DISCUSSION PAPER NO. 264

INCENTIVES IN PLANNING PROCEDURES FOR THE PROVISION OF PUBLIC GOODS

by

John Roberts

January 6, 1977

Incentives in Planning Procedures for the Provision of Public Goods*

by John Roberts

I. Introduction

In the past few years several studies have addressed the problem of designing decentralized mechanisms for determining an efficient allocation of resources in the presence of public goods. A crucial feature of such mechanisms is the incentives that are provided for the participants to supply correct information about their preferences, since, unless proper incentives are present, self-interested behavior may frustrate the realization of efficiency. As has long been recognized, the study of such self-interested, strategic behavior and of incentive problems in general must be essentially game-theoretic in nature. However, there are many possible specifications of the game that arises when agents under such a mechanism can decide whether or not to reveal their preferences correctly. Each such specification yields a different formalization of such concepts as self-interest and incentive compatibility. Correspondingly, very different patterns of behavior may emerge under different formulations of the game of preference revelation.

*I would like to thank John Ledyard, Don Saari, and the Editor of the Review of Economic Studies for their valuable suggestions and assistance, Henry Tulkens for his comments on an earlier draft, and Claude Henry and Françoise Schoumaker for giving me access to their unpublished work. The first version of this paper was prepared while I was at CORE in the spring of 1976 under the auspices of the Ford Foundation Program for Research in International Economic Order. My research was supported by the National Science Foundation under grants SOC 74-04076 and SOC 76-20953.

In this note we will be concerned with one aspect of this problem, namely, the specification of the payoff functions in the context of planning procedures with public goods. The games we will consider will be non-cooperative ones in which the players are the agents (consumers) in the process and the strategy sets consist of messages the agents can send regarding what their preferences are. (These messages may, in fact, be complete preference orderings.) This much is more or less common to most studies of incentive questions. However, in many iterative planning procedures, a proposed allocation is generated at each iteration. This gives rise to a choice as to the payoff functions in the game. On the one hand, one might hypothesize that agents will be concerned only with what they will finally receive. This leads to specifying the payoff to each player as the utility to him under his true preferences of the final allocation selected. We will refer to the game with this payoff as the <u>global</u> game. Alternatively, one might argue that it is more reasonable to suppose that agents simply try to do as well as they can for themselves at each iteration. This then leads to consideration of a separate game at each iteration, with the payoffs being the change in the utilities associated with the adjustment of the proposed allocation. We call games with this payoff <u>local</u> or <u>instantaneous</u>.

Both these alternatives have been considered in the literature. Samuelson's initial discussion of the public good problem would seem to be in terms of the former, global criterion. Hurwicz adopted this approach in his pathbreaking analysis of incentives [6] and most of the rapidly growing literature on the topic has followed this line.

On the other hand, Drèze and de la Vallée Poussin [3] and Malinvaud [8] have employed the local or instantaneous game in their analyses of the incentives under their planning procedure for public goods. Use of these different modelings has led to strongly contrasting results. Hurwicz showed that even in the context of simple exchange economies there could not exist a mechanism for selecting individually rational Pareto optima which was incentive compatible in the sense that correct revelation was always a Nash equilibrium in the global game. On the other side, Drèze and de la Vallée Poussin showed that under correct revelation their procedure selected individually rational Pareto optima with public goods and that at an equilibrium of the system correct revelation was the only Nash equilibrium in the local game. (They also showed that correct revelation was always a maximin strategy in these games, even out of equilibrium.)

After presenting basic definitions and setting up the model in Section 2 of this paper, we examine further the incentives under the Malinvaud-Drèze-de la Vallée Poussin (MDP) procedure using the local or instantaneous game formulation. This is done in Section 3, where we obtain the best replay strategies for the players in these games, solve these for the Nash equilibria and then consider the resulting path of adjustment under the process. It turns out that although correct revelation is a Nash equilibrium only at system equilibrium, equilibrium misrepresentation at each iteration does not keep the procedure from converging to an individually rational Pareto optimum. In Section 4 we turn to the global game and present an extension of Hurwicz's impossibility theorem to the public goods case. The final section consists of a brief discussion of some possible extensions and some concluding remarks.

2. The Model

To keep matters simple, we will work with economies with one public good, whose quantities are denoted by y, one private good, denoted by x, and N consumers. Each consumer i is characterized by a preference order $\succsim^i$ over R^2_+, the non-negative orthant of R^2, and by a non-negative endowment w^i of private good. The possibilities for producing the public good from the private are described by a production set Y. In Section 3 especially, we will adopt Drèze and de la Vallée Poussin's assumptions that each $\succsim^i$ is representable by a utility function U^i which is strictly concave and three times continuously differentiable, with $U^i_x \equiv \partial U^i(x,y)/\partial x > 0$, $U^i_y \equiv \partial U^i(x,y)/\partial y \geq 0$, and $U^i_y(0,y)/U^i_x(0,y) = 0$ for all x,y. In the same context, we assume that the production possibilities are described by a function G(y) which is also C^3, with $G'(y) \equiv \gamma(y) > 0$, $G''(y) < 0$.

The assumption of one public good is made largely as a matter of convenience, as will be argued in Section 5. The assumption of a single private good is, however, not so easily relaxed (see [3], pages 148-149, and [9]).

Throughout, superscripts denote agents. Given a vector $s = (s^1,\ldots,s^N)$ we will write Σs^j for $\Sigma_{j=1}^N s^j$, Πs^j for the product of s^1 through s^N, $s^{)i(}$ for $(s^1,\ldots,s^{i-1},s^{i+1},\ldots,s^N)$, $\Sigma_{)i(}s^j$ for the sum and $\Pi_{)i(}s^j$ for the product of the components of $s^{)i(}$, and $(s^{)i(},s^*)$ for $(s^1,\ldots,s^{i-1},s^*,s^{i+1},\ldots,s^N)$.

A feasible allocation is represented by a vector $(x^1,\ldots,x^N,y)$ such that $(\Sigma w^i - x^i, y) \in Y$, i.e., $\Sigma (w^i - x^i) = G(y)$. A feasible allocation $(x^1,\ldots,x^N,y)$ is individually rational if $(x^i,y) \succsim^i (w^i,0)$, $i=1,\ldots,N$, and Pareto optimal if there does not exist a feasible allocation $(\bar{x}^1_j,\ldots,\bar{x}^N,\bar{y})$ with $(\bar{x}^i,\bar{y}) \succsim^i (x^i,y)$ for all i and $(\bar{x}^j,\bar{y}) \succ^j (x^j,y)$ for some j.

An N-person game is a triple (I,S,v), where $I = \{1,\ldots,N\}$ is the set of players, S is the set of jointly admissible strategies and $v = (v^1,\ldots,v^N)$, $v^i\colon\ S \to R$, are the pay-off functions. We will assume throughout that $S = S^1 \times S^2 \times \ldots \times S^N$, and denote elements of S by vectors $(s^1,\ldots,s^N)$. A best replay function for player i is a function $h^i\colon\ S \to S^i$ such that, for all $s^* \in S^i$, $v^i(s^{)i(}, h^i(s)) \geq v^i(s^{)i(}, s^*)$. An N-tuple $s = (s^1,\ldots,s^N) \in S$ is a Nash equilibrium if $s^i = h^i(s)$, $i = 1,\ldots,N$.

3. The Local Incentive Game Under the MDP Procedure

The MDP procedure is a continuous time planning process under which the proposed allocation at time t is adjusted according to the marginal cost of the public good and information provided by the participants regarding their marginal rates of substitution. (See Drèze and de la Vallée Poussin [3] and Malinvaud [8] for further discussion). Given an allocation $(x^1_t,\ldots,x^N_t,y_t)$ at time t and announced values $\psi^1_t,\ldots,\psi^N_t$ of the marginal rates of substitution, the adjustment in y at time t is given by

$$\frac{dy}{dt} = \begin{cases} \Sigma\,\psi^j_t - \gamma_t, & y > 0 \\ \max\,[0,\ \Sigma\,\psi^j_t - \gamma_t], & y = 0, \end{cases}$$

$$\frac{dx^i}{dt} = -\,\psi^i_t\,\frac{dy}{dt} + \delta^i\left(\frac{dy}{dt}\right)^2, \qquad i = 1,\ldots,N, \tag{1}$$

$$= \begin{cases} -\psi^i_t(\Sigma\,\psi^j_t - \gamma_t) + \delta^i(\Sigma\psi^j_t - \gamma_t)^2, & y > 0 \\ -\psi^i_t(\max[0,\Sigma\psi^j_t - \gamma_t]) + \delta^i(\max[0,\Sigma\psi^j_t - \gamma_t])^2, & y = 0 \end{cases}$$

where $\gamma_t = \gamma(y_t) \equiv G'(y_t)$ is the marginal rate of transformation and the δ^i are positive constants, the "distributional coefficients", whose sum is one.

If the ψ_t^i are determined by the reference allocation at time t, then these equations specify a differential equation system. A solution from a given initial position $(x_0^1,\ldots,x_0^N,y_0)$ for this system is a continuous function from R_+ to R_+^{N+1} such that $(x^1(0),\ldots,x^N(0),y(0)) = (x_0^1,\ldots,x_0^N,y_0)$ and whose derivative on the right at each t is given by the equations of (1). A stationary point of the system in these circumstances is an allocation such that $\frac{dy}{dt} = \frac{dx^1}{dt} = \ldots = \frac{dx^N}{dt} = 0$. We will sometimes refer to such a point as an equilibrium of the system.

Define $\pi^i(x^i,y) \equiv U_y^i(x^i,y)/U_x^i(x^i,y)$. Consumer i is correctly revealing his preferences at time t if $\psi_t^i = \pi^i(x_t^i,y_t)$, where (x_t^i,y_t) is his consumption under the current reference allocation. Assuming correct revelation by all consumers for all t, Drèze and de la Vallée Poussin ([3], Section 2) showed that the system has a solution from any feasible initial allocation and that any solution to the system converges to a Pareto optimum. Moreover, again assuming correct revelation, if $(x^i(0),y(0)) \succsim^i (w^i,0)$ for all i, then any equilibrium is individually rational since $\dot{U}_t^i \equiv dU^i(x^i(t),y(t))/dt \geq 0$.

Drèze and de la Vallée Poussin considered the problem of the incentives for agents to reveal their preferences correctly under this process. They established two principal results in this connection. First, they showed that correct revelation is a maximin strategy in the game arising at each instant t in which $v^i = \dot{U}^i(x_t^i,y_t)$, and that for $y_t > 0$ it is the only such strategy ([3], Theorem 3). This is so because correct revelation guarantees that $\dot{U}_t^i$ is non-negative, while for any ψ_t^i announced by i there exists some

$\psi_t^{)i(}$ that can hold i to a non-positive payoff; moreover if $y > 0$ and $\psi_t^i \neq \pi_t^i$, then $\dot{U}_t^i$ can be made negative. Second, they showed that at a stationary point of the process with $y > 0$ correct revelation is the only Nash equilibrium ([3], Theorem 4). Intuitively this may be seen as follows. If one is at an equilibrium of the process and some agent is not revealing his preferences correctly, then by announcing $\pi^i(x^i,y)$ he could start the process in motion again and, since $\delta^i > 0$, he would share in the resulting social surplus. On the other hand, a stationary point with correct revelation is a Pareto optimum. It is thus impossible to increase one person's utility without decreasing that of someone else. But this in turn cannot happen, since $\psi^j = \pi^j$ insures that $\dot{U}^j$ will be non-negative.

These results on Nash equilibria in the local incentive game hold only at equilibria of the system. As Drèze and de la Vallée Poussin note, their results have little to say regarding the nature of best replay strategies away from system equilibrium, although they do note that if the system converges at all under misrepresentation it must still go to an optimum. In this regard, Malinvaud [8] suggested that convergence would still occur if at each t the ψ_t^i define a Nash equilibrium, but he did not supply a proof for this. In the remainder of this section we consider the nature of the Nash equilibria in the instantaneous games and the convergence of the resulting allocations.

Specifically, consider the game at each t in which the strategy for each of the N players is the choice of $\psi_t^i \in R$ and in which the payoff to i is

$$v^i(\psi^1,\ldots,\psi^N) = U_y^i[\Sigma\,\psi^j-\gamma] + U_x^i[-\psi^i(\Sigma\,\psi^j-\gamma)+\delta^i(\Sigma\,\psi^j-\gamma)^2]$$

where U_y^i, U_x^i and γ are evaluated at $(x_t^1,\ldots,x_t^N,y_t)$. Thus, v^i is simply the rate of change in U^i at time t if the adjustments in y_t and x_t^i are given by the MDP equations and $\psi^1,\ldots,\psi^N$ are the announced marginal rates of substitution.

Note that we are assuming here that $y_t > 0$. In fact, it is easy to construct examples in which if $y = 0$ then, even though $\Sigma\,\pi^j > \gamma$, it is a Nash equilibrium for each agent to under-report his MRS sufficiently that $\Sigma\,\psi^j < \gamma$. Such a point would then be a stationary point but not an optimum. For example, let $\gamma(0) = 1$, $\pi^1(w^1,0) = \pi^2(w^2,0) = 3/4$ and $\psi^1 = \psi^2 < 1/4$. However, if optimality involves positive amounts of the public good, then such problems are easily avoided by starting with $y > 0$, and we will assume throughout this section that the public good's level is positive. Note too that we allow ψ^i to be negative. We will return briefly to these matters in Section 5.

To obtain the best replay strategies, denoted Φ^i, we maximize v^i with respect to ψ^i. This yields

$$\Phi^i = h^i(\psi^1,\ldots,\psi^N) = \frac{1}{2(1-\delta^i)}[\pi^i-(1-2\delta^i)(\Sigma_{)i(}\psi^j-\gamma)].$$

We note immediately that if $\delta^i = 1/2$, then correct revelation is a dominant strategy for i, that is, the optimal choice for him no matter what the other consumers say. Since this case may have some claim to a central position when N = 2, it is perhaps of some interest. However, more generally, when $\delta^i < 1/2$ we readily verify that[1]

$$\Phi^i \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} \pi^i \quad \text{as} \quad \Sigma_{)i(}\psi^i + \pi^i - \gamma \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} \quad 0.$$

Thus correct revelation is a best replay if and only if the agent wishes no adjustment in the allocation. Otherwise, if he wants y increased he will under-report his MRS and if he wants y decreased he will over-report. However, it is simply a matter of substitution to show that $\Sigma_{j\neq i}\psi^{j} + \Phi^{i} - \gamma =$ $[\Sigma_{j\neq i}\psi^{j} + \pi^{i} - \gamma]/2(1-\delta^{i})$. Thus, the change in y that would result from the agent's best replay is a fraction (less than one if $\delta^{i} < 1/2$) of what it would have been if he told the truth. Note that he does not misrepresent enough to reverse the sign of the adjustment in the public good from that which he desires; instead, he merely acts to reduce it.

The best replay functions define a system $(I + A)\psi = \pi + \gamma a$ of N linear equations, where $a' = (a^{1},\ldots,a^{N})$, $a^{i} = (1-2\delta^{i})$, $\pi' = (\pi^{1},\ldots,\pi^{N})$, I is the $N \times N$ identity matrix and A is the matrix each of whose columns is the vector a. Inverting, we obtain $(I + A)^{-1} = I - A/(1 + \Sigma a^{j})$. But note that $\Sigma a^{j} = N - 2$, so that the inverse is simply $I - A/(N-1)$. Note too that for any vector z, $Az = (\Sigma z^{j})a$. Then, to obtain the Nash equilibrium vector Φ as a function of π, we simply evaluate:

$$
\begin{aligned}
\Phi &= (I + A)^{-1}(\pi + \gamma a), \\
&= (I - A/(N-1))(\pi + \gamma a), \\
&= \pi + \gamma a - \frac{1}{(N-1)}(\Sigma\pi^{j} + \gamma\Sigma a^{j})a, \\
&= \pi - \frac{1}{(N-1)}(\Sigma\pi^{j} - \gamma)a.
\end{aligned}
$$

Thus, the Nash equilibrium strategies exist and are unique and are given by

$$\Phi^{i} = \pi^{i} - \frac{(1-2\delta^{i})}{(N-1)}(\Sigma\pi^{j} - \gamma). \tag{2}$$

We observe immediately that for $y > 0$, then correct revelation is a Nash equilibrium in this local game if, and only if, the system would be at a stationary point under correct revelation or $\delta^i = \frac{1}{2}$ for all i. With the result of Drèze and de la Vallée Poussin noted earlier, this completely characterizes the relationship between Nash equilibria and correct revelation.

Now, the change in y if the agents use their Nash equilibrium strategies is just $\Sigma \Phi^i - \gamma$. This is easily seen to equal $(\Sigma \pi^j - \gamma)/N-1$. Thus, the adjustment in y from any given allocation is just 1/(N-1) times what it would have been had the agents revealed their preferences correctly. The corresponding adjustment in x^i is then given by

$$-\Phi^i \left(\frac{\Sigma \pi^j - \gamma}{N-1}\right) + \delta^i \left(\frac{\Sigma \pi^j - \gamma}{N-1}\right)^2$$

$$= -\left[\pi^i - \frac{(1-2\delta^i)}{N-1} (\Sigma \pi^j - \gamma)\right] \left(\frac{\Sigma \pi^j - \gamma}{N-1}\right) + \delta^i \left(\frac{\Sigma \pi^j - \gamma}{N-1}\right)^2$$

$$= \frac{1}{N-1}\left[-\pi^i(\Sigma \pi^j - \gamma) + \frac{(1-\delta^i)}{N-1} (\Sigma \pi^j - \gamma)^2\right] .$$

However, note that $0 < 1-\delta^i < 1$ and that $\Sigma(1-\delta^i) = N-1$. Thus, if we define $\sigma^i \equiv (1-\delta^i)/(N-1)$, we may write the system that emerges under the Nash equilibrium misrepresentation in the local game as

$$\frac{dy}{dt} = \frac{1}{N-1}[\Sigma \pi^j - \gamma],$$

$$\frac{dx^i}{dt} = \frac{1}{N-1}[-\pi^i(\Sigma \pi^j - \gamma) + \sigma^i(\Sigma \pi^j - \gamma)^2]. \tag{3}$$

Once we make the obvious change of variable this becomes, of course, exactly the MDP system (1) with correct revelation and distributional parameters σ^i. Thus, the effect of self-interested misrepresentation is merely to slow the speed of adjustment of the system and to alter the distribution of the surplus. This means, in particular, that under the assumptions made in Theorem 1 of [3], the system still is stable and still converges to a Pareto optimum.

In general, this will not be the same optimum as would result under correct revelation. However, two points are worth noting. First, if $\delta^i = 1/N$, which might well be considered to be distributionally just, then $\sigma^i = \delta^i$, and the paths of adjustment and system equilibrium are totally immune to misrepresentation. Secondly, since the mapping from δ to σ is one to one and onto, if the authorities have certain desired values δ^{i*} for the distribution parameters (see Champsaur [1]) and correctly expect the agents to adopt the strategic behavior studied here, they can nullify the effect of this misrepresentation on the final outsome by setting $\delta^i = 1 - (N-1)\ \delta^{i*}$, which is positive if δ^{i*} is greater than $1/(N-1)$.

Of course, the significance of these results depends to a great extent on whether one might expect the agents actually to reach the Nash equilibrium strategies (2) at each instant t. This leads to consideration of the stability of these Nash equilibria, which in turn requires specifying a dynamic system to describe the adjustment of the vector ψ away from Nash equilibrium. Within the context of a continuous time process for adjusting the proposed allocation, it seems appropriate also to specify the adjustment process for the announced MRS's at any instant as a continuous time one. The two most frequently considered adjustment processes of this type are that leading to the gradient system

$$\frac{d\psi^i}{ds} = k^i \frac{\partial v^i}{\partial \psi^i}, \qquad k^i > 0, \qquad i=1,\ldots,N,$$

and the "Cournot process"

$$\frac{d\psi^i}{ds} = k^i [h^i(\psi) - \psi^i], \quad k^i > 0, \quad i=1,\dots,N.$$

Since the k^i are arbitrary positive numbers and U_x^i is fixed in this context, the stability properties of either of these systems are the same as those of the homogeneous system

$$\frac{d\psi^i}{ds} = -k^i [\psi^i + \frac{(1-2\delta^i)}{2(1-\delta^i)} \Sigma_{j \neq i} \psi^j], \quad i = 1,\dots,N.$$

Consider the positive definite quadratic form $V(\psi) = \Sigma(1-\delta^i)(\psi^i)^2/k^i$.

Then

$$\begin{aligned}\frac{dV(\psi)}{ds} &= \Sigma\, 2\frac{(1-\delta^i)}{k^i} \psi^i \frac{d\psi^i}{ds} \\ &= -\Sigma (\psi^i)^2 - \Sigma (1-2\delta^i) \psi^i \Sigma \psi^j \\ &= -\Sigma (\psi^i)^2 - (\Sigma \psi^i - \Sigma\delta^i \psi^i)^2 + (\Sigma\delta^i \psi^i)^2 \\ &\leqq -\Sigma (\psi^i)^2 + (\Sigma\delta^i \psi^i)^2 \\ &\leq -\Sigma (\psi^i)^2 + \Sigma(\delta^i)^2 \Sigma(\psi^i)^2\end{aligned}$$

Since $0 < \Sigma(\delta^i)^2 < 1$, we thus have $dV/ds \leq 0$. Further, $dV(\psi)/ds < 0$ if $\psi \neq 0$, while $dV(0)/ds = 0$. Thus, V is a Lyapunov function and the Nash equilibrium is globally asymptotically stable if the adjustment of ψ is as specified.

On the other hand, if at each instant t we specify the discrete time process $\psi_{s+1}^i = h^i(\psi_s)$, i=1,...,N, as giving the path of adjustment of the announced marginal rates of substitution, then the system will not be stable. However, it is worth noting in this connection that the discrete time Cournot adjustment system specified here may be a very poor model of the actual process of revision of strategies. This point is discussed by Vernon Smith [11] in the context of his experiments with procedures designed to find an optimum with public goods. In these experiements, participants's strategies converged to the (Pareto optimal) Nash equilibrium even though this equilibrium would have been unstable if participants had maximized at time s+1 against the others' strategies at time s.

4. The Global Criterion: An Impossibility Result

The results in Section 3 indicate that if one models the incentive problem via the local or instantaneous game, then misrepresentation is not a serious problem in the continuous time MDP process. In this section we investigate the implications of using the utility of the final outcome as the payoff in the game.

In this context it is convenient to summarize the workings of the resource allocation system by a mapping f which assigns to each economy E the resulting final allocation f(E). Here E is described, as before, by the consumer preference orderings and endowments and by the production possibilities. Now, by behaving consistently as if his preferences were given by some ordering $\succsim^*$ different from his true ordering $\succsim$, consumer i can make the economy E actually appear to be another economy E^* which is, of course, obtained from E by specifying the preferences of consumer if to be $\succsim^*$ rather than $\succsim$. If the resource allocation system is sensitive to individual preferences, it may well be that the consumption that i receives when he makes the economy appear to be E^* is preferred by him under $\succsim$ to what he would have received revealing his true preferences. This gives rise to the incentive problem of preference misrepresentation.

This framework leads to consideration of a game for each resource allocation system f and each economy $E = ((\succsim^i, w^i)_{i=1}^N, Y)$. Specifically, we consider the game in which the players are the N consumers, the strategy set S^i for each i is a set of admissible preference orderings, and the payoff to i from a particular strategy N-tuple $(s^1,\ldots,s^N)$ is $U^i(x^i,y)$, where $(x^1,\ldots,x^N,y) = f((s^i,w^i)_{i=1}^N,Y)$ and U^i represents $\succsim^i$. In this context, we say that f is individually incentive compatible if for every E the N-tuple $(\succsim^1,\ldots,\succsim^N)$ of true preference orderings in E is a Nash equilibrium of the game corresponding to f and E.

The fundamental result in this area is due to Hurwicz [6], who showed that any f which always selects an individually rational Pareto optimum in pure exchange economies cannot be individually incentive compatible.

Ledyard and Roberts [7] extended Hurwicz's result to economies with public goods, also by presenting an economy in which no f yielding individually rational Pareto optima could result in correct revelation as a Nash equilibrium in the global game. However, their example involved non-differentiable preferences and constant returns to scale. While these features do not detract from the significance of their result, they do violate the Drèze-de la Vallée Poussin assumptions made in Section 3. Here we will consider an example which meets these conditions.

Consider the economy E with N = 2 in which $w^i = 5/2$ and $U^i(x^i,y) = x^i + 2\ \ln(y+1)$, $i = 1,2$, and in which $G(y) = y^2$. It is readily verified that all Pareto optima in this economy involve $y = 1$, so the contract curve (the set of individually rationa optima) consists of those allocations $(x^1,x^2,y) \in R^3_+$ such that $y = 1$, $x^1 + x^2 = 4$, and $x^i \geq 5/2 - 2\ \ln 2$. Under correct revelation, any f selecting individually rational Pareto optima must select a point f(E) on this contract curve. In particular, given any specification of the δ^i, the MDP procedure would have to converge under correct revelation to such an allocation, provided y(0) is selected positive and $x^i(0) + 2\ \ln(y(0) + 1) \geq 5/2$. Suppose then that the equilibrium under correct revelation is (x^1,x^2,y), where $y = 1$, and $5/2 - 2\ \ln 2 \leq x^1 \leq 2$. If consumer 1 were to have claimed that his preferences were actually given by $U^{i*}(x^i,y) = x^i + (2/9)\ \ln(y+1)$ and had consistently announced $\psi^i_t = 2/9\ (y_t+1)$, while the other reported truthfully, then the apparent contract curve would be given by $y = 2/3$, $x^1 + x^2 = 41/9$, $x^1 + (2/9)\ \ln(5/3) \geq 5/2$, $x^2 + 2\ \ln(5/3) \geq 5/2$. The equilibrium of the MDP procedure (or the outcome of any other system selecting individually rational Pareto optima) would then necessarily lie in this set. The worst point in this set for consumer 1 is $x^1 = 5/2 - (2/9)\ \ln(5/3)$,

$y = 2/3$. Thus, the worst he will obtain under this misrepresentation is a utility level of $x^1 + 2 \ln(5/3) = 2.5 - (2/9) \ln(5/3) + 2 \ln(5/3) = 3.40814$, while (by hypothesis) under correct revelation he receives at most a utility level of $2 + 2 \ln 2 = 3.38630$. Thus, correct revelation cannot be a Nash equilibrium with respect to the global game arising from the MDP process or from any other resource allocation process selecting individually rational Pareto optima.

Yet note that, by the result in Section 3, if $\delta^1 = \delta^2 = 1/2$, then correct revelation is not only the unique Nash equilibrium strategy for each player at each instant in the instantaneous games arising from the MDP procedure, it is in fact his best replay no matter what the other agent does. Thus, if the local game is the proper way to model the incentive problem, misrepresentation presumably will not arise, let alone cause any problems. If on the other hand the consumers are concerned with their final consumption bundles and the global game captures the essence of the incentive question, we must expect difficulties.

5. Possible Extensions and Concluding Remarks

The present analysis has left open a number of important issues, five of which we consider here. Four of these relate to the local game and one to the global game. The first of these concerns the effect of imposing non-negativity constraints on the announced marginal rates of substitution under the MDP procedure. This restriction, which amount to the authorities assuming *a priori* that the public good is never a "bad" for any individual and enforcing this assumption by refusing to accept negative announced marginal rates of substitution, might seem natural in view of the assumption (used in the proof of convergence) that $U^i_y \geqq 0$. Of course, our results indicate that there is no need to impose the $\psi^i \geqq 0$ constraint, since one obtains convergence to an optimum without it. Still, one might not want to countenance overt lying, and thus one might require non-negativity. In this regard Claude Henry [5] has shown that even with a non-negativity constraint, which introduces a non-linearity in the best replay functions, there is still a unique Nash equilibrium at each t in the instantaneous game. The resulting system, which gives dy/dt and dx^i/dt in terms of π, γ and δ, is more complicated than that arising in the unconstrained case, since the factor of proportionality between dy/dt and $\Sigma\pi^i - \gamma$ (i.e., the speed of adjustment) and the coefficients of the quadratic terms in the expressions for the dx^i/dt depend on the reference allocation at time t. However, this system still converges to a Pareto optimum. Thus, the principal results of Section 3 carry over to the constrained case.

The next issue concerns our assumption that the public good level is always positive, which justified the specification of dy/dt as $\Sigma\psi^i - \gamma$. If we relax this assumption, then we must specify dy/dt as $\max[0, \Sigma\pi^i - \gamma]$ when y is zero. In this case, the best replay function takes a more complicated form. If $y > 0$ or if y=0 and $\Sigma_{)i(}\psi^i + \pi^i - \gamma \geqq 0$, then $h^i(\psi)$ is as given earlier. But if y=0 and $\Sigma_{)i(}\psi^j + \pi^i - \gamma < 0$, then all values of ψ^i less than or equal to $\gamma - \Sigma_{)i(}\psi^j$ are best replays. In this situation it is appealing to take π^i as <u>the</u> best replay, both because this choice alone makes behavior continuous and because one might hypothesize that agents would not go out of their way to misrepresent their preferences if they do equally well through correct revelation. If one makes this extra behavioral assumption, then there is still a unique Nash equilibrium at each t. If $y > 0$ or if y=0 and $\Sigma\pi^i - \gamma \geqq 0$, this equilibrium is that defined in (2). If y=0 and $\Sigma\pi^i - \gamma \leq 0$, this equilibrium is π. (The proofs of these claims are sketched in the Appendix). In either case, we still can use the Drèze-de la Vallée Poussin theorem to obtain convergence to an optimum.

One would also like to relax the assumption of one public good. If we do so, however, the assumption that no public good level is ever zero becomes particularly difficult to maintain and would require rather stringent assumptions on preferences. However, if we adopt the assumption introduced in the preceding paragraph, that if correct revelation is a best replay it will be used, these difficulties can be avoided. Given this, extra public goods present no difficulties, since the best replay strategy for any individual and any public good are exactly as before. The problem is then completely separable between public goods, and the Nash equilibrium simply consists of a vector ψ^i for each i, where ψ^i_j, i's announced MRS for public good j, is given by (2) if $y_j > 0$ or if the correct revelation would lead to an increase in y_j, and is his correct MRS otherwise.

The final question relates to the nature and effects of misrepresentation of preferences in implementable, discrete time planning procedures. The results in [7], as discussed in the preceding Section, are applicable for the global game arising from such a procedure. For the local game at each iteration Françoise Schoumaker [10] has investigated a version of the Champsaur-Drèze-Henry discrete time procedure [2]. For the case of a single public good her results indicate that, although there may be multiple equilibria, all such equilibria result in convergence to an optimal allocation. The case of multiple public goods presents further complications, but convergence can still be obtained.

In the context of the global game, the Ledyard-Roberts result shows that correct revelation is not a Nash equilibrium, but it leaves open the question of the nature of the Nash equilibrium. In particular, the question of whether equilibria can be optima is largely open. In the Ledyard-Roberts example, the allocations corresponding to Nash equilibria form a set with non-empty interior relative to the space of feasible allocations, while the intersection of the equilibrium allocations with the Pareto optima, although non-empty, is a closed, nowhere-dense set. Thus, although Nash equilibria can yield optima, such optimal equilibria are very rare. One might conjecture that this is the general case, but the question is open.

The results of Sections 3 and 4 emphasize the sensitivity of the nature and outcomes of self-interested behavior to the particular formulation of the game used to model self-interest. This is further emphasized by consideration of the mechanism for optimally allocating public goods proposed by Groves and Ledyard [4]. They allow agents to (mis)state their marginal willingness to pay, but show that Nash equilibria give optimality. This is accomplished by giving up individual rationality of the allocations and by not allowing the agents to recognize the full impact of their messages. These considerations suggest that the information and computing ability the agents have relative to the complexity of the problem of evaluating various strategies ought to be taken into account in selecting a modeling.

FOOTNOTES

1. The case $\delta^i < 1/2$ is clearly the more interesting. However, if $\delta^i > 1/2$ then the relationship between the direction of the desired change in y and over- or under-reporting is simply reversed; i.e., if i wants y increased he will over-report his MRS.

2. This argument is due to Peter Hammond.

Appendix

Note first that if $\Sigma\pi^i - \gamma \leqq 0$, then $\psi = \pi$ is a Nash equilibrium, since $\Sigma_{)i(}\psi^i + \pi^i - \gamma = \Sigma\pi^i - \gamma \leqq 0$, for all i. Suppose then that $\psi \neq \pi$ is also a Nash equilibrium, and suppose that $\psi^i > \pi^i$ for some i. If $\Sigma\psi^j - \gamma \leqq 0$, then $\Sigma_{)i(}\psi^j + \pi^i - \gamma < \Sigma_{)i(}\psi^j + \psi^i - \gamma \leqq 0$ and π^i is i's best replay, which is a contradiction. If $\Sigma\psi^j - \gamma > 0$, and $\Sigma_{)i(}\psi^j + \pi^i - \gamma \leqq 0$, we get the same contradiction, while if $\Sigma_{)i(}\psi^j + \pi^i > 0$ then the best replay for i is $\psi^i = h^i(\psi) < \pi^i$, which is again a contradiction. Thus, if $\psi \neq \pi$ is a Nash equilibrium, we must have $\psi^i \leqq \pi^i$ for all i. But then $\Sigma_{)i(}\psi^i + \pi^i - \gamma < 0$ for all i, and again we get a contradiction.

Now, if $\Sigma\pi^i - \gamma > 0$, the equations (2) still define a Nash equilibrium Φ, since $\Sigma_{)i(}\Phi^j + \pi^i - \gamma = 2(1-\delta^i)(\Sigma\pi^j - \gamma)/(N-1) > 0$, so indeed Φ^i is a best replay against $\Phi^{)i(}$. Suppose then that $\psi \neq \Phi$ is also a Nash equilibrium. If $\Sigma_{)i(}\psi^j + \pi^i - \gamma > 0$ for all i, then $\psi = \Phi$, so we must have the reverse inequality holding for some agents. Denote the set of such consumers as S, and note that $\psi^i = \pi^i$ for $i \in S$. Now $S \neq \{1,\dots,N\}$, since if it did we would have $\psi = \pi$, from which follows $\Sigma\pi^i - \gamma \leqq 0$, which is a contradiction. Note that $\psi^i < \pi^i$ for those $i \notin S$. This follows from the definition of $h^i(\psi)$. Now suppose $\Sigma\psi^i - \gamma \leqq 0$. In this case $v^i(\psi) = 0$ for all i, since $dy/dt = 0$. But for any agent with $\Sigma_{)i(}\psi^j + \pi^i - \gamma > 0$, he could obtain $v^i > 0$ if he reported π^i, so ψ^i could not have been a best replay for him. On the other hand, if $\Sigma\psi^i - \gamma > 0$, then $\Sigma_{)i(}\psi^j + \pi^i - \gamma = \Sigma\psi^j - \gamma > 0$ for any $i \in S$, which is a contradiction. Thus Φ is the unique Nash equilibrium if $\Sigma\pi^i - \gamma > 0$.

REFERENCES

[1] Champsaur, P., Neutrality of Planning Procedures in an Economy with Public Goods. The Review of Economic Studies, 43 (1976), 293-300.

[2] Champsaur, P., J. Drèze and C. Henry, Stability Theorems with Economic Applications. Econometrica, 45 (1971), 273-294.

[3] Drèze, J. and D. de la Vallée Poussin, A Tatonnement Process for Public Goods. The Review of Economic Studies, 38 (1971), 133-150.

[4] Groves, T. and J. Ledyard, Optimal Allocation of Public Goods: A Solution of the 'Free Rider' Problem. Econometrica, 45 (1977), 783-810.

[5] Henry, C., "Vraisemblance des mensonges et convergence de la procédure M.D.P." manuscript, 1976.

[6] Hurwicz, L., On Informationally Decentralized Systems. in Decision and Organization: A Volume in Honor of Jacob Marschak, ed. by C. McGuire and R. Radner. North-Holland, Amsterdam, 1972.

[7] Ledyard, J. and D.J. Roberts, On the Incentive Problem with Public Goods. Center for Mathematical Studies in Economics and Management Science, Northwestern University, Discussion Paper 116, 1974.

[8] Malinvaud. E., A Planning Approach to the Public Good Problem. Swedish Journal of Economics, 73 (1971), 96-112.

[9] Malinvaud, E. Prices for Individual Consumption, Quantity Indicators for Collective Consumption. Review of Economic Studies, 39 (1972), 385-405.

[10] Schoumaker, F., Best Replay Strategies in the Champsaur-Dreze-Henry Procedure, Center for Mathematical Studies in Economics and Management Science, Northwestern University, Discussion Paper 262, 1976.

[11] Smith, V., Incentive Compatible Experimental Processes for the Provision of Public Goods. Mimeo, University of Arizona, 1976. Presented at the Madison, Wisconsin Meetings of the Econometric Society, June 1976.