DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

INCENTIVES AND A PROCESS CONVERGING TO THE CORE OF A PUBLIC GOODS ECONOMY

Parkash Chander
California Institute of Technology
Indian Statistical Institute

SOCIAL SCIENCE WORKING PAPER 677

June 1988

INCENTIVES AND A PROCESS CONVERGING TO THE CORE OF A PUBLIC GOODS ECONOMY

Parkash Chander

ABSTRACT

The paper considers economies involving one public good, one private good, and constant returns to scale. It is shown that the process proposed earlier in Chander (1983, 1987a, and 1987b) always converges to an allocation which is in the core of the economy. This is then interpreted as an incentive property of the process and it is shown that there exists no process which always converges to the core and in which truthtelling constitutes a dominance equilibrium of the 'local incentive game'.

INCENTIVES AND A PROCESS CONVERGING TO THE CORE OF A PUBLIC GOODS ECONOMY

Parkash Chander*

I. INTRODUCTION

A major concern of economic theory is the problem of choosing from among the set of all feasible alternatives those which are acceptable in some sense to every agent or every subset of agents. An often used tool of analysis in this regard is the theory of n-person cooperative games and the associated solution concept of a core. Though the path breaking work of H. Scarf established sufficient conditions for the non-emptiness of the core of an economy independently of the descriptive concept of a competitive equilibrium, the dynamic approaches to the core continue to be confined to only the competitive process that under certain conditions can be shown to converge to some particular allocations in the core. This is perhaps because the core as a solution concept involves the unrealistic assumption of complete information. Whereas the competitive process allows privacy or dispersion of information.

The case of economies with public goods, however, appears to be even more difficult because the competitive process can no longer be shown to lead to an efficient solution and there is no other informationally decentralized process available that converges to the core. The informational and computational requirements of achieving core allocations in economies with public goods therefore seem to be far beyond any accessibility.

Chander (1983, 1987a, and 1987b) propose a dynamic process for economies involving one public and one private good and constant returns to scale which exhibits a continuous path of feasible reallocations of commodities that converges from the initial endowments to an individually rational and Pareto optimal allocation. The process is informationally decentralized in the sense that only local information concerning the marginal rates of substitution of the agents is required. The process is also unbiased in the sense that any Pareto optimal allocation can be realized by a suitable redistribution of the initial endowments.

---

* California Institute of Technology and Indian Statistical Institute

I would like to thank Jacques Drèze and Henry Tulkens for their helpful suggestions. Thanks are also due to Jean-Jacques Laffont for his comments. I have also benefited from a presentation at the Caltech Workshop.

The present paper shows that the process in fact always converges from the initial endowments to some allocation in the core, i.e., to a subset of Pareto optimal and individually rational allocations.[1]

Apart from its theoretical interest as outlined above, the convergence property is of considerable interest from the point of view of the problem of incentives as well.

Hurwicz (1972), Ledyard and Roberts (1974), and Roberts (1979) prove an impossibility result that their exists no (static) mechanism whose outcomes are always Pareto optimal and in which truthtelling is a dominance strategy. Groves and Ledyard (1977), Hurwicz (1979) and Walker (1981) among others therefore design mechanisms by sacrificing dominance equilibria requirement and accept weaker type namely the noncooperative Nash equilibria.

On the other hand Fugigaki and Sato (1981) and Laffont and Maskin (1983) describe a class of dynamic processes which have the remarkable property that truthtelling constitutes a dominance equilibrium of the local incentive game (see Drèze and de la vallée Poussin (1971), Malinvaud (1971) and Roberts (1979)) and which converge to individually rational and Pareto optimal allocations. Chander (1987b) shows that truthtelling is not a dominance strategy of the local incentive game corresponding to his process, though unique Nash equilibrium strategies exist and if the agents adopt their Nash strategies then the so-defined 'Nash strategically stable' process, as seen below, also converges to the core.

This leads to the question whether there exists a process which converges to the core and in which truthtelling is a dominance strategy of the local incentive game. We show below that there exists no such process. This means there are limits to what can be achieved in terms of the various desirable properties of dynamic processes as well.

In terms of the existing literature on dynamic processes for public good economies, the present result may be described as follows. In their path breaking work Drèze and de la Vallée Poussin (1971) and Malinvaud (1970 and 1971) describe a process which always converges from the initial endowments to an individually rational and Pareto optimal allocation. Their process which has come to be termed the MDP process, has been at the centre of all later work on dynamic processes for public good economies during the last fifteen years or so. Chander (1987a and 1987b) give several theoretical justifications for his process in comparison with the MDP process, though as noted there the two processes are similar in nature. The results of this paper should therefore strengthen those justifications further.

We present both a mathematical as well as an intuitive geometrical proof of the main results. For the geometrical proofs we extend the Kolm triangle diagram beyond what has been done in the literature so far. The extension in itself is of some independent interest.

The contents of this paper are as follows: section 2 introduces the model and definitions. Section 3 proves the convergence result. Section 4 proves the impossibility result that there exists no process which *always* converges to the core and in which truthtelling is a dominance strategy of the local incentive game.

1. As a matter of fact this was conjectured already in Chander (1987a, footnote 8), but no proof was offered.

II. THE MODEL AND DEFINITIONS

We consider economies consisting of one public good, whose quantity is denoted by x, one private good, whose quantity is denoted by y, n consumers, and a single producer. Each consumer is characterized by his utility function u^i defined on R^2_+ (the nonnegative orthant of R^2) and by a positive endowment w^i of the private good. The producer is characterized by a cost function $g : R^1_+ \to R^1_+$ which associates with every quantity of public good the minimum cost (in terms of the private good) $g(x)$ needed in order to produce quantity x of the public good. Let $N = \{1, \cdots, n\}$ denote the set of consumers.

Assumption 1: For each $i \in N$

1.1 u^i is quasi–concave and at least twice continuously differentiable and strictly quasi–concave for at least one i.

1.2 $\partial u^i(x,y) / \partial x \geq 0$ and $\partial u^i(x,y) / \partial y > 0$ for all $(x,y) \in R^2_{++}$ (the positive orthant of R^2);

1.3 for any $x \geq 0$, $\lim\limits_{y \to 0} \dfrac{\partial u^i(x,y)/\partial x}{\partial u^i(x,y)/\partial y} = 0$

Assumption (1.2) implies that the utility functions are monotonic in terms of the public good and strictly monotonic in terms of the private good. Assumption (1.3) rules out the possibility of a consumer giving up every last bit of the private good. It is not essential for the analysis below, but is made only in order to avoid certain boundary problems that may arise otherwise.

Assumption 2: The cost function g is linear, i.e., $g(x) = x$.

An allocation is an $(n+1)$ tuple $(x, y^1, \cdots, y^n) \in R^{n+1}_+$ where (x, y^i) denotes the consumption bundle of consumer i. An allocation $(x, y^1, \cdots, y^n)$ is *feasible* if and only if $x + \Sigma y^i \leq \Sigma w^i$. An allocation $(x, y^1, \cdots, y^n)$ is *individually rational* if it is feasible and if $u^i(x, y^i) \geq u^i(0, w^i)$ for all $i \in N$, and *Pareto optimal* if it is feasible and if there does not exist a feasible allocation $(\bar{x}, \bar{y}^1, \cdots, \bar{y}^n)$ such that $u^i(\bar{x}, \bar{y}^i) \geq u^i(x, y^i)$ for all $i \in N$ and $u^j(\bar{x}, \bar{y}^j) > u^j(x, y^j)$ for some $j \in N$. Let

$$Z = \left\{ (x, y^1, \cdots, y^n) \in R^n_+ \ : \ x + \sum_{i \in N} y^i \leq \sum_{i \in N} w^i \right\}.$$

Then Z is the *set of all feasible allocations.* It is clearly compact.

We shall often denote consumer i's marginal rate of substitution by $\pi^i(x, y^i)$, that is, $\pi^i(x, y^i) = (\partial u^i(x, y^i) / \partial x) / (\partial u^i(x, y^i) / \partial y^i)$, a feasible allocation $(x, y^1, \cdots, y^n)$ by z, $u^i(x, y^i)$ by $u^i(z)$ and $\pi^i(x, y^i)$ by $\pi^i(z)$ whenever no confusion is possible. Clearly $\pi^i : R^2_+ \to R_+$ for all $i \in N$.

Assumption 3: $\sum\limits_{j \in N} \pi^j(0, w^j) \geq 1$.

The economies which do not satisfy this assumption are not of much interest.

We shall often adopt the following notation. Let $(\pi^1, \cdots, \pi^n)$ be some n-tuple of real numbers. Then $\pi^S = \sum_{i\in S} \pi^i$, $S \subseteq N$.

Lemma 1: An allocation $\bar{z} \equiv (\bar{x}, \bar{y}^1, \cdots, \bar{y}^n)$ is Pareto optimal if and only if $\bar{x} + \bar{y}^N = w^N$, $\pi^N(\bar{z}) = 1$, and $u^i(\bar{x}, \bar{y}^i) \geq u^i(x, y^i)$ for all $(x, y^1, \cdots, y^n)$ such that $\pi^i(\bar{z})x + y^i \leq \pi^i(\bar{z})\bar{x} + \bar{y}^i$, $i \in N$.

Proof: See Drèze and de la Vallée Poussin (1971).

An allocation $(x, y^1, \cdots, y^n)$ is *attainable* by a coalition $S \subseteq N$ if

$$x + y^S \leq w^S$$

An allocation $(x, y^1, \cdots, y^n)$ is a *core allocation* if there exists no coalition $S \subseteq N$ and an allocation $(\bar{x}, \bar{y}^1, \cdots, \bar{y}^n)$ attainable by S such that $u^i(\bar{x}, \bar{y}^i) > u^i(x, y^i)$ for each $i \in S$.

An allocation $(x, y^1, \cdots, y^n)$ is *coalitionally noncoercive* if it is feasible and if there exists no subset $S \subseteq N$ such that

$$x + y^S < w^S \quad .$$

Notice that an allocation $(x, y^1, \cdots, y^n)$ is coalitionally noncoercive if and only if $x + y^N = w^N$ and $y^i \leq w^i$ for all $i \in N$. If an allocation is not coalitionally noncoercive then $(x, y^i) >> (0, w^i)$ for some i, i.e., some consumer i has more of the public good as well as the private good.[2]

There is one more aspect of noncoercive allocations which is of interest from the point of view of informational decentralization. If there is privacy of information, i.e., if each consumer knows only about his own characteristics, then given an allocation a consumer by himself may not be able to figure out whether or not it is in the core. On the other hand given an allocation (and therefore the implicit net trade vector), each consumer can immediately figure out whether or not it was a coalitionally noncoercive allocation even when there is privacy of information.

Notice further that every core allocation is an individually rational, Pareto optimal and coalitionally noncoercive allocation, but every individually rational, Pareto optimal, and coalitionally noncoercive allocation is not necessarily a core allocation. The core is therefore contained in the set of all individually rational, Pareto optimal and coalitionally noncoercive allocations.

---

2. The concept of a coalitionally noncoercive allocation was first introduced in Chander (1987b). It can also be extended to pure exchange economies as follows. Let $(w_i = (w_{i1}, w_{i2}))_{i=1}^n$ be the initial endowments in a two-good and n-person pure exchange economy. Then an allocation $(x_i = (x_{i1}, x_{i2}))_{i=1}^n$ is coalitionally noncoercive if there exists no subset $S \subseteq N$ such that either

$$\sum_{i\in S} x_{i1} < \sum_{i\in S} w_{i1} \text{ and } \sum_{i\in S} x_{i2} \leq \sum_{i\in S} w_{i2} \text{ or } \sum_{i\in S} x_{i1} \leq \sum_{i\in S} w_{i1} \text{ and } \sum_{i\in S} x_{i2} < \sum_{i\in S} w_{i2} \; .$$

observe that every core allocation is a coalitionally noncoercive allocation.

We now illustrate these solution concepts for the two consumer case by extending the Kolm triangle diagram.[3] In this diagram a feasible allocation $z = (x, y^1, \cdots, y^n)$ is represented as a point Q in an equilateral triangle ABC of height $w_1 + w_2$, see Figure 1.

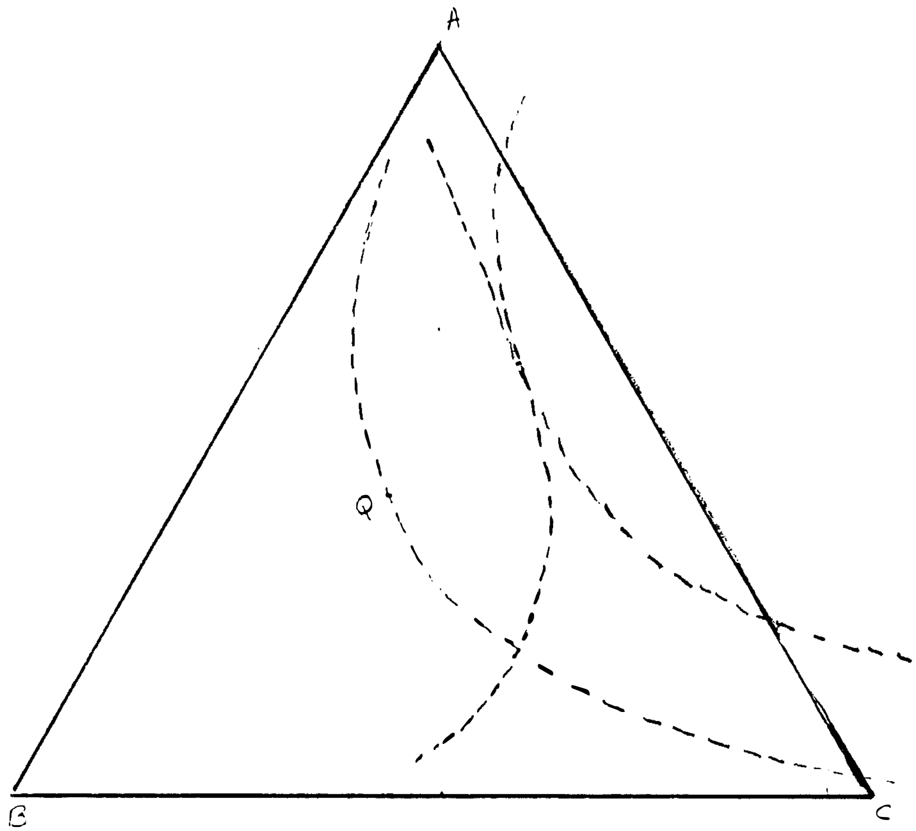

Figure 1

3. This diagram was first used by Kolm (1964) and later by Malinvaud (1971).

The vertical distance from the point Q to the horizontal side BC is taken as x : the public good quantity, the distance from Q to the side AB as y^1 : the private good allocation of agent 1, and the distance from Q to the side AC as y^2. We know from geometry that the sum of these distances must be equal to the vertical height of the triangle, i.e., $w^1 + w^2$ irrespective of the location of Q. Given this coordinate system the indifference curves of the two consumers can be represented as shown in Figure 1 and a Pareto optimal allocation as a point where the indifference curves of the two consumers are tangent to each other.

The extended diagram is shown in Figure 2.

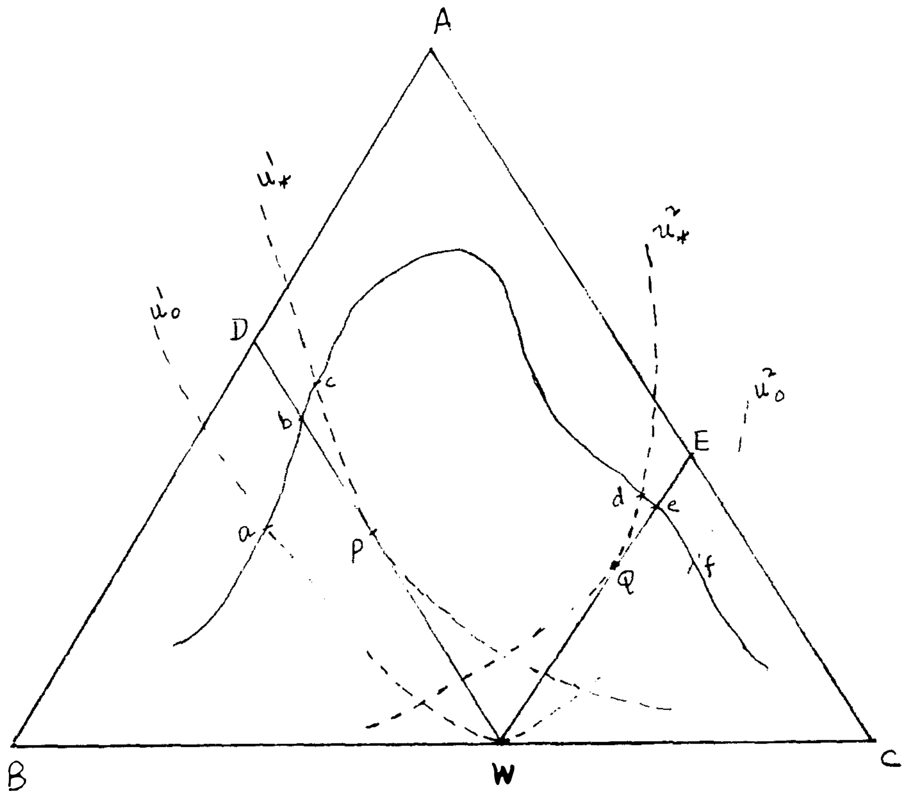

Figure 2

Let the point W on the horizontal line BC represent the initial endowment $(0,w^1,w^2)$. Then the points in the parallelogram $WDAE$ represent coalitionally noncoercive allocations. The point in the triangles DBW and EWC represent allocations attainable by consumer 1 and 2, respectively, and the points P and Q their utility maximizing attainable allocations. The points enclosed between the indifference curves u_0^1 and u_0^2 represent the individually rational allocations.

Let the set of points where the indifference curves of the two consumers are tangent to each other be represented by the curve passing through a, b, c, d, e and f. Then the points on the curve $\overline{af}$ represent individually rational and Pareto optimal allocations. The points on the curve $\overline{be}$ represent individually rational, Pareto optimal, and coalitionally noncoercive allocations. Finally the points on the curve $\overline{cd}$ represent the core allocations.

On this representation we immediately see how the three solution concepts of individually rational and Pareto optimal allocations, individually rational, Pareto optimal and coalitionally noncoercive allocations, and core allocations are related to each other.

III. THE PROCESS AND ITS CONVERGENCE

Starting at $t=0$ with $z(0) \equiv (x(0), y^1(0), \cdots, y^n(0)) = (0, w^1, \cdots, w^n)$, let $z(t) = (x(t), y^1(t), \cdots, y^n(t))$ be the feasible allocation at time $t \in [0,+\infty)$, then the process is described by the following differential equation system

$$\dot{x}(t) = \pi^N(t) - 1 \tag{1}$$

$$\dot{y}^i(t) = -(\pi^i(t)/\pi^N(t))\,\dot{x}(t), \quad i \in N, t \in [0,+\infty)$$

where $\pi^i(t) = \pi^i(z(t))$ and $\pi^N(t) = \sum_{i \in N} \pi^i(t)$. The n-tuple $((\pi^1(t)/\pi^N(t)), \cdots, (\pi^n(t)/\pi^N(t)))$ is called the *cost-sharing* rule.

The differential equation system (1) is of the form $\dot{z} = f(z)$. Since π^i's are continuously differentiable functions of $z, z \in Z$, the function f is defined on the set Z and is continuously differentiable. This means the function f satisfies the Lipschitz condition for uniqueness of solution.

A *solution* to the differential equation system (1) is a continuous function, say $h : [0,+\infty) \to Z$ such that for every $t \in [0,+\infty)$ the derivative $dh(t)/dt$ exists and satisfies equation system (1). The vector $h(0)$ is the initial value of the system.

Since the function f is continuously differentiable the existence and uniqueness of a solution $h : [0,+\infty) \to Z$ for the differential equation system (1) with $h(0) = (0, w^1, \cdots, w^n)$ follows from standard theorems on differential equations, for example, Coddington and Levinson (1955).

It is also clear from (1) that any stationary point is a Pareto optimal allocation. The process is *balanced* at each instant t, since $\dot{x}(t) + \sum_{i \in N} \dot{y}^i(t) = 0$ and *monotonic* in terms of each u^i, since

$$\dot{u}^i(t) \equiv \frac{du^i(t)}{dt} = \frac{\partial u^i(t)}{\partial x(t)}\dot{x}(t) + \frac{\partial u^i(t)}{\partial y^i(t)}\dot{y}^i(t) = \frac{\partial u^i(t)}{\partial y^i(t)}\left[\pi^i(t) - \frac{\pi^i(t)}{\pi^N(t)}\right](\pi^N(t)-1)$$

$$= \left[\frac{\partial u^i(t)}{\partial y^i(t)}\right]\left[\frac{\pi^i(t)}{\pi^N(t)}\right](\pi^N(t)-1)^2 \geq 0 \ , \ i \in N \ , t \in [0,+\infty)$$

where $u^i(t) = u^i(x(t), y^i(t))$.

These plus the compactness of the set of feasible allocations (see Assumption 2) and the continuity of the utility functions u^i implies that $\Sigma u^i(t)$ is a suitable Lyapunov function which means that the process is quasi–stable in the sense that any limit point of its trajectory is a stationary point. The fact that at least one u^i is strictly quasi–concave (Assumption 1) and that any stationary point of (1) is a Pareto optimum implies that the process, in fact, converges to a unique stationary point which is Pareto optimal and at least as good as the initial endowment in terms of the utility functions u^i, i.e., individually rational.

It is easily seen that it must also be coalitionally noncoercive. From the second equation in (1) and that $y^i(0) = w^i$, $\pi^i(t) \geq 0$ for $i \in N$, $t \in [0,+\infty)$, and $\Sigma \pi^j(0,w^j) \geq 1$ (Assumption 3), it follows that $\dot{x}(t) \geq 0$ and $\dot{y}^i(t) \leq 0$ for all $i \in N$ and $t \in [0,+\infty)$, that is, $y^i(t) \leq w^i$, $i \in N$ and $t \in [0,+\infty)$.

These properties of the process are summarized in the following lemma.

Lemma 2: The differential equation system (1) has a unique solution $h \ : \ [0,+\infty) \rightarrow Z$ which is continuous and such that $h(0) = (0, w^1, \cdots, w^n)$, $\dot{x}(t) \geq 0$, $\dot{y}^i(t) \leq 0$ and $\dot{u}^i(t) \geq 0$, $i \in N$ and $t \in [0,+\infty)$ and that the process converges to a unique stationary point which is individually rational, Pareto optimal, and coalitionally noncoercive.

Theorem 1 below proves that the process must in fact converge to a core allocation. This may first be seen intuitively from the representation in Figure 2 as follows. Since the process converges to a coalitionally noncoercive allocation, therefore, it cannot converge to any point outside the parallelogram *WDAE*. The process must converge thus to some point on the curve $\overline{be}$. However, since $\dot{y}^i(t) \leq 0$ for $i \in N$ and $t \in [0,+\infty)$, to converge to any point on the curve $\overline{be}$ the process must first reach some point in the area enclosed between the indifference curves u_*^1 and u_*^2 including the points on u_*^1 and u_*^2. But if that is so then the process cannot converge to some point on the curves $\overline{bc}$ and $\overline{de}$ except for the points $\underline{c}$ and $\underline{d}$ because the points on these curves except $\underline{c}$ and $\underline{d}$ correspond to utility levels lower than u_*^1 and u_*^2 and $\dot{u}^i(t) \geq 0$ for $i \in N$ and $t \in [0,+\infty)$. Therefore, the process must converge to some point on the curve $\overline{cd}$, i.e., to a core allocation point.

Theorem 1: The differential equation system (1) has a unique solution and the process converges to a core allocation.

Proof: Lemma 2 implies that there exists a unique solution $h \ : \ [0,+\infty) \rightarrow Z$ such that the process converges to a unique point, say $(\overset{*}{x}, \overset{*}{y}{}^1, \cdots, \overset{*}{y}{}^n)$ which is individually rational, Pareto optimal and coalitionally noncoercive.

Let $(\overset{*}{\pi}^1, \cdots, \overset{*}{\pi}^n) \in R_+^n$ be the marginal rates of substitution of the n consumers corresponding to the allocation $(\overset{*}{x}, \overset{*}{y}^1, \cdots, \overset{*}{y}^n)$. Then Lemma 1 implies

(a) $\overset{*}{\pi}^N = 1$; and

(b) $u^i(\overset{*}{x}, \overset{*}{y}^i) \geq u^i(x, y^i)$ for all $(x, y^1, \cdots, y^n) \in R_+^{n+1}$ such that $\overset{*}{\pi}^i x + y^i \leq \overset{*}{\pi}^i \overset{*}{x} + \overset{*}{y}^i, i \in N$.
Since $(\overset{*}{x}, \overset{*}{y}^1, \cdots, \overset{*}{y}^n)$ is coalitionally noncoercive,

(c) $\overset{*}{x} + \overset{*}{y}^N = w^N$ and $\overset{*}{x} + \overset{*}{y}^S \geq w^S$ for all $S \subseteq N$.

Suppose contrary to the assertion that there exists a coalition S and an allocation $(\bar{x}, \bar{y}^1, \cdots, \bar{y}^n)$ attainable by S such that $u^i(\bar{x}, \bar{y}^i) > u^i(\overset{*}{x}, \overset{*}{y}^i)$ for each $i \in S$. It follows from (b) above that $\overset{*}{\pi}^i \bar{x} + \bar{y}^i > \overset{*}{\pi}^i \overset{*}{x} + \overset{*}{y}^i$ for each $i \in S$. Therefore

(d) $\overset{*}{\pi}^S \bar{x} + \bar{y}^S > \overset{*}{\pi}^S \overset{*}{x} + \overset{*}{y}^S$ and $\bar{x} + \bar{y}^S \leq w^S$,
since $(\bar{x}, \bar{y}^1, \cdots, \bar{y}^n)$ is attainable by S. It follows from (c) and (d) that

$$\overset{*}{\pi}^S(\bar{x} - x^*) > \overset{*}{y}^S - \bar{y}^S \geq \bar{x} - x^* .$$

Since $0 \leq \overset{*}{\pi}^S \leq 1$ (see (a)), the above inequality can be true only if $x^* > \bar{x}$.

Since $h : [0,+\infty) \rightarrow Z$ is continuous, $h_1 : [0,+\infty) \rightarrow [0, \overset{*}{x}]$ is also continuous. Therefore, there exists a $t \in [0,+\infty)$ such that $h_1(t) \equiv x(t) = \bar{x}$. Which means for some $t \in [0,+\infty)$ we have $(x(t), y^1(t), \cdots, y^n(t))$ such that

$$x(t) = \bar{x} \text{ and } x(t) + y^N(t) = w^N .$$

Lemma 2 implies that $y^i(t) \leq w^i$ for each $i \in N$. Therefore, $x(t) = \bar{x}$ and $x(t) + y^S(t) \geq w^S$, i.e., $x(t) = \bar{x}$ and $y^S(t) \geq \bar{y}^S$. It follows that $u^i(\bar{x}, \bar{y}^i) \leq u^i(x(t), y^i(t))$ for at least one $i \in S$. Lemma 2 implies that $u^j(x(t), y^j(t)) \leq u^j(\overset{*}{x}, \overset{*}{y}^j)$ for all $j \in N$. Hence $u^i(\bar{x}, \bar{y}^i) > u^i(\overset{*}{x}, \overset{*}{y}^i)$ is not true for at least one $i \in S$. This is a contradiction of our supposition. Hence the theorem.

Notice that the proof does not depend on the particular form of the second equation in (1), but only on the fact that it leads to both $\dot{y}^i(t) \leq 0$ and $\dot{u}^i(t) \geq 0$ for all $t \in [0,+\infty)$. This implies there may exist other processes which also converge to the core. For example, the one below:

$$\dot{x}(t) = \frac{1}{\ell - \pi^L(t)} (\pi^L(t) - 1)$$

$$\dot{y}^i(t) = -\left[\frac{1 - \pi^L(t) + (\ell - 1)\pi^i(t)}{\ell - \pi^L(t)}\right] \dot{x}(t) \quad \text{for } i \in L \tag{2}$$

$$= 0 \quad \text{for } i \in N - L \ , \ t \in [0,+\infty)$$

where $L \equiv L(t)$ and $\ell \equiv \ell(t) = |L(t)|$ (the cardinality of $L(t)$) are such that

$$1 - \pi^L(t) + (\ell - 1)\pi^i(t) \geq 0 \quad \text{for all} \quad i \in L$$

and

$$1 - \pi^L(t) + (\ell - 1)\pi^i(t) < 0 \quad \text{for all} \quad i \in N - L \quad .$$

It is easily seen that this process also has a unique continuous solution and that $\dot{x}(t) \geq 0, \dot{y}^i(t) \leq 0$ and $\dot{u}^i(t) \geq 0$ for all $i \in N$ and $t \in [0, +\infty)$.

It may be noted that we have proved the result only in the case of a continuous–time process. However, this is more a matter of convenience rather than necessity. It is possible to design a discrete–time version as well which converges to the core.

Finally, note that the proof of Theorem 1 crucially depends on the fact that the public good is continuously substituted for the private good by all the consumers. A generalization of our result to pure exchange economies is therefore not straightforward. It however suggests that if one could somehow design a process for a pure exchange economy which is simultaneously coalitionally noncoercive and monotonic at *each instant* then it must converge to the core.

IV. INCENTIVES AND AN IMPOSSIBILITY RESULT

The term "free rider" has not been formally defined in the public goods literature, perhaps because its meaning is quite clear. A consumer is a free rider if he gets to enjoy the public good quantity without incurring any cost. This suggests the following definition.

A process admits *downright free riding* if for some economy it converges to a feasible allocation $(x, y^1, \cdots, y^n)$ which is such that $(x, y^i) >> (0, w^i)$ for some consumer $i \in N$, that is, if it converges to an allocation which is not coalitionally noncoercive.[4]

From the point of view incentives, therefore, it is desirable that a process must always converge to at least a coalitionally noncoercive allocation if not a core allocation. Particularly so because as noted earlier the allocations which are not coalitionally noncoercive can be easily identified by the consumers even under the assumption of privacy of information.

The core convergence property is thus an incentive property defined in terms of the payoffs associated with the final outcome of the process. On the other hand issues related to incentives in the context of dynamic processes have often been studied in the framework of what has come to be termed the local incentive game, see Drèze and de la Vallée Poussin (1971), Malinvaud (1971), Roberts (1979), Fugigaki and Sato (1981), and Laffont and Maskin (1983) among others. In this connection Chander (1987b) shows that though truthtelling does not constitute a dominance equilibrium, a unique Nash equilibrium exists for the local incentive game at each point of the trajectory of process (1) and if the agents adopt their Nash equilibrium strategies then the so-defined 'Nash strategically stable' process is process (2) described above, which means the incentive

4. This term was suggested to me by Professor Jacques Drèze.

Process (3) converges to a Pareto optimal allocation only if

$$F(\pi(t)) = 0 \;\Leftrightarrow\; \pi^1(t) + \pi^2(t) = 1 \quad ,t \in [0,+\infty)$$

Notice that if $F(\pi(t)) = 0$, then feasibility and monotonicity imply that $T^i(\pi(t)) = 0$, $i = 1, 2$; $t \in [0,+\infty)$.

Process (3) is *locally strongly incentive compatible* if and only if

$$T^1(\pi^1(t),\pi^2(t)) \geq (\pi^1(t) - \theta^1) F(\theta^1,\pi^2(t)) + T^1(\theta^1,\pi^2(t)) \text{ for all } \theta^1 \in R_+, t \in [0,+\infty)$$

and only if

$$\frac{\partial}{\partial \theta^1}((\pi^1(t) - \theta^1) F(\theta^1,\pi^2(t)) + T^1(\theta^1,\pi^2(t))\Big|_{\theta^1 = \pi^1(t)} = 0 \text{ for all } \theta^1 \in R_+ \text{ and } t \in [0,+\infty)$$

and similarly for consumer 2.

Theorem 2: Suppose the functions $F(\cdot), T^1(\cdot)$, and $T^2(\cdot)$ are such that (3) has a solution and the process is balanced, monotonic, locally strongly incentive compatible and converges to a Pareto optimal allocation. Then it does not *always* converge from the initial endowment to a coaltionally noncoercive allocation and therefore core allocation.

Let E be the class of economies in which consumer 1 has a strictly quasi–concave utility function and an initial endowment w^1 such that $\pi^1(0,w^1) > 1$ and consumer 2 has some utility function of the form $u^2(x,y^2) = \varepsilon x + y^2$ where $0 \leq \varepsilon < 1$ and an initial endowment $w^2 > 0$. The individually rational, Pareto optimal, and coalitionally noncoercive allocations and core allocations for an economy belonging to E are shown in Figure 3 by the curves $\overline{bd}$ and $\overline{cd}$, respectively.

property that the process converges to the core is not affected by myopic strategic behaviour of the consumers.

We now show that there exists no process which always converges to a coaltionally noncoercive allocation and in which truthtelling constitutes a dominance equilibrium of the local incentive game. This means there are limits to what can be achieved in terms of the various desirable properties of dynamic processes.

We consider only two-person processes, i.e., $N = \{1,2\}$. Generalization to more than two persons is straightforward and also not necessary, since we prove only an impossibility result.

Starting at time $t = 0$ with $(x(0), y^1(0), y^2(0)) = (0, w^1, w^2)$, let $(x(t), y^1(t), y^2(t))$ be the feasible allocation at time $t \in [0,+\infty)$, then a two-person process is defined as the following differential equation system.

$$\dot{x}(t) = F(\pi(t))$$

$$\dot{y}^i(t) = -\pi^i(t)\dot{x}(t) + T^i(\pi(t)) \tag{3}$$

$$= -\pi^i(t)F(\pi(t)) + T^i(\pi(t)),\ i = 1,2\,;\ t \in [0,+\infty)$$

where $\pi^i(t) = \pi^i(x,(t), y^i(t)), \pi(t) = (\pi^1(t), \pi^2(t))$ and $F(\cdot), T^1(\cdot)$ and $T^2(\cdot)$ are some arbitrary real valued and continuously differentiable functions on R^2_+. This means if (3) has a solution than it must be unique and continuous.

Process (3) is *balanced* if

$$\dot{x}(t) + \sum_{i=1}^{2} \dot{y}^i(t) = 0 \ ,\ t \in [0,+\infty) \ .$$

Equivalently, if

$$\sum_{i=1}^{2} T^i(\pi(t)) = \left[\sum_{i=1}^{2} \pi^i(t) - 1\right] F(\pi(t)),\ t \in [0,+\infty) \ .$$

Process (3) is *monotonic* in terms of the utilities if

$$\dot{u}^i(t) = \frac{\partial u^i(x(t), y^i(t))}{\partial y^i(t)} (\pi^i(t)\dot{x}(t) + \dot{y}^i(t)) \geq 0 \text{ for } i = 1,2, t \in [0,+\infty).$$

Equivalently, if

$$T^i(\pi(t) \geq 0 \ ,\ i = 1,2; t \in [0,+\infty) \ .$$

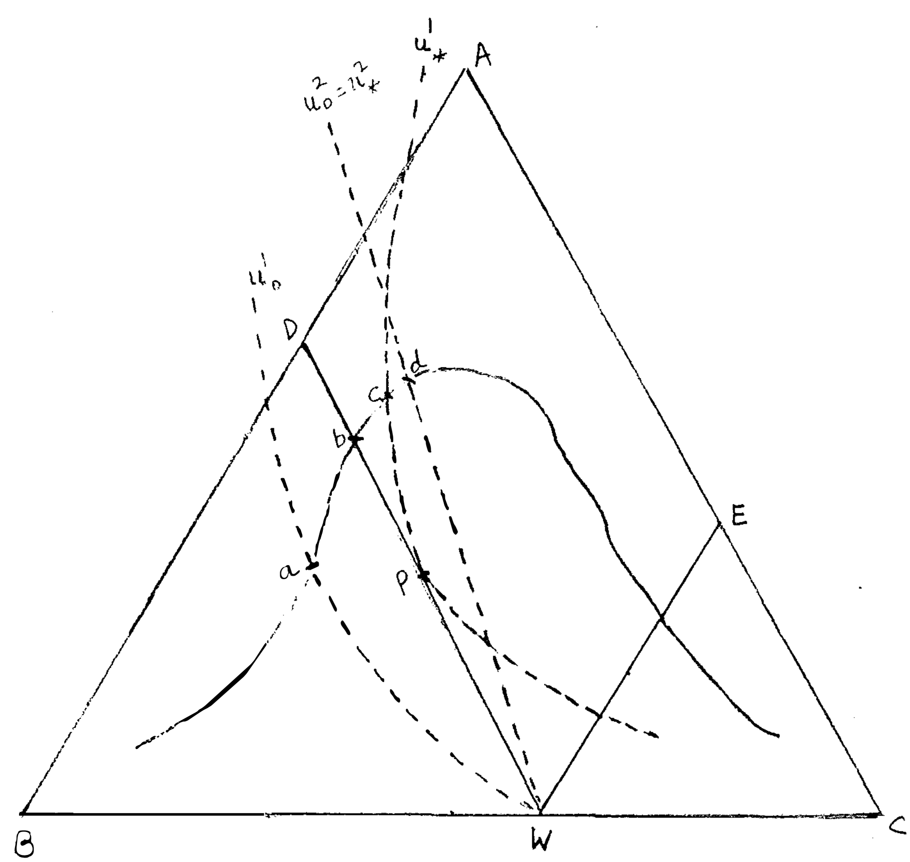

Figure 3

As ε decreases the indifference curve u_0^2 of consumer 2 rotates towards the left around point W and the set of individually rational, Pareto optimal, and coalitionally noncoercive allocations contracts. For $\varepsilon = 0$, u_0^2 coincides with the line DW and both the sets contract to a single point P. In such a case any process which is monotonic and converges to the core, for example process (1), must have all of its trajectory on the line PW. On the other hand, any other process which does not have all of its trajectory on PW and is monotonic must converge to some point in the area enclosed between the line DW and the curve u_0^1 excluding the points on DW, i.e., to some allocation which is not coalitionally non-coercive. Since the functions $F(\cdot), T^1(\cdot)$, and $T^2(\cdot)$ are continuous this must happen for at least sufficiently small but positive ε also.

This intuitively explains the theorem. We shall need the following lemma for a formal proof.

Lemma 3: Suppose the functions $F(\cdot), T^1(\cdot)$ and $T^2(\cdot)$ are such that the differential equation system (3) has a solution, and the process is balanced, monotonic, locally strongly incentive compatible and converges from the initial endowment to a Pareto optimal allocation. Then for the class of economics E

$$T^1(\pi^1,\varepsilon) = \int_{1-\varepsilon}^{\pi^1} F(\delta,\varepsilon)\, d\delta$$

and $F(\pi^1,\varepsilon)$ is weakly increasing in π^1 for all $\pi^1 \in [1,+\infty)$ and $0 \le \varepsilon < 1$.

Proof: Since (3) has a solution, it has a unique continuous solution. By assumption the process converges from the initial endowment to a Pareto optimal allocation for each economy in E, therefore, $\pi^1(t)$ must assume all values in the interval $[1-\varepsilon, \pi^1(0,w^1)]$ for each $0 \le \varepsilon < 1$ and $\pi^1(0,w^1) \ge 1$. Further, we must have

$$F(1-\varepsilon,\varepsilon) = 0 \text{ for all } \varepsilon \text{ such that } 0 \le \varepsilon < 1.$$

Since the process is strongly incentive compatible

$$\frac{\partial}{\partial\theta}((\pi^1-\theta)F(\theta,\varepsilon) + T^1(\theta,\varepsilon))\Big|_{\theta=\pi^1} = 0 \text{ for all } \pi^1 \in [1-\varepsilon,+\infty) \text{ and } 0 \le \varepsilon < 1.$$

That is

$$-F(\pi^1,\varepsilon) + \frac{\partial T^1(\pi^1,\varepsilon)}{\partial \pi^1} = 0 \text{ for all } \pi^1 \in [1-\varepsilon,+\infty) \text{ and } 0 \le \varepsilon < 1 \ .$$

This means

$$T^1(\pi^1,\varepsilon) = \int_{1-\varepsilon}^{\pi^1} F(\delta,\varepsilon)d\delta + K \text{ for all } \pi^1 \in [1-\varepsilon,+\infty) \text{ and } 0 \le \varepsilon < 1 \ .$$

Since $F(1-\varepsilon,\varepsilon) = 0$ for all $0 \le \varepsilon < 1$, balancedness and monotonicity imply that $T^1(1-\varepsilon,\varepsilon) = 0$. Therefore, $K = 0$. Hence

$$T^1(\pi^1,\varepsilon) = \int_{1-\varepsilon}^{\pi^1} F(\delta,\varepsilon)\, d\delta \text{ for all } \pi^1 \in [1,+\infty) \text{ and } 0 \le \varepsilon < 1 \ .$$

On the other hand, since the process is strongly incentive compatible

$$T^1(\pi^1,\varepsilon) \ge (-\Delta(\pi^1))\,(F(\pi^1+\Delta(\pi^1)),\varepsilon) + T^1(\pi^1+\Delta(\pi^1),\varepsilon)$$

and

$$T^1(\pi^1 + \Delta(\pi^1),\varepsilon) \geq \Delta(\pi^1)\, F(\pi^1,\varepsilon) + T^1(\pi^1,\varepsilon)$$

for all $\pi^1, \pi^1 + \Delta(\pi^1) \in [1-\varepsilon, +\infty)$ and $0 \leq \varepsilon < 1$. Adding, we get

$$\Delta(\pi^1)\,(F(\pi^1 + \Delta(\pi^1),\varepsilon) - F(\pi^1,\varepsilon) \geq 0 \text{ for all } \pi^1, \pi^1 + \Delta(\pi^1) \in [1-\varepsilon, +\infty) \text{ and } 0 \leq \varepsilon < 1 \;.$$

Hence $F(\pi^1,\varepsilon)$ is weakly increasing in π^1 for all $\pi^1 \in [1-\varepsilon, +\infty)$ and $0 \leq \varepsilon < 1$. This completes the lemma.

Proof of Theorem 2: From lemma 3 and the balancedness condition, it follows that for each economy in E,

$$\dot{y}^2(t) = \pi^1(t)\, F(\pi^1(t),\varepsilon) - T^1(\pi^1(t),\varepsilon) - F(\pi^1(t),\varepsilon)$$

$$= (\pi^1(t) - 1)\, F(\pi^1(t),\varepsilon) - \int_{1-\varepsilon}^{\pi^1(t)} F(\delta,\varepsilon)\, d\delta \text{ for all } t \varepsilon [0,+\infty).$$

Therefore,

$$\dot{y}^2(t) \begin{cases} > 0 & \text{if } \pi^1(t) > 1 \text{ and } \varepsilon = 0 \\ = 0 & \text{if } \pi^1(t) = 1 \text{ and } \varepsilon = 0 \;, \end{cases}$$

$t \varepsilon [0,+\infty)$. Therefore, if $\varepsilon = 0$ and $\pi^1(0) > 1$ then $y^2(t) > w^2$ for all $t \in [0,+\infty)$. Since the functions $F(\cdot), T^1(\cdot)$, and $T^2(\cdot)$ are continuous functions, there exists an $\varepsilon^*, 0 < \varepsilon^* < 1$ such that $y^2(t) > w^2$ for all $t \varepsilon [0,+\infty)$.

This proves that for any given utility function of consumer 1, there exists a utility function

$$u^2(x,y^2) = \varepsilon^* x + y^2$$

of consumer 2 such that process (3) converges to an allocation which is not coalitionally noncoercive. This completes the proof.

It may be noted that there does exist a process which has a unique continuous solution and which is balanced, monotonic, locally strongly incentive compatible, and which converges from the initial endowment $(0,w^1,w^2)$ to a Pareto optimal allocation, namely

$$\dot{x}(t) = \pi^1(t) + \pi^2(t) - 1$$

$$\dot{y}^i(t) = -\pi^i(t)\,\dot{x}(t) + \tfrac{1}{2}(\pi^1(t) + \pi^2(t) - 1)^2, i = 1,2 \text{ and } t \varepsilon [0,+\infty) \;.$$

See Roberts (1979), Fugigaki and Sato (1981), and Laffont and Maskin (1983).

Finally, note that the requirement that $\pi^1(0,w^1) > 1$ is not essential for the above result to obtain if there are more than two consumers.

REFERENCES

Chander, P. (1983): "Incentives and an alternate (dynamic) process for the allocation of public goods," Technical Report 8305 (May), Indian Statistical Institute; cited in J. Roberts (1987) "Incentives, information, and iterative planning" in: T. Groves, R. Radner and S. Reiter, eds., *Information, Incentives, and Economic Mechanisms* (*Essays in Honor of Leonid Hurwicz*), University of Minnesota Press, Minneapolis, pp. 349–374.

Chander, P. (1987a): "Cost–sharing local games in dynamic procedures for public goods," CORE Discussion Paper No. 8727 (June).

Chander, P. (1987b): "Surplus–sharing rules and incentives in planning procedures for public goods," CORE Discussion Paper No. 8726 (June).

Coddington, E. A. and N. Levinson (1955): *Theory of Ordinary Differential Equations*, New York: McGraw Hill.

Drèze J. and D. de la Vallée Poussin (1971): "A tatonnement process for public goods," *Review of Economic Studies* 37, pp. 133–150.

Fugigaki, Y. and R. Sato (1981): "Incentives in the generalized MDP procedure for the provision of public goods," *Review of Economic Studies* 48, pp. 473–486.

Green, J. and J.-J. Laffont (1979): *Incentives in Public Decision Making*, Amsterdam: North Holland.

Groves, T. and J. Ledyard (1977): "Optimal allocation of public goods: a solution to the 'free rider' problem," *Econometrica* 48, pp. 1487–1506.

Hurwicz, L. (1972): "On informationally decentralized systems," in R. Radner and C. B. McGuire, eds., *Decision and Organization* (*volume in Honor of J. Marschak*), Amsterdam: North Holland, pp. 297–336.

Hurwicz, L. (1979): "Outcome functions yielding Walrasian and Lindahl allocations at Nash equilibrium points," *Review of Economic Studies* 46, pp. 217–225.

Kolm, S. (1964): "Introduction à la théorie du rôle economique de l'état : Les fondements de l'économie publique," Paris: IFP.

Laffont J.-J. and E. Maskin (1983): "A characterization of SLIIC planning procedures with public goods," *Review of Economic Studies* 50, pp. 171–186.

Ledyard, J. and J. Roberts (1974): "On the incentive problem with public goods," Discussion Paper No. 116, Evanston: CMSEMS, Northwestern University. Also see Theorem 3.2 in T. Groves and J. Ledyard (1987) "Incentive compatibility since 1972" in: T. Groves, R. Radner and S. Reiter, eds., *Information, Incentives, and Economic Mechanisms* (*Essays in Honor of Leonid Hurwicz*), University of Minnesota, Minneapolis pp. 48–111.

Malinvaud, E. (1970): "Procedures for the determination of a program of collective consumption," *European Economic Review* 2, pp. 187–217.

Malinvaud, E. (1971): "A planning approach to the public good problem," *Swedish Journal of Economics* 73, pp. 96–112.

Roberts, J. (1979): "Incentives in planning procedures for the provision of public goods," *Review of Economic Studies* 46, pp. 283–292.

Roberts, J. (1987): "Incentives, information, and iterative planning" in: T. Groves, R. Radner and S. Reiter, eds., *Information, Incentives, and Economic Mechanisms* (*Essays in Honor of Leonid Hurwicz*), University of Minnesota Press, Minneapolis, pp. 349–374.

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

Scarf, M. (1967): "The Core of an N-person game," *Econometrica* 35, pp. 50–69.