INCENTIVES AND INCOMPLETE INFORMATION

Claude d'Aspremont Louis-André Gérard-Varet*

Abstract

The problem of incentives for correct revelation in a collective decision model is presented as a game with incomplete information. Two approaches to incomplete information are used, a first where the individual beliefs are not introduced and a second where they are. In the first approach it is recalled that the mechanisms for which the solution to the incentive problem is in dominant strategies lead in general to a budgetary problem for the central agency. For these mechanisms a uniqueness property is demonstrated. In the second approach it is shown that if a compatibility condition is imposed on the individual beliefs and if a Bayesian solution is given to the incentive problem, then it is possible to avoid the budgetary problem.

1 Introduction

In a collective decision context, a selection rule may be called decentralized if it relies, at least partially, on the information that each individual participant holds. With such a rule, some participant may find in his self-interest to distort the information on which is based the selection, in a way undetectable by the others.

Historically this incentive problem has been brought up in the theory of Public Expenditure and Taxation [see Wicksell (1896), Lindhal (1919) and Samuelson (1969)] and was also considered, more or less explicitly, in some discussions concerning the Lange-Lerner economic model. However, as well established by Hurwicz (1972), this problem may arise for any collective decision rule preserving some kind of informational decentralization.

In this paper we shall argue that the problem of incentives for correct revelation should be viewed as a game with incomplete information [see Harsanyi

*We express our gratitude to Hervé Moulin, John Roberts and Shmuel Zamir for helpful discussions. This research was supported in part by the Commission Interministérielle de la Politique Scientifique in Belgium, in part by the Ford Foundation (Program for Research in International Economic Order) and by Cordes 136–177.

Reprinted with permission from *Journal of Public Economics* Volume 11, pp. 25–45.

(1967-1968)]. That it could be viewed as a game was already recognized by Samuelson (1969) and non-cooperative game-theoretic concepts were introduced by Drèze-de la Vallée Poussin (1971) and Hurwicz (1971). In section 2, we shall in fact introduce two appraoches to incomplete information. The first approach, which does not take explicitly into account the partial information that each participant may have concerning the others, is developed in section 3. The results presented there are typically based on a class of transfer schemes among participants with a strong incentive property, which was analyzed by several authors and which leads in general to the budgetary problem of balancing the transfers. In section 4, the second approach, in which the participants' beliefs concerning each other are introduced, provides a particular way for solving the budget problem whenever these beliefs satisfy a compatibility condition and only a Bayesian incentive property is required.

2 The Model

2.1 The basic collective decision problem

We shall consider a set N of n *individual agents* or *players* plus a special agent called the *central agency*. To define the collective decision problem, let us assume that there exists a subset X of $\mathbb{R}^m$ describing all possible *physical outcomes* for the individual agents and which form the set of alternatives.[1] Assume also the existence of a commodity called *money*. The choice of an outcome $x \in X$ is supposed to be the responsibility of the central agency which has to define simultaneously a vector $y = (y_1, \cdots, y_i, \cdots, y_n) \in \mathbb{R}^n$ of monetary *transfers* for all individual agents. In the process of selecting an outcome and of defining a vector of transfers the central agency is restricted by the 'a priori information specifications' and by the rules of a given 'mechanism'.

2.2 The a priori information assumptions

Each individual agent $i \in N$ is supposed to be described by a k-dimensional vector of characteristics belonging to $\mathbb{R}^k$. When α_i is the value of agent i's characteristics, we say that agent i is of *type* α_i. Furthermore, to each agent $i \in N$, we associate a function $V_i(\cdot;\alpha_i)$ from $\mathbb{R}^{m+1}$ to $\mathbb{R}$ such that $V_i(x, y_i; \alpha_i)$ denotes the *payoff*, for player i, in the situation where $x \in X$ is the outcome selected and $y_i \in \mathbb{R}$ is his monetary transfer. For the following, we shall actually restrict to the case where, for every $i \in N$, there exists a real-valued function[2] $U_i(\cdot;\alpha_i)$ such that, for every $x \in X$ and every $y_i \in \mathbb{R}$:

[1]The set X may actually be interpreted as being the set of outcomes associated to joint strategies in a game or in an 'organization form' as defined by Groves (1975). It may also be the set of all 'possible public projects' in a public good allocation problem as studied by Green-Laffont (1976).

[2]Functions U_i are interpreted, according to the model considered, as utilities or profits.

$$V_i(x, y_i; \alpha_i) \; = \; U_i(x; \alpha_i) + y_i.$$

This *separability* requirement amounts, in game-theoretic terms, to admit unrestricted side-payments with full-tranferability.

In decentralized contexts, to which we want to restrict our attention, it is supposed that every agent has *incomplete information* concerning the types α_i, except his own type, which is his private information. Indeed we assume that every agent only knows that the type α_i of any other agent i belongs to some space $A_i \subseteq \mathbb{R}^k$, which is the space of description of all possible types of agent i. All the sets A_i are assumed to be of *common knowledge*,[3] in the Aumann (1975) sense, meaning that no agent can consciously disagree on what they are. Furthermore we assume that the functions V_i and U_i are respectively defined on $X \times \mathbb{R} \times A_i$ and $X \times A_i$ and are of common knowledge. In the language of probability theory, the type of agent i is for every other agent a random phenomenon and A_i is its sample space.

In this incomplete information framework, we introduce strategic considerations by allowing some kind of communication process between the agents. Specifically we assume that each agent i has to announce to the other agents some type $a_i \in \mathbb{R}^k$ as being his own type α_i. We shall call such an announcement by agent i a *message* of agent i. Moreover, as a *plausibility condition*, we shall require that every individual message a_i belongs to the space A_i of possible types. Hence A_i is both the ith sample space and the message space of individual i.

2.3 Definition of a mechanism

Let $\{A_i; i \in N\}$ be a given family of sets of messages of the individual agents. We will denote by $A = \times_{i\in N} A_i$ the set of all n-tuples $a = (a_1, \cdots, a_n)$ of individual messages and by $A_{-i} = \times_{j\in N, j\neq i} A_i$ the set of all $(n-1)$-tuples $(a_1, \cdots, a_{i-1}, a_{i+1}, \cdots, a_n)$ of individual messages. Before sending any message to the central agency each individual agent is supposed to know the mechanism ruling the agency behavior. Formally, we call *mechanism* any function $m = (d, t)$ from A to $X \times \mathbb{R}^n$ where:

1. d is a function from A to X called *decision rule* and such that $d(a) = x$ is the outcome selected by the central agency whenever $a \in A$ is the n-tuple of messages received from the individual agents.

2. t is a function from A to $\mathbb{R}^n$ called *transfer scheme* and such that $t(a) = (t_1(a), \cdots, t_n(a))$ is the vector of individual transfer $y_i = t_i(a)$ determined by the central agency whenever $a \in A$ is the n-tuple of messages received from the individual agents.

[3] Clearly player i's payoff function is fully known only when his type α_i is known.

In the following, we shall restrict the set of all possible mechanisms to a certain admissible subset M. A first requirement for M that we shall use in most results is an *outcome efficiency* requirement:

$$\forall a \in A,\, d(a) \in \mathcal{P}(a) \stackrel{\text{def}}{=} \{x^* \in X : \sum_{i \in N} U_i(x^*; a_i) = \max_{x \in X} \sum_{i \in N} U_i(x; a_i)\}.$$

For any m (or d) satisfying this condition we shall say that m (or d) is outcome efficient.

In order for such a requirement to be meaningful, we shall impose one of the following alternative *regularity hypotheses*:

H_1: X is compact and $\forall a \in A$, $\sum_{i \in N} U_i(\cdot; a_i)$ is upper semicontinuous on X.

H_2: X is open convex and $\forall a \in A$, $\sum_{i \in N} U_i(\cdot; a_i)$ is a differentiable strictly concave function having a critical point in X.

Condition H_1 implies that, for every $a \in A$, $\mathcal{P}(a) \neq \emptyset$ and condition H_2 ensures in addition that $\mathcal{P}(a)$ is single-valued. Other conditions could be used.

2.4 The communication game: Alternative approaches to incomplete information

Suppose that a mechanism $m \in M$ has been chosen and that $\alpha \in A$ is the n-tuple of the individual agents' types. A communication process between the individual agents and the center can be formalized as an n-person game in normal form[4] conditional to $\alpha \in A$. In this game, the strategy space of each $i \in N$ is his message space A_i. Given $\alpha \in A$, the payoff functions are:

$$\begin{aligned} \forall i \in N,\, \forall a \in A,\, W_i^m(a; \alpha_i) &\stackrel{\text{def}}{=} V_i(m(a); \alpha_i) \\ &= U_i(d(a); \alpha_i) + t_i(a). \end{aligned}$$

Thus, for every $\alpha \in A$, we have the normal form game:

$$\Gamma^m(\alpha) \;=\; \{\{A_i; i \in N\},\, \{W_i^m(\cdot; \alpha_i); i \in N\}\}.$$

However, because of the incomplete information framework, every player ignores what game $\Gamma^m(\alpha)$ is to be played. Hence, when he wants to characterize the behavior of any other player j, player i must consider not only what message a_j player j announces but also what type α_j could be player j's true type.

This consideration is essential in the formulation of the incentive problem that will be examined in the subsequent sections. Indeed the question will be to determine whether every player has interest to reveal his true type. In

[4]See Luce and Raiffa (1957, p. 157).

this context, every player may want to characterize, for every other player, the particular behavior consisting in revealing his true type.

For this reason, we shall introduce a more sophisticated strategy concept. For every i, we introduce the notion of a *normalized strategy of player i* to be a decision rule a_i^* associating a unique strategy choice to each of his possible types. Formally a_i^* is a function from A_i to A_i. We denote A_i^* the set of all admissible normalized strategies for i. The strategy, consisting in declaring the true value of his parameter in the communication game, is a normalized strategy for each player.

Now to treat the incentive problem we shall distinguish two appraoches, each one associated to a different definition of the communication process. The first approach, treated in section 3, considers that, for every player $i \in N$, the other players' space of types A_{-i} is a space of states of nature for which player i satisfies, as a decision-maker, the *'complete ignorance' postulate* [Luce and Raiffa (1957, p.294)]. Let, for $m \in M$, $G(m) = \{\{\Gamma^m(\alpha); \alpha \in A\}, \{A_i^*; i \in N\}\}$. According to the first approach, all the games belonging to $G(m)$ have to be considered simultaneously by *all* agents, i.e. by the players *and* by the central agency, for every matter related to the mechanism $m \in M$. Consequently, we say of $G(m)$ that it is the *standard form of the communication game associated to the mechanism $m \in M$ under the complete ignorance postulate*.

The second approach, treated in section 4, considers that every player i, whatever his type $\alpha_i \in A_i$, has some 'beliefs' concerning the others types. We shall assume[5] that the beliefs of player i are represented by a real-valued function p_i defined over $B_{-i} \times A_i$, where B_{-i} is the Borel σ-algebra on A_{-i}, and such that for every $\alpha_i \in A_i$, $p_i(\cdot|\alpha_i)$ is a probability on (A_{-i}, B_{-i}) with full support. All functions p_i are of common knowledge. However, player i beliefs are fully known only when his type $\alpha_i \in A_i$ is also known.

In this model, we may associate to a mechanism $m \in M$ not only the family $G(m)$ of games, but also the family $\{p_i; i \in N\}$ of beliefs. We get the *standard form of the communication game associated to the mechanism $m \in M$ under the (Bayesian) probabilistic postulate*[6] which may be denoted:

$$\Gamma(m) \;=\; \{\{\Gamma^m(\alpha), \alpha \in A\},\, \{a_i^*; i \in N\},\, \{p_i; i \in N\}\}.$$

The basic difference with the complete ignorance framework is that players have now beliefs about which one is the true game in $G(m)$. Notice that the central agency, which is not a player, does not have any beliefs.

---

[5] In section 4, we shall see that some additional assumptions should be put on these functions p_i like requiring that $p_i(\cdot|\alpha_i)$ be discrete or, more generally, that p_i be a transition probability between the spaces (A_i, B_i) and (A_{-i}, B_{-i}) [see Neveu (1970, p.69)], where B_i is the Borel σ-algebra on A_i.

[6] As introduced by Harsanyi (1967-1968, Part I, p.72). Notice that $\Gamma(m)$ is not a normal form game.

3 Incentives under the Complete Ignorance Postulate

Let us consider first the case where the incomplete information game which formalizes the communication process associated to a mechanism $m \in M$ is defined under the (nonprobabilitic) complete ignorance postulate.

The study of the solution of the incentive problem will result from the analysis of the behavior of the players in the game $G(m)$.

3.1 Incentive and strongly incentive compatible mechanisms

Take any mechanism $m \in M$ and consider the communication game $G(m)$ associated to m under the complete ignorance postulate. For notational convenience we let: $\forall a \in \mathbb{R}^{nk}$, $\forall i \in N$, $\forall \overline{a}_i \in \mathbb{R}^k$,

$$\begin{aligned}(\overline{a}_i, a_{-i}) &= (a_1, \cdots, a_{i-1}, \overline{a}_i, a_{i+1}, \cdots, a_n),\\ A^* &= \underset{i \in N}{\times} A_i^*,\ a^*(\alpha) \ = \ (a_1^*(\alpha_1), a_1^*(\alpha_2), \cdots, a_n^*(\alpha_n)),\\ A_{-i}^* &= \underset{\substack{j \in N \\ j \neq i}}{\times} A_j^*,\ a_{-i}^*(\alpha_{-i}) \ = \ (a_1^*(\alpha_1), \cdots, a_{i-1}^*(\alpha_{i-1}), a_{i+1}^*(\alpha_{i+1}), \cdots, a_n^*(\alpha_n)).\end{aligned}$$

We say that an n-tuple of normalized strategies $a^* \in A^*$ is a *Nash Equilibrium (locally) for some* $\alpha \in A$ if and only if $a^*(\alpha)$ is a Nash Equilibrium for the game $\Gamma^m(\alpha)$, i.e.

$$\forall i \in N,\ \forall a_i \in A_i,\ W_i^m(a_i, a_{-i}^*(\alpha_{-i}); \alpha_i) \ \leq W_i^m(a^*(\alpha); \alpha_i).$$

An n-tuple $a^* \in A^*$ is a *uniform equilibrium* if and only if it is a Nash equilibrium for every possible $\alpha \in A$. It is clear now that the notion of normalized strategy is needed here so that each player can characterize the behavior of the others. We denote by $E(m)$ the subset of A^* of all uniform equilibria for the game $G(m)$.

With these notions we are able to define a first solution concept to the incentive problem (due primarily to Hurwicz (1972)). Define the particular normalized strategy $\hat{a}_i^*$ for player i by

$$\forall \alpha_i \in A_i,\ \hat{a}_i^*(\alpha_i) \ = \ \alpha_i.$$

This is the normalized strategy, for player i, consisting in always revealing his true type. We say that the mechanism m is *(locally) incentive compatible for some* $\alpha \in A$ if and only if $\hat{a}^*$ is a Nash equilibrium for that α. Of course one is generally interested to show that this local property holds for a large subset of A. For this reason, we say that a mechanism m is *incentive compatible* if and only if it is incentive compatible for every $\alpha \in A$, i.e. $\hat{\alpha}^* \in E(m)$.

Some authors[7] have introduced a stronger concept of incentive compability. To introduce this other concept we recall that a *dominant strategy* for any player i for a game $\Gamma^m(\alpha)$ is a strategy $\overline{a}_i \in A_i$ such that:

$$\forall a_{-i} \in A_{-i},\, \forall a_i \in A_i,\, W_i^m(a_i, a_{-i}; \alpha_i) \leq W_i^m(\overline{a}_i, a_{-i}; \alpha_i).$$

We say that an n-tuple $a^* \in A^*$, is a *dominant uniform equilibrium* if and only if, for every $\alpha \in A$, every $a_i^*(\alpha_i)$ is a dominant strategy for player i. We denote by $E_D(m)$ the subset of A^* of all dominant uniform equilibria for the game $G(m)$. Then we may say that a mechanism m is *strongly incentive compatible* if and only if $\hat{a}^* \in E_D(m)$.

It is clear that this concept of incentive compatibility is stronger than the previous one, since, for every mechanism m, $E_D(m) \subseteq E(m)$. However under the plausibility condition we have introduced we get:

Theorem 1. *Any mechanism $m \in M$ is strongly incentive compatible if and only if it is incentive compatible.*

Proof: We only prove sufficiency. If $m \in M$ is incentive compatible we may write:

$$\forall i \in N,\, \forall \alpha_i \in A_i,\, \forall \alpha_{-i} \in A_{-i},\, \forall a_i \in A_i,\, W_i^m(a_i, \hat{a}_{-i}^*(\alpha_{-i}); \alpha_i) \leq W_i^m(\hat{a}^*(\alpha); \alpha_i).$$

This implies:

$$\forall i \in N,\, \forall \alpha_i \in A_i,\, \forall a_{-i} \in A_{-i},\, \forall a_i \in A_i,\, W_i^m(a_i, a_{-i}; \alpha_i) \leq W_i^m(\alpha_i, a_{-i}; \alpha_i).$$

i.e., m is strongly incentive compatible. ∎

3.2 A class of strongly incentive compatible mechanisms and the budget problem

We say of a mechanism $m \in M$ that it is a *distribution mechanism* if and only if the transfer scheme is such that:

$$\forall i \in N,\, \forall a \in A,\, t_i(a) \;=\; \sum_{\substack{j \in N \\ j \neq i}} U_j(d(a); a_j) - f_i(a),$$

where f_i is, for every $i \in N$, a real-valued function defined over A. In this case player $i \in N$ receives from (pays to) the central agency the difference between the amounts $\sum_{j \in N, j \neq i} U_j(d(a); a_j)$ and $f_i(a)$ both defined in terms of the declared types and of the corresponding decision rule. The n-tuple $(f_1, \cdots, f_i, \cdots, f_n)$ is a *distribution rule*. In particular, a distribution rule is said to be *discretionary* if and only if:

[7]See, for instance, Groves (1973, 1975), Groves and Loeb (1975) and Green and Laffont (1976).

$$\forall i \in N, \forall a_{-i} \in A_{-i}, \forall a_i \in A_i, \forall a_i' \in A_i, \; f_i(a_i, a_{-i}) \;=\; f_i(a_i', a_{-i}).$$

In that case, for every player, the distribution rule is constant with respect to the message that the player sends to the agency. Accordingly, we shall say of a mechanism that it is a *discretionary mechanism* whenever it is a distribution mechanism with discretionary distribution rule.

We are interested by mechanisms which are simultaneously outcome efficient and discretionary. (They are sometimes called *Groves-mechanisms*.[8]) The reason for restricting to this class of mechanisms is that it turns out to coincide with the class of strongly incentive mechanisms. More precisely, we have the characteristization resulting from the next two theorems:[9]

Theorem 2. *(a) If H_1 holds, then any mechanism which is outcome efficient and discretionary is strongly incentive compatible. (b) If H_2 holds, any mechanism which is strongly incentive compatible and outcome efficient is discretionary.*

Note that, by theorem 1, theorem 2 characterizes also incentive compatible mechanisms.

The above result is clearly due to the structure of the transfers in a discretionary mechanism. However these transfers are made through the budget of the central agency. Hence, an important consideration is to know whether the structure of the transfers makes it possible for the agency to balance its budget. More generally we shall say that a mechanism $m = (d, t)$ is *budget balancing* if and only if:

$$\sum_{i \in N} t_i(\cdot) \;\equiv\; 0.$$

The question is therefore to know whether an outcome efficient discretionary mechanism can be budget balancing. However, as noted in Groves-Loeb (1975) the answer is often negative[10] except when the utility functions are of a particular quadratic type.

In Groves-Ledyard (1975) model of a general equilibrium economy with public goods, where the utility functions are not supposed to be separable but in which the agents are only required to communicate marginal willingness to pay functions, the budget problem is similarly treated by quadratic approximation.

---

[8]This terminology which is used by Green-Laffont (1975) and (1976) is based on Groves (1973) and Groves and Loeb (1975) papers. Groves mechanisms are in fact analogous to Vickrey (1961) and Clarke (1971) mechanisms and identical to Smets (1973) compensation principle.

[9]Theorem 2(a) is proved by Groves-Loeb (1975) and Theorem 2(b) is proved by Green-Laffont (1977). Both of them are presented in Groves (1975) and Green-Laffont (1976).

[10]Proposition 4, p.27, in d'Aspremont–Gérard-Varet (1975) and Theorem 9, p.39, in Green-Laffont (1976) demonstrate for different particular cases, the impossibility of having a budget balancing outcome efficient discretionary mechanism.

In a partial equilibrium approach with only two alternative outcomes, Green and Laffont (1976) assume that the willingness to pay individual values are randomly sampled from a continuous distribution of a given law.[11] With such an assumption, they show that the budget differences, in the strongly incentive compabile mechanism they use, may become negligible in expected value when the sample size is increased.

In section 4, we shall treat the budget problem in an alternative way by using the Bayesian approach.

3.3 Almost strictly incentive compatible mechanisms

It is clear from the definitions that even for incentive compatible mechanisms there may be other uniform equilibria, for the corresponding game $G(m)$, than the truth normalized strategy n-tuple $\hat{a}^*$. The simplest way to solve this problem would be to restrict oneself to mechanisms for which $\hat{a}^*$ is the only uniform equilibrium for the game $G(m)$. Without being so restrictive, we define a mechanism to be *almost strictly incentive compatible* if and only if: (i) $\hat{a}^* \in E(m)$, (ii) $\forall a^* \in E(m)$, $\forall \alpha \in A$, $d(\hat{a}^*(\alpha)) = d(a^*(\alpha))$. This means that for any other uniform equilibrium in the game the outcome remains unchanged. Now, we have the following theorem.

Theorem 3. *If H_2 holds, then any mechanism which is outcome efficient and discretionary is almost strictly incentive compatible.*

Proof: Suppose $m = (d, t)$ is outcome efficient and discretionary. Then we have the following three facts:

(1) $\forall \overline{a}^* \in E(m)$, $\forall \alpha \in A$, $\forall i \in N$, $W_i^m(\overline{a}^*(\alpha); \alpha_i) \ = \ W_i^m(\alpha_i, \overline{a}^*_{-i}(\alpha_{-i}); \alpha_i)$.
Indeed, since $\hat{a}^* \in E_D(m)$ by theorem 1, we have

$$\forall a^* \in A^*,\ \forall \alpha \in A,\ \forall i \in N,\ W_i^m(\alpha_i, a^*_{-i}(\alpha_{-i}); \alpha_i) \ \geq \ W_i^m(a^*(\alpha); \alpha_i),$$

and, by definition,

$$\forall \overline{a}^* \in E(m),\ \forall \alpha \in A,\ \forall i \in N,\ W_i^m(\overline{a}^*(\alpha); \alpha_i) \ \geq \ W_i^m(\alpha_i, \overline{a}^*_{-i}(\alpha_{-i}); \alpha_i).$$

(2) $\forall \overline{a}^* \in E(m)$, $\forall \alpha \in A$, $\forall i \in N$, $d(\alpha_i, \overline{a}^*_{-i}(\alpha_{-i})) \ = \ d(\overline{a}^*(\alpha))$.
Indeed, since m is defined by a discretionary distribution rule, (1) is equivalent to:

$$\begin{aligned} & \forall \overline{a}^* \in E(m),\ \forall \alpha \in A,\ \forall i \in N,\ U_i(d(\overline{a}^*(\alpha)); \alpha_i) \ + \ \sum_{j \neq i} U_j(d(\overline{a}^*(\alpha)); \overline{a}^*_j(\alpha_j)) \\ = \quad & U_i(d(\alpha_i, \overline{a}^*_{-i}(\alpha_{-i})); \alpha_i) \ + \ \sum_{j \neq i} U_j(d(\alpha_i, \overline{a}^*_{-i}(\alpha_{-i})); \overline{a}^*_j(\alpha_j)). \end{aligned}$$

Hence, since by H_2 the function $U_i(x, a_i) + \sum_{j \neq i} U_j(x; \overline{a}^*_j(\alpha_j))$ has a unique maximum, we get (2).

[11] See also Green-Kohlberg-Laffont (1976).

(3) $\forall \overline{a}^* \in E(m)$, $\forall \alpha \in A$, $d(\overline{a}^*(\alpha)) \; = \; d(\alpha)$.
Indeed, by H_2 and (2), $\forall \overline{a}^* \in E(m)$, $\forall \alpha \in A$, $\forall i \in N$, the first derivative

$$DU_i(d(\overline{a}^*(\alpha)); \alpha_i) = -\sum_{j \neq i} DU_j(d(\overline{a}^*(\alpha)); \overline{a}_j^*(\alpha_j)) = DU_i(d(\overline{a}^*(\alpha)); \overline{a}_i^*(\alpha_i)).$$

Hence: $\forall \overline{a}^* \in E(m)$, $\forall \alpha \in A$,

$$\sum_{i \in N} DU_i(d(\overline{a}^*(\alpha)); \alpha_i) \; = \; \sum_{i \in N} DU_i(d(\overline{a}^*(\alpha)); \overline{a}_i^*(\alpha_i)) \; = \; 0.$$

Therefore, by H_2 again, (3) follows. ■

4 Incentives under the Bayesian Postulate

We want now to turn to the case where for any mechanism $m \in M$ the behavior of the individual agents satisfies the Bayesian postulate. Accordingly the relevant standard form of the communication game is $\Gamma(m)$ which includes the players' beliefs. The concept of incentive compatibility may also be reformulated since each player is now supposed to maximize a mathematical expectation of his payoffs in terms of his subjective probability.

4.1 Bayesian and strongly Bayesian incentive compatible mechanisms

In a game $\Gamma(m)$, the message of every player i is going to be determined by his *expected-payoff* conditional to α_i and relative to the choice of normalized strategy by every other player. Hence we shall write $\forall i \in N$, $\forall \alpha_i \in A_i$, $\forall a_{-i}^* \in A_{-i}^*$, $\forall a_i \in A_i$,

$$\begin{aligned} \overline{W}_i^m(a_i, a_{-i}^*; \alpha_i) &= \int_{A-i} W_i^m(a_i, a_{-i}^*(\alpha_{-i}); \alpha_i) p_i(d\alpha_{-i}|\alpha_i) \\ &= \int_{A-i} [U_i(d(a_i, a_{-i}^*(\alpha_{-i})); \alpha_i) + t_i(a_i, a_{-i}^*(\alpha_{-i}))] p_i(d\alpha_{-i}|\alpha_i). \end{aligned}$$

In order for such an expression to be well defined we shall add a *measurability restriction* (R): every message space must be bounded and measurable and all the normalized strategies, the decision rules and the transfer schemes are restricted to be measurable functions.

Under the measurability assumption, we may define a weaker notion of equilibrium for the communication game.

We shall say that a *Bayesian equilibrium for* $\Gamma(m)$ is an n-tuple of normalized strategies $\overline{a}^* \in A^*$ such that

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

We denote by $\mathcal{B}(m)$ the subset in A^* of all Bayesian equilibria for $\Gamma(m)$.

Notice that, as in the complete ignorance case (see 3.1), the equilibrium notion is defined with respect to every possible individual type. Consider the n-tuple of normalized strategies $\hat{a}^* \in A^*$ (the 'truth' strategies) such that: $\forall i \in N$, $\forall \alpha_i \in A_i$, $\hat{a}_i^*(\alpha_i) = \alpha_i$.

We now say of a (measurable) mechanism m that it is *Bayesian incentive compatible* if and only if $\hat{a}^* \in \mathcal{B}(m)$. Clearly this is defined by the condition: $\forall i \in N$, $\forall \alpha_i \in A_i$, $\forall a_i \in A_i$,

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

In other words, for every player $i \in N$ and every possible type $\alpha_i \in A_i$, sending as a message this information to the center dominates every other possible message $a_i \in A_i$, whenever the other players have presumably the same behavior.

Like in the complete ignorance case a stronger incentive compatible notion is obtained if such a dominance property has to hold for every player whatever the behavior of the $n-1$ other players. A strategy $\overline{a}_i^* \in A_i^*$ is *dominating for player* $i \in N$ if and only if:

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

An n-tuple of strategies $\overline{a}^* \in A^*$ dominating for every player is called a *strong Bayesian equilibrium for* $\Gamma(m)$. Let $\mathcal{B}_S(m)$ be the set of all strong Bayesian equilibria with respect to $m \in M$. We may now say of a (measurable) mechanism m that it is *strongly Bayesian incentive compatible* if and only if $\hat{a}^* \in \mathcal{B}_S(m)$, or:

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

Now it is clear that if a (measurable) mechanism m is strongly incentive compatible then it is Bayesian incentive compatible. In addition, we have:

Theorem 4. *Under Assumption R, a mechanism m is strongly Bayesian incentive compatible if and only if it is strongly incentive compatible.*

Proof: (1) Assume m is strongly Bayesian incentive compatible but not strongly incentive compatible, i.e. $\forall i \in N$, $\forall \alpha_i \in A_i$, $\forall a_i^* \in A_i^*$, $\forall a_{-i}^* \in A_{-i}^*$,

$$\int_{A-i} W_i^m(a_i^*(\alpha_i), a_{-i}^*(\alpha_{-i}); \alpha_i) p_i(d\alpha_{-i}|\alpha_i) \leq \int_{A-i} W_i^m(\alpha_i, a_{-i}^*(\alpha_{-i}); \alpha_i) p_i(d\alpha_{-i}|\alpha_i),$$

but, $\exists i \in N$, $\exists \overline{\alpha}_i \in A_i$, $\exists \overline{a}_{-i} \in A_{-i}$, $\exists \overline{a}_i \in A_i$ such that:

$$W_i^m(\overline{a}_i, \overline{a}_{-i}; \overline{\alpha}_i) > W_i^m(\overline{\alpha}_i, \overline{a}_{-i}; \overline{\alpha}_i).$$

Now define: $\forall k \in N$, $\forall \alpha_k \in A_k$, $\overline{a}_k^*(\alpha_k) = \overline{a}_k$. Clearly $\overline{a}_k^* \in A_k^*$. Then

$$W_i^m(\overline{a}_i, \overline{a}_{-i}; \overline{\alpha}_i) \;=\; \int_{A-i} W_i^m(\overline{a}_i^*(\overline{\alpha}_i), \overline{a}_{-i}^*(\alpha_{-i}); \overline{\alpha}_i) p_i(d\alpha_{-i}|\overline{\alpha}_i)$$

$$\leq \int_{A-i} W_i^m(\overline{\alpha}_i, \overline{a}_{-i}^*(\alpha_{-i}); \overline{\alpha}_i) p(d\alpha_{-i}|\overline{\alpha}_i) \;=\; W_i^m(\overline{\alpha}_i, \overline{a}_{-i}; \overline{\alpha}_i),$$

which is a contradiction.

(2) Assume m is strongly incentive compatible; then $\forall i \in N$, $\forall \alpha_i \in A_i$, $\forall a_i^* \in A_i^*$, $\forall a_{-i}^* \in A_{-i}^*$,

$$W_i^m(a_i^*(\alpha_i), a_{-i}^*(\alpha_{-i}); \alpha_i) \;\leq W_i^m(\alpha_i, a_{-i}^*(\alpha_{-i}); \alpha_i),\ \forall \alpha_{-i} \in A_{-i}.$$

Hence

$$\int_{A-i} W_i^m(a_i^*(\alpha_i), a_{-i}^*(\alpha_{-i}); \alpha_i) p(d\alpha_{-i}|\alpha_i) \leq \int_{A-i} W_i^m(\alpha_i, a_{-i}^*(\alpha_{-i}); \alpha_i) p_i(d\alpha_{-i}|\alpha_i),$$

and m is strongly Bayesian incentive compatible. ■

As a corollary, we may get a theorem analogous to theorem 3 since we know, by theorem 2(b), that every strongly incentive compatible mechanism, which is outcome efficient, is discretionary.

Corollary 1. *If H_2 and R hold, then any mechanism m which is outcome efficient and strongly Bayesian incentive compatible is almost strictly incentive compatible.*

4.2 Bayesian incentive compatible mechanisms and a solution to the budget problem

For any distribution mechanism satisfying R, we say that the associated distribution rule f is *subjectively discretionary* if and only if: $\forall i \in N$, $\forall \alpha_i \in A_i$, $\forall a_i \in A_i$, $\forall a_i' \in A_i$,

$$\int_{A-t} f_i(a_i, \alpha_{-i}) p_i(d\alpha_{-i}|\alpha_i) \;=\; \int_{A-i} f_i(a_i', \alpha_{-i}) p_i(d\alpha_{-i}|\alpha_i).$$

A *subjectively discretionary mechanism* is a distribution mechanism for which the associated distribution rule is subjectively discretionary. In contrast with discretionary distribution rules, the restriction introduced here on f imposes only that, for every $i \in N$, the expected value of f_i must be constant with respect to i's messages.

The interest we have for the class of subjective discretionary mechanisms comes not only from the fact that it includes the class of discretionary mechanisms, but also because it is included in the class of Bayesian incentive compatible mechanisms as shown by the next result paralleling Theorem 2a.

Theorem 5. *Under Assumptions H_1 and R, any mechanism $m \in M$ which is outcome efficient and subjectively discretionary is Bayesian incentive compatible.*

Proof: By Assumption R the individual payoffs are well-defined. We want to show that for the mechanism m:

$$\forall i \in N,\, \forall \alpha_i \in A_i,\, \forall a_i \in A_i,\, \overline{W}_i^m(a_i, \hat{a}^*_{-i}; \alpha_i) \ \leq \overline{W}_i^m(\alpha_i, \hat{a}^*_{-i}; \alpha_i). \tag{1}$$

Since m is outcome efficient: $\forall \alpha \in A$, $\forall i \in N$, $\forall A_i \in A_i$,

$$U_i(d(\alpha);\alpha_i) \ + \sum_{\substack{j\in N\\ j\neq i}} U_j(d(\alpha);\alpha_j) \ \geq \ U_i(d(a_i,\alpha_{-i});\alpha_i) + \sum_{\substack{j\in N\\ j\neq i}} U_j(d(a_i,\alpha_{-i});\alpha_j),$$

which implies:

$$\int_{A-i} U_i(d(\alpha);\alpha_i)p_i(d\alpha_{-i}|\alpha_i) + \int_{A_{-i}} \sum_{j\neq i} U_j(d(\alpha);\alpha_j)p_i(d\alpha_{-i}|\alpha_i) \tag{2}$$

$$\geq \int_{A-i} U_i(d(a_i,\alpha_{-i});\alpha_i)p_i(d\alpha_{-i}|\alpha_i) + \int_{A-i} \sum_{j\neq i} U_j(d(a_i,\alpha_{-i});\alpha_j)p_i(d\alpha_{-i}|\alpha_i).$$

Since m is subjectively discretionary:$\forall i \in N$, $\forall \alpha_i \in A_i$, $\forall a_i \in A_i$,

$$\int_{A-i} f_i(\alpha_i,\alpha_{-i})p_i(d\alpha_{-i}|\alpha_i) \ = \ \int_{A-i} f_i(a_i,\alpha_{-i})p_i(d\alpha_{-i}|\alpha_i). \tag{3}$$

Clearly, in light of (3), (2) is equivalent to (1).

■

We now want to study the balanced budget problem. The first result we have concerning this problem uses a strong assumption on the various beliefs of the individual players, called the *independence condition*.[12] It requires that for every player i:

$$\forall \alpha_i \in A_i,\, \forall \alpha'_i \in A_i,\, p_i(\cdot|\alpha_i) \ = \ p_i(\cdot|\alpha'_i) \ \stackrel{\text{def}}{=} \ \pi_i(\cdot)\,.$$

This condition is, in terms of information, very restrictive since it implies in fact that the true beliefs of any agent is of common knowledge.

Theorem 6. *Let H and R hold. If the independence condition is satisfied, then the mechanism $m = (d,t)$, where d is any outcome efficient decision rule and t is such that, $\forall i \in N$, $\forall a \in A$,*

$$\begin{aligned} t_i(a) \quad &= \quad \int_{A-i}\left[\sum_{j\neq i} U_j(d(a);a_j)\right]\pi_i(da_{-i}) \\ &- \quad \frac{1}{n-1}\sum_{j\neq i}\int_{A-j}\left[\sum_{k\neq j} U_k(d(a);a_k)\right]\pi_j(da_{-j}) \end{aligned}$$

[12]It has been shown in d'Aspremont and Gérard-Varet (1975) that a converse theorem to theorem 5 holds if the independence condition holds.

is both budget balancing and Bayesian incentive compatible.

Proof: By construction $\sum_{i \in N} t_i(\cdot) = 0$, hence m is budget balancing. Also we may rewrite m as the following distribution mechanism:

$$\forall i \in N,\, \forall a \in A,\; f_i(a) \;=\; \sum_{j \neq i} U_j(d(a); a_j) - g_i(a_i) + g_{-i}(a_{-i}),$$

where

$$g_i(a_i) \;=\; \int_{A-i} \left[\sum_{j \neq i} U_j(d(a); a_j) \right] \pi_i(da_{-i})$$

and

$$g_{-i}(a_{-i}) \;=\; \frac{1}{n-1} \sum_{j \neq i} g_j(a_j).$$

Then we have: $\forall \alpha \in A$, $\forall i \in N$, $\forall a_i \in A_i$,

$$\int_{A-i} f_i(a_i, \alpha_{-i}) \pi_i(d\alpha_{-i}) \;=\; \int_{A-i} g_{-i}(\alpha_{-i}) \pi_i(d\alpha_{-i}),$$

which is a constant. Hence m is subjectively discretionary and so Bayesian incentive compatible by theorem 5.

■

Since the independence condition is a very restrictive sufficient condition it seems important to establish whether or not it is also necessary for the result of theorem 6 to hold. In this paper we answer this question only in the finite case. Indeed, in that case, we may show that the result of theorem 6 holds for a much larger class than the class of independent beliefs.

Let, for every $i \in N$, the set A_i be finite and for every $\alpha_i \in A_i$, $p_i(\cdot|\alpha_i)$ be a discrete probability of full support over A_{-i}. For the following let also: $\forall i \in N$, $C_i \stackrel{\text{def}}{=} \{(a_i, \alpha_i) \in A_i \times A_i : \alpha_i \neq a_i\}$, and $\Lambda \stackrel{\text{def}}{=} \{\lambda = (\lambda_1, \cdots, \lambda_i, \cdots, \lambda_n)$: $\forall i \in N,\, \lambda_i \in \mathbb{R}_+^{C_i}\}$. The *compatibility condition* we use in the next theorem can be written as:

$\forall \kappa \in \mathbb{R}^A$, if $\kappa \neq 0$ then there is no $\lambda \in \Lambda$ such that, $\forall i \in N$, $\forall \alpha \in A$,

$$p_i(\alpha_{-i}|\alpha_i) \sum_{\substack{a_i \in Ai \\ a_i \neq \alpha_i}} \lambda_i(\alpha_i, a_i) \;=\; \kappa(\alpha) + \sum_{\substack{a_i \in Ai \\ a_i \neq \alpha_i}} \lambda_i(\alpha_i, a_i) p_i(\alpha_{-i}|a_i).$$

That this new condition is implied by the independence condition is immediate. In addition, the following example ensures easily that the converse does not hold. For $n = 2$, let p_1 and p_2 satisfy: $\forall \alpha_1 \in A_1$, $\exists \alpha_2 \in A_2$ such that

(i) $\qquad \forall a_1 \in A_1,\, a_1 \neq \alpha_1,\, p_1(\alpha_2|a_1) \;>\; p_1(\alpha_2|\alpha_1)$

and

(ii) $\forall a_2 \in A_2,\, a_2 \neq \alpha_2,\, p_2(\alpha_1|a_2) \;<\; p_2(\alpha_1|\alpha_2).$

Indeed for any $\kappa \in \mathbb{R}^A$ such that $\kappa(\overline{\alpha}) \neq 0$, for some $\overline{\alpha} \in A$, and for any $\lambda \in \Lambda$ such that:

$$p_1(\overline{\alpha}_2|\overline{\alpha}_1) \sum_{\substack{a_1 \in A_1 \\ a_1 \neq \overline{\alpha}_1}} \lambda_1(\overline{\alpha}_1, a_1) \;-\; \sum_{\substack{a_1 \in A_1 \\ a_1 \neq \overline{\alpha}_1}} \lambda_1(\overline{\alpha}_1, a_1) p_1(\overline{\alpha}_2|a_1) \;=\; \kappa(\overline{\alpha}) \neq 0,$$

we must have

$$\sum_{\substack{a_1 \in A_1 \\ a_1 \neq \overline{\alpha}_1}} \lambda_1(\overline{\alpha}_1, a_1) \;>\; 0.$$

Hence for α_2 such that $(\overline{\alpha}_1, \alpha_2)$ satisfies (i) and (ii) we get:

$$p_1(\alpha_2|\overline{\alpha}_1) \sum_{\substack{a_1 \in A_1 \\ a_1 \neq \overline{\alpha}_1}} \lambda_1(\overline{\alpha}_1, a_1) \;-\; \sum_{\substack{a_1 \in A_1 \\ a_1 \neq \overline{\alpha}_1}} \lambda_1(\overline{\alpha}_1, a_1) p_1(\alpha_2|a_1) \;<\; 0,$$

and

$$p_2(\overline{\alpha}_1|\alpha_2) \sum_{\substack{a_2 \in A_2 \\ a_2 \neq \alpha_2}} \lambda_2(\alpha_2, a_2) \;-\; \sum_{\substack{a_2 \in A_2 \\ a_2 \neq \alpha_2}} \lambda_2(\alpha_2, a_2) p_2(\overline{\alpha}_1|a_2) \;\geq\; 0.$$

We can now state our last theorem:

Theorem 7. *Let H hold and, for every $i \in N$, let A_i be finite and p_i be discrete. If the compatibility condition is satisfied, then there exists an outcome efficient mechanism which is both budget balancing and Bayesian incentive compatible.*

Proof: Let d be any outcome efficient decision rule. For every $i \in N$, we write:

$$\forall a \in A,\, \forall i \in N,\, \forall \alpha_i \in A_i,\, u_i(a; \alpha_i) \;\stackrel{\text{def}}{=}\; U_i(d(a); \alpha_i)$$

and

$$\forall a_i \in A_i,\, \forall \alpha_i \in A_i,\, \overline{u}_i(a_i, \alpha_i) \;\stackrel{\text{def}}{=}\; \sum_{\alpha_{-i} \in A_{-i}} [u_i(a_i, \alpha_{-i}; \alpha_i) - u_i(\alpha_i, \alpha_{-i}; \alpha_i)] p_i(\alpha_{-i}|\alpha_i).$$

Consider the following system of linear inequalities where $z \in \mathbb{R}^{nA}$ is taken as variable:

$$\begin{aligned} \forall i \;\in\; & N,\, \forall (a_i, \alpha_i) \in C_i,\, \sum_{\alpha_{-i} \in A_{-i}} [z_i(\alpha_i, \alpha_{-i}) - \frac{1}{n-1} \sum_{j \neq i} z_j(\alpha_i, \alpha_{-i}) \\ & -z_i(a_i, \alpha_{-i}) + \frac{1}{n-1} \sum_{j \neq i} z_j(a_i, \alpha_{-i})] p_i(\alpha_{-i}|\alpha_i) \;\geq\; \overline{u}_i(a_i, \alpha_i). \quad (1) \end{aligned}$$

Clearly if the system (1) has a solution $z \in \mathbb{R}^{nA}$ then the mechanism $m = (d, t)$, where,

$$\forall a \in A, \forall i \in N, \, t_i(a) \stackrel{\text{def}}{=} z_i(a) - \frac{1}{n-1} \sum_{j \neq i} z_j(a),$$

is both budget balancing and Bayesian incentive compatible. Thus we have to show that the system (1) is consistent. The proof is divided into three steps.

Step 1. $\forall \lambda \in \Lambda$, *if*

$$\forall \alpha \in A, \forall i \in N, \, p_i(\alpha_{-i}|\alpha_i) \sum_{\substack{a_i \in A_i \\ a_i \neq \alpha_i}} \lambda_i(a_i, \alpha_i) \; = \; \sum_{\substack{a_i \in A_i \\ a_i \neq \alpha_i}} \lambda_i(\alpha_i, a_i) p_i(\alpha_{-i}|a_i), \quad (2)$$

then

$$\sum_{i \in N} \sum_{(a_i, \alpha_i) \in C_i} \lambda_i(a_i, \alpha_i) \overline{u}_i(a_i, \alpha_i) \; \leq 0. \quad (3)$$

Proof of step 1. Since d is outcome efficient, we have, $\forall i \in N$, $\forall \alpha_i \in A_i$, $\forall a_i \in A_i$,

$$\overline{u}_i(a_i, \alpha_i) \; \leq \sum_{\alpha_{-i} \in A_{-i}} \left[\sum_{j \neq i} u_j(\alpha_i, \alpha_{-i}; \alpha_j) - \sum_{j \neq i} u_j(a_i, \alpha_{-i}; \alpha_j) \right] p_i(\alpha_{-i}|\alpha_i).$$

This implies that for any $\lambda \in \Lambda$ we get:

$$\begin{aligned}
&\sum_{i \in N} \sum_{(a_i, \alpha_i) \in C_i} \lambda_i(a_i, \alpha_i) \overline{u}_i(a_i, \alpha_i) \\
&\leq \sum_{i \in N} \sum_{\alpha_i \in A} \sum_{\alpha_{-i} \in A_{-i}} \sum_{j \neq i} u_j(\alpha_i, \alpha_{-i}; \alpha_j) p_i(\alpha_{-i}|\alpha_i) \left[\sum_{\substack{a_i \in A_i \\ a_i \neq \alpha_i}} \lambda_i(a_i, \alpha_i) \right] \\
&- \sum_{i \in N} \sum_{a_i \in A_i} \sum_{\alpha_{-i} \in A_{-i}} \sum_{j \neq i} u_j(a_i; \alpha_{-i}; \alpha_j) \left[\sum_{\substack{\alpha_i \in A_i \\ \alpha_i \neq a_i}} \lambda_i(a_i, \alpha_i) p_i(\alpha_{-i}|\alpha_i) \right] = \\
&\sum_{i \in N} \sum_{\alpha \in A} \sum_{j \neq i} u_j(\alpha_i, \alpha_{-i}; \alpha_j) \left[p_i(\alpha_{-i}|\alpha_i) \sum_{\substack{a_i \in A_i \\ a_i \neq \alpha_i}} \lambda_i(a_i, \alpha_i) - \sum_{\substack{a_i \in A_i \\ a_i \neq \alpha_i}} \lambda_i(\alpha_i, a_i) p_i(\alpha_{-i}|a_i) \right].
\end{aligned}$$

Assuming (2) we get (3).

Step 2. $\forall \lambda \in \Lambda$, *if*

$$\forall z \in \mathbb{R}^{nA}, \sum_{i \in N} \sum_{(a_i, \alpha_i) \in C_i} \lambda_i(a_i, \alpha_i) \sum_{\alpha_{-i} \in A_{-i}} \left[z_i(\alpha_i, \alpha_{-i}) - \frac{1}{n-1} \sum_{j \neq i} z_j(\alpha_i, \alpha_{-i}) \right.$$

$$-z_i(a_i, \alpha_{-i}) + \frac{1}{n-1} \sum_{j \neq i} z_j(a_i, \alpha_{-i}) \Bigg] p_i(\alpha_{-i}|\alpha_i) = 0, \tag{4}$$

then, $\exists \kappa \in \mathbb{R}^A$ *such that :*

$$\forall \alpha \in A, \forall i \in N, \; p_i(\alpha_{-i}|\alpha_i) \sum_{\substack{a_i \in A_i \\ a_i \neq \alpha_i}} \lambda_i(a_i, \alpha_i) - \sum_{\substack{a_i \in A_i \\ a_i \neq \alpha_i}} \lambda_i(\alpha_i, a_i) p_i(\alpha_{-i}|a_i) = \kappa(\alpha). \tag{5}$$

Proof of step 2. Take $z \in \mathbb{R}^{nA}$ such that for some $k \in N$ and $\alpha_0 \in A$ we have $z_k(\alpha^0) > 0$ and $\forall j \neq k$, $\forall \alpha \neq \alpha^0$, $z_j(\alpha^0) = 0$. Then, by (4), we get:

$$z_k(\alpha^0) \Bigg[p_k(\alpha^0_{-k}|\alpha^0_k) \sum_{\substack{a_k \in A_k \\ a_k \neq \alpha^0_k}} \lambda_k(a_k, \alpha^0_k) - \frac{1}{n-1} \sum_{i \neq k} p_i(\alpha^0_{-i}|\alpha^0_i) \sum_{\substack{a_i \in A_i \\ a_i \neq \alpha^0_i}} \lambda_i(a_i, \alpha^0_i)$$

$$- \sum_{\substack{\alpha_k \in A_k \\ \alpha_k \neq \alpha^0_k}} \lambda_k(\alpha^0_k, \alpha_k) p_k(\alpha^0_{-k}|\alpha_k) + \frac{1}{n-1} \sum_{i \neq k} \sum_{\substack{\alpha_i \in A_i \\ \alpha_i \neq \alpha^0_i}} \lambda_i(\alpha^0_i, \alpha_i) p_i(\alpha^0_{-i}|\alpha_i) \Bigg] = 0$$

or, equivalently,

$$p_k(\alpha^0_{-k}|\alpha^0_k) \sum_{\substack{a_k \in A_k \\ a_k \neq \alpha^0_k}} \lambda_k(a_k, \alpha^0_k) - \sum_{\substack{\alpha_k \in A_k \\ \alpha_k \neq \alpha^0_k}} \lambda_k(\alpha^0_k, \alpha_k) p_k(\alpha^0_{-k}|\alpha_k)$$

$$= \frac{1}{n} \left[\sum_i p_i(\alpha^0_{-i}|\alpha^0_i) \sum_{\substack{a_i \in A_i \\ a_i \neq \alpha_i}} \lambda_i(a_i, \alpha^0_i) - \sum_i \sum_{\substack{\alpha_i \in A_i \\ \alpha_i \neq \alpha^0_i}} \lambda_i(\alpha^0_i, \alpha_i) p_i(\alpha^0_{-i}|\alpha_i) \right].$$

Since the same argument holds for every $k \in N$ and every $\alpha_0 \in A$, we may conclude that (4) implies (5).

Step 3. By the compatitibility condition, for any $\lambda \in \Lambda$, *(5) implies (2).*

Proof of step 3. First we note that if $\lambda \in \Lambda$ satisfies (5), then $\forall i \in N$, $\forall \alpha \in A$,

$$\sum_{\substack{a_i \in A_i \\ a_i \neq \alpha_i}} \lambda_i(a_i, \alpha_i) - \sum_{\substack{a_i \in A_i \\ a_i \neq \alpha_i}} \lambda_i(\alpha_i, a_i) =$$

$$\sum_{\alpha_{-i} \in A_{-i}} \left[p_i(\alpha_{-i}|\alpha_i) \sum_{\substack{a_i \in A_i \\ a_i \neq \alpha_i}} \lambda_i(a_i, \alpha_i) - \sum_{\substack{a_i \in A_i \\ a_i \neq \alpha_i}} \lambda_i(\alpha_i, a_i) p_i(\alpha_{-i}|a_i) \right] = \sum_{a_{-i} \in A_{-i}} \kappa(\alpha_{-i}, \alpha_i)$$

$$= \sum_{a_{-i} \in A_{-i}} \left[p_j(\alpha_{-j}|\alpha_j) \sum_{\substack{a_j \in A_j \\ a_j \neq \alpha_j}} \lambda_j(a_j, \alpha_j) - \sum_{\substack{a_j \in A_j \\ a_j \neq \alpha_j}} \lambda_j(\alpha_j, a_j) p_j(\alpha_{-j}|a_j) \right] = 0.$$

Now suppose that $\lambda \in \Lambda$ is such that (5) holds but not (2). Then $\exists \kappa \in \mathbb{R}^A$

$$\forall \alpha \in A, \forall i \in N, \; p_i(\alpha_{-i}|\alpha_i) \sum_{\substack{a_i \in A_i \\ a_i \neq \alpha_i}} \lambda_i(\alpha_i, a_i) - \sum_{\substack{a_i \in A_i \\ a_i \neq \alpha_i}} \lambda_i(\alpha_i, a_i) p_i(\alpha_{-i}|a_i) = \kappa(\alpha),$$

but $\kappa \neq 0$. But this contradicts the compatibility condition.

In summary, we thus get that, for any $\lambda \in \Lambda$, (4) implies (5) (by step 2), (5) implies (2) (by step 3) and (2) implies (3) (by step 1). Hence, $\forall \lambda \in \Lambda$, (4) implies (3). But this is a well-known sufficient condition for system (1) to be consistent [see for example Fan (1956, theorem 1)].

■

The compatibility condition and the argument of theorem 7 can be stated for the general class of beliefs we have been considering previously. However to get a theorem analogous to theorem 7 we need to introduce an assumption like H_2 (instead of H_1) and to use some representation theorem of functionals in terms of conjugate spaces. We propose to discuss this matter elsewhere.

5 Conclusion

The main contribution of this paper is to show that by using the Bayesian approach to incomplete information, one may find mechanisms to solve efficiently a collective decision problem, which ensure simultaneously incentive compatibility and budget equilibrium. This positive result however relies on a compatibility condition which is imposed on the beliefs of the agents. Whether this condition is not only sufficient but also necessary remains an open question the answer to which would give a corresponding impossibility result.

The compatibility condition includes, as a particular case, the requirement of independence of the players' beliefs with respect to their own type. This independence condition implies that all players' beliefs are fully known. However, one can also consider it as associated to a statistical experiment, as Green, Laffont and Kohlberg (1976, p.384) who 'assume that each of the individuals in the society believes that all of the others are drawn indepently from a normal population with zero mean'.[13] More generally, on the basis of some preliminary empirical evidence, all agents may agree on some class of individual beliefs satisfying some compatibility condition and ask the central agency to reject any announcement outside this restricted class.

[13]With this assumption and in a public good model where, under their proposed mechanism the collected taxes are distributed equally to all players, the authors show that the best reply (in expectations) of each player and for each type, in the communication game, approaches the truth as the number of players increases.

References

Aumann, R.J. (1975), "Agreeing to disagree," Technical Report no.187, Institute for Mathematical Studies in the Social Sciences, Stanford University.

Clarke, E.H. (1971), "Multipart pricing of public goods," *Public Choice* II, 19–33.

d'Aspremont, C., and L.-A. Gérard-Varet (1975), "Individual incentives and collective efficiency for an externality game with incomplete information," CORE Discussion Paper 7519, Université catholique de Louvain.

Drèze, J.H. and D. de la Vallée Poussin (1971), "A tâtonnement process for public goods," *Review of Economic Studies* 38, 133–150.

Fan, K. (1956), "On systems of linear inequalities," in H.W. Kuhn and A.W. Tucker (eds). *Linear inequalities and related systems*. Princeton University Press, Princeton.

Groves, T. (1973), "Incentives in teams," *Econometrica* 14, 617–631.

Groves, T. (1975), "Incentives compatible control of decentralized organization," Discussion Paper no.166, Center for Mathematical Studies in Economics and Management Science, Northwestern University.

Groves, T. and M. Loeb (1975), "Incentives and public inputs," *Journal of Public Economics* 4, 211–226.

Groves, T. and J.O. Leyard (1977), "Optimal allocation of public goods: a solution to the free-rider problem," *Econometrica* 45, 783–809.

Green, J. and J.J. Laffont (1977), "Characterization of satisfactory mechanisms for the revelation of preferences for public goods," *Econometrica* 45, 427–438.

Green, J. and J.J. Laffont (1976), "Révélation des préférences pour les biens publics - 1ère partie: caractérisation des mécanismes satisfaisants," A125-0176, Laboratoire d'Econométrie de l'Ecole Polytechnique, Paris.

Green, J., E. Kohlberg and J.J. Laffont (1976), "Partial equilibrium approach to the free-rider problem," *Journal of Public Economics* 6, 375–394.

Harsanyi, J.C. (1967-68), "Games with incomplete information played by 'Bayesian' players, Parts I-III," *Management Science* 14, 158–182, 320–334, 486–502.

Hurwicz, L. (1971), "Centralization and decentralization in economic processes," in A. Eckstein (ed). *Comparison of economic systems: theoretical and methodological approaches*. University of California Press, Berkeley, Ch. 2.

Hurwicz, L. (1972), "On informationally decentralized systems," in R. Radner and C.D. McGuire (eds). *Decision and organization*. North-Holland, Amsterdam.

Lindhal, E. (1919), "Die Gerechtigkeit der Besteuerung," translated in R. Musgrave and A.T. Peacock (eds), 1958. *Classics in the theory of public finance*. MacMillan, London.

Luce, R.D. and H. Raiffa (1975), *Games and decisions*. L. Wiley, New York.

Neveu, J. (1970), *Bases mathématiques du calcul des probabilités*. Masson, Paris.

Samuelson, P.A. (1969), "Pure theory of public expenditure and taxation," in J. Margolis and H. Guitton (eds). *Public economics*. MacMillan, London.

Smets, H. (1973), "Le principe de la compensation réciproque. Un instrument économique pour la solution de certains problèmes de pollution transfrontière," O.C.D.E., Direction de l'Environnement.

Vickrey, W. (1961), "Counterspeculation, auctions and competitive sealed tenders," *Journal of Finance* 16, 8–37.

Wicksell, K. (1896), "Finanztheoretische Untersuchungen und das steuerwesen Schweden's," translated in R. Musgrave and A.T. Peacock (eds), 1958. *Classics in the theory of public finance*. MacMillan, London.