Collective choice mechanisms and individual incentives *

Claude d'Aspremont
Louis-André Gérard-Varet†

Contemporary economic analysis has recognized the importance of "asymmetric information" not only for markets, but also for many different organizations. The resulting "theory of incentives" aims to define rules, or institutional mechanisms, likely to lead individual agents to make collectively optimal choices and to reveal all private information necessary for an efficient collective choice. A new chapter in game theory has thus been opened: the strategic analysis of economic institutions, their stability, or the conditions under which they may be changed.

The driving forces at the origin of the theory of incentives go well back in the history of economic thought. They can be traced in the first observations regarding possible market failures (externalities, public goods). They are also found in the well-known "controversy over socialism" of the 1930s, which highlighted not only the question of whether collective ownership was compatible with decentralization, but also, with von Hayek (1945), the informational dimension of the workings of markets. Whenever information is dispersed, "equilibrium" requires a process of communication. It may then be in the agents' interest to manipulate such a process and not to reveal fully the information they have at hand.

Wicksell (1896) followed by Samuelson (1954) both held the similar view that, in order to achieve an optimal level of collective goods, it was necessary to centralize some information but that: "it is in the selfish interest of each person to give false signals, to pretend to have less interest in a given collective-consumption activity than he really has" (Samuelson, 1954, 388–9). However, Samuelson believed that in a competitive economy, with only privately owned goods, the problem of encouraging agents to reveal their preferences correctly could not arise. It was Hurwicz (1960, 1972) who showed that this problem, as well as the strategic phenomena associated with it, are not due to the introduction of public goods, at least when the number of agents is finite. The theoretical analysis of planning such as by Drèze and de la Vallée Poussin (1971) and Malinvaud (1971), raises similar questions irrespective of whether goods are owned privately or publicly. And it was indeed Vickrey (1961) who provided an essential impetus in our understanding of incentives in his study of the strategic aspects of auctions or tenders, viewed as exchange mechanisms in which the value of the good exchanged is "uncertain".

Informational problems play a significant role in any market involving bilateral contracts. This is true for insurance markets where nowadays a well-established terminology has in fact originated. An insurance company faces the problem of having to cover particular risks which are not fully known. The insurer then offers a set of contracts so as to allow each insured party to self-select from that set the contract that corresponds best to his own type of risks. This is the problem of *adverse selection.* Moreover, the insured party may also control a "self-protection" variable which is not observable by

∗. Chapter 8 in C. Schmidt (ed.), *Game Theory and Economic Analysis. A Quiet Revolution in Economics*, London, Routledge, 151–171, 2002 – Reprint 1592.

†. Louis-André Gérard-Varet dies 2001.

the insurer. The insured party's choice of a value for this variable affects his risk of accident, and may indeed be incompatible with the risk considered by the insurer. Such insufficient self-protection is called *moral hazard*. Other examples illustrating the problems of moral hazard are the management of collective property rights, or sharing a collective outcome (Alchian and Demsetz, 1972). It is a priori in everyone's interest to cooperate. But if each participant can gain an individual advantage by acting as a "free-rider," then agreement may be impossible or unachievable.

This chapter aims to show how the two problems of moral hazard and adverse selection may be integrated and treated within the general theory of non-cooperative games. To ensure that the determination of players' behavior is understood by all and is not to impute irrationality to any player, the outcome of such games must be a non-cooperative (Nash) equilibrium. This is a state of individual strategies in which no participant is able to take his own alternative route and expect to achieve a better outcome. Moral hazard reflects situations of "bad equilibria," that is, equilibria which, due to strategic externalities, are socially unsatisfactory. But game theory can also propose a solution: some *cooperative transformation* may be introduced creating a new game with equilibria having better welfare properties. Such a transformation can come about through a "regulation," a "mediation" or an "audit." It may be obtained by "repeating" the game, by adding a "communication scheme," or by "contractually" modifying the original payoff structure. The latter is the aspect we shall concentrate on here. However, cooperative transformation can also give rise to multiple equilibria. It may thus be necessary to introduce an additional contractual structure not only to define new rules, but also to select a collectively efficient equilibrium.

One needs to be careful about the definition of an equilibrium in a model which integrates both adverse selection and moral hazard, and therefore relies on the analysis of "games with incomplete information." A model of a game with incomplete information must specify not only what each player "can do," that is the actions, but also what he "knows," that is his private information, and his "beliefs" as regards other players' information. In this framework, an individual strategy consists of a plan of action that defines what action is selected in every contingency, given the player's private information (or what he acquires during the course of the game). Thus, a necessary requirement is the notion of a "Bayesian equilibrium" (Harsanyi, 1967-68) generalizing the Nash equilibrium to situations of incomplete information: The players' "conjectures" about their mutual behavior are confirmed by the decisions taken by each on the basis of their private information.[1]

First, we will introduce a general model allowing the analysis of incentives constraints in situations where adverse selection and moral hazard are both present. We will however make a strong assumption: players assign to strategic issues utilities which are quasi-linear with respect to a numeraire (or money). This hypothesis facilitates the analysis in two ways. On the one hand, the cooperative transformation of the game can be limited to the introduction of a single instrument, namely monetary transfers. On the other hand, it considerably simplifies the definition of the welfare criterion in a world with incomplete information, by allowing us to retain the criterion of Pareto optimality in its most standard form.[2]

Next, we will obtain further simplification by applying the "revelation principle." The Bayesian equilibrium in a game with incomplete information, which is defined by the players' individual plan of action contingent on their private information, can indeed be viewed as the equilibrium of a "game of revelation" where each player's choice is related to the information he reveals. Incentive

1. The notion of "conjecture" may have different meanings. We take here the more restrictive where the information structure of the game is common knowledge, and where each player can predict other players' strategies at equilibrium.
2. We will come back to this issue. See Forges (1994) for a more detailed discussion.

constraints characterize those issues which guarantee that it is optimal for each player to communicate his information sincerely: this is the meaning of the "revelation principle."[3] A game of revelation may involve an extremely complicated mechanism. From the viewpoint of incentive constraints, we will ignore these complexities in order to concentrate on the issues which are compatible with the sincere revelation of private information by participants. This framework allows us to relate our general model to examples in the literature. It will also allow us finally to define in the third section, a large class of situations where individual incentives and collective efficiency are compatible. In conclusion, we will discuss other ways of integrating moral hazard and adverse selection, as well as other criteria of collective efficiency.

1. A general model

We consider strategic situations where players have imperfect information with respect to the choices of possible courses of action and also have incomplete information regarding their individual characteristics by other players. The problem of moral hazard derives from imperfect information. Adverse selection is due to incomplete information. In such situations, players strategies have two components: an "action" which is non-observable and a "signal" which is observable. A vector of players' strategies generates a probability distribution over outcomes. The assumption that individual utilities are quasi-linear with respect to money allows us to analyze the conditions for a Bayesian equilibrium to be collectively efficient in terms of "transfers" between players.

1.1 The environment: imperfect and incomplete information

There is a finite set N of players facing a finite set X of collective *outcomes*, which are publicly observable. The *utility* (or payoff) of player $i \in N$ derived from outcome $x \in X$ is given by the real-valued function $u_i(x, d_i, \alpha_i)$, where d_i is a choice of action from a finite set D_i, and $\alpha_i \in A_i$ is a state of *private information* also in a finite set. Players make their choice of actions simultaneously, so that no one can observe the complete chosen profile of actions $d = (d_1, \ldots, d_i, \ldots, d_n) \in D \stackrel{\text{def}}{=} X_{i \in N} D_i$. The information at the disposal of player i only concerns his own characteristics. A_i can also be viewed as the set of all possible *types* of players. The sets A_i and D_i, as well as the set X_i are common knowledge. On the other hand, only player i knows the type in A_i which describes his true characteristics. Since the types are not observable, no player i can tell which state

$$\alpha_{-i} = (\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \alpha_n) \in A_{-i} \stackrel{\text{def}}{=} X_{j \neq I} A_j$$

describes the characteristics of the other players.[4]

We restrict ourselves to the case where an individual's utility depends only on his choice of action and on his type, but is independent of the full profile of actions and of the complete state of types. In other words, we consider here only the case of *private values* with *individual concern*.[5] Furthermore, individual utilities are measured in units which can be freely transferred between players: this is the assumption of quasi-linearity.

3. See Myerson (1985) for a discussion and more references.
4. The literature refers often to models with "hidden information" or "hidden action." A chosen action d_i is hidden if it is not publicly observable. Similarly, a type $\alpha_i \in A_i$ is hidden if it cannot be verified on the basis of common knowledge.
5. In the case of *common values* with *collective concern*, utilities would take the form $u_i(x, d, \alpha)$. The question of collective concern (as regards actions) is less problematic than that of common values (as regards types).

Incomplete information implies that each $i \in N$ is uncertain about the characteristics of others. In a bayesian framework however, he acts on the basis of the likelihood he assigns to the others. Thus, to each player i is associated a function $p_i; A_i \rightarrow D(A_{-i})$ where $D(A_{-i})$ is the set of all probability distributions defined over A_{-i}. It gives the probability $p_i(\alpha_{-i} \mid \alpha_i)$ attached by $i \in N$ of type $\alpha_i \in A_i$ to other players being of type $\alpha_{-i} \in A_{-i}$. The p_i functions are common knowledge; but the α_i types are private information, and the same is true for the "beliefs" $p_i(\cdot \mid \alpha_i)$. We assume for simplicity[6] that all individual beliefs are derived from the same joint probability distribution p defined over A, such that:

$$p_i(\alpha_{-i} \mid \alpha_i) = p(\alpha_{-i} \mid \alpha_i) = \frac{p(\alpha_i, \alpha_{-i})}{\sum_{\alpha - i} p(\alpha_i, \alpha_{-i})}, i \in N.$$

(We also assume that p has the full support of A.)

The distribution p gives the *structure of beliefs*. The environment is therefore characterized by two elements: the profile of utilities $(u_i)_{i \in N}$ and the structure of beliefs p.

The strategic description is however not yet exhausted. We further assume that each player i can select a *signal* s_i from a finite set S_i. Unlike the action d_i, s_i is made public. Individual behavior thus has a dual dimension: while the configuration of actions $d \in D$ is not observable, that of the signals $s \in S \overset{\text{def}}{=} X_{i \in N} S_i$, is public information. The assumptions underlying the game are then incorporated into a function g defined over $D \times S$ and with value in the set $D(X)$ of all probability distributions over X. This function which is called the *outcome function* gives the probability $g(x \mid d, s)$ of getting result x conditional on the players having chosen the configuration (d, s) of actions and signals.

Player i of type α_i evaluates a configuration (d, s) on the basis of his payoffs which are given by:

$$U_i^g(d, s, \alpha_i) = \sum_x U_i(x, d_i, \alpha_i) g(x \mid d, s) \tag{1}$$

However, under incomplete information and in the absence of any other communication than what is carried (implicitly) by the signals, a player only knows the characteristics which are attached to his type. Player i, who knows his type α_i but doesn't know other types α_j, $j \neq i$, imputes to each j a "decision rule," that is a plan or a function which we define as $(\underset{\sim}{d}_j, \underset{\sim}{s}_j) : A_j \rightarrow D_j \times S_j$.[7] This function determines a strategy $(\underset{\sim}{d}_j(\alpha_j), \underset{\sim}{s}_j(\alpha_j))$ giving the action and the signal implemented by player j if his type were α_j. Such a decision rule is called a *normalized strategy*. The set of all normalized strategies that can be imputed to j is $(\underset{\sim}{D}_j \times \underset{\sim}{S}_j)$ with $\underset{\sim}{D}_j \overset{\text{def}}{=} D_j^{A_i}$ and $\underset{\sim}{S}_j \overset{\text{def}}{=} S_j^{A_j}$. Player i of type α_i who has to choose a strategy $(d_i, s_i) \in D_i \times S_i$ will conjecture normalized strategies $(\underset{\sim}{d}_{-i}, \underset{\sim}{s}_{-i}) \in \underset{\sim}{D}_{-i} \times \underset{\sim}{S}_{-i}$, $\underset{\sim}{D}_{-i} \overset{\text{def}}{=} X_{j \neq i} \underset{\sim}{D}_j$, $\underset{\sim}{S}_{-i} \overset{\text{def}}{=} X_{j \neq i} \underset{\sim}{S}_j$, and evaluate the situation on the basis

6. The properties derived from the model do not depend on this assumption.

7. We assume, without loss of generality, deterministic decision rules. One could consider functions of A_j mapped into the set D (D_J ò S_j) of probability distributions over D_J ò S_j.

of conditional expected payoffs, also called *interim payoffs*:

$$\begin{aligned} & U_i^g((d_i,s_i),(\underset{\sim}{d}_{-i},\underset{\sim}{s}_{-i}),\alpha_i) \\ =&\sum_{\alpha-i} U_i^g(d_i,\underset{\sim}{d}_{-i}(\alpha_{-i}),s_i,\underset{\sim}{s}_{-i}(\alpha_{-i}),\alpha_i)p(\alpha_{-i}\mid\alpha_i) \\ =&\sum_{\alpha-i}\left(\sum_x u_i(x,d_i,\alpha_i)g(x\mid d_i,\underset{\sim}{d}_{-i}(\alpha_{-i}),s_i,\underset{\sim}{s}_{-i}(\alpha_{-i}))\right)p(\alpha_{-i}\mid\alpha_i). \end{aligned} \tag{2}$$

We therefore have a "Bayesian game," as introduced by Harsanyi (1967-68). A choice of normalized strategies $(\underset{\sim}{d}^*,\underset{\sim}{s}^*)\in\underset{\sim}{D}\times\underset{\sim}{S}$, where $\underset{\sim}{D}=X_{i\in N}\underset{\sim}{D}$, $\underset{\sim}{S}=X_{i\in N}\underset{\sim}{S}$, is called a Bayesian equilibrium if: $\forall\, i\in N,\forall\,\alpha_i\in A_i,\forall\, d_i\in S_i,\forall\, s_i\in S_i$,

$$U_i^g((d_i,s_i),(\underset{\sim}{d}^*_{-i},\underset{\sim}{s}^*_{-i}),\alpha_i)\le U_i^g((\underset{\sim}{d}^*_i(\alpha_i),\underset{\sim}{s}^*_i(\alpha_i)),(\underset{\sim}{d}^*_{-i},\underset{\sim}{s}^*_{-i}),\alpha_i).$$

The strategy $(\underset{\sim}{d}^*_i(\alpha_i),\underset{\sim}{s}^*_j(\alpha_i))$ guarantees that, in equilibrium, player i of type α_i will get the highest conditional payoff $\overline{U}_i^g(\cdot,(\underset{\sim}{d}^*_i,\underset{\sim}{s}^*_{-i}),\alpha_i)$, given his conjectures $(\underset{\sim}{d}^*_{-i},\underset{\sim}{s}^*_{-i})$ on the strategies of others. In equilibrium, these conjectures are automatically satisfied. This concept generalizes to the case of incomplete information, the notion of a Nash equilibrium while incorporating a principle of "rational expectations."[8] Since a configuration $(\underset{\sim}{d}^*,\underset{\sim}{s}^*)$ is a Bayesian equilibrium on the basis of the *interim* payoffs given by (2), it is also a Nash equilibrium on the basis of *ex ante* payoffs. Indeed, if we denote by p_{A_i} the marginal distribution of p with respect to A_i, we have:

$$\begin{aligned} &\forall\, i\in N,\forall\,\underset{\sim}{d}_i\in\underset{\sim}{D}_i,\forall\,\underset{\sim}{s}_i\in\underset{\sim}{S}_i, \\ &\sum_{\alpha_i}\overline{U}_i^g((\underset{\sim}{d}_i(\alpha_i),\underset{\sim}{s}_i(\alpha_i)),(\underset{\sim}{d}^*_{-i},\underset{\sim}{s}^*_{-i}),\alpha_i)p_{A_i}(\alpha_i) \\ \le&\sum\overline{U}_i^g((\underset{\sim}{d}^*_i(\alpha_i),\underset{\sim}{s}^*_i(\alpha_i)),(\underset{\sim}{d}^*_{-i},\underset{\sim}{s}^*_{-i}),\alpha_i)p_{A_i}(\alpha_i). \end{aligned}$$

1.2 Collective efficiency: introducing transfers

For a given state configuration of types α, a profile $(\underset{\sim}{d},\underset{\sim}{s})$ of normalized strategies in $\underset{\sim}{D}\times\underset{\sim}{S}$ leads to a strategic outcome $(\underset{\sim}{d}(\alpha),\underset{\sim}{s}(\alpha))=(\underset{\sim}{d}_i(\alpha)_i,\underset{\sim}{s}_i(\alpha_i))_{i\in N}$ which, through the outcome function, generates a probability distribution over the outcomes denoted $g(\cdot\mid\underset{\sim}{d}(\alpha),\underset{\sim}{s}(\alpha))$. A classical notion of collective efficiency consists of selecting a profile $(\underset{\sim}{d}^*,\underset{\sim}{s}^*)$ of normalized strategies such that, irrespective of the state α, the outcome $(\underset{\sim}{d}^*(\alpha),\underset{\sim}{s}^*(\alpha))$ is Pareto optimal for that state. Under incomplete information, other notions are possible.[9] With the hypothesis that utilities are quasi-linear, it would simply require the maximization of the collective surplus, given the configuration of the utilities $(u_i)_{i\in N}$ and the outcome function g.

8. See the discussion on this framework in d'Aspremont et al. (1995b).
9. See Holmström et al. (1983), Forges (1994).

Thus, a profile (of normalized strategies) $(\underset{\sim}{d}^*, \underset{\sim}{s}^*) \in (\underset{\sim}{D} \times \underset{\sim}{S})$ is said to be *collectively efficient* (with respect to $(u_i)_{i\in N}$ and g) if:

$$\forall\, \alpha \in A, (\underset{\sim}{d}^*(\alpha)\underset{\sim}{s}^*(\alpha)) \in \arg\max_{(d,s)} \sum_i U_i^g(d, s, \alpha_i). \tag{3}$$

The collective surplus, evaluated *ex post* relative to the private information (although *ex ante* with respect to the outcome) is given by:

$$\begin{aligned} &\sum_{\text{base}}^{\text{top}} U_i^g(\underset{\sim}{d}^*(\alpha), \underset{\sim}{s}^*(\alpha), \alpha_i) \\ =&\sum_x \left(\sum_i u_i(x, \underset{\sim}{d}^*_i(\alpha_i), \alpha_i)) g(x \mid \underset{\sim}{d}^*(\alpha), \underset{\sim}{s}^*(\alpha) \right). \end{aligned} \tag{4}$$

For the environments which are considered here, a Bayesian equilibrium does not in general yield a collectively efficient configuration of actions and signals. This conflict of rationality leads us to question the possibility of defining "new rules of the game." Since we limit ourselves to the static case, and without introducing new possibilities of communication between players, we will only consider the instruments given by transfers of utilities in numéraire. Such transfers can be managed by a mediator. They have also to be based on observable variables only: physical results or signals, but not actions or types.

A *transfer scheme* is a function $t : X \times S \to \mathbb{R}^n$ which determines the amount of numéraire $t_i(x, s)$ which is paid or received (according to its sign) by player $i \in N$, if the outcome $x \in X$ is observed and when the signals $s \in S$ have been made public. The introduction of a transfer scheme, accepted by all players, allows us to introduce a new game with incomplete information, which is a cooperative transformation of the original game, with the payoff functions given by:

$$\sum_x (u_i(x, d_i, \alpha_i) + t_i(x, s)) g(x \mid d, s) = U_i^g(d, s, \alpha_i) + T_i(d, s) \tag{5}$$

where we use (1) and $T_i(d, s) \overset{\text{def}}{=} \sum_x t_i(x, s) g(x \mid d, s)$. It is assumed here that the structure of beliefs is unchanged.

A transfer scheme must fulfill admissibility constraints. To achieve collective efficiency, the strongest constraint is that the whole surplus be divided amongst the players. Thus, a transfer scheme t is said to satisfy *budget balance* if:

$$\forall\, x \in X, \forall\, s \in S, \sum_i t_i(x, s) = 0. \tag{6}$$

In the rest of this chapter, we assume that this condition holds.[10]

Given a configuration of utilities $(u_i)_{i\in N}$, a structure of beliefs p and an outcome function g, a (balanced) transfer scheme t is said to implement or *support* a profile of (collectively efficient) normalized strategies $(\underset{\sim}{d}^*, \underset{\sim}{s}^*)$, if this profile is a Bayesian equilibrium with respect to the expected

10. An alternative condition, called *expected budget balance* postulates that: for every d and every s, $\sum_i T_i(d, s) = 0$ This is a weaker condition, because it only balances the budget on average with respect to the distribution given by g. It does allow the reinforcement of later results. See d'Aspremont and Gérard-Varet (1998).

interim payoffs deduced from (5), in such a way that the following *self-selection constraints* are satisfied:

$$\begin{aligned}
&\forall\, i \in N, \forall\, \alpha_i \in A_i, \forall\, d_i \in S_i, \forall\, s_i \in S_i,\\
&\sum_{\alpha_{-i}} \left(\sum_x [u_i(x, d_i, \alpha_i) + t_i(x, s_i, \underset{\sim}{s}^*_{-i}(\alpha_{-i}))] g(x \mid d_i, \underset{\sim}{d}^*_{-i}(\alpha_{-i}), s_i, \underset{\sim}{s}^*_{-i}(\alpha_{-i}))) \right) p(\alpha_{-i} \mid \alpha_i)\\
\leq &\sum_{\alpha_{-i}} \left(\sum_x [u_i(x, \underset{\sim}{d}^*_i(\alpha_i), \alpha_i) + t_i(x, \underset{\sim}{s}^*(\alpha))] g(x \mid \underset{\sim}{d}^*(\alpha), \underset{\sim}{s}^*(\alpha)) \right) p(\alpha_{-i} \mid \alpha_i).
\end{aligned} \tag{7}$$

In what follows, we try to identify belief structures and outcome functions such that a collectively efficient structure is achieved (as a Bayesian equilibrium) by balanced transfers, whatever be the configuration of utilities. In the next section however, we see how the discussion is made simple by reducing self-selection constraints to "incentive constraints."

We have not mentioned up to now the possibility open to each agent to refuse to participate in the mechanism and be satisfied with an alternative utility level. Assuming that this level is equal to zero irrespective of the agent's type, and that the a priori expected total surplus is non-negative, it is possible to add *ex ante participation (or individual rationality) constraints*:

$$\forall\, i \in N, \sum_\alpha \left(\sum_x [u_i(x, \underset{\sim}{d}^*_i(\alpha_i), \alpha_i) + t_i(x, \underset{\sim}{s}^*(\alpha))] g(x \mid \underset{\sim}{d}^*(\alpha), \underset{\sim}{s}^*(\alpha)) \right) p(\alpha) \geq 0. \tag{8}$$

Indeed, if transfers t satisfy (6) and (7) but not (8), we can always define new transfers t' such that: $\forall\, x \in X, \forall\, \alpha \in A$,

$$t'_i(x, \underset{\sim}{s}^*(\alpha)) = t_i(x, \underset{\sim}{s}^*(\alpha)) - \sum_a [U^g_i(\underset{\sim}{d}^*(\alpha), \underset{\sim}{s}^*(\alpha), \alpha_i) + T_i(\underset{\sim}{d}^*(\alpha), \underset{\sim}{s}^*(\alpha))] p(\alpha)$$

for all $i \neq i$, and

$$\begin{aligned}
&t'_1(x, \underset{\sim}{s}^*(\alpha))\\
=&t_1(x, \underset{\sim}{s}^*(\alpha)) + \sum_{i \neq 1} \sum_\alpha [U^g_i(\underset{\sim}{d}^*(\alpha), \underset{\sim}{s}^*(\alpha), \alpha_i) + T_i(\underset{\sim}{d}^*(\alpha), \underset{\sim}{s}^*(\alpha))] p(\alpha).
\end{aligned}$$

The transfers t' clearly satisfy (6), (7) and (8). However, the constraints (8) apply to payoffs which are *ex ante* relative to outcomes and types. Other constraints of participation (for example *interim constraints*, that are conditional on each individual's private information) are harder to satisfy. We will return to this issue.

2. Incentive constraints

We have so far adopted a very general formulation of the incentives problem. Given a configuration of utilities, a structure of beliefs and an outcome function, a mediator defines balanced transfers in such a way that a configuration of normalized strategies, leading to a Pareto optimal outcome in each state, be self-selected (as defined by constraints (7)). Such a problem may prove difficult to solve and the mediator is forced to rely a priori on everything that is common knowledge. It is however possible to simplify the problem, using the "revelation principle." This principle implies

that players exchange information about their types, thus changing the self-selection constraints into incentive constraints. Once we define the canonical form of the incentive constraints, the revelation principle also allows us to discuss two polar cases: models with pure moral hazard and models with pure adverse selection.

2.1 The revelation principle

Let us assume that, to determine the transfers, the mediator can collect information about the players' private characteristics. These data, which are obtained from all players simultaneously, are substitutes for the public signals. We assume that A_i itself, is the set of all messages that player i can communicate regarding his type. In this new setup, the outcome is given by an outcome function $\gamma : D \times A \to D(X)$, where $\gamma(x \mid d, a)$ is the probability of observing the outcome $x \in X$, given that $d \in D$ is the course of actions that is selected and $a = (a_1, \ldots, a_i, \ldots, a_n) \in A$ is the profile of messages that are communicated. We obtain a "*revelation game*."

In this kind of game, each player $i \in N$ selects at the same time a *decision rule* $\underset{\sim}{d}_i$ which determines his action conditional on his type, and a *strategy of announcement* which is a function $\underset{\sim}{a}_i : A_i \to A_i$, where $\underset{\sim}{a}_i(\alpha_i) \in A_i$ is the type he declares given that $\alpha_i \in A_i$ is his true type. We denote by $\underset{\sim}{A}_i = A_i^{A_i}$ the set of all possible strategies of announcement of the player. Given a profile $(\underset{\sim}{d}_{-i}, \underset{\sim}{a}_{-i}) \in \underset{\sim}{D}_{-i} \times \underset{\sim}{A}_{-i}$, player $i \in N$ of type $\alpha_i \in A_i$ selects an action $d_i \in D_i$ and a message $\alpha_i \in A_i$, on the basis of his expected interim payoffs:

$$\begin{aligned} & U_i^g((d_i, a_i), (\underset{\sim}{d}_{-i}, \underset{\sim}{a}_{-i}), \alpha_i) \\ =& \sum_{\alpha_{-i}} \left(\sum_x u_i(x, d_i, \alpha_i) \gamma(x \mid d_i, \underset{\sim}{d}_{-i}(\alpha_{-i}), a_i, \underset{\sim}{a}_{-i}(\alpha_{-i})) \right) p(\alpha_{-i} \mid \alpha_i). \end{aligned} \tag{2'}$$

One special announcement strategy, denoted by $\hat{a}_i \in \underset{\sim}{A}_i$, consists of player i "telling the truth," that is we have $\hat{a}_i(\alpha_i) = \alpha_i$ for all possible values of the type $\alpha_i \in A_i$ ($\hat{a}_i$ is the identity function of A_i into A_i). We now seek conditions under which $(\underset{\sim}{d}, \hat{a}) \in \underset{\sim}{D} \times \underset{\sim}{A}$ is a Bayesian equilibrium: each player is induced to reveal his private characteristics.

In a revelation game, collective efficiency leads to a profile of decision rules $\underset{\sim}{d}^* \in \underset{\sim}{D}$ satisfying:

$$\forall\, \alpha \in A, \underset{\sim}{d}^*(\alpha) \in \arg\max_d \sum_x \left(\sum_i u_i(x, d_i, \alpha_i) \right) \gamma(x \mid d, \alpha). \tag{3'}$$

A transfer scheme is now a function $\tau : X \times A \to \mathbb{R}^n$ which is balanced if:

$$\forall\, x \in X, \forall\, \alpha \in A, \sum_i \tau_i(x, \alpha) = 0. \tag{6'}$$

Given a configuration of utilities $(u_i)_{i \in N}$, a structure of beliefs p and an outcome function γ, then a balanced transfer rule τ and a collectively efficient decision rule $\underset{\sim}{d}^*$ form an *incentive compatible*

mechanism if the following incentives constraints hold:

$$\begin{aligned}
&\forall\, i \in N, \forall\, \alpha_i \in A_i, \forall\, a_i \in A_i, \forall\, d_i \in D_i, \\
&\sum_{\alpha_{-i}} \left(\sum_x [u_i(x, d_i, \alpha_i) + \tau_i(x, a_i, \alpha_{-i})] \gamma(x \mid d_i, \underset{\sim}{d}^*_{-i}(\alpha_{-i}), a_i, \alpha_{-i}) \right) p(\alpha_{-i} \mid \alpha_i) \\
\leq &\sum_{\alpha_{-i}} \left(\sum_x [u_i(x, \underset{\sim}{d}^*_i(\alpha_i), \alpha_i) + \tau_i(x, \alpha)] \gamma(x \mid \underset{\sim}{d}^*(\alpha), \alpha) \right) p(\alpha_{-i} \mid \alpha_i).
\end{aligned} \tag{7'}$$

In other words, $(\underset{\sim}{d}^*, \hat{a})$ is a Bayesian equilibrium of the associated game.

We can now define the *revelation principle*. Given a utility profile $(u_i)_{i \in N}$ and a structure of beliefs p, if for a given outcome function g there exists a balanced transfer scheme t which leads to a collectively efficient configuration $(\underset{\sim}{d}^*, \underset{\sim}{s}^*)$, then the assumptions that $S_i = A_i$, $\forall\, i \in N$, and $g^*(\cdot \mid d, \alpha) = g(\cdot \mid d, \underset{\sim}{s}^*(\alpha))$ imply that we can find a balanced transfer scheme τ which together with $\underset{\sim}{d}^*$ constitutes an incentive compatible mechanism. Furthermore, $\underset{\sim}{d}^*$ remains collectively efficient with respect to g^*. Therefore, if the self-selection constraints are satisfied, the mediator can evaluate, on the basis of what is common knowledge, the signals that players would have communicated in equilibrium depending on their type. The mediator would obtain an incentive compatible mechanism by choosing the transfers $\tau_i(x, \alpha) = t_i(x_i, \underset{\sim}{s}^*(\alpha))$, proposing the outcome function g^* and recommending the decision rule $\underset{\sim}{d}^*$. No player will have an advantage in lying about his type or in disobeying secretly the recommended course of action. In fact, in a revelation game where the incentive constraints do not hold, at least one player would find it to his advantage to lie to himself before implementing his signaling strategy, or to deviate from the implementation of his individual action. The self-selection constraints would no longer be binding either.

2.2 Moral hazard and adverse selection: two polar cases

Our general model involves both incomplete information (with respect to types) and imperfect information (with respect to actions). A model of pure moral hazard corresponds to complete information and a model of pure adverse selection to perfect information.

Let us first consider the case of complete information: here, the space of types is reduced to a single element $A_i = \{\alpha_i^0\}$ for all players $i \in N_i$, and the certainty (i.e. $p(\alpha^0) = 1$) that the true types are $\alpha^0 = (\alpha_i^0, \ldots, \alpha_i^0, \ldots, \alpha_i^0)$ becomes common knowledge. Since information is complete, the expression of types can be implicitly incorporated into the utilities. The latter can be written as functions $u_i(x_i, d_i)$ of the observable result x and the non-observable action d_i. Even the signaling strategies may be eliminated. Only the actions D remain significant. The structure of the game is now reduced to an outcome function $g : D \to D(X)$ which gives the probability of getting an outcome $x \in X$ conditional on the players implementing a profile $d \in D$ of actions. *The problem of moral hazard is characterized by the remaining imperfect information with regard to actions.*

A balanced transfer scheme for a problem of moral hazard is written as:

$$t : X \to \mathbb{R}^n, \forall\, x \in X, \sum_i t_i(x) = 0,$$

which involves only the observable outcomes. Let us now consider when a transfer scheme t can sustain a configuration of actions $d^* \in D$, by satisfying the self-selection constraints:

$$\sum_x (u_i(x, d_i) + t_i(x))g(x \mid d_i, d^*_{-i}) \leq \sum_x (u_i(x, d^*_i) + t_i(x))g(x \mid d^*).$$

Furthermore, collective efficiency would require choosing a profile of actions such that:

$$d^* \in \arg\max_d \sum_x \left(\sum_i u_i(x, d_i) \right) g(x \mid d).$$

Using the arguments of the previous section, it is easy to verify that one can impose individual rationality.

A canonical example of the problem of pure moral hazard has been exhibited by Holmström (1982) and discussed by Radner et al. (1986), Radner and Williams (1994) or Legros and Matsushima (1991) and more recently by Fudenberg et al. (1994) and d'Aspremont and Gérard-Varet (1998). The n players form a team (or a partnership) and together contribute to achieve a collective outcome. This outcome is an observable "output," whereas the individual actions are non-observable "inputs." The outcome function is treated like a "production function." Members of the team must share out the monetary output of the collective action, and the division must be totally balanced. One can then derive (generic) conditions on the outcome function g which guarantee that a collectively efficient profile of actions can be based on balanced transfers, irrespective of the utilities $(u_i)_{i \in N}$. The conditions given by d'Aspremont and Gérard-Varet (1998) are more general than the ones introduced by Fudenberg et al. (1994).

The second polar case arises when imperfect information about the actions is removed, so that only incomplete information remains, characterizing the problem of adverse selection. Let us suppose that players' choice of actions are predetermined such that for all $i \in N$ we have $D_i = \{d_i^0\}$. The utility functions $u_i(x_i, \alpha_i)$ now implicitly incorporate the actions d_i^0 and are therefore functions of the public outcome $x \in X$ and the type $\alpha_i \in A_i$. Furthermore, the application of the revelation principle brings us to a situation where, for each player $i \in N$, the set of signals coincides with that of types, i.e. $S_i = A_i$. A mechanism, composed of an outcome function $\gamma : A \to D(X)$ and a transfer scheme $t : X \times A \to \mathbb{R}^n$, is incentive compatible if:

$$\begin{aligned}
&\forall\, i \in N_i, \forall\, \alpha_i \in A_i, \forall\, \alpha_i \in A_i, \\
&\sum_{\alpha_{-i}} \left(\sum_x [u_i(x, \alpha_i) + t_i(x, \alpha_i, \alpha_{-i})] \gamma(x \mid \alpha_i, \alpha_{-i}) \right) p(\alpha_{-i} \mid \alpha_i) \\
\leq &\sum_{\alpha_{-i}} \left(\sum_x [u_i(x, \alpha_i) + t_i(x, \alpha)] \gamma(x \mid \alpha) \right) p(\alpha_{-i} \mid \alpha_i).
\end{aligned}$$

Collective efficiency as defined above is satisfied, but can be strengthened by choosing the function γ which stipulates:

$$\forall\, \alpha \in A, \forall\, x \in X, \gamma(x \mid \alpha) > 0 \text{ implies } \sum_i u_i(x', \alpha_i) \leq \sum_i (x, \alpha_i), \forall\, x' \in X.$$

One canonical example of the problem of adverse selection is the case where X is a set of public projects and $u_i(x, \alpha_i)$ represents what player i is willing to pay for project x Clarke (1971); Groves

(1973); d'ASPREMONT and Gérard-Varet (1976), Green and Laffont (1979). d'ASPREMONT and Gérard-Varet (1976), d'Aspremont et al. (1990, 1995a) and Johnson et al. (1990) (generic) conditions are put on the structure of beliefs p which entail the existence of balanced transfers, irrespective of the profile of utilities and the *deterministic* rule[11] of selecting collectively efficient outcomes. Such transfers, together with the deterministic rule, form an incentive compatible mechanism.

Problems of bargaining with incomplete information provide other interesting examples (see Kennan and Wilson, 1993, for a review). Let us consider two players $N = \{1, 2\}$. Player 1 has one unit of a good which he believes is worth $\alpha_1 > 0$ units of money. He wants to sell the good. Player 2 attributes the value $\alpha_2 > 0$ to the ownership of the good, and is the buyer. Both α_1 and α_2 are private information. Utilities[12] are given by $u_1(1, \alpha_1) = \alpha_1$ and $u_2(1, \alpha_2) = \alpha_2$ (with $u_1(0, \alpha_1) = u_2(0, \alpha_2) = 0$). A mechanism is a function $\gamma : A_1 \times A_2 \to [0, 1]$ giving the probability $\gamma(\alpha_1, \alpha_2)$ that a trade will take place, given the players' announced valuations α_1 and α_2 respectively. The mechanism is completed[13] with a transfer scheme $t : A_1 \times A_2 \to \mathbb{R}$, where $t(a_1, a_2)$ is the monetary transfer from the buyer to the seller, given their announced valuations α_1 and α_2. The incentive constraints are:

$$\begin{aligned}
&\forall\, \alpha_1 \in A_1, \forall\, a_1 \in A_1,\\
&\sum_{\alpha_2}(t(a_1, \alpha_2) - \gamma(a_1, \alpha_2)\alpha_1)p(\alpha_2 \mid \alpha_1) \leq \sum_{\alpha_2}(t(\alpha_1, \alpha_2) - \gamma(\alpha_1, \alpha_2)\alpha_1)p(\alpha_2 \mid \alpha_1)\\
&\forall\, \alpha_2 \in A_2, \forall\, a_2 \in A_2,\\
&\sum_{\alpha_1}(\gamma(\alpha_1, a_2)\alpha_2 - t(\alpha_1, a_2))p(\alpha_1 \mid \alpha_2) \leq \sum_{\alpha_1}(\gamma(\alpha_1, \alpha_2)\alpha_2 - t(\alpha_1, \alpha_2))p(\alpha_1 \mid \alpha_2).
\end{aligned}$$

Furthermore, collective efficiency imposes $\gamma(a_1, a_2) = 1$ (resp. $= 0$) if and only if $a_2 - a_1 > 0$ (resp. < 0).

This simple model of bargaining between a buyer and a seller highlights an inherent difficulty with any problem of adverse selection. Even under private values and with transferable utilities, individual incentives can be inconsistent with the *interim* individual rationality constraints:

$$\begin{aligned}
&\forall\, i \in N, \forall\, \alpha_i \in A_i,\\
&\sum_{\alpha_{-i}}\left(\sum_{x}[u_i(x, \alpha_i) + t_i(x, \alpha)]\gamma(x \mid \alpha)\right)p(\alpha_{-i} \mid \alpha_i) \geq u_i^0(\alpha_i),
\end{aligned}$$

where $u_i^0(\alpha_i)$ is the reservation valuation of player i of type α_i. This is in fact the core of the argument in Myerson and Satterthwaite (1983). Examples in other contexts are provided in d'Aspremont and Gérard-Varet (1979b) or Laffont and Maskin (1979).

11. Unlike our general model where *stochastic* mechanisms are considered, most of the literature has discussed *deterministic* mechanisms, i.e. functions $g : A \to X$. In that case, transfers are given by a function $t : A \to \mathbb{R}^n$. Johnson et al. (1990) introduce observable actions and discuss the case of common values.

12. We follow here Myerson and Satterthwaite (1983) in taking a model with private values. Akerlof (1970) considers a very different model with common values, where $u_2(1, \alpha_1) = \alpha_1$ and $u_2(1, \alpha_2) = \alpha_1, \alpha_2$ being irrelevant.

13. This model can also be applied to sequential mechanisms. The probability $\gamma(\alpha_1, \alpha_2)$ is interpreted as the probability of reaching arrangement discounted on the basis of the time period in which it takes place and $t(\alpha_1, \alpha_2)$ is the expected present value of the transfers. Thus, when a procedure is likely to go onto infinity, and if after exchange the gains are discounted at the rate δ, we have $\gamma(\alpha_1, \alpha_2) = \delta^{\tau(\alpha_1, \alpha_2)}$, with $\tau(\alpha_1, \alpha_2)$ being the period during which transaction takes place (when $\tau(\alpha_1, \alpha_2) = \not\infty$, exchange never takes place). Such an interpretation is acceptable when the negotiating parties continue to bargain as long as there are unexploited gains from trade.

3. Environments where individual incentives and collective efficiency are compatible

The general model we consider here is an extension of the classical problems of moral hazard and adverse selection. These two problems are special cases of our model. It is possible to identify certain conditions on beliefs (which characterize adverse selection) as well as the outcome function (characterizing moral hazard) which, at least given transferable individual utilities and private values, guarantee the compatibility of collective efficiency and individual incentives. In this section, we present these conditions, using duality arguments. The discussion of the previous section allows us to restrict to a revelation game.

3.1 A necessary and sufficient condition

Let us consider a utility profile $(u_i)_{i\in N}$ and a structure of beliefs p. Let $\underset{\sim}{d}^* : A \to D$ be a decision rule and $\gamma : D \times A \to \Delta(X)$ an outcome function. Saying that there exist balanced transfers which form an incentive compatible mechanism together with $\underset{\sim}{d}^*$ is equivalent to solving a linear system of inequalities in the transfers derived from (6') and (7'), that is:

$$
\begin{aligned}
&\exists \tau \in \mathbb{R}^{X\times A} \text{ such that } \forall\, i \in N, \forall\, \alpha_i \in A_i, \forall\, a_i \in A_i,\ \forall\, d_i \in D_i \mid \\
&\sum_{\alpha_{-i}} \left(\sum_x [\tau_i(x,\alpha)\gamma(x \mid \underset{\sim}{d}^*(\alpha),\alpha) - \tau_i(x,a_i,\alpha_{-i})\gamma(x \mid d_i, \underset{-i}{\underset{\sim}{d}^*}(\alpha_{-i}), a_i, \alpha_{-i})] \right) p(\alpha_{-i} \mid \alpha_i) \\
\geq &\sum_{\alpha_{-i}} \left(\sum_x [u_i(x, \underset{i}{\underset{\sim}{d}}(\alpha_i),\alpha_i)\gamma(x|d_i|, \underset{-i}{\underset{\sim}{d}^*}(\alpha_i), a_i, \alpha_{-i}) - u_i(x, \underset{i}{\underset{\sim}{d}^*}, (\alpha_i), \alpha_i)\gamma(x \mid d^*(\alpha),\alpha)] \right) \\
&p(\alpha_{-i} \mid \alpha_i) \\
\text{and} &\forall\, \alpha \in A, \forall\, x \in X, \sum_i \tau_i(x,\alpha) = 0.
\end{aligned}
$$

Duality theorems (or the "theorem of the alternative," see Fan (1956) for example) can be used to derive a necessary and sufficient condition.

Denote by $\lambda_i(\alpha_i, a_i, d_i) \geq 0$, $i \in N$, $d_i \in D_i$, the dual variables associated with the incentive constraints, and $\mu(\alpha, x), \alpha \in X, x \in A$, the dual variables associated with the budget constraints. The necessary and sufficient condition can be written as: either there is no $\lambda \in X_{i\in N}\mathbb{R}^{A_i\times D_i}, \lambda \neq 0$, and no $\mu \in \mathbb{R}^{X\times A}$ such that:

$$
\begin{aligned}
&\forall\, i \in N, \forall\, \alpha \in A, \forall\, x \in X, \\
&\gamma(x \mid \underset{\sim}{d}^*(\alpha),\alpha)p(\alpha_{-i} \mid \alpha_i)\sum_{a_i}\sum_{d_i} \lambda_i(a_i,\alpha_i,d_I) \\
&-\sum_{a_i}\sum_{d_i} \lambda_i(\alpha_i,\alpha_i d_i)\gamma(x \mid d_i, \underset{-i}{\underset{\sim}{d}^*}(\alpha_i)\alpha)p(\alpha_{-i} \mid \alpha_i) + \mu(x,\alpha) = 0
\end{aligned}
\tag{8'}
$$

or we have:

$$
\begin{aligned}
\sum_i \Bigg\{ \sum_\alpha \cdot \sum_x \sum_{a_i} \sum_{d_i} \lambda_i(a_i,\alpha_i,d_i)[u_i(x_i,d_i,\alpha_i)\gamma(x \mid d_i, \underset{-i}{\underset{\sim}{d}^*}(\alpha_{-i}), \alpha_{-i}, \alpha_{-i}) \\
- u_i(x, \underset{i}{\underset{\sim}{d}^*}(\alpha_i), \alpha_i)\gamma(x \mid \underset{\sim}{d}^*(\alpha),\alpha)]p(\alpha_{-i} \mid \alpha_i) \Bigg\} \leq 0.
\end{aligned}
\tag{9'}
$$

This necessary and sufficient condition, for the existence of balanced and incentive compatible transfers is about all variables simultaneously (utilities, beliefs, the outcome function). We shall now concentrate on beliefs and the outcome function.

3.2 An informational based condition

Apart from covering all the elements of the problem, the last condition does not take advantage of the requirement of collective efficiency. Such a requirement, given a configuration of utilities $(u_i)_{i\in N}$, implies in the present context the choice of a decision rule $\underset{\sim}{d}^*$ satisfying (3').

We will now consider a condition on beliefs and the outcome function which will ensure the possibility of setting up an incentive compatible and collectively efficient mechanism, with balanced transfers, irrespective of the utility profile $(u_i)_{i\in N}$.

Condition $C^*_{\underset{\sim}{d}^*}$ Let $\underset{\sim}{d}^* \in \underset{\sim}{D}$ be a profile of decision rules. A structure of beliefs p and an outcome function γ satisfy the following condition:
either there is no $\lambda \in X_{i\in N}\mathbb{R}^2_i$, $\lambda \neq 0$, such that (8') is satisfied,
or we have:

$$\begin{aligned} &\forall\, i \in N, \forall\, \alpha \in A, \forall\, x \in X \\ &\gamma(x \mid \underset{\sim}{d}^*(\alpha), \alpha) p(\alpha_{-i} \mid \alpha_i) \sum_{(\alpha_i,\alpha_i)} \sum_{d_i} \lambda_i(a_i, \alpha_i, d_i) \\ &- \sum_{(\alpha_i,\alpha_i)} \sum_{d_i} \lambda_i(\alpha_i, \alpha_i, d_i) \gamma(x \mid d_i, \underset{-i}{\underset{\sim}{d}^*}(\alpha_{-i}), \alpha) p(\alpha_{-i} \mid \alpha_{-i}) = 0 \end{aligned} \tag{10}$$

We have the following:

Theorem 1 *Assume that the structure of beliefs p and the outcome function γ satisfy condition $C^*_{\underset{\sim}{d}^*}$ for a profile of decision rules $\underset{\sim}{d}^* \in \underset{\sim}{D}$ and that $\gamma(x \mid d, \alpha)$ is constant in α. Then for any utility profile $(u_i)_{i\in N}$ for which the rule $\underset{\sim}{d}^*$ is collectively efficient given γ, there exist balanced transfers which, together with this rule, form an incentive compatible, balanced, ex ante individually rational and collectively efficient mechanism.*

This result may be proved in many ways. One is to show that condition $C^*_{\underset{\sim}{d}^*}$ implies the previous necessary and sufficient condition, using collective efficiency.[14] Another is to use a second formulation,[15] which is equivalent to the condition $C^*_{\underset{\sim}{d}^*}$, but is now its "primal" version.[16]

14. This method is used in d'Aspremont and Gérard-Varet (1979a, 1998) respectively for the problem of pure adverse selection and that of pure moral hazard.
15. See d'Aspremont et al. (1995a).
16. To see this, simply apply the theorem of the alternative.

Condition $C_{\underset{\sim}{d}^*}$ Let $\underset{\sim}{d}^* \in \underset{\sim}{D}$ be a profile of decision rules. For any function

$$\beta : X \times A \to \mathbb{R}, \exists\, \tau \in \mathbb{R}^{X \times A}, \text{ such that } \forall\, i \in N, \forall\, a_i \in A_i, \forall\, \alpha_i \in A_i, \forall\, d_i \in D_i,$$
$$\sum_{\alpha_{-i}} \sum_{x} [\tau_i(x, \alpha)\gamma(x \mid \underset{\sim}{d}^*(\alpha), \alpha) - \tau_i(x, a_i, \alpha_{-i})\gamma(x \mid d_i, \underset{-i}{\underset{\sim}{d}^*}(\alpha_{-i}), a_i\alpha_{-i})]p(\alpha_{-i} \mid \alpha_i) \geq 0$$
$$\text{and, } \forall\, \alpha \in A, \forall\, x \in X, \sum_i \tau_i(x, \alpha) = \beta(x, \alpha).$$

The proof goes as follows. Let $(u_i)_{i \in N}$ be a profile of utility functions and $\underset{\sim}{d}^* \in \underset{\sim}{D}$, a profile of decision rules which is collectively efficient for an outcome function γ. Following Clarke, Groves and Vickrey, we introduce the transfers given by:

$$t_i^0(x, \alpha) \overset{\text{def}}{=} \sum_{j \neq i} u_j(x, \underset{j}{\underset{\sim}{d}^*}(\alpha_j), \alpha_j)\tilde{\gamma}(x \mid \underset{\sim}{d}^*(\alpha)),$$

with $\tilde{\gamma}(x \mid \underset{\sim}{d}^*(\alpha)) = \gamma(x \mid \underset{\sim}{d}^*(\alpha), \alpha)$, assumed to be constant in α.

Collective efficiency means that such transfers generate an incentive compatible mechanism.[17]

Clearly, there is no guarantee that the transfers are balanced: $\sum_i t_i^0(x, \alpha)$ may be either positive or negative. The condition $C_{\underset{\sim}{d}^*}$ in fact states that the total, be it a surplus or a deficit, may be allocated amongst all players in a way which is incentive compatible. All we need is the requirement that:

$$\beta(x, \alpha) = -\sum_i t_i^0(x, \alpha),$$

and the application of the condition: the transfers $t = \tau + t^0$ are both balanced and incentive compatible. Finally, the earlier discussion may be repeated to obtain ex ante individual rationality.

Thus, condition $C_{\underset{\sim}{d}^*}$ is sufficient to obtain balanced transfers that allow an efficient and incentive compatible mechanism for an efficient profile of decision rules. It is possible to strengthen this property by considering a stronger condition that holds *irrespective* of the decision rule $\underset{\sim}{d}^*$ likely to be selected by introducing a condition similar to condition B in d'Aspremont and Gérard-Varet (1982, 1998).

The general model presented here can be used to analyze many particular situations, relevant to markets and organizations. Examples include relations between the regulator and the management of a public enterprise (or between the shareholders and the management of a private enterprise), those between producers and consumers of a public good, tendering or auctioning procedures, insurance contracts, and even employment contracts between an employer and an employee. Challier and Gérard-Varet (1994) follow the same lines of argument for a discussion in the context of employment contracts.

It is however important to recognize that the conditions under which incentive compatible and collectively efficient (in its classical sense) mechanisms exist in the presence of moral hazard and

17. In fact a strongly incentive compatible mechanism: this property continues to hold with dominant strategies. This is the point made by Groves (1973).

adverse selection, are crucially linked to two central assumptions: "transferability of utilities" and "private values."

To begin with, even when utilities are transferable, the application of the condition $C_{\underset{\sim}{d}^*}$ and the existence of incentive compatible and collectively efficient mechanisms are put into question by the introduction of common values. This is indeed the main lesson to be drawn from the famous example in Akerlof (1970). However, d'Aspremont and Gérard-Varet (1982, 1998) suggest a stronger condition for the existence of incentive compatible mechanisms in this case (condition B). The ensuing mechanisms do not have to satisfy efficiency in its traditional sense. As suggested in Forges (1994), one may indeed question whether the classical definition of efficiency is relevant outside a world of private values.

The assumption of transferable utilities opens the possibility of using transfers between players as an instrument to fulfill individual incentives. To abandon it, would not only eliminate this possibility, but also introduce all the difficulties, which we have ignored here, and which relate to the combination of incentives and risk sharing. One may envisage adopting a sequential approach. In the case of pure moral hazard, "repetition" helps reconcile collective efficiency with the principles of non-cooperative individual behaviour (see Fudenberg et al., 1994; d'Aspremont and Gérard-Varet, 1998). However, with adverse selection as well as moral hazard, the question is no longer so simple, mainly due to problems of individual rationality under incomplete information. Furthermore, if there is repetition, it would be difficult to ignore the selection of equilibrium on the basis of a criterion of "sequential rationality" which implies that each player's strategy is the best response to the strategies of others, not only during the course of the game, but also for every possible contingency he may find himself in and for the rest of the game.

A minimal sequential approach, without going as far as repetition, combined with some coordination can be introduced through "communication mechanisms" as suggested in Forges (1986) and Myerson (1985, 1991). In this case, one has to consider a (finite) set of messages R_i that each player $i \in N$ receives from the mechanism. These messages are distinct from the signals S_i the player transmits. Let $R = X_{i \in N} R_i$ be the set of all possible configurations of messages, and $S = X_{i \in N} S_i$ the set of all possible configurations of signals. A *communication mechanism* is a function π of S to the set $D(R)$ of all possible probability distributions over R.

We can add communication mechanisms to our model. Let $D_i : R_i \to D_i$ be a strategy that player i chooses (privately) on the basis of a message τ_i communicated to him (secretly). The resulting action is denoted by $\delta_i(r_i)$. Let $D_i \overset{\text{def}}{=} D_i^{R_i}$ be the set of all such strategies and assume $D = X_{i \in N} D_i$. Given a communication scheme π and an outcome function g, we can form a new outcome function:

$$g^{\pi}(x \mid \delta, s) = \sum_{r} g(x \mid \delta(r), s) \pi(r \mid s).$$

The analysis may then continue with this outcome function. The self-selection constraints now apply to decision rules $\underset{\sim}{s} : A \to S$ and $\underset{\sim}{d} : A \to \delta$, and the "revelation principle" allows the identification of A_i and Δ_i on the one hand, and R_i and D_i on the other. We are then back to the incentive constraints discussed by Myerson (1991). In fact, we have considered here communication relative to "inputs" (signals), whereas Myerson (1991) argues the case for extending communication to "outputs" (messages).

A basic conflict lies at the heart of the question of combining collective efficiency with individual incentives. Individual incentives depend on what each person knows privately, whereas collective

efficiency assumes the pooling of this information. A second best approach, called "interim incentive efficiency," has been put forward to deal with this difficulty (Holmström et al., 1983). An incentive compatible mechanism (in general a mechanism which is supported by a non-cooperative equilibrium) which dominates it uniformly as far as the interim payoffs are concerned. However, it is possible for such mechanisms to select Pareto dominated outcomes, once the intrinsic communication has taken place (Forges, 1994).

We must underline finally the limitations of the revelation principle upon which we have relied so much. This principle indicates that, by studying the revelation mechanisms, it is possible to determine an efficiency frontier in the set of all feasible mechanisms for which, in equilibrium, each agent reveals sincerely his private information. It does not however preclude the dependence of efficient mechanisms on all elements which are common knowledge, but often contingent, such as the form of beliefs or individual utilities, maybe even the number of participants. In practice, the question is to find a procedure – either static like an auction or dynamic such as a negotiation – and equilibrium strategies which allow an efficient mechanism to be sustained (or "implemented"). The procedures which are actually used, often lack such contingency. Forges (1990) defines a general class of "universal mechanisms." As underlined by Wilson (1985), revelation mechanisms use complex rules, taking contingencies into account, but lead to an equilibrium which is easy to calculate ("to tell the truth"), whereas procedures used in practice have simple rules, though the players on the other hand must have complex equilibrium strategies, which integrate all contingencies they know about. Theory will gain much by taking more into consideration such "games" where practical organizational procedures determine the rules.

References

George Akerlof. The market for lemons: qualitative uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84:488–500, 1970.

Armen A Alchian and Harold Demsetz. Production, information costs, and economic organization. *American Economic Review*, 62(5):777–795, 1972.

MC Challier and LA Gérard-Varet. Incentive payment schemes in teams with moral hazard and adverse selection. Technical report, Mimeo, 1994.

Edward H Clarke. Multipart pricing of public goods. *Public choice*, II:19–33, 1971.

C d'ASPREMONT and LA Gérard-Varet. Un modèle de négociation pour certains problèmes internationaux de pollution. *Revue d'Économie Politique*, 4:547–562, 1976.

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

Claude d'Aspremont and Louis-André Gérard-Varet. On bayesian incentive compatible mechanisms. In J.J. Laffont, editor, *Aggregation and Revelation of Preferences*. North Holland, Amsterdam, 1979b.

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. Mécanismes de marchés, concurrence et contraintes d'incitation. In J. Cartelier, editor, *La formation des grouleurs Économiques*, pages 83–126. PUF, Paris, 1989.

Claude d'Aspremont and Louis-André Gérard-Varet. Linear inequality methods to enforce partnerships under uncertainty: An overview. *Games and Economic Behavior*, 25(2):311–336, 1998.

Claude d'Aspremont, Jacques Crémer, and Louis-André Gérard-Varet. Incentives and the existence of pareto-optimal revelation mechanisms. *Journal of Economic Theory*, 51(2):233–254, 1990.

Claude d'Aspremont, Jacques Crémer, and Louis-André Gérard-Varet. Correlation, independence, and Bayesian incentives. Technical report, Mimeo, 1995a.

Claude d'Aspremont, Rodolphe Dos Santos Ferreira, and Louis-André Gérard-Varet. Fondements stratégiques de l'équilibre en économie: coordination, rationalité individuelle et anticipations. In L.-A. Gérard-Varet and J.C. Passeron (eds.), editors, *Le modèle et l'enquête*, pages 447–468. Editions de l'EHESS, Paris, 1995b.

Jacques H Drèze and Dominique de la Vallée Poussin. A tâtonnement process for public goods. *Review of Economic Studies*, 38:133–150, 1971.

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

Françoise Forges. An approach to communication equilibria. *Econometrica*, 54:1375–1385, 1986.

Françoise Forges. Universal mechanisms. *Econometrica*, 58:1341–1364, 1990.

Françoise Forges. Posterior efficiency. *Games and Economic Behavior*, 6(2):238–261, 1994.

Drew Fudenberg, David K Levine, and Eric Maskin. The folk theorem with imperfect public information. *Econometrica*, 62(5):997–1039, 1994.

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

Sanford J Grossman and Oliver D Hart. An analysis of the principal-agent problem. *Econometrica*, 51:7–45, 1983.

Theodore Groves. Incentives in teams. *Econometrica*, 31:617–663, 1973.

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. Moral hazard in teams. *The Bell Journal of Economics*, 13:324–340, 1982.

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

Leonid Hurwicz. Optimality and informational efficiency in resource allocation processes. In S. Karlin K. Arrow and P. Suppes, editors, *Mathematical Methods in the Social Sciences*, pages 27–46. Stanford University Press, Stanford, CA, 1960.

Leonid Hurwicz. On informationally decentralized systems. In R. Radner and C.D. McGuire, editors, *Decision and Organization*, pages 297–336. North Holland, Amsterdam, 1972.

Scott Johnson, John W Pratt, and Richard J Zeckhauser. Efficiency despite mutually payoff-relevant private information: The finite case. *Econometrica*, 58:873–900, 1990.

John Kennan and Robert Wilson. Bargaining with private information. *Journal of Economic Literature*, 31(1):45–104, 1993.

Jean-Jacques Laffont and Eric Maskin. A differential approach to expected utility maximizing mechanisms. In J. J. Laffont, editor, *Aggregation and Revelation of Preferences*, pages 289–305. North Holland, Amsterdam, 1979.

Patrick Legros and Hitoshi Matsushima. Efficiency in partnerships. *Journal of Economic Theory*, 55 (2):296–322, 1991.

Edmond Malinvaud. A planning approach to the public good problem. *Swedish Journal of Economics*, 173:96–112, 1971.

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.

Roger B Myerson. *Game theory: analysis of conflict. 1991*. Harvard University Press, Cambridge (MA), 1991.

Roger B Myerson and Mark A Satterthwaite. Efficient mechanisms for bilateral trading. *Journal of economic theory*, 29(2):265–281, 1983.

Roy Radner and Steven R Williams. Efficiency in partnership when the joint output is uncertain. In J. Ledyard, editor, *The Economics of Informational Decentralization*. Kluwer, London, 1994.

Roy Radner, Roger Myerson, and Eric Maskin. An example of a repeated partnership game with discounting and with uniformly inefficient equilibria. *Review of Economic Studies*, 53(1):59–69, 1986.

Paul A Samuelson. The pure theory of public expenditure. *Review of Economics and Statistics*, 36 (4):387–389, 1954.

William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *Journal of finance*, 16(1):8–37, 1961.

Friedrich August von Hayek. The use of knowledge in society. *American Economic Review*, 35: 519–530, 1945.

Knut Wicksell. Finanztheoretische untersuchungen and das steuerwesen schweden's. In R. Musgrave and A.T. Peacock, editors, *Classics in the Theory of Public Finance*. Macmillan, London, 1896.

Robert Wilson. Efficient trading. In G. Feiwel, editor, *Issues of Contemporary Microeconomics and Welfare*, pages 169–208. State University of New York Press, Albany, 1985.