Equilibrium Participation in Public Goods Allocations*

Paul J. Healy†

Abstract

We consider a notion of voluntary participation for mechanism design in public goods economies in which mechanisms select public goods allocations and individuals then choose whether or not to submit their requested transfer to the central planner. The set of allocations that are robust to non-participation is shown to be sub-optimal in a wide variety of environments and may shrink to the endowment as the economy is replicated. When agents become small as the economy becomes large, *any* non-trivial mechanism suffers from non-participation when agents cannot be coerced to contribute.

Keywords: Public goods, mechanism design, voluntary participation.
JEL: C62, C72, H41

*The author thanks Chris Chambers, Matt Jackson, Hideo Konishi, John Ledyard, Tom Palfrey, Yoshi Saijo, and William Thomson for valuable comments.

†1945 N. High St., Columbus, OH 43210-1120. healy.52@osu.edu.

> [The strategies in a mechanism] are the "natural" alternatives to which participants might resort if it were to their advantage. Are there then any limits to how broad these domains of alternative strategies might be?
> —Leonid Hurwicz (1972, p. 321)

1 Introduction

In the design of mechanisms it is common to apply an individual rationality constraint guaranteeing that the selected outcome be weakly preferred to the endowment for every agent. In economies with private goods, Hurwicz (1972, p. 327) assumes that the mechanism designer must allow the agents a 'no-trade option', forcing any incentive compatible mechanism to select allocations that are preferred to the endowment by all agents.[1] In the subsequent mechanism design literature explicit reference to a 'no-trade option' is infrequent, with most authors directly requiring the mechanism to select outcomes that (weakly) Pareto dominate the endowment point—a requirement frequently called 'individual rationality'.

With externalities, however, a 'no-trade option' is quite different from individual rationality because an agent who does not trade may still be affected by the trades of others. This leads Green and Laffont (1979, p. 121) to argue instead that the individual rationality constraint in the public goods model is founded on the ethical belief that each agent has a natural right to her endowment and the welfare its consumption would generate.

The current paper reconsiders the mechanism design problem with public goods when the mechanism designer must explicitly allow a no-trade option after the mechanism is run and does not have a credible punishment strategy, such as a budget breaker, to punish non-contributors. The resulting constraint, called *equilibrium participation*, requires the mechanism to select an outcome such that every agent prefers to contribute their requested transfer payment rather than withhold it. We assume that if an agent withholds her transfer payment then the level of the public good is reduced to that which can be feasibly produced with the remaining transfers, and the withholding agent is not excluded from its consumption.

In order to induce all agents to choose participation over non-participation, a

[1] Hurwicz (1972) also allows for an arbitrary but pre-specified and publicly-known redistribution of the endowment, much in the spirit of the second welfare theorem.

mechanism can satisfy equilibrium participation by making those agents with the strongest free-riding incentive responsible for the largest share of the production inputs. Such an agent would prefer to contribute since withholding will cause a significant reduction in output. This is demonstrated in example 1 of section 2.3. If several agents have strong free-riding incentives, however, they cannot all be made responsible for the lion's share of production. This problem is exacerbated in larger economies. This is the intuition behind the two main results of this paper: (1) there are many finite economies in which only the endowment satisfies equilibrium participation, and (2) as a classical public goods economy is replicated, if agents become 'small' then the set of outcomes satisfying equilibrium participation converges to the endowment.

The negative results of this paper imply that coercion is absolutely necessary for mechanisms to successfully implement desirable outcomes. If an agent opts out of the mechanism outcome by withholding her transfer, some punishment system must be in place so that the dissenting agent cannot free ride on the production of others. This can be obtained explicitly through fines and sanctions, or implicitly by threatening to produce nothing if any agent defects. If explicit coercion is unavailable and implicit threats incredible, then mechanism design cannot avoid the standard free-rider problem.

The notation and key definition of the paper are provided in the following section. General properties of the set of allocations satisfying equilibrium participation are explored in section 3, followed by an analysis of the constraint in classical, quasi-concave economies with convex technology in section 4. The main result on convergence to the endowment in large economies is proven in section 5. A brief review of relevant literature appears in section 6, followed by open questions and concluding thoughts in section 7.

2 Notation & Definitions

For simplicity, we consider an economy with one private and one public good. Let $\mathcal{I} = \{1, \ldots, I\}$ be the set of consumers, with $I \geq 2$. An allocation of the private good is denoted by $\boldsymbol{x} = (x_1, \ldots, x_I) \in \mathbb{R}_+^I$ and a level of the public good is denoted by $y \in \mathbb{R}_+$. We let $\boldsymbol{z} = (\boldsymbol{x}, y)$ represent a complete allocation and the set $\mathcal{Z} \subseteq \mathbb{R}_+^{I+1}$ represents the set of allocations. The initial endowment is given by $(\boldsymbol{\omega}, 0) \in \mathcal{Z}$, where

$\omega_i > 0$ for each $i \in \mathcal{I}$.[2] Thus, we assume no public good exists initially. For any allocation $\boldsymbol{z}$, let $t_i = \omega_i - z_i$ represent the net transfer of private good paid by agent i. The vector of transfers is given by $\boldsymbol{t} = (t_1, \ldots, t_I)$, their sum by T, and $T_{-i} := T - t_i$. Each agent i has a complete, transitive and reflexive preference relation $\succeq_i$ on $\mathcal{Z} \times \mathcal{Z}$. Let $\succ_i$ and $\sim_i$ represent the asymmetric and symmetric parts of $\succeq_i$, respectively. Where convenient, we further assume that there exists a utility representation of $\succeq_i$ given by $u_i : \mathcal{Z} \to \mathbb{R}$.

Given an allocation $(\boldsymbol{x}, y)$, the *allocation consumed by* i is simply (x_i, y). Since preferences are assumed to be 'selfish' (meaning that $x_i = x_i'$ implies $(\boldsymbol{x}, y) \sim_i (\boldsymbol{x'}, y)$ for every y), we occasionally abuse notation and let $\succeq_i$ represent the projection of i's preference relation onto $\mathbb{R}^2_+ \times \mathbb{R}^2_+$; this allows for statements such as $(x_i, y) \succeq_i (x_i', y')$ to be well-defined. Similar abuses will be applied to $\succ_i$ and $\sim_i$.

The production technology is represented by the set $\mathcal{Y} \subseteq \mathbb{R}_- \times \mathbb{R}$, which is assumed to be non-empty, closed, comprehensive ($\mathcal{Y} - \mathbb{R}^2_+ \subseteq \mathcal{Y}$), and satisfies $\mathcal{Y} \cap \mathbb{R}^2_+ = \{\mathbf{0}\}$.[3] The production set $\mathcal{Y}$ can be described by a production function $F(T)$ such that $\mathcal{Y} = \{(-T, y) : y \leq F(T)\}$, or, equivalently, by a cost function $c(y)$ such that $\mathcal{Y} = \{(-T, y) : T \geq c(y)\}$.[4]

An economy is a collection of agents, possible allocations, endowments, preferences, and production possibilities. The set of all admissible economies with I agents is given by $\mathcal{E}_I$, with typical element $\boldsymbol{e} = (\{\succeq_i\}_{i=1}^I, \mathcal{Z}, \boldsymbol{\omega}, \mathcal{Y})$. An allocation $\boldsymbol{z} = (\boldsymbol{x}, y)$ is *feasible* if $(-T, y) \in \mathcal{Y}$ and $\boldsymbol{t} \leq \boldsymbol{\omega}$.[5] The set of all feasible allocations for an economy $\boldsymbol{e}$ is denoted by $\mathcal{Z}(\boldsymbol{e}) \subseteq \mathcal{Z}$. An allocation is *balanced* if it is feasible and $y = F(T)$. The set of balanced outcomes is denoted $\bar{\mathcal{Z}}(\boldsymbol{e})$.

The following additional assumptions are used at various points in the paper.

A1 (Monotonicity) If $(x_i', y') \geq (x_i, y)$ then $(\boldsymbol{x}', y') \succeq_i (\boldsymbol{x}, y)$.

A1' (Strict Monotonicity) If $(x_i', y') > (x_i, y)$ then $(\boldsymbol{x}', y') \succ_i (\boldsymbol{x}, y)$.

A2 (Convexity) If $\boldsymbol{z}' \succeq_i \boldsymbol{z}$, then $\alpha \boldsymbol{z}' + (1-\alpha)\boldsymbol{z} \succeq_i \boldsymbol{z}$ for all $\alpha \in (0, 1)$.

---

[2] The term 'initial endowment' will be used interchangeably to mean either $\boldsymbol{\omega}$ or $(\boldsymbol{\omega}, 0)$.

[3] $\mathbf{0}$ represents the origin in Euclidean space.

[4] Note that this is not the standard notion of a cost function, which would also incorporate the market input prices; this function only identifies the needed input for a given level of output.

[5] If $\boldsymbol{x}$ and $\boldsymbol{x}'$ are in $\mathbb{R}^n$, then $[\boldsymbol{x} \geq \boldsymbol{x}'] \Leftrightarrow [x_i \geq x_i' \text{ for all } i]$, $[\boldsymbol{x} > \boldsymbol{x}'] \Leftrightarrow [\boldsymbol{x} \geq \boldsymbol{x}' \text{ and } x_i > x_i' \text{ for some } i]$, and $[\boldsymbol{x} \gg \boldsymbol{x}'] \Leftrightarrow [x_i > x_i' \text{ for all } i]$.

A3 (Continuity) For every $\boldsymbol{z} \in \mathcal{Z}(\boldsymbol{e})$, $\{\boldsymbol{z}' \in \mathcal{Z}(\boldsymbol{e}) : \boldsymbol{z}' \succeq_i \boldsymbol{z}\}$ and $\{\boldsymbol{z}' \in \mathcal{Z}(\boldsymbol{e}) : \boldsymbol{z}' \preceq_i \boldsymbol{z}\}$ are closed.

A4 (Nondecreasing marginal cost) $\mathcal{Y}$ is convex.

A5 (Differentiable utility) Preferences $\succeq_i$ can be represented by a differentiable utility function u_i.

A6 (Differentiable production) The function F is differentiable.

Denote the set of 'classical' economies satisfying A1 through A4 by $\mathcal{E}_I^C$. Let $\mathcal{E}_I^D$ denote the set of differentiable economies satisfying A1 through A6. Note that under A4 and A6, $c'(y) = 1/F'(T)$.

2.1 Mechanisms

A *mechanism* in this environment is a mapping from a strategy space $\mathcal{S} = \times_i \mathcal{S}_i$ into the set of allocations $\mathcal{Z}$. Given a strategy profile $\boldsymbol{s} \in \mathcal{S}$, the *transfer function* $\boldsymbol{\tau}(\boldsymbol{s})$ identifies a vector of net transfers, one for each agent, and the *outcome function* $\eta(\boldsymbol{s})$ identifies the level of the public good. The mechanism is denoted by $\Gamma = (\mathcal{S}, \boldsymbol{\tau}, \eta)$ and its space of outcomes is given by

$$\mathcal{O}_\Gamma(\boldsymbol{e}) = \{(\boldsymbol{x}, y) \in \mathcal{Z} : \exists\, \boldsymbol{s} \in \mathcal{S} \text{ s.t. } \boldsymbol{x} = \boldsymbol{\omega} - \boldsymbol{\tau}(\boldsymbol{s}) \ \&\ y = \eta(\boldsymbol{s})\}.$$

A *solution correspondence* (or *equilibrium correspondence*) $\mu_\Gamma(\boldsymbol{e})$ is a mapping from the environment $\boldsymbol{e}$ into a set of possible strategy profiles in $\mathcal{S}$, which depends on the mechanism Γ. For example, $\mu_\Gamma(\boldsymbol{e})$ may select all Nash equilibrium strategy profiles of the game induced by Γ. Given μ_Γ, we can define the space of *outcomes under* μ_Γ by

$$\mathcal{O}_\Gamma^\mu(\boldsymbol{e}) = \{(\boldsymbol{x}, y) \in \mathcal{Z} : \exists\, \boldsymbol{s} \in \mu_\Gamma(\boldsymbol{e}) \text{ s.t. } \boldsymbol{x} = \boldsymbol{\omega} - \boldsymbol{\tau}(\boldsymbol{s}) \ \&\ y = \eta(\boldsymbol{s})\}.$$

We say that Γ is *decisive under* μ if $\mathcal{O}_\Gamma^\mu(\boldsymbol{e}) \neq \emptyset$ for every $\boldsymbol{e} \in \mathcal{E}_I$, *feasible under* μ if $\mathcal{O}_\Gamma^\mu(\boldsymbol{e}) \subseteq \mathcal{Z}(\boldsymbol{e})$ for every $\boldsymbol{e}$, and *balanced under* μ if $\mathcal{O}_\Gamma^\mu(\boldsymbol{e}) \subseteq \bar{\mathcal{Z}}(\boldsymbol{e})$ for every $\boldsymbol{e}$.

An allocation $\boldsymbol{z}'$ Pareto dominates $\boldsymbol{z}$ if $\boldsymbol{z}' \succeq_i \boldsymbol{z}$ for all i and $\boldsymbol{z}' \succ_j \boldsymbol{z}$ for some j. The set of *Pareto optimal* allocations for $\boldsymbol{e}$ is given by

$$\mathcal{PO}(\boldsymbol{e}) = \{\boldsymbol{z} \in \mathcal{Z}(\boldsymbol{e}) : \nexists \boldsymbol{z}' \in \mathcal{Z}(\boldsymbol{e}) \text{ s.t. } \boldsymbol{z}' \text{ Pareto dominates } \boldsymbol{z}\}.$$

We say that Γ is *Pareto efficient under* μ if $\mathcal{O}^{\mu}_{\Gamma}(\boldsymbol{e}) \subseteq \mathcal{PO}(\boldsymbol{e})$ for every $\boldsymbol{e}$. Note that if preferences are strictly monotonic (assumption **A1'**), then $\mathcal{PO}(\boldsymbol{e}) \subseteq \bar{\mathcal{Z}}(\boldsymbol{e})$, so any mechanism that is Pareto efficient (under μ) must be balanced.[6]

2.2 Implementation

In general, a *social choice correspondence (SCC)* $\mathcal{G}$ is a mapping from economies $\boldsymbol{e}$ into a set of (feasible) allocations in $\mathcal{Z}(\boldsymbol{e})$. A mechanism Γ *(fully) implements* $\mathcal{G}$ *under* μ if $\mathcal{O}^{\mu}_{\Gamma}(\boldsymbol{e}) = \mathcal{G}(\boldsymbol{e})$ for every $\boldsymbol{e} \in \mathcal{E}_I$ and *partially implements* $\mathcal{G}$ *under* μ if $\mathcal{O}^{\mu}_{\Gamma}(\boldsymbol{e}) \subseteq \mathcal{G}(\boldsymbol{e})$ for every $\boldsymbol{e} \in \mathcal{E}_I$.

The mapping $\mathcal{PO}(\boldsymbol{e})$ defined above is an example of a particular SCC that identifies all Pareto optimal allocations of any economy $\boldsymbol{e}$. Thus, a mechanism partially implements $\mathcal{PO}(\boldsymbol{e})$ under μ if and only if it is Pareto efficient under μ. Another example of a SCC is the 'individually rational' SCC given by $\mathcal{IR}(\boldsymbol{e}) = \bigcap_i \mathcal{IR}_i(\boldsymbol{e})$, where

$$\mathcal{IR}_i(\boldsymbol{e}) = \{\boldsymbol{z} \in \mathcal{Z}(\boldsymbol{e}) : (\boldsymbol{x}, y) \succeq_i (\boldsymbol{\omega}, 0)\}.$$

This SCC selects all outcomes that are weakly preferred to the endowment by all agents. If some Γ implements $\mathcal{IR}(\boldsymbol{e})$ under μ, then all agents are made (weakly) better off—relative to the endowment—by participating in Γ and playing a strategy profile in $\mu_{\Gamma}(\boldsymbol{e})$.

Under certain assumptions on the class of admissible preferences and on the continuity of the mechanism, Hurwicz (1979) shows that if a mechanism implements $\mathcal{PO}(\boldsymbol{e}) \cap \mathcal{IR}(\boldsymbol{e})$ in Nash equilibrium, then $\mathcal{O}^{\mu}_{\Gamma}(\boldsymbol{e})$ is exactly the set of (interior) Walrasian (or Lindahl) allocations.[7]

2.3 The Participation Game

The standard view in mechanism design is that agents in economy $\boldsymbol{e}$ participate in a mechanism Γ by choosing strategy profile $\boldsymbol{s}$, which generates the outcome $(\boldsymbol{\omega} - \boldsymbol{\tau}(\boldsymbol{s}), \eta(\boldsymbol{s}))$. At this point, the game ends and the selected outcome is consumed by the agents.

[6] In the sequel, we drop the qualifier 'under μ' where there is no confusion or where the particular choice of μ is irrelevant.

[7] See also Hurwicz (1972) and Ledyard and Roberts (1975).

Suppose instead that each agent i, upon learning the mechanism outcome, has the freedom either to contribute $\tau_i(\boldsymbol{s})$ units of the private good or to exercise a 'no-trade' option by withholding his requested contribution $(\tau_i(\boldsymbol{s}))$, in which case he consumes his initial endowment ω_i of the private good. By withholding his contribution, i may reduce size of the public good, but since it is a pure public good he cannot be excluded from enjoying what is produced. After participation decisions are made, those who opt to contribute form a *contribution coalition*. Once we specify the level of public good that would result for each of the 2^I possible contribution coalitions, we have a well-defined *participation game* in which each agent simultaneously decides whether to contribute $\tau_i(\boldsymbol{s})$ or not. If a mechanism selects an outcome whose resulting participation game does not have every agent contributing in the participation game, then a mechanism designer who lacks coercive power should not expect the mechanism's outcome to be realized.

The focus of this paper is on those allocations for which full participation is a Nash equilibrium outcome. The following example shows how the participation game is constructed and how some allocations may not have full participation as an equilibrium outcome.

Example 1. Let $I = 2$ and define

$$u_1(\boldsymbol{x}, y) = x_1 + 21y - 2y^2$$

and

$$u_2(\boldsymbol{x}, y) = x_2 + 77y - 9y^2.$$

Fix $\omega_i = 50$ for each i and let $F(T) = T/10$. Any balanced allocation in which four units of the public good are provided is Pareto optimal. The unique Lindahl allocation for this economy (where each agent pays a per-unit price for the public good equal to their marginal valuation) corresponds to the equal-tax allocation $(t_1, t_2, y) = (20, 20, 5)$.

Suppose a social planner uses an incentive compatible mechanism to identify (correctly) the Lindahl allocation for this economy. The planner would then request that each agent pay 20 units of the private good. If both agents contribute as requested then the planner can build 4 units of the public good and the Lindahl allocation is realized. If only one agent contributes while the other does not, the realized public good level is reduced to $y = 2$. If neither contribute, no public good is produced. This simultaneous decision by the agents to contribute or not constitutes a two-player, two-strategy game whose payoff matrix is given in panel (a) of Figure I. Here it is a dominant strategy for agent 2 to contribute, but then agent 1 prefers to withhold her contribution, resulting in a suboptimal outcome of $y = 2$ in equilibrium.

t_1 \ t_2	20	0
20	82, 194	64, 168
0	84, 148	50, 50

(a)

t_1 \ t_2	10	0
30	72, 204	65, 200
0	69, 108	50, 50

(b)

Figure I: The induced participation game for Example 1 from (a) the equal-price Lindahl allocation, and (b) an unequal-price Pareto optimal allocation.

Now consider another efficient mechanism that selects $y = 4$ but splits the cost asymmetrically, so that $\boldsymbol{t} = (30, 10)$. If only agent 1 contributes then $y = 3$, and if only agent 2 participates then $y = 1$. In the resulting participation game, shown in panel (b) of Figure I, it is a dominant strategy for both agents to participate. By giving agent 1 a larger share of the production responsibility, this second mechanism has increased her benefit of non-contribution (30 units of private good) but has also increased her cost of non-contribution (3 units of public good). The concavity of preferences in the public good (and linearity of preferences for the private good) ensures that the larger public-good cost outweighs the larger private-good benefit.

Although this redistribution of production responsibility is an effective trick to offset free-riding incentives when preferences are convex, feasibility constraints limit how many agents can have their tax burden sufficiently increased in this way. Furthermore, some agents may prefer always to defect, regardless of how much of the burden they must bear. These difficulties are key to the negative results of the paper.

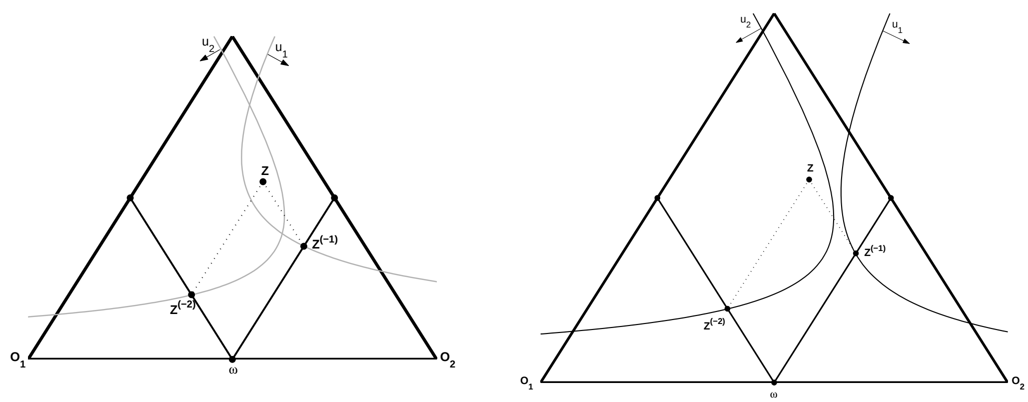
Consider now any economy with two players and a constant marginal cost. If an allocation z is proposed such that $t_i > 0$ for each i and $F(T) > 0$, then the allocation that obtains when agent 1 opts out is given by

$$\boldsymbol{z}^{(-1)} = (\omega_1, x_2, y^{(-1)}),$$

where

$$y^{(-1)} = F(t_2).$$

The opt-out point $\boldsymbol{z}^{(-2)}$ is similarly defined. Panel (a) of Figure II provides a graphical example of these points in the Kolm triangle diagram (Kolm (1970); see Thomson (1999) for a detailed exposition.) For the proposal $\boldsymbol{z}$ to satisfy *equilibrium participation*, both agents must prefer $\boldsymbol{z}$ to their 'opt-out' points $\boldsymbol{z}^{(-i)}$, as in the figure. In panel (b) of the figure, agent 1 prefers to opt out, resulting in the allocation $\boldsymbol{z}^{(-1)}$.



(a) (b)

Figure II: Opt-out allocations $\boldsymbol{z}^{(-1)}$ and $\boldsymbol{z}^{(-2)}$ where the original point $\boldsymbol{z}$ is preferred by both agents (panel (a)), and where one agent prefers to opt out (panel (b)).

In the case where $t_1 < 0$ while $t_2 > 0$, then $y^{(-2)} = 0$ since negative quantities of the public good are not admissible and $y^{(-1)} = y$ since agent 1 is not asked to contribute any private good. In this case, it is assumed that the negative transfer rejected by agent 1 is either redistributed among the other agents or destroyed, rather than affecting the level of the public good.[8] Under A1, agent 1 will always prefer participation when $t_1 < 0$ and agent 2 will prefer participation only if $(\boldsymbol{x}, y) \in \mathcal{IR}_2(\boldsymbol{e})$.

Generalizing this concept to allow for more players and arbitrary production technologies provides the key definition of this paper.

Definition 1. For any $I = 1, 2, \ldots$ and any economy $\boldsymbol{e} \in \mathcal{E}_I$, a feasible allocation $(\boldsymbol{x}, y) \in \mathcal{Z}(\boldsymbol{e})$ (with $\boldsymbol{t} = \boldsymbol{\omega} - \boldsymbol{x}$) satisfies *equilibrium participation for agent i* (EP$_i$) if

[8] Whether the transfer is redistributed or destroyed will not affect the i's participation decision since $\succeq_i$ depends only on x_i and y.

and only if

$$(\boldsymbol{x}, y) \succeq_i (\boldsymbol{x}^{(-i)}, y^{(-i)}),$$

where

$$x_i^{(-i)} = \omega_i,$$

$$y^{(-i)} = \begin{cases} F(T_{-i}) & \text{if} \quad t_i \geq 0,\, T_{-i} \geq 0, \text{ and } y \geq F(T_{-i}) \\ 0 & \text{if} \quad T_{-i} < 0 \\ y & \text{otherwise} \end{cases}, \tag{1}$$

and

$$(\boldsymbol{x}^{(-i)}, y^{(-i)}) \in \mathcal{Z}(\boldsymbol{e}).$$

The allocation $(\boldsymbol{x}, y) \in \mathcal{Z}(\boldsymbol{e})$ satisfies *equilibrium participation (EP)* if and only if it satisfies EP_i for all $i \in \mathcal{I}$.

There are three cases considered in this definition. When $t_i \geq 0$, $T_{-i} \geq 0$, and $y \geq F(T_{-i})$, removing agent i's transfer necessarily reduces production, but not to zero. If $T_{-i} < 0$, then $t_i > 0$ and removing i's transfer results in $y^{(-i)} = 0$. If $t_i < 0$ or $y < F(T_{-i})$, then y can be produced in the absence of i's transfer, so $y^{(-i)} = y$.

For any economy $\boldsymbol{e} \in \mathcal{E}_I$, let

$$\mathcal{EP}_i(\boldsymbol{e}) = \{\boldsymbol{z} \in \mathcal{Z}(\boldsymbol{e}) : \boldsymbol{z} \text{ satisfies } \mathrm{EP}_i\},$$

and define

$$\mathcal{EP}(\boldsymbol{e}) = \bigcap_{i \in \mathcal{I}} \mathcal{EP}_i(\boldsymbol{e}).$$

Thus, $\mathcal{EP}(\boldsymbol{e})$ is a well-defined SCC. Referring back to the example of Figure II, $\boldsymbol{z} \in \mathcal{EP}(\boldsymbol{e})$ in panel (a), but in panel (b), $\boldsymbol{z} \notin \mathcal{EP}_1(\boldsymbol{e})$, so $\boldsymbol{z} \notin \mathcal{EP}(\boldsymbol{e})$.

The SCC $\mathcal{EP}(\boldsymbol{e})$ represents the set of allocations that are stable (in the sense of Nash equilibrium) when agents are free to exercise a 'no-trade' alternative. If a social planner wishes to implement some other SCC $\mathcal{G}(\boldsymbol{e})$ but must also allow agents this no-trade alternative, then the social planner must choose a mechanism that (at least partially) implements $\mathcal{G}(\boldsymbol{e}) \cap \mathcal{EP}(\boldsymbol{e})$. (This can be compared to the more standard requirement that the planner implement $\mathcal{G}(\boldsymbol{e}) \cap \mathcal{IR}(\boldsymbol{e})$.) Obviously, a mechanism (partially) implements $\mathcal{G}(\boldsymbol{e}) \cap \mathcal{EP}(\boldsymbol{e})$ only if it partially implements $\mathcal{EP}(\boldsymbol{e})$, but the following results indicate that $\mathcal{EP}(\boldsymbol{e})$ has some undesirable properties; specifically, for many interesting SCC's $\mathcal{G}$ the correspondence $\mathcal{G}(\boldsymbol{e}) \cap \mathcal{EP}(\boldsymbol{e})$ is either empty for many values of $\boldsymbol{e}$ or converges to $\{\boldsymbol{\omega}\}$ (the endowment allocation) as the economy becomes large.

It is natural to consider more complex participation games where agents have more options than simply "contribute or not". For example, agents may be free to choose *any* level of contribution after observing the mechanism's recommended allocation (as in Bergstrom, Blume, and Varian, 1986, e.g.). In general, if the actual participation game includes additional strategies, then allocations that were supportable as equilibrium outcomes of the participation game with two strategies may no longer be supportable when more strategies are added. It must be the case, however, that if an allocation $\boldsymbol{z}$ does *not* satisfy EP when there are only two strategies, then it also cannot be supported as an equilibrium outcome of a participation game with a larger strategy space. Thus, the set of supportable allocations when more strategies are available is a subset of $\mathcal{EP}(\boldsymbol{e})$. Since most of the results regarding $\mathcal{EP}(\boldsymbol{e})$ are negative, further restricting this correspondence will only lead to stronger negative results.

3 Properties of $\mathcal{EP}(\boldsymbol{e})$

The shaded region of Figure III demonstrates a typical equilibrium participation set for agent 1 in a two-agent classical economy. Note that with continuous preferences (A3) each set $\mathcal{EP}(\boldsymbol{e})$ is closed and has a continuous boundary, but need not be convex. Clearly, $\mathcal{EP}(\boldsymbol{e})$ is non-empty for every $\boldsymbol{e} \in \mathcal{E}_I$ and every I since $(\boldsymbol{\omega}, 0) \in \mathcal{EP}(\boldsymbol{e})$.

As an alternative to equilibrium participation, consider the environment of Bergstrom, Blume, and Varian (1986) in which agents can freely (and simultaneously) choose $t_i \in [0, \omega_i]$, resulting in $y = F(T)$. The set of Nash equilibrium allocations of this larger participation game is given by

$$\begin{aligned}\mathcal{NE}(\boldsymbol{e}) = \{(\boldsymbol{x}^*, y^*) \in \mathcal{Z}(\boldsymbol{e}) : \boldsymbol{x}^* \leq \boldsymbol{\omega} \text{ and} \\ [\forall i \in \mathcal{I}][\forall \boldsymbol{t}' \geq 0]\ (\boldsymbol{x}^*, y^*) \succeq_i (\boldsymbol{\omega} - \boldsymbol{t}', F(T^*_{-i} + t'_i))\}.\end{aligned}$$

The notion of equilibrium participation is now shown to be more stringent than the standard notion of individual rationality, but less restrictive than the Nash equilibrium requirement.

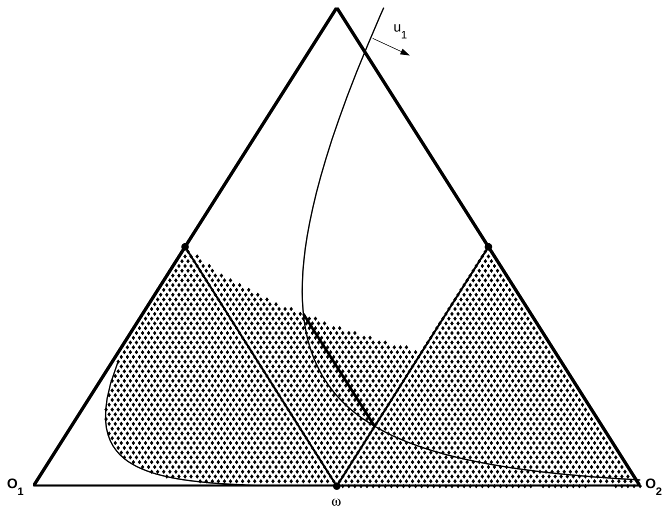


Figure III: An example of the set $\mathcal{EP}_1(\boldsymbol{e})$ for a two-agent economy with convex preferences.

Proposition 1. Under monotone increasing preferences (A3), all allocations satisfying equilibrium participation also satisfy individual rationality ($\mathcal{EP}(\boldsymbol{e}) \subseteq \mathcal{IR}(\boldsymbol{e})$ for all $\boldsymbol{e}$.)

Proof. Consider a point $(\boldsymbol{x}, y)$ such that $(\boldsymbol{x}, y) \succeq_i (\boldsymbol{\omega}, y^{(-i)})$ for all $i \in \mathcal{I}$. Note that $y^{(-i)} \geq 0$ for each i, so A3 implies that $(\boldsymbol{\omega}, y^{(-i)}) \succeq_i (\boldsymbol{\omega}, 0)$. The result follows by transitivity. $\square$

Proposition 2. All Nash equilibria of the voluntary contributions game satisfy equilibrium participation ($\mathcal{NE}(\boldsymbol{e}) \subseteq \mathcal{EP}(\boldsymbol{e})$ for all $\boldsymbol{e}$.)

Proof. From any Nash equilibrium point, the allocation consumed by i in the participation game should i choose unilaterally to drop out is simply $(\omega_i, F(T^*_{-i}))$. Since the definition of Nash equilibrium requires that $(x^*_i, y^*) \succeq_i (\omega_i, F(T^*_{-i}))$ for all i (by letting $t'_i = 0$), the point $(\boldsymbol{x}^*, y^*)$ must satisfy equilibrium participation. $\square$

The most common assumption in the mechanism design literature with public goods is that the social planner aims to implement $\mathcal{PO}(\boldsymbol{e})$. There exist several mechanisms that implement this SCC under Nash equilibrium, including Groves and Ledyard (1977) and Walker (1981); however, if the outcomes of these mechanism fail to satisfy equilibrium participation (*i.e.*, if they don't partially implement

$\mathcal{PO}(\boldsymbol{e}) \cap \mathcal{EP}(\boldsymbol{e})$) then their desirable properties are of little use in environments where agents cannot be coerced to submit their transfers. The following example shows the potential difficulty of finding points in $\mathcal{PO}(\boldsymbol{e}) \cap \mathcal{EP}(\boldsymbol{e})$.

Example 2. Let $I \geq 2$. Define $u_i(\boldsymbol{x}, y) = v_i(y) + x_i$, where each v_i is concave and differentiable with $v_i'(y) \in (0, \kappa)$ for all $i \in \mathcal{I}$ and $y \geq 0$ and let the unique maximizer of $\sum_i v_i(y)$ be some $y^o > 0$ satisfying $\sum_{j \neq i} \omega_j < \kappa y^o$ for each i. Assume a constant marginal cost of production κ, so that $F(T) = T/\kappa$.

In this setting no agent is willing to unilaterally fund any amount of the public good at any level (since $v'(y) < \kappa$ for all y) and therefore any agent asked to contribute a positive amount will refuse to contribute in any participation game unless the proposed allocation is the endowment. To see this, pick any allocation $(\boldsymbol{x}, y) \neq (\boldsymbol{\omega}, 0)$, so $\boldsymbol{t} \neq \boldsymbol{0}$ and $\sum_i t_i \geq 0$. Let i be any agent such that $t_i > 0$ (there must be at least one). If i participates he receives

$$u_i(\boldsymbol{x}, y) = v_i(y) + \omega_i - t_i.$$

If i withholds his transfer he receives

$$u_i(\boldsymbol{x}^{(-i)}, y^{(-i)}) = v_i(y^{(-i)}) + \omega_i.$$

If $T_{-i} < 0$ then $y^{(-i)} = 0$ by Definition 1. Thus,

$$u_i(\boldsymbol{x}, y) - u_i(\boldsymbol{x}^{(-i)}, y^{(-i)}) = \int_0^y v_i'(\tilde{y}) \; d\tilde{y} - t_i,$$

which is strictly less than $\kappa y - t_i$ since $v_i' < \kappa$ everywhere. But feasibility requires that $t_i > \kappa y - T_{-i}$, so the difference in utilities is strictly negative. Therefore, i strictly prefers not to participate.

If $t_i \geq 0$, $T_{-i} \geq 0$, and $y \geq F(T_{-i})$ then $y^{(-i)} = F(T_{-i})$. Here,

$$u_i(\boldsymbol{x}, y) - u_i(\boldsymbol{x}^{(-i)}, y^{(-i)}) = \int_{F(T_{-i})}^y v_i'(\tilde{y}) \; d\tilde{y} - t_i,$$

which is strictly less than $\kappa(y - F(T_{-i})) - t_i$ because $v_i' < \kappa$ everywhere. Since $\kappa F(T_{-i}) = T_{-i}$, this expression becomes $\kappa y - T$, which must be weakly negative by feasibility. Thus, i strictly prefers not to participate.

Finally, if neither of the two above conditions holds then $y^{(-i)} = y$ and i strictly prefers not to participate since his participation has no effect on the public good but participation means surrendering some positive quantity of the private good.

In any economy with these specifications no allocation other than the endowment can satisfy EP_i for every i, so $\mathcal{EP}(\boldsymbol{e}) = \{(\boldsymbol{\omega}, 0)\}$.

The above example proves the following proposition.

Proposition 3. For every $I \geq 2$, there exists an open set of economies in $\mathcal{E}_I^C$ such that no allocation except the endowment satisfies equilibrium participation ($\mathcal{EP}(\boldsymbol{e}) = \{(\boldsymbol{\omega}, 0)\}$).

Given this negative result about $\mathcal{EP}(\boldsymbol{e})$, a social planner interested in implementing $\mathcal{G}(\boldsymbol{e})$ might instead look to pick an allocation $\boldsymbol{z}$ such that a Nash equilibrium of the participation games leads to a new allocation $\boldsymbol{z'}$ with $\boldsymbol{z'} \in \mathcal{G}(\boldsymbol{e})$. In other words, the planner might foresee agents' non-participation and try to pick an allocation such that non-participation will lead to a desirable result. Unfortunately, the following proposition verifies that this trick gains the planner no additional flexibility since the allocation that will obtain after the participation game is played must satisfy equilibrium participation.

Proposition 4. If an allocation $\boldsymbol{z}$ is a pure strategy equilibrium outcome of the (two-strategy) participation game in economy $\boldsymbol{e}$, then $\boldsymbol{z}$ satisfies $\mathcal{EP}(\boldsymbol{e})$.

Proof. For any allocation $\boldsymbol{z}$ and any subset of agents $\mathcal{S} \subseteq \mathcal{I}$, let $\boldsymbol{z}^{\mathcal{S}}$ be the allocation that obtains if all $i \in \mathcal{S}$ contribute and all $j \in \mathcal{I} \setminus \mathcal{S}$ do not. For any pure strategy equilibrium of the game induced by $\boldsymbol{z}$, the outcome is an allocation $\boldsymbol{z}^{\mathcal{S}}$ for some subset $\mathcal{S}$ of agents who chose to contribute. Since $\mathcal{S}$ is a Nash equilibrium participation coalition, we have $\boldsymbol{z}^{\mathcal{S}} \succeq_i \boldsymbol{z}^{\mathcal{S}\setminus\{i\}}$, thus $\boldsymbol{z}^{\mathcal{S}} \in \mathcal{EP}_i(\boldsymbol{e})$ for each $i \in \mathcal{S}$. Since $\boldsymbol{z}^{\mathcal{S}} = \boldsymbol{z}^{\mathcal{S}\setminus\{j\}}$ for each $j \notin \mathcal{S}$ we also have $\boldsymbol{z}^{\mathcal{S}} \in \mathcal{EP}_j(\boldsymbol{e})$ for each $j \notin \mathcal{S}$. Thus, $\boldsymbol{z}^{\mathcal{S}} \in \mathcal{EP}(\boldsymbol{e})$. $\square$

Combining Propositions 3 and 4 and recalling that the endowment in Example 2 is not Pareto optimal gives the following negative result.

Proposition 5. For every $I \geq 2$, there exists an open set of economies in $\mathcal{E}_I^C$ in which no allocation can be selected such that the equilibrium of the resulting participation game is Pareto optimal.

Since Proposition 5 indicates that EP is often inconsistent with Pareto optimality, it is natural to ask whether there can exist *any* non-trivial mechanisms that satisfy this constraint.[9] In other words, is there a mechanism and a μ that implements $\mathcal{EP}(\boldsymbol{e})$ in μ? It is simple to show that $\mathcal{EP}(\boldsymbol{e})$ satisfies the definition of monotonicity from Maskin (1999), giving the following result.[10]

Proposition 6. The set of allocations satisfying equilibrium participation ($\mathcal{EP}(\boldsymbol{e})$) can be non-trivially implemented in Nash equilibrium when $I \geq 3$.

The proof of this proposition for full implementation relies on Maskin's mechanism, which is not a particularly natural game form. Proposition 2 shows that $\mathcal{EP}(\boldsymbol{e})$ can be partially implemented using the voluntary contribution mechanism since $\mathcal{NE}(\boldsymbol{e}) \subseteq \mathcal{EP}(\boldsymbol{e})$ or by using the trivial mechanism since $(\boldsymbol{\omega}, 0) \in \mathcal{EP}(\boldsymbol{e})$ for every $\boldsymbol{e}$. Note that in the economies like those of Example 2, $\mathcal{EP}(\boldsymbol{e}) = \{(\boldsymbol{\omega}, 0)\}$, so any mechanism that partially implements $\mathcal{EP}(\boldsymbol{e})$ must pick only the initial endowment in those cases.

4 Quasi-Concave Economies

4.1 Necessary and Sufficient Conditions

The additional structure gained by adding assumptions A1 through A6 allows for the derivation of separate necessary and sufficient conditions for an allocation to satisfy equilibrium participation. Although these conditions are not tight, they require only 'local' information about marginal rates of substitution and the marginal cost of production.

Proposition 7. For any economy in $\mathcal{E}_I^D$, if equilibrium participation is satisfied at a point $(\boldsymbol{x}, y) = (\boldsymbol{\omega} + \boldsymbol{t}, y)$, then for each $i \in \mathcal{I}$ such that $t_i, T_{-i} \geq 0$ and $y \geq F(T_{-i})$,

$$\frac{\partial u_i(\boldsymbol{\omega}, F(T_{-i}))/\partial y}{\partial u_i(\boldsymbol{\omega}, F(T_{-i}))/\partial x_i} \geq c'(y^{(-i)}). \tag{2}$$

[9] A trivial mechanism selects the endowment in every economy; a non-trivial mechanism is not trivial.

[10] The other sufficient condition, 'no-veto power' is trivially satisfied in economic environments such as this.

See the appendix for proofs of Propositions 7, 8, and 9.

A similar condition is now shown to be sufficient for a point to satisfy equilibrium participation. Whereas the necessary condition compares the marginal rate of substitution to marginal costs at the drop-out point, the sufficient condition compares these quantities at the proposed allocation.

Proposition 8. For any economy in $\mathcal{E}_I^D$, equilibrium participation is satisfied at $(\boldsymbol{x}, y)$ if for all i with $t_i, T_{-i} \geq 0$ and $y \geq F(T_{-i})$,

$$\frac{\partial u_i(\boldsymbol{x}, y)/\partial y}{\partial u_i(\boldsymbol{x}, y)/\partial x_i} \geq c'(y), \tag{3}$$

and for all j such that $T_{-j} < 0$,

$$u_j(\boldsymbol{x}, y) \geq u_j(\boldsymbol{\omega}, 0). \tag{4}$$

Unlike the necessary condition, equation (4) implies that information about the utilities of some agents at both the suggested allocation and the endowment is needed. This may be undesirable from the standpoint of mechanism design since additional information is necessary to determine that the condition is met.[11] The following condition shows how equation (4) could be replaced by a stronger version of equation (3) to give a single condition sufficient for all agents that uses only information about preferences and costs at the selected allocation.

Proposition 9. For any economy in $\mathcal{E}_I^D$, if a point $(\boldsymbol{x}, y) = (\boldsymbol{\omega} + \boldsymbol{t}, y)$ satisfies

$$\frac{\partial u_i(\boldsymbol{x}, y)/\partial y}{\partial u_i(\boldsymbol{x}, y)/\partial x_i} \geq \frac{t_i}{F(T)} \tag{5}$$

for all i, then equilibrium participation is satisfied at $(\boldsymbol{x}, y)$.

Figure IV demonstrates the interpretation of these conditions. The quantity $(\partial u_i/\partial y)/(\partial u_i/\partial x_i)$ is the slope of the gradient of u_i, while c' is the slope of the normal to the production possibilities frontier. In the figure, F is horizontally shifted so that its graph represents the production possibilities set for agent i given her endowment and T_{-i}. If agent i withholds t_i, then the allocation $z^{(-i)}$ results. In

[11] Of course, there could exist mechanisms whose outcomes satisfy equilibrium participation without satisfying this sufficient condition.

this case, i will prefer the Pareto optimal point z to $z^{(-i)}$, so z satisfied equilibrium participation for agent i.

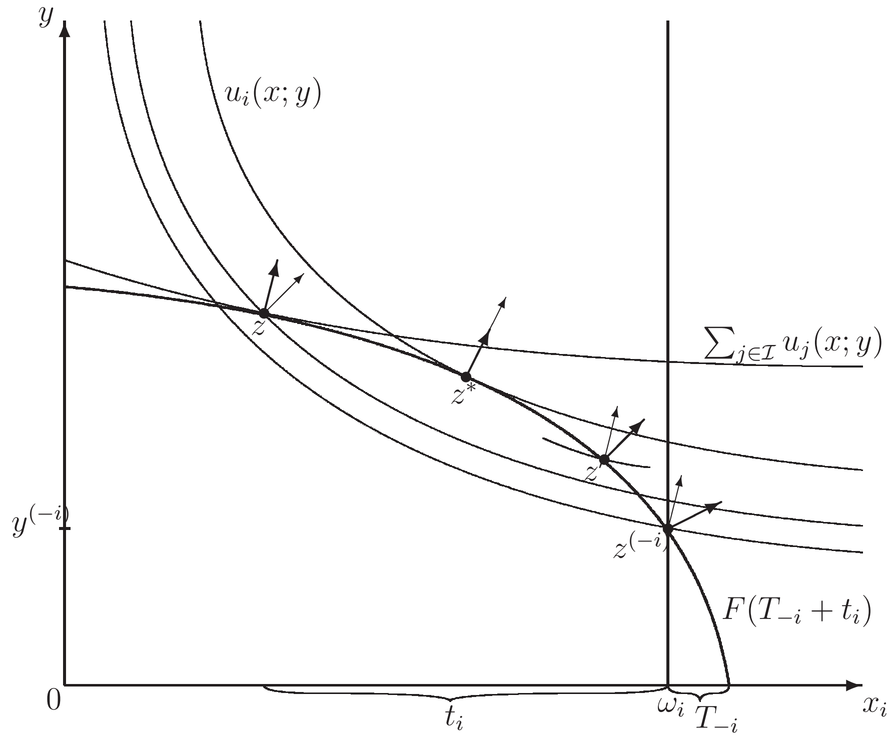


Figure IV: An example with quasi-concave utilities and convex production sets. z is Pareto optimal, z^* is i's most-preferred feasible allocation, and $z^{(-i)}$ is i's drop-out point. $z^{(-i)}$ satisfies the sufficient condition for EP. z^* and z' satisfy the sufficient condition.

The necessary condition for equilibrium participation is satisfied at z in the figure since the gradient of utility has a steeper slope than the normal to F at $\boldsymbol{z}^{(-i)}$. The sufficient condition is satisfied at $\boldsymbol{z}'$ since the gradient of utility is steeper than the normal to F at $\boldsymbol{z}'$, but this condition fails at the optimal point z. In fact, the sufficient condition is satisfied for any point along F between $\boldsymbol{z}^{(-i)}$ and $\boldsymbol{z}^*$, but nowhere left of $\boldsymbol{z}^*$. This is intuitive; $\boldsymbol{z}'$ is closer to $\boldsymbol{z}^*$ (i's most preferred point) than $\boldsymbol{z}^{(-i)}$, so i will not opt out of $\boldsymbol{z}'$.

The Samuelson (1954) condition for an interior optimum forces $\boldsymbol{z}$ to be to the left of $\boldsymbol{z}^*$, where the sufficient condition fails. Thus, equilibrium participation requires that $\boldsymbol{z}^{(-i)}$ be sufficiently to the right of $\boldsymbol{z}^*$, causing t_i to be large. As in the opening example, large transfers are needed to incentivize participation, but feasibility may constrain how large the transfer can be or how many agents can have these inflated transfers. Clearly, this constraint will be more restrictive in larger economies, as will be demonstrated in Section 5.

4.2 Quasi-Linear Preferences

By assuming quasilinear preferences, one can focus on the size of transfer necessary to satisfy a particular constraint, such as an incentive constraint, without concern for how that transfer will affect preferences via wealth effects. In this way, quasilinear environments are useful in identifying the transfer needed to exactly satisfy the participation constraint implied by the equilibrium participation concept.

Assume agents have utility functions $u_i(\boldsymbol{x}, y) = v_i(y) + x_i$, where $v_i' > 0$ and $v'' \leq 0$, and let the production function be strictly increasing and concave, so $c(y)$ is strictly increasing and convex. Let y_i^* be the unique solution to $c'(y) = v_i'(y)$.

Equilibrium participation at a public good level of $\hat{y}$ requires that $v_i(\hat{y}) - t_i \geq v_i(\hat{y}^{(-i)})$, or

$$t_i \leq \int_{\hat{y}^{(-i)}}^{\hat{y}} v_i'(y)dy.$$

Feasibility requires that if t_i is non-negative then $c(\hat{y}) \leq c(\hat{y}^{(-i)}) + t_i$, or

$$\int_{\hat{y}^{(-i)}}^{\hat{y}} c'(y)dy \leq t_i,$$

with equality if the allocation is non-wasteful. Putting these together, if $\hat{y} > y_i^*$ and $\hat{y}$ satisfies equilibrium participation, then

$$\int_{\hat{y}^{(-i)}}^{\hat{y}} \left(v_i'(y) - c'(y)\right) dy \geq 0,$$

or

$$\int_{\hat{y}^{(-i)}}^{y_i^*} (v_i'(y) - c'(y))dy \geq \int_{y_i^*}^{\hat{y}} (c'(y) - v_i'(y))dy, \tag{6}$$

both of which are non-negative quantities.

For an optimal allocation y^o (for which we know that $y^o > y_i^*$), equation (6) provides an exact requirement on how 'far' $y^{(-i)}$ must be from y_i^* to guarantee equilibrium participation. This is demonstrated in Figure V, where $y^{(-i)}$ is the largest value satisfying (6) for the optimal point y^o. The necessary and sufficient conditions from equations (2) and (3) are also intuitive in this figure; if $y^{(-i)} > y_i^*$, then the necessary condition fails because marginal costs are everywhere larger than the marginal benefit between $y^{(-i)}$ and y^o, and the sufficient condition is satisfied for any $y \in [y^{(-i)}, y_i^*)$ since marginal costs are everywhere less than the marginal benefit between y and $y^{(-i)}$.

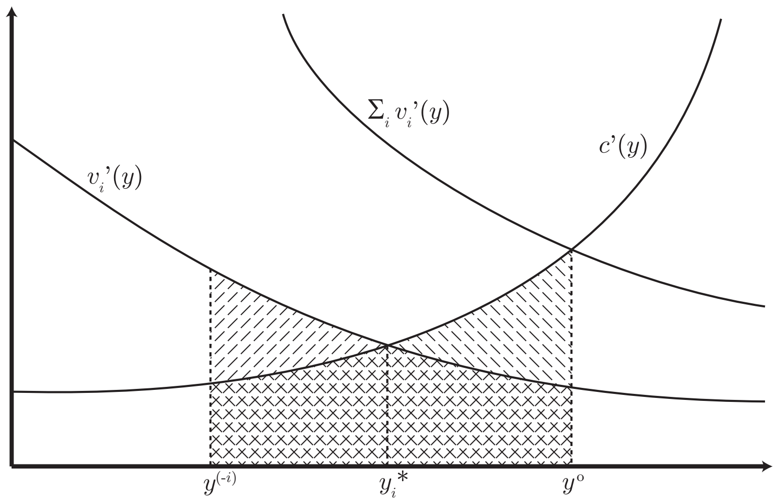


Figure V: The largest $y^{(-i)}$ such that the Pareto optimal public good level y^o satisfies EP_i.

5 Equilibrium Participation in Large Economies

The typical approach for analyzing large economies is to consider replications of some finite economy that converges to an economy with a continuum of agents. With public goods, however, allocations in the sequence of replica economies do not converge to equivalent allocations in the continuum economy. For example, if individual contributions in the sequence of replica economies is bounded away from zero for an unbounded set of agents then public good levels must diverge to infinity as the sequence grows. If, on the other hand, the public good level stays bounded between zero and infinity then individual contributions must be bounded away from zero for only a finite set of agents. But in the limit economy the contributions of a finite set of agents has no impact on production and the public good level is zero. Only when no public goods are produced in the limit does the allocation in the continuum economy represent a sensible limit of the sequence of allocations in the replica economies.

Given these issues, we analyze continuum economies separately from replica economies. The results on equilibrium participation for continuum economies are both trivial and pessimistic. Following the definition of a continuum economy given by Muench (1972), let the set of agents be given by $\mathcal{I} = [0, 1]$ and the endowments and private good contributions by $\omega(i)$ and $x(i)$, respectively. Feasibility requires that the public good level y satisfy $\int_{\mathcal{I}}(\omega(i) - x(i))di \geq c(y)$. Since each agent i's contribution has an infinitesimal effect on y, no agent will contribute any positive amount to the public good when preferences for the private good are strictly monotonic. This gives our first result for large economies:

Proposition 10. In any limit economy with a continuum of agents and strictly increasing preferences (A1'), the only allocation satisfying equilibrium participation is the endowment ($\mathcal{EP}(\boldsymbol{e}) = \{(\omega(i), 0)\}$).

To analyze finitely replicated economies, we start with some base economy $\boldsymbol{e}^1 \in \mathcal{E}_I$ with I agents, where $\boldsymbol{e}^1 = (\{\succeq_i\}_{i=1}^{I}, \mathcal{Y}, \boldsymbol{\omega})$. A replica economy $\boldsymbol{e}^R$ is then defined by replicating each agent R times. Let the r^{th} replicate (for $r = 1, \ldots, R$) of the i^{th} agent be denoted by (i, r). Agent (i, r) is endowed with $\omega_{i,r} = \omega_i$ units of the private good and has preferences $\succeq_{i,r}$ such that $(\boldsymbol{x}, y) \succeq_{i,r} (\boldsymbol{x}', y)$ if and only if $(\boldsymbol{x}_{\cdot r}, y) \succeq_i (\boldsymbol{x}'_{\cdot r}, y)$, where $\boldsymbol{x}_{\cdot r} = (x_{1,r}, \ldots, x_{I,r})$ when $\boldsymbol{x} \in \mathbb{R}^{I \cdot R}$. Let $\boldsymbol{\omega}^R = (\boldsymbol{\omega}, \ldots, \boldsymbol{\omega})$ be the $R \times I$-dimensional vector of endowments for $\boldsymbol{e}^R$. The production technology remains unchanged as the economy is replicated.

The following proposition gives a result quite the opposite from Proposition 10; here, the size of the equilibrium participation can only grow as an economy is replicated.

Proposition 11. For any base economy $\boldsymbol{e}^1$, public good level $y \geq 0$, and number of replications $R \in \{1, 2, \ldots\}$, if there is some $\boldsymbol{x} \in \mathbb{R}^{I \cdot R}$ such that $(\boldsymbol{x}, y) \in \mathcal{EP}\left(\boldsymbol{e}^R\right)$, then there is some $\boldsymbol{x}' \in \mathbb{R}^{I \cdot (R+1)}$ such that $(\boldsymbol{x}', y) \in \mathcal{EP}\left(\boldsymbol{e}^{R+1}\right)$.

The proof of the proposition is in fact trivial; if $(\boldsymbol{x}, y)$ satisfies equilibrium participation in $\boldsymbol{e}^R$ and new agents are added, simply do not change the contribution levels for the old agents and do not ask for contributions from the new ones.

One objection to Proposition 11 is that it uses allocations that violate the equal-treatment property; old agents consume allocations that are significantly different from their offspring. The following example shows, however, that non-trivial levels of the public good can be produced in any finite economy while satisfying both equilibrium participation and the equal treatment property.

Example 3. Let $\boldsymbol{e}^1$ be an economy of I identical agents with $u_i(x_i, y) = \ln(y) + x_i$ and let $c(y) = y$. For each replication R of $\boldsymbol{e}^1$, consider the allocation where each agent contributes $t_R^* = \ln(RI/(RI-1))$ units of the private good. Since $\ln(RIt_R^*) - t_R^* = \ln(RIt_R^* - t_R^*)$, all agents are indifferent between contributing t_R^* and not. Therefore, full participation is an equilibrium of the induced participation game for every R. As R grows, t_R^* approaches zero but $y = RIt_R^*$ converges to one from above.

Thus, pure public goods may be achievable in arbitrarily large finite economies without coercion.

Milleron (1972) suggests an alternative notion of replication in which a fixed endowment is split into successively smaller shares as the economy is replicated.[12] This forces private good consumption to shrink rather than allowing public good levels to diverge. It also captures the idea that each agent becomes "insignificant" as the economy grows large.

As before, we begin with a base economy $\boldsymbol{e}^1$ and replicate it R times, with the r^{th} replica of the i^{th} agent being denoted by (i, r). The difference here is that $\omega_{i,r} = \omega_i/R$, so that the total endowment of the private good always sums to $\sum_i \omega_i$.

Milleron (1972) also adjusts the preferences of the replicates to compensate for the fact that their feasible consumption set shrinks as R grows. Specifically, he defines $\succeq_{i,r}$ by $(\boldsymbol{x}_r, y) \succeq_{i,r} (\boldsymbol{x}'_r, y')$ if and only if $(R\boldsymbol{x}, y) \succeq_i (R\boldsymbol{x}', y')$. With this specification, agents value their private good consumption only as a fraction of their total endowment rather than on an absolute scale. This is useful if, for example, i's marginal utility for the private good diverges to infinity as his consumption drops to zero; without rescaling replicates' preferences, all agents would eventually be driven to demand nothing but private goods.

We show that when agents' endowments shrink with replication, each truly be-

[12]This replication method is also used in Furusawa and Konishi (2007).

comes insignificant in the production of the public good and will strictly prefer non-participation in large enough economies.

Proposition 12. For any economy satisfying A1, A3, and A4 (continuous, monotone preferences and increasing, continuous production technology) and with endowment-splitting replication, the set of allocations satisfying equilibrium participation converges to the initial endowment as the economy is infinitely replicated.

The proof of this theorem, available in the appendix, demonstrates how the shrinking endowment restricts the amount any agent can be asked to pay in the limit. This, in turn, limits the agent's effect on production. Since agents in large economies care about small changes in their private goods consumption but not in the level of the public good, agents eventually prefer to opt-out as their individual effect on production vanishes.

In many base economies the set of Pareto optimal allocations remains bounded away from the endowment as the economy grows; thus, equilibrium participation not only precludes efficient allocations in these settings, but often precludes any notion of approximate efficiency. For these large economies, it is necessary that the social planner have the power of coercion in order to overcome the free-rider problem.

6 Related Literature

The notion of equilibrium participation is similar in spirit to the core of a public goods economy. Both define a stability property based on comparisons between the proposed allocation and feasible defections. There are many possible definitions of the core, however, because it is unclear what allocations might result (or, how non-dissenting coalitions might behave) when a coalition blocks an allocation.

In the original definition of the core by Foley (1970), only the dissenting coalition may produce the public good; non-dissenters withdraw their contributions to production. This maximizes the threat to dissenters and many allocations remain in the core.[13] Richter (1974) assumes that non-dissenting agents select levels of production that are 'rational' for themselves (under various meanings) and finds that the subsequent definition of the core may be empty.

[13] Muench (1972) shows that Foley's core does not converge to the set of Lindahl equilibria in large economies.

Champsaur, Roberts, and Rosenthal (1975) define the φ-core as the allocations that remain unblocked when blocking coalitions are given the power to tax the remaining agents an amount up to φ, which depends on the *proposed* blocking allocation. If φ were a function of the *original* allocation, then this notion of blocking (for single-agent coalitions) could encompass the definition of equilibrium participation. Though the results for both definitions are similarly negative, they are logically independent.

Saijo (1991) analyzes the mechanism design problem if the utility of autarkic production is used as a welfare lower bound instead of the utility of the endowment. His notion of *autarkic individual rationality* requires each agent's final utility level to be weakly greater than that which the agent could achieve in isolation with his endowment and access to the production technology. Whereas Ledyard and Roberts (1975) demonstrate that the standard notion of individual rationality is incompatible with incentive compatibility among the class of Pareto optimal mechanisms, Saijo (1991) shows that autarkic individual rationality is incompatible with incentive compatibility for *all* mechanisms, optimal or not.

There have been other papers examining explicit outside options of agents in mechanism design. The most general of these is Jackson and Palfrey (2001), where an unspecified function maps from any given outcome to another (possibly identical) outcome. The necessary and sufficient conditions of Maskin (1999) are then extended in a simple way to accommodate this 'reversion function'. This approach unifies several existing attempts to model renegotiation and participation in the outcomes of mechanisms in private goods settings, such as Ma, Moore, and Turnbull (1988), Maskin and Moore (1999), and Jackson and Palfrey (1998). It also encompasses pubic goods models with an exogenous status quo outcome or mechanism, as in Perez-Nievas (2002). A selection from the equilibrium correspondence of the participation game defined in this paper would be another example of a reversion function, and, for particular environments, the conditions of Jackson and Palfrey (2001) could be analyzed to determine if the resulting allocations (after participation decisions are made) are implementable.

Saijo and Yamato (1999) and Saijo and Yamato (2008) focus on the question of whether forward-looking agents would want to participate in a mechanism if its designer did have coercive power in enforcing the mechanism outcome. Here, if an agent opts out of a mechanism that implements Lindahl allocations then her preferences will not be used to determine the optimal public good level, but she also will not be required to pay any transfers. Saijo and Yamato (2008) show that any mechanism that implements Lindahl allocations will suffer from participation problems when the domain of preferences is sufficiently rich. Thus, the results on participation in a mechanism are similar in spirit to our results on participation in a mechanism's outcome.

Finally, it is worth noting that concepts such as dominant strategy incentive compatibility and ex-post equilibrium do not encompass the definition of equilibrium participation. Although these concepts do require that the mechanism outcome be preferred by each individual to all other outcomes in the range of the mechanism, there is no guarantee that the allocation obtaining after an agent opts out is in the mechanism's range. Indeed, most 'standard' public goods mechanisms (such as those of Groves and Ledyard (1977) or Groves (1973)) do not include the opt-out points in their range. Therefore, the fact that an allocation is selected as part of an equilibrium decision does not preclude the possibility that agents will later prefer to free-ride on the contributions of others.

7 Conclusion

If a mechanism is to implement a desired social choice correspondence with public goods when agents have available a no-trade alternative, it must select an allocation impervious to agents withdrawing their transfers. The incompatibility between equilibrium participation and Pareto optimality is established through simple quasi-linear examples, indicating that optimality is unobtainable under the standard assumptions used in mechanism design. In many economies, only the initial endowment is insusceptible to agents withdrawing. Even in those economies for which non-trivial allocations satisfy equilibrium participation, the set of equilibrium participation allocations eventually shrinks to the endowment as the economy is replicated.

The above analysis leaves open important questions about participation in public goods allocations. Perhaps it is possible to characterize those economies for which optimality is not inconsistent with equilibrium participation. If this class of such economies is reasonable to assume as the set of possible economies, then the negative results may be avoided with small numbers of agents. Similarly, there may exist a wide range of economies for which Pareto optimality may be well approximated under equilibrium participation. If such 'approximately desirable' outcomes could be identified, perhaps there exists a more natural mechanism that can implement these outcomes in Nash equilibrium. Given that the equilibrium participation constraint can be thought of as a restriction on the size of transfers, it is conceivable that a total transfer maximizing solution to this system of restrictions may be identified and used to maximize the total size of the public good in a given economy.

Finally, empirical observation demonstrates that non-trivial quantities of public goods are regularly provided in large economies. Governments and other voluntarily established methods of coercion exist as enforcement devices to guarantee that welfare improving allocations are attained. This suggests a natural next step; the study of the endogenous selection of enforcement systems for the provision of public goods.

A Appendix

Proof of Proposition 7. Pick any agent i such that $t_i, T_{-i} \geq 0$ and $y \geq F(T_{-i})$Equilibrium participation implies that

$$u_i(\omega_i - t_i, F(T_{-i} + t_i)) \geq u_i(\omega_i, F(T_{-i})).$$

By quasi-concavity of u_i,

$$\nabla u_i(\omega_i, F(T_{-i})) \cdot (-t_i, F(T_{-i} + t_i) - F(T_{-i})) \geq 0,$$

or

$$\frac{F(T_{-i} + t_i) - F(T_{-i})}{t_i} \geq \frac{\partial u_i(\omega_i, F(T_{-i}))/\partial x_i}{\partial u_i(\omega_i, F(T_{-i}))/\partial y}.$$

Thus, by concavity of F,

$$\frac{\partial u_i(\omega_i, F(T_{-i}))/\partial x_i}{\partial u_i(\omega_i, F(T_{-i}))/\partial y} \leq F'(T_{-i}).$$

Inverting this inequality gives the necessary condition. □

Proof of Proposition 8. By monotonicity, equilibrium participation is trivially satisfied for all j such that $t_j < 0$ or $y < F(T_{-j})$. Equation (4) guarantees equilibrium participation when $T_{-j} < 0$. Now consider some $i \in \mathcal{I}$ such that $t_i, T_{-i} \geq 0$ and

$y \geq F(T_{-i})$, but for whom equilibrium participation fails. For this agent,

$$u_i(\omega_i, F(T_{-i})) > u_i(\omega_i - t_i, F(T_{-i} + t_i)), \tag{7}$$

so that

$$\nabla u_i(\boldsymbol{x}, y) \cdot (t_i, F(T_{-i}) - F(T_{-i} + t_i)) > 0.$$

This is equivalent to

$$\frac{\partial u_i(\boldsymbol{x}, y)/\partial x_i}{\partial u_i(\boldsymbol{x}, y)/\partial y} > \frac{F(T_{-i} + t_i) - F(T_{-i})}{t_i}, \tag{8}$$

so applying the concavity of F at $T_{-i} + t_i$ and inverting the resulting relationship gives

$$\frac{\partial u_i(\boldsymbol{x}, y)/\partial y}{\partial u_i(\boldsymbol{x}, y)/\partial x_i} < \frac{1}{F'(T_{-i} + t_i)}.$$

Equation (3) implies that (7) cannot hold, so by the contrapositive of this argument, $(\boldsymbol{x}, y)$ must satisfy EP_i. □

Proof of Proposition 9. For agents with $T_{-i} < 0$, $y^{(-i)} = 0$, but $F(T_{-i}) < 0$. By replacing $F(T_{-i})$ with zero in the proof of Proposition 8, the argument is identical through equation (8). At this point, the subsequent relationship with $F'(T)$ cannot be derived from $F(T)/t_i$ when $T_{-i} < 0$, so inverting (8) gives the alternative sufficient condition

$$\frac{\partial u_i(\boldsymbol{x}, y)/\partial y}{\partial u_i(\boldsymbol{x}, y)/\partial x_i} \geq \frac{1}{F(T)/t_i} \tag{9}$$

for all i such that $T_{-i} < 0$. Since this is a stronger condition than (3), it is also sufficient *every* agent. □

Proof of Proposition 12. By way of contradiction, assume that there exists some economy $\boldsymbol{e}$ and some sequence $\{(\boldsymbol{x}^R, \hat{y}^R)\}_{R=1}^{\infty}$ in $\mathcal{EP}(\boldsymbol{e}^R)$ for each R such that $|\hat{y}^R|$ fails to converge to zero. For each (i, r), let $t_{i,r}^R = \omega_{i,r}^R - x_{i,r}^R$. For any $(\boldsymbol{x}^R, \hat{y}^R) \in \mathcal{EP}(\boldsymbol{e}^R)$, if $\hat{y}^R < F(\sum_{i,r} t_{i,r}^R)$, then by monotonicity, $(\boldsymbol{x}^R, y^R) \in \mathcal{EP}(\boldsymbol{e}^R)$, where $y^R = F(\sum_{i,r} t_{i,r}^R)$. In other words, if a wasteful allocation $(\boldsymbol{x}, \hat{y})$ satisfies equilibrium participation, so does the transfer-equivalent non-wasteful allocation $(\boldsymbol{x}, y)$. (This is trivially true if $t_i \leq 0$; if $t_i > 0$ and $T_{-i} < 0$ then $\hat{y}^{(-i)} = y^{(-i)} = 0$ and it is true; if $t_i > 0$, $T_{-i} \geq 0$, and $\hat{y} < F(T_{-i})$ then it is vacuously true since $(\boldsymbol{x}, \hat{y})$ would not satisfy equilibrium participation; and the case of $t_i \geq 0$, $T_{-i} \geq 0$, and $\hat{y} \geq F(T_{-i})$ is true since $\hat{y}^{(-i)} = y^{(-i)} = F(T_{-i})$.) Thus, the sequence $\{(x^R, y^R)\}_{R=1}^{\infty}$ satisfies equilibrium participation for each R and $\{|y^R|\}_{R=1}^{\infty}$ also fails to converge to zero. This implies that there exists some $\varepsilon > 0$ and an infinite subsequence $\{(x^{R_k}, y^{R_k})\}_{k=1}^{\infty}$ such that for all $k = 1, 2, \ldots$ we have $\left|y^{R_k}\right| \geq \varepsilon$.

Let $\bar{y} = \lim_{k\to\infty} y^{R_k}$ if it exists. If it does not, there must exist a convergent subsequence since $\left\{y^{R_k}\right\}_k$ resides in the compact set $[\varepsilon, F(\sum_i \omega_i)]$, and so we can redefine $\{(x^{R_k}, y^{R_k})\}_{k=1}^{\infty}$ to be this convergent subsequence and let $\bar{y} = \lim_{k\to\infty} y^{R_k}$.

Letting $c(y)$ represent the minimal cost of producing y (which is the inverse of F), non-convergence guarantees that $c(y^{R_k}) \geq c(\varepsilon) > 0$ for each k since c is an increasing function and $\mathcal{Y} \cap \mathbb{R}_+^2 = \{0\}$. For any k, if $c(\varepsilon) > (\max_{i\in\mathcal{I}} \omega_i)/R_k$, then no one agent can unilaterally fund $y^{R(k)}$ using $t_{i,r}^{R(k)}$. Let

$$k^* = \max\{k \in \mathbb{N} : R(k) \leq \frac{1}{c(\varepsilon)} \max_{i\in\mathcal{I}} \omega_i\},$$

and consider any sequence of agents $\{(i_k, r_k)\}_{k=1}^{\infty}$ such that, for each k, $(i_k, r_k) \in \arg\max_{(i,r)} t_{i,r}^{R_k}$. Each (i_k, r_k) in this sequence contributes at least as much as the average contribution, so

$$\begin{aligned} t_{i_k,r_k}^{R_k} &\geq \frac{c\left(y^{R_k}\right)}{R_k\, I} \\ &> \frac{c(\varepsilon)}{R_k\, I} \\ &> 0 \end{aligned}$$

for each k and $T_{-(i_k,r_k)} > 0$ for all $k > k^*$. Since each (x^{R_k}, y^{R_k}) satisfies equilibrium participation for all (i, r), it must be the case that, for every k,

$$(\omega_{i,r} - t_{i_k,r_k}^{R_k}, y^{R_k}) \succeq_{i_k,r_k} (\omega_{i,r}, (y^{R_k})^{-(i_k,r_k)}),$$

or

$$\left(\omega_i - R_k\, t_{i_k,r_k}^{R_k}, y^{R_k}\right) \succeq_{i_k} \left(\omega_i, \left(y^{R_k}\right)^{-(i_k,r_k)}\right). \tag{10}$$

Note that for $k > k^*$,

$$(y^{R_k})^{-(i_k,r_k)} = F\left(\sum_{(j,s)\neq(i_k,r_k)} t_{j,s}^{R_k}\right).$$

By continuity of the production function, $(y^{R(k)})^{-(i_k,r_k)}$ becomes arbitrarily close to $y^{R(k)}$ as k grows, so that $\lim_{k\to\infty}\left(y^{R_k}\right)^{-(i_k,r_k)} = \lim_{k\to\infty} y^{R_k} = \bar{y}$.

Since $t^{R_k}_{i_k,r_k} > c(\varepsilon)/(R_k I)$ for each k, monotonicity of preferences and equation (10) imply that

$$\left(\omega_i - \frac{c(\varepsilon)}{I}, y^{R_k}\right) \succeq_{i_k} \left(\omega_i, \left(y^{R_k}\right)^{-(i_k,r_k)}\right). \tag{11}$$

Strict monotonicity also implies that, for each k,

$$\left(\omega_i - \frac{c(\varepsilon)}{I}, \bar{y}\right) \prec_{i_k} (\omega_i, \bar{y}),$$

and so, by continuity of preferences, there exists a $\delta_{i_k} > 0$ for each k such that

$$\left(\omega_i - \frac{c(\varepsilon)}{I}, \bar{y} + \delta_{i_k}\right) \prec_{i_k} (\omega_i, \bar{y} - \delta_{i_k}).$$

Lett $\delta^* = \min_k \{\delta_{i_k}\}$, which is well-defined and positive since there are only a finite number of types i_k, we have

$$\left(\omega_i - \frac{c(\varepsilon)}{I}, \bar{y} + \delta^*\right) \prec_{i_k} (\omega_i, \bar{y} - \delta^*).$$

But since $y^{R_k} \to \bar{y}$ and $\left(y^{R_k}\right)^{-(i_k,r_k)} \to \bar{y}$, there is some k such that $y^{R_k} \leq \bar{y} + \delta^*$ and $\left(y^{R_k}\right)^{-(i_k,r_k)} \geq \bar{y} - \delta^*$, so that, by monotonicity,

$$\begin{aligned}\left(\omega_i - \frac{c(\varepsilon)}{I}, y^{R_k}\right) &\preceq_{i_k} \left(\omega_i - \frac{c(\varepsilon)}{I}, \bar{y} + \delta^*\right) \\ &\prec_{i_k} (\omega_i, \bar{y} - \delta^*) \preceq_{i_k} \left(\omega_i, \left(y^{R_k}\right)^{-(i_k,r_k)}\right).\end{aligned}$$

But this contradicts equation (11).

Since there cannot be an infinite subsequence of allocations with $|y^{R_k}| > \varepsilon$ for any $\varepsilon > 0$, it must be the case that $y^R \to 0$ as $R \to \infty$. Feasibility then requires that $\|x^R - \omega^R\|_\infty \to 0$, completing the proof. $\square$

References

BERGSTROM, T., L. BLUME, AND H. VARIAN (1986): "On the private provision of public goods," 29, 25–49.

CHAMPSAUR, P., D. J. ROBERTS, AND R. W. ROSENTHAL (1975): "On cores of economies with public goods," *International Economic Review*, 16, 751–764.

FOLEY, D. K. (1970): "Lindahl's solution and the core of an economy with public goods," *Econometrica*, 38, 66–72.

FURUSAWA, T., AND H. KONISHI (2007): "A Theory of Endogenous Lobby Formation," Boston College working paper.

GREEN, J. R., AND J.-J. LAFFONT (1979): *Incentives in public decision-making*. North-Holland, Amsterdam.

GROVES, T. (1973): "Incentives in Teams," *Econometrica*, 41, 617–633.

GROVES, T., AND J. O. LEDYARD (1977): "Optimal Allocation of Public Goods: A Solution to the 'Free-Rider' Problem," *Econometrica*, 45, 783–809.

HURWICZ, L. (1972): "On Informationally Decentralized Systems," in *Decision and Organization: A Volume in Honor of Jacob Marschak*, ed. by C. B. McGuire, and R. Radner. North-Holland, Amsterdam.

——— (1979): "On allocations attainable through Nash equilibria," *Journal of Economic Theory*, 21, 140–165.

JACKSON, M. O., AND T. R. PALFREY (1998): "Efficiency and voluntary implementation in markets with repeated pairwise bargaining," *Econometrica*, 66, 1353–1388.

——— (2001): "Voluntary implementation," *Journal of Economic Theory*, 98, 1–25.

KOLM, S.-C. (1970): *La Valeur Publique*. Dunod, Paris.

LEDYARD, J. O., AND J. ROBERTS (1975): "On the incentive problem with public goods," Northwestern University Center for Mathematical Studies in Economics and Management Science, Discussion Paper No. 116.

Ma, C.-T., J. Moore, and S. Turnbull (1988): "Stopping agents from cheating," *Journal of Economic Theory*, 46, 355–372.

Maskin, E. (1999): "Nash equilibrium and welfare optimality," *Review of Economic Studies*, 66, 23–38.

Maskin, E., and J. Moore (1999): "Implementation and renegotiation," *Review of Economic Studies*, 66, 39–56.

Milleron, J.-C. (1972): "Theory of value with public goods: A survey article," *Journal of Economic Theory*, 5, 419–477.

Muench, T. J. (1972): "The core and the Lindahl equilibrium of an economy with a public good: An example," *Journal of Economic Theory*, 4, 241–255.

Perez-Nievas, M. (2002): "Interim efficient allocation mechanisms," Ph.D. thesis, Universidad Carlos III de Madrid.

Richter, D. K. (1974): "The core of a public goods economy," *Int. Econ. Rev.*, 15, 131–142.

Saijo, T. (1991): "Incentive compatibility and individual rationality in public good economies," *Journal of Economic Theory*, 55, 203–212.

Saijo, T., and T. Yamato (1999): "A voluntary participation game with a non-excludable public good," *Journal of Economic Theory*, 84, 227–242.

——— (2008): "Voluntary Participation in the Design of Non-excludable Public Goods Provision Mechanisms," Mimeo.

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

Thomson, W. (1999): "Economies with public goods: An elementary geometric exposition," *Journal of Public Economic Theory*, 1, 139–176.

Walker, M. (1981): "A Simple Incentive Compatible Scheme for Attaining Lindahl Allocations," *Econometrica*, 49, 65–71.