Incentives and the Existence of Pareto-Optimal Revelation Mechanisms *

Claude d'Aspremont
CORE, Université catholique de Louvain, Louvain-la-Neuve, Belgium
Jacques Crémer
Virginia Polytechnic Institute and State University, Blacksburg, Virginia
Louis-André Gérard-Varet
GREQE, Ecole des Hautes Etudes en Sciences Sociales, Marseille, France

Abstract

From the characterization of strongly and Bayesian incentive compatible Pareto-optimal mechanisms with transferable utilities, we derive the following results. If there are only two types per individual then a strongly incentive compatible Pareto-optimal mechanism exists. If there are only two individuals (with more than three types) then there are sets of beliefs (open in the class of all beliefs) for which no Bayesian incentive compatible Pareto-optimal mechanism exists. If there are more than two individuals then the class of beliefs for which such mechanisms exist is open and dense in the class of all beliefs.

JEL Classification: 020, 026, 027

1. Introduction

The present work belongs to a long series of papers studying conditions under which correct revelation of individual characteristics is obtained for efficient collective decision-making when information is incomplete and can be used strategically by the agents. More specifically we are interested in getting first-best (or full information) Pareto-optimality using redistributive transfers so as to elicit truthful information from individuals (see Arrow, 1979; d'Aspremont and Gérard-Varet, 1975, 1979, 1982; Bensaid, 1986; Créemer and McLean, 1985).

Besides the assumption that individuals behave on the basis of perfectly transferable utilities, a major restriction in the literature is about the beliefs (or subjective probabilities) that each one may have about any other's characteristics. In previous works there were conditions given on these beliefs such that, whatever may be the (perfectly transferable) individual utilities, first-best Pareto-optimality could be obtained under "Bayesian" incentive constraints. The so-called "independence" assumption is only one such condition. However, in d'Aspremont and Gérard-Varet (1979, 1982) a weaker

∗. Reprinted from *Journal of Economic Theory*, 51(2), 233–254, 1990.
Part of this work was done while the authors were visiting (respectively) the GSB and the IMSSS at Stanford University, the Laboratoire d'Econométrie de l'Ecole Polytechnique de Paris, and the Economics Department of the University of California at San Diego. We are grateful to these institutions for their hospitality as well as to the CIM, Belgium, the NSF (Grant SES-8408962), the Commissariat Général du Plan (MDPE 26/1985), and the Direction Générale de la Technologie du Ministère de la Recherche et de l'Enseignement Supérieur, Paris, for their financial support. Thanks are also due to S. Williams and M. Whinston for fruitful discussions.

(sufficient) condition was introduced, but its degree of generality remained unclear. Our purpose here is to clarify the matter. We show two main results. On one hand, when there are two agents, there exist a non-negligible set of beliefs and utilities such that it is impossible to find a first-best mechanism under Bayesian incentive constraints. On the other hand, when there are at least three agents, for "nearly all" beliefs, such a mechanism can be found.

It should be recalled that the explicit introduction of beliefs was motivated by the difficulty of solving the revelation problem without them. In that case the incentive constraints were formulated as a "dominant strategy equilibrium," thus avoiding completely the specification of these beliefs. Then the problem could be solved only by violating the budget constraint imposed on the transfers (see e.g., Harsanyi 1967-68, Groves and Ledyard 1987, and the survey given in Groves and Loeb 1975), except in some special cases. In the following, we shall give such a case: the one where every individual can be of two possible types only. This first result will be proved in Section 2 and will be based on the characterization of the dominant strategy incentive constraints in terms of dual variables. Also in this section we shall characterize analogously the "Bayesian" incentive constraints, those using explicitly the individual beliefs, and recall conditions on them leading to a Pareto-optimal solution. Then in Section 3 we shall show the contrast between having two and having more than two individuals. More specifically, when there are two individuals, one can always find pairs of utility functions such that there exist no Pareto-optimal solution for open sets of individual beliefs in the space of all beliefs; by contrast, when there are more than two individuals, there exist Pareto-optimal solutions, whatever the utility functions, for an open dense subset of individual beliefs.

2. Incentive compatibility as a finite system of linear inequalities

In the following we shall consider two kinds of decentralized collective decision problems. For both kinds, we start from a set $N = \{1, 2, \ldots, i, \ldots, n\}$ of individuals who have to choose an alternative in a compact set X of admissible *outcomes*. For concreteness, one may imagine that X is a set of public projects, but this is just an example. Every individual i is characterized by a parameter α_i, describing all his private information and taking value in a set A_i, that we assume finite. Such an α_i will be called the *type* of individual i. Individual i is supposed to know his own type but to have incomplete information about the types of the others. Moreover each individual i, of type α_i, evaluates each outcome x through a utility function $u_i(x;\alpha_i)$, assumed to be continuous in x, and the quantity $u_i(x;\alpha_i)$ representing the willingness to pay of agent i of type α_i for outcome x; the utility of agent i is $u_i(x;\alpha_i) + t_i$, where t_i is any monetary transfer. A *mechanism* is a pair (s,t), where s is an *outcome function* from $A \equiv \times_{i=1}^{n} A_i$ to X and where t is a *transfer function* from A to $\mathbb{R}^N$. The outcome function is *efficient* if, for all $\alpha \in A$,

$$\sum_{i=1}^{n} u_i(s(\alpha);\alpha_i) = \max_{x\in X} \sum_{i=1}^{n} u_i(x;\alpha_i),$$

and the transfer function is *balanced*, if, for all $\alpha \in A$,

$$\sum_{i=1}^{n} t_i(\alpha) = 0.$$

When s is efficient and t balanced, then the mechanism (s,t) is called *Pareto-optimal*. For any individual i, a decision rule $a_i(\cdot)$ is a function from A_i to itself, associating to every "true" type α_i the "announced" type $\tilde{\alpha}_i = a_i(\alpha_i)$. The problem is to design a Pareto-optimal mechanism where each individual i finds in his self-interest to use the "truth-telling" strategy, that is the identity function $(a_i(\alpha_i) = \alpha_i$, for all $\alpha_i \in A_i)$. However, the self-interest criterion may vary. In the literature two kinds of criteria have been used leading to two different versions of this revelation problem.

In a first version, one supposes that every i is "completely ignorant" of the vector $\alpha_{-i} = (\alpha_1, \alpha_2, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_n)$ in $A_{-i} = \bigtimes_{j \neq i} A_j$, and hence one requires that i announce the truth whatever the types of the other agents. Denoting by $\mathcal{A}_i^2$ the set $\{(\tilde{\alpha}_i, \alpha_i) \in A_i \times A_i : \tilde{\alpha}_i \neq \alpha_i\}$ one gets a first system of incentive compatibility inequalities that may be imposed on a mechanism (s, t):

$$\begin{aligned}
&\forall\, i \in N, \forall\, (\tilde{\alpha}_i, \alpha_i) \in \mathcal{A}_i^2, \forall\, \alpha_{-i} \in A_{-i},\\
&u_i\big(s(\alpha_i, \alpha_{-i}); \alpha_i\big) + t_i(\alpha_i, \alpha_{-i}) \geq u_i\big(s(\tilde{\alpha}_i, \alpha_{-i}); \alpha_i\big) + t_i(\tilde{\alpha}_i, \alpha_{-i}).
\end{aligned} \tag{I}$$

We can interpret these inequalities by looking at all the games in normal form (with complete information) generated by all possible parameter values α in A, taking for every i the payoff function $u_i(s(\cdot); \alpha_i) + t_i(\cdot)$ and the strategy space A_i; we see that the strategy $\tilde{\alpha}_i = \alpha_i$ is a dominant strategy for i. Following Green and Laffont (1979), we call a mechanism satisfying (I) *strongly incentive compatible* (SIC). This first version of the revelation problem will be denoted RP. It is summarized by

$$\mathrm{RP} = (X; A_1, \ldots, A_n; u_1, \ldots, u_n).$$

The second version of the revelation problem is obtained from a Bayesian point of view. Following Harsanyi (1967-68), we associate to every mechanism (s, t) a game with incomplete information including explicitly the beliefs (or subjective probabilities) of every individual i concerning the others' types.[1] This is given, for every i, by a probability transition $p_i(\alpha_{-i} \mid \alpha_i)$ from A_i to A_{-i}. For some results in the sequel we shall impose the reasonable assumption that these beliefs are *consistent* (see Holmström and Myerson, 1983): the individual conditional beliefs are derived from some joint probability distribution p on A; then $p_i(\cdot \mid \alpha_i)$ is the probability distribution computed from p by conditionalizing on α_i the known parameter of i. A family of beliefs $\{p_i; i \in N\}$ is also called an *information structure* (IS). The payoff of player i of type α_i is evaluated, for every strategy vector $(a_1(\cdot), \ldots, a_n(\cdot))$, as the conditional expected utility

$$\sum_{\alpha_{-i}} \big[u_i(s(a_1(\alpha_1), \ldots, a_n(\alpha_n)); \alpha_i) + t_i(a_1(\alpha_1), \ldots, a_n(\alpha_n))\big] p_i(\alpha_{-i} \mid \alpha_i).$$

A mechanism (s, t) is *Bayesian Incentive Compatible* (BIC) if the truthtelling strategy is a Bayesian equilibrium in the sense of Harsanyi; it satisfies the following inequalities

$$\begin{aligned}
&\forall\, i \in N, \forall\, (\tilde{\alpha}_i, \alpha_i) \in \mathcal{A}_i^2,\\
&\sum_{\alpha_{-i}} \big[u_i(s(\alpha_i, \alpha_{-i}); \alpha_i) + t_i(\alpha_i, \alpha_{-i})\big] p_i(\alpha_{-i} \mid \alpha_i)\\
\geq &\sum_{\alpha_{-i}} \big[u_i(s(\tilde{\alpha}_i, \alpha_{-i}); \alpha_i) + t_i(\tilde{\alpha}_i, \alpha_{-i})\big] p_i(\alpha_{-i} \mid \alpha_i).
\end{aligned} \tag{II}$$

This yields a second version of the revelation problem, where the beliefs of the individuals play an important role. Accordingly we call it a *Bayesian Revelation Problem*; we denote it BRP and summarize it by

$$\mathrm{BRP} = (X; A_1, \ldots, A_n; u_1, \ldots, u_n; p_1, \ldots, p_n).$$

1. For more properties see d'Aspremont and Gérard-Varet (1979, Section 3.1).

Clearly, whenever a mechanism (s,t) satisfies (I), it also satisfies system (II) for any IS with the same sets of types. Therefore if for a given RP we can find a Pareto-optimal SIC-mechanism, we an also find a Pareto-optimal BIC-mechanism for every BRP with the same sets of types. However, as will be seen later, Pareto-optimal SIC-mechanisms do not exist in general.

When we fix an outcome function s (efficient or not), looking for a balanced transfer function t satisfying the BIC inequalities (or the SIC inequalities) amounts to solving a (finite) system of linear inequalities. Hence we can easily characterize the class of revelation problems and outcome functions for which the system is consistent, using standard results on linear inequalities. This is stated in two lemmata that will turn out to be very useful in the sequel.

Lemma 1 *Let* $\Lambda \overset{def}{=} \times_{i=1}^{n} \mathbb{R}_+^{\mathcal{A}_i^2}$. *For any* BRP $= (X; A_1, \ldots, A_n; u_1, \ldots, u_n; p_1, \ldots, p_n)$ *and any outcome function s there exists a balanced transfer function t such that (s,t) is a BIC-mechanism if and only if,* $\forall \lambda \in \Lambda$, *if*

$$\begin{aligned}
&\forall\, i, j \in N, \forall\, \alpha \in A, \\
&\quad p_i(\alpha_{-i} \mid \alpha_i) \sum_{\tilde{\alpha}_i \neq \alpha_i} \lambda_i(\tilde{\alpha}_i, \alpha_i) - \sum_{\tilde{\alpha}_i \neq \alpha_i} \lambda_i(\alpha_i, \tilde{\alpha}_i) p_i(\alpha_{-i} \mid \tilde{\alpha}_i) \\
&\qquad = p_j(\alpha_{-j} \mid \alpha_j) \sum_{\tilde{\alpha}_j \neq \alpha_j} \lambda_j(\tilde{\alpha}_j, \alpha_j) - \sum_{\tilde{\alpha}_j \neq \alpha_j} \lambda_j(\alpha_j, \tilde{\alpha}_j) p_j(\alpha_{-j} \mid \tilde{\alpha}_j)
\end{aligned} \tag{1}$$

then

$$\sum_{i=1}^{n} \sum_{(\tilde{\alpha}_i, \alpha_i) \in \mathcal{A}_i^2} \lambda_i(\tilde{\alpha}_i, \alpha_i) \sum_{\alpha_{-i}} \Big[u_i(s(\tilde{\alpha}_i, \alpha_{-i}); \alpha_i) - u_i\big(s(\alpha_i, \alpha_{-i}); \alpha_i\big) \Big] p_i(\alpha_{-i} \mid \alpha_i) \leq 0. \tag{2}$$

Proof The set of inequalities

$$\begin{aligned}
&\sum_{\alpha_{-i}} \big[t_i(\alpha_i, \alpha_{-i}) - t_i(\tilde{\alpha}_i, \alpha_{-i}) \big] p_i(\alpha_{-i} \mid \alpha_i) \\
&\qquad\qquad \geq \sum_{\alpha_{-i}} \big[u_i\big(s(\tilde{\alpha}_i, \alpha_{-i}); \alpha_i\big) - u_i\big(s(\alpha_i, \alpha_{-i}); \alpha_i\big) \big] p_i(\alpha_{-i} \mid \alpha_i),
\end{aligned}$$

for $i \in N$, $(\tilde{\alpha}_i, \alpha_i) \in \mathcal{A}_i^2$, together with $\sum_{i=1}^{n} t_i(\alpha) = 0$, for $\alpha \in A$, can have no solution t if there exist $\lambda \in \Lambda$ and $\mu \in \mathbb{R}^A$ such that

$$\begin{aligned}
&\forall\, i \in N, \forall\, \alpha \in A, \\
&\quad p_i(\alpha_{-i} \mid \alpha_i) \sum_{\tilde{\alpha}_i \neq \alpha_i} \lambda_i(\tilde{\alpha}_i, \alpha_i) - \sum_{\tilde{\alpha}_i \neq \alpha_i} \lambda_i(\alpha_i, \tilde{\alpha}_i) p_i(\alpha_{-i} \mid \tilde{\alpha}_i) + \mu(\alpha) = 0
\end{aligned} \tag{1.bis}$$

and

$$\sum_{i=1}^{n} \sum_{(\tilde{\alpha}_i, \alpha_i) \in \mathcal{A}_i^2} \lambda_i(\tilde{\alpha}_i, \alpha_i) \sum_{\alpha_{-i}} \Big[u_i\big(s(\tilde{\alpha}_i, \alpha_{-i}); \alpha_i\big) - u_i\big(s(\alpha_i, \alpha_{-i}); \alpha_i\big) \Big] p_i(\alpha_{-i} \mid \alpha_i) > 0.$$

Otherwise, taking the solution t, multiplying both sides of (1.bis) by $t_i(\alpha)$ and summing over i and α, one would get

$$0 = \sum_{i=1}^{n} \sum_{\alpha} t_i(\alpha) \Big[p_i(\alpha_{-i} \mid \alpha_i) \sum_{\tilde{\alpha}_i \neq \alpha_i} \lambda_i(\tilde{\alpha}_i, \alpha_i) \\ - \sum_{\tilde{\alpha}_i \neq \alpha_i} \lambda_i(\alpha_i, \tilde{\alpha}_i) p_i(\alpha_{-i} \mid \tilde{\alpha}_i) + \mu(\alpha) \Big] \times \sum_{\alpha_{-i}} \big[t_i(\alpha_i, \alpha_{-i}) - t_i(\tilde{\alpha}_i, \alpha_{-i}) \big] p_i(\alpha_{-i} \mid \alpha_i) > 0,$$

a contradiction.

Since (1) and (1.bis) are equivalent, the result follows from a variant of Farkas lemma (see Gale, 1960, pp. 44–47). ■

By the same reasoning we can derive a similar statement about RPS.

Lemma 2 *Let* $M \overset{def}{=} \times_{i=1}^{n} \mathbb{R}_{+}^{\mathcal{A}_i^2 \times A_{-i}}$. *For any RP=* $(X; A_1, \ldots, A_n; u_1, \ldots, u_n)$ *and any outcome function* s *there exists a balanced transfer function* t *such that* (s, t) *is a SIC-mechanism if and only if,* $\forall\, \mu \in M$*, if*

$$\forall\, i, j \in N, \forall\, \alpha \in A, \sum_{\tilde{\alpha}_i \neq \alpha_i} \mu_i(\tilde{\alpha}_i, \alpha_i, \alpha_{-i}) - \sum_{\tilde{\alpha}_i \neq \alpha_i} \mu_i(\alpha_i, \tilde{\alpha}_i, \alpha_{-i}) \\ = \sum_{\tilde{\alpha}_j \neq \alpha_j} \mu_j(\tilde{\alpha}_j, \alpha_j, \alpha_{-j}) - \sum_{\tilde{\alpha}_j \neq \alpha_j} \mu_j(\alpha_j, \tilde{\alpha}_j, \alpha_{-j}) \tag{3}$$

then

$$\sum_{i=1}^{n} \sum_{(\tilde{\alpha}_i, \alpha_i) \in \mathcal{A}_i^2} \sum_{\alpha_{-i}} \mu_i(\tilde{\alpha}_i, \alpha_i, \alpha_{-i}) \times [u_i(s(\tilde{\alpha}_i, \alpha_{-i}); \alpha_i) - u_i(s(\alpha_i, \alpha_{-i}); \alpha_i)] \leq 0. \tag{4}$$

To illustrate immediately the usefulness of this "dual" approach, we give two applications of these lemmata.

2.1 Application of Lemma 2

Using this last result, we can prove first that there exists a Pareto-optimal SIC-mechanism whenever each individual has only two types, that is:

Theorem 1 *If* $|A_i| = 2$ *for* $i = 1, 2, \ldots, n$*, then, whatever the utility functions and the beliefs, there exists a Pareto-optimal SIC-mechanism.*

Proof Take any $\mu \in M$ and assume that (3) holds. We have to show that (4) also holds. But since $|A_i| = 2, i = 1, 2, \ldots, n$, there must be, by (3), some $\alpha^0 \in A$ such that for $\tilde{\alpha}_i \neq \alpha_i^0$, $i = 1, 2, \ldots, n$, and some $\rho \in \mathbb{R}_+$,

$$\begin{aligned} &\mu_1(\tilde{\alpha}_1, \alpha_1^0, \alpha_{-1}^0) - \mu_1(\alpha_1^0, \tilde{\alpha}_1, \alpha_{-1}^0) \\ =&\mu_2(\tilde{\alpha}_2, \alpha_2^0, \alpha_{-2}^0) - \mu_2(\alpha_2^0, \tilde{\alpha}_2, \alpha_{-2}^0) = \ldots \\ =&\mu_n(\tilde{\alpha}_n, \alpha_n^0, \alpha_{-n}^0) - \mu_n(\alpha_n^0, \tilde{\alpha}_n, \alpha_{-n}^0) = \rho \geq 0. \end{aligned}$$

Clearly, if we permute $\tilde{\alpha}_i$ with α_i^0 for one (or an odd number of) individual(s) we get negative equalities ($= -\rho$). Similarly, if we permute $\tilde{\alpha}_i$ with α_i^0 for an even number of individuals we get positive equalities ($= \rho$). Let us therefore define $[\alpha^0]$ as the set

$$\{\alpha \in A : \text{for some } S \subset N \text{ with } |S| \text{ even}, \alpha_i \neq \alpha_i^0, \text{if } i \in S, \text{ and } \alpha_i = \alpha_i^0, \text{ if } i \in N - S\}.$$

We have $\alpha \in [\alpha^0]$ if and only if, for $\tilde{\alpha}_i \neq \alpha_i$, $i = 1, 2, \ldots, n$,

$$\mu_1(\tilde{\alpha}_1, \alpha_1, \alpha_{-1}) - \mu_1(\alpha_1, \tilde{\alpha}_1, \alpha_{-i}) = \ldots = \mu_n(\tilde{\alpha}_n, \alpha_n, \alpha_{-n}) - \mu_n(\alpha_n, \tilde{\alpha}_n, \alpha_{-n}) = \rho \geq 0.$$

Thus (4) can be written as follows: for $\tilde{\alpha}_i \neq \alpha_i$, $i = 1, 2, \ldots, n$,

$$\sum_{i=1}^{n} \sum_{\alpha \in [\alpha^0]} \Big\{ \mu_i(\alpha_i, \tilde{\alpha}_i, \alpha_{-i}) \, [u_i(s(\tilde{\alpha}_i, \alpha_{-i}); \alpha_i) - u_i(s(\alpha_i, \alpha_{-i}); \alpha_i) \\ + u_i(s(\alpha_i, \alpha_{-i}); \tilde{\alpha}_i) - u_i(s(\tilde{\alpha}_i, \alpha_{-i}); \tilde{\alpha}_i)] \\ + \rho[u_i(s(\tilde{\alpha}_i, \alpha_{-i}); \alpha_i) - u_i(s(\alpha_i, \alpha_{-i}); \alpha_i)] \Big\} \leq 0.$$

Observe that, for all α and all i, using the efficiency of s and the convention $\tilde{\alpha}_i \neq \alpha_i$, we have

$$\begin{aligned} & u_i(s(\tilde{\alpha}_i, \alpha_{-i}); \alpha_i) - u_i(s(\alpha_i, \alpha_{-i}); \alpha_i) \\ & \qquad + u_i(s(\alpha_i, \alpha_{-i}); \tilde{\alpha}_i) - u_i(s(\tilde{\alpha}_i, \alpha_{-i}); \tilde{\alpha}_i) \\ & \leq \; - \sum_{j \neq i} u_j(s(\tilde{\alpha}_i, \alpha_{-i}); \alpha_j) + \sum_{j \neq i} u_j(s(\alpha_i, \alpha_{-i}); \alpha_j) \\ & \qquad - \sum_{j \neq i} u_j(s(\alpha_i, \alpha_{-i}); \alpha_j) + \sum_{j \neq i} u_j(s(\tilde{\alpha}_i, \alpha_{-i}); \alpha_j) = 0, \end{aligned}$$

so that, in order to get (4), it only remains to show that

$$\sum_{i=1}^{n} \sum_{\alpha \in [\alpha^0]} u_i(s(\tilde{\alpha}_i, \alpha_{-i}); \alpha_i) \leq \sum_{i=1}^{n} \sum_{\alpha \in [\alpha^0]} u_i(s(\alpha_i, \alpha_{-i}); \alpha_i).$$

In the case where n is odd, we get the result immediately since, by the definition of $[\alpha^0]$

$$\sum_{i=1}^{n} \sum_{\alpha \in [\alpha^0]} u_i(s(\tilde{\alpha}_i, \alpha_{-i}) = \sum_{i=1}^{n} \sum_{\alpha \in [\alpha^0]} u_i(s(\tilde{\alpha}); \alpha_i)$$

and by the efficiency of s

$$\sum_{\alpha \in [\alpha^0]} \sum_{i=1}^{n} u_i(s(\tilde{\alpha}); \alpha_i) \leq \sum_{\alpha \in [\alpha^0]} \sum_{i=1}^{n} u_i(s(\alpha); \alpha_i).$$

Besides, by simple changes of variables and the efficiency of s we can write

$$
\begin{aligned}
&\sum_{i=1}^{n} \sum_{\alpha\in[\alpha^0]} u_i(s(\tilde{\alpha}_i, \alpha_{-i}); \alpha_i) \\
&\quad = \sum_{i=1}^{n} \sum_{(\tilde{\alpha}_i, \alpha_{-i})\in[\alpha^0]} u_i(s(\alpha); \tilde{\alpha}_i) \\
&\quad = \sum_{\alpha\in A-[\alpha^0]} \sum_{i=1}^{n} u_i(s(\alpha); \tilde{\alpha}_i) \leq \sum_{\alpha\in A-[\alpha^0]} \sum_{i=1}^{n} u_i(s(\tilde{\alpha}); \tilde{\alpha}_i) \\
&\quad = \sum_{i=1}^{n} \sum_{(\tilde{\alpha}_i, \alpha_{-i})\in A-[\alpha^0]} u_i(s(\alpha_i, \tilde{\alpha}_{-i}); \alpha_i) \\
&\quad = \sum_{i=1}^{n} \sum_{\alpha\in[\alpha^0]} u_i(s(\alpha_i, \tilde{\alpha}_{-i}); \alpha_i).
\end{aligned}
$$

Therefore, in the case where n is even, by the definition of $[\alpha^0]$ and efficiency, we get for some $j \in N$,

$$
\begin{aligned}
&\sum_{i=1}^{n} \sum_{\alpha\in[\alpha^0]} u_i(s(\tilde{\alpha}_i, \alpha_{-i}); \alpha_i) \\
&\quad \leq \sum_{i\neq j}^{n} \sum_{\alpha\in[\alpha^0]} u_i(s(\tilde{\alpha}_i, \alpha_{-i}); \alpha_i) + \sum_{\alpha\in[\alpha^0]} u_j(s(\alpha_j, \tilde{\alpha}_{-j}); \alpha_j) \\
&\quad = \sum_{i\neq j}^{n} \sum_{\alpha\in[\alpha^0]} u_i(s(\alpha_j, \tilde{\alpha}_{-j}); \alpha_i) + \sum_{\alpha\in[\alpha^0]} u_j(s(\alpha_j, \tilde{\alpha}_{-j}); \alpha_j) \\
&\quad \leq \sum_{i=1}^{n} \sum_{\alpha\in[\alpha^0]} u_i(s(\alpha); \alpha_i).
\end{aligned}
$$

The result follows. ■

Since any SIC-mechanism is also a BIC-mechanism, we immediately get

Corollary 1 *If $|A_i| = 2$, for $i = 1, 2, \ldots, n$, then, whatever the utility functions and the beliefs, there is a Pareto-optimal BIC-mechanism.*

Maskin (1986b, Theorem 2) states this corollary and provides a proof only for the case $n = 2$. The corollary confirms his intuition for $n > 2$. Furthermore it is interesting to notice that using duality has led us to a stronger result: the existence of Pareto-optimal SIC-mechanisms (and not only BIC-mechanisms) for any n (provided each i has only two types). This might seem surprising since we know (see Green and Laffont, 1979; Walker, 1980) that in general there exists no balanced Groves mechanism, that balanced Groves mechanisms are Pareto-optimal SIC-mechanisms, and that, in many environments, Pareto-optimal SIC-mechanisms must necessarily be Groves mechanisms.[2]

2. Recall that a Groves mechanism is a mechanism satisfying $\forall\, i \in N, \forall\, \alpha \in A, t_i(\alpha) = \sum_{j\neq i} u_j(s(\alpha); \alpha_j) - f_i(\alpha_{-i})$ for some f_i, a real-valued function defined on A_{-i}.

Hence, the Pareto-optimal SIC-mechanism that exists when each agent has only two types may have to be chosen outside the class of Grove mechanisms. The following example shows this clearly.

EXAMPLE 1. Let $n = 2$, $A_i = \{\alpha_i^+, \alpha_i^-\}$, and $X = \{0, 1\}$. Suppose

(i) $u_i(0; \alpha_i) = 0,\ \forall\, \alpha_i \in A, i = 1, 2,$

(ii) $u_1(1; \alpha_1) + u_2(1; \alpha_2) > 0$ if and only if $\alpha_1 = \alpha_1^+$ and $\alpha_2 = \alpha_2^+$,

(iii) $u_1(1; \alpha_1^+) > u_1(1; \alpha_1^-) > 0 > u_2(1; \alpha_2^+) > u_2(1; \alpha_2^-)$.

The efficient outcome function s satisfies

$$s(\alpha_1, \alpha_2) = 1 \quad \text{if and only if} \quad u_1(1; \alpha_1) + u_2(1; \alpha_2) > 0.$$

The existence of a balanced transfer function t satisfying the SIC-inequalities can be checked directly: Indeed we need

$$u_1(1; \alpha_1^+) + t_1(\alpha_1^+, \alpha_2^+) \geq t_1(\alpha_1^-, \alpha_2^+)$$

and

$$\begin{aligned} u_2(1; \alpha_2^+) - t_1(\alpha_1^+, \alpha_2^+) &\geq -t_1(\alpha_1^+, \alpha_2^-), \\ t_1(\alpha_1^-, \alpha_2^+) &\geq u_1(1; \alpha_1^-) + t_1(\alpha_1^+, \alpha_2^+) \end{aligned}$$

and

$$\begin{aligned} -t_1(\alpha_1^-, \alpha_2^+) &\geq -t_1(\alpha_1^-, \alpha_2^-), \\ t_1(\alpha_1^-, \alpha_2^-) &\geq t_1(\alpha_1^+, \alpha_2^-) \end{aligned}$$

and

$$\begin{aligned} -t_1(\alpha_1^-, \alpha_2^-) &\geq -t_1(\alpha_1^-, \alpha_2^+), \\ t_1(\alpha_1^+, \alpha_2^-) &\geq t_1(\alpha_1^-, \alpha_2^-) \end{aligned}$$

and

$$-t_1(\alpha_1^+, \alpha_2^-) \geq u_2(1; \alpha_2^-) - t_1(\alpha_1^+, \alpha_2^+).$$

This system reduces to

$$\begin{aligned} t_1(1; \alpha_1^-, \alpha_2^+) = t_1(\alpha_1^-, \alpha_2^-) = t_1(\alpha_1^+, \alpha_2^-) = \tau, \\ u_1(1; \alpha_1^-) \leq \tau - t_1(\alpha_1^+, \alpha_2^+) \leq u_1(1; \alpha_1^+), \end{aligned}$$

and

$$-u_2(1; \alpha_2^+) \leq \tau - t_1(\alpha_1^+, \alpha_2^+) \leq -u_2(1; \alpha_2^-).$$

By (ii) and (iii) a possible solution is to fix $t_1(\alpha_1^+, \alpha_2^+)$ and let

$$\tau = \min\{u_1(1; \alpha_1^+), -u_2(1; \alpha_2^-)\} + t_1(\alpha_1^+, \alpha_2^+).$$

Although the previous system admits a solution, it cannot lead to a Groves mechanism for the transfer function would have to satisfy

$$\begin{aligned} t_1(\alpha_1, \alpha_2) &= u_2(s(\alpha_1, \alpha_2); \alpha_2) + f_1(\alpha_2) \\ &= -t_2(\alpha_1, \alpha_2) = -u_1(s(\alpha_1, \alpha_2); \alpha_1) - f_2(\alpha_1), \end{aligned}$$

implying that the functions f_1 and f_2 satisfy the system

$$\begin{aligned} f_1(\alpha_2) + f_2(\alpha_1) &< 0 \text{ if } \alpha_1 = \alpha_1^+ \text{ and } \alpha_2 = \alpha_2^+, \\ &= 0 \text{ otherwise} \end{aligned}$$

which is obviously inconsistent.

Theorem 1 is a positive result on the possibility of finding Pareto-optimal SIC-mechanisms. However the case of two types per agent is very special and, as already mentioned, one cannot expect a general existence result for such mechanisms (e.g., see Green and Laffont, 1979, pp. 90–96).

2.2 Application of Lemma 1

Turning now to Bayesian revelation problems, we may, following d'Aspremont and Gérard-Varet (1979), apply Lemma 1 to derive various conditions on the individual beliefs sufficient to ensure the existence of Pareto-optimal BIC-mechanisms. In the next section we shall see that some of these conditions are restrictive in the two-agent case, but hardly at all for more than two individuals. In that case indeed the main condition will be shown to hold generically.

Before presenting our results, let us briefly review the various sufficient conditions. The most general condition bearing only on individual beliefs $p_1, \ldots, p_n$, is called *condition C* (see d'Aspremont and Gérard-Varet, 1979, 1982). It requires that for every λ in Λ,

$$\begin{aligned} &\forall\, i \in N, \forall\, \alpha \in A, \\ &p_i(\alpha_{-i} \mid \alpha_i) \sum_{\tilde{\alpha}_i \neq \alpha_i} \lambda_i(\tilde{\alpha}_i, \alpha_i) - \sum_{\tilde{\alpha}_i \neq \alpha_i} \lambda_i(\alpha_i, \tilde{\alpha}_i) p_i(\alpha_{-i} \mid \tilde{\alpha}_i) = 0 \end{aligned} \tag{5}$$

whenever (1) of Lemma 1 holds. Indeed taking any BRP $= (X; A_1, \ldots, A_n; u_1, \ldots, u_n; p_1, \ldots, p_n)$ and any efficient outcome function s, we observe that (2) of Lemma 1 is implied by (5), since, using the efficiency of s,

$$\begin{aligned} &\sum_{i=1}^{n} \sum_{(\tilde{\alpha}_i, \alpha_i) \in \mathcal{A}_i^2} \lambda_i(\tilde{\alpha}_i, \alpha_i) \times \sum_{\alpha_{-i}} \big[u_i(s(\tilde{\alpha}_i, \alpha_{-i}); \alpha_i) - u_i(s(\alpha_i, \alpha_{-i}); \alpha_i)\big] p_i(\alpha_{-i} \mid \alpha_i) \\ \leq &\sum_{i=1}^{n} \sum_{(\tilde{\alpha}_i, \alpha_i) \in \mathcal{A}_i^2} \lambda_i(\tilde{\alpha}_i, \alpha_i) \times \sum_{\alpha_{-i}} \left[\sum_{j \neq i} u_j(s(\alpha_i, \alpha_{-i}); \alpha_j) - \sum_{j \neq i} u_j(s(\tilde{\alpha}_i, \alpha_{-i}); \alpha_j)\right] p_i(\alpha_{-i} \mid \alpha_i) \\ = &\sum_{i=1}^{n} \sum_{\alpha} \sum_{j \neq i} u_j(s(\alpha); \alpha_j) \times \left[p_i(\alpha_{-i} \mid \alpha_i) \sum_{\tilde{\alpha}_i \neq \alpha_i} \lambda_i(\tilde{\alpha}_i, \alpha_i) - \sum_{\tilde{\alpha}_i \neq \alpha_i} \lambda_i(\alpha_i, \tilde{\alpha}_i) p_i(\alpha_{-i} \mid \tilde{\alpha}_i)\right] \\ = &0 \qquad \text{(by (5))}. \end{aligned}$$

Hence, under condition C, there exists a Pareto-optimal BIC-mechanism.

Condition C is difficult to interpret. However, there are stronger conditions that are easier to understand. A first one is condition F, which requires that at least one agent have "free beliefs", namely

$$\exists\, i \in N \text{ such that } \forall\, (\tilde{\alpha}_i, \alpha_i) \in \mathcal{A}_i^2, p(\cdot \mid \tilde{\alpha}_i) \equiv p(\cdot \mid \alpha_i). \tag{6}$$

This amounts to saying that the beliefs of agent i are common knowledge. That it is stronger than condition C is clear since, whenever (1) holds for some λ, summing on both sides of (1) on α_{-i}, we get

$$\sum_{\tilde{\alpha}_i \neq \alpha_i} \lambda_i(\tilde{\alpha}_i, \alpha_i) - \sum_{\tilde{\alpha}_i \neq \alpha_i} \lambda_i(\alpha_i, \tilde{\alpha}_i) = 0 \quad \forall\, i \in N, \tag{7}$$

so that (1) and (6) imply (5). In fact the first positive result obtained for BRPs (see Arrow, 1979; d'Aspremont and Gérard-Varet, 1975), was based on an even more restrictive condition, the so-called "independent case" whereby all players are assumed to have free beliefs.

3. On the existence of Pareto-optimal BIC-mechanisms

We have just seen that, with transferable utilities, the search for Pareto-optimal BIC-mechanisms appears much more promising than the one for Pareto-optimal SIC-mechanisms. However, logically the situation is not altogether clear. It is known from d'Aspremont and Gérard-Varet (1982) that condition C reduces to the independent case when $n = 2$ and the beliefs are consistent. It is also known from d'Aspremont and Gérard-Varet (1982) that condition C is not a necessary condition: a two-agent example is exhibited where a Pareto-optimal BIC-mechanism exists but condition C is not satisfied. In this section we shall try to clarify the scope of Condition C. First, in the case $n = 2$, we shall show that there is no hope for a general positive result. In a world of two individuals there exists a subset of positive measure in the space of beliefs such that no Pareto-optimal BIC-mechanism exists. However, the scope of Condition C is much larger when the number of individuals is increased. Indeed, in the case $n > 3$, we shall prove that condition C is generic: The set of beliefs satisfying condition B contains an open and dense subset of the set of all consistent beliefs.

3.1 The case $n = 2$: A class of counterexamples

To construct a class of counterexamples we shall proceed as follows. First we will consider a particular revelation problem RP for which no Pareto-optimal SIC-mechanism exists. We will then use this RP to construct a Bayesian revelation problem, BRP, for which no Pareto-optimal BIC-mechanism exists. Furthermore, we will show that this non-existence also holds in an open neighborhood of the constructed problem.

Consider a RP with two individuals, $X = \{0, 1\}$ and $A_1 = A_2 = \{1, 2, 3\}$ (in view of the results of Section 2, we must have at least three types). Say that, for both players, the utility of decision 0 is equal to 0, whatever their type. The utilities of the agents, when the decision is 1, are

$$\begin{aligned} &u_1(1;1) = -10; \quad u_1(1;2) = -5; u_1(1;3) = 1; \\ &u_2(1;1) = -6; \qquad u_2(1;2) = 4; u_2(1;3) = 6. \end{aligned}$$

With such utilities, the efficient outcome function $s(\alpha_1, \alpha_2)$ is 0 in the states of the world (1,1), (1,2), (1,3), (2,1), (2,2), (3,1), and 1 in the others. It may be quickly shown that there is no balanced transfers allowing for a Pareto-optimal SIC-mechanism. Indeed, let $t_1(\cdot) = -t_2(\cdot) = t(\cdot)$, and write the SIC constraints

$$\begin{aligned} t(\alpha) - t(\tilde{\alpha}_1, \alpha_2) &\geq u_1(s(\tilde{\alpha}_1, \alpha_2); \alpha_1) - u_1(s(\alpha); \alpha_1), \\ t(\alpha_1, \tilde{\alpha}_2) - t(\alpha) &\geq u_2(s(\alpha_1, \tilde{\alpha}_2); \alpha_2) - u_2(s(\alpha); \alpha_2), \end{aligned}$$

for $(\tilde{\alpha}_i, \alpha_1, \alpha_2)$ equal to (3,1,1), (1,2,3), and (2,3,3) for $i = 1$ and equal to (1,3,2), (3,1,1), and (2,3,3) for $i = 2$. We thus get the system of inequalities

$$\begin{aligned}
t(1,1) - t(3,1) &\geq 0 \\
t(3,1) - t(3,2) &\geq -4 \\
t(3,2) - t(3,3) &\geq 0 \\
t(3,3) - t(2,3) &\geq 0 \\
t(2,3) - t(1,3) &\geq 5 \\
t(1,3) - t(1,2) &\geq 0.
\end{aligned}$$

Clearly there is a cycle ($t(\alpha, \beta)$ which appears only twice, once with coefficient +1, once with coefficient -1). By summation the left-hand sides cancel out where the right-hand sides sum up to 1. This shows that the system is inconsistent.

Using this RP (which admits no Pareto-optimal SIC-mechanism) we now construct a BRP and show that it has no Pareto-optimal BIC-mechanism. The basic idea is to model the dominant strategy problem in a Bayesian framework. Let there still be two agents and let $X = \{0, 1\}$. Define, for each agent, the identical set of types B, selecting in the product $A_1 \times A_2$ the points generated by the indices of the above SIC-constraints (namely, (3,1,1), (1,2,3), and (2,3,3) for $i = 1$, and (1,3,2), (3,1,1), and (2,3,3) for $i = 2$):

$$\begin{aligned}
B &= \{(3,1), (1,3), (2,3), (1,1), (3,3), (3,2)\} \\
&= \{\beta = (\beta_1, \beta_2) \in \{1,2,3\} \times \{1,2,3\} : \text{such that either, for some } \alpha_1 \in A_1, (\beta_1, \alpha_1, \beta_2) \\
&\quad \text{(or } (\alpha_1, \beta_1, \beta_2)) \text{ is equal to } (3,1,1), (1,2,3), \text{or } (2,3,3), \text{or, for some} \alpha_2 \in A_2, (\beta_2, \beta_1, \alpha_2) \\
&\quad \text{(or } (\alpha_2, \beta_1, \beta_2)) \text{ is equal to } (1,3,2), (3,1,1), \text{ or } (2,3,3)\}.
\end{aligned}$$

Also define the utilities $U_1(x; \beta)$ and $U_2(x; \beta)$ on $X \times B$ such that

$$U_1(x; \beta) = u_1(x; \beta_1) \quad \text{and} \quad U_2(x; \beta) = u_2(x; \beta_2),$$

and the conditional probabilities $P_1(b \mid \beta), P_2(b \mid \beta)$ to be zero whenever $b \neq \beta$.

We interpret this framework as follows: Agent 1 is of type (3,1) if his utility is $u_1(\cdot; 3)$ and he "knows" that agent 2's utility is $u_2(\cdot; 1)$. Similarly, agent 2 is of type (3,1) if his utility is $u(\cdot; 1)$ and he "knows" that agent 1's utility is $u_1(\cdot; 3)$. It is therefore natural to have

$$P_1((3,1) \mid (3,1)) = P_2((3,1) \mid (3,1)) = 1$$

but

$$P_1((3,1) \mid (\alpha, \beta)) = 0 \text{ for any } (\alpha, \beta) \neq (3,1).$$

These utilities and probabilities are chosen in order that the constructed BRP mimic the initial RP and will only be perturbed later. Observe that, for any $(b, \beta) \in B \times B$,

$$\sup_x [U_1(x; b) + U_2(x; \beta)] = \sup_x [u_1(x; b_1) + u_2(x; \beta_2)].$$

Thus, to the efficient outcome function s for the original RP corresponds uniquely an efficient outcome function S for the constructed BRP, which is $S(b, \beta) = s(b_1, \beta_2), (b, \beta) \in B \times B$.

The goal is now to choose a subset of the BIC-constraints and to get an inconsistent subsystem in the balanced transfers: $T_1(\cdot, \cdot) \equiv -T_2(\cdot, \cdot) \equiv T(\cdot, \cdot)$ defined on $B \times B$. We start by taking the

BIC-constraints corresponding to the retained SIC constraints and complete the system thus obtained in order to get a cycle, namely to get that every $T(b, \beta)$ appears only twice, once with coefficient $+1$ and once with coefficient -1. (The added constraints are singled out by a (*)). Written in matrix form, the final system is

$$
\begin{array}{c}
\\ (*) \\ (*) \\ \\ (*) \\ \\ \\ (*) \\ \\ (*) \\ (*) \\ \\
\end{array}
\begin{bmatrix}
+1 & -1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
0 & +1 & -1 & \cdot & & & & & & & & \cdot \\
\cdot & \cdot & +1 & -1 & \cdot & & & & & & & \cdot \\
\cdot & & \cdot & +1 & -1 & \cdot & & & & & & \cdot \\
\cdot & & & \cdot & +1 & -1 & \cdot & & & & & \cdot \\
\cdot & & & & \cdot & +1 & -1 & \cdot & & & & \cdot \\
\cdot & & & & & \cdot & +1 & -1 & \cdot & & & \cdot \\
\cdot & & & & & & \cdot & +1 & -1 & \cdot & & \cdot \\
\cdot & & & & & & & \cdot & +1 & -1 & \cdot & \cdot \\
\cdot & & & & & & & & \cdot & +1 & -1 & 0 \\
0 & & & & & & & & & \cdot & +1 & -1 \\
-1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & +1
\end{bmatrix}
\begin{bmatrix}
T((1,1),(1,1)) \\
T((3,1),(1,1)) \\
T((3,1),(3,1)) \\
T((3,2),(3,1)) \\
T((3,2),(3,2)) \\
T((3,3),(3,2)) \\
T((3,3),(2,3)) \\
T((2,3),(3,3)) \\
T((2,3),(2,3)) \\
T((1,3),(2,3)) \\
T((1,3),(1,3)) \\
T((1,1),(1,3))
\end{bmatrix}
\geq
\begin{bmatrix}
0 \\ 0 \\ 0 \\ -4 \\ 0 \\ 0 \\ 0 \\ 0 \\ 5 \\ 0 \\ 0 \\ 0
\end{bmatrix}
$$

Observe that the right-hand sides of the starred equations in this system are automatically zero, so that the sum of all the right-hand sides is equal to the sum of the right-hand sides of the original subsystem of SIC inequalities; that is, 1. Since the sum of the left-hand sides of the present subsystem is identically zero, it is inconsistent. Hence the whole system of BIC inequalities is inconsistent.

The probabilities (P_1, P_2) used in our constructed example are degenerate. We shall now show that they can be arbitrarily but slightly perturbed while preserving the inconsistency of the whole system. Clearly in this whole system we have 60 inequalities and 36 variables. We may rewrite it in the compact form

$$AT \geq \overline{U},$$

where the matrix A of coefficients is a 60 $\times$ 36 matrix. Each line of the left-hand side is either of the form $[T(\beta, \beta) - T(b, \beta)]$ or of the form $[T(\beta, b) - T(\beta, \beta)]$, with $b, \beta \in B$, implying that each line of the matrix A has only two nonzero coefficients, one equal to +1, the other equal to -1. Consequently if we fix an arbitrary variable, say $T(\bar{b}, \bar{\beta})$, to the value zero, and consider the resulting homogeneous system in 35 variables, $AT = 0$, it can only have the zero solution. Therefore the rank of the matrix A is equal to 35. In other words, A has a minor determinant of order 35 which is different from zero. Note that, if we perturb slightly the matrix A to the matrix $\hat{A}$, obtained by taking probabilities $(\hat{P}_1, \hat{P}_2)$ close (for the Euclidean norm) to the probabilities (P_1, P_2), this determinant will still be different from zero and the rank of $\hat{A}$ to 35.

Now by the Farkas lemma the inconsistency of the system $AT \geq \bar{U}$ is equivalent to the consistency (see Green and Laffont, 1979, Chap. 1) of the following (dual) system in $\lambda \in \mathbb{R}^{\mathcal{B}} \times \mathbb{R}^{\mathcal{B}}$, with $\mathcal{B} = \{(b, \beta) \in B \times B : b \neq \beta\}$:

$$A'\lambda = 0, \qquad \lambda\bar{U} > 0, \qquad \text{and} \qquad \lambda \geq 0.$$

Here A' is the transpose of the matrix A (and so has rank equal to 35). Moreover the dual system has a strictly positive solution λ^*. Indeed, taking $\lambda^1 \equiv 1$, we have $A'\lambda^1 = 0$, since any column of A contains five elements equal to $+1$, and five elements equal to -1, if it corresponds to a variable $T(\beta, \beta)$, and one element equal to $+1$ and one element equal to -1, if it corresponds to a variable $T(b, \beta)$. So for any solution $\bar{\lambda}$ to the dual system we may find θ large enough so that

$$\theta(\bar{\lambda}\bar{U}) + \lambda^1\bar{U} > 0 \quad \text{and} \quad \theta\bar{\lambda} + \lambda^1 > 0.$$

To get the result it remains to show that the "perturbed" dual system

$$\hat{A}'\lambda = 0, \quad \lambda\bar{U} > 0, \quad \text{and} \quad \lambda \geq 0$$

has also a solution. This results from the following lemma (the proof of which is given in appendix) on matrices, where the norm of a matrix is taken to be the square root of the sum of the squares of its elements.

Lemma 3 *Let $\mathcal{F}$ be the set of $\mu \times \nu$ matrices of rank $\nu - m$, for some $m > 0$. Let $D \in \mathcal{F}$ and $y \in \mathbb{R}^\nu$ such that $Dy = 0$. Then for any $\bar{\eta} > 0$, there exists $\bar{\varepsilon} > 0$ such that $C \in \mathcal{F}$, $\|D - C\| < \bar{\varepsilon}$ implies the existence of a vector v such that $Cv = 0$ and $\|v - y\| < \bar{\eta}$.*

In this way we get (in the space of beliefs) an open subset of beliefs $(\hat{P}_1, \hat{P}_2)$ in a neighborhood of (P_1, P_2), such that no Pareto-optimal BIC-mechanism exists. This class of counterexamples was constructed starting from a given RP admitting no Pareto-optimal SIC-mechanism. It can be shown that the same procedure can be repeated for any such RP (a proof is available from the authors). Therefore we really have a "large" class of counterexamples in the case $n = 2$.

3.2 The case $n > 2$: The genericity of condition C

The argument of the preceding section does not carry over to the case $n > 2$. In the approach followed above, Bayesian implementation is akin to Nash implementation, not dominant strategy implementation, as each agent can announce the preferences of the others. When $n = 2$ this is of no consequence. When n is at least equal to 3, the classical result by Maskin (1986b) gives hope that we might obtain more positive results.[3] This is indeed the case.

Let us fix the finite sets $A_1, \ldots, A_n$ of individual types and assume that the individual beliefs are consistent. The space of beliefs $\mathcal{P}$ is then simply the unit simplex of $\mathbb{R}^A_+$, the set of joint probability vectors p, and the individual beliefs are the conditional probabilities $p(\alpha_{-i} \mid \alpha_i)$. Also let us denote by $\mathcal{C}$ the set of joint probabilities in $\mathcal{P}$ satisfying condition C. The theorem we want to get is

Theorem 2 *For $n > 2$, $\mathcal{C}$ contains an open and dense subset of $\mathcal{P}$.*

To prove this theorem we shall actually introduce another condition that can be imposed on the beliefs and we shall show in a sequence of lemmata that this condition determines an open and dense subset of $\mathcal{P}$ and that it is stronger than condition C. The condition denoted C°, requires that $\exists\, i, j \in N, i \neq j$, such that $|A_j| \geq |A_i|$ and $\lambda_i \equiv 0$ is the unique solution to

$$\forall\, \alpha_i, \forall\, \alpha_j, \sum_{\tilde{\alpha}_i \neq \alpha_i} [p(\alpha_j \mid \alpha_i) - p(\alpha_j \mid \tilde{\alpha}_i)]\lambda_i(\alpha_i, \tilde{\alpha}_i) = 0, \tag{8}$$

where $p(\alpha_j \mid \tilde{\alpha}_i) = \sum_{(\alpha_k)_{i \neq k \neq j}} p(\alpha_{-i} \mid \tilde{\alpha}_i)$, whenever (1) of Lemma 1 has a nonzero solution in Λ.

It is clear that this condition can hold only for $n > 2$. Denoting by $\mathcal{C}^\circ$ the set of p in $\mathcal{P}$ satisfying $\mathcal{C}^\circ$ we have:

Lemma 4 *For $n > 2$, $\mathcal{C}^\circ$ is open in $\mathcal{P}$.*

---

3. We are thankful to M. Whinston and S. Williams for this insight.

Proof Take $p^\circ \in \mathcal{C}^\circ$. If on one hand the homogeneous system (1) of Lemma 1 admits only the identically zero solution (so that C$^\circ$ holds trivially), it must contain a subsystem of $m = \sum_{i=1}^{n} |\mathcal{A}_i^2|$ equations with a nonsingular $m \times m$ matrix of coefficients. By taking p close to p° we perturb this matrix slightly, so that it remains nonsingular, and the corresponding subsystem of (1) still only admits the zero solution (and C$^\circ$ still holds trivially). On the other hand, suppose system (1) has a nonzero solution in Λ and, for some pair $\{i, j\} \subset N$, with $|A_j| \geq |A_i|$, the homogeneous system (8) admits only the identically zero solution. Then we can choose a non-singular matrix of coefficients defined by a subsystem of (8) and, for any p close to p°, the corresponding perturbed matrix remains nonsingular. Therefore $p \in \mathcal{C}^\circ$. ■

Lemma 5 *For $n > 2$, $\mathcal{C}^\circ$ is dense in $\mathcal{P}$.*

Proof Take $p^\circ \in \mathcal{P} - \mathcal{C}^\circ$. We want to show that, for any $\varepsilon > 0$, there is some $p^\varepsilon \in \mathcal{P}$ such that $\|p^\varepsilon - p^\circ\| < \varepsilon$ and, for some pair $\{i, j\}$ with $|A_j| \geq |A_i|$, system (8) admits only the identically zero solution (hence $p^\varepsilon \in \mathcal{C}^\circ$). We may as well suppose that p° has full support, since, otherwise we could start from $\bar{p}^\circ$ arbitrarily close, with full support, and correspondingly take a smaller ε. Also we may suppose that for all $p \in \mathcal{P}$, such that $\|p - p^\circ\| < \varepsilon$, system (1) of Lemma 1 admits a nonzero solution in Λ (otherwise the result follows immediately). Let $m = |A_i| - 1$. There must exist some $\tilde{\alpha}_i \in A_i$ such that any $(m \times m)$-matrix of the form

$$[p^\circ(\alpha_j^h \mid \tilde{\alpha}_i) - p^\circ(\alpha_j^h \mid \alpha_i^k)]_{h,k=1,2,\ldots,m} \quad \alpha_i^k \in A_i - \{\tilde{\alpha}_i\}, \alpha_j^h \in A_j$$

which is singular. We may then perturb p° and get a nonsingular matrix as follows. Let for every $\alpha_{-i-j} \in \bigtimes_{i \neq \ell \neq j} A_\ell$ and for $\alpha_j^{m+1} \in A_j - \{\alpha_j^h : h = 1, \ldots, m\}$,

$$\begin{aligned}
\tilde{p}(\alpha_i^k, \alpha_j^h, \alpha_{-i-j}) &= p^\circ(\alpha_i^k, \alpha_j^h, \alpha_{-i-j}) - \tilde{\varepsilon}_{hk} > 0 && h, k = 1, \ldots, m, \\
\tilde{p}(\alpha_i^k, \alpha_j^{m+1}, \alpha_{-i-j}) &= p^\circ(\alpha_i^k, \alpha_j^{m+1}, \alpha_{-i-j}) + \sum_{h=1}^{m} \tilde{\varepsilon}_{hk} > 0 && k = 1, \ldots, m,
\end{aligned}$$

and

$$\tilde{p}(\alpha_i, \alpha_j, \alpha_{-i-j}) = p^\circ(\alpha_i, \alpha_j, \alpha_{-i-j}) \quad \text{otherwise},$$

with $|\tilde{\varepsilon}_{hk}| < \tilde{\delta}, h, k = 1, \ldots, m$ and $\tilde{\delta}$ small enough so that $\tilde{p} \in \mathcal{P}$. Then the associated conditional probabilities are given by

$$\begin{aligned}
\tilde{p}(\alpha_j^h \mid \alpha_i^k) &= \frac{\displaystyle\sum_{\alpha_{-i-j}} p^\circ(\alpha_i^k, \alpha_j^h, \alpha_{-i-j}) - \tilde{\varepsilon}_{hk}}{\displaystyle\sum_{\alpha_{-i}} p^\circ(\alpha_i^k, \alpha_{-i})} > 0 && h, k = 1, \ldots, m \\
\tilde{p}(\alpha_j^h \mid \tilde{\alpha}_i) &= p^\circ(\alpha_j^h \mid \alpha_i) && h = 1, \ldots, m;
\end{aligned}$$

so that for $h, k = 1, \ldots, m$,

$$\tilde{p}(\alpha_j^h \mid \tilde{\alpha}_i) - \tilde{p}(\alpha_j^h \mid \alpha_i^k) = p^\circ(\alpha_j^h \mid \tilde{\alpha}_i) - p^\circ(\alpha_j^h \mid \alpha_i^k) + \frac{\tilde{\varepsilon}_{hk}}{\sum_{\alpha_{-i}} p^\circ(\alpha_i^k, \alpha_{-i})}.$$

Clearly, for $\tilde{\delta}$ small enough, every $\tilde{\varepsilon}_{hk}$ can be chosen arbitrarily in the open interval $(-\tilde{\delta}, \tilde{\delta})$, and hence we can perturb the above matrix in any direction. Therefore, for some direction (i.e., for some $\tilde{\delta}$ and some $\tilde{\varepsilon}_{hk}$'s) the matrix

$$[\tilde{p}(\alpha_j^h \mid \tilde{\alpha}_i) - \tilde{p}(\alpha_j^h \mid \alpha_i^k)]_{h,k=1,\ldots,m} \quad \text{is non-singular.}$$

We may repeat the same perturbation argument for any $\alpha_i' \neq \tilde{\alpha}_i$ and any associated $m \times m$ singular matrix of the above form (with α_i' replacing $\tilde{\alpha}_i$). Taking δ' and the ε_{hk}''s small enough we may preserve the non-singularity of the perturbed matrix associated to $\tilde{\alpha}_i$. Choosing them appropriately we get the non-singularity of the perturbed matrix associated to α_i'. Since $|A_i|$ is finite we get a finite sequence of δ's, ε_{hk}'s, and probabilities in $\mathcal{P}$ such that taking the δ's and the ε_{hk}'s small enough the last probability in the sequence, say p^ε, is in $\mathcal{C}^\circ$ and such that $\|p^\varepsilon - p^\circ\| < \varepsilon$. ■

Lemma 6 $\mathcal{C}^\circ \subset \mathcal{C}$.

Proof We only consider the case $n > 2$. Suppose $p \in \mathcal{C}^\circ$ and, for $\lambda^\circ \in \Lambda$, $\lambda^\circ \neq 0$, that (1) of Lemma holds. We have to show that (5) holds. Since $p \in \mathcal{C}^\circ$, there must be a pair in N, say $\{1, 2\}$, such that $|A_2| \geq |A_1|$ and $\lambda_1 \equiv 0$ is the unique solution to (8).

Now, observe that, whenever (1) of Lemma 1 holds for some λ°,

$$\sum_{\tilde{\alpha}_i \neq \alpha_i} \lambda_i^\circ(\tilde{\alpha}_i, \alpha_i) = \sum_{\tilde{\alpha}_i \neq \alpha_i} \lambda_i^\circ(\alpha_i, \tilde{\alpha}_i),$$

for all $i \in N$. (See (7) above.)

So, by (1) and since $n > 2$, $\forall\, (\alpha_1, \alpha_2) \in A_1 \times A_2$,

$$\begin{aligned}
&\sum_{\alpha_{-1-2}} \left[p(\alpha_{-i} \mid \alpha_1) \sum_{\tilde{\alpha}_1 \neq \alpha_1} \lambda_1^\circ(\tilde{\alpha}_1, \alpha_1) - \sum_{\tilde{\alpha}_1 \neq \alpha_1} \lambda_1^\circ(\alpha_1, \tilde{\alpha}_1) p(\alpha_{-1} \mid \tilde{\alpha}_1) \right] \\
=& \sum_{\alpha_{-1-2-i}} \left[\sum_{\alpha_i} \sum_{\tilde{\alpha}_i \neq \alpha_i} p(\alpha_{-i} \mid \alpha_i) \lambda_i^\circ(\tilde{\alpha}_i, \alpha_i) - \sum_{\alpha_i} \sum_{\tilde{\alpha}_i \neq \alpha_i} p(\alpha_{-i} \mid \tilde{\alpha}_i) \lambda_i^\circ(\alpha_i, \tilde{\alpha}_i) \right] \\
=&0, \quad \text{for } 1 \neq i \neq 2 \text{ and } \alpha_{-1-2-i} \in \left\{ \times_j A_j : j \in N \setminus \{1, 2, i\} \right\},
\end{aligned}$$

that is, $\forall\, (\alpha_1, \alpha_2) \in A_1 \times A_2$,

$$p(\alpha_2 \mid \alpha_1) \sum_{\tilde{\alpha}_1 \neq \alpha_1} \lambda_1^\circ(\alpha_1, \tilde{\alpha}_1) - \sum_{\tilde{\alpha}_1 \neq \alpha_1} \lambda_1^\circ(\alpha_1, \tilde{\alpha}_1) p(\alpha_{-1} \mid \alpha_1) = 0.$$

But this is nothing else than (8) for $\{i, j\} = \{1, 2\}$, so that if λ° satisfies (1) of Lemma 1 we must have $\lambda^\circ \equiv 0$. Therefore, $\forall\, \alpha \in A, \forall\, i \neq 1$,

$$\begin{aligned}
&p(\alpha_{-i} \mid \alpha_i) \sum_{\tilde{\alpha}_i \neq \alpha_i} \lambda_i^\circ(\tilde{\alpha}_i, \alpha_i) - \sum_{\tilde{\alpha}_i \neq \alpha_i} \lambda_i^\circ(\alpha_i, \tilde{\alpha}_i) p(\alpha_{-i} \mid \tilde{\alpha}_i) \\
=&p(\alpha_{-1} \mid \alpha_1) \sum_{\tilde{\alpha}_1 \neq \alpha_1} \lambda_1^\circ(\tilde{\alpha}_1, \alpha_1) - \sum_{\tilde{\alpha}_1 \neq \alpha_1} \lambda_1^\circ(\alpha_1, \tilde{\alpha}_1) p(\alpha_{-1} \mid \alpha_1) = 0.
\end{aligned}$$

So (5) holds. ■

This concludes the proof of Theorem 2. Observe that, in the proof of the lemmata, we have not used the consistency of p in any crucial way. Therefore a similar theorem holds for the nonconsistent case. The advantage of considering the consistent case (apart from notational simplicity) is that the case $n = 2$ is clearer: condition C then coincides with independence.

4. Conclusion

We have started the paper by emphasizing the difficulty of finding Pareto-optimal SIC-mechanisms. It should be noted that it is this same difficulty which motivated in the literature a second-best approach applying other efficiency criteria, such as "interim efficiency" (see, in particular, Holmström and Myerson, 1983; Wilson, 1978). The central result of the present paper actually indicates that, "essentially", incomplete information may not generate inefficiencies in the first-best sense, at least when individual utilities are transferable and for a finite number of types and more than two agents.

There remain areas around this result requiring some complementary work. The proof of the theorem strongly relies upon the finiteness of the sets of individual types; the nonfinite case deserves a special treatment. Also, as already observed in d'Aspremont and Gérard-Varet (1982), the sufficient condition on individual beliefs we have considered here, although quite "general", is not necessary for having a Pareto-optimal BIC-mechanism (under transferability). More has to be said about those individual beliefs on the "boundary" of the set given by that sufficient condition.

It should be stressed, to conclude, that there are other routes to Pareto-optimal BIC-mechanisms. In particular, one route is to consider restrictions on both utilities and beliefs. One can, as in Maskin (1986b); Bensaid (1986), consider, for a one-dimensional continuous type-variable, a monotonicity condition on the marginal utilities together with a negative correlation on the beliefs. More general conditions can certainly be considered along these lines of ideas. Going to nontransferable utility worlds is also a major issue (see d'Aspremont and Gérard-Varet, 1989).

Appendix: Proof of Lemma 4

Lemma 4 *Let $\mathcal{F}$ be the set of $\mu \times \nu$ matrices of rank $v - m$ for some $m > 0$. Let $D \in \mathcal{F}$ and $y \in \mathbb{R}^{\nu}$ such that Dy is equal to 0. Then, for any $\bar{\eta} > 0$, there exists $\bar{\varepsilon} > 0$ such that $C \in \mathcal{F}$, $\|D - C| < \bar{\varepsilon}$ implies the existence of a vector v such that $Cv = 0$ and $v - y < \bar{\eta}$.*

Proof Assume that the result were not true. There would exist a $\delta > 0$ and a sequence of matrices $D^1, D^2, \dots, D^n, \dots$ in $\mathcal{F}$ converging to D such that, for all n, $D^n y^n = 0$ would imply $\|y^n - y\| > \delta$. We will show that this leads to a contradiction.

For each n, let us choose an orthonormal basis $\bar{z}^n = \{z^{n,1}, z^{n,2}, \dots, z^{n,m}\}$ of the kernel of D^n. There exists a subsequence along which $\bar{z}^n$ converges to $\bar{z} = \{z^1, z^2, \dots, z^m\}$. For all n we have $D^n z^{n,j} = 0$ for all $j \in \{1, 2, \dots, m\}$, $z^{n,i} z^{n,j} = 0$ for all $i \neq j$, and $\|z^{n,j}\| = 1$ for all j. By continuity, we get $Dz^j = 0$ for all j, $z^i z^j = 0$ for $i \neq j$, and $\|z^j\| = 1$ for all j. Hence $\bar{z}$ is also an orthonormal basis of the kernel of D.

There exist $\mu^1, \mu^2, \dots, \mu^m$ such that y is equal to $\sum_{i=1}^m \mu^i z^i$. Let y^n be equal to $\sum_{i=1}^m \mu^i z^{n,i}$. The vector y^n belongs to the kernel of D^n, and a subsequence of $\{\dots, y^n, \dots\}$ converges to y, which establishes the contradiction. ■

References

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

B Bensaid. Bayesian incentive compatible mechanisms under a negative dependence assumption. Discussion paper, Université Paris-I, 1986.

Jacques Créemer and Richard P McLean. Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica*, 53:345–361, 1985.

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.

Claude d'Aspremont and Louis-André Gérard-Varet. Bayesian incentive compatible beliefs. *Journal of Mathematical Economics*, 10(1):25–45, 1982.

Claude d'Aspremont and Louis-André Gérard-Varet. Utilitarian fundamentalism and limited information. CORE Discussion Paper 8935, Université catholique de Louvain, 1989.

David Gale. *The theory of linear economic models*. Mcgraw-Hill, New York, 1960.

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

Theodore Groves and John Ledyard. Incentive compatibility ten years later. In R. Radner T. Groves and S. Reiter, editors, *Information, Incentives and Economic Mechanisms*. University of Minnesota Press, Minneapolis, 1987.

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

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.

Bengt Holmström and Roger Myerson. Efficient and durable decision rules with incomplete information. *Econometrica*, 51(6):1799–1819, 1983.

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.

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

Eric S Maskin. The theory of implementation in Nash equilibrium: A survey. In D. Schmeidler L. Hurwicz and H. Sonnenschein, editors, *Social Goals and Social Organization, Essays in Memory of Elisha Pazner*. Cambridge University Press, Cambridge (UK), 1986a.

Eric S Maskin. Optimal Bayesian mechanism. In R.M. Starr W.P. Heller and D.A. Starrett, editors, *Uncertainty, Information and Communication, Essays in Honor of K.J. Arrow, Vol. 3*. Cambridge University Press, Cambridge (UK), 1986b.

Paul R Milgrom and Robert J Weber. A theory of auctions and competitive bidding. *Econometrica*, pages 1089–1122, 1982.

Roger B Myerson. Bayesian equilibrium and incentive-compatibility: An introduction. In D. Schmeidler L. Hurwicz and H. Sonnenschein, editors, *Social Goals and Social Organization, Essays in Memory of Elisha Pazner*. Cambridge University Press, Cambridge (UK), 1985.

Mark Walker. On the impossibility of a dominant-strategy mechanism to optimally decide public question. *Econometrica*, pages 1521–1540, 1980.

Robert Wilson. Information, efficiency, and the core of an economy. *Econometrica*, pages 807–816, 1978.