B24 E 2-3

BEBR

FACULTY WORKING
PAPER NO. 882

Perfection in Repeated Two-person Nonzero-sum
Games of Asymmetric Information

Jonathan A. K. Cave

BEBR

Perfection in Repeated Two-person Nonzero-sum
Games of Asymmetric Information

Jonathan A. K. Cave, Assistant Professor
Department of Economics

Abstract

S. Hart has recently obtained an analogue of the Folk Theorem for two-player non-zero-sum games where one player is informed of the true game and the other is not. In this note, we begin the proof of a perfectness theorem for the same set-up, presenting a sufficient condition for perfect equilibrium. In Hart's case, the limit of individual rationality for the uniformed player is the greatest convex function dominated by the value of the zero-sum game given by player II's payoffs; in our case it is the greatest quasi-convex function dominated by the value of the "expected game".

Perfection in Repeated Two-person Nonzero-sum Games of Asymmetric Information

This note reports some preliminary results towards the proof of a perfectness theorem for games of asymmetric information. It uses the model developed by Hart [1981] in his recent characterization of the set of equilibria of such games.

In this model, Nature chooses an integer, k, from a finite set K, according to a known probability distribution $p \in \Delta^K$. Player I is informed of k, while II is not. Corresponding to k are two nxm matrices, A^k, representing player I's payoffs, and B^k, representing the payoffs to player II. The game is played repeatedly, with both players using behavioral strategies. The generic strategy for player I is $\sigma:K \to F_1$, and we will write $\sigma(k) = \sigma$, when no confusion will result. At each play of the game, players employ mixed moves $\sigma_1(h) \in \Delta^n$ for player I and $\tau_2(h) \in \Delta^m$ for player II, and where h is the history of realizations of the moves of the players up to time t. The generic strategy for player II is $\tau \in F_2$, and does not depend on k.

We define cumulative average payoffs by noticing that the result of strategies σ and τ is a sequence of random variables taking values (i,j), i = 1,...,n; j=1,...,m. Letting $\underset{\sim}{a}_t^k(\sigma,\tau)$ denote the random member of the matrix A^k selected at time t, we define

$$\overline{\underset{\sim}{a}}_T^k = \frac{1}{T} \underset{t \leq T}{\Sigma} \underset{\sim}{a}_t^k(\sigma,\tau)$$

$$\overline{\underset{\sim}{b}}_T = \underset{k \in K}{\Sigma} p(k) \left[\underset{t \leq T}{\Sigma} \frac{1}{T} \underset{\sim}{b}_t^k \right]$$

Thus $\overline{a}_T^k$ represents player I's conditional (on k) expected (over σ,τ)
average payoff up to time T, while $\overline{b}_T$ is player II's expected (over
σ, τ, and k) average payoff to time T. We now define equilibrium.

1.1 <u>Definition</u>: The pair $(\sigma*, \tau*)$ is an <u>equilibrium</u> iff, for all
σ', τ';

    i) $\lim_{T} \inf E[\overline{a}_T^k:k, \sigma*, \tau*] \geq \lim_{T} \sup E[\overline{a}_T^k:k, \sigma', \tau*]$ for all k, and

    ii) $\lim_{T} \inf E[\overline{b}_T:\sigma*, \tau*] \geq \lim_{T} \sup E[\overline{b}_T:\sigma*, \tau']$

Setting $\sigma* = \sigma'$ in i) and $\tau* = \tau'$ in ii), we observe that there exists
a vector $(a,b) \in R^K x\ R$ s.t.

$$\lim_{T} E[\overline{a}_T^k: k, \sigma*, \tau*] = a_k \text{ and } \lim_{T} E[\overline{b}_T: \sigma*, \tau*] = b$$

So that equilibrium sequences are convergent sequences. The Folk
Theorem for the undiscounted, complete information supergame states
that any feasible, individually-rational payoff of the stage game can
be obtained as the outcome of an equilibrium of the supergame, and
conversely. The Perfectness Theorem for the same model states that
the set of outcomes of equilibria and of perfect equilibria coincide.

Hart has shown that the Folk Theorem goes through, when we make
the appropriate translations of the concepts of "feasibility" and
"individual rationality". We begin with the latter.

1.2 <u>Definition</u>: let $(a,b) \in R^K x\ R$. (a,b) is <u>individually-rational</u>
<u>for player I</u> iff, for every $p \in \Delta^K$,

(1) $pa \geq val\ A(p)$

where $A(p)$ is the "expected game" for player I: an nxm matrix whose entries are

$$A(p)(i,j) = \sum_{k \in K} p(k) A^k(i,j)$$

Defining $B(p)$ similarly, and letting val and vex( ) have their usual interpretations as the value and the greatest convex function everywhere less than or equal to ( ), respectively, we say that (a,b) is individually-rational for II at p^* $\in \Delta^K$ iff

(2) $b \geq$ vex val $B(p^*)$

(σ,τ) is said to be a uniform equilibrium point if, $\forall \sigma', \tau'$,

(3) $\lim \inf E[\overline{a}_T^k: k, \sigma, \tau] \geq \lim \sup [\sup_{\sigma'} E[\overline{a}_T^k: k, \sigma', \tau]$ for all k, and

(4) $\lim \inf E[\overline{b}_T: \sigma,\tau] \geq \lim \sup [\sup_{\tau'} E[\overline{b}_T: \sigma,\tau']]$

By reversing the orders of the sup and lim sup it can be seen that (3) and (4) imply i) and ii) of definition 1.1, so that every uniform e.p. is an equilibrium point (e.p.). Hart has shown that the reverse implication holds as well.

As a first approach to feasibility, we know that the outcomes of stationary e.p.'s must be feasible: if conditions i) and ii) are satisfied for $a = A^k(i,j): k \in K$, and $b = A(p^*)(i,j)$; then (a,b) are feasible. By a slight extension, it can be seen that any convex combination of these stationary, nonrevealing equilibria is feasible.

Now let $F = CH \{[[a,b](i,j)]\}$: this is the convex hull of the pure-strategy payoffs, considered as a subset of R^{2K}. We can bound this set by the following device: let $\overrightarrow{\alpha}^k = \max_{i,j} A^k(i,j)$, $\overrightarrow{\beta}^k = \max_{i,j} B^k(i,j)$, $\underline{\alpha}^k = \min_{i,j} A^k(i,j)$, and $\underline{\beta}^k = \min_{i,j} B^k(i,j)$. Now define

$R_1^K = \{x \in R^K: \underline{\alpha} \leq x \leq \overline{\alpha}\}$; $R_2^K = \{x \in R^K: \underline{\beta} \leq x \leq \overline{\beta}\}$, $\overline{\overline{\beta}} = \max_k \overrightarrow{\beta}^k$, and

$\underline{\underline{\beta}} = \min_k \underline{\beta}^k$. Then $R_2 = \{x \in R: \underline{\underline{\beta}} \leq x \leq \overline{\overline{\beta}}\}$. We now define the set of feasible and individually-rational actual average outcomes.

1.3 <u>Definition</u>: $G = \{(a,b,p) \in R_1^K \times R_2 \times \Delta^K:$ there exist $(c,d) \in F$ with

    i) $a \geq c$ and $pa = pc$

   ii) for all $q \in \Delta^K$, $qa \geq val A(q)$

  iii) $b \geq vex val B(p)\}$.

G is the set of payoffs corresponding to non-revealing equilibria: ii) and iv) are individual-rationality conditions, while i) and iii) are "representation" conditions: in i) we see that the vector a agrees with the vector c on all games k with positive probability, while iii) guarantees that there is some collection of payoffs in the true games with expectation b.

When we turn to the study of equilibria in which there is some transmission of information, the notion of feasibility becomes more complicated. In their pioneering work on this model, Aumann, Maschler and Stearns [1968] developed a class of equilibria involving transmission of information from the informed to the uninformed player, followed by completely non-revealing play from that point on. In essence, what

Hart did was to extend this to obtain certain randomized collections of such nonrevealing "plans". This extends the set G since the posterior distribution after communication may differ from the prior.

However, such a process of communication must always end in G. This motivates the following definition:

1.4 <u>Definition</u>: let $g \in R_1^K \times R_2 \times \Delta^K$. A sequence $[g_n : n = 1,\ldots,]$ of $R_1^K \times R_2 \times \Delta^K$ - valued random variables is called a <u>G-sequence</u> <u>starting at g</u> iff

    i) $g_1 = g$ a.s.;

    ii) there exists a nondecreasing sequence $[F_n : n = 1,\ldots]$ of finite fields s.t. g_n is a martingale w.r.t. F_n;

    iii) if $g_n \xrightarrow{a.s.} g^*$, then $g^* \in G$; and

    iv) for each n <u>either</u> $a_{n+1} = a_n$ a.s. or $p_{n+1} = p_n$ a.s.

Property ii) says that g_n is F_n-measurable and that $E(g_{n+1} : F_n) = g_n$ a.s. for all n. Combining this with i) tells us that $E(g_n) = g$ a.s. for all n. Since the space $R_1^K \times R_2 \times \Delta^K$ is compact, the limit g^* in iii) exists. Thus property iii tells us that g^* satisfies conditions 1.3.i-iv a.s. Property iv is called the bi-martingale property; at every step either player I's (vector-valued) payoff or player II's posterior must remain fixed, while the other may change in such a manner that the conditional expectation is unaffected by iii).

1.5 <u>Definition</u>: $G^* = \{(a,b,p) \in R_1^K \times R_2 \times \Delta^K : $ there exists a G-process starting at $(a,b,p)\}$.

We remark that in the absence of conditions i) and iv) we would have $G* = CH[G]$. Hart's Theorem is that $G*$ is the Graph of the Equilibrium Correspondence.

II. Perfection for Nonrevealing Equilibria

Let us consider a triple (a,b,p) in G. We know that the grim punishments can be used to hold the players to their individual rationality levels, but that these grim punishments are not worth carrying out in the event. In the complete information game, this problem is solved by punishing defectors for a long but finite period of time, after which play is to return to the cooperative sequence. Defection at any stage of this process, either on the part of the original defector or the punishers, is met with renewed punishment. In the limit, no player can hope to profit by defection: either he defects for at most a finite number of periods, and ultimately returns to an infinite play of the cooperative sequence, or he defects forever, and is punished forever, providing him with his minmax payoff. Unfortunately, even the equilibrium strategies here have some revelation; to hold player II to vex val $B(p*)$, player I must reveal some information, unless vex val $B(p*) = $ val $B(p*)$. In the equilibrium situation this causes no problem, as the nonrevealing punishments associated with each possible posterior of player II are balanced in such a way that player II's expected minmax level is vex val $B(p*)$, but this is inconsistent with perfection.

The reason is simply that, if the nonrevealing strategy calls for $(a,b,p*)$, where $b < $ val $B(p*)$, player II will respond to a promise of long but finite punishment by immediate defection: after punishment

begins, player II will revise his posterior to some $p' \neq p*$, and will begin to respond optimally in $B(p')$. If it happens that val $B(p') > b$, then player II has no incentive to return to the cooperative sequence. In this section, we make the simple point that if this phenomenon cannot arise, we can still construct a perfect equilibrium. In other words, if the punishment can be arranged so that player II never forms a posterior p' with the property that val $B(p') > b$, then we can construct long-but-finite punishment sequences for player II. Note also that there is no problem in policing the actions of player I.

2.1 <u>Definition</u>: Let a function H: $\Delta^K \rightarrow R_2$ be defined by

$$H(p) = \inf \{b \in R_2: \text{ there exist } p_t \in \Delta^K, t = 1,\ldots,K+1, \text{ and } q \in \Delta^{K+1} \text{ s.t.}$$

i) val $B(p_t) \leq b$ for all t; and

ii) $\sum_{t=1}^{K+1} q_t p_t = p$

We then define an analogue of G for the perfect nonrevealing case by:

2.2 <u>Definition</u>: $G_p = \{(a,b,p) \in R_1^K \times R_2 \times \Delta^K: \text{ there exist } (c,d) \in F$ with

    i) $a \geq c$ and $pa = pc$;

    ii) for all $q \in \Delta^K$, $qa \geq$ val $A(q)$;

    iii) $b = pd$; and

    iv) $b \geq H(p)\}$

It will be noted that G_p differs from G only in the individual rationality condition for player II, which has been strengthened. Analogously with G*, we can define a G_p-process starting at point g, and set

G_p^* = the set of points g s.t. there is a G_p-process starting with g.
We can then demonstrate:

2.3 <u>Theorem</u>: If $(a,b,p) \in G_p^*$, then there is a perfect equilibrium with
payoff (a,b).

The nature of the strategies used in these "essentially nonrevealing"
equilibria is as follows: from the initial point $(a,b,p) \in G_p^*$, the
players follow the Hart strategies, including communications [signalling
and jointly controlled lotteries] and payoff accumulation periods, fol-
lowed by an eventual play of nonrevealing equilibrium in G_p. There is
no need to punish players during the communications periods, as Hart
shows. During the other periods, we make use of the lemma that Hart
calls Proposition 3.18: if (a_n, b_n, p_n) is an $R_1^K \times R_2 \times \Delta^K$-valued
martingale converging a.s. to (a, b, p) then a satisfies (1) a.s. iff
a_n satisfies (1) a.s. for all n, and (b,p) satisfies (2) a.s. iff
(b_n, p_n) satisfies (2) a.s. for all n. In other words, the elements of
the sequences used in the G_p-process are individually-rational. In
our case we need the following

2.4 <u>Lemma</u>: if (a_n, b_n, p_n) is a martingale convering a.s. to (a, b, p),
then (b, p) satisfies 2.2 iv a.s. if (b_n, p_n) satisfies 2.2 iv a.s. for
all n.

<u>Proof</u>:   noting that H( ) is a continuous function, take lim.
                                                                    $n \to \infty$
Unfortunately, the "only if" part of this lemma is untrue. Hart's proof
works for G-processes, and uses Jensen's inequality and thus the con-
vexity of vex val B. In our case, H(p) is not convex. However;

2.5 <u>Lemma</u>: $H(p)$ is the greatest quasiconvex function bounded above by val $B(p)$. This result follows from Definition 2.1.

Now we can sidestep the problem by a slight restriction on the sequence of fields that makes the G_p-process into a martingale. Recall that the reason we needed Lemma 2.4 was to guarantee that the threatened punishments would be sufficient during the payoff-accumulation periods, played with expected payoffs to player II of b_n, and a posterior distribution of p_n. Thus, we need to have inequality 2.2.iv satisfied for all n.

2.6 <u>Definition</u>: $D = \{p \in \Delta^K: H(p) \neq \text{val } B(p)\}$

Clearly, D can be represented uniquely as a collection of disjoint open sets. We must choose the field that we use in the construction of a G_p-process in such a way that the part of the fields that changes can be made finite. This requirement ensures that the strong convergence results we need go through, and also that the fields can be generated by partial (finite) histories of play using player I's strategy. Before defining the fields we will use, we must therefore assure ourselves that this finiteness condition can be met.

2.7 <u>Lemma</u>: D is the union of finitely many disjoint open sets.
Proof: the proof is in two parts. First, we show that Val $B(p)$ is a piecewise monotonic function of p. Then, we show that the quasi-convexification of a piecewise monotonic function differs from the original function on a finite number of disjoint open sets.

First, consider the definition of val $B(p)$.

$$\text{val } B(p) = \max_{\mu_1} \min_{\mu_2} \sum_{k=1}^{K} p(k) \sum_{i=1}^{n} \sum_{j=1}^{m} \mu_1(i)\mu_2(j)b_{ij}^k$$

so that val $B(p)$ is the minimum of finitely-many functions of μ_1 and p. Define

$$W(\mu_1,p,j) = \sum_{k=1}^{K} p(k) \sum_{i=1}^{n} \mu_1(i)b_{ij}^k$$

and let $W^*(\mu_1,p) = \min_{j} W(\mu_1,p,j)$. Then each $W(\mu_1,p,j)$ is a linear function of p, and $W^*(\mu_1,p)$ is the minimum of finitely-many such linear functions (M of them). It is therefore continuous and piece-wise monotonic, in both arguments. Since the μ_1 are drawn from the compact set Δ^K, we are assured of the continuity of val $B(p)$ in p. Piecewise monotonicity means that there are a finite number of regions within each of which W^* is monotonic in each of the M variables $\mu_1(i)$ and the K (or K-1) variables $p(k)$. Since these regions exhaust the domain $\Delta^M \times \Delta^K$, it follows that the maximizing choice of μ_1 will always lie on a boundary of such a region. Projecting these boundaries onto Δ^K we obtain the desired result.

Now suppose that $W(p)$ is a piecewise monotonic continuous func-tion, and let $w\#(p)$ be its "quasi-convexification": the greatest quasi-convex function everywhere less than or equal to $W(p)$. Consider first the case where p is a scalar (this is the case K = 2). It is easy to see that the condition

$$w\#(p) \neq W(p)$$

holds only between minima of W. In the case where all such local minima are isolated, the number of such regions (connected intervals

bounded on one side by a local minimum, and on the other side by the closest p s.t. W(p) = the value of W at the local minimum) is bounded above by the number of such local minima + 2 (for the boundaries of Δ^2). If there is a continuum (interval) of such local minima, we need consider only the extreme ones. It is also obvious that if the domain of W is divided into finitely-many intervals on each of which W is monotone, there will be at most finitely-many such extreme or isolated local minima. Thus, the set D consists of finitely many disjoint open sets, on which H(p) is equal to the value of W at the local minimum attained at the boundary of the set.

Now consider the K-dimensional case, for $K \geq 3$. In this case, we cannot say that H(p) is constant on each subset of D. However, for each p in D there will be a line segment crossing the subset of D containing p along which H is constant. In other words, if P is in D, there exists a unique maximal connected open subset of D containing p. Call this E(p). Then there is p' in the boundary of E(p) s.t. val B(p') = H(p). Moreover, if A(p) denotes the line containing p and p', then for all $p'' \in A(p) \cap E(p)$ we have H(p") = H(p) = val B(p'). Along the direction of A(p), the point p' must be a local minimum. With respect to this line, the previous result goes through. It follows that, fixing any "direction" (linear submanifold of Δ^K), the previous result applies. Therefore the image of D projected to any lower-dimensional face of the simplex is a finite union of disjoint open sets, so D must be such a finite union.   QED

Remark: another way to formulate this is directly by induction on the dimension K. We have seen that the result is true for K = 2. Now

suppose that $K = 3$; we have seen that for every line (of dimension 1) through the 3-simplex (which is of dimension 2), the set obtained as the intersection of the line with D is finite union of open sets; therefore D is finite. The induction set is:

2.8 <u>Lemma</u>: Let D be a union of disjoint open sets in Δ^K. D is finite iff for every $K-2$ dimensional linear manifold M in R^K, $D \cap M$ is a finite union of disjoint open sets.
[proof is obvious].

2.9 <u>Definition</u>: Let F be a field. If F is associated with an $R_1^K \times R_2 \times \Delta^K$ - valued random variable, we say that F is <u>admissibly</u> <u>finite</u> iff:

    i) The partition on Δ^K induced by F always refines the partition of the simplex into $\Delta^K - D$ and the partition of D into finitely many disjoint open sets as described above; and

    ii) the restriction of the partition induced by F to D is finite.

2.10 <u>Definition</u>: a $\overline{G}_p$-<u>process</u> <u>starting</u> <u>at</u> <u>g</u> is a G_p-process starting at g where the fields F_n of 1.4.ii (or the analogous G_p condition) are all admissibly finite.

2.11 <u>Lemma</u>: if $[a_n, b_n, p_n]$ is a $\overline{G}_p$-process terminating at $[a,b,p]$ (in the sense of a.s. convergence) then $b \geq H(p)$ iff $b_n \geq H(p_n)$ for all n.
proof: $b_n = E[b:F_n] \geq E[H(p):F_n]$. Fix any event $A \in F_1$, and denote by $H_A(\ )$ the corestriction of H to $[a_n, b_n, p_n]$ that have positive probability density conditional on A. Notice that $H_A(\ )$ is convex. Therefore

$$E[H_A(p):F_n] \geq H_A(E[p:F_n]) = H_A(p_n) \qquad \text{QED}$$

We remark that this result does not go through if either of the players has an infinite number of moves or if the number of games is infinite, although it should be possible to show that the main theorem is still valid with an infinite (countable) number of finite-move games.

2.12 <u>Definition</u>: $\overline{G}^*_p = \{g \in G^*_p: \text{ there is a } \overline{G}_p\text{-process starting at g}\}$

<u>Remark</u>: $\overline{G}^*_p = G^*_p$.

As a result of this construction, Hart's proof can be adapted directly to prove Theorem 2.3, which tells us that in a perfect equilibrium it is sufficient to have player II (the uniformed player) receive at least $H(p)$.

<u>References</u>:

Aumann, R. J., Maschler, M. and Stearns, R. Repeated Games of Incomplete
    Information: the Zero-sum Extensive Case; Mathematica, Inc. 1955.

Hart, S. Nonzero-sum Two-player Repeated Games with Incomplete Information, mimeo.

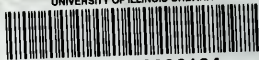<u>Acknowledgements</u>: