Discussion Paper No. 116

ON THE INCENTIVE PROBLEM WITH PUBLIC GOODS *

by

John Ledyard and John Roberts

revised

January 1975

* The authors wish to thank Jacques Dreze for his comments on an earlier draft without in any way implicating him in the faults remaining. This research, which was supported by the National Science Foundation under grant No. SOC 74-04076, was completed while Roberts was at CORE.

ON THE INCENTIVE PROBLEM WITH PUBLIC GOODS

by

John Ledyard
Department of Economics
Northwestern University

and

Donald John Roberts
Department of Managerial Economics and Decision Sciences
Northwestern University

1. INTRODUCTION

While questions of incentives have long been a central concern of economics, it has been only recently that formal methods have been developed which permit rigorous analysis of such problems. A central result of this development is an impossibility theorem due to Hurwicz [7]. He showed that, if consumers' preferences are not directly observable, it is not possible to find a system for allocating resources in pure exchange economies which yields Pareto optimal allocations preferred by each consumer to his initial position and which is individually incentive compatible in the sense that no consumer will ever find it to his advantage to misrepresent his preferences. Although Hurwicz's result applies to classical, pure exchange situations without externalities, one would expect that such a result must also be true in the presence of public goods. Indeed, the problem of obtaining correct revelation of preferences has typically been considered to be primarily one which arises with public goods [see especially Samuelson [11]]. However, no analogue of the Hurwicz theorem exists as yet with public goods. We are thus in a somewhat anomalous position. We know that incentive compatibility, efficiency and a measure of equity cannot be simultaneously achieved in the absence of public goods, when incentive problems have been considered minimal, while we do not have a corresponding result for exactly that case where it has been assumed that incentive problems were most acute. The first objective of this paper, then, is to establish an analogue of Hurwicz's result for public goods economies. This is done in Section 4 .

This impossibility result raises further problems, however. Malinvaud [9,10] and Dreze and Vallee Poussin [1] have proposed a mechanism for allocating public goods which achieves optima preferred by everyone to an initial situation with zero levels of the public goods and which is incentive compatible in the sense that "at an equilibrium, all consumers have an incentive to reveal their preferences correctly for those goods which are positively produced" [1 ,p. 147]. Moreover , even out of equilibirum, the process enjoys a measure of incentive-compatibility, since it is a minimax strategy to tell the truth [1, p. 144]. The second purpose of this paper is to resolve the apparent contradiction between these results and our impossibility theorem.

It turns out that the apparent inconsistency lies in the use of different concepts of incentive compatibility: the Hurwicz result and our extension are based on a global notion of individual incentive compatibility relating to the final outcome of the process, while Dreze and Vallee Poussin employ a local notion based on the utility gain arising from a differential adjustment of a proposed allocation. We establish this point by providing a version of the Malinvaud-Dreze - Vallee Poussin procedure which inherits the essential properties of the latter and in which the Hurwicz criterion can be applied. We then show that the global version of incentive compatibility does not obtain for this mechanism although the local version does hold. This leads to the final section of the paper, which contains some thoughts on the significance of the two criteria.

2. ECONOMIES

We consider economies with K pure public goods, indexed $k = 1,\ldots,K$, and a single private good. There are N consumers indexed $i = 1,\ldots,N$. A consumption plan for consumer n is a vector $(x,y^i) \in R_+^{K+1}$ where x is his public goods consumption and y^i is private goods consumption. The preferences of consumer i are represented by a strictly concave utility function u^i which is continuously differentiable on R_+^{K+1} and for which $\partial u^i/\partial y^i > 0$. His initial endowment consists of a non-negative amount w^i of the private good.

The production sector is modeled by a continuously differentiable convex function $g: R_+^{K+1} \to R_+$. We interpret $g(x)$ as the input of private good necessary to produce the vector x of public goods.

3. THE M-D-P MODEL

The model introduced by Dreze and Vallee Poussin and Malinvaud (the M-D-P procedure) is a continous time planning procedure for determining the allocation and financing of public goods levels. It can be expressed in two differential equations which depend directly on communications received from the consumers and producer. The messages of the participants are elements of R^K. We let $\psi^i(t)$ be the message sent by consumer i at time t to the planning board, and interpret $\psi_k^i(t)$ as i's marginal willingness to pay for an additional unit of the public good k in terms of the private good. We let $\lambda(t) \in R^K$ be the message from the productive sector and interpret λ_k as the marginal cost of producing another unit of k in terms of the private good.

The M-D-P procedure is defined by:

$$(1) \qquad \dot{x}_k = \begin{cases} \Sigma_{i=1}^{N} \Psi_k^i(t) - \lambda_k(t) \text{ if } x_k > 0 \\ \max[0, \Sigma_i \Psi_k^i(t) - \lambda_k(t)] \text{ if } x_k = 0 \end{cases}$$

for $k = 1,\ldots,K$, and $t \geq 0$, and

$$(2) \qquad \dot{y}^i = -\Psi^i(t)\cdot\dot{x} + \delta_i[\dot{x}'\dot{x}]$$

for $i = 1,\ldots,N$, where $\delta^i > 0$ for all i and $\Sigma_i \delta^i = 1$.

Equation (1) represents the allocation decision between private and public goods while (2) represents the financing, or taxing, decision. We will assume throughout that the production sector is always truthful. That is, $\lambda(t) = \nabla g[x(t)]$ where $\nabla g = [\partial g/\partial x_1,\ldots,\partial g/\partial x_k]$. With this assumption, it is easy to see that

$$(3) \qquad \dot{y} = \Sigma_i \dot{y}^i = -\lambda(t)\cdot\dot{x},$$

and, therefore, (1) and (2) produce a feasible change in plans. Equations (1)-(3) are identical with [(1),(3),(7)] in [1].

The main question of interest is whether consumers have the incentive to reveal their preferences correctly. In the M-D-P model, this amounts to asking whether $\Psi^i(t)$ will equal $\pi^i[x(t),y^i(t)]$, where $\pi_k^i(x,y^i) \equiv (\partial u^i/\partial x_k)/(\partial u^i/\partial y^i)$ is i's true marginal rate of substitution of y^i for x_k. Dreze and Vallee Poussin prove [1, p. 145], Theorem: If $[\bar{x},\bar{y}^1,\ldots,\bar{y}^N,\bar{\Psi}^1,\ldots,\bar{\Psi}^N]$ is an equilibrium of [(1)(2)], (i.e., $\dot{x} = \dot{y}^i = 0$ at those values), then $\frac{du^i}{dt} < 0$ for $\bar{\Psi}^i \neq \pi^i(\bar{x},\bar{y}^i)$ whenever $\bar{x} > 0$.

This is the result referred to in the introduction of this paper. They also show that if $(x,y^1,\ldots,y^N,\Psi^1,\ldots,\Psi^N)$ is not an equilibrium, then $\frac{du^i}{dt} \geqq 0$ for $\Psi^i = \pi^i(x^i,y^i)$, i.e. each agent's utility is always non-decreasing under the adjustment if he reveals his true preferences.

3. THE HURWICZ MODEL

The model introduced by Hurwicz [7] to analyze the incentive-compatibility of allocation mechanisms also relies on communication among the participants. Although this model was originally conceived [6] for economies with only private goods, it is a simple matter to incorporate public goods in the model.

In particular, we will assume there is a language, M, in which consumers communicate. An element $m^i \in M$ will be a message of consumer i. A mechanism is specified by a triple $\langle f, \Phi, M\rangle$. Here, $f^i: M^N \times E \to M$, $i = 1,\ldots,N$, where E is the class of economies. We interpret $f^i(m^1,\ldots,m^N, e)$ as the message i is to send if m is the vector of messages received and e is the environment (the specification of tastes, endowments, and production possibilities). We say that $\bar{m}$ is an equilibrium message if $\bar{m} = f(\bar{m},e)$, and write $\bar{M}(e)$ to denote the set of equilibrium messages. The function $\Phi: \bar{M}(e) \to R^{K+N}$ is called an outcome rule, and $\Phi(m) = [x,y^1,\ldots,y^N]$ is the equilibrium allocation. A privacy respecting mechanism is one for which f^i depends on e only through $e^i = \langle u^i,w^i\rangle$, the characteristics of the i^{th} agent in the economy.

Now let G^i be the set of functions g^i from M^N into M such that $g^i(m) = f^i(m,\bar{e})$ for some $\bar{e} \in E$. G^i is a set of strategies (false response rules) which can be "rationalized" as correct behavior and whose use thus cannot be identified by an outside observer as misbehavior. Let $u^i(g^1,\ldots,g^N) \equiv u^i[\Phi(m(g^1,\ldots,g^N))]$, where $m(g^1,\ldots,g^N) \in \{m \mid g(m) = m\}$, be the payoff to i if the participants use the response rules g and let $g^{*i}(m) = f^i(m;e)$ for all m where e is the true economy. We say an agent who chooses to announce g^{*i} is revealing his true preferences. If $u^{*i}(g^*) \geq u^{*i}(g^*/g^i)$ for all $g^i \in G^i$, where $(g^*/g^i) = [g^{*1},\ldots,g^{*-1},g^i,g^{*i+1},\ldots, g^{*N}]$, then Hurwicz calls the mechanism $\langle f, \Phi, M\rangle$ individually incentive compatible [6, p. 321], or, in the language of the introduction of this note, it is in the self-interest of the participants to reveal their correct preferences.

Using this framework, Hurwicz obtained his impossibility result for economies without externalities by displaying a very "nice" pure-exchange economy for which there is no privacy respecting resource allocation mechanism which selects Pareto optima which all agents prefer to the initial allocation and which is individually incentive compatible for all agents in this economy. In the next section we present a parallel example with public goods.

4. HURWICZ'S IMPOSSIBILITY RESULT WITH PUBLIC GOODS

To establish the general impossibility of finding a resource allocation mechanism which yields individually-rational Pareto-optima in the presence of public goods, which is privacy respecting and which is individually incentive compatible for all agents, we follow the approach used by Hurwicz in [7, pp. 328-32]. Specifically, we construct a simple economy and show that no mechanism can meet these requirements when applied to this given economy. *A fortiori*, these requirements cannot be simultaneously met by any mechanism designed to work on a larger class of environments.

The economy we construct has two identical consumers, one private good, and one public good which can be produced from the private good under constant returns to scale. Then, by our choice of units, we can set $g(x) = x$. Each agent holds one unit of private good and has preferences which are given by the indifference map in figure 1.

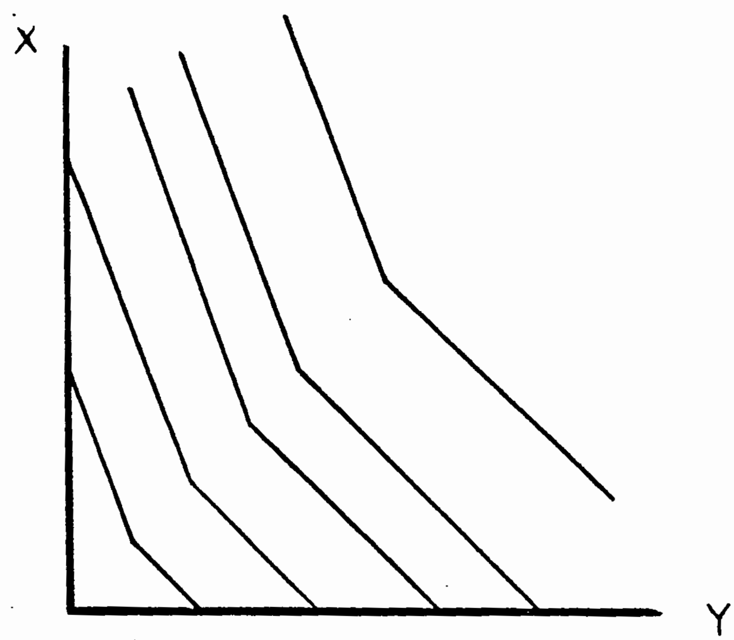


FIGURE 1

For $x < y$, the indifference curves have slope of -1, while for $x > y$, the slope is -3.[1]

It is convenient to represent this economy graphically by means of an analogue of the Edgeworth box diagram. This construction was used by Malinvaud [9], who attributes it to Kolm.

The equilateral triangle in figure 2 has height 2. Since the sum of the distances from any point in the triangle to the three sides is a constant, and since a feasible allocation (x,y^1,y^2) in this economy satisfies $x+y^1+y^2 = 2 = w^1+w^2$, there is a one-to-one correspondence between points in the triangle and the feasible allocations: using the point B as the origin for the first agent and C as that for the second, a point such as S corresponds to an allocation where x is the distance from S to BC, y^1 is the distance from S to AB and y^2 is the distance from S to AC. The initial position (0,1,1) is then the point W on BC. Sample indifference curves for the two agents are shown. Pareto optima correspond to "double tangencies", and thus the Pareto optima are the points along DEF.

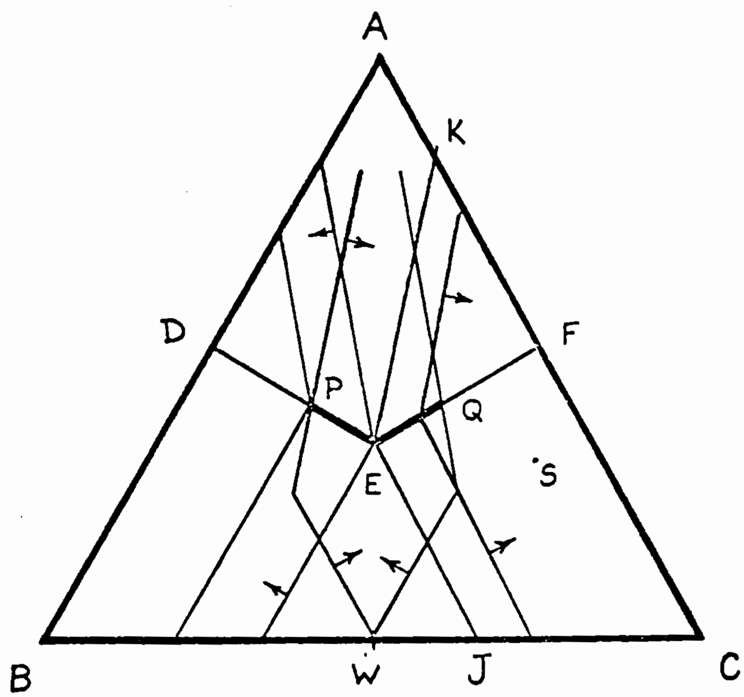


FIGURE 2

The points on PEQ are the Pareto optima in this economy which are preferred or indifferent for each agent to the initial allocation, W. We refer to the set of Pareto optima which are individually rational in this sense as the contract curve.

Any mechanism which selects allocations on the contract curve must select some point on PEQ if the agents reveal their true preferences. Suppose the outcome were on the segment PE. Then, if the second consumer reveals his true preferences, the first agent will be better off if he can, by misrepresenting his preferences, shift the apparent contract curve into the region to the right of JEK.

Clearly he can do this. For example, he can use the strategy which can be rationalized as being the true response of an agent with preferences given by straight line indifference curves with slope -3. This is illustrated in figure 3, where the apparent contract curve is now GT. Since the final allocation must be on GT, it is not individually incentive compatible for the first agent to reveal his true preferences (i.e., the g^{*i} strategies do not constitute a Nash equilibrium).

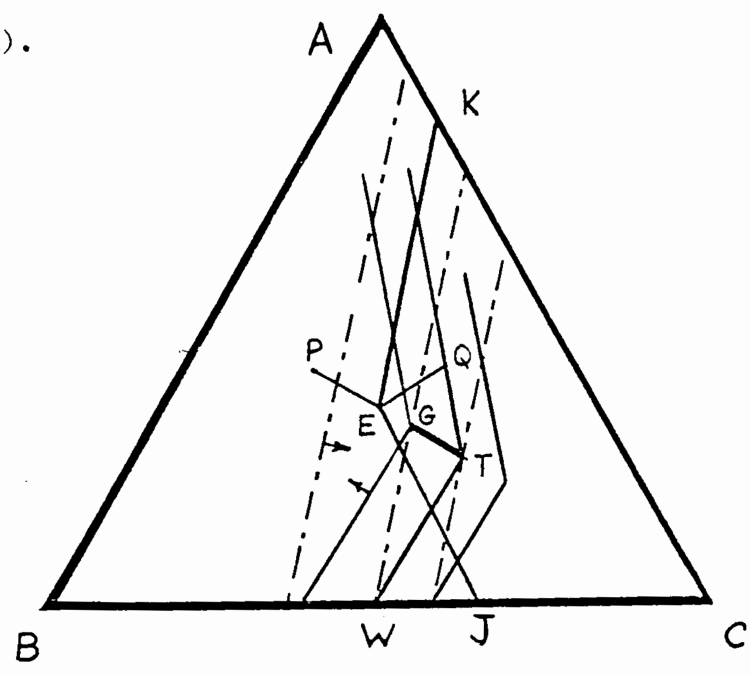


FIGURE 3

This result is, of course, what one would have expected: it ought not to be any easier to obtain incentive compatibility with public goods than in their absence. The question that now arises is, however, that of reconciling this result and those of Dreze and Vallee Poussin.

5. THE MDPH MECHANISM [2]

There is an obvious answer to the question of the relationship of these results. Since, relative to the revealed preferences, the MDP procedure generates paths of adjustment that converge monotonically in utility space to Pareto optima, our theorem indicates that their approach cannot enjoy incentive compatibility in the sense of Hurwicz. Thus the notions of incentive compatibility must, in fact, be different. To examine and illustrate this difference, however, it would be desirable to have a direct analysis of the global incentive properties of the Malinvaud-Dreze-Vallee Poussin approach. Unfortunately, the form in which the MDP procedure is specified, that of a differential equation system, does not lend itself to such an analysis.

One conceivable approach to resolving this difficulty would be to try to solve the system [(1),(2),(3)] of differential equations. This task is formidable even for the simplest cases, and we will not attempt it. Rather, we present a resource allocation mechanism of the Hurwicz type which we believe is a reasonable representation of the Malinvaud-Dreze-Vallee Poussin approach in this framework. This mechanism, which we call the MDPH mechanism, inherits the essential properties of the MDP procedure, and a reasonable computational procedure for employing the mechanism is identical with the MDP process. It may thus be of some independent interest. We then use this formulation to examine the global incentive properties of the MDP approach.

Formally, let $M = \{m^i \mid m^i : R^K_+ \to R\}$ be the language. We interpret $m^i(x)$ as the maximum amount of the private good consumer i is willing to pay for the public goods levels x, while $m^{N+1}(x)$ is the minimum amount of private good required by the production sector to produce x. Given the messages $m = (m^1, \ldots, m^{N+1})$

the planning board chooses an allocation of public goods, x, and a vector of taxes, $[c^1,\ldots,c^N]$. We assume that all profits of the production sector are collected by the board. The MDPH allocation rule is defined by:

(4) x(m) is the solution to the problem; choose $x \in R_+^K$ to maximize

$$\Sigma_{i=1}^N m^i(x) - m^{N+1}(x) .$$

The tax rule is, for i = 1,...,N,

(5) $c^i(m, x) = m^i(x) - \delta^i[\Sigma_{\ell=1}^N m^\ell(x) - m^{N+1}(x)]$.

Equation (4) requires the board to choose those levels of public goods which maximize the reported social surplus. Equation (5) assesses each consumer for the public goods level x at the rate he says he is willing to pay minus a dividend equal to a share δ^i of the reported surplus.[3]

The outcome rule, Φ, for the DPH process is defined by:

(6) $\Phi(m) = [x(m), y^1(m),\ldots,y^N(m)]$

where x(m) is given by (4) and $y^i(m) = w^i - c^i[m,x(m)]$.

We will continue to assume that producers are truthful, i.e. $m^{N+1} = g$. It is then easy to see that $\Phi(m)$ is feasible, since $\Sigma c^i[m,x(m)] = \Sigma_i m^i(x(m)) - \Sigma_i \delta^i[\Sigma_\ell m^\ell(x(m)) - g(x(m))] = g(x(m))$.

To complete a description of the process we must specify the response rules $f=[f^1,\ldots,f^{N+1}]$. For $N+1$, the producer, as indicated above, we let

(7) $f^{N+1}(m,e) = g$

where $g(x)$ is defined in section 1.

For consumer $i(i = 1,\ldots,N)$ we let

(8) $f^i(m, e^i) = \hat{m}^i$

where $\hat{m}^i(x)$ solves (for each $x \in R_+^K$) the equation $u^i[\bar{x}, w^i - c^i(m,\bar{x})] = u^i(x, w^i - \hat{m}^i(x))$, and $\bar{x} = x(m)$.

Note that f^i depends on m only through c^i and $\bar{x}$, the <u>current</u> allocation and taxing decisions, and on e only through $e^i = \langle u^i, w^i\rangle$ which implies the MDPH process is privacy respecting.

We consider a dynamic version of the MDPH process which is defined by: for $t \geq 0$,

(9) $m_t^i = f^i(m_{t-1}, e^i)$ for $i = 1,\ldots,N+1$,

(10) $s_t = [x_t, y_t^1,\ldots,y_t^N] = \Phi(m_t)$.

We can now indicate why we believe the MDPH mechanism to be adequate representation of the Malinvaud-Dreze-Vallee Poussin approach. First, consider the planning board of the MDPH model which is faced with solving, at each iteration t, the equations (4) and (5). Instead of computing the solutions directly, they could use some form of an iterative procedure. One which instantly comes to mind for solving the maximum problem in (4) is a gradient procedure. In particular, they might use the system of equations:

$$(11)\quad \dot{x}_k = \begin{cases} \Sigma_i\, m_k^i(x) - m_k^{N+1}(x) & \text{if } x_k > 0, \\ \text{Max}\,[\,0,\ \Sigma_i m_k^i(x) - m_k^{N+1}(x)] & \text{if } x_k = 0. \end{cases}$$

Once having decided on (11), they could then compute (5) by using:

for $i = 1,\ldots,N$,

$$(12)\quad \dot{c}^i = m_x^i(x)\cdot\dot{x} - \delta^i\,[\,\Sigma_i m_x^i(x)\cdot\dot{x} - m_x^{N+1}(x)\dot{x}]$$

Finally, they would realize that at any iteration they would need to know only $m_x^i[x(t)]$ and not the entire function $m_t^i(x)$. Thus, they could ask consumers to send only $\Psi^i(t) = m_{t,x}^i(x(t))$ where $m_t^i(x)$ solves $u^i[x(t),y^i(t)] = u^i[x,w^i - m_t^i(x)]$. This is equivalent to asking for $\Psi^i(t) = \pi^i[x(t), y^i(t)]$. Substituting $\Psi_t^i(t)$ for $m_{t,x}^i(x)$ and λ (t) for $m_{t,x}^{N+1}(x)=g_x(x)$ in (11) and (12) yields precisely equations (1) and (2), which are the MDP procedure. Thus, the MDPH mechanism reduces to the MDP procedure when viewed in this context or, alternatively [(1),(2)] can be thought of as an iterative (tatonnement) procedure for locating an equilibrium allocation[4] of the MDPH process [(6),(7),(8)].

Moreover, it is easily verified that a message m is an equilibrium of the MDPH mechanism if and only if the outcome $\Phi(m)$ is Pareto optimal and, further, that if $\Phi(m_t)$ is not an equilibrium, then $u^i(\Phi(m_{t+1})) > u^i(\Phi(m_t))$. [See appendix]. Thus, the essential properties of the MDP procedure, namely that it achieves optima and that utility is increasing for all agents at non-equilibria, hold for the MDPH mechanism.

We now examine the incentive properties of the MDPH mechanism. Since the outcome resulting from using the mechanism (and the MDP procedure) depends on the initial conditions we must treat two cases depending on whether or not $\Phi(m_0)$ is Pareto-optimal in the true economy.

Lemma 1: If $\Phi(m_0)$ is Pareto-optimal then $(g^{*1},\ldots,g^{*N})$ is a Nash-equilibrium of the game described in section 3. That is, under these conditions, the MDPH mechanism is incentive compatible in the sense of Hurwicz.

Proof: If one could select any (false) response function g^i (rather than only one such that $g^i(m) = f^i(m,\hat{e}^i)$ for some $\hat{e}^i$), then one would choose such that $\bar{g}^i(m) = \tilde{m}^i$ where $\tilde{m}^i(x)$ solves $\bar{u}^i = u^i[x, w^i - \tilde{m}^i(x)]$ where $\bar{u}^i = \max_x u^i[x, w^i + \Sigma_{h\neq i} m^h(x)]$. That is, one sends a message which absorbs all the other participants' stated consumers' surplus. It is easy to see that, if $\Phi(m_0)$ is Pareto-optimal then $\bar{g}^i(m_0) = f^i(m_0; e^i)$, which proves the lemma. QED

This result is another form of Theorem 4 of Dreze-Vallee Poussin. It is also consistent with Hurwicz' findings which require the initial position to be non-optimal if misrepresentation is to be profitable.

Lemma 2: If $(\bar{m}_0)$ is not Pareto-optimal, then the DPH mechanism is not incentive compatible.

Proof: We provide an example in which the equilibria of MDPH and MDP are the same. Assume there are two goods and that $g(x) = x$. Also let $u^i(x, y^i) = \alpha^i \ell n(x+1) + y^i$ for $i = 1,\ldots,N$. Given m_0, $x_0 = x(m_0)$ and $c_0^i = c^i(m_0, x(m_0))$, we have $f^i(m_0; e^i) = \bar{m}^i$, where:

(13) $\bar{m}^i(x) = \alpha^i \ \ell n\ (x+1) - \ell n\ (x_0+1) + c_0^i.$

We will restrict selections of false response functions to those which are consistent with the above economy except that α^i may be altered; that is, $\hat{e}^i = (\hat{u}^i, w^i)$ where $\hat{u}^i = \hat{\alpha}^i \ell n\ (x+1) + y^i$. The DPH procedure is not incentive compatible if it can be shown that $(\alpha^1,\ldots,\alpha^N)$ is not a Nash Equilibrium of the following game: $\{1,\ldots,N\}$ are the players, $A^i = [0,\infty)$ is the strategy space of i (where $\hat{\alpha}^i \in A^i$), and $V^i(\alpha)$ is the pay-off to i where $V^i(\hat{\alpha}) = u^i\ (x(\hat{\alpha}), w^i - c^i(\hat{\alpha}, x(\hat{\alpha})) = \alpha^i \ell n [\sum_\ell \hat{\alpha}^\ell] + w^i - \hat{\alpha}^i\ \ell n\ (\sum_\ell \hat{\alpha}^\ell) + \delta^i [(\sum_\ell \hat{\alpha}^\ell) \ell n\ (\sum_\ell \hat{\alpha}^\ell) - (\sum_\ell \hat{\alpha}^\ell)]$. This is the payoff to i if $c_0^i = x_0 = 0$ and if each $\ell = 1,\ldots N$ acts as if his u^ℓ is characterized by $\hat{\alpha}^\ell$. Consider $V^i_{\alpha^i}(\hat{\alpha}) = (\alpha^i / \sum_\ell \hat{\alpha}^\ell) - (1-\delta^i)\ \ell n\ (\Sigma \hat{\alpha}^\ell) - (\hat{\alpha}^i / \Sigma \hat{\alpha}^\ell)$. If α is a Nash equilibrium then $V^i_{\alpha^i}(\alpha) = 0$.

Assume that, for the true environment, $\Sigma\, \alpha^i > 1$. Then $V^i_{\alpha^i}(\alpha) = 0$ if and only if $(1 - \delta^i)\ \ell n(\Sigma \alpha^i) = 0$. But $\ell n(\Sigma \alpha_i) > 0$. Therefore, the true parameters, α, are not a Nash equilibrium (in fact the Nash equilibrium values of $\hat{\alpha}$ are such that $\hat{\alpha}^i < \alpha^i$. QED.

This lemma establishes that the Hurwicz result (referred to in the introduction) does apply to the MDP procedure as we would expect. That is if non-myopic strategic behavior is assumed, a participant in the MDP procedure can gain by sending messages $\psi^i(t)$ such that $\psi^i(t) = \hat{\alpha}^i/(x+1) \neq \Pi^i[x(t), y^i(t)] = \alpha^i/(x+1)$ and, therefore, has an incentive to misrepresent his preferences.

7. CONCLUDING COMMENTS

Hurwicz himself suggested that the contrasting results obtained by Dreze and Vallee Poussin and by himself might "be due to the local and instantaneous nature of the Dreze-Poussin pay-off function, since their criterion is whether $du^i/dt < 0$ for a participant departing from the prescribed strategy", while Hurwicz followed Samuelson in considering "the relevant pay-off to be the utility of the final (equilibrium) allocation" [7, p. 324n]. This insight is verified by Lemma 2 , which shows that when non-myopic strategic behavior is posited, the MDPH mechanism is not generally incentive compatible.

In general, one would expect that if individual tax rates are to be determined by the preferences revealed by the agents, then there will be a tendency to under-report marginal rates of substitution. An individual adopting a best replay strategy against the preferences announced by the other agents would only be willing to reveal preferences which would lead to production of the public goods at levels where his share of their marginal costs equaled his true marginal valuations of these goods. This would generally correspond to under-reporting of marginal valuations and will result in allocations which are not Pareto optimal. [5]

Malinvaud has considered the possibility that each participant in the D-P process might announce marginal rates of substitution at each instant t that correspond to a best replay (using the du^i/dt criterion) against the others' announced marginal rates. He notes that this will result in under-reporting, but concludes that the process will still converge to a Pareto optimum [10, pp. 110-111]. However, it is easy to see that the Nash equilibria with respect to the non-myopic Hurwicz criterion in the economy in Lemma 2 do not correspond to Pareto optima. Again, assuming non-myopic strategic behavior destroys either the efficiency of the process or its incentive properties (or both) enjoyed by the process under myopic behavior.

Since the local and global criteria for incentive compatibility lead to such divergent results, it is important to distinguish the two and to consider each in somewhat more detail. The following remarks are intended to open such a consideration and to suggest some of the factors we consider to be important in the analysis of incentives.

First, it seems that if one could establish a property of global incentive compatibility for some mechanism in some interesting class of situations, then one would have a very strong result. In fact, Groves and Loeb [4] have established an even stronger property for a system they developed for determining levels of and financing for public inputs, since they show that revealing one's true marginal valuations of these inputs is a best replay strategy no matter whether or not the other agents are being truthful. However, it would seem that, in general, global incentive compatibility may be too much to ask for both from the viewpoint of proving theorems (as Hurwicz's and our results here indicate) and from that of modeling behavior. A global criterion is certainly appropriate if the agents are able to predict the outcomes that will result from their choices of strategies. However, being able to make such predictions would often imply the agents' having very great amounts of information about the other agents, and vast computing ability. Not only would this be unrealistic to assume in many situations, but also it is somewhat at odds with the spirit of the basic assumption of the literature on resource allocation mechanisms that information about individual characteristics is initially decentralized. [6]

The local criterion is somewhat less susceptible to these criticisms, but, with its assumption that agents in the context of an iterative procedure will base their strategic choices on the utility of the next proposed allocation, this criterion is also surely less than completely satisfactory. Out of equilibrium, the proposed allocations are of little intrinsic interest, since they presumably will not be put into effect, and it would seem that the utility of what one actually will receive is the relevant object. If, on the other hand, equilibrium has been reached, then to disrupt the equilibrium by strategic behavior an agent must alter his response to the equilibrium values of the planning indicators and thus reveal an inconsistency that would easily be detected.

The problem, of course, is to develop some better criterion. This does not appear to be an easy task. However, it would seem to us that one fruitful direction for work is to attempt to recognize the uncertainty that an agent making strategic choices faces as a result of his lack of information about other agents' true characteristics and their strategic choices and about the workings of the mechanism. Indeed, some beginning on this task has been made by Ledyard [8] and by Gerard Varet and d'Apremont [5], using Harsanyi's construct of an incomplete information game. A special advantage of recognizing this uncertainty is that it should facilitate positive analysis of the nature of incentive compatible behavior in the context of various allocation systems.

APPENDIX

The following lemmata apply to the MDPH mechanism.

Lemma 3: If $\bar{m} = f(\bar{m}, e)$, then $\Phi(\bar{m})$ is Pareto-optimal in e.

Proof: By (8), $\bar{m}^i_x(x) = \Pi^i[\bar{x}, \bar{y}^i]$. By (7), $\bar{m}^{N+1}_x(\bar{x}) = g_x(\bar{x})$. Hence, by (4), $\Sigma_i \Pi^i_k(\bar{x}, \bar{y}^i) - g_k(\bar{x}) \leq 0$ for $k = 1,\ldots,K$ where $<$ holds only if $\bar{x}_k = 0$. Since u^i is quasi-concave, g is convex, and $\Phi(\bar{m})$ is feasible the result follows.

Lemma 4: If $\Phi(m_t)$ is Pareto - optimal, then $m_t = f(m_t, e)$, (i.e., m_t is an equilibrium.).

Proof: By (8), $m^i_{t+1,k}(x_t) = \Pi^i_k(x_t, y^i_t)$. By Pareto-optimality, $\Sigma_i \Pi^i_k(x_t, y^i_t) - g_k(x\ ) \leq 0$ where $<$ holds only if $x_{t,}\ = 0$. Therefore, $x(m_{t+1}) = x_t$. Also $m^i_{t+1}(x_t) = c^i(m_t, x\ )$. Hence it remains to show that $\Sigma_i m^i_{t+1}(x_t) - g(x_t) = 0$. But this follows easily since $\Sigma_i c^i(m_t, x_t) = g(x_t)$ since $\Phi(m_t)$ is feasible. Therefore, $c^i[m_{t+1}, x_t] = c^i[m_t, x_t]$ and the lemma is proven. QED

Lemma 5: If $\Phi(m_t)$ is not Pareto-optimal in e and $m_{t+1} = f(m_t, e)$, then $u^i[\Phi^i(m_{t+1})] > u^i[\Phi^i(m_t)]$ for all $i = 1,\ldots,N$ and all $t \geq 0$.

Proof: If $\Phi(m_t)$ is not optimal, $\exists\ \hat{s} = (\hat{x}, \hat{y}^1, \ldots, \hat{y}^N)$ such that $u^i(\hat{x}, \hat{y}^i) \geq u^i(x_t, w^i - c^i_t)$ for all i and $>$ holds for some i. By (8), $u^i[\hat{x}, w^i - m^i_{t+1}(\hat{x})] = u^i[x_t, w^i - c^i_t] \leq u^i(\hat{x}, \hat{y}^i)$. Therefore, since $u^i_y > 0$, $\hat{y}^i \geq w^i - m^i_{t+1}(\hat{x})$. Summing over i, we find that $\Sigma\, m^i_{t+1}(\hat{x}) - g(\hat{x}) \geq 0$ since $\Sigma_i(\hat{y}^i - w^i) + g(\hat{x}) = 0$. Now, $u^i[\Phi^i(m_{t+1})] = u^i[x_{t+1}, w^i - c^i_{t+1}] = u^i[x_{t+1}, w^i - m^i_{t+1}(x_{t+1}) + \delta^i(\Sigma\, m^i_{t+1} - g(x_{t+1}))] > u^i(x_{t+1}, w^i - m^i_{t+1}(x_{t+1})) = u^i[\ x_t, w^i - c^i_t\ ] = u^i[\Phi^i(m_t)]$. QED

FOOTNOTES

1. These preferences, although convex, monotone and locally non-satiated, are not strictly convex or smooth. However, as will be seen, imposing these conditions would create no essential difficulties although it would complicate computations.
2. This section depends heavily on the work of Groves [2] and Groves and Ledyard [3].
3. See [1, p. 139, bottom paragraph] for the origin of these taxing rules.
4. It is unlikely that [(1), (2)] will lead to the same equilibrium as [(9), (10)] due to the path-dependence of the solutions. However in the example to be presented later, the equilibra are identical.
5. See, however, Groves and Ledyard [3], where a mechanism for allocating resources in the presence of public goods is presented which has the property that if individuals take private goods prices as given, then the allocation resulting from a Nash equilibrium of reported preferences is Pareto-optimal.
6. Hurwicz has argued, however, that one should distinguish the information needed by an analyst to determine whether a given set of strategies constitute a Nash equilibrium and that needed by a player to upset a non-equilibrium position. If this point is accepted, then to the extent that the latter requires less information, our criticisms are somewhat blunted.

BIBLIOGRAPHY

[1] Dreze, J. and D. de la Vallee Poussin, "A Tatonement Process for Public Goods", Review of Economic Studies, October, 1971.

[2] Groves, T., "Incentives in Teams", Econometrica

[3] Groves, T., and J. Ledyard, "An Incentive Mechanism for Efficient Resource Allocation in General Equilibrium with Public Goods", Center for Mathematical Studies in Economics and Management Science, Discussion Paper No. 119, Northwestern University 1974.

[4] Groves, T. and M. Loeb, "Incentives and Public Inputs", Journal of Public Economics, forthcoming.

[5] Gerard Varet, L., and C. d'Aspremont, "Cost-sharing and Incentives in a Pollution Game with Incomplete Information", Document de travail CB - 1/12, CORE, 1974.

[6] Hurwicz, L., "Optimality and Informational Efficiency in Resource Allocation Processes", in Mathematical Methods in the Social Sciences, Arrow, Karlin and Suppes (editors), Stanford University Press, Stanford, 1960.

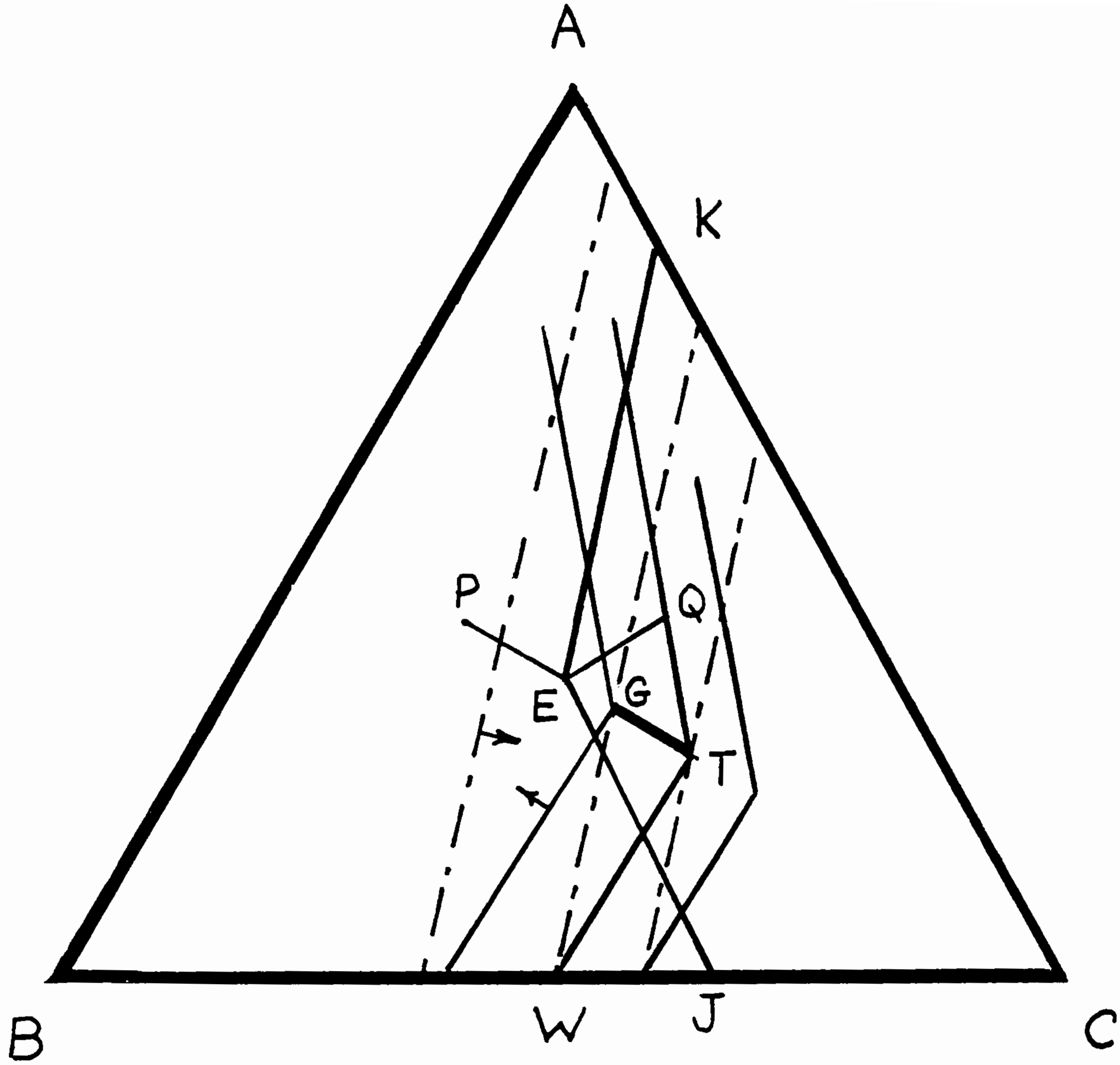
[7] Hurwicz, L., "On Informationally Decentralized Systems", in Decision and Organization, Radner, R., and B. McGuire (editors), North Holland Press, Amsterdam, Holland, 1972.

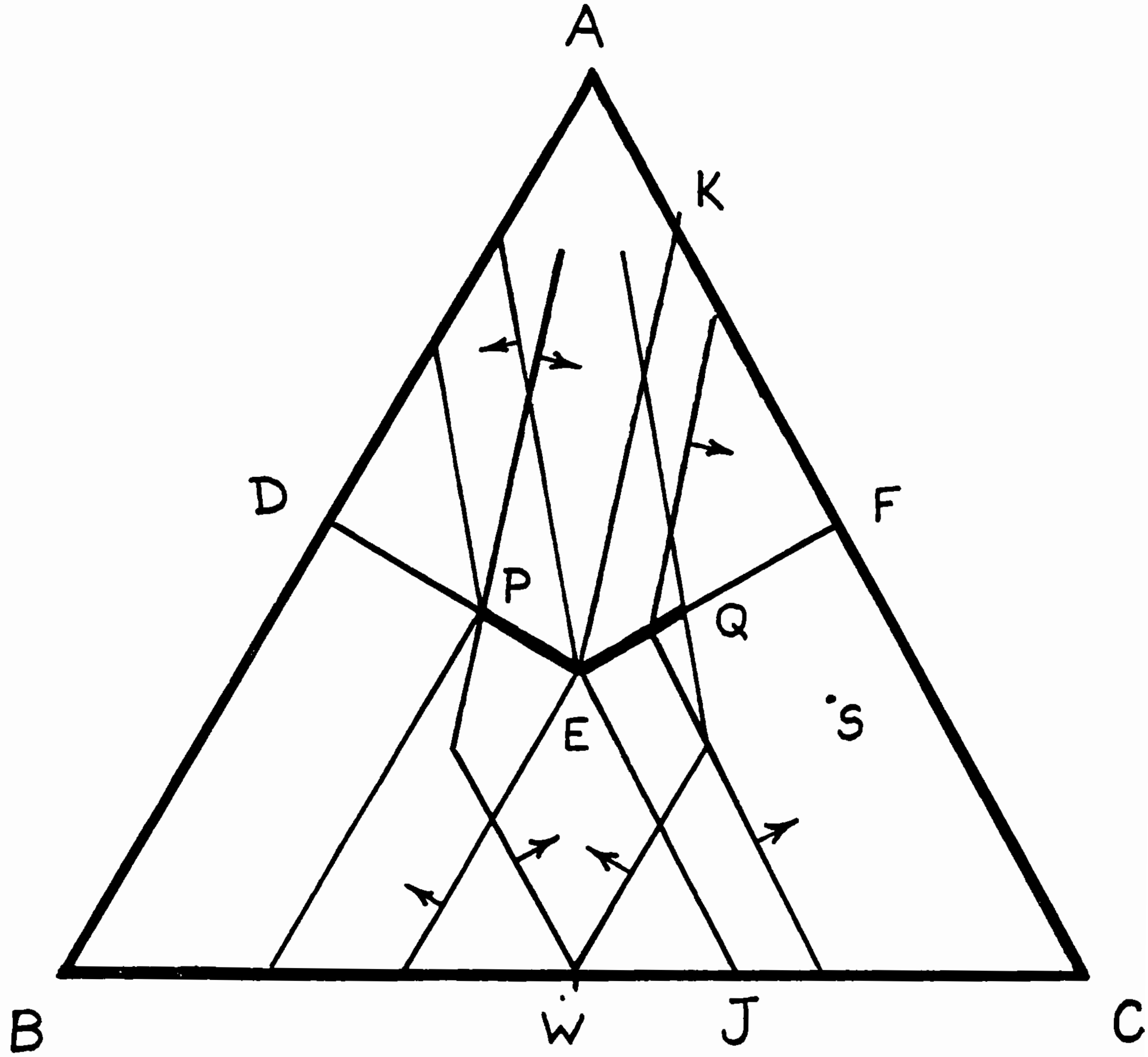
[8] Ledyard, J., "The Incentive Properties of the Competitive Mechanism with Incomplete Information", working paper.

[9] Malinvaud, E., "A Planning Approach to the Public Goods Problem", Swedish Journal of Economics, March 1971.

[10] Malinvaud, E., "Procedures for the Determination of a Program of Collective Consumption", European Economic Review, Winter 1970-71.

[11] Samuelson, P. A., "The Pure Theory of Public Expenditures", Review of Economics and Statistics, 1954.


A
K
P
Q
E
G
T
B
W
J
C


A
K
D
F
P
Q
E
S
B
W
J
C