^{T} [h_i(a^t(f)(i), a^t(f)) - h_i(a^t(f|g_i)(i), a^t(f|g_i))]$$

$$\leqq \limsup_{T\to\infty} T^{-1} \sum_{\substack{t\in B(T)\\ t\geqq T_i}} [h_i(a^t(f)(i), a^t(f)) - h_i(b_i, a^t(f))] \tag{3}$$

by the definition of g_i, boundedness of h_i, and (iii). Now, by conditions (1) and (2), (3) is less than or equal to

$$\limsup_{T\to\infty} T^{-1} \sum_{\substack{t\in B(T)\\ t\geqq T_i}} [h_i(a(i), a) - h_i(b_i, a) + 2\delta_i/3]$$

$$=\limsup_{T\to\infty} T^{-1} \sum_{\substack{t\in B(T)\\ t\geqq T_i}} (-\delta_i/3) \leqq (-\delta_i/3) \liminf_{T\to\infty} \#B(T)/T. \tag{4}$$

Hence, if $\liminf_{T\to\infty} \#B(T)/T>0$, (4) contradicts the hypothesis that $f\in F^*$.

(2) $X(f)\subset(A^*\cup\bar{Z}(f))$ implies $f\in F^*$.

Let $f\in F$ be s.t. $X(f)\subset(A^*\cup\bar{Z}(f))$. In order to show that $f\in F^*$, we must show that for a.e. $i\in I$ and all $g_i\in F_i$, i does not gain by employing g_i. Accordingly, it is sufficient to limit our attention to

$$B_i=\{t\in N: h_i(a^t(f\,|\,g_i)(i), a^t(f\,|\,g_i))>h_i(a^t(f)(i), a^t(f))\}.$$

Next, we consider the (countable, at most) subsets $\bar{Z}(f)$ and $Y(f)\equiv (X(f)\cap A^*)-\bar{Z}(f)$ of limit points of converging subsequences of $\{a^t(f)\}_{t\in N}$. Let σ_k be the kth element of $Y(f)$ and z_ℓ be the ℓth element of $\bar{Z}(f)$. Associated with each σ_k and z_ℓ are the time indices B_k and C_ℓ of the subsequences converging to σ_k and z_ℓ, respectively. Without loss of generality, we may assume that the families $\{B_k\}$ and $\{C_\ell\}$ are both pairwise disjoint.

Now,

$$\begin{aligned}
&H_i(f\,|\,g_i)-H_i(f)\\
&\leqq \limsup_{T\to\infty} T^{-1}\sum_{t=1}^{T}[h_i(a^t(f\,|\,g_i)(i), a^t(f\,|\,g_i))-h_i(a^t(f)(i), a^t(f))]\\
&\leqq \limsup_{T\to\infty} T^{-1}\Bigg\{\sum_{\ell}\sum_{t\in C_\ell(T)\cap B_i}[h_i(a^t(f\,|\,g_i)(i), a^t(f\,|\,g_i))\\
&\qquad\qquad -h_i(a^t(f)(i), a^t(f))]\\
&\qquad +\sum_{k}\sum_{t\in B_k(T)\cap B_i}[h_i(a^t(f\,|\,g_i)(i), a^t(f\,|\,g_i))\\
&\qquad\qquad -h_i(a^t(f)(i), a^t(f))]\Bigg\}\equiv\Delta
\end{aligned}$$

since the remaining terms are non-positive. For any $\varepsilon>0$, let T_{ik} be s.t. $\forall\, t\geqq T_{ik}$, $t\in B_k$

$$\left|h_i(\sigma_k(i), \sigma_k) - h_i(a^t(f)(i), a^t(f))\right| < \varepsilon/2 \tag{5}$$

and

$$\left|h_i(a^t(f\,|g_i)(i), a^t(f)) - h_i(a^t(f\,|g_i)(i), \sigma_k)\right| < \varepsilon/2. \tag{6}$$

Let

$$y_i = \sup\{\left|h_i(a_i, a) - h_i(b_i, b)\right| : a_i, b_i \in A_i;\, a, b \in A\}.$$

Then

$$\begin{aligned}
\Delta &\leqq \limsup_{T\to\infty} T^{-1}\Bigg\{\sum_{\ell}\sum_{t\in C_\ell(T)\cap B_i} y_i + \sum_k \min\{\#B_k(T), T_{ik}\}y_i \\
&\qquad + \sum_k \sum_{\substack{t\in B_k(T)\cap B_i\\ t\geqq T_{ik}}} [h_i(a^t(f\,|g_i)(i), a^t(f)) - h_i(a^t(f)(i), a^t(f))]\Bigg\} \\
&\leqq \sum_\ell \limsup_{T\to\infty} T^{-1}\,\#C_\ell(T)y_i + \sum_k \limsup_{T\to\infty} T^{-1}T_{ik}y_i \\
&\quad + \limsup_{T\to\infty} T^{-1}\sum_k \sum_{\substack{t\in B_k(T)\cap B_i\\ t\geqq T_{ik}}} [h_i(a^t(f\,|g_i)(i), \sigma_k) - h_i(\sigma_k(i), \sigma_k) + \varepsilon]
\end{aligned}$$

from (5) and (6). But since $\sigma_k \in A^*$ and $\limsup_{T\to\infty} \#C_\ell(T)/T = 0$, this expression is bounded above by

$$\limsup_{T\to\infty} T^{-1}\sum_k \#B_k(T)\varepsilon \leqq \varepsilon.$$

Since ε was arbitrary, this implies $f \in F^*$.

References

Aumann, Robert J., 1981, Survey of repeated games, in: Essays in game theory and mathematical economics in honor of Oskar Morgenstern (Bibliographisches Institut, Mannheim) 11–42.

Dubey, Pradeep and Mamoru Kaneko, 1984, Information patterns and Nash equilibria in extensive games: I, Mathematical Social Sciences 8, 111–139.

Kaneko, Mamoru, 1982, Some remarks on the folk theorem in game theory, Mathematical Social Sciences 3, 281–290; also Erratum in 5, 1983, 233.

Peleg, Bezalel, 1969, Equilibrium points for games with infinitely many players, Journal of the London Mathematical Society 44, 292–294.

Rubinstein, Ariel, 1979, Equilibrium in supergames with the overtaking criterion, Journal of Economic Theory 21, 1–9.

Schmeidler, David, 1973, Equilibrium points of nonatomic games, Journal of Statistical Physics 7, 295–300.