*Chapter 16*

# A DIFFERENTIAL APPROACH TO EXPECTED UTILITY MAXIMIZING MECHANISMS*

JEAN-JACQUES LAFFONT and ERIC MASKIN

## 1. Introduction

Good collective decision-making often requires discovering personal characteristics, including preferences, known only by private agents. Unfortunately it may not be in the interest of an agent to reveal this information if he knows how it will be used. Consequently, a fundamental problem of collective decision-making is to design procedures which both elicit private information and make good decisions. Formally speaking the problem is that of designing a game form the equilibria of which coincide with the optima of the chosen welfare criterion. There are several alternative equilibrium concepts available, A game form satisfying the above coincidence property may be called incentive compatible with respect to the corresponding equilibrium notion.

The strongest notion is that of dominant strategies. Following Vickrey [10], Groves [5] and Clarke [3] have exhibited classes of dominant strategy game forms ('mechanisms') whose dominance property is due to appropriate transfers of a private good among agents. Unfortunately, as shown in Laffont and Maskin [8], these transfers sum to zero only for exceedingly restricted classes of utility functions. To overcome this lack of balance one must therefore drop the requirement of dominant strategies and appeal to weaker equilibrium notions.

One possibility is Nash-equilibrium, an approach explored in Groves and Ledyard [6], Hurwicz [7], Maskin [9] and elsewhere. In another line of research d'Aspremont and Gérard-Varet [1] and Arrow [2] have applied the expected utility equilibrium, pioneered by Harsanyi, to the design problem. In this paper we extend the differential approach of our above cited paper to characterize and study the family of individually incentive compatible expected utility maximizing mechanisms.

We shall throughout confine our discussion to a model of one private and

*This work was supported by Cordes no. 136–77.

*Aggregation and Revelation of Preferences, edited by Jean-Jacques Laffont*

one public good, which we develop in section 2. In section 3 we characterize dominant strategy mechanisms for arbitrary decision functions as a technical preliminary. Section 4 provides a complete characterization of the expected utility maximizing mechanisms. We also show that only in a very limited family – namely the linear family – of welfare criteria based on valuation functions can one in general construct individually incentive compatible expected utility maximizing mechanisms. In section 5 the mechanisms studied by d'Aspremont and Gérard-Varet [1] and Arrow [2] are shown to be special cases of the class exhibited in section 4. In section 6 we show that none of these mechanisms satisfies individual rationality but that a property called individual rationality on average obtains in a large class of cases. In section 7 we briefly explore the possibility of designing mechanisms which allow for coalition formation. We conclude with a consideration of the case where agents' expectations about others depend on their own characteristics.

## 2. The model

We consider an economy with $n$ ($n \geq 2$) consumers (indexed by $i = 1, \ldots, n$) and two commodities one public and one private. The utility function of consumer $i$, $u_i(K, x_i)$ is additively separable in the public good $K$ and the private good $x_i$, $i = 1, \ldots, n$, and is written

$$v_i(K, \theta_i) + x_i \tag{1}$$

where $\theta_i$ belongs to an open interval $\Theta_i$ in $\mathbf{R}$.[1]

The theory developed below can be trivially extended to utility functions of the type

$$v_i(K, \theta_i) + \phi_i(x_i), \quad \text{with } \phi_i'(\cdot) > 0,$$

when the functions $\phi_i(\cdot)$ are known by the designer.

We assume that $K$ is a real number; generalizations to multidimensional project spaces are entirely straightforward.

The designer is supposed to know the valuation *functions* $v_i(\cdot, \cdot)$ (possibly identical for all consumers) but is ignorant of the true value $\hat{\theta}_i$ of the parameter $\theta_i$, which identifies agent $i$'s tastes for the public good. The purpose of a mechanism is to choose an 'optimal' level of the public good within this framework of incomplete information. More formally, a *mechanism* is a mapping, $f(\cdot)$, from the *strategy spaces*

$$\Theta = \prod_{i=1}^{n} \Theta_i \text{ into } \mathbf{R} \times \mathbf{R}^n,$$

composed of a decision function, $d(\cdot)$, from $\Theta$ into $\mathbf{R}$ and of an $n$-tuple of transfer functions, $t(\cdot) = [t_1(\cdot), \ldots, t_n(\cdot)]$, from $\Theta$ into $\mathbf{R}^n$. $d(\cdot)$ associated to any

[1] Generalizations to multidimensional parameter spaces are straightforward.

$n$-tuple $\theta$ of announced parameters a level $d(\theta)$ of public good, while $t_i(\theta)$ is a transfer of the private good to agent $i$, $i = 1, \ldots, n$. The mechanism is said to be *continuously differentiable* or $C^1$ when the function $f(\cdot)$ is continuously differentiable.

Since the endowments of the private good play no role in the paper we will henceforth write the utility function of an agent $i$, for a mechanism $(d(\cdot), t(\cdot))$, as

$$v_i(d(\theta), \hat{\theta}_i) + t_i(\theta), \qquad i = 1, \ldots, n.$$

Faced with a mechanism the agent chooses his answer $\tilde{\theta}_i$. We may be interested in a number of solution concepts for this game form.

We may require, for instance, that $\tilde{\theta}_i$ be a dominant strategy for each agent. In Laffont and Maskin [8] we characterize and study the mechanisms for which the truth is a dominant strategy, i.e. for which $\forall i$, $\forall \theta \in \Theta$

$$v_i(d(\hat{\theta}_i, \theta_{-i}), \hat{\theta}_i) + t_i(\hat{\theta}_i, \theta_{-i}) \geq v_i(d(\theta_i, \theta_{-i}), \hat{\theta}_i) + t_i(\theta_i, \theta_{-i}),$$

where $\theta_{-i} = (\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_n)$ and for which the decision function maximizes the sum of the valuation functions.

Here we consider a more general decision function $K^*(\theta_1, \ldots, \theta_n)$ which maximizes the welfare criterion

$$F = F(v_1(K, \theta_1), \ldots, v_n(K, \theta_n))$$

and we make the following assumption.

**Assumption 1.** For $i = 1, \ldots, n$, let $v_i : \mathbf{R} \times \Theta_i \to \mathbf{R}$ be a continuously differentiable function such that for any $\theta \in \Theta$, there exists $K^*(\theta)$ such that (i) $K^*(\theta)$ maximizes $F(v_1(K, \theta_1), \ldots, v_n(K, \theta_n))$, and (ii) $K^*(\theta)$ is continuously differentiable.

In the next section we adapt the results of Laffont and Maskin [8] to this slightly more general framework. We find that it is impossible in general to obtain balanced mechanisms (i.e. such that $\sum_{i=1}^{n} t_i(\theta) \equiv 0$) if one insists on the dominant strategy requirement.

A possible solution to the balance problem, first explored by d'Aspremont and Gérard-Varet [1], is to weaken the solution concept and require only that the truth be an equilibrium with respect to expected utility maximization.

Let $f_i(\theta_{-i}; \hat{\theta}_i)$ be the probability density function reflecting agent $i$'s expectations about the strategies, $\theta_{-i}$, of the other agents when his own parameter is $\hat{\theta}_i$. One case where $f_i$ might depend on $\hat{\theta}_i$ is when $f_i$ represents the conditional distribution derived from the joint and presumably publicly known distribution of $\theta$.

A mechanism $[d(\cdot), t(\cdot)]$ is said to be individually incentive compatible with respect to expected utility maximization (EIIC) iff for any $i$

$$\int_{\Theta_{-i}} [v_i[\mathrm{d}(\hat{\theta}_i, \theta_{-i}), \hat{\theta}_i] + t_i(\hat{\theta}_i, \theta_{-i})] f_i(\theta_{-i}; \hat{\theta}_i)\, \mathrm{d}\theta_{-i}$$

$$\geq \int_{\Theta_{-i}} [v_i[\mathrm{d}(\theta_i, \theta_{-i}), \hat{\theta}_i] + t_i(\theta_i, \theta_{-i})] f_i(\theta_{-i}; \hat{\theta}_i)\, \mathrm{d}\theta_{-i}, \quad \text{for any } \theta_i.$$

Throughout the paper (but see section 8), we shall assume that the designer knows the expectations of individuals, not just the functional forms ($f_i$). We can therefore assume that the density function $f_i$ does not depend on the unknown parameter $\hat{\theta}_i$ (i.e. $f_i(\theta_{-i}; \hat{\theta}_i)$ can be written as $f_i(\theta_{-i})$).

## 3. Dominant strategy mechanisms

We present here some preliminary results on dominant strategies adapted from Laffont and Maskin [8]. We consider the problem (dominant strategy problem, DSP) of choosing transfer functions $t_i(\theta)$, $i = 1, \ldots, n$ so that

$$\forall i = 1, \ldots, n, \forall \hat{\theta}_i, \forall \theta_{-i},$$

$$\theta_i = \hat{\theta}_i \text{ maximizes } v_i(K^*(\theta), \hat{\theta}_i) + t_i(\theta) \text{ on } \Theta_i,$$

where $K^*(\theta)$ is an arbitrary decision function satisfying Assumption 1.

**Proposition 3.1.** Under Assumption 1, if there exists a $C^1$ solution $t = (t_1, \ldots, t_n)$ to the DSP, then $t_i(\theta)$ must be of the form

$$t_i(\theta) = -\int \frac{\partial v_i}{\partial K} \frac{\partial K^*}{\partial \theta_i}\, \mathrm{d}\theta_i + h_i(\theta_{-i}), \qquad i = 1, \ldots, n$$

where $h_i(\cdot)$ is an arbitrary function of $\theta_{-i}$.

*Proof.* Let $t$ be a solution to the DSP. If $\bar{\theta}_i$ is an optimal strategy for the player $i$ given $\hat{\theta}_i$ and $\theta_{-i}$, then

$$\frac{\partial v_i}{\partial K}(K^*(\bar{\theta}_i, \theta_{-i}), \hat{\theta}_i) \frac{\partial K^*}{\partial \theta_i}(\bar{\theta}_i, \theta_{-i}) + \frac{\partial t_i}{\partial \theta_i}(\bar{\theta}_i, \theta_{-i}) = 0, \qquad i = 1, \ldots, n.$$

Since the optimal solution of agent $i$'s problem must be the truth for any $\theta_{-i}$ and any $\hat{\theta}_i$ we have the identity

$$\frac{\partial t_i}{\partial \theta_i}(\theta) \equiv -\frac{\partial v_i}{\partial K}(K^*(\theta), \theta_i) \frac{\partial K^*}{\partial \theta_i}(\theta).$$

Hence

$$t_i(\theta) = -\int \frac{\partial v_i}{\partial K}(K^*(\theta), \theta_i) \frac{\partial K^*}{\partial \theta_i}(\theta)\, \mathrm{d}\theta_i + h_i(\theta_{-i}),$$

where $h_i(\cdot)$ is an arbitrary function of $\theta_{-i}$. Q.E.D.

**Proposition 3.2.** Under Assumption 1, if $K^*(\theta)$ maximizes $\Sigma_{i=1}^{n} \lambda_i v_i(K, \theta_i)$, $\lambda_i > 0$, $i = 1, \ldots, n$, then a solution to the DSP exists and is such that

$$t_i(\theta) = \frac{1}{\lambda_i} \sum_{j \neq i} \lambda_j v_j(K, \theta_j) + \bar{h}_i(\theta_{-i}),$$

where $\bar{h}_i(\theta_{-i})$ is an arbitrary function of $\theta_{-i}$.

*Proof.* $t_i(\theta)$ is a solution to the DSP since

$$v_i(K^*(\hat{\theta}_i, \theta_{-i}), \hat{\theta}_i) + \frac{1}{\lambda_i} \sum_{j \neq i} \lambda_j v_j(K^*(\hat{\theta}_i, \theta_{-i}), \theta_j)$$
$$\geq v_i(K^*(\theta_i, \theta_{-i}) + \frac{1}{\lambda_i} \sum_{j \neq i} \lambda_j v_j(K^*(\theta_i, \theta_{-i}), \theta_j)$$

from the definition of $K^*(\theta)$ and, therefore, $\theta_i = \hat{\theta}_i$ is a dominant strategy for agent $i$. Thus,

$$-\int \frac{\partial v_i}{\partial K} \frac{\partial K^*}{\partial \theta_i} \mathrm{d}\theta_i \text{ differs from } \frac{1}{\lambda_i} \sum_{j \neq i} \lambda_j v_j(K, \theta_j)$$

only by a function of $\theta_{-i}$, so that

$$\frac{1}{\lambda_i} \sum_{j \neq i} \lambda_j v_j(K, \theta_j) + \bar{h}_i(\theta_{-i})$$

provides a description of all the solutions to the DSP. Q.E.D.

## 4. Characterization of balanced EIIC mechanisms

Consider the problem of choosing the functions $t_1, \ldots, t_n$ such that $\Sigma_{i=1}^{n} t_i \equiv 0$ and so that for all $i$ and all $\hat{\theta}_i$, $\theta_i = \hat{\theta}_i$ maximizes

$$\int_{\theta_{-i}} [v_i(K^*(\theta), \hat{\theta}_i) + t_i(\theta)] f_i(\theta_{-i}) \, \mathrm{d}\theta_{-i},$$

where $K = K^*(\theta)$ (assuming it exists) maximizes the function $F = F(v_1(K, \theta_1), \ldots, v_n(K, \theta_n))$. Call this the *expected utility problem* (EUP) for $F$.

In the following proposition we characterize the set of transfer function vectors $t = (t_1, \ldots, t_n)$ which solve the EUP if a solution exists. We shall subsequently (Proposition 4.2) characterize the family of $F$'s for which the EUP can be solved.

**Proposition 4.1.** Under Assumption 1, if there exists a $C^1$ solution to the

EUP for $F$ then $t$ is a solution iff

$$\forall i\; t_i(\theta) = \int_{\Theta_{-i}} \varphi_i(\theta) f_i(\theta_{-i})\,\mathrm{d}\theta_{-i} - \frac{1}{n-1}\sum_{j\neq i} \int_{\Theta_{-j}} \varphi_j(\theta) f_j(\theta_{-j})\,\mathrm{d}\theta_{-j} + H_i(\theta), \tag{2}$$

where

$$\varphi_j(\theta) = -\int \frac{\partial v_j}{\partial K}(K^*(\theta), \theta_j)\frac{\partial K^*}{\partial \theta_j}(\theta)\,\mathrm{d}\theta_j$$

and where $\{H_j\}$ is a collection of functions such that $\Sigma_{j=1}^n H_j \equiv 0$ and

$$\int_{\Theta_{-j}} \frac{\partial H_j}{\partial \theta_j}(\theta) f_j(\theta_{-j})\,\mathrm{d}\theta_{-j} = 0, \quad \text{for all } \theta_j \in \Theta_j.$$

*Proof.* Suppose that $t^*$ solves the EUP. From incentive compatibility, we have $\forall i$, $\forall \hat{\theta}_i \in \Theta_i$

$$\frac{\partial}{\partial \theta_i}\int_{\Theta_{-i}} t_i^*(\theta) f_i(\theta_{-i})\,\mathrm{d}\theta_{-i} = -\frac{\partial}{\partial \theta_i}\int_{\Theta_{-i}} v_i(K^*(\theta), \hat{\theta}_i) f_i(\theta_{-i})\,\mathrm{d}\theta_{-i} \tag{3}$$

at the point $\theta_i = \hat{\theta}_i$.

Reversing the order of integration and differentiation in (3) we obtain

$$\frac{\partial}{\partial \theta_i}\int_{\Theta_{-i}} t_i^*(\hat{\theta}_i, \theta_{-i}) f_i(\theta_{-i})\,\mathrm{d}\theta_{-i}$$
$$= \int_{\Theta_{-i}} \frac{\partial v_i}{\partial K}(K^*(\hat{\theta}_i, \theta_{-i}), \hat{\theta}_i)\frac{\partial K^*}{\partial \theta_i}(\hat{\theta}_i, \theta_{-i}) f_i(\theta_{-i})\,\mathrm{d}\theta_{-i}. \tag{4}$$

Replacing $\hat{\theta}_i$ by $\theta_i$, reintegrating (4) with respect to $\theta_i$ and reversing the order of integration yields

$$\int_{\Theta_{-i}} t_i^*(\theta) f_i(\theta_{-i})\,\mathrm{d}\theta_{-i} = -\int_{\Theta_{-i}} \varphi_i(\theta) f_i(\theta_{-i})\,\mathrm{d}\theta_{-i} + C_i^*, \tag{5}$$

where $C_i^*$ is a constant. Consider the set of equations

$$\int_{\Theta_{-i}} t_i(\theta) f_i(\theta_{-i})\,\mathrm{d}\theta_{-i} = -\int_{\Theta_{-i}} \varphi_i(\theta) f_i(\theta_{-i})\,\mathrm{d}\theta_{-i} + C_i, \tag{6}$$

where $C_i$ is a constant.

The set of solutions to (6) consists of $\bar{t} = (\bar{t}_1, \ldots, \bar{t}_n)$ such that

$$\bar{t}_i(\theta) = -\int_{\Theta_{-i}} \varphi_i(\theta) f_i(\theta_{-i})\,\mathrm{d}\theta_{-i} + \bar{H}_i(\theta),$$

where $\bar{H}_i$ is a function such that

$$\int_{\Theta_{-i}} \frac{\partial \bar{H}_i(\cdot)}{\partial \theta_i} f_i(\theta_{-i})\, d\theta_{-i} = 0.$$

It is immediate that if $\bar{t}$ solves (6), then $\bar{t}$ solves the EUP iff $\Sigma_{i=1}^{n} \bar{t}_i = 0$. Choose

$$H_i = \bar{H}_i + \frac{1}{n-1} \sum_{j \neq i} \int_{\Theta_{-j}} \varphi_j(\theta) f_j(\theta_{-j})\, d\theta_{-j}.$$

Notice that

$$\int_{\Theta_{-i}} H_i(\theta) f_i(\theta_{-i})\, d\theta_{-i}$$

is a constant and that

$$\bar{t}_i(\theta) = \int_{\Theta_{-i}} \varphi_i(\theta) f_i(\theta_{-i})\, d\theta_{-i} - \frac{1}{n-1} \sum_{j \neq i} \int_{\Theta_{-j}} \varphi_j(\theta) f_j(\theta_{-j})\, d\theta_{-j} + H_i(\theta). \quad (7)$$

Then, if $\bar{t}_i$ satisfies (7) for all $i$, $\bar{t}$ solves the EUP for $F$ iff $\Sigma_i H_i = 0$.

Q.E.D.

## 5. Existence of balanced EIIC mechanisms

D'Aspremont and Gérard-Varet [1] have exhibited a class of balanced EIIC mechanisms which can be derived as follows. Suppose one restricts the transfer functions to be additively separable, i.e.

$$t_i(\theta) = \sum_{j=1}^{n} r_{ij}(\theta_j)$$

and suppose that the decision function $K^*(\theta)$ maximizes $\Sigma_{i=1}^{n} \lambda_i v_i(K, \theta_i)$.

Revelation of the truth $\hat{\theta}_i$, under expected utility maximizing behavior, then leads to the necessary condition[2]

$$E_{\theta_{-i}} \frac{\partial v_i}{\partial K} \frac{\partial K^*}{\partial \theta_i} + E_{\theta_{-i}} \frac{\partial r_{ii}(\theta_i)}{\partial \theta_i} \equiv 0 \quad (8)$$

or

$$r_{ii}(\theta_i) = -\int \left( E_{\theta_{-i}} \frac{\partial v_i}{\partial K} \frac{\partial K^*}{\partial \theta_i} \right) d\theta_i + C_i.$$

Switching the operators $\int$ and $E_{\theta_{-i}}$, we have

$$r_{ii}(\theta_i) = -E_{\theta_{-i}} \int \frac{\partial v_i}{\partial K} \frac{\partial K^*}{\partial \theta_i}\, d\theta_i + C_i.$$

[2] $E_{\theta_{-i}}(\cdot) = \int_{\Theta_{-i}} (\cdot) f_i(\theta_{-i})\, d\theta_{-i}$.

But from Proposition 3.2 we get

$$r_{ii}(\theta_i) = E_{\theta_{-i}}\left[\frac{1}{\lambda_i}\sum_{j\neq i}\lambda_j v_j(K^*(\theta),\theta_j) + h_i(\theta_{-i})\right]$$
$$= E_{\theta_{-i}}\frac{1}{\lambda_i}\sum_{j\neq i}\lambda_j v_j(K^*(\theta),\theta_j) + \bar{C}_i.$$

The $r_{ij}(\theta_j)$, $j\neq i$, are irrelevant for incentive compatibility. They can be chosen to balance the budget in, for example, the following simple and symmetric way:

$$r_{ij}(\theta_j) = -\frac{1}{n-1}r_{jj}(\theta_j).$$

Deleting the constants ($\bar{C}_i$) one obtains

$$t_i(\theta) = E_{\theta_{-i}}\frac{1}{\lambda_i}\sum_{j\neq i}\lambda_j v_j(K^*(\theta),\theta_j) - \frac{1}{n-1}\sum_{l\neq i}E_{\theta_{-l}}\frac{1}{\lambda_l}\sum_{j\neq l}\lambda_j v_j(K^*(\theta),\theta_j). \tag{9}$$

The fact that this solution to the necessary condition (8) is a solution to the expected utility incentive compatibility problem is derived immediately from Proposition 3.2.[3]

The mechanisms specified by (9) coincide with the d'Aspremont Gérard-Varet (AGV) mechanisms when all the ($\lambda_j$) are equal. The preceding argument provides, therefore, a constructive way of obtaining the AGV mechanisms, which appear as a special case of the general class of expected utility maximizing mechanisms exhibited in section 4.[4]

We have just shown that when $F$ is a weighted sum of valuation functions, then solutions to the EUP for $F$ exist. We shall now demonstrate that if $F$ is anything else, then if the class of valuation functions is sufficiently large, the EUP problem has no solution.

For convenience, we shall make the following assumption about $F$.

**Assumption 2.** $F$ is strictly concave and $\partial F/\partial v_i > 0\ \forall i$.

Assumption 2 is not actually necessary, but one can argue that most interesting nonlinear welfare criteria will satisfy its stipulations anyway.

We shall work with a particular small class $V_Q$ of valuation functions; namely, one where players 2 through $n$ have quadratic valuation functions and player 1 has a quadratic valuation function with an additional constant term. That is

$$V_Q = \left\{\theta_1 K - K^2 + \alpha, \theta_2 K - K^2, \ldots, \theta_n K - K^2 \,\middle|\, \theta \in \Theta = \prod_{i=1}^{n}\Theta_i, \alpha \in \mathbf{R}\right\},$$

where $\Theta_i$ is a bounded open interval of the real line.

[3]If the valuation functions are all strictly concave, one can show directly that the truth is the best strategy with a transfer function as defined by (9). The argument consists of proving that the second order conditions of the expected utility maximization problem are fulfilled at all critical points.

[4]See also Arrow [2].

The reasons for considering the family $V_Q$ are, that (1) it is a very small class of preferences and one which would be likely to be contained in most design problems of interest, and that (2) quadratic preferences are ordinarily the most promising class for positive results (see, for example, section 7 on coalitions); therefore, it is especially interesting when results are negative for this class (or, as in this case, a slightly modified version of this class). We shall make one more simplifying assumption.

**Assumption 3.** $\forall \theta \in \Theta$, $\forall \alpha_1 \in \mathbf{R}$ there exists a unique $K$ which maximizes $F(\bar{v}_1(K, \theta_1, \alpha_1), v_2(K, \theta_2), \ldots, v_n(K, \theta_n))$, where the valuation functions are those of class $V_Q$.

We can now state our impossibility result.

**Proposition 5.1.** If $F$ and $V_Q$ satisfy Assumptions 1, 2 and 3, there exists no solution to the EUP for $F$.

**Remark.** This result may seem to contradict Arrow's finding (2) that, for any welfare criterion $F$ for which the welfare maximizing choice $K^*$ satisfies $(\partial K^*/\partial \theta_i)(\partial^2 v_i/\partial K \partial \theta_i) \geq 0$ for all $i$, there exists an EIIC mechanism. The resolution of the discrepancy is the observation that the condition $(\partial K^*/\partial \theta)(\partial^2 v_i/\partial K \partial \theta_i) \geq 0$ (which Arrow calls 'positive responsiveness') cannot be satisfied in general unless $F$ is a weighted sum of individual valuation functions.

*Proof.* We shall begin by reparameterizing $v_1$. Take $v_1 = \bar{\theta}_1 K - K^2 + \beta \bar{\theta}_1 + \gamma$, where $\beta, \gamma \in \mathbf{R}$. Obviously the class of valuation functions $\{v_1\}$ is the same as the class $\{\bar{v}_1\}$, so the reparameterization affects the problem in no substantive way. Take $\theta_1 = (\bar{\theta}_1, \beta, \gamma)$. Now, suppose that $t$ solves the EUP for $F$. From Proposition 4.1

$$\forall \theta_1, \int\limits_{\Theta_{-1}} \frac{\partial t_1}{\partial \bar{\theta}_1}(\theta) f_1(\theta_{-1})\, \mathrm{d}\theta_{-1} = -\int\limits_{\Theta_{-1}} \frac{\partial v_1}{\partial K}(K^*(\theta_1), \theta_1) \frac{\partial K^*}{\partial \bar{\theta}_1}(\theta) f_1(\theta_{-1})\, \mathrm{d}\theta_{-1}.$$

The second order conditions for player 1's maximization problem imply, therefore, $\forall \hat{\theta}_1, \hat{\beta}, \hat{\gamma}$

$$\frac{\partial^2}{\partial \bar{\theta}_1^2}\left(\int\limits_{\Theta_{-1}} [v_1(K^*(\theta_1, \beta, \gamma, \theta_{-1}), \hat{\theta}_1, \hat{\beta}_1, \hat{\gamma}) + t_1(\theta_1, \beta, \gamma, \theta_{-1})] f_1(\theta_{-1})\, \mathrm{d}\theta_{-1}\right)\Bigg|_{\hat{\bar{\theta}}_1}$$

$$= -\int\limits_{\Theta_{-1}} \frac{\partial K^*}{\partial \bar{\theta}_1}(\hat{\theta}_1, \theta_{-1}) \frac{\partial^2 v_1}{\partial K \partial \bar{\theta}_1}(K^*(\hat{\theta}_1, \theta_{-1}), \hat{\theta}_1) f_1(\theta_{-1})\, \mathrm{d}\theta_{-1} < 0. \qquad (10)$$

We may solve for $\partial K^*/\partial\bar{\theta}_1$ as follows:

$$\sum_{j=1}^{n} \frac{\partial F}{\partial v_j}((v_j(K^*(\theta), \theta_j))) \frac{\partial v_j}{\partial K}(K^*(\theta), \theta_j) = 0, \tag{11}$$

because $K = K^*$ maximizes $F$. Differentiating (11) with respect to $\bar{\theta}_1$ yields

$$\sum_{j=1}^{n} \frac{\partial F}{\partial v_j} \frac{\partial^2 v_j}{\partial K^2} \frac{\partial K^*}{\partial \bar{\theta}_1} + \frac{\partial F}{\partial v_1} \frac{\partial^2 v_1}{\partial K \partial \bar{\theta}_1} + \sum_r \sum_s \frac{\partial v_r}{\partial K} \frac{\partial^2 F}{\partial v_r \partial v_s} \frac{\partial v_s}{\partial K} \frac{\partial K^*}{\partial \bar{\theta}_1}$$
$$+ \sum_{j=1}^{n} \frac{\partial v_j}{\partial K} \frac{\partial^2 F}{\partial v_j \partial v_1} \frac{\partial v_1}{\partial \bar{\theta}_1} = 0, \tag{12}$$

where, for convenience, the arguments have been suppressed. Solving for $\partial K^*/\partial\bar{\theta}_1$ in (12), we obtain

$$\frac{\partial K^*}{\partial \bar{\theta}_1} = \frac{-\left(\dfrac{\partial F}{\partial v_1} \dfrac{\partial^2 v_1}{\partial K \partial \bar{\theta}_1} + \displaystyle\sum_{j=1}^{n} \dfrac{\partial v_j}{\partial K} \dfrac{\partial^2 F}{\partial v_j \partial v_1} \dfrac{\partial v_1}{\partial \bar{\theta}_1}\right)}{\displaystyle\sum_{j=1}^{n} \dfrac{\partial F}{\partial v_j} \dfrac{\partial^2 v_j}{\partial K^2} + \sum_r \sum_s \dfrac{\partial v_r}{\partial K} \dfrac{\partial^2 F}{\partial v_r \partial v_s} \dfrac{\partial v_s}{\partial K}}. \tag{13}$$

The denominator of (13) does not vanish because of Assumption 2. Substituting (13) into (10), we get

$$\int_{\theta_{-1}} \frac{\left(\dfrac{\partial F_1}{\partial v_1} \dfrac{\partial^2 v_1}{\partial K \partial \bar{\theta}_1} + \displaystyle\sum_{j=1}^{n} \dfrac{\partial v_j}{\partial K} \dfrac{\partial^2 F}{\partial v_j \partial v_1} \dfrac{\partial v_1}{\partial \bar{\theta}_1}\right) \dfrac{\partial^2 v_1}{\partial K \partial \bar{\theta}_1}}{\displaystyle\sum_{j=1}^{n} \dfrac{\partial F}{\partial v_j} \dfrac{\partial^2 v_j}{\partial K^2} + \sum_r \sum_s \dfrac{\partial v_r}{\partial K} \dfrac{\partial^2 F}{\partial v_1 \partial v_s} \dfrac{\partial v_s}{\partial K}} f_1(\theta_{-1}) \, d\theta_{-1} < 0. \tag{14}$$

Or, letting $D(\hat{\theta}_1, \theta_{-1})$ be the denominator in (14) and substituting the explicit functions in (14),

$$\int_{\theta_{-1}} \left[\frac{\partial F}{\partial v_1} + \sum_{j \neq 1} (\theta_j - 2K^*) \frac{\partial^2 F}{\partial v_j \partial v_1} (K^* + \hat{\beta}) + (\hat{\bar{\theta}}_1 - 2K^*) \frac{\partial^2 F}{\partial v_1 \partial v_1} (K^* + \hat{\beta})\right]$$
$$\times \frac{f_1(\theta_{-1})}{D(\hat{\theta}_1, \theta_{-1})} \, d\theta_{-1} < 0. \tag{15}$$

Let

$$G(\hat{\theta}_1) = \int_{\theta_{-1}} \frac{\displaystyle\sum_{j \neq i} (\theta_j - 2K^*) \frac{\partial^2 F}{\partial v_j \partial v_1} + (\hat{\bar{\theta}}_1 - 2K^*) \frac{\partial^2 F}{\partial v_1^2}}{D(\hat{\theta}_1, \theta_{-1})} f_1(\theta_{-1}) \, d\theta_{-1}.$$

Choose $\hat{\theta}_1$ such that $G(\hat{\theta}_1) \neq 0$. Such a choice is possible by the strict concavity of $F$ and the $v_j$'s. Let

$$H(\hat{\theta}_1) = \int_{\theta_{-1}} \frac{\dfrac{\partial F}{\partial v_1} + \displaystyle\sum_{j \neq i} (\theta_j - 2K^*) \frac{\partial^2 F}{\partial v_j \partial v_1} K^* + (\hat{\bar{\theta}}_1 - 2K^*) \frac{\partial^2 F}{\partial v_1^2} K^*}{D(\hat{\theta}_1, \theta_{-1})} f_1(\theta_{-1}) \, d\theta_{-1}.$$

Then (15) can be rewritten as

$$\hat{\beta}G(\hat{\theta}_1) + H(\hat{\theta}_1) < 0. \tag{16}$$

If (16) holds, choose $\hat{\hat{\beta}}$ to reverse the inequality. Note that such a choice is possible without changing the values of $G(\hat{\theta}_i)$ and $H(\hat{\theta}_i)$ if one simultaneously substitutes $\hat{\hat{\gamma}}$ for $\hat{\gamma}$ so that

$$\hat{\hat{\gamma}} + \hat{\hat{\beta}}\hat{\hat{\theta}}_1 = \hat{\gamma} + \hat{\beta}\hat{\hat{\theta}}_1.$$

Indeed, writing $\hat{\hat{\theta}}_1 = (\hat{\hat{\theta}}_1, \hat{\hat{\beta}}, \hat{\hat{\gamma}})$ we have

$$v_1(K^*(\hat{\hat{\theta}}_1, \theta_{-1}), \hat{\hat{\theta}}_1) = v_1(K^*(\hat{\theta}_1, \theta_{-1}), \hat{\theta}_1), \quad \forall \theta_{-1}.$$

Therefore, second order conditions are violated at the truth and incentive compatibility cannot hold. Q.E.D.

## 6. Individual rationality

By the 'individual rationality' of a game is meant the property that all players are ex ante at least as well off from playing the game as from refraining from play. For the expected utility game, this property amounts to requiring that

$$\forall i, \forall \hat{\theta}_i\, E_{\theta_{-i}}[v_i(K^*(\theta), \hat{\theta}_i) + t_i(\theta)] \geq v_i(\bar{K}, \hat{\theta}_i),$$

where $\bar{K}$ is the level of $K$ which would prevail if the game were not played. For convenience, we assume that $\bar{K} = 0$. From Proposition 4.1, we have, in equilibrium,

$$\forall i, \forall \hat{\theta}_i\, E_{\theta_{-i}}[v_i(K^*(\theta), \hat{\theta}_i) + t_i(\theta)]$$

$$= E_{\theta_{-i}}\left[v_i(K^*(\hat{\theta}_i, \theta_{-i}), \hat{\theta}_i) + \varphi_i - \frac{1}{n-1}\sum_{j \neq i} E_{-\theta_j}\varphi_j + H_i(\theta)\right].$$

Consider the case of consistent expectations, i.e. expectations derived from the same joint probability distribution over $\theta$. Then

$$E_\theta = E_{\theta_i}E_{\theta_{-i}} = E_{\theta_j}E_{\theta_{-j}}, \quad \forall i, j.$$

Since transfers are always balanced we have

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

and therefore

$$E_\theta\left[\sum_{i=1}^{n}\left(\varphi_i - \frac{1}{n-1}\sum_{j \neq i} E_{-\theta_j}\varphi_j + H_i(\theta)\right)\right] = 0$$

or

$$\sum_{i=1}^{n} E_\theta H_i(\theta) = \sum_{i=1}^{n} E_{\theta_{-i}} H_i(\theta) = 0,$$

since $E_{\theta_{-i}}H_i(\theta)$ is constant for every $i$ from the characterization theorem.

But suppose, as is surely possible in general, that for any $i$ there exists $\hat{\theta}_i$ such that

$$E_{\theta_{-i}} v_i[K^*(\hat{\theta}_i, \theta_{-i}), \hat{\theta}_i] < v_i(0, \hat{\theta}_i).$$

Then

$$\sum_{i=1}^{n} E_{\theta_{-i}}\left[v_i(K^*(\hat{\theta}_i, \theta_{-i}), \hat{\theta}_i) + \varphi_i - \frac{1}{n-1}\sum_{j \neq i} E_{\theta_{-j}}\varphi_j + H_i(\theta)\right] < \sum_{i=1}^{n} v_i(0, \hat{\theta}_i).$$

The above inequality implies that there exists for some $j$ a $\hat{\theta}_j$ such that

$$E_{\theta_{-j}}[v_j(K^*(\hat{\theta}_j, \theta_{-j}), \hat{\theta}_j) + t_j(\hat{\theta}_j, \theta_{-j})] < v_j(0, \hat{\theta}_j).$$

Therefore it is clear that individual rationality cannot be guaranteed in general.

On the other hand, there is a kind of individual rationality which can be ensured in a wide class of cases, particularly when expectations are consistent. Suppose that we think of society as facing a large number of public decisions in the future, and of agents asking themselves if they should accept the mechanism for this set of decisions. Individuals are not certain what the exact nature of these projects will be, but they do have probabilistic beliefs. One way of modelling this situation is to suppose that individuals are unsure of their *own* characteristics as well as those of others. In this case individual rationality becomes *individual rationality on average*, a concept introduced by Green and Laffont [4] for dominant strategy mechanisms. In the simple case of consistent expectations, this property can be written as

$$E_{\hat{\theta}_i} E_{\theta_{-i}}[v_i(K^*(\hat{\theta}_i, \theta_{-i}), \hat{\theta}_i) + t_i(\hat{\theta}_i, \theta_{-i})] \geq E_{\hat{\theta}_i} v_i(0, \hat{\theta}_i), \qquad i = 1, \ldots, n.$$

Suppose that $K^*(\theta)$ maximizes the weighted sum $\Sigma_{i=1}^{n} \lambda_i v_i(K^*(\theta), \theta_i)$, we can then prove:

**Proposition 6.1.** Given Assumption 1 and consistent expectations, there exists a solution to the EUP which is individually rational on average.

*Proof.* Under Assumption 1 we know that there exists a general solution $t_i(\theta)$ given by Proposition 3.2. Since expectations are consistent we may write without ambiguity

$$A = E_\theta \sum_{i=1}^{n} \lambda_i v_i(K^*(\theta), \theta_i).$$

Observe that by definition of $K^*(\theta)$

$$A \geq E_\theta \sum_{i=1}^{n} \lambda_i v_i(0, \theta_i). \tag{17}$$

Suppose that $t = (t_1, \ldots, t_n)$ is a vector of transfer functions such that $t$ is a solution to the EUP. Then, take $A_i = E_\theta(v_i(K^*(\theta), \theta_i) + t_i)$. Notice that

$$\sum_{i=1}^{n} A_i = A. \tag{18}$$

From (17) and (18) we can choose constants $T_1, \ldots, T_n$ such that $A_i + T_i \geq E_\theta(v_i(0, \theta_i))$ and $\Sigma_{i=1}^{n} T_i = 0$. Therefore $(t_1 + T_1, \ldots, t_n + T_n)$ is still a solution to the EUP and ensures individual rationality on average. Q.E.D.

## 7. Coalition incentive compatibility

To approach the question of cooperation in games of incomplete information with side payments, it is necessary to make assumptions about the information exchanges and the types of agreements which take place between agents when they cooperate. Within the framework of our revelation game, two hypotheses appear attractive.

Suppose first that there exists a class $\mathscr{C}$ of disjoint coalitions within which knowledge of the parameters $\theta_i$ of the coalition's agents is common knowledge. Within each coalition a group of agents can choose or choose not to formulate a joint strategy.

Let $C$ be a coalition and let $I_C$ be any subset of $C$. Expectations of the members of the coalition about those outside $C$ are identical (since we assume that they share their private information).

Let $f_C(\theta_{-C})$ be the density function reflecting coalition $C$'s expectations. It will also be the density function of any subcoalition of agents, since their information is the same as that of the entire coalition.

Coalition incentive compatibility for a coalition $C$ means then

$$\forall I_C \subset C, \sum_{i \in I_C} \int_{\theta_{-C}} \{v_i[K^*(\hat{\theta}_C, \theta_{-C}), \hat{\theta}_i] + t_i(\hat{\theta}_C, \theta_{-C})\} f_C(\theta_{-C})\, d\theta_{-C}$$

$$\geq \sum_{i \in I_C} \int_{\theta_{-C}} \{v_i[K^*(\theta_{I_C}, \hat{\theta}_{C \setminus I_C}, \theta_{-C}), \hat{\theta}_i] + t_i(\theta_{I_C}, \hat{\theta}_{C \setminus I_C}, \theta_{-C})\} f_C(\theta_{-C})\, d\theta_{-C}, \tag{19}$$

for any $\hat{\theta}_{C \setminus I_C}$ and for any $\theta_{I_C}$.

Let us call

$$V_i[\theta_{I_C}, \hat{\theta}_{C \setminus I_C}, \hat{\theta}_i] = \int_{\theta_{-C}} v_i[K^*(\theta_{I_C}, \hat{\theta}_{C \setminus I_C}, \theta_{-C}), \hat{\theta}_i] f_C(\theta_{-C})\, d\theta_{-C},$$

$$T_i[\theta_{I_C}, \hat{\theta}_{C \setminus I_C}] = \int_{\theta_{-C}} t_i(\theta_{I_C}, \hat{\theta}_{C \setminus I_C}, \theta_{-C}) f_C(\theta_{-C})\, d\theta_{-C}.$$

Then (19) can be rewritten as

$$\forall I_C \subset C, \sum_{i\in I_C} (V_i[\hat{\theta}_{I_C}, \hat{\theta}_{C\setminus I_C}, \hat{\theta}_i] + T_i(\hat{\theta}_{I_C}, \hat{\theta}_{C\setminus I_C}))$$

$$\geq \sum_{i\in I_C} (V_i[\theta_{I_C}, \hat{\theta}_{C\setminus I_C}, \hat{\theta}_i] + T_i(\theta_{I_C}, \hat{\theta}_{C\setminus I_C})),$$

for any $\hat{\theta}_{C\setminus I_C}, \theta_{I_C}, \hat{\theta}_{I_C}$.

Then, within coalition $C$, we are strictly back in the framework of dominant strategy mechanisms for the functions $V_i(\cdot)$. The requirements on the $T_i(\cdot)$ functions are the same as those for the construction of a dominant strategy mechanism which is coalition incentive compatible with respect to all sub-coalitions of $C$. The negative results obtained in Laffont and Maskin [8] apply; in particular, Corollary (10) implies the general nonexistence of such mechanisms.

Another framework of interest is the situation in which the information structure is not defined before the game starts. We shall assume that when a subset of agents decides to form a coalition and to share private information, members of the coalition are constrained to adhere to the jointly determined strategy vector, i.e. they cannot renege.

Let $\mathscr{C}$ be the class of potential coalitions. The requirement for coalition incentive compatibility is then

$$\forall C \in \mathscr{C}$$

$$\sum_{i\in C} \int_{\theta_{-C}} [v_i[K^*(\hat{\theta}_C, \theta_{-C}), \hat{\theta}_i] + t_i(\hat{\theta}_C, \theta_{-C})] f_C(\theta_{-C})\, d\theta_{-C}$$

$$\geq \sum_{i\in C} \int_{\theta_{-C}} [v_i[K^*(\theta_C, \theta_{-C}), \hat{\theta}_i] + t_i(\theta_C, \theta_{-C})] f_C(\theta_{-C})\, d\theta_{-C},$$

for any $\theta_C, \theta_{-C}$.

Let $\mathscr{C}_{ik} = \{\{i, k\}, \{i\}, \{k\}\}$.

**Proposition 7.1.** Under Assumption 1 a necessary condition for coalition incentive compatibility for the class $\mathscr{C}_{ik}$ is

$$E_{\theta_{-i-k}} \frac{\partial^2 v_i}{\partial K \partial \theta_i} \frac{\partial K^*}{\partial \theta_i} = E_{\theta_{-i-k}} \frac{\partial^2 v_k}{\partial K \partial \theta_k} \frac{\partial K^*}{\partial \theta_k}. \tag{20}$$

*Proof.* From individual incentive compatibility

$$t_i(\theta) = E_{\theta_{-i}} \varphi_i - \frac{1}{n-1} \sum_{l\neq i} E_{\theta_{-l}} \varphi_l + H_i(\theta), \tag{21}$$

$$t_j(\theta) = E_{\theta_{-j}} \varphi_j - \frac{1}{n-1} \sum_{l\neq j} E_{\theta_{-l}} \varphi_l + H_j(\theta). \tag{22}$$

When coalition $(i, j)$ forms, $i$ and $j$ share their information so that coalition incentive compatibility requires

$$E_{\theta_{-i-j}}\left[\frac{\partial t_i}{\partial \theta_i}+\frac{\partial t_j}{\partial \theta_i}\right]=-E_{\theta_{-i-j}}\left[\frac{\partial v_i}{\partial K}\frac{\partial K^*}{\partial \theta_i}+\frac{\partial v_j}{\partial K}\frac{\partial K^*}{\partial \theta_i}\right]. \tag{23}$$

$$E_{\theta_{-i-j}}\left[\frac{\partial t_i}{\partial \theta_j}+\frac{\partial t_j}{\partial \theta_j}\right]=-E_{\theta_{-i-j}}\left[\frac{\partial v_i}{\partial K}\frac{\partial K^*}{\partial \theta_j}+\frac{\partial v_j}{\partial K}\frac{\partial K^*}{\partial \theta_j}\right]. \tag{24}$$

Substituting (21) and (22) into (23) and (24) and switching the operators $E_{\theta_{-i-j}}$ and $\partial/\partial\theta_i$, $\partial/\partial\theta_j$ gives

$$\frac{\partial}{\partial \theta_i}E_{\theta_{-i-j}}[H_i(\theta)+H_j(\theta)]$$
$$=-\frac{n-2}{n-1}E_{\theta_{-i}}\frac{\partial \varphi_i}{\partial \theta_i}-E_{\theta_{-i-j}}\left[\frac{\partial v_i}{\partial K}\frac{\partial K^*}{\partial \theta_i}+\frac{\partial v_j}{\partial K}\frac{\partial K^*}{\partial \theta_i}\right], \tag{25}$$

$$\frac{\partial}{\partial \theta_j}E_{\theta_{-i-j}}[H_i(\theta)+H_j(\theta)]$$
$$=-\frac{n-2}{n-1}E_{\theta_{-j}}\frac{\partial \varphi_j}{\partial \theta_j}-E_{\theta_{-i-j}}\left[\frac{\partial v_i}{\partial K}\frac{\partial K^*}{\partial \theta_j}+\frac{\partial v_j}{\partial K}\frac{\partial K^*}{\partial \theta_j}\right]. \tag{26}$$

A necessary and sufficient condition for the existence of functions $H_i(\theta)+H_j(\theta)$ satisfying these conditions is, by Poincaré's theorem, the equality of the second cross derivatives, which is equivalent to (20). Q.E.D.

This condition (20), which is a joint condition on expectations, utility functions and decision functions, is a very stringent stipulation, and is likely not to be satisfied in most cases. Moreover, it is only a necessary condition because $H_i(\theta)$ and $H_j(\theta)$ must also be such that

$$E_{\theta_{-i}}\frac{\partial H_i}{\partial \theta_i}(\theta)=E_{\theta_{-j}}\frac{\partial H_j}{\partial \theta_j}(\theta)=0. \tag{27}$$

In addition, second order conditions must be satisfied.

Nevertheless, condition (20) is always fulfilled if utility functions are of the quadratic type, i.e.

$$\theta_i K-\tfrac{1}{2}K^2$$

and if

$$E_{\theta_{-i-k}}\frac{\partial K^*}{\partial \theta_i}=E_{\theta_{-i-k}}\frac{\partial K^*}{\partial \theta_k}, \tag{28}$$

which is a kind of anonymity requirement for the decision function $K^*$.

Notice that (28) is fulfilled if $K^*$ maximizes the sum of the (quadratic) valuations. This does not yet establish the existence of coalition incentive compatible mechanisms for $\mathscr{C}_{ik}$. However, we will now show that by examining more general EIIC mechanisms than the AGV mechanisms, we can indeed construct EIIC mechanisms (for our quadratic family) which are coalition incentive compatible for $\mathscr{C}_{ik}$.

Let

$$v_i(K, \theta_i) = \theta_i K - \tfrac{1}{2}K^2, \qquad i = 1, \ldots, n$$

then

$$K^*(\theta) = \frac{\sum_{i=1}^{n} \theta_i}{n}.$$

If each agent $i$'s marginal distribution for any agent $j$'s strategy is represented by a normal distribution $\mathcal{N}(0, \sigma^2)$, then

$$\int \frac{\partial v_i}{\partial K} \frac{\partial K^*}{\partial \theta_i} \mathrm{d}\theta_i = \frac{\theta_i^2}{2n}\left(1 - \frac{1}{n}\right) - \sum_{l \neq i} \frac{\theta_l \theta_i}{n^2} + h_i(\theta_{-i})$$

and

$$E_{\theta_{-i}} \varphi_i = -\frac{\theta_i^2}{2n}\left(1 - \frac{1}{n}\right) + C_i$$

and

$$t_i(\theta) = -\frac{\theta_i^2}{2n}\left(1 - \frac{1}{n}\right) + C_i + \frac{1}{n-1} \sum_{l \neq i} \frac{\theta_l^2}{2n}\left(1 - \frac{1}{n}\right) - \frac{1}{n-1} \sum_{l \neq i} C_l + H_i(\theta).$$

Then (25) and (26) become

$$\frac{\partial}{\partial \theta_i} E_{\theta_{-i,-j}}[H_i(\theta) + H_j(\theta)] = -\frac{n-2}{n^2} \theta_j$$

and

$$\frac{\partial}{\partial \theta_j} E_{\theta_{-i,-j}}[H_i(\theta) + H_j(\theta)] = -\frac{n-2}{n^2} \theta_i.$$

Choosing

$$H_i(\theta) = -\frac{1}{2} \frac{n-2}{n^2} \theta_i \theta_j,$$

$$H_h(\theta) = -\frac{1}{2} \frac{n-2}{n^2} \theta_i \theta_j,$$

we observe that

$$E_{\theta_{-i}} \frac{\partial H_i}{\partial \theta_i} = E_{\theta_{-j}} \frac{\partial H_j}{\partial \theta_j} = 0.$$

To have a balanced mechanism it is enough to choose an agent $k \neq i \neq j$ and have

$$H_k(\theta) = -H_i(\theta) - H_j(\theta).$$

Since neither $H_i(\theta)$ nor $H_j(\theta)$ depend on $\theta_k$, obviously

$$E_{\theta_{-k}} \frac{\partial H_k}{\partial \theta_k} = 0.$$

Moreover, the second order conditions are also satisfied for the coalition.

## 8. Dependent case

Throughout the paper we have assumed that for any $i$, agent $i$'s expectations about the other agents, represented by $f_i(\theta_{-i})$ were independent of his own characteristics $\theta_i$. Of course, the dependent case is also of interest; it occurs in particular when an agent's expectations are derived from joint knowledge, represented by a general joint density function $f(\theta)$. Then

$$f_i(\theta_{-i}) = f(\theta_{-i};\theta_i) = \frac{f(\theta_{-i},\theta_i)}{\int_{\Theta_{-i}} f(\theta_{-i},\theta_i)\,\mathrm{d}\theta_{-i}}.$$

We denote by $E_{\theta_{-i}/\theta_i}$ the expectation operator with respect to $f(\theta_{-i};\theta_i)$. In this section we shall deal only with the welfare criterion $F(v_1,\ldots,v_n)=\Sigma_{i=1}^n v_i$.

Suppose that one restricts the transfer functions to be additively separable, as in section 5. The first order condition for incentive compatibility becomes

$$E_{\theta_{-i}/\hat{\theta}_i}\left(\frac{\partial v_i}{\partial K}\frac{\partial K^*}{\partial \theta_i}+\frac{\partial r_{ii}}{\partial \theta_i}\right)=0. \tag{29}$$

Therefore $\theta_i^*=\hat{\theta}_i$ must be the solution of (29) for every $\hat{\theta}_i$, so that we have the identity

$$\frac{\partial r_{ii}}{\partial \theta_i}\equiv-\int_{\Theta_{-i}}\frac{\partial v_i}{\partial K}\frac{\partial K^*}{\partial \theta_i}f(\theta_{-i};\theta_i)\,\mathrm{d}\theta_{-i}.$$

Hence,

$$r_{ii}(\theta_i)=-\int\int_{\Theta_{-i}}\frac{\partial v_i}{\partial K}\frac{\partial K^*}{\partial \theta_i}f(\theta_{-i};\theta_i)\,\mathrm{d}\theta_{-i}\,\mathrm{d}\theta_i. \tag{30}$$

We can construct a balanced mechanism by choosing

$$t_i(\theta)=r_{ii}(\theta_i)-\frac{1}{n-1}\sum_{j\neq i}r_{jj}(\theta_j).$$

It remains to see what happens with the second order conditions.

The second derivative is

$$E_{\theta_{-i}/\hat{\theta}_i}\left[\frac{\partial^2 v_i}{\partial K^2}\frac{\partial K^{*2}}{\partial \theta_i}+\frac{\partial v_i}{\partial K}\frac{\partial^2 K^*}{\partial \theta_i^2}\right]$$
$$-E_{\theta_{-i}/\theta_i}\left[\frac{\partial^2 v_i}{\partial K^2}\frac{\partial K^{*2}}{\partial \theta_i}+\frac{\partial v_i}{\partial K}\frac{\partial^2 K}{\partial \theta_i^2}\right]-E_{\theta_{-i}/\theta_i}\frac{\partial^2 v_i}{\partial K\partial\theta_i}\frac{\partial K^*}{\partial \theta_i}$$
$$-\int_{\Theta_{-i}}\frac{\partial v_i}{\partial K}\frac{\partial K^*}{\partial \theta_i}\frac{\partial f}{\partial \theta_i}(\theta_{-i};\theta_i)\,\mathrm{d}\theta_{-i}.$$

If the utility functions $v_i(K,\theta_i)$ are strictly concave in $K$

$$-E_{\theta_{-i}/\theta_i}\frac{\partial^2 v_i}{\partial K\partial\theta_i}\frac{\partial K^*}{\partial \theta_i}=E_{\theta_{-i}/\theta_i}\frac{(\partial^2 v_i/\partial K\partial\theta_i)^2}{\sum_{j=1}^n(\partial^2 v_j/\partial K^2)}<0.$$

At the truth $\theta_i = \hat{\theta}_i$, the second order condition reduces to

$$\int_{\theta_{-i}} \frac{\partial v_i}{\partial K} \frac{\partial K^*}{\partial \theta_i} \frac{\partial f}{\partial \theta_i}(\theta_{-i}; \hat{\theta}_i)\, \mathrm{d}\theta_{-i} \geq 0. \tag{31}$$

We can exhibit a large class of cases in which (31) can be expected to hold. Consider a class of valuation functions $V = \{(v(\cdot, \theta_1), \ldots, v(\cdot, \theta_n)) \mid \theta \in \Theta\}$, where $\Theta_1 = \cdots = \Theta_n$ is an open interval such that

$$\forall K, \forall \theta_i, \frac{\partial^2 v}{\partial K^2}(K, \theta_i) < 0 \quad \text{and} \quad \frac{\partial^2 v}{\partial K \partial \theta_i}(K, \theta_i) > 0$$

with a welfare criterion equal to the sum of the valuation functions, we have

$$\sum_{j=1}^{n} \frac{\partial v_j}{\partial K}(K^*(\theta), \theta_j) = 0. \tag{33}$$

Therefore,

$$\sum_{j=1}^{n} \frac{\partial^2 v}{\partial K^2}(K^*, \theta_j) \frac{\partial K^*}{\partial \theta_i} + \frac{\partial^2 v}{\partial K \partial \theta_i}(K^*, \theta_i) = 0$$

and so

$$\frac{\partial K^*}{\partial \theta_i} = -\frac{\partial^2 v(K^*, \theta_i)}{\partial K \partial \theta_i} \Big/ \sum_{j=1}^{n} \frac{\partial^2 v(K^*, \theta_j)}{\partial K^2} > 0. \tag{34}$$

Now, suppose that each individual i thinks that his tastes are 'different' from those of others. This feeling of difference can be expressed formally as follows. Each realization $\theta$ of agents' characteristics induces a configuration

$$\frac{\partial v}{\partial K}(K^*, \theta_1), \ldots, \frac{\partial v}{\partial K}(K^*, \theta_n)$$

of partial derivatives, the sum of which, by (33), is zero. Let $\theta^*(\theta)$ be defined so that

$$\forall \theta, \frac{\partial v}{\partial K}(K^*(\theta), \theta^*(\theta)) = 0.$$

Then $\theta^*(\theta)$ may be thought of as the 'average' of the realized $\theta_j$'s. For an agent i to feel that he is different might be taken to mean that as $\theta_i$ moves away from $\theta^*(\theta)$, i.e. away from the 'average', the density $f_i(\theta_{-i}; \theta_i)$ increases. Formally, this amounts to

$$\begin{aligned} &\frac{\partial f_i}{\partial \theta_i}(\theta_{-i}; \theta_i) \geq 0, \quad \text{if } \theta_i \geq \theta^*(\theta), \\ &\frac{\partial f}{\partial \theta_i}(\theta_{-i}; \theta_i) \leq 0, \quad \text{if } \theta_i \leq \theta^*(\theta). \end{aligned} \tag{35}$$

Notice that, from (32) and the definition of $\theta^*(\theta)$,

$$\frac{\partial v}{\partial K}(K^*(\theta), \theta_i) < 0, \quad \text{if } \theta_i < \theta^*(\theta)$$
$$\frac{\partial v}{\partial K}(K^*(\theta), \theta_i) > 0, \quad \text{if } \theta_i > \theta^*(\theta). \tag{36}$$

Therefore, together (34), (35) and (36) imply that

$$\frac{\partial v}{\partial K}(K^*(\theta), \theta_i)\frac{\partial K^*}{\partial \theta_i}\frac{\partial f}{\partial \theta_i}(\theta_{-i}; \hat{\theta}_i) > 0, \quad \text{for all } \theta$$

and so

$$\int_{\theta_{-i}} \frac{\partial v}{\partial K}\frac{\partial K^*}{\partial \theta_i}\frac{\partial f}{\partial \theta_i}(\theta_{-i}; \hat{\theta}_i)\, \mathrm{d}\theta_{-i} > 0. \tag{37}$$

Incentive compatibility is therefore possible, at least locally.[5]

In contrast with (35), which leads to positive results, belief by an agent that he is 'average' may lead to negative results. The belief that one is average amounts to

$$\frac{\partial f}{\partial \theta_i}(\theta_{-i}; \theta_i) \leq 0, \quad \text{if } \theta_i \geq \theta^*(\theta),$$
$$\frac{\partial f}{\partial \theta_i}(\theta_{-i}; \theta_i) \geq 0, \quad \text{if } \theta_i \leq \theta^*(\theta). \tag{38}$$

Notice that this condition leads to the reversal of the inequality (37) and therefore to the impossibility of incentive incompatibility.

These positive and negative results agree well with economic intuition. Roughly speaking we expect the free-rider problem to be aggravated when agents believe that they are similar. An agent who, for example, has a strong liking for a public project and whose liking is shared by most other agents can understate his preference and still feel confident that the project will go through. Not so for the agent whose preferences are atypical.

## References


[1] d'Aspremont, C. and L.A. Gérard-Varet, 'Incentives and Incomplete Information', CORE Discussion Paper, 1976.

[2] Arrow, K., 'The Property Rights Doctrine and Demand Revelation under Incomplete Information', D.P. no. 252, Stanford University.


[5] We must include the qualification 'locally' because we have succeeded in showing only that the second order conditions hold at the truth and, therefore, only that the truth is a locally maximizing strategy.

[3] Clarke, E.H., 'Multipart Pricing of Public Goods', *Public Choice*, II (1971) 19–33.
[4] Green, J. and J.J. Laffont, *Incentives in Public Decision Making*, Studies in Public Economics, vol. 1 (North-Holland, Amsterdam, 1978).
[5] Groves, T., 'Incentives in Teams', *Econometrica*, 41 (1973) 617–631.
[6] Groves, T. and J. Ledyard, 'Optimal Allocation of Public Goods: A Solution to the Free Rider Problem', *Econometrica*, 45 (1977) 783–810.
[7] Hurwicz, L., 'On Allocations Attainable through Nash Equilibria', Chapter 22, this volume.
[8] Laffont, J.J. and E. Maskin, "A Differential Approach to Dominant Strategy Mechanisms', Ecole Polytechnique, D.P., 1977, to appear in *Econometrica*.
[9] Maskin, E., 'Nash Equilibrium and Welfare Optimality', mimeo, 1977.
[10] Vickrey, W., 'Counterspeculation, Auctions and Competitive Sealed Tenders', *Journal of Finance*, 16 (1961) 8–37.