Stackelberg-solvable games and pre-play communication *

Claude d'Aspremont
SMASH, Facultés Universitaires Saint-Louis and CORE, Université catholique de Louvain, Louvain-la-Neuve, Belgium
Louis-André Gérard-Varet
GRASCE, Faculté des Sciences de Luminy, Université d'Aix-Marseille, Marseille, France

Abstract

The concept of Stackelberg-solvable games is introduced and analyzed as a generalization of 2-person zero-sum games. Then, the problem of sincere pre-play communication is examined and the incentive compatibility of the Nash-equilibrium selection is shown to correspond to the Stackelberg-equilibrium property.

As already recognized by Harsanyi (1967-68) and Aumann (1974), the introduction of inter-player communication into the analysis of non-cooperative games, where the players restrict their choice to Nash-equilibrium strategies, may change substantially the interpretation of a solution. Indeed, even if in such games no individual commitment is irrevocable and no collective agreement enforceable, some pre-play communication allows the players to coordinate their strategic choice.

However, pre-play communication generates additional strategic considerations: strategic considerations of a higher order. First, in the case where the players have complete information about the data of the game, communication is limited to the strategies to be played. But each player may consider the advantage to commit himself to some specific strategic choice, having some expectation concerning the others' reaction, namely a best-reply behavior.

In the case of incomplete information, we suppose that each player has to communicate to the others the data of the game which he privately knows. In this framework the additional strategic considerations are linked to the fact that the players may find it in their self-interest to distort the information they reveal to each other.

The purpose of this paper is to characterize the class of games which can be solved according to the Nash-Equilibrium concept, in spite of these additional strategic considerations.

In Section 1, we define such a class of games using a generalization of the concept of Stackelberg-point from duopoly theory and we analyze the complete information case. In Section 2, we treat the incomplete information problem.

1. Stackelberg-solvable games

1.1. Let us consider an n-person game in normal form denoted

$$\Gamma \stackrel{\text{def}}{=} \{(X_i, U_i) \mid i \in N\},$$

$*$. Reprinted from *Journal of Economic Theory*, 23(2), 201–217, October 1980 (CORE Reprint no. 425).

where $N = \{1, \dots, i, \dots, n\}$ is the (finite) set of all players, X_i is player i's strategy space and U_i is his payoff function. We assume each X_i to be a compact space and each U_i to be continuous on $X \overset{\text{def}}{=} \times_{i \in N} X_i$.

Denote for every $i \in N$, $x_{-i} \overset{\text{def}}{=} (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$, $X_{-i} \overset{\text{def}}{=} \times_{j \in N, j \neq i} X_j$ and

$$\forall\, \overline{x}_{-i} \in X_{-i}, \mu_i(\overline{x}_{-i}) \overset{\text{def}}{=} \{\overline{x}_i \in X_i \mid U_i(\overline{x}) = \max_{x_i \in X_i} U_i(x_i, \overline{x}_{-i})\}.$$

The correspondence μ_i from X_{-i} to X_i is player i's best reply correspondence in game Γ. We also consider, for every $i \in N$, the following set:

$$G_i \overset{\text{def}}{=} \{x \in X \mid \forall\, j \in N - \{i\}, x_j \in \mu_j(x_{-j})\}.$$

We recall that for a game Γ, a *Nash-equilibrium* is any n-tuple of strategies $x^* \in X$ such that

$$\forall\, i \in N,\ x_i^* \in \mu_i(x_{-i}^*).$$

Note that, for any $x \in G_i$, the $(n-1)$-tuple x_{-i} is a Nash equilibrium in the $(n-1)$-person subgame, where player i is excluded and the payoffs of the other players are defined by supposing that i plays his strategy x_i.

1.2. Following an idea introduced by Von Stackelberg (1934) for duopoly theory, we define a different concept of equilibrium.[1] For a game Γ we call *Stackelberg-point for player* $i \in N$ any n-tuple of strategies $\overline{x}^{(i)} \in G_i$ such that

$$U_i(\overline{x}^{(i)}) = \max_{x \in G_i} U_i(x).$$

This definition may be interpreted in the present framework of a game with pre-play communication. Indeed we may view $U_i(\overline{x}^{(i)})$ as the maximal payoff player i can expect if he commits himself to $\overline{x}_i^{(i)}$ and supposes the others react noncooperatively. The others' behavior is described by the definition of G_i: they choose the Nash-equilibrium $\overline{x}_{-i}^{(i)}$ in the subgame where i plays $\overline{x}_i^{(i)}$. The definition[2] of G_i we use is thus justified because we want to consider the maximal payoff player i can expect in this way.

Now, we call a *Stackelberg-equilibrium* for game Γ any n-tuple of strategies $\overline{x} \in X$ which is a Stackelberg-point for every player $i \in N$. Clearly, any Stackelberg-equilibrium is a Nash-equilibrium. Also, we say that a game Γ is a *Stackelberg-solvable game* iff it possesses at least one Stackelberg-equilibrium.

In a Stackelberg-solvable game no player has interest to block the communication process by imposing a strategy of his choice without any coordination with the others. Indeed the Stackelberg-equilibrium n-tuple of strategies is a Nash-equilibrium which gives to every player the maximal payoff he can expect from such a behavior.

1. See Von Stackelberg (1934) or Fellner (1949).
2. Other interesting generalizations of von Stackelberg definition would be to let (a) $G_i' = \{x \in X; \forall\, j \neq i, x_j \in \mu_j(x_{-j})$ and $\forall\, x' \in X$ such that $x_i' = x_i$ and $\forall\, j \neq i$, $x_j' \in \mu_j(x_{-j}'), U_i(x') \geq U_i(x)\}$. (b) $G_i'' = \{x \in X; \forall\, j \neq i, x_j \in \mu_j(x_{-j})$ and $\forall\, x' \in X$ such that $x_i' = x_i$ and $\forall\, j \neq i$, $x_j' \in \mu_j(x_{-j}')$, if $U_\ell(x') > U_\ell(x)$, $\ell \neq i$, then $U_k(x') < U_k(x)$, $k \neq i\}$ and define a "Stackelberg-point for player i" accordingly.

1.3. In the following we give first an important property of Stackelberg-solvable games and then analyse few examples.

Given a game Γ, write $U = (U_1, \ldots, U_n)$. For any $x \in X$ a vector $u = U(x) \in \mathbb{R}^n$ is called a *feasible payoff*. If $x^* \in X$ is a Nash-equilibrium then $u^* = U(x^*) \in \mathbb{R}^n$ is called a *Nash-equilibrium payoff*. If $\overline{x} \in X$ is a Stackelberg-equilibrium then $v = U(\overline{x}) \in \mathbb{R}^n$ is called a *Stackelberg-equilibrium payoff*.[3] For any two feasible payoff u and u', we say that u' *dominates* u iff: $\forall\, i \in N, u'_i \geq u_i$.

Theorem 1 *In an n-person Stackelberg-solvable game Γ, there is a unique Stackelberg-equilibrium payoff v which dominates any Nash-equilibrium payoff.*

Proof The first part of the theorem is immediate since, if $\overline{x} \in X$ and $\overline{\overline{x}} \in X$ are two Stackelberg-equilibria, then we have

$$\forall\, i \in N,\; U_i(\overline{x}) = U_i(\overline{\overline{x}}) = \max_{x \in G_i} U_i(x) = v_i.$$

For the second part, if $x^* \in X$ is a Nash-equilibrium then $x^* \in \cap_{i \in N} G_i$, and hence

$$\forall\, i \in N,\; U_i(x^*) \leq \max_{x \in G_i} U_i(x) = v_i.$$

■

If a game is Stackelberg-solvable, the unique Stackelberg-equilibrium payoff v will be called the *Stackelberg-value* of the game.

To illustrate these concepts we now study few examples.

Example A Given a game Γ, player i has a *strictly dominating strategy* $\tilde{x}_i \in X_i$ iff: $\forall\, x_{-i} \in X_{-i}, \forall\, x_i \in X_i, x_i \neq \tilde{x}_i \Rightarrow U_i(x_i, x_{-i}) < U_i(\tilde{x}_i, x_{-i})$. If $(\tilde{x}_1, \ldots, \tilde{x}_i, \ldots, \tilde{x}_n)$ is an n-tuple of strictly dominating strategies then it is the unique Nash-equilibrium and Stackelberg-equilibrium. Accordingly, (2.2) is the Stackelberg-value of the following 2-person game:

	x_2^1	x_2^2
x_1^1	3,0	0,1
x_1^2	5,0	2,2

Example B A two-person game $\Gamma = \{X_1, X_2, U_1, U_2\}$ is called strictly competitive iff:

$$\begin{aligned}&\forall\, x \in X,\; \forall\, x' \in X,\\ &U_1(x') > U_1(x) \text{ iff } U_2(x') < U_2(x)\end{aligned}$$

and

$$U_1(x') = U_1(x) \text{ iff } U_2(x') = U_2(x).$$

Concerning this class of games we have the following statement: *If Γ is a 2-person strictly competitive game, then every Nash-equilibrium is a Stackelberg-equilibrium.*

3. If $\overline{x}^{(i)} \in X$ is a Stackelberg point for player $i \in N$, then $v_i^i = U_i(\overline{x}^{(i)}) \in \mathbb{R}$ denotes *player i's Stackelberg payoff*. However, $(v_1^1, \ldots, v_n^n) \in \mathbb{R}^n$ is not necessarily a Nash equilibrium payoff and is not even necessarily feasible.

Proof If Γ is a strictly competitive game, we may write

$$\begin{aligned} G_1 &= \{x \in X : U_2(x) = \max_{\tilde{x}_2 \in X_2} U_2(x_1, \tilde{x}_2)\} \\ &= \{x \in X : U_1(x) = \min_{\tilde{x}_2 \in X_2} U_1(x_1, \tilde{x}_2)\} \\ G_2 &= \{x \in X : U_1(x) = \max_{\tilde{x}_1 \in X_1} U_1(\tilde{x}_1, x_2)\} \\ &= \{x \in X : U_2(x) = \min_{\tilde{x}_1 \in X_1} U_2(\tilde{x}_1, x_2)\}. \end{aligned}$$

Also if $\overline{x}$ is a Nash-equilibrium in Γ we must have

$$\begin{aligned} U_1(\overline{x}) &= \max_{x_1 \in X_1} \min_{x_2 \in X_2} U_1(x_1, x_2) = \max_{x \in G_1} U_1(x) \\ U_2(\overline{x}) &= \max_{x_2 \in X_2} \min_{x_1 \in X_1} U_2(x_1, x_2) = \max_{x \in G_2} U_2(x). \end{aligned}$$

This with the fact $\overline{x} \in G_1 \cap G_2$ implies that $\overline{x}$ is a Stackelberg-equilibrium.

This proves that in some sense the notion of Stackelberg-solvable games is a generalization of the notion of 2-person strictly competitive games, (i.e., games equivalent to a 2-person 0-sum game). ■

Example C Consider two firms producing the same good in nonnegative quantity q_1 and q_2 at a cost $a_1 q_1$ and $a_2 q_2$, respectively, and selling on a market where the demand function is piecewise linear, i.e., for $B > b > 0$, the price p is given by

$$\begin{aligned} p(q_1, q_2) &= B - b(q_1 + q_2) \qquad && q_1 + q_2 \leq B/b \\ &= 0 && q_1 + q_2 > B/b. \end{aligned}$$

We suppose that a_1 and a_2 belong to the open interval $(0, B)$. For each value of the parameters a_1 and a_2 we can construct a duopoly game $\Gamma(a)$ as follows. Let $X_1 = X_2 = [0, B/b]$ and

$$\begin{aligned} U_1(q_1, q_2) &= (B - a_1)q_1 - bq_1^2 - bq_2 q_1 \qquad && q_1 + q_2 \leq B/b \\ &= -a_1 q_1 && q_1 + q_2 > B/b \\ U_2(q_1, q_2) &= (B - a_2)q_2 - bq_2^2 - bq_1 q_2 && q_1 + q_2 \leq B/b \\ &= -a_2 q_2 && q_1 + q_2 > B/b. \end{aligned}$$

In the Appendix, we give the computation of the Stackelberg-point $q^{(1)}$ for player 1. The result is

$$\begin{aligned} & q_1^{(1)} = 0; \quad q_2^{(1)} = \frac{B - a_2}{2b}, && \text{if } (B - a_1) \leq \frac{1}{2}(B - a_2), \\ & q_1^{(1)} = \frac{B - a_1}{2b}; \quad q_2^{(1)} = 0, && \text{if } (B - a_1) \geq 2(B - a_2), \\ & q_1^{(1)} = \frac{1}{b}\left[(B - a_1) - \frac{B - a_2}{2}\right]; q_2^{(1)} = 0; && \text{if } \frac{3}{2}(B - a_2) < (B - a_1) < 2(B - a_2) \\ & q_1^{(1)} = \frac{1}{b}\left[(B - a_1) - \frac{B - a_2}{2}\right]; && \\ & q_2^{(1)} = \frac{1}{2b}\left[\frac{3}{2}(B - a_2) - (B - a_1)\right], && \text{otherwise.} \end{aligned}$$

The Stackelberg-point for player 2, $q^{(2)}$ may be written symmetrically. Now, if the values of the parameters are such that $\frac{1}{2}(B-a_2) < (B-a_1) < 2(B-a_2)$, then the unique Nash-equilibrium q^* is given by:

$$q_1^* = \frac{1}{b}\left[\frac{2}{3}(B-a_1) - \frac{(B-a_2)}{3}\right];$$
$$q_2^* = \frac{1}{b}\left[\frac{2}{3}(B-a_2) - \frac{(B-a_1)}{3}\right].$$

For this Nash-equilibrium to be a Stackelberg-equilibrium one should have

$$\frac{2}{3}(B-a_1) - \frac{B-a_2}{3} = (B-a_1) - \frac{B-a_2}{2}, \quad i.e.\ \frac{B-a_2}{2} = B-a_1$$

and

$$\frac{2}{3}(B-a_2) - \frac{B-a_1}{3} = (B-a_2) - \frac{B-a_1}{2}, \quad i.e.\ B-a_1 = 2(B-a_2),$$

which is impossible. Hence, for values of the parameters such that the Nash-equilibrium quantities are positive for each duopolist,[4] the game $\Gamma(a)$ is not Stackelberg-solvable and the situation, as recognized in Stackelberg original analysis, is very unstable, if both desire to be leader.

1.4. Let us now consider the relationship between Stackelberg-solvable games, almost-strictly-competitive games as introduced by Aumann (1961) and a-cooperative games as defined by Moulin (1976a).

First recall that a *twisted equilibrium* for a game Γ is any n-tuple $\hat{x} \in X$ such that

$$\forall\, i \in N,\ \forall\, x_{-i} \in X_{-i},\ U_i(\hat{x}) \leq U_i(\hat{x}_i, x_{-i}).$$

In a 2-person game, a twisted equilibrium is such that no player can decrease the other player's payoff by a unilateral change of strategy. If $\hat{x} \in X$ is a twisted equilibrium then $\hat{u} = U(\hat{x}) \in \mathbb{R}^n$ is a *twisted equilibrium payoff*.

A game Γ is *almost-strictly-competitive* (for short a.s.c.) iff there exists an n-tuple of strategies which is both a Nash-equilibrium and a twisted-equilibrium and the set of Nash-equilibrium payoffs coincides with the set of twisted equilibrium payoffs. In an a.s.c. game, there is a unique Nash-equilibrium payoff which is also the unique twisted equilibrium payoff.[5]

All 2-person zero-sum-games are a.s.c. Other games are both a.s.c. and Stackelberg-solvable as shown by the "prisoners' dilemma."

Example D

	x_2^1	x_2^1
x_1^1	4,4	0,5
x_1^2	5,0	2,2

However, the class of a.s.c. games is different from the class of Stackelberg-solvable games. On one hand, Example A provides a game which is Stackelberg-solvable but is not a.s.c. since the unique Nash-equilibrium is not a twisted equilibrium. On the other hand, there exist games which are a.s.c. but are not Stackelberg-solvable, as shown by Example E.

4. For other values of the parameters, we have $q^* = q^{(1)} = q^{(2)}$ and one of the duopolists produces nothing.
5. This result forms an easy extension to the n-person case of Theorem A in Aumann (1961) which is proved for the 2-person case. Notice that other properties of 2-person a.s.c. games do not remain valid in the general case.

Example E

	x_2^1	x_2^1
x_1^1	4,4	0,3
x_1^2	5,0	1,1

In this game (x_1^2, x_2^2) is the unique Nash-equilibrium and the unique twisted-equilibrium. However, it is not a Stackelberg-equilibrium.

Given a game Γ, we say of an n-tuple $x \in X$ that it is *Pareto-optimal* iff there is no other n-tuple of strategies giving at least as much to every player and more to some player, i.e.,

$$\nexists\, x' \in X \text{ such that } \forall\, i \in N,\ U_i(x') \geq U_i(x)$$

and

$$\exists\, j \in N \text{ such that } U_j(x') > U_j(x).$$

If $x \in X$ is a Pareto-optimal n-tuple then $u = U(x) \in \mathbb{R}^n$ is a *Pareto-optimal payoff.*

A game Γ is *a-cooperative* iff it has at least one Pareto-optimal twisted equilibrium. The class of a-cooperative games (strictly) contains the class of 2-person zero-sum games and is (strictly) contained in the class of a.s.c. games.[6] Moreover,

Theorem 2 *In a game Γ, any Pareto-optimal twisted equilibrium is a Stackelberg-equilibrium.*

Proof Suppose $\hat{x} \in X$ is a Pareto-optimal twisted-equilibrium. First we may show that $\hat{x}$ is a Nash-equilibrium. Indeed if such was not the case we would have

$$\exists\, j \in N,\ \exists\, x_j \in X_j \text{ such that } U_j(\hat{x}) < U_j(x_j, \hat{x}_{-j}).$$

And, since $\hat{x}$ is a twisted-equilibrium, we get, in addition,

$$\forall\, i \in N - \{j\},\ U_i(\hat{x}) \leq U_i(x_j, \hat{x}_{-j}),$$

a contradiction to the Pareto-optimality of $\hat{x}$.

Hence, we must have $\hat{x} \in \cap_{i \in N} G_i$. Moreover, if we don't have $U(\hat{x}) = v$, i.e., $\exists\, j \in N$, $\exists\, \overline{x} \in G_j$ such that $U_j(\hat{x}) < U_j(\overline{x}) = \max_{x \in G_j} U_j(x)$, then, by the Pareto-optimality of $\hat{x}$ and the fact that $\overline{x}_j \in G_j$, we get

$$\exists\, k \in N - \{j\} \text{ such that } U_k(\hat{x}) > U_k(\overline{x}) \geq U_k(\hat{x}_k, \overline{x}_{-k}),$$

a contradiction to $\hat{x}$ being a twisted equilibrium. ■

Theorem 2 shows that any a-cooperative game is a Stackelberg-solvable game. The prisoner's dilemma (Example D) provides an example of a game which is Stackelberg-solvable but is not a-cooperative.

Remark 1 *It is interesting to note that Theorems 1 and 2 imply that in any zero-sum game having a twisted equilibrium, there is a unique Nash-equilibrium payoff. Indeed in such a game any twisted equilibrium is Pareto-optimal and, by Theorem 2, it is also Stackelberg-equilibrium which, by Theorem 1, gives the unique Nash-equilibrium payoff.*

6. Characterizations of a-cooperative games have been provided for the 2-person case by Moulin (1976a,b). Example D gives a game which is a.s.c. but not a-cooperative.

As a last example we give the "variable-threat bargaining game."

Example F Consider a compact convex set X of "outcomes" on which are defined two continuous utility functions U_1 and U_2. Each player i is supposed to have a compact convex set[7] of available "threat" strategies X_i such that $X_1 \times X_2 \subset X$. Then an "arbitration scheme" is any function σ from $X_1 \times X_2$ to X. The following minimal requirements are usually imposed on σ:

(i) Individual Rationality: $\forall\, x \in X, \forall\, i \in \{1,2\}$,

$$U_i(\sigma(x)) \geq U_i(x).$$

(ii) Pareto Optimality: $\forall\, x \in X, \forall\, x' \in X, \forall\, i \in \{1,2\}, \forall\, j \in \{1,2\}$,

$$U_i(x') > U_i(\sigma(x)) \Rightarrow U_j(x') < U_j(\sigma(x)).$$

For any arbitration scheme σ one may associate a game, also called the "threat-game" and denoted

$$\Gamma_\sigma = \{X_1, X_2, U_1^\sigma, U_2^\sigma\},$$

where

$$U_i^\sigma \stackrel{\text{def}}{=} U_i(\sigma(\cdot)),\ i = 1, 2.$$

As remarked in Aumann (1961) if σ satisfies (i) and (ii) and if the game Γ_σ has a Nash-equilibrium then it is almost strictly competitive. Hence there is a unique Nash-equilibrium and twisted-equilibrium payoff which by condition (ii) is Pareto-optimal. So we may conclude then that Γ_σ is a-cooperative and, by Theorem 2, that it is Stackelberg-solvable.

2. Incentives and complete ignorance in non-cooperative games

2.1. The game Γ introduced in the previous section is a game with *complete information*: all payoff functions and all strategy spaces are "common knowledge" in the sense that no player can consciously disagree with some other about what they are. However, in practice, some of this information might be initially "private information" and then some kind of communication between players might have to take place before the game. In this context, the complete information assumption appears as equivalent to a sincere pre-play communication assumption. Our purpose now is to weaken this assumption considerably. We proceed as follows.

We suppose first that every player i is described by some parameter α_i of finite or infinite dimension. To introduce *incomplete information* we assume then that every player only knows his parameter and that the parameter value of any other player j, α_j, belongs to some space A_j, which is the space of descriptions of all possible values.[8]

Also we assume that the functions U_i are respectively defined on $X \times A_i$, i.e., if α_i is the parameter value for player i, $U_i(\cdot;\alpha_i)$ is his payoff function. Hence for every n-tuple $\alpha = (\alpha_1, \ldots, \alpha_n) \in A$, where $A = \bigtimes_{i \in N} A_i$, we get a game $\Gamma(\alpha) = \{(X_i, U_i(\cdot;\alpha_i)); i \in N\}$ of the type described in Section 1.1. Player i's best reply correspondence in game $\Gamma(\alpha)$ can be denoted $\mu_i(\cdot;\alpha_i)$ and accordingly

7. As in Nash (1953), Raiffa (1953) and Kalai and Rosenthal (1978), we could have started with a two-person finite game, consider Γ as its mixed extension and X as the set of correlated pairs of strategies.

8. In probabilistic terms, for every player $j \neq i$, α_i is a random phenomenon and A_i is its sample space.

every set G_i will be recognized to depend on $\alpha_{-i} = (\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_n) \in A_{-i}$, where $A_{-i} = \bigtimes_{j \in N, j \neq i} A_j$.

In such an incomplete information framework, pre-play communication can no longer be considered as straightforward. For our purpose we shall assume that during the pre-play communication period the players have to settle two problems:

(i) They have to agree on a *selection rule* **x** from A to X giving the unique n-tuple of strategies $\mathbf{x}(\alpha) \in X$ which would have to be played if the game is $\Gamma(\alpha)$.

(ii) Each player i has to announce a value $a_i \in A_i$ of his parameter which is taken by the other players as his true value. Each possible a_i is called a *message* of player i. Hence A_i can also be viewed as the message space of player i. In fact, to solve problem (ii), known as the "incentive problem", the players have to play another kind of game. More specifically we define the *pre-play communication game given the selection* **x** as the set of normal form games – one for every $\alpha \in A$ – denoted

$$\mathcal{G}(\mathbf{x}; \alpha) = \{(A_i, U_i(\mathbf{x}(\cdot); \alpha_i); i \in N\},\ a \in A,$$

where the players strategy spaces A_i are their message space and where their payoff functions are the composition of their initial payoff functions with **x**. This is a game with incomplete information in which no account is taken of the beliefs of the players regarding each other parameters. This is in contrast to Harsanyi (1967-68) Bayesian approach and, by similarity with the well-known distinction in the theory of decision under uncertainty,[9] the present approach may be called the "complete ignorance" approach (see d'Aspremont and Gérard-Varet, 1979).

In this approach one may define the following solution concept: an n-tuple of messages $a^* \in A$ is a *Nash-equilibrium for* α (or for the game $\mathcal{G}(\mathbf{x},\alpha)$) iff:

$$\forall\, i \in N,\ \forall\, a_i \in A_i,\ U_i(\mathbf{x}(a_i, a^*_{-i}), \alpha_i) \leq U_i(\mathbf{x}(a^*); \alpha_i).$$

Now, following Hurwicz (1972), we may say that a selection **x** is *incentive compatible for* α iff α is a Nash-equilibrium for α, i.e.,

$$\forall\, i \in N,\ \forall\, a_i \in A_i,\ U_i(\mathbf{x}(a_i, \alpha_{-i}); \alpha_i) \leq U_i(\mathbf{x}(\alpha); \alpha_i).$$

Of course such a property is very weak since it is a "local" property. Hence, one has to look for classes of values α for which the property holds.[10]

2.2. In the present paper we assume that for every $a \in A$ the revealed game $\Gamma(a)$ is expected to be played non-cooperatively and hence that the selection **x** gives some Nash-equilibrium $\mathbf{x}(a)$ for the game $\Gamma(a)$. Such a selection should be viewed as an agreement between the players which is not enforceable. Actually, we call **x** a *Nash selection (undominated)* iff for every $a \in A$, $\mathbf{x}(a)$ is some Nash-equilibrium for $\Gamma(a)$ (which is dominated by no other Nash-equilibrium).

We have now the following:

9. See Luce and Raiffa (1953, Chap. 13).

10. We consider the following stronger property in d'Aspremont and Gérard-Varet (1979). We say that a selection **x** is *uniformly incentive compatible* iff it is incentive compatible for every $\alpha \in A$. We show that if **x** is uniformly incentive compatible then it is true that for every $\alpha \in A$ and every player $i \in N$, α_i is a dominating strategy in the game $\mathcal{G}(x; \alpha)$.

Theorem 3 *If $\Gamma(\alpha)$ is a Stackelberg-solvable game and if for every $i \in N$ and every $a_i \in A_i$, $\mathbf{x}(a_i, \alpha_{-i})$ is some undominated Nash-equilibrium for $\Gamma(a_i, \alpha_{-i})$ then the selection $\mathbf{x}$ is incentive compatible for α.*

Proof Since $x(\alpha)$ is an undominated Nash-equilibrium for $\Gamma(\alpha)$, for every $i \in N$, $U_i(\mathbf{x}(\alpha); \alpha_i)$ $= v_i$, where v is the Stackelberg-value of $\Gamma(\alpha)$ (by Theorem 1). Moreover, since $\forall\, i \in N, \forall\, a_i \in A_i$, $\mathbf{x}(a_i, \alpha_{-i})$ is a Nash-equilibrium for the game $\Gamma(a_i, \alpha_{-i})$, we get: $\mathbf{x}(a_i, \alpha_{-i}) \in G_i(\alpha_{-i})$.

But then:

$$\forall\, i \in N,\ \forall\, a_i \in A_i,\ \ U_i(\mathbf{x}(a_i, \alpha_{-i}); \alpha_i) \leq \max_{x \in G_i(\alpha - i)} U_i(x; \alpha_i) = v_i = U_i(\mathbf{x}(\alpha); \alpha_i).$$

■

By using Theorem 2, a corollary of Theorem 3 is that the same incentive compatibility property holds at α whenever $\Gamma(\alpha)$ is an a-cooperative game.

Let us now consider the possibility of a converse result. We need for that matter a reasonable condition ensuring the relative "richness" of each space A_i of possible reported values of the parameters. In some sense to get the most general result, one should allow each player to announce any kind of payoff function. However in the present context where we restrict our attention to Nash selections, any player i, knowing the other players' parameters α_{-i}, cannot expect to reach a point outside $G_i(\alpha_{-i})$. The following theorem is based on a condition taking into account these considerations.

Theorem 4 *Assume a selection rule $\mathbf{x}$ such that, for some $\alpha \in A$ and every $i \in N$, $\{\mathbf{x}(a_i, \alpha_{-i}) : a_i \in A_i\} = G_i(\alpha_{-i})$. If in addition the selection $\mathbf{x}$ is incentive compatible for α, then $\mathbf{x}(\alpha)$ is a Stackelberg-equilibrium in $\Gamma(\alpha)$.*

Proof By the assumption on $\mathbf{x}$, we can find, for the game $\Gamma(\alpha)$ and every $i \in N$, some $\overline{a}_i \in A_i$ such that

$$\max_{x \in G_i(\alpha_{-i})} U_i(x; \alpha_i) = U_i(\mathbf{x}(\overline{a}_i, \alpha_{-i}); \alpha_i).$$

Moreover, since $\mathbf{x}$ is incentive compatible for α,

$$U_i(\mathbf{x}(\overline{a}_i, \alpha_{-i}); \alpha_i) \leq U_i(\mathbf{x}(\alpha); \alpha_i).$$

But, by assumption,

$$\forall\, i \in N,\ \mathbf{x}(\alpha) \in G_i(\alpha_{-i}).$$

Therefore, $\mathbf{x}(\alpha)$ is a Stackelberg-equilibrium for $\Gamma(\alpha)$. ■

It is clear that if we require the selection to be a Nash selection, then this converse result holds, with a much weaker "richness" condition. Actually, the following characterization results by the arguments used in Theorems 3 and 4.

Theorem 5 *Assume an undominated Nash selection such that, for some $\alpha \in A$ and every $i \in N$,*

$$\mathbf{x}(a_i, \alpha_{-i}) = \overline{x}^i,$$

where $a_i \in A_i$ and $\overline{x}^i$ is some Stackelberg-point for i in $\Gamma(\alpha)$. Then the selection $\mathbf{x}$ is incentive compatible for α iff $\mathbf{x}$ is a Stackelberg-equilibrium in $\Gamma(\alpha)$.

2.3. To illustrate the richness assumptions of Theorems 4 and 5, consider the particular case where there are only two players and where every $\mu_i(\cdot;\alpha_i)$ is a function. In this case we may draw a figure of the kind of Figure 1. The condition used in Theorem 4 requires that player 1 can displace his best-reply function μ_1 so that nay point on $\mu_2(x;\alpha_2) = G_1(\alpha_2)$ may be obtained as a Nash-equilibrium. The one used in Theorem 5 only requires that in addition to x^* the Stackelberg-point $\overline{x}^{(1)}$ may be obtained. Similarly for player 2. It is easy to see then that the second condition is weaker than the first. Returning to Example C we may check that the two conditions are satisfied.

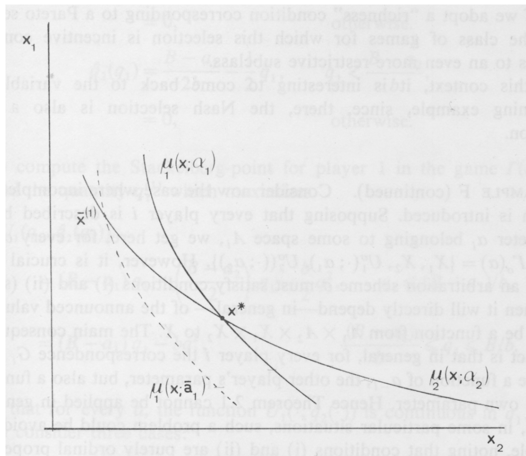

Figure 1

Example C (continued) It is enough to consider any $\alpha \in A$ such that $\frac{1}{2}(B-\alpha_2) < (B-\alpha_1) < \frac{3}{2}(B-\alpha_2)$ and let, say for player 1, $\overline{a}_1$ be such that

$$(B-\overline{a}_1) = \frac{3}{2}(B-\alpha_1) - \frac{1}{4}(B-\alpha_2).$$

Then, we get

$$\frac{1}{2}(B-a_2) < (B-\overline{a}_1) < 2(B-\alpha_2)$$

and

$$q_1^*(\overline{a}_1,\alpha_2) = q_1^{(1)}(\alpha).$$

In conclusion, we should remark that a result like Theorem 5 is of the same family as Hurwicz (1972) "Impossibility Theorem": one is concerned with a Nash selection, the other with a Pareto selection. Theorem 5 shows that, if we adop a "richness" assumption corresponding to a Nash selection, then the class of games for which this selection is incentive compatible reduces to a very particular subclass of all games, namely the class of Stackelberg-solvable games. Similarly, Hurwicz's argument can be interpreted by saying that, if we adopt a "richness" condition corresponding to a Pareto selection, then the class of games for which this selection is incentive compatible reduces to an even more restrictive subclass.

In this context, it is interesting to come back to the variable-threat bargaining example, since, there, the Nash selection is also a Pareto selection.

Example F (continued) Consider now the case where incomplete information is introduced. Supposing that every player i is described by some parameter α_i belonging to some space A_i, we get here, for every $\alpha \in A$, a game $\Gamma_\sigma(\alpha) = \{X_1, X_2, U_1^\sigma(\cdot;\alpha_1), U_2^\sigma(\cdot;\alpha_2)\}$. However, it is crucial to note that if an arbitration scheme σ must satisfy conditions (i) and (ii) (see page 9), then it will directly depend – in general – of the announced value a, i.e., σ will be a function from $A_1 \times A_2 \times X_1 \times X_2$ to X. The main consequence of this fact is that in general, for every player i the correspondence G_i will not only be a function of a_{-i}, the other player's parameter, but also a function of a_i, his own parameter. Hence Theorem 3 cannot be applied in general. Of course, in some particular situations, such a problem could be avoided. For example, noting that conditions (i) and (ii) are purely ordinal properties on the set X, we could let, for every $i \in N$, A_i be the set of all monotone strictly increasing transformation from $\mathbb{R}$ to itself such that, for every transformation $\alpha_i \in A_i$ and every $x \in X$

$$U_i(x;\alpha_i) = \alpha_i(u_i(x)),$$

where $u_i(\cdot)$ is some given function from X to $\mathbb{R}$. Then, any scheme σ satisfying (i) and (ii) for some arbitrary $\alpha \in A$ necessarily satisfies (i) and (ii) for any other $a \in A$. But, this provides a trivial example since such arbitration schemes are constant in $\alpha \in A$. Moreover, as shown by Shapley (1969), any arbitration scheme which would depend only on the utility payoffs of the players cannot be invariant to any pair of order-perserving transformations of the two players payoffs. However, this is not equivalent to the impossibility of finding an arbitration scheme which satisfies (i) and (ii) and depends on $\alpha \in A$ in such a way that the argument of Theorem 3 still holds.

Appendix: Computations for Example C

The "reaction curves" are described by

$$\begin{aligned}
\hat{q}_1(q_2) &= \frac{B-a_1}{2b} - \frac{1}{2}q_2, && q_2 < \frac{B-a_1}{b},\\
&= 0, && \text{otherwise.}\\
\hat{q}_2(q_1) &= \frac{B-a_2}{2b} - \frac{1}{2}q_1, && q_1 < \frac{B-a_2}{b},\\
&= 0, && \text{otherwise.}
\end{aligned}$$

Now to compute the Stackelberg-point for player 1 in the game $\Gamma(a)$, we must find the quantity $q_1^{(1)}$ which maximizes

$$\begin{aligned}
U_1(q_1, \hat{q}_2(q_1)) &= (B-a_1)q_1 - \left(\frac{B-a_2}{2}\right) q_1 - \frac{b}{2}q_1^2, && q_1 \leq (B-a_2)/b,\\
&= (B-a_1)q_1 - bq_1^2, && \frac{(B-a_2)}{b} \leq q_1 \leq B/b.
\end{aligned}$$

We see that for every a, the function $U_1(\cdot, \hat{q}_2(\cdot))$ is continuous in q_1. Now we shall consider three cases:

(i) $(B-a_1)/2 < (B-a_1) < 2(B-a_2)$.

In this case, for $(B-a_2)/b \leq q_1 < B/b$,

$$U_1'(q_1, \hat{q}_2(q_1)) = (B-a_1) - 2bq_1 \leq (B-a_1) - 2(B-a_2) < 0$$

and hence $U_1(q_1, \hat{q}_2(q_1))$ is decreasing for $q_1 \geq (B - a_2)/b$.
Therefore in this case: $0 \leq q_1^{(1)}(a) < (B - a_2)/b$.
Now the function $[(B - a_1)q_1 - (B - a_2/2)q_1 - (b/2)q_1^2]$ is strictly concave and its maximum is for $q_1 = (B - a_1)/b - (B - a_2)/2b$, which is in this case both positive and less than $(B - a_1)/b$. Therefore $q_1^{(1)} = (B - a_1)/b - (B - a_2)/2b$.

(ii) $(B - a_2)/2 \geq (B - a_1)$.
Again $U_1(q_1, \hat{q}_1(q_1))$ is decreasing for $q_1 \geq (B - a_2)/B$. In addition

$$\left[(B - a_1) - \left(\frac{B - a_2}{2}\right)\right] q_1 - \frac{b}{2}q_1^2 \leq 0,$$

and so

$$U_1(q_1, \hat{q}_2(q_1)) \leq 0, \quad \forall\, q_1.$$

Therefore $q_1^{(1)} = 0$.

(iii) $(B - a_1) \geq 2(B - a_2)$.
The maximum value (for $q_1 \geq 0$) of $[(B - a_1)q_1 - (B - a_2/2)q_1 - (b/2)q_1^2]$ is $(1/2b)[(B - a_1) - (B - a_2)/2]^2$. On the other hand the function $[(B - a_1)q_1 - bq_1^2]$ is strictly concave, obtains its maximum for $q_1 = (B - a_1)/2b$ and reaches there the value $[(1/2b)(B - a_1)^2]$ which, in this case, is greater than $(1/2b)[(B - a_1) - (B - a_2)/2]^2$. Hence $q_1^{(1)} = (B - a_1)/2b$.
In conclusion we get

$$\begin{array}{ll}
q_1^{(1)} = \dfrac{(B - a_1)}{2b}; \quad q_2^{(1)} = 0 & \text{if } (B - a_1) \geq 2(B - a_2) \\[2ex]
q_1^{(1)} = 0; \quad q_2^{(1)} = \dfrac{(B - a_2)}{2b} & \text{if } \dfrac{1}{2}(B - a_2) \geq (B - a_1) \\[2ex]
q_1^{(1)} = \dfrac{1}{2}\left[(B - a_1) - \dfrac{(B - a_2)}{2}\right]; q_2^{(1)} = 0; & \text{if } \dfrac{3}{2}(B - a_2) < (B - a_1) < 2(B - a_2) \\[2ex]
q_1^{(1)} = \dfrac{1}{b}\left[(B - a_1) - \dfrac{(B - a_2)}{2}\right]; & \\[2ex]
q_2^{(1)} = \dfrac{1}{2b}\left[\dfrac{3}{2}(B - a_2) - (B - a_1)\right], & \text{otherwise.}
\end{array}$$

Acknowledgments

We would like to thank Yossi Greenberg for several valuable suggestions. Comments by an unknown referee are also gratefully acknowledged.

References

Robert J Aumann. Almost strictly competitive games. *Journal of the Society for Industrial and Applied Mathematics*, 9(4):544–550, 1961.

Robert J Aumann. Subjectivity and correlation in randomized strategies. *Journal of mathematical Economics*, 1(1):67–96, 1974.

Claude d'Aspremont and Louis-André Gérard-Varet. Incentives and incomplete information. *Journal of Public economics*, 11(1):25–45, 1979.

William Fellner. *Competition among the Few*. Kelley, New York, 1949.

John C Harsanyi. Games with incomplete information played by 'Bayesian' players. Parts I–III. *Management science*, 14:158–182, 320–334, 486–502, 1967-68.

Leonid Hurwicz. On informationally decentralized systems. In R. Radner and C.D. McGuire, editors, *Decision and Organization*. North Holland, Amsterdam, 1972.

Ehud Kalai and Robert W Rosenthal. Arbitration of two-party disputes under ignorance. *International Journal of Game Theory*, 7(2):65–72, 1978.

R Duncan Luce and Howard Raiffa. *Games and decisions: Introduction and critical survey*. L. Wilsey, New York, 1957.

Hervé Moulin. Correlated and prudent equilibria for a 2-person game. Cahiers Mathématiques de la Décision 7605, Université de Paris IX-Dauphine, 1976a.

Hervé Moulin. Les contrats d'un jeu à 2 joueurs. Cahiers Mathématiques de la Décision 7609, Université de Paris IX-Dauphine, 1976b.

John Nash. Two-person cooperative games. *Econometrica*, 21:138–160, 1953.

Howard Raiffa. Arbitration schemes for generalized two-person games. In H.W. Kuhn and A.W. Tucker, editors, *Contributions to the Theory of Games*, volume II, Annals of Mathematics, No. 28, pages 361–387. Princeton University Press, Princeton, NJ, 1953.

Lloyd S Shapley. Utility comparison and the theory of games. In *La décision*, pages 251–263. Editions du C.N.R.S., 1969.

Heinrich Von Stackelberg. *Marktform und gleichgewicht*. Springer, Wien und Berlin, 1934.