Bayesian incentive compatible beliefs*

Claude d'Aspremont
CORE, Université catholique de Louvain
Louis-André Gérard-Varet
Université Louis Pasteur, Strasbourg I and G.R.E.Q.E. - Era CNRS983, Universités Aix-Marseille.†

Abstract

The problem of incentives for correct revelation is studied as a game with incomplete information where players have individual beliefs concerning others' types. General conditions on the beliefs are given which are shown to be sufficient for the existence of a Pareto-efficient mechanism for which truth-telling is a Bayesian equilibrium.

1. Introduction

Much attention has been devoted recently to a game-theoretic analysis of the strategic use of private information by individual agents in a decentralized mechanism, designed to reach a Pareto optimal outcome (For a recent bibliography, see Green and Laffont, 1979). In a previous paper (d'Aspremont and Gérard-Varet, 1979) we have distinguished two approaches to such games of incomplete information. The first approach, the 'complete ignorance' approach, relies on strong equilibrium concepts, such as the 'uniform equilibrium' or the 'dominant strategy equilibrium', which are independent of any particular representation of the players' expectations. This is not the case for the second approach which is based on Harsanyi (1967-68) concept of 'Bayesian equilibrium'. One basic result for this approach (see Arrow, 1979; d'Aspremont and Gérard-Varet, 1975, 1979) is that under a transferability assumption on the utility functions and an independence assumption on the beliefs (or probability distributions) of the players, there exist Pareto-efficient mechanisms (with feasible transfer schemes) for which 'truth-telling' is a Bayesian equilibrium. However, it has also been shown that the same result could be obtained under a weaker condition – called the 'compatibility condition' – but for discrete probability distributions.

In the present work, we shall show how to extend to a more general framework the result obtained under the assumption of discrete beliefs. We shall also provide a detailed discussion[1] of the 'compatibility condition', implying in particular that it is not a necessary condition for the existence of a Pareto-efficient mechanism for which 'truth-telling' is a Bayesian equilibrium. Finally we shall show that a much less restrictive condition is required if Pareto-efficiency is weakened to a property in expected value.

∗. Reprinted from *Journal of Mathematical Economics*, 10, 83–103, 1982 (CORE Reprint no. 481).

†. The materials presented in this paper were worked out when the author was visiting the Institut für Gesellschafts- und Wirtschaftswissenschaften, Universität Bonn, Germany. Support from the Institut is gratefully acknowledged.
The authors want to thank J.-P. Florens, G. Laffond, J.-F. Mertens, H. Moulin, S. Sorin and all participants at the 1980 Stanford I.M.S.S.S. Summer Workshop for helpful comments.

1. In particular we shall give stronger but more interpretable conditions.

2. The communication game of incomplete information

In its simplest form a mechanism based on preference revelation may be described as follows. Suppose a finite set $N = \{1, \ldots, i, \ldots, n\}$ of players and a set X of *outcomes*. To give examples (see Green and Laffont, 1977; Groves and Loeb, 1975) X can be a set of possible public projects or a set of quantity levels of a public good (or bad) or, more generally, the set of joint strategies in some n-person game. Assume that the set of admissible preference orderings of X for any player i in N can be represented by the family of *payoff functions* $U_i(x; \alpha_i)$, where α_i is a general parameter varying in some space A_i and summarizing all the private information of player i. We also say that A_i is the space of possible *types* of player $i \in N$. Furthermore, we shall suppose that all payoff transfers among players are permitted and that every $U_i(\cdot, \alpha_i)$ is so calibrated that, for any x in X and any transfer y in $\mathbb{R}$, the payoff of player $i \in N$ of parameter $\alpha_i \in A_i$ is equal to $U_i(x; \alpha_i) + y$. In other words we assume unrestricted side-payments with full-transferability.

The sets X and $\{A_i; i \in N\}$ and the functions $\{U_i; i \in N\}$ are common knowledge, but player $i \in N$ is privately informed only about his own type in A_i. Let us assume that any player $i \in N$ has to independently announce to the others some type as being his own type and that A_i is the space of all his possible *messages*. A *mechanism* is then defined as a *selection rule* s, which is a function from the message n-tuple $a = (a_1, a_2, \ldots, a_n)$ in $A = \times_{i=1}^{n} A_i$ to the set of outcomes, and a *transfer scheme* t, which associates to each n-tuple of messages $a \in A$ a certain vector of transfers $(t_1(a), \ldots, t_n(a))$ in $\mathbb{R}^n$ made to the individual players. So, for any n-tuple of messages $a \in A$, the payoff of player $i \in N$ of type $\alpha_i \in A_i$ under the mechanism $m = (s, t)$ can be defined as

$$W_i^m(a; \alpha_i) = U_i(s(a); \alpha_i) + t_i(a).$$

Actually, we shall restrict our attention to transfer schemes t which may be called *budget-balancing* in the sense that

$$\forall\, a \in A,\ \sum_{i \in N} t_i(a) = 0.$$

Also, in the following, we shall be concerned by selection rules which are *efficient*, namely such that

$$\forall\, a \in A, \sum_{i \in N} U_i(s(a); a_i) = \max_{x \in X} \sum_{i \in N} U_i(x; a_i).$$

A mechanism $m = (s, t)$ is called efficient if its selection rule is efficient and budget-balancing if its transfer scheme is budget-balancing. These two properties ensure that a mechanism has to be feasible and Pareto-optimal under 'full information' (see Harris and Townsend, 1981, for a discussion of the notion of 'full-information' optimality, also sometimes called 'ex-post' optimality). Hence we shall call *Pareto-efficient* a mechanism which is both efficient and budget-balancing.

In the approach we take here, every player $i \in N$ is supposed to have only partial information about the value of all players' types. This generates a game of incomplete information. More specifically we shall associate to each player $i \in N$ a measurable space $(A_i, \mathcal{A}_i)$ and a transition probability P_i from this space $(A_i, \mathcal{A}_i)$ towards the $n-1$ product space denoted $(A_{-i}, \mathcal{A}_{-i})$, where

$$A_{-i} = \mathop{\times}_{\substack{j \in N \\ j \neq i}} A_j \text{ and } \mathcal{A}_{-i} = \bigotimes_{\substack{j \in N \\ j \neq i}} \mathcal{A}_j.$$

This transition probability, also denoted for simplicity $P_i = \{P_i(\cdot \mid \alpha_i); \alpha_i \in A_i\}$, represents the *beliefs* of player i about the other players' types conditional to his own type.[2] All transitions P_i, taken as functions, are of common knowledge. However, player i beliefs are fully known only when his type $\alpha_i \in A_i$ is also known.

For the sequel, we shall maintain all the following assumptions:

- For every $i \in N$, A_i is a compact metric space and $\mathcal{A}_i$ denotes its Borel σ-algebra.
- X is a compact metric space.
- For every $i \in N, U_i$ is a continuous function from $X \times A_i$ to $\mathbb{R}$.
- For every $i \in N, P_i$ is a continuous function from A_i to the space of probability measures over $(A_{-i}, \mathcal{A}_{-i})$, endowed with the topology of pointwise convergence. Every $P_i(\cdot \mid \alpha_i)$ is of full support.
- The space X and all functions U_i are such that there exists a continuous efficient-selection rule.

We shall call *continuous efficient mechanisms* these mechanisms $m = (s, t)$ which are continuous and where s is efficient.

One particular case, to be called the *discrete case*, arises when for *every* player $i \in N$, A_i is a finite set. Then, we define $\mathcal{A}_i$ as the set of all subsets of A_i and continuity is considered, trivially, with respect to the discrete topology. In the discrete case, every $P_i(\cdot \mid \alpha_i)$ will simply denote a function from A_{-i} to $\mathbb{R}_+$ such that $\sum_{\alpha_{-i} \in A_{-i}} P_i(\alpha_{-i} \mid \alpha_i) = 1$.

We may now introduce the game of incomplete information which can be associated to a mechanism m.

We first define, for every player $i \in N$, a set of (normalized) strategies as the set A_i^* of measurable functions (decision rules) $a_i^* : A_i \to A_i$ associating to each possible type $\alpha_i \in A_i$ a message (or revealed type) $a_i = a_i^*(\alpha_i)$. We then define, for every player $i \in N$, for every type $\alpha_i \in A_i$, and for any choice $a_{-i}^* = (a_1^*, \ldots, a_{i-1}^*, a_{i+1}^*, \ldots, a_n^*)$ of strategies by the other players, the conditional expected payoff which player i is supposed to maximize on A_i,

$$\overline{W}_i^m(a_i, a_{-i}^*; \alpha_i) = \int_{A_{-1}} W_i^m(a_i, a_{-i}^*(\alpha_{-i}); \alpha_i) P_i(d\alpha_{-i} \mid \alpha_i),$$

where, for notational convenience,

$$a_{-i}^*(\alpha_{-i}) = (a_1^*(\alpha_1), \ldots, a_{i-1}^*(\alpha_{i-1}), a_{i+1}^*(\alpha_{i+1}), \ldots, a_n^*(\alpha_n)).$$

A *Bayesian equilibrium* in such a game is a n-tuple of strategies $\overline{a}^* \in A^* = \times_{i \in N} A_i^*$ such that, for every $i \in N$ and every $\alpha_i \in A_i$,

$$\forall\, a_i \in A_i,\ \overline{W}_i^m(a_i, \overline{a}_{-i}^*; \alpha_i) \leq \overline{W}_i^m(\overline{a}_i^*(\alpha_i), \overline{a}_{-i}^*; \alpha_i).$$

We are interested here by those mechanisms which are such that every player has 'interest to reveal' the information he is privately controlling. Consider, for that matter, the *truth-telling* strategy denoted

2. The notion of a transition probability is defined in Neveu (1970). Notice that P_i may be considered as a regular version of the conditional probability given $\mathcal{A}_i$ arising from a basic probability measure over $(A, \mathcal{A}) = (\times_{i \in N} A_i, \bigotimes_{i \in N} \mathcal{A}_i)$.

$\hat{a}_i^*$ for $i \in N$ and defined by $\forall\, \alpha_i \in A_i,\ \hat{a}_i^*(\alpha_i) = \alpha_i$ (i.e., $\hat{a}_i^*$ is the identity function). Thus, a mechanism m is said to be *Bayesian Incentive Compatible* (for short BIC) if and only if the n-tuple $\hat{a}^* \in A^*$ of truth-telling strategies is a Bayesian equilibrium in the game of incomplete information associated to m.

3. Pareto-efficiency and the BIC-property

The difficulty associated to the determination of efficient BIC-mechanisms arises essentially when one requires, in addition, that the transfers cancel each other (budget-balance). In a previous paper (d'Aspremont and Gérard-Varet, 1979), we have given a 'compatibility condition' imposed on the players' beliefs which, in the discrete case (every A_i is finite), is sufficient for the existence of a Pareto-efficient BIC-mechanism.

The present section is devoted to the more general case. In that framework we first give a condition, involving both individual beliefs and payoffs, which is necessary and sufficient for the existence of a Pareto-efficient mechanism having a slightly weaker BIC-property. We finally obtain a sufficient condition, concerning the beliefs alone which coincides, for the discrete case, with the 'compatibility condition'.

We say that, for every $\varepsilon > 0$, a mechanism m satisfies the *ε-BIC-property* iff

$$\forall\, i \in N,\ \forall\, \alpha_i \in A_i,\ \forall\, a_i \in A_i,\ \overline{W}_i^m(a_i, \hat{a}_{-i}^*; \alpha_i) - \overline{W}_i^m(\alpha_i, \hat{a}_{-i}^*; \alpha_i) \leq \varepsilon,$$

i.e., in words, the n-tuple $\hat{a}^* \in A^*$ of truth-telling strategies is a *Bayesian ε-equilibrium* of the game associated to m.

In order to investigate the possibility of finding Pareto-efficient ε-BIC-mechanisms let us introduce some more notation. Let $C(A)$ be the space of real-valued continuous function on A. Let T denote the space $C(A)^n$ of continuous functions from A to $\mathbb{R}^n$ and $\tilde{T} = \{t \in T; \sum_{i \in N} t_i(\cdot) \equiv 0\}$. Let, for every $i \in N$, G_i be the linear application from $C(A)$ to $\mathbb{R}^{A_i^2}$ defined by

$$\forall\, z \in C(A),\ G_i(z)(a_i, \alpha_i) \overset{\text{def}}{=} \int_{A_{-i}} [z(\alpha) - z(a_i, \alpha_{-i})] P_i(d\alpha_{-i} \mid \alpha_i).$$

To interpret this definition, suppose $z \in C(A)$ is a 'payoff function' of some player i, defined on the space A of all types. Then the quantity $G_i(z)(a_i, \alpha_i)$ appears as an expected gain (or loss) obtained by player i *ex ante* for deviating from the situation where all players use their truth-telling strategies: this expected gain results from his announcing of $a_i \in A_i$, while being of type α_i. Thus it may be called player i's *expected gain from (unilateral) deviation* at (a_i, α_i) w.r.t. z.

Now, for any selection rule s and $i \in N$, we denote by

$$\overline{U}_i^s(a_i, \alpha_i) \overset{\text{def}}{=} \int_{A_{-i}} \big[U_i(s(a_i, \alpha_i); \alpha_i) - U_i(s(\alpha); \alpha_i)\big] P_i(d\alpha_{-i} \mid \alpha_i).$$

With this notation the problem of finding a continuous Pareto-efficient ε-BIC mechanism amounts to find a solution $t \in \tilde{T}$ to the following system of linear inequalities:

$$\forall\, i \in N,\ \forall\, \alpha_i \in A_i,\ \forall\, a_i \in A_i,\ a_i \neq \alpha_i,\ G_i(t_i)(a_i, \alpha_i) > \overline{U}_i^s(a_i, \alpha_i) - \varepsilon, \qquad \text{(I)}$$

where s is assumed to be efficient and continuous.

Let Λ denote the set of vectors of finite Borel-measures $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n) \neq 0$, where each λ_i is a measure on $(A_i \times A_i, \mathcal{A}_i \otimes \mathcal{A}_i$. Then we can state:

Theorem 1 *A necessary and sufficient condition for the existence of a continuous Pareto-efficient ε-BIC-mechanism is:* $\forall \lambda \in \Lambda$, *if*

$$\forall z \in C(A),\ \forall i,j \in N,\ \int_{A_i^2} G_i(z)(a_i,\alpha_i)\lambda_i(da_i,d\alpha_i) = \int_{A_j^2} G_j(z)(a_j,\alpha_j)\lambda_j(da_j,d\alpha_j), \quad (1)$$

then

$$\sum_{i\in N}\int_{A_i^2}\left[\overline{U}_i^s(a_i,\alpha_i) - \varepsilon\right]\lambda_i(da_i,d\alpha_i) < 0. \quad (2)$$

Proof We proceed in two steps:

Step 1 First we can show that if, for any $z \in C(A)$ and $i \in N$, $G_i(z)$ is a continuous function defined on A_i^2 then the result derives directly from the necessary and sufficient condition given in Theorem 15 in Fan (1956). Indeed letting Y denote the product space $\times_{i\in N} C(A_i^2)$ with $\|y\| = \max_{N\times A_i^2}|y_i(a_i,\alpha_i|$, $y \in Y$? and $\|t\| = \max_{N\times A}|t_i(\alpha)|$, $t \in T$, the application $G = (G_1,\ldots,G_i,\ldots,G_n)$ from $\tilde{T}$ to Y is continuous and the set $\operatorname{Int} Q \stackrel{\text{def}}{=} \operatorname{Int}\{y \in Y : y \geq 0\} \neq \emptyset$. So Ky Fan's condition may be written: $\forall g \neq 0$ belonging to the conjugate convex cone Q^* of Q (i.e., $Q^* = \{g \in Y^* : \forall y \in Q, g(y) \geq 0\}$) if

(a) $\forall t \in \tilde{T},\ g(G(t)) = 0$, then

(b) $g(\overline{U}^s - \overline{\varepsilon}) < 0$ for $(\overline{U}^s - \overline{\varepsilon}) \stackrel{\text{def}}{=} (\overline{U}_1^s + \varepsilon,\ldots,\overline{U}_i^s - \varepsilon,\ldots,\overline{U}_n^s - \varepsilon)$. Now $\forall y \in Y, \forall g \in Q^*$, $g(y) = \sum_{i\in N} g_i(y_i)$, where every g_i is the bounded linear functional defined on $C(A_i^2)$ by $g_i(y_i) = g(0,\ldots,y_i,\ldots,0)$. Clearly $g_i(y_i) \geq 0$ if $y_i \geq 0$. Hence, by Riesz representation theorem, to every such g_i there corresponds a unique finite Borel measure λ_i on A_i^2 such that

(c) $\forall y_i \in A_i,\ g_i(y_i) = \int_{A_i^2} y_i(a_i,\alpha_i)\lambda_i(da_i,d\alpha_i)$.

It follows immediately that (b) is equivalent to (2). To get the equivalence of (a) and (1), suppose that (1) holds for some $\lambda \in \Lambda$. Then for any $t \in \tilde{T}$ and any $i \in N$,

$$\int_{A_i^2} G_i(t_i)(a_i,\alpha_i)\lambda_i(da_i,d\alpha_i) = \int_{A_i^2} G_1(t_i)(a_1,\alpha_1)\lambda_1(da_1,d\alpha_1),$$

and hence we may associate some $g \in Q^*$, $g \neq 0$, such that g satisfies (c) and

$$\begin{aligned}\forall t \in \tilde{T},\ g(G(t)) &= \sum_{i\in N}\int_{A_i^2} G_i(t_i)(a_i,\alpha_i)\lambda_i(da_i,d\alpha_i)\\ &= \int_{A_1^2} G_1\left(\sum_{i\in N} t_i\right)(a_1,\alpha_1)\lambda_1(da_1,d\alpha_1) = 0.\end{aligned}$$

To prove the converse take any $g \in Q^*$, $g \neq 0$, such that, $\forall t \in \tilde{T}$, $g(G(t)) = 0$ and take any $z \in C(A)$. If (a) holds we may then define, for any pair $\{i,j\} \subset N$, some $\hat{t} \in \tilde{T}$ such that $\hat{t}_i = z, \hat{t}_j = -z$ and $\hat{t}_k \equiv 0$, for $k \notin \{i,j\}$, and thus get

$$\begin{aligned} 0 &= g(G(\hat{t})) = g_i(G_i(z)) - g_j(G_j(z))\\ &= \int_{A_i^2} G_i(z)(a_i,\alpha_i)\lambda_i(da_i,d\alpha_i) - \int_{A_i^2} G_j(z)(a_j,\alpha_j)\lambda_j(da_j,d\alpha_j),\end{aligned}$$

where $\lambda \in \Lambda$ is defined according to (c).

Therefore the two conditions are equivalent and to prove the result it remains only to show the following.

Step 2 $\forall\, i \in N,\ \forall\, z \in C(A),\ G_i(z) \in C(A_i^2)$.

Take any $z \in C(A)$. By the definition of G_i, it is enough to show that both $\int_{A_{-i}} z(\alpha_i, \alpha_{-i}) P_i(d\alpha_{-i} \mid \alpha_i)$ and $\int_{A_{-i}} z(a_i, \alpha_{-i}) P_i(d\alpha_{-i} \mid \alpha_i)$ belong to $C(A_i^2)$. Since A_i^2 is compact z is the limit of sums of the form $\sum_{k=1}^{K} f_K^k(\alpha_i) g_K^k(\alpha_{-i})$, where every $f_K^k \in C(A_i)$ and every $g_K^k \in C(A_{-i})$. Take any sequence (a_i^m, α_i^m) in A_i^2 which is convergent to (a_i^0, α_i^0). Clearly, by assumption or definition, we have that

$$\lim_{m\to\infty} P_i(\cdot \mid \alpha_i^m) = P_i(\cdot \mid \alpha_i^0),\ \lim_{m\to\infty} f_K^k(a_i^m) = f_K^k(a_i^0), \lim_{m\to\infty} f_K^k(\alpha_i^m) = f_K^k(\alpha_i^0).$$

Hence, by a general convergence argument (see e.g. Proposition 18 in Royden, 1968, p. 232)

$$\lim_{m\to\infty} \int_{A_{-i}} g_K^k(\alpha_{-i}) P_i(d\alpha_{-i} \mid \alpha_i^m) = \int_{A_{-i}} g_K^k(\alpha_{-i}) P_i(d\alpha_{-i} \mid \alpha_i^0),$$

and, by Lebesgue convergence theorem,

$$\begin{aligned}\forall\, \alpha_i \in A_i, \quad & \int_{A_{-i}} \left[\lim_{K\to\infty} \sum_{k=1}^{K} f_K^k(\alpha_i) g_K^k(\alpha_{-i})\right] P_i(d\alpha_{-i} \mid \alpha_i) \\ &= \lim_{K\to\infty} \int_{A_{-i}} \left[\sum_{k=1}^{K} f_K^k(\alpha_i) g_K^k(\alpha_{-i})\right] P_i(d\alpha_{-i} \mid \alpha_i).\end{aligned}$$

Finally, using properties of limits, we get

$$\begin{aligned}&\lim_{m\to\infty} \int_{A_{-i}} z(\alpha_i^m, \alpha_{-i}) P_i(d\alpha_{-i} \mid \alpha_i^m) \\ &= \lim_{m\to\infty} \lim_{K\to\infty} \sum_{k=1}^{K} \int_{A_{-i}} f_K^k(\alpha_i^m) g_K^k(\alpha_{-i}) P_i(d\alpha_{-i} \mid \alpha_i^m) \\ &= \lim_{K\to\infty} \sum_{k=1}^{K} \lim_{m\to\infty} \left[f_K^k(\alpha_i^m) \int_{A_{-i}} g_K^k(\alpha_{-i}) P_i(d\alpha_{-i} \mid \alpha_i^m)\right] \\ &= \lim_{K\to\infty} \sum_{k=1}^{K} f_K^k(\alpha_i^0) \int_{A_{-i}} g_K^k(\alpha_{-i}) P_i(d\alpha_{-i} \mid \alpha_i^0) \\ &= \int_{A_{-i}} z(\alpha_i^0, \alpha_{-i}) P_i(d\alpha_{-i} \mid \alpha_i^0).\end{aligned}$$

Similarly,

$$\begin{aligned}\lim_{m\to\infty} \int_{A_{-i}} z(a_i^m, \alpha_{-i}) P_i(d\alpha_{-i} \mid \alpha_i^m) &= \lim_{k\to\infty} \sum_{k=1}^{K} f_K^k(a_i^0) \int_{A_{-i}} g_K^k(\alpha_{-i}) P_i(d\alpha_{-i} \mid \alpha_i^0) \\ &= \int_{A_{-i}} z(a_i^0, \alpha_{-i}) P_i(d\alpha_{-i} \mid \alpha_i^0).\end{aligned}$$

■

An unsatisfactory feature of the necessary and sufficient condition stated in this theorem is that it involves not only the given transition probabilities but also the given payoff functions. It is a joint condition on the beliefs and on the utilities of the players. Our purpose now is to state another condition which will concern only the beliefs of the players. This seems to be more conform to the classical problem of preference revelation. We shall see later that this condition is equivalent, in the discrete case, to the 'compatibility condition' stated in d'Aspremont and Gérard-Varet (1979).

Condition C $\forall\, \lambda \in \Lambda$, *if*

$$\begin{aligned} &\forall\, z \in C(A),\ \forall\, i \in N,\ \forall\, j \in N, \\ &\int_{A_i^2} G_i(z)(a_i, \alpha_i)\lambda_i(da_i, d\alpha_i) = \int_{A_i^2} G_j(z)(a_j, \alpha_j)\lambda_j(da_j, d\alpha_j), \end{aligned} \tag{3}$$

then

$$\forall\, z \in C(A),\ \forall\, i \in N,\ \int_{A_i^2} G_i(z)(a_i, \alpha_i)\lambda_i(da_i, d\alpha_i) = 0. \tag{4}$$

To enunciate Condition C verbally, interpret $\lambda = (\lambda_1, \ldots, \lambda_i, \ldots, \lambda_n)$ as being a system of individual scaling rules where λ_i gives weights which are attributed *ex ante* to player i for every value of his private information and declaration. Thus $\int_{A_i^2} G_i(z)(a_i, \alpha_i)\lambda_i(da_i, d\alpha_i)$ is player i weighted sum, under the system λ, of all possible expected gains from deviation w.r.t. $z \in C(A)$. Condition C says that if for some system $\lambda \in \Lambda$ of individual scaling rules the weighted sums of all possible expected gains from deviation are made comparable and identical, across payoff functions $z \in C(A)$ and players, then they should be identically zero.

Assume that all players met before knowing of any mechanism and before having been privately informed about their own type. They plan to cooperate against the mechanism. Condition C says that there is no room for such cooperation in the sense that there does not exist a system $\lambda \in \Lambda$ of individual scaling rules making the weighted sum of conditional expected gains from deviation of every player both comparable and equal to some non-zero-amount.

In order to show that Condition C is sufficient to give the existence of a solution $t \in \tilde{T}$ to system (I), we need the following lemma where a continuous efficient BIC-mechanism is constructed:

Lemma 1 *Let s be a continuous efficient selection rule.* $\exists\, t^* \in T$ *such that*

$$\forall\, i \in N,\ \forall\, a_i \in A_i,\ \forall\, \alpha_i \in A_i,\ G_i(t_i)(a_i, \alpha_i) \geq \overline{U}_i^s(a_i, \alpha_i).$$

Proof Define

$$\forall\, i \in N, \forall\, a \in A,\ t_i^*(a) = \sum_{j \neq i} U_j(s(a); a_j).$$

Since every U_i and s are continuous functions, $t^* \in T$. Moreover by the efficiency of s,

$$\begin{aligned} &\forall\, i \in N,\ \forall\, a_i \in A_i,\ \forall\, \alpha_i \in A_i,\ \forall\, \alpha_{-i} \in A_{-i}, \\ &U_i(s(\alpha); \alpha_i) + t_i^*(\alpha) \geq U_i(s(a_i, \alpha_{-i}); \alpha_i) + t_i^*(a_i, \alpha_{-i}). \end{aligned}$$

Integrating on both sides the result follows. ■

With this lemma we can state and prove the following theorem:

Theorem 2 *If the given beliefs $\{P_i, i \in N\}$ satisfy Condition C, then there exists a continuous, Pareto-efficient ε-BIC-mechanism.*

Proof We show that, with Lemma 1, Condition C implies the necessary and sufficient condition of Theorem 1. First notice that this last condition has the same antecedent as Condition C. Also, by Lemma 1, we have, for any $\lambda \in \Lambda$, if

$$\forall\, t \in T, \sum_{i \in N} \int_{A_i^2} G_i(t_i)(a_i, \alpha_i)\lambda_i(da_i, d\alpha_i) = 0, \tag{5}$$

then

$$\sum_{i \in N} \int_{A_i^2} \overline{U}_i^s(a_i, \alpha_i)\lambda_i(da_i, d\alpha_i) \leq 0. \tag{6}$$

Indeed, for $t^* \in T$ as defined in Lemma 1, we get

$$0 = \sum_{i \in N} \int_{A_i^2} \int_{A_i} G_i(t_i^*)(a_i, \alpha_i)\lambda_i(da_i, d\alpha_i) \geq \sum_{i \in N} \int_{A_i^2} \overline{U}_i^s(a_i, \alpha_i)\lambda_i(da_i, d\alpha_i).$$

Clearly (5) above is equivalent to the consequent (4) of Condition C. Therefore (6) must hold, which implies the consequent (2) of the condition in Theorem 1, since the vector measure λ is different from zero and $\varepsilon > 0$. The result follows. ■

We shall now show that Condition C is equivalent to a statement about individual beliefs. For that let $\mathcal{A}^0 = \{E \in \mathcal{A}; E = \bigtimes_{i \times N} E_i \text{ and } \forall\, i \in N, E_i \in \mathcal{A}_i\}$ be the semi-algebra of rectangles in $\mathcal{A}$. Then we show:

Theorem 3 *A family of beliefs $\{P_i, i \in N\}$ satisfies Condition C if and only if $\forall\, \lambda \in \Lambda$, if*

$$\begin{aligned} &\forall\, E \in \mathcal{A}^0,\ \forall\, i \in N,\ \forall\, j \in N, \\ &\int_{A_i \times E_i} P_i(E_{-i} \mid \alpha_i)\lambda_i(da_i, d\alpha_i) - \int_{E_i \times A_i} P_i(E_i \mid \alpha_i)\lambda_i(da_i, d\alpha_i) \\ =&\int_{A_j \times E_j} P_j(E_{-j} \mid \alpha_j)\lambda_j(da_j, d\alpha_j) - \int_{E_j \times A_j} P_j(E_{-j} \mid \alpha_j)\lambda_j(da_j, d\alpha_j), \end{aligned} \tag{7}$$

then

$$\begin{aligned} &\forall\, E \in \mathcal{A}^0,\ \forall\, i \in N, \\ &\int_{A_i \times E_i} P_i(E_{-i} \mid \alpha_i)\lambda_i(da_i, d\alpha_i) - \int_{E_i \times A_i} P_i(E_{-i} \mid \alpha_i)\lambda_i(da_i, d\alpha_i) = 0. \end{aligned} \tag{8}$$

Proof Let, for every $i \in N$ and $\lambda \in \Lambda$, F_i^λ denote the following bounded linear functional on $C(A)$:

$$\forall\, z \in C(A),\ F_i^\lambda(z) = \int_{A_i^2} G_i(z)(a_i, \alpha_i)\lambda_i(da_i, d\alpha_i).$$

By Riesz representation theorem there corresponds a *unique* finite signed Borel measure Q_λ^i on A such that, $\forall\, z \in C(A)$, $F_i^\lambda(z) = \int_A z(\alpha)Q_i^\lambda(d\alpha)$.

Now take, for every $i \in N$ and $\lambda \in \Lambda$, the signed measure P_i^λ on $(A, \mathcal{A})$ defined by the property

$$\begin{aligned}\forall\, E \in \mathcal{A}^0,\ P_i^\lambda(E) = &\int_{A_i \times E_i} P_i(E_{-i} \mid \alpha_i)\lambda_i(da_i, d\alpha_i) \\ &- \int_{E_i \times A_i} P_i(E_{-i} \mid \alpha_i)\lambda_i(da_i, d\alpha_i).\end{aligned}$$

We see that

$$\begin{aligned}&\forall\, z \in C(A), \\ &\int_A z(\alpha) P_i^\lambda(d\alpha) = \int_{A_{-i}} \int_{A_i^2} z(\alpha) P_i(d\alpha_{-i} \mid \alpha_i)\lambda_i(da_i, d\alpha_i) \\ &\qquad - \int_{A_{-i}} \int_{A_i^2} z(a_i, \alpha_{-i}) P_i(d\alpha_{-i} \mid \alpha_i)\lambda_i(da_i, d\alpha_i) \\ &\qquad = \int_{A_i^2} G_i(z)(a_i, \alpha_i)\lambda_i(da_i, d\alpha_i) = F_i^\lambda(z) \\ &\qquad = \int_A z(\alpha) Q_i^\lambda(d\alpha),\end{aligned}$$

i.e., $P_i^\lambda = Q_i^\lambda$. The theorem follows. ■

In the discrete case, Theorem 3 shows that Condition C is equivalent to the 'compatibility condition' already mentioned.[3] Therefore by Theorem 7 in d'Aspremont and Gérard-Varet (1979), Theorem 2 can be strengthened in the discrete case by putting $\varepsilon = 0$.

4. On compatible beliefs

Our purpose in the present section is to provide a detailed analysis of the compatibility condition – Condition C – determining families of beliefs which are admissible for the existence of Pareto-efficient mechanisms satisfying the BIC-property. We first consider two conditions which are stronger than Condition C, but such that each one of them has a more specific interpretation. Second, we study connections between Condition C and another restriction upon beliefs known in the literature as the 'consistency' condition. In that case a weaker budget-balancing property can be defined and a stronger positive result can be obtained. Last, we show that Condition C is not necessary by exhibiting a family of discrete beliefs which do not fulfill the requirement and however is such that, whatever may be the players' payoffs (under the full-transferability assumption), there exist efficient budget-balancing BIC-mechanisms.

Take the beliefs of a particular player, say player $i \in N$. We say that player i beliefs are free of any dependency with respect to his own types or, for short, that he has *free beliefs* iff[4]

$$\forall\, \alpha_i \in A_i,\ \forall\, \overline{\alpha}_i \in A_i,\ P_i(\cdot \mid \alpha_i) = P_i(\cdot \mid \overline{\alpha}_i).$$

3. One may provide for the statement of Theorem 3 an interpretation similar to the one given for Condition C. Indeed, in a Bayesian set-up, the beliefs P_i give the bets $P_i(E_{-i} \mid \cdot)$ of player $i \in N$ on events $E_{-i} \in \mathcal{A}_{-i}$ concerning the others' information type, conditional to the player's own information $\mathcal{A}_i$.

4. Thus, $\mathcal{A}_{-i}$ and $\mathcal{A}_i$ are σ-algebras which are independent with respect to the probability over $(A, \mathcal{A})$ generating P_i.

In terms of information, free beliefs mean that the true beliefs of player $i \in N$ are common knowledge.

Consider now the whole family $\{P_i; i \in N\}$ of beliefs. We state:

Condition F *There exists at least one player who has free beliefs.*

We shall now see that Condition F is sufficient for having Condition C, and, as a consequence, the result of Theorem 2. We first prove:

Lemma 2 *Given a family of beliefs $\{P_i; i \in N\}$, for every $\lambda \in \Lambda$, if*

$$\begin{aligned}
&\forall\, i \in N,\ \forall\, j \in N,\ \forall\, E \in \mathcal{A}^0,\\
&\int_{A_i \times E_i} P_i(E_{-i} \mid \alpha_i)\lambda_i(da_i, d\alpha_i) - \int_{E_i \times A_i} P_i(E_{-i} \mid \alpha_i)\lambda_i(da_i, d\alpha_i)\\
=&\int_{A_j \times E_j} P_j(E_{-j} \mid \alpha_j)\lambda_j(da_j, d\alpha_j) - \int_{E_j \times A_j} P_j(E_{-j} \mid \alpha_j)\lambda_j(da_j, d\alpha_j)
\end{aligned}$$

then

$$\forall\, i \in N,\ \forall\, E_i \in \mathcal{A}_i,\ \lambda_i(E_i \times A_i) = \lambda_i(A_i \times E_i).$$

Proof Take any $i \in N$ and consider $E = E_i \times A_{-i}$ in $\mathcal{A}^0$ in the antecedent. We get, for $j \neq i$,

$$\begin{aligned}
\lambda_i(A_i \times E_i) - \lambda_i(E_i \times A_i) = &\int_{A_j \times A_j} P_j(E_{-j} \mid \alpha_j)\lambda_j(da_j, d\alpha_j)\\
&- \int_{A_j \times A_j} P_j(E_{-j} \mid \alpha_j)\lambda_j(da_j, d\alpha_j) = 0.
\end{aligned}$$

■

Theorem 4 *Given a family of beliefs $\{P_i; i \in N\}$, Condition F implies Condition C.*

Proof Suppose w.l.o.g. that player 1 has free beliefs, i.e., $\forall\, \alpha_1, \forall\, \overline{\alpha}_1, P_1(\cdot \mid \alpha_1) = P_1(\cdot \mid \overline{\alpha}_1) = \pi_1(\cdot)$ and that Condition C does not hold. Then, by Theorem 3, there exist $\lambda \in \Lambda$ and $E \in \mathcal{A}^0$, such that $\forall\, j \in N - \{1\}$,

$$\begin{aligned}
&\pi_1(E_{-1})[\lambda_1(A_1 \times E_1) - \lambda_1(E_1 \times A_1)]\\
=&\int_{A_j \times E_j} P_j(E_{-j} \mid \alpha_j)\lambda_j(da_j, d\alpha_j) - \int_{E_j \times A_j} P_j(E_{-j} \mid \alpha_j)\lambda_j(da_j, d\alpha_j) \neq 0.
\end{aligned}$$

This is a contradiction to Lemma 2. ■

Since Condition F implies Condition C, its only advantage is that it provides an interpretation for an important subcase, in which the beliefs of at least one player are fully known. An even stronger condition is to require that *every* player has free beliefs. We shall call it *Condition F** (it is called the independence condition in d'Aspremont and Gérard-Varet, 1979). In this case we obtain a stronger result, namely that there exists a Pareto-efficient BIC-mechanism. However other interpretable subcases may be considered.

An n-person normal form game over the spaces $\{A_i; i \in N\}$, taken as strategy spaces, is a function ω from A to $\mathbb{R}^n$ where ω_i is the payoff function of player $i \in N$. The game is *zero-sum* if $\sum_{i \in N} \omega_i \equiv 0$. Moreover we say that the family $\{P_i; i \in N\}$ provides a *strict correlated equilibrium* for the game iff

$$\forall\, i \in N,\ \forall\, \alpha_i \in A_i,\ \forall\, a_i \in A_i,\ a_i \neq \alpha_i,$$
$$\int_{A_{-i}} \omega_i(a_i, \alpha_{-i}) P_i(d\alpha_{-i} \mid \alpha_i) < \int_{A_{-i}} \omega_i(\alpha_i, \alpha_{-i}) P_i(d\alpha_{-i} \mid \alpha_i).$$

This non-cooperative equilibrium notion has to be considered as referring to an underlying random mechanism designed by the players for selecting their strategies. Thus the P_i's represent the beliefs of the players about the outcome of this mechanism (see Aumann, 1974).

We now consider the following new condition relative to the family of beliefs $\{P_i; i \in N\}$. In order to match Condition C, we give a continuous version of this condition:

Condition B *There exists a zero-sum game, $\{(A_i, \omega_i); i \in N\}$, with continuous payoff functions and for which the family of beliefs $\{P_i, i \in N\}$ provides a strict correlated equilibrium.*
The statement of Condition B may be written as follows:

$$\exists\, \omega \in \tilde{T},$$

such that

$$\forall\, i \in N,\ \forall\, \alpha_i \in A_i,\ \forall\, a_i \in A_i,\ G_i(\omega_i)(a_i, \alpha_i) > 0.$$

Consider on the other hand any family of payoff functions $\{U_i; i \in N\}$ and take an efficient selection rule s. Define, as in the beginning of Section 3, the numbers $\overline{U}_i^s(a_i, \alpha_i)$, $i \in N$, $a_i \in A_i$, $\alpha_i \in A_i$. From Condition B, each ω_i being multiplied by a constant correctly chosen, one easily deduce the solvability of system (I) with $\varepsilon = 0$, i.e.

$$\exists\, t \in \tilde{T},$$

such that

$$\forall\, i \in N,\ \forall\, \alpha_i \in A_i,\ \forall\, a_i \in A_i,\ G_i(t_i)(a, \alpha_i) \geq \overline{U}^s(a_i, \alpha_i).$$

Thus, we prove:

Theorem 5 *Assuming Condition B, there exists a (continuous) efficient and budget-balancing BIC-mechanism.*

One interpretation of Theorem 5 is that the individual beliefs characterized by Condition B must introduce sufficient 'disagreements' between the players in order to make 'correlation' worthwhile in a game having a strong competitive character [5]

Comparisons of Condition B with Conditions F, F* and C, are based on the following characterization:

Lemma 3 *Given $\{P_i; i \in N\}$, Condition B holds if and only if there is no $\lambda \in \Lambda$, such that*

$$\forall\, z \in C(A),\ \forall\, i \in N, \forall\, j \in N,$$
$$\int_{A_i^2} G_i(z)(a_i, \alpha_i) \lambda_i(da_i, d\alpha_i) = \int_{A_i^2} G_j(z)(a_j, \alpha_j) \lambda_j(da_j, d\alpha_j).$$

5. See (Aumann, 1974, p .67, and example 2.1, p. 68). This property captures features similar to those which are noticed in an example given in (Laffont and Maskin, 1979, p. 307).

Proof Using, as in Theorem 1, Fan (1956), a necessary and sufficient condition for solving system (I), in which we take $\overline{U}_i^s(\cdot) - \varepsilon$ identically null, is simply: there is no $g \neq 0$ in Q^* such that, $\forall\, t \in \tilde{T}$, $g(G(t)) = 0$. By the same reasoning as in Theorem 1 we get the result. ■

Theorem 6 *Given a family of beliefs $\{P_i; i \in N\}$, if Condition B holds, then*

(i) Condition C holds.

(ii) Condition F does not hold.

Proof (i) is immediate from Lemma 3. To prove (ii), suppose w.l.o.g. that player 1 has free beliefs represented by π_1. Then Condition B implies

$$\exists\, \omega \in \tilde{T},$$

such that

$$\forall\, \alpha_i \in A_i,\ \forall\, a_i \in A_i,\ \int_{A_{-i}} \omega_i(\alpha_i, \alpha_{-i})\pi(d\alpha_{-i}) > \int_{A_{-i}} \omega_i(a_i, \alpha_{-i})\pi(d\alpha_{-i}),$$

which forms a set of contradictory inequalities. ■

Theorem 6 shows that the beliefs satisfying Condition F and those satisfying Condition B are two distinct subclasses of those satisfying Condition C. However, the size of the class of beliefs satisfying C but neither B nor F, remains an open question.

Our next objective is to compare Condition C with another condition known as 'consistency' (see, for example Harsanyi, 1967-68, for a detailed discussion of this condition). Let, for that matter, $\mathcal{M}_i$ be the set of all probability measures over $(A_i, \mathcal{A}_i)$ and $\mathcal{M}$ be the set of all probability measures over $(A, \mathcal{A})$. We say that $\{P_i; i \in N\}$ is a *consistent family of beliefs* if

$$\exists\, \nu \in \textstyle\times_{i \in N} \mathcal{M}_i,\ \exists\, R \in \mathcal{M},\ R \neq 0,$$

such that

$$\forall\, E \in \mathcal{A}^0,\ \forall\, i \in N,\ \int_{E_i} P_i(E_{-i} \mid \alpha_i)\nu_i(d\alpha_i) = R(E).$$

Thus the P_i's are, for the respective players, the different conditionalizations of the same joint probability.

Our first result is for the two player case:

Theorem 7 *Assuming $N = \{1, 2\}$ and given $\{P_1, P_2\}$, a pair of consistent beliefs, Condition C is equivalent to Condition F*.*

Proof Given $\{P_1, P_2\}$, by consistency we have

$$\exists \nu_1 \in \mathcal{M}_1,\ \exists\, \nu_2 \in \mathcal{M}_2,\ \exists\, R \in \mathcal{M},\ R \neq 0,$$

such that

$$\forall\, E \in \mathcal{A}^0,\ \forall\, i, j \in \{1, 2\},\ i \neq j,\ \int_{E_i} P_i(E_j \mid \alpha_i)\nu_i(d\alpha_i) = R(E).$$

Also, the negotiation of Condition C is, by Theorem 3, equivalent to

$$\exists\, \lambda \in \Lambda,$$

such that

$$\int_{A_i \times E_i} P_i(E_j \mid \alpha_i)\lambda_i(da_i, d\alpha_i) - \int_{E_i \times A_i} P_i(E_j \mid \alpha_i)\lambda_i(da_i, d\alpha_i) \stackrel{\text{def}}{=} \kappa(E)$$

is equal for both i and all $E \in \mathcal{A}^0$ and different from zero for some $\overline{E}$ in $\mathcal{A}^0$.

Define $\lambda \in \Lambda$ to be

$$\forall\, i \in \{1, 2\},\ \forall\, E_i, F_i \in \mathcal{A}_i,\ \lambda_i(E_i \times F_i) = \nu_i(E_i) \cdot \nu_i(F_i).$$

Thus,

$$\begin{aligned}
&\forall\, E \in \mathcal{A}^0,\ \forall\, i \in \{1, 2\},\\
&\int_{A_i \times E_i} P_i(E_j \mid \alpha_i)\lambda_i(da_i, d\alpha_i) - \int_{E_i \times A_i} P_i(E_j \mid \alpha_i)\lambda_i(da_i, d\alpha_i)\\
=&\int_{E_i} P_i(E_j \mid \alpha_i)\nu_i(d\alpha_i) - \left[\int_{A_i} P_i(E_j \mid \alpha_i)\nu_i(d\alpha_i)\right]\left[\int_{E_i} \nu_i(da_i)\right]\\
=&R(E) - R(A_i \times E_j) \cdot \nu_i(E_i)\\
=&R(E) - \nu_j(E_j) \cdot \nu_i(E_i).
\end{aligned}$$

If Condition F^* does not hold ten it is well known that $R(E) - \nu_1(E_1) \cdot \nu_2(E_2) \neq 0$ for some $E \in \mathcal{A}^0$, and so Condition C does not hold either. Theorem 4 implies the other direction. ∎

However, this negative result only holds in the case of two players, as the following example demonstrates:

Theorem 8 *For $|N| > 2$, there exist consistent families of beliefs which satisfy Condition C and not Condition F^*.*

Proof According to Theorem 2, it is sufficient to find a consistent family of beliefs for which only one player has free beliefs (i.e., satisfying Condition F).

Consider $N = \{1, 2, 3\}$, $A_1 = \{\alpha_1, \alpha_2\}$, $A_2 = \{\beta_1, \beta_2\}$, $A_3 = \{\gamma_1, \gamma_2\}$, the following beliefs satisfy the requirements:

For player 1, P_1

	$\beta_1\gamma_1$	$\beta_2\gamma_1$	$\beta_1\gamma_2$	$\beta_2\gamma_2$
α_1	4/15	1/15	8/15	2/15
α_2	1/9	2/9	2/9	4/9

For player 2, P_2

	$\alpha_1\gamma_1$	$\alpha_2\gamma_1$	$\alpha_1\gamma_2$	$\alpha_2\gamma_2$
β_1	4/15	1/15	8/15	2/15
β_2	1/9	2/9	2/9	4/9

For player 3, P_3

	$\alpha_1\beta_1$	$\alpha_1\beta_2$	$\alpha_2\beta_1$	$\alpha_2\beta_2$
γ_1	12/24	3/24	3/24	6/24
γ_2	12/24	3/24	3/24	6/24

■

As a last example, we shall take the simplest case of two players, each having only two types, i.e.,

$$N = \{1, 2\}, \;\; A_1 = \{\alpha_1, \alpha_2\}, \;\; A_2 = \{\beta_1, \beta_2\}.$$

We shall, for this case, consider the following class of beliefs, choosing $\delta \neq \frac{1}{2}$, $0 \leq \delta \leq 1$:

P_1

	β_1	β_2
α_1	δ	$1-\delta$
α_2	$1-\delta$	δ

P_2

	α_1	α_2
β_1	δ	$1-\delta$
β_2	$1-\delta$	δ

Notice that *none* of these beliefs are free and that it is a *consistent* family of beliefs. Consider now any set X of outcomes and any family of two payoff functions $\{U_1, U_2\}$ satisfying the side-payment full-transferability assumption. Take any selection rule s which is efficient. As in the proof of Theorem 1, finding a balanced transfer scheme $t \in \mathbb{R}^{N \times A}$, such that $m(s, t)$ is a BIC-mechanism, consists in showing that, $\forall (\lambda_1(\alpha_1, \alpha_2),\ \lambda_1(\alpha_2, \alpha_1),\ \lambda_2(\beta_1, \beta_2), \lambda_2(\beta_2, \beta_1)) \in \mathbb{R}^4_+ - \{0\}$, if $\forall\, t \in \tilde{T}$,

(i) $$\lambda_1(\alpha_1, \alpha_2) \sum_{b \in A_2} [t_1(\alpha_2, b) - t_1(\alpha_1, b)] P_1(b \mid \alpha_2)$$
$$+\lambda_1(\alpha_2, \alpha_1) \sum_{b \in A_2} [t_1(\alpha_1, b) - t_1(\alpha_2, b)] P_1(b \mid \alpha_1)$$
$$+\lambda_2(\beta_1, \beta_2) \sum_{a \in A_1} [t_2(a, \beta_2) - t_2(a, \beta_1)] P_2(a \mid \beta_2)$$
$$+\lambda_2(\beta_2, \beta_1) \sum_{a \in A_1} [t_2(a, \beta_1) - t_2(a, \beta_2) P_2(a \mid \beta_1) = 0,$$

then

(ii) $\lambda_1(\alpha_1,\alpha_2)\overline{U}_1^s(\alpha_1,\alpha_2)+\lambda_2(\alpha_2,\alpha_1)\overline{U}_1^s(\alpha_2,\alpha_1)$
$+\lambda_2(\beta_1,\beta_2)\overline{U}_2^s(\beta_1,\beta_2)+\lambda_2(\beta_2,\beta_1)\overline{U}_2^s(\beta_2,\beta_1)\geq 0.$

Now, assuming (i), we get

$$\lambda_1(\alpha_2,\alpha_1)=\lambda_1(\alpha_1,\alpha_2)=\lambda_1, \qquad \lambda_2(\beta_2,\beta_1)=\lambda_2(\beta_1,\beta_2)=\lambda_2,$$

and

$$[P_1(\beta_1\mid\alpha_1)-P_1(\beta_1\mid\alpha_2)]\lambda_1=[P_2(\alpha_1\mid\beta_1)-P_2(\alpha_1\mid\beta_2)]\lambda_2.$$

But, for the given family of beliefs, we have

$$P_1(\beta_1\mid\alpha_1)-P_1(\beta_1\mid\alpha_2)=P_2(\alpha_1\mid\beta_1)-P_2(\alpha_1\mid\beta_2)=2\delta-1,$$

implying

(iii) $\lambda_1=\lambda_2=\lambda_0.$

Now, given (iii) and the assumed family of beliefs, we have in (ii)

$$\begin{aligned}
&\lambda_0[\overline{U}_1^s(\alpha_1,\alpha_2)+\overline{U}_1^s(\alpha_2,\alpha_1)+\overline{U}_2^s(\beta_1,\beta_2)+\overline{U}_2^s(\beta_2,\beta_1)]\\
=&(1-\delta)\lambda_0[U_1(s(\alpha_1,\beta_1),\alpha_2)+U_1(s(\alpha_2,\beta_2),\alpha_1)+U_2(s(\alpha_1,\beta_1),\beta_2)\\
&+U_2(s(\alpha_2,\beta_2),\beta_1)-U_1(s(\alpha_2,\beta_1),\alpha_2)-U_1(s(\alpha_1,\beta_2),\alpha_1)\\
&-U_2(s(\alpha_1,\beta_2),\beta_2)-U_2(s(\alpha_2,\beta_1),\beta_1)]\\
&+\delta\lambda_0[U_1(s(\alpha_1,\beta_2),\alpha_2)+U_1(s(\alpha_2,\beta_1),\alpha_1)+U_2(s(\alpha_2,\beta_1),\beta_2)\\
&+U_2(s(\alpha_1,\beta_2),\beta_1)-U_1(s(\alpha_2,\beta_2),\alpha_2)-U_1(s(\alpha_1,\beta_1),\alpha_1)\\
&-U_2(s(\alpha_2,\beta_2),\beta_2)-U_2(s(\alpha_1,\beta_1),\beta_1)].
\end{aligned}$$

By applying twice the efficiency property, this is smaller than or equal to

$$\begin{aligned}
&(1-\delta)\lambda_0[U_1(s(\alpha_2,\beta_2),\alpha_1)+U_2(s(\alpha_2,\beta_2),\beta_1)\\
&-U_1(s(\alpha_1,\beta_1),\alpha_1)-U_2(s(\alpha_1,\beta_1),\beta_1)]\\
&+\delta\lambda_0[U_1(s(\alpha_1,\beta_2),\alpha_2)+U_2(s(\alpha_1,\beta_2),\beta_1)\\
&-U_1(s(\alpha_2,\beta_1),\alpha_2)-U_2(s(\alpha_2,\beta_1),\beta_1)]\leq 0.
\end{aligned}$$

Thus, there exists, for the given family of beliefs, an efficient budget-balancing mechanism which is BIC. Also, since $\delta\neq\frac{1}{2}$, Condition C does not hold (see Theorem 3). Therefore, we may state:

Theorem 9 *Condition C is not necessary for the existence of an efficient budget-balancing mechanism which is BIC.*

The last situation which we shall consider using the same approach is the one where we keep the consistency condition but without requiring Condition C. In this case we can only find efficient-BIC-mechanisms satisfying a weak budget-balancing property, namely,

$$\int_A\left[\sum_{i\in N}t_i(\alpha)\right]R(d\alpha)=0.$$

For the sake of simplicity we shall consider this new property, called the *expected-budget-balancing* property, only in the discrete case:

Theorem 10 *In the discrete case and for any consistent family of beliefs, there exists an efficient BIC-mechanism which satisfies the expected-budget-balancing property.*

Proof of Theorem 9 We have to show the existence of $t \in T$ solving the following system of inequalities:

$$\forall\, i \in N,\ \forall\, \alpha_i \in A_i,\ \forall\, a_i \in A_i,\ G_i(t_i)(a_i, \alpha_i) \geq \overline{U}_i^s(a_i, \alpha_i),$$

and

$$\sum_{\alpha \in A} \sum_{i \in N} t_i(\alpha)\, R(\alpha) = 0, \tag{5-II}$$

where s is some efficient selection rule.

A well-known (see e.g. Fan, 1956) necessary and sufficient condition for the existence of a solution of (5-II) is, $\forall\, \lambda \in \Lambda, \forall\, \mu \in \mathbb{R}$, if

$$\forall\, t \in T,\ \sum_{i \in N} \sum_{a_i \in A_i} \sum_{\alpha_i \in A_i} G_i(t_i)(a_i, \alpha_i)\lambda_i(a_i, \alpha_i) = \mu \sum_{\alpha \in A} \sum_{i \in N} t_i(\alpha)\, R(\alpha), \tag{9}$$

then

$$\sum_{i \in N} \sum_{a_i \in A_i} \sum_{\alpha_i \in A_i} \overline{U}_i^s(a_i, \alpha_i)\lambda_i(a_i, \alpha_i) \leq 0. \tag{10}$$

But

$$\begin{aligned}
&\sum_{i \in N} \sum_{a_i \in A_i} \sum_{\alpha_i \in A_i} G_i(t_i)(a_i, \alpha_i)\lambda_i(a_i, \alpha_i) \\
=&\sum_{i \in N} \sum_{a_i \in A_i} \sum_{\alpha_i \in A_i} \sum_{\alpha_{-i}} [t_i(\alpha) - t_i(a_i, \alpha_{-i})] P_i(\alpha_{-i} \mid \alpha_i)\lambda_i(a_i, \alpha_i) \\
=&\sum_{\alpha \in A} \sum_{i \in N} t_i(\alpha) \sum_{\alpha_i \in A_i} [P_i(\alpha_{-i} \mid \alpha_i)\lambda_i(a_i, \alpha_i) - P_i(\alpha_{-i} \mid a_i)\lambda_i(\alpha_i, a_i)].
\end{aligned}$$

So that (1) is equivalent to the following:

$$\begin{aligned}
&\forall\, i \in N,\ \forall\, \alpha_i \in A_i, \\
&\sum_{a_i \in A_i} [P_i(\alpha_{-i} \mid \alpha_i)\lambda_i(a_i, \alpha_i) - P_i(\alpha_{-i} \mid a_i)\lambda_i(\alpha_i, a_i)] = \mu\, R(\alpha).
\end{aligned} \tag{5-9'}$$

Since, for some $\alpha_{-i} \in A_{-i}$, $\sum_{\alpha_i} R(\alpha_{-i}, \alpha_i) \neq 0$, we get in (5-9') that $\mu = 0$. Therefore the necessary and sufficient condition becomes, $\forall \lambda \in \Lambda$, if

$$\forall\, t \in T,\ \sum_{i \in N} \sum_{a_i \in A_i} \sum_{\alpha_i \in A_i} G_i(t_i)(a_i, \alpha_i)\lambda_i(a_i, \alpha_i) = 0, \tag{11}$$

then

$$\sum_{i \in N} \sum_{a_i \in A_i} \sum_{\alpha_i \in A_i} \overline{U}_i^s(a_i, \alpha_i)\lambda_i(a_i, \alpha_i) \leq 0. \tag{12}$$

This holds by Lemma 1 (see the argument in the proof of Theorem 2). ■

In conclusion, the first contribution of this section is to provide some interpretable conditions on the beliefs implying Condition C, and thus representing sufficient conditions for the existence of a

Pareto-efficient BIC-mechanism [whatever the (transferable) payoff functions]. They require that the beliefs of at least one player be fully known or that the beliefs of all players be sufficiently in 'disagreement' to make correlation, in some sense, profitable. However, the Compatibility Condition C is not necessary. There are families of beliefs which do not satisfy Condition C but have such strong symmetry properties that it remains possible to design a Pareto-optimal feasible mechanism which is Bayesian incentive compatible.

The second contribution concerns the much more general 'consistency' condition. Under this condition we prove (for the discrete case) that it is always possible to find an efficient BIC-mechanism, provided the budget is only balanced in expected value.

References

Kenneth J Arrow. The property rights doctrine and demand revelation under incomplete information. In *Economics and Human Welfare*. Academic Press, New York, 1979.

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. Individual incentives and collective efficiency for an externality game with incomplete information. CORE Discussion Paper 7519, Université catholique de Louvain, 1975.

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

Ky Fan. On systems of linear inequalities. In H.W. Kuhn and A.W. Tucker, editors, *Linear Inequalities and Related Systems*. Princeton University Press, Princeton, 1956.

Jerry Green and Jean-Jacques Laffont. Characterization of satisfactory mechanisms for the revelation of preferences for public goods. *Econometrica: Journal of the Econometric Society*, 45:427–438, 1977.

Jerry R Green and Jean-Jacques Laffont. *Incentives in public decision making*. North-Holland, 1979.

Theodore Groves and Martin Loeb. Incentives and public inputs. *Journal of Public Economics*, 4(3): 211–226, 1975.

Milton Harris and Robert M Townsend. Resource allocation under asymmetric information. *Econometrica: Journal of the Econometric Society*, pages 33–64, 1981.

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.

Jean-Jacques Laffont and Eric Maskin. A differential approach to expected utility maximizing mechanisms. In J. J. Laffont, editor, *Aggregation and Revelation of Preferences*. North Holland, Amsterdam, 1979.

Jacques Neveu. *Bases mathématiques du calcul des probabilités*. Masson, Paris, 1970.

Halsey L Royden. *Real Analysis. Macmillan*. Macmillan, New York, 1968.