NOISY OBSERVATION IN ADVERSE SELECTION MODELS

by B. CAILLAUD
R. GUESNERIE
P. REY

N° 8921

January 1988
revision July 1989

NOISY OBSERVATION IN ADVERSE SELECTION MODELS+

by B. Caillaud*
R. Guesnerie**
P. Rey***

* CEPREMAP, Paris
** DELTA, Paris
*** ENSAE, Paris and MIT, Cambridge

+ This article appeared in a very first and incomplete version in Caillaud's thesis at EHESS, Paris, June 1985. The final version benefited from the comments of participants at Seminars in Bonn, LSE and MIT, and from remarks of two referees and an Associate Editor. We also gratefully acknoweldge helpful discussions with P.A. Chiappori, O.D. Hart, B. Jullien and J. Tirole.

ABSTRACT

NOISY OBSERVATION IN ADVERSE SELECTION MODELS

We consider a principal-agent contracting problem under incomplete information where some of the agent's actions are imperfectly observable. Contracts take the form of noisy reward schedules, where transfers depend on the observable signals. We first review situations where an optimal noisy reward schedule allows the principal to reach his pure adverse selection utility despite the imperfection of observation. Among all possible schedules, those linear with respect to the imperfectly observed variables implement a given adverse selection mechanism independently of the specification of the noise and are thus desirable. We exhibit sufficient conditions under which linear reward schedules implement a given mechanism. Finally, we characterize necessary conditions for a mechanism to be implementable under noisy observation by a linear schedule, and by quadratic schedules. We give the geometric intuition behind all results.

Journal of Economic Literature classification : 026

Key words : Principal-Agent, Contracts, Adverse Selection, Moral Hazard, Linear Reward Schedules.

RESUME

OBSERVATION IMPARFAITE DANS LES MODELES DE SELECTION

On considère un problème de contrat entre un principal et un agent en information incomplète et où certaines des actions prises par l'agent ne sont qu'imparfaitement observables. Les contrats prennent la forme de schémas de récompense, où les transferts dépendent des signaux observables. Dans un premier temps, on passe en revue des situations où un schéma de récompense optimal permet au principal d'atteindre la même satisfaction que dans le simple modèle de sélection, ceci malgré l'imperfection de ses observations. Parmi tous les schémas possibles, les schémas de récompense linéaires sont attractifs car ils concrétisent un mécanisme de sélection indépendamment de la spécification du bruit d'observation. On trouve alors des conditions suffisantes pour qu'un tel schéma concrétise un mécanisme donné. Enfin, on caractérise des conditions nécessaires pour qu'un mécanisme soit réalisable en situation d'observation imparfaite, grâce à un schéma de récompense linéaire ou quadratique. On insiste sur l'intuition géométrique des résultats.

Mots clés : principal-agent, contrats, sélection, aléa moral, schémas de rémunérations linéaires.

I. INTRODUCTION

This paper is a contribution to principal-agent contracting theory. The starting point of the analysis is a pure adverse selection problem, which has been studied in many different contexts (see Guesnerie-Laffont [1984]). We modify the standard model by assuming that some of the action variables on which the agent's reward is based, are not perfectly observable by the principal or the third party who is supposed to enforce the contract. The pure adverse selection framework is thus extended to introduce the possibility of errors in the observation of the actions of the agent. The model combines the "hidden knowledge" aspect of adverse selection problems and the "hidden action" aspect of pure moral hazard contexts. However, by considering risk neutral agents, one eliminates the insurance question that characterizes moral hazard problems. Thus the situation can be better described as a "noisy" adverse selection situation.

This type of situation is the focus of a recent strand of contract theory, pioneered by Laffont-Tirole [1986] and developed by Picard [1987], Melumad-Reichelstein [forthcoming] and Rogerson [1988] among others. Its applications and relevance to real world analysis are obvious. When the manager of a firm is controled by a regulatory agency or a group of shareholders, the costs and profits achieved are merely a noisy estimator of the actual decisions made by the manager. A worker's production depends upon his intrinsic productivity, but also on the random, unobserved quality of the materials he uses and can therefore only provide an imperfect assessment of the worker's quality.

In situations of noisy adverse selection, the pure adverse selection optimum, i.e. in the absence of noise, provides a desirable benchmark for the principal. The literature has focused on situations where this benchmark is achievable despite the noise of observation. In this paper, we take advantage of the fact that if some variables are subject to errors of observation, some others may be perfectly observed. We show that in general there exist many ways of achieving a pure adverse selection benchmark. We then investigate the possibility of imposing an additional property on the mechanism that attains this benchmark, namely that it be robust to our imperfect knowledge of the noise of observation. We characterize cases where the benchmark can be achieved despite the fact that the principal ignores all the characteristics of the noise (except that it is unbiased) : we call this property "universal implementation".

Again, our reflections on universal implementation are in line with Laffont-Tirole's initial attempt. However, by considering a general setting, the analysis casts the subject in an improved perspective. Existing positive results of previous authors are obtained in limit cases of more general favorable situations and the role of the different parameters of the contracting problem is ascertained. Therefore, these reflections on universal implementation contribute to the understanding of the role of linear schemes in economic contracting.

The paper is constructed as follows. Section II presents our simple framework where some actions are perfectly observed and some actions are observed with noise. In section III we review the body of existing results that guarantee that the principal can achieve under noisy observation, the same utility as in the pure adverse selection situation. This short survey on implementation under noisy observation encompasses more general settings as presented in Melumad-Reichelstein [forthcoming] (hereafter MR) and Caillaud-Guesnerie-Rey [1988] (hereafter CGR). It shows that in our specific framework, there may be a large number of possibilities of implementation under noisy observation, some being very demanding on the knowledge of the noise distribution, some others being more robust to the imperfection of this knowledge.

In section IV, which constitutes the core of the paper, we analyze in detail the condition under which a principal can implement an adverse selection optimum under noisy observation, when he ignores the distribution of the noise of observation of the agent's actions (except the zero mean) and we insist on geometric intuition. We identify the only possible reward schedule that can possibly yield universal implementation (IV A), we exhibit sufficient conditions for this schedule to be indeed an acceptable reward schedule, i.e. sufficient conditions for universal implementation (IV B), and we focus on local incentive compatibility required from reward schedules to exhibit a necessary condition for universal implementation (IV C). Finally, in subsection IV D we relax our requirement of universality, to focus on reward schedules that only require the knowledge of the variance (and of the zero mean) of the distribution of the noise of observation.

II. THE MODEL

We consider a standard principal-agent model under asymmetric information, with a multi-dimensional action space and a one-dimensional information space. The principal designs a contract with the agent who has private information on one characteristic denoted by θ; θ is assumed to belong to Θ , a compact interval of $\mathbb{R}$. The contract bears on a multidimensional action ℓ that the agent can take and on the amount of transfers $t \in \mathbb{R}$ that he can receive from the principal ; ℓ is restricted to be a pair of sub-actions $\ell = (\ell_1, \ell_2)$ where $\ell_i \in L_i$, L_i is a compact subset of $\mathbb{R}^{n_i}$ and $L = L_1 \times L_2$ a compact subset of $\mathbb{R}^n$, $n = n_1 + n_2$, $n_1 \geqslant 1$, $n_2 \geqslant 1$. The agent's VNM utility function depends on action ℓ and characteristic θ and exhibits risk neutrality w.r.t. revenue; it is denoted by $t + U(\ell\ ;\ \theta) = t + U(\ell_1, \ell_2\ ;\ \theta)$.

The pure adverse selection (or hidden knowledge) framework refers to the situation where both actions ℓ_1 and ℓ_2 are observable and verifiable (therefore contractable). It is well known from the Revelation Principle that, when designing the contract, the principal can restrict attention to the set of Direct Incentive Compatible Mechanisms (DICM):

Definition 1 : A DICM is a pair of functions $(\ell(.),t(.))$ mapping the set of characteristics Θ into $L \times \mathbb{R}$, such that for any $(\theta,\theta') \in \Theta^2$:

$$t(\theta) + U(\ell(\theta);\theta) \geq t(\theta') + U(\ell(\theta');\theta)$$

For a comprehensive analysis of this type of adverse selection problems we refer the reader to Guesnerie-Laffont [1984]. In particular, as a consequence of the "taxation principle" (Hammond [1979], Guesnerie [1981]), a DICM is equivalent to a non-linear tax schedule φ :

Proposition 1 : The pair $(\ell(.),t(.))$ is a DICM if and only if there exists a mapping φ from $\mathbb{R}^n$ to $\mathbb{R}$ such that:

$$\forall\theta\in\Theta,\ \ell(\theta)\in \underset{\ell\in L}{\mathrm{Argmax}}\,[\varphi(\ell)+U(\ell;\theta)]$$
$$t(\theta)=\varphi(\ell(\theta))$$

Such a function φ will be called a <u>(ℓ,t)-associated schedule</u>.

In the present problem, the locus of points (ℓ_1,ℓ_2,t) in $\mathbb{R}^{n+1}$ such that $(\ell_1,\ell_2,t) = (\ell_1(\theta),\ell_2(\theta),t(\theta))$ for some $\theta\in\Theta$, is generally a one-dimensional curve, called the "contract curve". Given a DICM $(\ell(.),t(.))$, a (ℓ,t)-associated schedule corresponds to an "hyper surface" in $\mathbb{R}^{n+1}$ which contains the contract curve. Moreover, for each type $\theta\in\Theta$, the set $P(\theta)$ of actions (ℓ_1,ℓ_2) strictly preferred by θ to the contractual actions $(\ell_1(\theta),\ell_2(\theta))$ has an empty intersection with the (ℓ,t)-associated schedule: for all θ, these sets $P(\theta)$ must lie above (in the sense of increasing transfers) the (ℓ,t)-associated surface. The situation is illustrated on Figure 1 for $n_1 = n_2 = 1$.

Note that a (ℓ,t)-associated schedule is uniquely defined only for $(\ell_1,\ell_2)\in\ell(\Theta)$, i.e. for the set of contractual actions. There is thus much freedom in the construction of an associated schedule outside $\ell(\Theta)$. In the paper, we will often make the following assumption:

(B) There exists a large enough penalty M^* such that:

$$\underset{\ell\in L}{\mathrm{Max}}\ U(\ell;\theta) - M^* \le \underset{\theta\in\Theta}{\mathrm{Max}}\ \{t(\theta)+U(\ell(\theta);\theta)\}$$

Assumption (B) allows us to build "M-associated" schedules, i.e. associated schedules that act as a penalty M against deviations far away from the

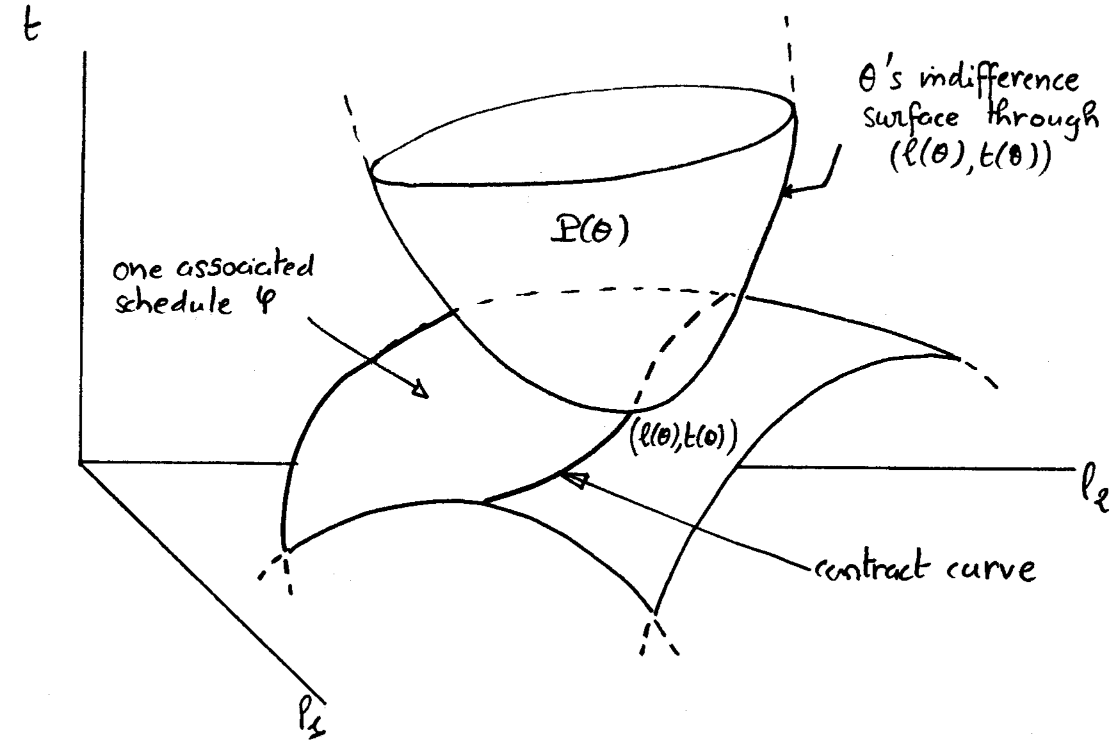
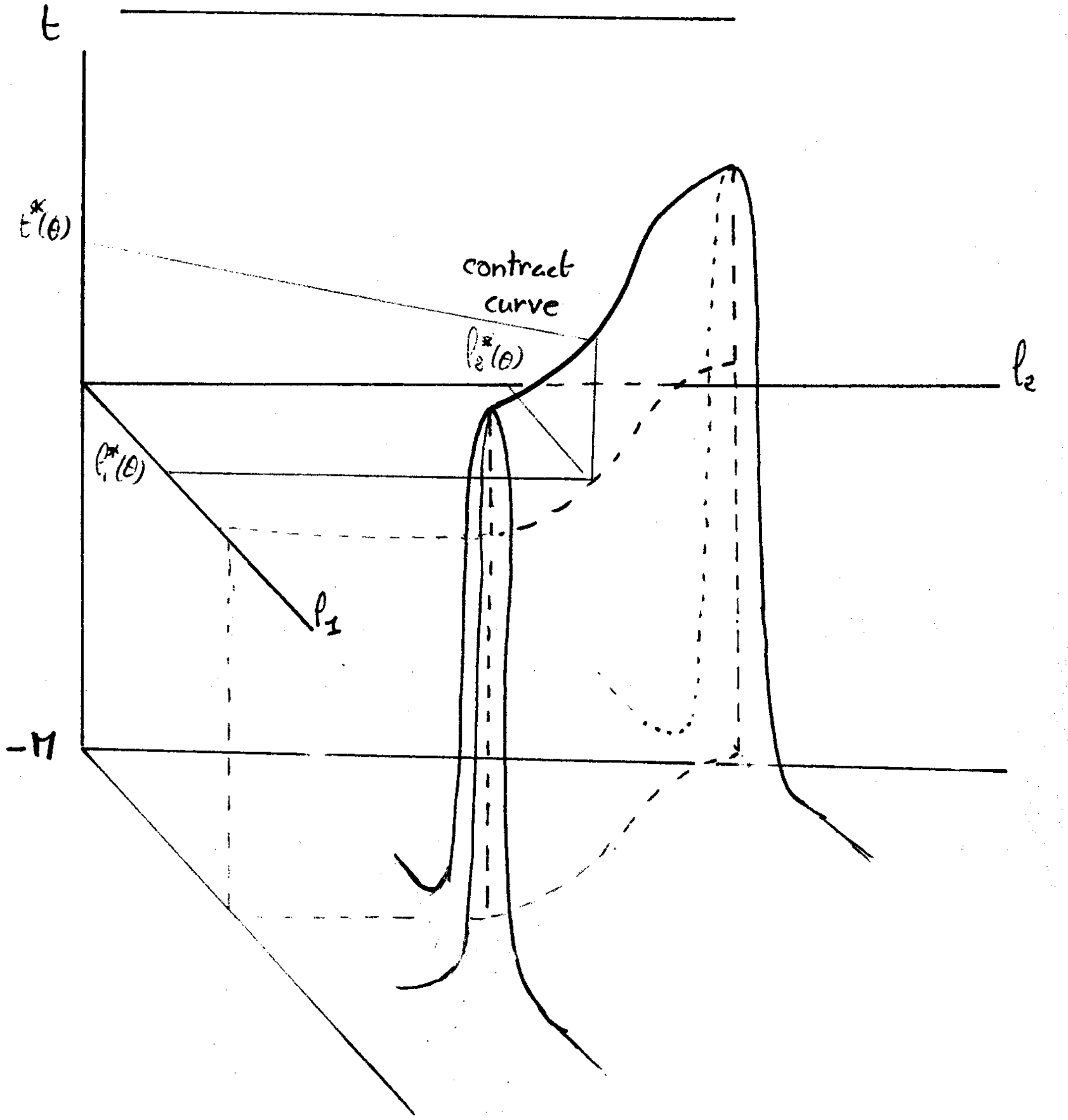


FIGURE 1 : AN ASSOCIATED SCHEDULE

FIGURE 2 : A M-ASSOCIATED SCHEDULE



contract curve.[1] In the limit, we can choose the associated schedule to be $\varphi(\ell) = -M$ with $M \geqslant M^*$ everywhere outside the support $\ell(\Theta)$; this would deter any agent of type θ, when facing the associated schedule $\varphi(.)$, to choose a pair of actions (ℓ_1, ℓ_2) outside $\ell(\Theta)$. This particular associated schedule is called a "knife-edge" associated schedule (In Figure 2, we present an associated schedule that is close to the knife-edge schedule).

Definition 2 : The M-knife edge schedule is defined by:

$$\varphi(\ell) = t(\theta) \text{ if there exists } \theta\epsilon\Theta \text{ such that } \ell = \ell(\theta),^2$$
$$= -M \text{ otherwise}$$

The (partially) noisy observation framework relaxes the assumption of perfect observability of the agent's actions. More precisely we assume that ℓ_2 is perfectly observed whereas ℓ_1 is not . In stead of ℓ_1 , the principal observes the signal $\ell_1' = \ell_1 - \varepsilon$, where the additive noise ε is distributed independently of θ, according to a conditional density $f(.|\ell)$ on $\mathbb{R}$, with mean 0. Moreover in Section III, we specialize to the case where ε is distributed independently of ℓ.[3]

Loosely speaking, the question addressed here is whether introducing a noise in the observation of some action variables induces a welfare loss for the principal compared to the pure adverse selection optimum (given by the DICM or any associated schedule). Throughout the paper, we will assume that the principal is income-risk neutral. With this assumption, we will say that the imperfection of observation is innocuous for the principal's welfare if he can induce any type θ of agent to choose the same actions $(\ell_1(\theta), \ell_2(\theta))$ as in the pure adverse selection framework, while rewarding him with the same expected transfers $t(\theta)$. This idea is reflected in the following notion of implementation under noisy observation:

Definition 3 : A DICM (ℓ, t) is implementable under noisy observation of ℓ_1, if there exists a "noisy reward schedule" ψ from $\mathbb{R}^n$ to $\mathbb{R}$, such that:[4]

$$\forall\theta\in\Theta,\ (\ell_1(\theta),\ell_2(\theta)) \in \underset{\ell\in L}{\mathrm{Argmax}}\ \mathbb{E}[\psi(\ell_1-\varepsilon,\ell_2)\,|\,\ell_1,\ell_2] + U(\ell_1,\ell_2;\theta)$$

$$t(\theta) = \mathbb{E}[\ \psi(\ell_1(\theta)-\varepsilon,\ell_2(\theta))\,|\,\ell_1(\theta),\ell_2(\theta)]\ .$$

This notion of implementation relies on an extension of the notion of associated schedule to the framework with imperfect observation : the reward received by the agent depends on the observable variables (ℓ_1',ℓ_2). In the present setting, the most general mechanism that the principal could propose, takes the form of a revelation mechanism, where an announcement $\theta'\in\Theta$ would determine an observable action $\ell_2(\theta')$ to be taken by the agent, and a payment $t = H(\ell_1',\theta')$ as a function of the observable signal ℓ_1'. An adapted version of Proposition 1 can be shown to hold, i.e. such a revelation mechanism is equivalent to a noisy reward schedule; hence the use of a noisy reward schedule does not entail any restriction in the principal's set of instruments.

Now the question addressed in this paper can be decomposed in two parts:

(a) When is a DICM implementable under noisy observation in the sense of Definition 3 ? How can a noisy reward schedule ψ be computed for a given specification of the noise ? Are there many noisy reward schedules ?

(b) Can we find reward schedules when we have only partial information on the noise distribution (e.g. its variance) ? Could there exist, and under what conditions, "universal" reward schedules, i.e. schedules that implement a DICM whatever the distribution of the observation error ?

III. IMPLEMENTATION VIA NOISY REWARD SCHEDULES : EXISTING GENERAL RESULTS.

In this section, we start from the existence of a DICM (ℓ^*,t^*), and we analyze its implementability under noisy observation in the sense of Definition 3. We view this section as a short summary of results that appeared independently in MR or in CGR.

The next proposition shows that the problem is equivalent to solving a functional equation that guarantees that, in expectation over ε, the noisy reward schedule gives the same incentives to the agent as a (ℓ^*,t^*) - associated schedule.

Proposition 2 : A DICM (ℓ^*,t^*) is implementable (under noisy observation) via a noisy reward schedule ψ if and only if there exists φ, a (ℓ^*,t^*) -associated schedule, such that for all $(\ell_1,\ell_2)\in L$

(1) $$\varphi(\ell_1,\ell_2) = \int \psi(\ell_1 - \varepsilon, \ell_2)\ f(\varepsilon \mid \ell_1,\ell_2)d\varepsilon .$$

Proof : i) Suppose $(\ell^*(\theta),t^*(\theta))$ is implementable via ψ, then :

$$\begin{cases} \ell^*(\theta) \in \underset{(\ell_1,\ell_2)\in L}{\text{Argmax}}\ \mathbb{E}[U(\ell_1,\ell_2,\theta) + \psi(\ell_1 - \varepsilon,\ell_2) \mid \ell_1,\ell_2] \\ \\ t^*(\theta) = \mathbb{E}\left[\psi(\ell_1^*(\theta) - \varepsilon,\ \ell_2^*(\theta) \mid \ell_1^*(\theta),\ell_2^*(\theta)\right] \end{cases}$$

Call $\varphi(\ell_1,\ell_2) = \mathbb{E}[\psi(\ell_1 - \varepsilon,\ell_2) \mid \ell_1,\ell_2]$ then φ satisfies (1) and is an associated schedule.

ii) Suppose there exists φ, a (ℓ^*,t^*) - associated schedule, such that there exists ψ solution of (1). Then for any (ℓ_1,ℓ_2), $\varphi(\ell_1,\ell_2) = \mathbb{E}[\psi(\ell_1-\varepsilon,\ell_2) \mid \ell_1,\ell_2]$ so that ψ implements (ℓ^*,t^*).

Q.E.D.

The associated schedule φ only reflects pure adverse selection considerations. On the contrary, ψ also takes into account the observation error. Given a DICM (ℓ^*,t^*), one finds the noisy reward schedules that implement (ℓ^*,t^*) by first choosing some (ℓ^*,t^*) - associated schedule and then solving the resulting functional equation (1). Proposition 2 holds also for multidimensional Θ, and/or if there is no perfectly observed action ℓ_2, i.e. $n_2 = 0$ (see CGR)[5] ; it can take the form of a more general functional equation when ℓ_1' and ℓ_1 are supposed to be of different dimensions (see MR).

Depending on the properties of the possible associated schedules and of the noise distribution, equation (1) may or may not admit a solution, i.e. the DICM may or may not be implementable under noisy observation ; a noisy reward schedule may even be explicitly computable. In the following, we mention a number of cases in which a DICM (ℓ^*, t^*) is implementable under noisy observation when $n_1 = 1$ and the noise is distributed <u>independently of ℓ</u>, according to the density f(.). Some of these results could be extended in more general frameworks, e.g. when Θ is multidimensional, or $n_1 > 1$ or $n_2 = 0$ (See CGR) ; MR also provides technical results when ε is not independent of ℓ or θ.

- If f(.) is uniform on a compact [-a,a], and there exists a continuous M-extended (ℓ^*, t^*)-associated schedule φ, which admits almost everywhere a partial derivative with respect to ℓ_1, then (1) is solvable explicitly by :

$$\psi(\ell_1', \ell_2) = 2a\left[\sum_{k=0}^{\infty} \frac{\partial \varphi}{\partial \ell_1}(\ell_1' - (2k+1)a, \ell_2)\right] + G(\ell_1', \ell_2) - M$$

where G is a function periodical in ℓ_1' (period 2a) such that $\int_{-a}^{+a} G(x, \ell_2)dx = 0$ for any ℓ_2. (See CGR).

- More generally, if f is continuously differentiable on a compact support [a,b] and there exists a continuously differentiable associated schedule φ, then (1) is solvable (under a mild assumption on the relative size of [a,b] and L, see MR).

- One can find more technical conditions (see CGR) that guarantee that f admits a Fourier transform $\mathcal{F}(f)$ (e.g. f normal), and that there exists a (ℓ^*, t^*) -associated schedule φ with Fourier transform $\mathcal{F}(\varphi)$ such that $\mathcal{F}(\varphi)/\mathcal{F}(f)$ admits an inverse Fourier transform $\mathcal{F}^{-1}(\mathcal{F}(\varphi)/\mathcal{F}(f))$. Under these conditions, (1) is solvable, and a solution is $\psi = \mathcal{F}^{-1}(\mathcal{F}(\varphi)/\mathcal{F}(f))$.

- When there exists an associated schedule that is a polynomial in ℓ_1, then Proposition 3 below indicates the conditions under which (1) is solvable and the way to compute explicitly the solution.

Note that we could depart from the exact resolution of equation (1) and adopt a concept of almost implementation in the spirit of MR. Basically, such extension consists in taking the closure of the above results (and of Proposition 3 below) under the topology of the uniform convergence. But it is possible to adopt other topologies, in particular if we allow for unbounded, or even infinite-valued functions. As an example, when f has compact support [a,b], ℓ_2^* is one-to-one (for simplicity), and φ is the M-knife schedule, the solution of (1) converges, when M goes to infinity, towards a "Mirrlees scheme" : $\psi(\ell_1',\ell_2) = t^*(\theta)$ if $\ell_2 = \ell_2^*(\theta)$ and $\ell_1' \in \left[\ell_1^*(\theta)-b, \ell_1^*(\theta)-a\right]$, $= -\infty$ otherwise. This limit reward schedule implements the DICM under the noise ε. (As is known, this construction extends for example to normal distributions).

The main conclusion of these results is that there is much leeway in the choice of noisy reward schedules that implement a DICM for two reasons. There is first much freedom in solving the convolution equation (1) for a given (ℓ^*,t^*)-associated schedule. Second, one can choose freely any associated schedule provided it satisfies some regularity and integrability properties. This last requirement is intuitively weak, and we will give a more precise assessment of this fact later.

The question is now : can we exploit the freedom in the choice of reward schedules to impose on them some additional desirable properties. An interesting property [6] would be that the noisy reward schedules do not heavily depend on the specification of the noise density f(.). Indeed, the information on f(.) required to compute a noisy reward schedule, may crucially depend on the associated schedule we are starting from. This fact is clearly illustrated by the next proposition.

Proposition 3 : Let us assume that for every ℓ_2, the associated schedule φ is a polynomial in ℓ_1 of degree smaller or equal to m, then there exists a noisy reward schedule ψ which, for a given ℓ_2, is a polynomial of degree smaller or equal to m (in ℓ_1'), the coefficients of which only depend on the moments of the noise distribution of order smaller or equal to m.

Proof : Suppose $\varphi(\ell_1,\ell_2) = \sum_{p\leq m} a_p(\ell_2)\ell_1^p$ and let us look for a solution of the form $\psi(\ell_1', \ell_2) = \sum_{p \leq m} b_p(\ell_2)\ell_1'^{\,p}$. The convolution equation can be written :$\forall(\ell_1, \ell_2)$

$$\sum_{p \leq m} a_p(\ell_2)\ell_1^p = \int \sum_{p \leq m} b_p(\ell_2)(\ell_1 - \varepsilon)^p f(\varepsilon)d\varepsilon = \sum_{p \leq m} \int b_p(\ell_2)\left[\sum_{q=0}^{p} C_p^q \ell_1^q (-\varepsilon)^{p-q}\right] f(\varepsilon)d\varepsilon$$

$$= \sum_{r \leq m} \ell_1^r \left\{\sum_{s=r}^{m} C_s^r \, b_s(\ell_2)\int (-\varepsilon)^{s-r} f(\varepsilon)d\varepsilon\right\}$$

Identifying each coefficient in this triangular, non-degenerate, linear system only requires the knowledge of all moments of ε of order at most m.

Q.E.D.

Proposition 3 shows that when a DICM admits as an associated schedule, a surface such that all sections by an hyperplane $\ell_2 = C^t$ are polynomial of order smaller than m, then one only needs to know the <u>m first moments</u> of the distribution of ε to compute a noisy reward schedule that implements the DICM. This is much less demanding than knowing the whole distribution f(.) of ε, as would be required to compute a reward schedule using e.g. a Fourier transform.

The less exhaustive the information on f(.) required to compute a reward schedule, the more <u>robust</u> the reward schedule to mispecifications in the observation disturbance. Robustness is typically a major concern in contract theory lately, since applications to the real world cannot reasonably assume a complete knowledge of the distribution of noises in the economy. There are then strong motivations to finding simple associated schedules, and thereby simple noisy reward schedules, so that implementation under noisy observation be possible despite a limited knowledge of the distribution f(.). In particular, a noisy reward schedule derived from a linear-in-section (affine in ℓ_1 for each ℓ_2) associated schedule is itself linear-in-section (i.e. affine in ℓ_1') and can always be computed even with no information on f(.) (the mean of ε is known to be zero). The schedule has then some "universal" validity, a point emphasized in the next section. Similarly, a noisy reward schedule derived from a quadratic-in-section (quadratic in ℓ_1 for each ℓ_2) associated schedule is itself quadratic in section (in ℓ_1') and can be computed when only the mean and the variance of ε are known. These cases, which have weak and thereby appealing informational requirements, are studied in the next section.

IV. NECESSARY AND SUFFICIENT CONDITIONS FOR UNIVERSAL IMPLEMENTATION

IV.A/Ruled Schedules and Universal Implementation

The previous section led us to emphasize the appeal of linear-in-section associated schedules in view of their universal validity. The central question considered in this section is then : under which conditions does there exist such a schedule ?

For simplicity we will assume $n_1 = n_2 = 1$, but now again ε and ℓ may be correlated. We will also assume that the agent's preferences are smooth. Precisely :

(D) U(.) is twice continuously differentiable.

We will also focus on a DICM (ℓ^*, t^*) that satisfies the following monotonicity property :

(MO) $\ell_2^*(.)$ is one-to-one from Θ to $\ell_2^*(\Theta) \subset L_2$.

Assumption (MO) is made partly for ease of presentation : we shall later discuss its precise role in the analysis.

In the present setting, there can be only one candidate for being a linear-in-section associated schedule. We call it the "ruled schedule" and define it as follows.

Definition 4 : Under (D), (MO) and (B), the "<u>M-ruled schedule</u>" is defined for $M \geqslant M^*$ as the surface

$$t = \varphi(\ell_1, \ell_2) \equiv t^*(\theta) - \partial_1 U(\ell^*(\theta);\theta)\left[\ell_1 - \ell_1^*(\theta)\right] \quad \text{if } \ell_2 = \ell_2^*(\theta) \text{ for some } \theta \in \Theta$$
$$\equiv -M \quad \text{otherwise}$$

The ruled schedule is built as follows. For every ℓ_2 that can be attained by the DICM, i.e. that is a contractual action for some type θ (unique by (MO)), consider the intersection between the tangent hyperplane to this type θ's indifference surface passing through $(\ell^*(\theta), t^*(\theta))$, and the hyperplane of coordinate ℓ_2. The intersection defines a line indexed by ℓ_2. As ℓ_2 varies in $\ell_2^*(\Theta)$, these lines generate a ruled surface. This ruled surface is then extended using a penalty as described in the discussion of (B). This construction is visualized on Figure 3.

From Proposition 3, given that $\mathbb{E}[\varepsilon \mid \ell] = 0$ independently of ℓ, we know that <u>if this ruled surface is an associated schedule</u>, then there exists a noisy reward schedule derived from (1) which is linear-in-section and in fact coincides with the ruled surface. Then, the ruled schedule can be viewed as a noisy reward schedule as well as an associated schedule. As a noisy reward schedule, the ruled schedule is universal, i.e. <u>it implements the DICM for any unbiased noise of observation</u>. (This latter property would hold even if ε were correlated to θ provided it is unbiased.) The next proposition shows that the ruled schedule is the only possible schedule leading to universal implementation.

Proposition 4 : Under (B), (D) and (MO), the ruled schedule is the only possible universal noisy reward schedule.

Proof : For any density function f, a universal schedule ψ must satisfy

(2) $$\forall \theta \in \Theta, \int \psi(\ell_1^*(\theta) - \varepsilon,\ \ell_2^*(\theta)) f(\varepsilon \mid \ell^*(\theta)) d\varepsilon = t^*(\theta)$$

By choosing a sequence of density functions that converge uniformly to the Dirac distribution at $\varepsilon = 0$, the previous equality implies :

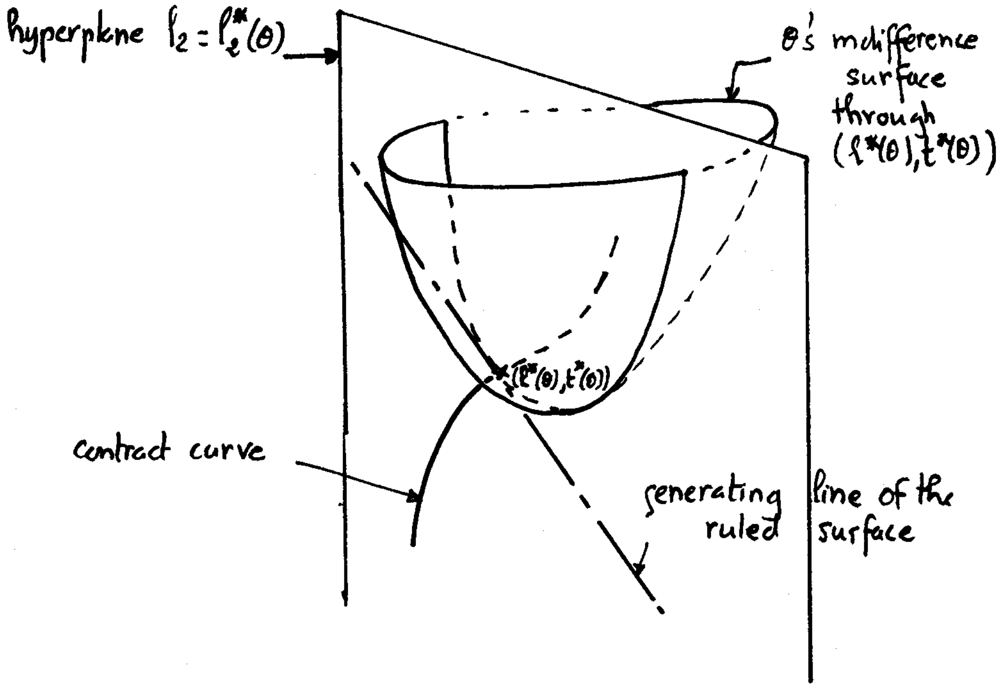


FIGURE 3 : THE RULED SURFACE

(3) $\forall\theta\in\Theta,\ \psi(\ell_1^*(\theta),\ell_2^*(\theta)) = t^*(\theta)$

Next, it is easy to show that if ψ is universal, then it must be affine in ℓ_1', i.e. for any $(\varepsilon_1,\varepsilon_2)\in\mathbb{R}^2_{+*}$, the points $\left\{\ell_1^*(\theta),\psi(\ell_1^*(\theta),\ell_2^*(\theta))\right\}$, $\left\{\ell_1^*(\theta)-\varepsilon_1,\psi(\ell_1^*(\theta)-\varepsilon_1,\ell_2^*(\theta))\right\}$ and $\left\{\ell_1^*(\theta)+\varepsilon_2,\psi(\ell_1^*(\theta)+\varepsilon_2,\ell_2^*(\theta))\right\}$ are on the same line. If this were not the case, then one could find a probability distribution putting almost all the weight on neighborhoods of ε_1 and ε_2, of zero mean, and such that (2) is wrong.

So ψ is universal only if : $\psi(\ell_1',\ell_2^*(\theta)) = t^*(\theta)+k(\theta)(\ell_1'- \ell_1^*(\theta))$.

Finally the condition that a θ-agent chooses $\ell^*(\theta)$ (i.e. that ψ be a noisy reward schedule) implies that $k(\theta) = -\partial_1 U(\ell^*(\theta);\theta)$.

Conversely, the ruled surface is obviously universal if it is an associated schedule.

Q.E.D.

IV.B/Sufficient Conditions for Universal Implementation

We can now come to the heart of the analysis : For a given DICM when is the ruled schedule an associated schedule ?

We first give sufficient conditions for having this property.

Proposition 5 : Assume that (B), (D), (MO) hold. If one of the following set of properties holds, then the ruled schedule is a (ℓ^*,t^*) -associated schedule (and hence is a universally implementable noisy reward schedule) :

either : i1) Preferences are independent of ℓ_2.

i2) The DICM (ℓ^*,t^*) is continuously differentiable.

i3) $\exists$ an associated schedule $\bar{\varphi}$ independent of ℓ_2 and convex in ℓ_1 on L_1.

or : ii1)The ruled surface is a connected portion of an hyperplane.

ii2)Preferences of the agents are convex.

Proof : As the agent's utility does not depend on ℓ_2, we simply write it as $U(\ell_1,\theta)$. Let us define

$$D(\theta) = \left\{(\ell_1,t)\in\mathbb{R}^2 \mid t = t^*(\theta) - \partial_1 U(\ell_1^*(\theta);\theta)\left[\ell_1 - \ell_1^*(\theta)\right]\right\}$$

$$P(\theta) = \left\{(\ell_1,t)\in\mathbb{R}^2 \mid t + U(\ell_1;\theta) > t^*(\theta) + U(\ell_1^*(\theta);\theta)\right\}$$

$P(\theta)$ is the set of points (ℓ_1,t) that are strictly preferred to $(\ell_1^*(\theta),t^*(\theta))$ by a θ-agent. $D(\theta)$ is the projection in the plane (ℓ_1,t) of the generating line of the ruled surface passing through $(\ell^*(\theta),t^*(\theta))$. Since U is independent of ℓ_2, the agent's indifference surfaces are cylinders parallel to the ℓ_2-axis ; then, the ruled schedule is an associated schedule if and only if : $D \cap P = \varnothing$, where $D = \bigcup_{\theta\in\Theta} D(\theta)$ and $P = \bigcup_{\theta\in\Theta} P(\theta)$. (See Figure 4).

Consider the associated schedule $\bar{\varphi}$, and let $\bar{Z}$ be its epigraph on L_1, (i.e. $\left\{(\ell_1,t)\in L_1\times\mathbb{R} \mid t \geqslant \bar{\varphi}(\ell_1)\right\}$). Using the definition of an associated schedule, it can be shown by contradiction that $P(\theta)\subset \mathrm{int}\bar{Z}$.

Moreover from (D) and the regularity of the DICM, $D(\theta)$ is the tangent line to the lower boundary of $\bar{Z}$ at point $(\ell_1^*(\theta),t^*(\theta))$. Then from the convexity of $\bar{Z}$, $D(\theta) \cap \mathrm{Int}\,\bar{Z}$ is empty. It follows that for any θ $D(\theta) \cap P = \varnothing$ and then $D \cap P = \varnothing$.

ii) If an agent chooses $(\ell^*(\theta),t^*(\theta))$, where $\ell_2^*(\theta)\in \mathrm{Int}\,(\ell_2^*(\Theta))$, the corresponding indifference surface is tangent to the hyperplane. As $P(\theta)$ is convex for all θ, it never intersects the hyperplane.

In the case where $(\ell^*(\theta),t^*(\theta))$ is such that $\ell_2^*(\theta)$ is on the boundary of $\ell_2^*(\Theta)$, the conclusion remains if M is large enough.

Q.E.D.

In 5i), the requirement that preferences be independent of ℓ_2 is restrictive, but the assumption that $\bar{\varphi}$ does not depend on ℓ_2 is innocuous. One could take the projection of the contract curve on an hyperplane of

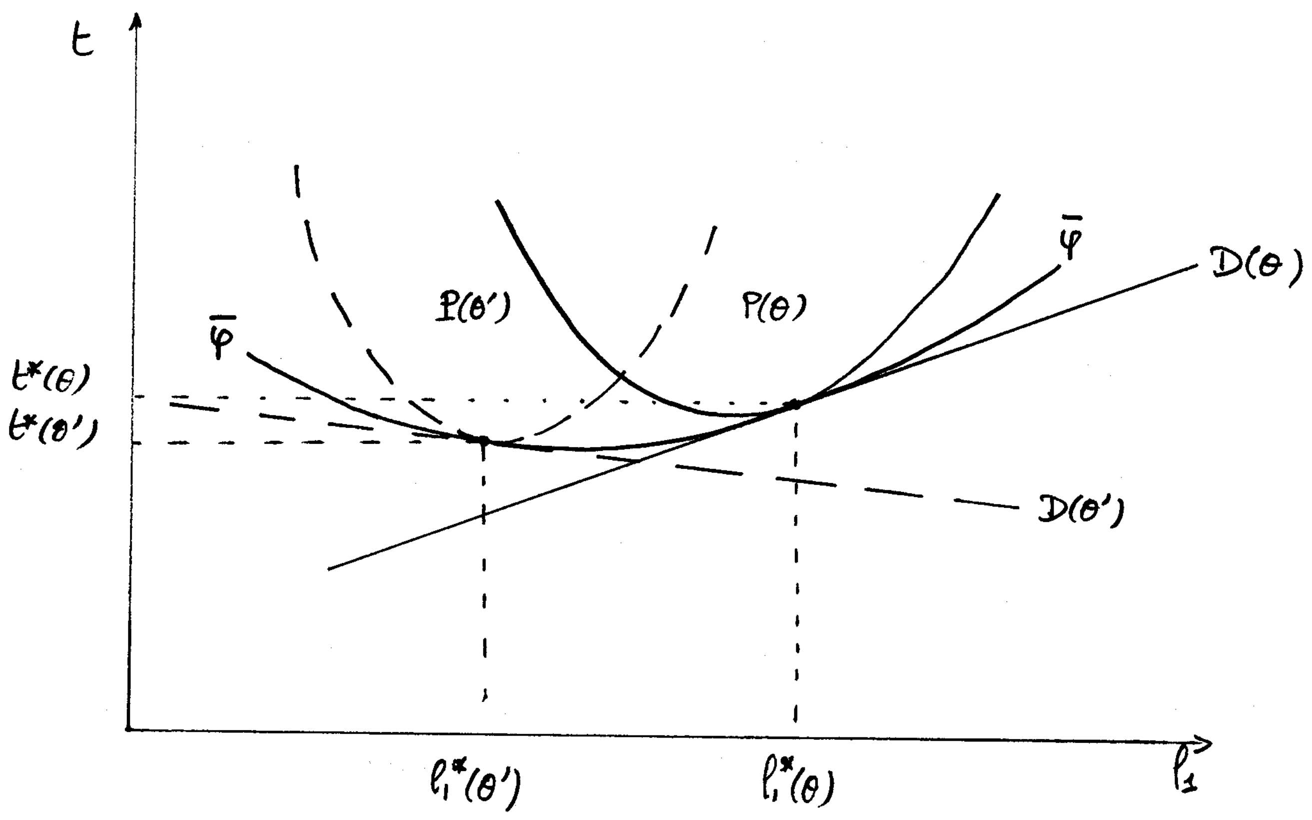


FIGURE 4 : PROPOSITION 5i)

coordinate $\ell_2 = \overset{t}{C^{-}}$, and construct a cylinder with identical section and orthogonal to that hyperplane ; this cylinder would be an associated schedule independent of ℓ_2. A crucial assumption however, is that $\bar{\varphi}$ be convex (in ℓ_1) . Given that any generating line of the ruled surface, tangent to the section of θ's indifference surface in the hyperplane $\ell_2 = \overset{t}{C^{=}}$, is also tangent to $\bar{\varphi}$ in $(\ell_1^*(\theta), t^*(\theta))$, the convexity of $\bar{\varphi}$ guarantees that no type θ' has a preferred point below (or on) the generating line, therefore below (or on) the ruled surface. The independence of preferences on ℓ_2 makes our model equivalent to Picard [1986] 's unidimensional problem of implementation via a family of linear schedules, and our condition of convexity is similar to his condition. It should be noted that the case is more general than it might appear. Indeed if $U(\ell;\theta) = U_1(\ell_1, \ell_2) + U_2(\ell_1;\theta)$, a similar argument would apply : what matters is that the cross derivative of U with respect to ℓ_2 and θ be zero, for in this case, the problem is essentially a one-dimensional incentive problem with action ℓ_1 and transfers $t + U_1(\ell_1, \ell_2)$.

In Laffont-Tirole's [1986] model of optimal regulation, the manager of the regulated firm has objectives : $t - H(e)$, where e is an effort that can reduce the cost of production of the good : $C = (\theta - e)q + \nu$, ν being an independent noise. Since q and C are observable, the model can be written :

$$\ell_1 = \theta - e, \quad \ell_2 = q \ , \quad \varepsilon = \frac{\nu}{q} \text{ unbiased}$$

$$U(\ell_1, \ \ell_2 \ ; \ \theta) = - H(\theta - \ell_1) \text{ independent of } \ell_2$$

For a continuously differentiable mechanism, Laffont and Tirole show that $\frac{d\ell_1}{d\theta} \geqslant 1$ $\left(\frac{de}{d\theta} \leqslant 0 \text{ in their proposition 4, p 622}\right)$ which is equivalent to 5i3). Hence Proposition 5i) applies, because the problem is actually one-dimensional, and there exists a convex associated schedule.

Proposition 5ii) shows that when preferences depend on ℓ_2, only strong conditions on the ruled schedule, i.e. the fact that it is an hyperplane, can ensure that it is an associated schedule. This can be understood as follows : the ruled schedule allows choices far away from the contract curve, so that a θ-agent may choose $\ell^*(\theta)$ as a local maximizer of utility but may prefer to jump to a distant point. This possibility will occur if the generating line does not rotate "properly" with θ along the contract curve i.e. at a pace similar to the one at which the indifference surfaces rotate with θ. In the case we have identified in 5ii), the contract curve is included in a plane and a "proper" rotation is no rotation at all (when preferences are convex).

IV.C/Necessary Condition for Universal Implementation

Thus we have identified a small subset of problems for which universal implementation holds. We now characterize a broader class of problems where this property might hold. These problems are those for which a necessary condition for implementation is satisfied. The intuition behind this necessary condition is first explained.

Consider the hyperplane $\ell_2 = \ell_2^*(\theta)$ and the sections in this hyperplane of all θ'-indifference surfaces going through $(\ell^*(\theta'),\ t^*(\theta'))$, for θ' close to θ. Three of these sections are depicted on Figure 5, one for θ and two for θ' and θ'' close to θ. Note that these sections are (implicitly) defined by equation (4) :

(4) $$t \equiv T(\ell_1,\theta') \equiv t^*(\theta') + U(\ell^*(\theta');\theta') - U(\ell_1,\ell_2^*(\theta)\ ;\ \theta')$$

The diagram suggests that the family of curves so defined has a lower envelope. The necessary condition which we will stress is that the lower envelope be locally above the section of the ruled schedule, i.e. above the tangent in $(\ell_1^*(\theta), t^*(\theta))$ to the section of θ's indifference surface in $(\ell^*(\theta), t^*(\theta))$. The reason why this condition is needed for universal implementation is immediate from Figure 5 . The necessary condition will express the requirement that the epigraph of the lower envelope be convex i.e. that the second derivative of the function defining the envelope be positive.

$$\begin{cases} \theta' = \hat{\theta}(\rho_1') \\ \theta'' = \hat{\theta}(\rho_1'') \end{cases}$$

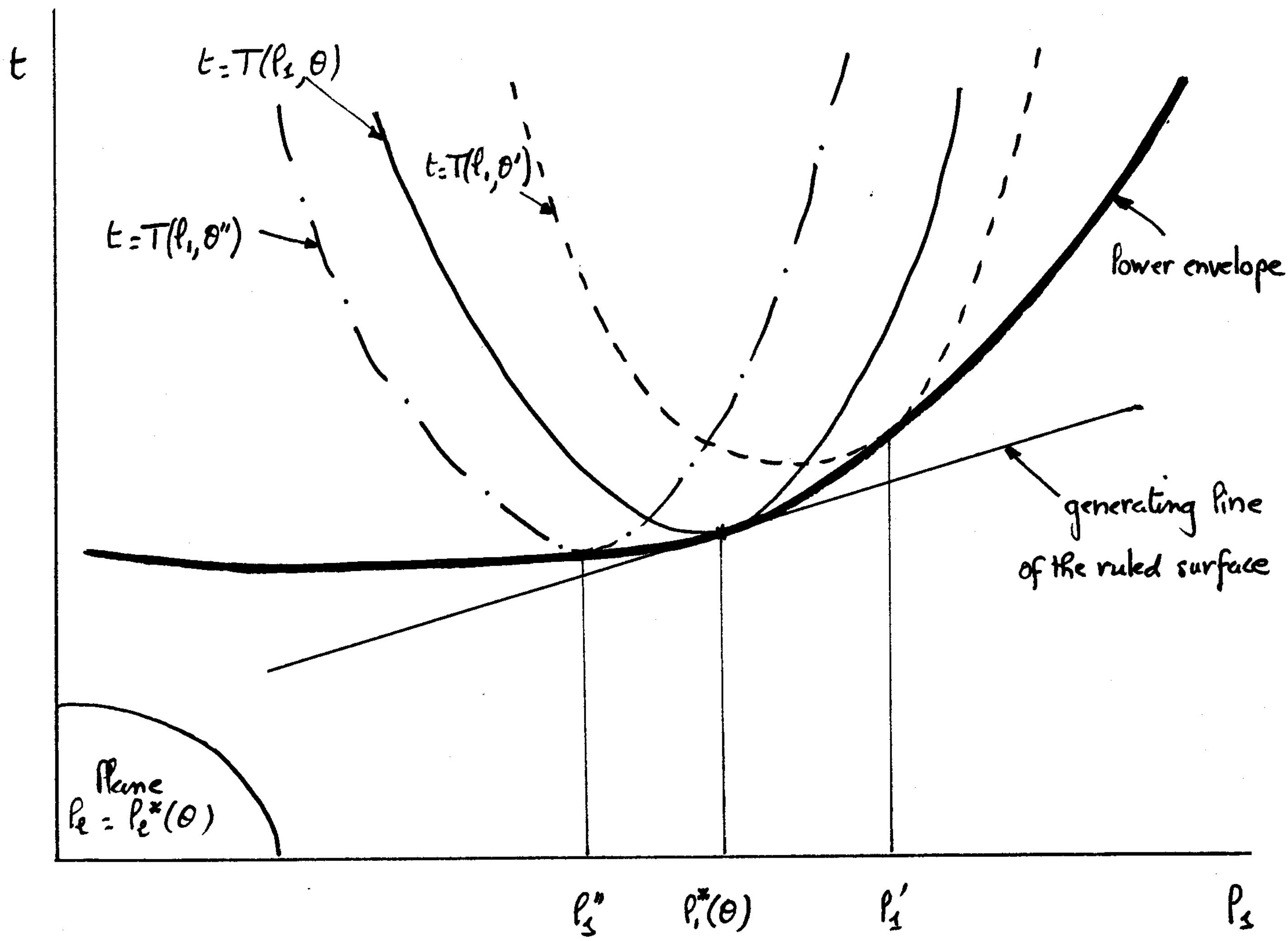


FIGURE 5 : THE ENVELOPE CURVE

Assuming enough regularity (and simply denoting ℓ_1 by ℓ), we know that the equation of the lower envelope is $t = T(\ell,\hat{\theta}(\ell))$, where $\hat{\theta}(\ell)$ is implicitly defined by (5) :

(5) $$\partial_{\theta'} T(\ell,\hat{\theta}(\ell)) = 0$$

Using this fact, we can now compute the second derivative of the equation of the lower envelope, which we denote $\frac{d^2 t}{d\ell^2}$. With straightforward notation, we have :

$$\frac{d^2 t}{d\ell^2} = \frac{d}{d\ell}\left[\partial_\ell T + \partial_\theta T \frac{d\hat{\theta}}{d\ell}\right]_{(\ell,\hat{\theta}(\ell))} = \frac{d}{d\ell}[\partial_\ell T]_{(\ell,\hat{\theta}(\ell))}$$

$$= \partial_{\ell\ell} T(\ell,\hat{\theta}(\ell)) + \partial_{\ell\theta'} T(\ell,\hat{\theta}(\ell)) \frac{d\hat{\theta}}{d\ell}$$

Moreover, differentiating (5) yields :

$$\partial_{\theta',\ell} T(\ell,\hat{\theta}(\ell)) + \partial_{\theta',\theta'} T(\ell,\hat{\theta}(\ell)) \frac{d\hat{\theta}}{d\ell} = 0$$

So that :

(6) $$\frac{dt^2}{d\ell^2} = \partial_{\ell\ell} T(\ell,\hat{\theta}(\ell)) - \frac{\left[\partial_{\ell\theta'} T(\ell,\hat{\theta}(\ell))\right]^2}{\partial_{\theta',\theta'} T(\ell,\hat{\theta}(\ell))}$$

an expression that only depends on the derivatives of T and therefore of U. By differentiating repeatedly (4) and using the incentive compatibility constraint on the contract curve, one finds :

$$\frac{d^2 t}{d\ell^2} = - \partial_{11} U(\ell^*(\theta);\theta) - \frac{\left[\partial_{1\theta} U(\ell^*(\theta);\theta)\right]^2}{\Gamma(\theta)}$$

where : $\Gamma(\theta) = \partial_{1\theta} U(\ell^*(\theta);\theta) \frac{d\ell_1^*}{d\theta}(\theta) + \partial_{2\theta} U(\ell^*(\theta);\theta) \frac{d\ell_2^*}{d\theta}(\theta)$ and recalling that incentive compatibility implies that $\Gamma(\theta)$ is positive (See Guesnerie-Laffont [1984], Theorem 1, p 336).

The proof sketched above is made completely rigorous in the Appendix, and justifies our main theorem :

Theorem : Assume that (B), (D) and (MO) hold. Consider a DICM (ℓ^*, t^*) which is continuously differentiable and a.e. twice differentiable. The ruled schedule is an associated schedule (allowing universal implementation) <u>only</u> if the following condition holds :

$$\forall\ \theta\in\Theta,\ \left[\partial_{1\theta}U(\ell^*(\theta),\theta)\right]^2 \leqslant -\ \partial_{11}U(\ell^*(\theta),\theta)\ \Gamma(\theta) \tag{7}$$

where $\Gamma(\theta) = \partial_{1\theta}U(\ell^*(\theta),\theta)\ \dfrac{d\ell_1^*}{d\theta} + \partial_{2\theta}U(\ell^*(\theta),\theta)\ \dfrac{d\ell_2^*}{d\theta}$.

Before going further, we should convince ourselves that the above formula fits the intuition of the phenomenon, as developed throughout the proof. For that let us consider the three terms in (7) and check that they indeed play in the direction one should expect.

First the LHS of (7) is non negative, and if (ℓ^*,t^*) is a DICM, then necessarily $\Gamma(\theta)$ is non negative. So the term $\left[-\partial_{11}U(\ell^*(\theta);\theta)\right]$ must be non negative for the ruled schedule to be a (ℓ^*,t^*)-associated schedule. Indeed, if the indifference surface of a θ-agent had a section in the hyperplane $\ell_2 = \ell_2^*(\theta)$ which were not locally concave around the contractual point $(\ell^*(\theta),t^*(\theta))$, then the ruled surface would not be an associated schedule since it would require the agent to choose an action ℓ_1 that minimizes (locally) his utility. Now, given that $\left[-\partial_{11}U(\ell^*(\theta);\theta)\right]$ must be positive, it increases with the curvature of the indifference surface of the agent in ℓ_1. Everything else being given, the higher the curvature of the family of curves, the better the chance that the lower envelope of this family be itself convex. When the term $\left[-\partial_{11}U(\ell^*(\theta');\theta')\right]$ increases, a θ'-agent will require a higher compensation to deviate to the point $\ell^*(\theta)$ for θ' close to θ, and, for highly convex preferences in ℓ_1 , local implementation by the ruled schedule will be possible.

Second, the term $\Gamma(\theta)$ has already been shown to be non-negative. Consider now the transfer needed to make a θ'-agent choose $(\ell^*(\theta),t^*(\theta))$ instead of his contractual point : it is given by

$$\{t^*(\theta') + U(\ell^*(\theta')\ ;\ \theta')\} - \{t^*(\theta) + U(\ell^*(\theta);\theta')\}$$

For θ' close to θ, this expression is equal to $\Gamma(\theta)(d\theta)^2$ to the second order. Thus $\Gamma(.)$ "measures" the distance in terms of transfers between the point chosen by θ-agents and the indifference curve of neighbor agents.

Increasing this distance is a favorable factor for the convexity of the epigraph of the lower envelope of all these curves, i.e. for the ruled schedule not to allow this amount of transfers that would trigger local deviations. This phenomenon conforms to expression (7) in the theorem.

Finally $|\partial_{1\theta}U(\ell^*(\theta);\theta)|$ must be as small as possible for the ruled schedule to be an associated schedule. The intuition here can be understood as follows. $\partial_1 U(.)$ measures the slope of the ruled surface or/and of the indifference surface of θ-agents in the hyperplane $\ell_2^*(\theta)$, so that $|\partial_{1\theta}U|$ measures the speed of rotation of the generating lines of the ruled schedule when θ varies. Given the (local) convexity of θ's indifference surface, it is clear that a high speed of rotation will imply that close to $(\ell^*(\theta),t^*(\theta))$, the generating lines of the ruled surface will intersect θ's indifference surface, thereby ruling out the ruled surface as an associated schedule.

To obtain a better understanding of the theorem, it is useful to specialize it to the case where preferences are independent of ℓ_2. We know from Proposition (5i) that if the DICM is continuously differentiable, the convexity in ℓ_1 of an associated schedule is a sufficient condition for the ruled schedule to implement the DICM as an associated schedule. The next corollary shows that it is equivalent to condition (7) so that (7) (or the convexity of an associated schedule) is a necessary and sufficient condition for universal implementation.

<u>Corollary 1</u>. Under the condition of the theorem and if preferences are independent of ℓ_2, condition (7) is a necessary and sufficient condition for universal implementation (by the ruled surface).

<u>Proof</u> : (7) is a necessary condition for universal implementation by the previous theorem. Next we prove (7) implies the existence of an associated schedule satisfying 5i3) (convex in ℓ_1, independent of ℓ_2).

For that consider the projection of the contract curve in the plane (ℓ_1,t) : it is a C^1-manifold. Moreover as $\ell_1^*(\theta_1)=\ell_1^*(\theta_2) \Longrightarrow t^*(\theta_1)=t^*(\theta_2)$, it can be written as a function $t = \mathfrak{T}(\ell_1)$. From incentive compatibility $\frac{dt^*}{d\theta} + \partial_1 U\left(\ell_1^*(\theta);\theta\right) \frac{d\ell_1^*}{d\theta} = 0$. First, since $\partial_1 U$ is bounded (on the compact $L_1 \times \Theta$), $\mathfrak{T}$ is C^1. Second, for almost all ℓ_1, the local inverse $\theta^*(\ell_1)$ of $\ell_1^*(\theta)$ is defined, and $\mathfrak{T}'(\ell_1) = -\partial_1 U\left(\ell_1, \theta^*(\ell_1)\right)$ and $\mathfrak{T}''(\ell_1) = \left[-(\partial_{11} U(\ell_1, \theta^*(\ell_1)) \frac{d\ell_1^*}{d\theta} - \partial_{1\theta} U(\ell_1, \theta^*(\ell_1))\right] / \frac{d\ell_1^*}{d\theta}$ exists. Condition (7) is equivalent to $\mathfrak{T}''(\ell_1) \geqslant 0$ for almost every ℓ_1, given that $\partial_{1\theta} U(\ell_1, \theta^*(\ell_1)) \frac{d\ell_1^*}{d\theta} \geqslant 0$ by incentive compatibility.

Now $\mathfrak{T}$ being C^1 and with a positive second derivative almost every where is globally convex.

It is now easy to take the cylinder generated by the section $t = \mathfrak{T}(\ell_1)$, i.e. the cylinder embedding the contract curve. It obviously is an associated schedule, and we showed that it is convex. Hence it satisfies 5i3).

Now applying proposition 5i), we know that the ruled surface is an associated schedule, hence universal implementation holds. So (7) is also sufficient.

Q.E.D.

The necessary requirement for universal implementation in the case where preferences are independent of ℓ_2 coincides with the sufficient condition of Proposition (5i) and can be satisfied only by special DICM (those with associated schedules which are convex in the sense of Proposition 5). This stringent restriction relates to the special form of preferences under consideration. (In fact, for any a priori given DICM, one can find preferences for which the DICM can be universally implemented.)

Starting from a situation depicted by Corollary 1, we can now turn back to the general case. We are going to show that a "reasonable" dependence on ℓ_2 is a favorable factor for the possibility of universal implementation.

Consider the family of agent's preferences, indexed by $\alpha \in \mathbb{R}^+$:

(8) $$U_\alpha(\ell_1, \ell_2; \theta) = V(\ell_1; \theta) + \alpha\, W(\ell_2; \theta)$$

where V and W are fixed functions such that

(P) $$\partial_{\ell_1 \theta} V > 0 \quad , \quad \partial_{\ell_2 \theta} W > 0$$

In this setting it is known (cf Guesnerie-Laffont [1984]) that the piecewise continuously differentiable functions (ℓ_1, ℓ_2) which satisfy

$$\frac{d\ell_1}{d\theta} > 0 \;,\; \frac{d\ell_2}{d\theta} > 0$$

are implementable, for any positive α . Let us focus attention on this set $\mathcal{L}$ of implementable functions which are common to all preference parameters α.

Condition (7) takes the following form

(7') $$\left[\partial_{\ell_1 \theta} V\right]^2 \leqslant \left[-\partial_{\ell_1 \ell_1} V\right] \left[\left(\partial_{\ell_1 \theta} V\right) \frac{d\ell_1}{d\theta} + \alpha \left(\partial_{\ell_2 \theta} W\right) \frac{d\ell_2}{d\theta}\right]$$

It is easy to check that if for a given (ℓ_1, ℓ_2), condition (7') holds for α, it also holds for any $\alpha' > \alpha$. Thus :

<u>Corollary 2</u> : Consider the class of functions indexed by α defined by (8) and assume in addition that they satisfy (B) (D) and (P).Then the set $\mathcal{L}_\alpha = \{(\ell_1, \ell_2) \in \mathcal{L} \mid \text{condition(7') holds for } \alpha\}$ is increasing in α.

Then Corollary 2 provides a precise and striking illustration of the fact that the dependence of preferences on ℓ_2 is a favorable factor for universal implementation [7] . Loosely speaking, among preferences which rationalize a given DICM, the more likely to pass the necessary test for universal implementation are those which depend strongly (and correctly) on ℓ_2 . Our test which is very demanding for preferences independent of ℓ_2 can be less stringent when preferences depend upon ℓ_2. In counterpart, the criterion which we have stated and which is (almost) sufficient for preferences independent of ℓ_2 is no longer sufficient when preferences depend upon ℓ_2.

Let us summarize our findings. For a given DICM, condition (7) provides a description, which we have shown to be intuitively plausible, of the border between preferences for which universal implementation is possible and those for which it cannot work. In some sense, the addition of the dimension ℓ_2 to our problem, when it is effective, increases the power of the ruled schedule for universal implementation.

As final comments, it should be noted first that the condition we have exhibited, although being only necessary for universal implementation, is necessary and sufficient when one considers the implementation problem under small noises of observation (small variance and small support). In CGR, we developed a theory based on the consideration of truncated ruled schedules for problems with small disturbances.

Second, the fact that ℓ_2 is one-dimensional does not seem crucial to our analysis. The important point is the equality between the dimension of the space of observable variables ℓ_2 and of the space of characteristics.

Finally, the assumption that ℓ_2^* is one-to-one has an ambiguous effect on the result of our analysis. The local considerations leading to our necessary condition hold independently of this assumption, but in the absence of such a condition of strong monotonicity, universal implementation is likely to be hopeless. The condition remains relevant however, for the analysis of implementation in the presence of small noises.

IV.D/Implementation via Quadratic-in-Section Schedules.

Implementation via noisy reward schedules which are quadratic-in-section is less appealing, although easier, than universal implementation. The design of a quadratic-in-section schedule requires the knowledge of the variance of the noise ; may be more importantly, it has the inconvenience of leading to a rather high variance of the agents remuneration, a fact which the recent theory of contracts has shown undesirable for self enforcement of contracts. To limit this inconvenience, it is important to characterize the smallest possible curvature of the quadratic schedules. Our analysis provides a lower bound to the curvature of such schedules.

Proposition 6 : Assume (B), (D) and (MO) hold, and that (ℓ^*, t^*) is C^1 a.e. twice differentiable DICM. A necessary condition for a M-extended, quadratic-in-section schedule to be an associated schedule, is that the curvature of its section in the hyperplane $\ell_2^*(\theta)$, be at least larger than : $\frac{1}{2}\left\{\frac{\left[\partial_{1\theta}U(\ell^*(\theta);\theta)\right]^2}{\Gamma(\theta)} + \partial_{11}U(\ell^*(\theta);\theta)\right\}$

The expression is reminiscent of condition (7) and the proof of Proposition 6 is a by-product of the proof of the Theorem. The differentiability assumptions guarantee that the lower envelope presented in the previous subsection is smooth around $(\ell^*(\theta);t^*(\theta))$, so that its curvature is bounded from below. Any parabola passing through $(\ell^*(\theta),t^*(\theta))$, tangent to the envelope at this point, and with high enough curvature will lie below this envelope curve : the surface generated by all these parabolas fulfills all local incentive compatibility constraints. Proposition 6 gives a precise bound for the minimal curvature.

Proposition 6 has two consequences. Since a formal statement would be somewhat heavy, we only provide the intuition of the straightforward arguments underlying them.

First, for small noises, there always exists a quadratic-in-section associated schedule : any one generated by a family of parabolas which satisfy the curvature condition of the proposition (See CGR).

Second, for general noises, Proposition 6 only gives the right local curvature, but global incentive compatibility constraints (i.e. between θ and θ' not close) might not hold. However, the reader will convince himself that "generically" one can increase the curvature of the parabola beyond the lower bound exhibited above, up to a point where the corresponding quadratic-in-section schedule is globally incentive compatible and therefore an associated schedule (compactness, uniform continuity).

This remark as well as the above proposition generalizes the results obtained by Picard [1987] for one-dimensional action variable, principal-agent problems.

V. CONCLUSION

The earlier literature on contracts has been faced with the contrast between the complexity of theoretically optimal schemes and the rough simplicity of many real world arrangements (See Hart-Holmstrom [1987]). The later literature has provided a number of explanations which mitigate the discrepancy between theory and practice. In particular Holmstrom-Milgrom [1987] have shown that simple linear schemes may be optimal in pure moral hazard problems where the action variable can be continuously corrected to respond to the accrual of information. The present paper (as well as the previous literature on which it relies) can be viewed as developing a different argument for the usefulness of linear schemes. The argument is that linear schemes provide a robust implementation of an optimum in a context of multidimensional adverse selection problems where "noise" significantly affects a subset of the contractual variables. It is hoped that the analysis has informed the reader both of the significance and of the limits of this argument.

FOOTNOTES

1. For this, consider any (ℓ,t) - associated schedule φ, and take two compact sets K_1 and K_2, such that the interior of K_1 contains K_2 and $K_2 \supset \ell(\Theta)$. Consider then the function $\tilde{\varphi}$ defined by :

$$\tilde{\varphi}(\ell) = \varphi(\ell),\ \forall\ \ell \in K_2$$
$$\tilde{\varphi}(\ell) = -M,\ \forall\ \ell \in \mathbb{R}^n/K_1$$

and by piecing together these parts on K_1/K_2. It is obvious that $\tilde{\varphi}$ is an (ℓ,t) - associated schedule, and it can be constructed to be as regular as desired on $\mathbb{R}^n/K_2$.

2. Incentive compatibility on the DICM implies :

$$\ell(\theta) = \ell(\theta') \implies t(\theta) = t(\theta')$$

3. Given that the signal has the same dimension as ℓ_1 and is unbiased, the additive form involves no loss of generality.

4. $\mathbb{E}$ denotes the expectation operator with respect to the distribution of ε.

5. When there is no perfectly observed action ℓ_2, however, the consideration of reward schedules $t = \varphi(\ell_1')$ rather than general revelation mechanisms $t = H(\ell_1',\theta')$ may entail a restriction on the set of instruments of the principal ; the resulting loss in welfare is called the value of communication in MR.

6. Another desirable property would be that the variance of transfers actually paid to the agent be limited, or that transfers be stochastically bounded. Unlimited transfers (as in a Mirrlees Scheme) or highly variable

transfers may limit the enforceability of contracts, and will cease to be optimal when a small amount of risk aversion is introduced, so that it is reasonable to try to implement an allocation without using them. This requirement also leads us to favor linear rather than quadratic schedules, as witnessed in section IV.

7. Some technical comments are in order : first condition (P) and the monotonicity of the DICM could be generalized to a constant sign assumption (see Guesnerie-Laffont [1984]). Second, the set $\mathcal{L}$ is only a subset of all implementable functions. It is the only one, in the present state of knowledge, for which one can be sure that it belongs to the intersection over α, of the sets of implementable functions.

APPENDIX

Consider the quadratic-in-section surface defined by :

$$\text{(A1)}\quad \begin{cases} \ell_1 = u \\ \ell_2 = \ell_2^*(v) \\ t = t^*(v)-\partial_1 U(\ell^*(v);v)\left[u-\ell_1^*(v)\right]-q(v)\left[u-\ell_1^*(v)\right]^2 \end{cases}$$

which is parametrized by $(u,v) \in \ell_1^*(\Theta) \times \Theta$ and quadratic in u. Let us consider some agent θ in Θ and call $F_{q,\theta}(u,v)$ the value of his utility over the surface : $F_{q,\theta}(u,v) = t+U(\ell_1,\ell_2;\theta)$, where t, ℓ_1 and ℓ_2 are derived from u and v by (A1).

It is first obvious that $(\ell_1^*(\theta),\theta)$ is a stationary point of $F_{q,\theta}$. Computing the Hessian quadratic form of the function F and using the incentive compatibility constraints fulfilled on the contract curve, one gets :

$$\partial^2_{u\,u}F_{q,\theta}(\ell_1^*(\theta);\theta) = \partial^2_{11}U(\ell^*(\theta),\theta) - 2q(\theta),$$

$$\partial^2_{u\,v}F_{q,\theta}(\ell_1^*(\theta),\theta) = -\left[\partial^2_{11}U(\ell^*(\theta),\theta) - 2q(\theta)\right]\frac{d\ell_1^*}{d\theta}(\theta) - \partial^2_{1\theta}U(\ell^*(\theta),\theta)$$

$$\partial^2_{v\,v}F_{q,\theta}(\ell_1^*(\theta),\theta) = \left[\partial^2_{11}U(\ell^*(\theta),\theta) - 2q(\theta)\right]\left[\frac{d\ell_1^*}{d\theta}\right]^2(\theta) + \frac{d\ell_1^*}{d\theta}(\theta)\ \partial^2_{1\theta}U(\ell^*(\theta),\theta) - \frac{d\ell_2^*}{d\theta}(\theta)\ \partial^2_{2\theta}U(\ell^*(\theta),\theta)$$

A necessary condition for the surface defined by (A1) to implement the DICM $(\ell^*(.),\ t^*(.)))$ is that the Hessian quadratic form be negative semi definite :

$$\partial^2_{11}U(\ell^*(\theta),\theta) - 2q(\theta) \leqslant 0 \tag{A2}$$

$$-\left[\partial^2_{11}U(\ell^*(\theta),\theta) - 2q(\theta)\right]\left[\frac{d\ell_1^*}{d\theta}(\theta)\ \partial^2_{1\theta}U(\ell^*(\theta),\theta) + \frac{d\ell_2^*}{d\theta}(\theta)\ \partial^2_{2\theta}U(\ell^*(\theta),\theta)\right] \geqslant \left[\partial_{1\theta}U(\ell^*(\theta),\theta)\right]^2 \tag{A3}$$

From second-order incentive compatibility condition of the DICM : $\forall\theta\in\Theta$, $\frac{d\ell_1^*}{d\theta}(\theta)\ \partial^2_{1\theta}U(\ell^*(\theta),\theta) + \frac{d\ell_2^*}{d\theta}(\theta)\ \partial^2_{2\theta}U(\ell^*(\theta),\theta) \geqslant 0$, so that (A3) implies (A2).

Take $q(.) \equiv 0$. (A3) is a necessary condition for the ruled schedule to implement the DICM, and is identical to (7). Hence the theorem.

Take $q(.)$ non null, (A3) is a necessary condition for the quadratic-in-section schedule to be an associated schedule, and it gives a lower bound for the value of the curvature $q(.)$. Hence Proposition 6.

REFERENCES

CAILLAUD B., GUESNERIE R.and P. REY [1988],
"Noisy Observation in Adverse Selection Models", Mimeo INSEE 8802.

GUESNERIE R. [1981],
"On taxation and Incentives : Further Reflections on the Limit to Redistribution", to appear in Contribution to the theory of Taxation, ch. I.

GUESNERIE R. and J.J. LAFFONT [1984],
"A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm", Journal of Public Economics, vol. 25, n° 3, pp. 329-370.

HAMMOND P. [1979],
"Straightforward Individual Incentive Compatibility in Large Economies", Review of Economic Studies, Symposium, pp 263-282.

HART O. and B. HOLMTROM [1987],
"The Theory of Contracts", in Advance in Economic Theory, Fifth World Congress, ed. by T. Bewley, Cambridge U.P.

HOLMSTROM B. and P. MILGROM [1987],
"Aggregation and Linearity in the Provision of Intertemporal Incentives", Econometrica, vol 55, pp 303-328.

LAFFONT J.J. and J. TIROLE [1986],
"Using Cost Observation to Control a Public Firm", Journal of Political Economy, vol. 94, pp. 614-641.

MELUMAD N. and S. REICHELSTEIN [1986],
"Value of Communication in Agencies", mimeo GSB Stanford University, forthcoming Journal of Economic Theory.

PICARD P. [1987],
"On the Design of Incentive Contracts Under Moral Hazard and Adverse Selection", Journal of Public Economics, 33.

ROGERSON W. [1988],
"On the Optimality of Menus of Linear Contracts", mimeo Northevestern University.