Int J Game Theory (2000) 29:359–364

International Journal of Game Theory

Sequence of opponents and reduced strategies

Jean-Pierre Beaud[1] **and Sylvain Sorin**[1,2]

Received: March 1999/Revised: May 2000

Abstract. We consider the framework of repeated two-person zero-sum games with lack of information on one side. We compare the equilibrium payoffs of the informed player in two cases: where he is facing either a) a single long-lived uninformed player, or b) a sequence of short-lived uninformed players. We show: 1) that situation b) is always (weakly) better than a), 2) that it can be strictly better in some cases, 3) that the two cases are equivalent if the long uninformed player has an optimal strategy independent of his own moves.

Key words: incomplete information, sequence of opponents

1. Presentation

We consider a two person zero-sum repeated game with incomplete information on one side ([1]). There is a finite state space K. For each $k \in K$, a $I \times J$ real payoff matrix A^k is given, as well as two $I \times J$ signalling matrices $H^{1,k}$ and $H^{2,k}$ with values in finite signal sets S and T.

The game G is played as follows: k is selected according to a publicly known distribution p. Player 1 (the maximizer) is informed on the outcome but not Player 2. Inductively at stage m, Player 1 chooses an action i_m in I and Player 2 chooses an action j_m in J. The stage payoff is $x_m = A^k_{i_m j_m}$ but is not announced. Rather the signal $s_m = H^{1,k}_{i_m j_m}$ (resp. $t_m = H^{2,k}_{i_m j_m}$) is told to Player 1 (resp. Player 2). We assume that for each player his signal contains

[1] Laboratoire d'Econométrie, Ecole Polytechnique, 1 rue Descartes, 75005 Paris, France ⟨*e-mail: beaud@poly.polytechnique.fr*⟩
[2] MODALX and THEMA, UFR SEGMI, Université Paris X, 200 Avenue de la République, 92001 Nanterre, France ⟨*e-mail: sorin@poly.polytechnique.fr*⟩

his move. A behavioral strategy for Player 1 is described by a sequence $\sigma = \{\sigma_m\}$ of mappings from $K \times S^{m-1}$ to the set of mixed actions $\Delta(I)$. Similarly, but taking into account his lack of information, a behavioral strategy of Player 2 is a sequence $\tau = \{\tau_m\}$ of mappings from past signals T^{m-1} to $\Delta(J)$. Denote by Σ and $\mathcal{T}$ the corresponding sets. Any triple (p, σ, τ) in $\Delta(K) \times \Sigma \times \mathcal{T}$ induces a probability distribution on the set of plays $K \times (S \times T)^{\infty}$ and $E_{p,\sigma,\tau}$ denotes the induced expectation.

The game Γ is played the same way as the game G but with a different Player 2 at every stage. However each of these players knows the history of signals sent to his predecessors. We call $2.m$ the Player 2 active at stage m, hence his strategy's set is also described by a mapping θ_m from T^{m-1} to $\Delta(J)$.

For each class of games we consider the finitely repeated and discounted version:

$G_n(p)$ is the n stage two person zero sum game with payoff for Player 1 given by $g_n^1(p)(\sigma, \tau) = E_{p,\sigma,\tau}\left(\frac{1}{n}\sum_{m=1}^{n} x_m\right)$. $G_n(p)$ has a value $v_n(p)$.

$G_\lambda(p)$ is the λ-discounted two person zero sum game with payoff for Player 1 given by $g_\lambda^1(p)(\sigma, \tau) = E_{p,\sigma,\tau}(\lambda \sum_{m=1}^{\infty} (1-\lambda)^{m-1} x_m)$. $G_\lambda(p)$ has a value $v_\lambda(p)$.

$\Gamma_n(p)$ is a $(n+1)$ player game with payoff $g_n^1(p)(\sigma, \theta) = g_n^1(p)(\sigma, \theta_1, \ldots, \theta_n)$ for Player 1 and $\gamma^{2,m}(p)(\sigma, \theta) = E_{p,\sigma,\theta}(-x_m)$ for Player $2.m$, $m = 1, \ldots, n$. (Note that this last payoff depends only upon $(\sigma_\ell, \theta_\ell)$ for $\ell \leq m$). $\mathcal{E}_n(p)$ denotes the set of equilibrium payoffs of Player 1 in $\Gamma_n(p)$.

$\Gamma_\lambda(p)$ is a game with an infinite family of Player 2. The payoff is $g_\lambda^1(p)(\sigma, \theta)$ for Player 1 and again $\gamma^{2,m}(p)(\sigma, \theta) = E_{p,\sigma,\theta}(-x_m)$ for Player $2.m$, $m \geq 1$. The existence of equilibria in $\Gamma_\lambda(p)$ follows from standard approximation by truncated games with finitely many Player 2. Let $\mathcal{E}_\lambda(p)$ denote the corresponding set of equilibrium payoffs for Player 1.

The purpose of this note is to compare $v_n(p)$ and $\mathcal{E}_n(p)$ (respectively $v_\lambda(p)$ and $\mathcal{E}_\lambda(p)$), namely to examine whether it is profitable or not for the informed player to face a single long lived opponent or a sequence of short lived opponents.

2. A first result

Player 1 is always better off facing a sequence of short lived opponents since we have:

Proposition 1. $\forall p, \forall n \geq 1, \forall y \in \mathcal{E}_n(p)$:

$$y \geq v_n(p)$$

Similarly $\forall p, \forall \lambda \in (0,1), \forall z \in \mathcal{E}_\lambda(p)$:

$$z \geq v_\lambda(p)$$

Proof: The proof is the same in both cases.

Player 1 faces the same set of strategies $\mathcal{T}$ in G and Γ. Thus the use by Player 1 in $\Gamma_n(p)$ or $\Gamma_\lambda(p)$ of an optimal strategy in $G_n(p)$ or $G_\lambda(p)$ implies the result. ■

3. Reduced strategies

Definition. τ is a reduced strategy of player 2 if it depends upon his signals only through the moves of Player 1.

In other words τ_m is measurable with respect to the algebra generated by $(i_1,\dots,i_{m-1})$. Explicitely: say that $(i_1,\dots,i_{m-1})$ is compatible with $(t_1,\dots,t_{m-1})$ if there exists k with $t_\ell = H^{2,k}_{i_\ell, j_\ell}$ for $\ell = 1,\dots,m-1$, (recall that t_ℓ determines j_ℓ). Now if some $(i_1,\dots,i_{m-1})$ is compatible both with $(t_1,\dots,t_{m-1})$ and $(t'_1,\dots,t'_{m-1})$ then $\tau_m(t_1,\dots,t_{m-1}) = \tau_m(t'_1,\dots,t'_{m-1})$.

Theorem 2. *If Player 2 has a reduced optimal strategy in $G_n(p)$ (resp. $G_\lambda(p)$), then $\mathscr{E}_n(p)$ is reduced to $\{v_n(p)\}$ (resp. $\mathscr{E}_\lambda(p)$ is reduced to $\{v_\lambda(p)\}$.*

Proof: Let (σ,θ) be an equilibrium in $\Gamma_n(p)$. Let $\tau^\star$ be a reduced optimal strategy of Player 2 in $G_n(p)$. The equilibrium condition for Player 2.m gives:

$$\gamma^{2,m}(\sigma,\theta) \geq \gamma^{2,m}(\sigma,\theta_{-m},\tau^\star_m)$$

The idea is to define inductively a fictitious strategy of Player 1, $\tilde\sigma$, as a function of σ and θ such that, for all $m = 1,\dots,n$:

$$\gamma^{2,m}(\sigma,\theta_{-m},\tau^\star_m) = \gamma^{2,m}(\tilde\sigma,\tau^\star).$$

$\tilde\sigma$ is a sequence of maps $\tilde\sigma_m$ from $K \times I^{m-1}$ to $\Delta(I)$ (hence a strategy of Player 1 independent of the moves of Player 2) defined inductively as follows:

$$\tilde\sigma^k_1 = \sigma^k_1$$

$$\tilde\sigma^k_2(i_1) = \sum_{s_1 \in S,\, j_1 \in J} \theta_1(j_1)\mathbf{1}_{\{H^{1,k}(i_1,j_1)=s_1\}}\sigma^k_2(s_1)$$

In words, Player 1 is playing at stage 2 the expectation of his initial strategy with respect to the distribution on signals induced by θ_1. Similarly, at stage m, $\tilde\sigma_m$ is defined by taking the conditional expectation on the signals of Player 1, given his moves:

$$\begin{aligned}\tilde\sigma^k_m(i_1,\dots,i_{m-1}) &= E(\sigma^k_m(s_1,\dots,s_{m-1}) \mid i_1,\dots,i_{m-1})\\ &= \sum_{s_1,\dots,s_{m-1}} P^k(s_1,\dots,s_{m-1} \mid i_1,\dots,i_{m-1})\sigma^k_m(s_1,\dots,s_{m-1})\end{aligned}$$

where $P^k(s_1,\dots,s_{m-1} \mid i_1,\dots,i_{m-1})$ is the probability (under σ^k and θ) of $s_1,\dots,s_{m-1}$ given $i_1,\dots,i_{m-1}$ (recall that s determines i).

We now have:

$$\gamma^{2,m}(p)(\sigma,\theta_{-m},\tau^\star_m) = \sum_{k\in K} p^k \sum_{s_1,t_1,\dots,s_{m-1},t_{m-1}} P^k(s_1,t_1,\dots,s_{m-1},t_{m-1})$$

$$\sigma^k_m(s_1,\dots,s_{m-1})(-A^k)\tau^\star_m(t_1,\dots,t_{m-1})$$

Taking conditional expectation with respect to $(i_1, \ldots, i_{m-1})$ we have, since $\tau^\star$ is reduced, hence measurable w.r.t. $(i_1, \ldots, i_{m-1})$ (and E^k stands for the expectation w.r.t. σ^k and θ):

$$\begin{aligned}\gamma^{2,m}(p)(\sigma, \theta_{-m}, \tau^\star_m) &= \sum_{k \in K} p^k E^k(\sigma^k_m(-A^k)\tau^\star_m) \\ &= \sum_{k \in K} p^k E^k(E^k(\sigma^k_m(-A^k)\tau^\star_m \mid i_1, \ldots, i_{m-1})) \\ &= \sum_{k \in K} p^k E^k(E^k(\sigma^k_m \mid i_1, \ldots, i_{m-1})(-A^k)\tau^\star_m) \\ &= \sum_{k \in K} p^k E^k(\tilde{\sigma}^k_m(-A^k)\tau^\star_m)\end{aligned}$$

Since the distribution of $(i_1, \ldots, i_{m-1})$ is the same under $(\sigma, \theta_{-m}, \tau^\star_m)$ and $(\tilde{\sigma}, \tau^\star)$, we obtain:

$$\gamma^{2,m}(p)(\sigma, \theta_{-m}, \tau^\star_m) = \gamma^{2,m}(p)(\tilde{\sigma}, \tau^\star)$$

It follows that:

$$\begin{aligned}ng^1_n(p)(\sigma, \theta) &= -\sum_{m=1}^{n} \gamma^{2,m}(p)(\sigma, \theta) \leq -\sum_{m=1}^{n} \gamma^{2,m}(p)(\sigma, \theta_{-m}, \tau^\star_m) \\ &= -\sum_{m=1}^{n} \gamma^{2,m}(p)(\tilde{\sigma}, \tau^\star) = ng^1_n(p)(\tilde{\sigma}, \tau^\star) \leq nv_n(p)\end{aligned}$$

Hence

$$v_n(p) \geq g^1_n(p)(\sigma, \theta)$$

which was to be shown.

The proof is analogous in the discounted case. ■

Definition. A *non strategic signalling* structure for Player 2 satisfies:

$$H^{2,k}(i, j) = H^{2,k}(i', j) \Leftrightarrow H^{2,k}(i, j') = H^{2,k}(i', j') \quad \forall j, j'$$

In words, the information that Player 2 obtains on Player 1's move is independent of his own move. A typical example is standard signalling where $H^{2,k}(i, j) = (i, j)$.

Theorem 3. *Assume non strategic signalling for Player 2. Then $\mathscr{E}_n(p)$ is reduced to $\{v_n(p)\}$ (resp. $\mathscr{E}_\lambda(p)$ is reduced to $\{v_\lambda(p)\}$.*

Proof: The proof follows from Theorem 2 and the fact that under non strategic signalling Player 2 has a reduced optimal strategy.

This property can be obtained using the recursive formula for the (primal) game with incomplete information on one side ([1], [4]): since Player 1 knows the posterior distribution of Player 2, he does not have to know the moves. The recursive formula for the dual game now implies the result (see [2], [5]).

An alternative direct approach is as follows (see [3], [6]). Assume that Player 1 plays σ, independent of the moves of Player 2. Then any τ will induce at each stage the same total distirbution on moves than its expectation τ' with respect to Player 2's previous moves. In particular the payoffs given (σ, τ) and (σ, τ') will coïncide. On the other hand, if Player 2 uses a reduced strategy τ, the payoff of Player 1 is the same for a strategy σ or for σ', its expectation (given τ) with respect to Player 2's moves. Denoting by Σ' and $\mathscr{T}'$ the corresponding sets of reduced strategies one thus has:

$$\min_{\mathscr{T}} \max_{\Sigma'} = \min_{\mathscr{T}'} \max_{\Sigma'}$$

$$\min_{\mathscr{T}'} \max_{\Sigma} = \min_{\mathscr{T}'} \max_{\Sigma'}$$

hence:

$$\min_{\mathscr{T}'} \max_{\Sigma} \leq \min_{\mathscr{T}} \max_{\Sigma} (= \max_{\Sigma} \min_{\mathscr{T}})$$

So that Player 2 has an optimal reduced strategy. ■

Remark: The same result holds if Player 2 receives a random signal independent of his moves. However in this case Player 1 may be unable to compute the posterior distribution of Player 2 and there is no recursive formula for the value on $\Delta(K)$.

4. An example

The following example, due to Ponssard and Sorin, was initially constructed to show the impact of the signalling structure on the speed of convergence of the values $v_n(p)$, (see [4] ex. 9, p. 303). The data are:

$$I = \{T, B\}, \quad J = \{L, M, R\}, \quad K = \{k_1, k_2\}$$

$$A^{k_1} = \begin{pmatrix} 0 & 2 & 2 \\ 0 & 2 & -3 \end{pmatrix} \quad A^{k_2} = \begin{pmatrix} 2 & 0 & -3 \\ 2 & 0 & 2 \end{pmatrix}$$

$H^{1,k}(i, j) = (i, j)$ (standard signalling). The signalling matrix for Player 2 is deterministic, independent of the state and given by:

$$H^2 = \begin{pmatrix} a & b & c \\ a & b & f \end{pmatrix}$$

Then one has (with $p = p^1$):

$$v_\infty(p) = \begin{cases} \frac{4}{7} & \text{on } [\frac{2}{7}, \frac{5}{7}], \\ 2\min\{p, 1-p\} & \text{otherwise.} \end{cases}$$

$$v_1(p) = 2\min\{p, 1-p\}.$$

Proposition. *For $p = \frac{1}{2}$ and n large enough, there exists an equilibrium (σ, θ) of $\Gamma_n(p)$ satisfying:*

$$g_n^1(p)(\sigma, \theta) > v_n(p)$$

A similar property holds for $\Gamma_\lambda(p)$.

Proof: Simply define σ^1 as always T, σ^2 as always B and $\theta_m = (\frac{1}{2}, \frac{1}{2}, 0)$. The one shot game $G_1(\frac{1}{2})$ is

$$\begin{array}{c} \\ TT \\ TB \\ BT \\ BB \end{array} \begin{array}{c} \begin{array}{ccc} L & M & R \end{array} \\ \begin{pmatrix} 1 & 1 & -\frac{1}{2} \\ 1 & 1 & 2 \\ 1 & 1 & -3 \\ 1 & 1 & -\frac{1}{2} \end{pmatrix} \end{array}$$

σ is clearly a best reply to θ.

At stage m, given σ and θ, the posterior probability on K is still $(\frac{1}{2}, \frac{1}{2})$. Player 2.$m$ is thus facing the game $G_1(\frac{1}{2})$ where R is a dominated move. Hence θ_m is a best reply and $g_n^1(p)(\sigma, \theta) = \frac{1}{2}$.

Since $v_n(\frac{1}{2})$ converges to $v_\infty(\frac{1}{2}) = \frac{4}{7}$ the result follows.

The proof for Γ_λ is similar. ■

Comments. A single uninformed Player 2 has a long term incentive to buy the information in playing the move Right. Obviously such a strategy is dominated for a short lived Player.

Acknowledgments. We thank the referees for precise remarks.

References

[1] Aumann RJ, Maschler M, with the collaboration of Stearns RE (1995) Repeated games with incomplete information. MIT Press

[2] De Meyer B (1996) Repeated games and partial differential equations. Mathematics of Operations Research 21:209–236

[3] De Meyer B (1996) Repeated games, duality and the Central Limit theorem. Mathematics of Operations Research 21:237–251

[4] Mertens JF, Sorin S, Zamir S (1994) Repeated Games. CORE DP 9420-9421-9422

[5] Rosenberg D (1998) Duality and markovian strategies. International Journal of Game Theory 27:577–597

[6] Rosenberg D (1998) Sur les jeux répétés à somme nulle. Thèse, Université Paris X-Nanterre