



## Provision of Public Goods: Fully Implementing the Core through Private Contributions

Mark Bagnoli; Barton L. Lipman

*The Review of Economic Studies*, Vol. 56, No. 4 (Oct., 1989), 583-601.

Stable URL:

<http://links.jstor.org/sici?sici=0034-6527%28198910%2956%3A4%3C583%3APOPFGFI%3E2.0.CO%3B2-P>

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

*The Review of Economic Studies* is published by The Review of Economic Studies Ltd.. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/resl.html>.

---

*The Review of Economic Studies*

©1989 The Review of Economic Studies Ltd.

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact [jstor-info@umich.edu](mailto:jstor-info@umich.edu).

©2003 JSTOR

# Provision of Public Goods: Fully Implementing the Core through Private Contributions

MARK BAGNOLI  
*University of Michigan*

and

BARTON L. LIPMAN  
*Carnegie-Mellon University*

First version received June 1987; final version accepted May 1989 (Eds.)

Standard economic intuition would say that private provision of public goods will be inefficient due to free-rider problems. This view is in contrast to the results in the literature on full implementation where it is shown that (under certain conditions) games exist which only have efficient equilibria. The games usually used to demonstrate existence are quite complex and seem “unnatural”, possibly leading to the perception that implementation requires a central authority to choose and impose the game. In a simple public goods setting, we show that a very natural game—similar to one often used elsewhere in the literature to model private provision—in fact fully implements the core of this economy in undominated perfect equilibria. More specifically, we consider a complete information economy with one private good and two possible social decisions. Agents voluntarily contribute any non-negative amount of the private good they choose and the social decision is to provide the public good iff contributions are sufficient to pay for it. The contributions are refunded otherwise. The set of undominated perfect equilibrium outcomes of this game is exactly the core of the economy. We give some extensions of this result, discuss the role of perfection and alternative equilibrium notions, and discuss the intuition and implications of the results.

## I. INTRODUCTION

The standard economic intuition regarding private provision of public goods has been largely verified by a series of recent papers on the subject.<sup>1</sup> In some of these papers, private provision is modeled as direct purchase—agents unilaterally put up streetlights, give money to the poor, or otherwise pay for some of the public good. As Bergstrom, Blume, and Varian (1986) observe, this is equivalent to having a “collector” whose sole function is to use the money contributed to purchase the largest possible quantity of the public good. Examples of such collectors might include the United Way, political action committees, and other nonprofit charitable organizations. In other papers, the collector takes a more active role in provision, such as refunding some contributions in certain circumstances. For instance, Palfrey and Rosenthal (1984) give the example of a person designated to collect money for an office coffee club, where contributions are refunded if insufficient. In all of these models, the equilibrium outcome typically has an inefficiently low level of the public good, confirming the standard view of the free-rider problem.

1. See, for example, Warr (1983), Palfrey and Rosenthal (1984, 1988), Bergstrom, Blume, and Varian (1986), and Andreoni (1988). Related papers include Bliss and Nalebuff (1985), Cornes and Sandler (1985), and Bernheim (1986).

By contrast, the literature on public provision has shown that a central authority with the ability to choose and administer the game can achieve efficient outcomes. In fact, the literature on full implementation has demonstrated the existence of games for which all equilibria are efficient.<sup>2</sup> Most of the games analyzed in the implementation literature are constructed for the purpose of proving existence in some very general setting, rather than for their plausibility. For this reason, these mechanisms are often quite complex and appear to require a central authority for their use.<sup>3</sup> For example, the way these games are usually described is that the agents send messages to a social planner or mediator who chooses the social decision based on the messages. In addition, these games often use seemingly artificial devices to guarantee efficiency. One such device is to have the agents name integers and, if their messages contradict one another in particular ways, the agent picking the largest integer is rewarded.

We show that these parts of the literature have more in common than has been recognized by providing an example of an efficient mechanism which is quite similar to games used to model private provision. The simplest environment we consider is a complete information economy with a single private good and a discrete public good. The mechanism allows each agent to voluntarily contribute any non-negative amount of the private good he chooses. The public good is provided iff contributions are sufficient to pay for it and the contributions are refunded otherwise. Surprisingly, this game is an efficient mechanism in the strong sense that *every* equilibrium outcome is efficient. In fact, we will show that the set of equilibrium outcomes is exactly equal to the core of the economy. In our game, the collector's only role is to purchase the largest quantity of the public good that the contributions will cover and to refund if necessary. Hence it seems quite reasonable to interpret this game as a model of private provision, a view we will emphasize. Under this interpretation, our results imply that private provision of public goods can be efficient.

An alternative interpretation of our game is that it is a simple mechanism for public provision of public goods, one much simpler than the mechanisms studied in the conventional implementation literature. Consequently, our results suggest that the boundary between mechanisms for private provision and mechanisms for public provision is not clear. The distinction between these classes of mechanisms is surely not the existence of a central figure like a collector. We would hardly view a collector like United Way as a "central authority". On the other hand, a mechanism which gives the collector the ability to alter the allocation of goods in any fashion must surely be thought of as a mechanism for public provision only.

We also consider the case of a public good that can take on finitely many values. The most obvious extension of our game would have one round of contributions where the amount of the public good provided is the maximum that the contributions will pay for. Unfortunately, as we discuss below, this game does not guarantee efficiency. Instead, we consider a sequential game with potentially many rounds of contributions. In addition, we will see that a more complex refinement is needed to guarantee efficiency. Hence, our results suggest that achieving efficiency through voluntary contributions is not impossible in this setting, but is more difficult.

Since our sequential game works for any finite number of values for the public good, we can reinterpret this case as approximating the case where the public good can take on any value in a continuum. Interestingly, as the increment between levels of the public

2. See Maskin (1977), Moore and Repullo (1988), and Palfrey and Srivastava (1987).

3. There are some exceptions to this. For instance, More and Repullo give several examples of natural-looking games with efficient equilibria for various settings.

good goes to zero, the game form converges to a repeated version of the game considered by Bergstrom, Blume, and Varian. While the limiting outcomes are efficient, the outcomes at the limit are not. This occurs because efficiency requires that the sum of the agents' marginal rates of substitution (MRS) equal marginal cost. Of course, this implies that any given agent's MRS is strictly less than marginal cost and so if a \$1 decrease in his payments reduces the level of the public good by exactly \$1 worth, he is better off. To avoid this problem, it must be true that a small decrease in one's contribution leads to a relatively large decrease in the amount of the public good. As we discuss, our game has this property for any strictly positive increment between values of the public good, but does not when the increment is zero. This provides an intriguing view of the role of discontinuities in the game form, a point noted in a different context by Aghion (1985).

One caveat that must be mentioned is that we assume complete information throughout. This is, of course, a very strong assumption and may limit the applicability of our results. Intuitively, incomplete information may lead to underprovision as agents trade off their contribution against the probability that the public good is provided. Much has been written regarding the problems incomplete information poses for public provision (see, for example, Mailath and Postlewaite (1988)) and on optimal mechanisms in such environments (e.g. d'Aspremont and Gérard-Varet (1979) or Laffont and Maskin (1982)). On the other hand, our results for the simplest case are not affected much by the inclusion of small amounts of incomplete information. In this case, all the equilibria with outcomes in the core are strong Nash equilibria. As van Damme (1983) has shown, every such equilibrium is "close" to an equilibrium with a small amount of incomplete information.

In the next section, we set out the model and give definitions. In Section III, we consider the simplest case where the set of social decision is  $\{0, 1\}$ , which we interpret as choosing whether or not to build a streetlight. We show that the simple game described above fully implements the core. In Section IV, we extend these results to the case where the set of social decisions is  $\{0, 1, \dots, M\}$  for some finite  $M$ . We interpret this case as the choice of the number of streetlights to build. This analysis extends trivially to the case where the decision set is  $\{0, \delta, 2\delta, \dots, M(\delta)\}$  where  $\delta > 0$  and  $M(\delta)$  is the largest multiple of  $\delta$  less than or equal to  $M$ . Viewing this decision set as an approximation of  $[0, M]$ , this fact gives us a way to approximately fully implement the core. In Section V we offer some concluding remarks. All proofs are in the Appendix.

## II. THE ECONOMY AND DEFINITIONS

We consider an economy with  $I$  agents indexed by  $i \in \mathcal{I} = \{1, \dots, I\}$ . There is one private good, which we will refer to as wealth. Agent  $i$ 's endowment of wealth is denoted  $w_i$  and the vector of endowments  $w$ , where we assume  $w \in \mathbf{R}_{++}^I$ . The agents must choose a decision  $d$  from the set  $D \subset \mathbf{R}_+$ , where we assume  $0 \in D$ . For the moment, we will not impose extra structure on  $D$ . A state of the economy, which we will denote  $\omega$ , specifies each agent's utility function. The set of possible states will be denoted  $\Omega$ . We will write the utility function for the  $i$ -th agent in state  $\omega$  as  $u_i(d, w_i | \omega)$ . We assume that  $u_i$  is strictly increasing in  $d$  for all  $i$  and all  $\omega \in \Omega$ . We also assume that  $u_i$  is continuous and strictly increasing in  $w_i$  for all  $i$  and  $\omega$ . We will impose some further conditions on  $\Omega$  in the subsequent sections.

For convenience, we define the cost function for the social decisions as a function  $c: \mathbf{R}_+ \rightarrow \mathbf{R}_+$ . Of course,  $c$  is only relevant on  $D$ . We assume that the cost function is strictly increasing and convex and that  $c(0) = 0$ . Finally, we assume complete information

so that all of the above (including the state) is common knowledge among the  $I$  agents at each state.

We will refer to a social decision and an allocation of the private good among the agents as an outcome. More precisely, an outcome is a point,  $\theta$ , in  $D \times \mathbf{R}_+^I$  and the set of feasible outcomes is

$$\Theta = \{(d, x) \in D \times \mathbf{R}_+^I \mid \sum_i x_i \leq \sum_i w_i - c(d)\}$$

The core of this economy is a mapping  $C: \Omega \rightarrow P(\Theta)$  where  $P(A)$  is the power set of  $A$ . To define the core, we must first define what a coalition can achieve. We will write  $\Theta_T$  as the set of feasible outcomes for coalition  $T \subseteq \mathcal{I}$ . That is,

$$\Theta_T = \{(d, x) \in D \times \mathbf{R}_+^I \mid \sum_{i \in T} x_i \leq \sum_{i \in T} w_i - c(d)\}$$

In the usual terminology, we will say that the coalition  $T \in P(\mathcal{I})$  can block the outcome  $\theta \in \Theta$  in state  $\omega$  if there exists some  $\theta' = (d', x') \in \Theta_T$  such that

$$u_i(d', x'_i \mid \omega) \geq u_i(d, x_i \mid \omega)$$

for all  $i \in T$  with a strict inequality for some  $i \in T$ . An outcome  $\theta$  is in  $C(\omega)$  iff there is no coalition that can block it in state  $\omega$ .

A game form,  $G$ , is a pair  $(S, O)$  where  $S = S_1 \times \cdots \times S_I$  and  $O: S \rightarrow \Theta$ . A game form together with a state define a game in normal form where the payoffs associated with the strategy combination  $\sigma \in S$  are given by  $u(O(\sigma) \mid \omega)$ . We will say that the normal form game  $\Gamma(\omega) = (S, u(O \mid \omega))$  is induced by  $G$  in state  $\omega$ . Thus, for a given  $G$ , the set of equilibrium strategies must be defined as a correspondence from  $\Omega$  into  $S$ . (We will be more explicit about the definition of this correspondence below.) For our purposes, a more useful correspondence is the set of equilibrium outcomes. For a game form  $G$ , let  $E_G^*(\omega)$  be the set of equilibrium strategy tuples in  $S$ . Then the set of equilibrium outcomes under  $G$ ,  $E_G$ , is defined by  $E_G(\omega) = O(E_G^*(\omega))$ . We say that  $G$  fully implements the core iff  $E_G(\omega) = C(\omega)$  for all  $\omega \in \Omega$ . That is, a game fully implements the core if the set of equilibrium outcomes exactly coincides with the core.

Most work in implementation theory uses the notion of Nash equilibrium to define the set  $E_G(\omega)$ . One important part of our analysis is that we work with refinements of Nash equilibrium. To make as clear as possible the role that the exact choice of refinement plays, we will discuss the outcomes under various equilibrium notions, the definitions of which are given below. To be as precise as possible in our statements about implementation, when the correspondence  $E_G$  has been defined using (for example) perfect equilibrium and  $E_G(\omega) = C(\omega)$  for all  $\omega$ , we will say that  $G$  fully implements the core in perfect equilibrium.

We work with two basic concepts: elimination of dominated strategies and perfect equilibria. This may seem redundant since a dominated strategy cannot be played in a perfect equilibrium. However, a perfect equilibrium can be supported by trembles to dominated strategies so that performing this elimination explicitly before applying perfection does affect the set of equilibria.<sup>4</sup> Our theorems use two different ways of combining these concepts. Theorem 1 uses what we will call undominated perfect equilibrium or UPE for brevity. This equilibrium concept eliminates dominated strategies and applies the notion of (trembling-hand) perfection to the resulting game. This result, as we discuss below, also holds for many stronger equilibrium concepts, including the one used in Theorem 2. There we use successive elimination of dominated strategies and then apply

4. See the classic example in Myerson (1978).

strict perfection to the resulting game. We will refer to this as successively undominated strictly perfect equilibrium or SUSPE.<sup>5</sup>

To define our equilibrium notions completely rigorously requires a great deal of cumbersome notation. To avoid this, we will sketch the ideas.<sup>6</sup> Our purpose is to define, for an arbitrary game form, a correspondence from  $\Omega$  to  $S$  giving the equilibrium strategy tuples. So for a given game form and state, consider the induced game, say  $\Gamma = (S, u)$ . In Section III, we will consider a simultaneous move game and so will work with the normal form. In Section IV, we will have a multistage game. To define perfection for such games, it is customary to work with the agent-normal form. That is, one replaces the original set of players with the "agents" who represent the players, where each player has a different agent working for him at each information set. So when the definitions below are applied to this game,  $\Gamma$  is taken to be the agent-normal form.

The set of actions available to each player (or each agent as the case may be) will be an interval of the real line and thus uncountable. Since perfection and its variations are generally defined for games with finite strategy sets, we must necessarily either consider a sequence of approximating finite games or accept some technical complexities in defining completely mixed strategies. We adopt the former approach.<sup>7</sup> We approximate the game by replacing the uncountable strategy sets with finite ones. The way we choose these finite sets is by picking a grid size,  $\mu_n$ , and for each  $i$ , a smallest element,  $a_i(n)$ , within  $\mu_n$  of the smallest element of the uncountable strategy set,  $S_i$ . The finite set,  $S_i(n)$ , will be  $a_i(n)$  plus any point of the form  $a_i(n) + k\mu_n$  which is less than the largest element of  $S_i$ . We require that  $\mu_n \downarrow 0$  as  $n \rightarrow \infty$ . Call the approximating game  $\Gamma(n)$ . For each equilibrium notion we use, we will say  $\sigma$  is an equilibrium of  $\Gamma$  if it is the limit of a sequence  $\{\sigma(n)\}$  of equilibrium points of a sequence of approximating games  $\{\Gamma(n)\}$ .

So now we only need to define our equilibrium notions for normal form games with finite strategy sets. We will say that a strategy for  $i$ ,  $\sigma_i \in S_i$ , is dominated if there exists  $\sigma'_i \in S_i$  such that for all  $\sigma_{-i} \in S_{-i}$ ,

$$u_i(\sigma'_i, \sigma_{-i}) \geq u_i(\sigma_i, \sigma_{-i})$$

with a strict inequality for some  $\sigma_{-i} \in S_{-i}$ . (As usual,  $\sigma_{-i}$  refers to a vector of strategies for the players other than  $i$ .) Let  $R^1(S_i)$  denote the set of strategies for  $i$  which are not dominated and let  $R^1(S) = R^1(S_1) \times \cdots \times R^1(S_1)$ . Then let  $R^1(\Gamma)$  denote the game  $(R^1(S), u)$ . An undominated perfect equilibrium of  $\Gamma$  is a perfect equilibrium of  $R^1(\Gamma)$ . Similarly, when we refer to an equilibrium notion such as undominated proper equilibrium or undominated strictly perfect equilibrium, we will mean the proper or strictly perfect equilibria of  $R^1(\Gamma)$ . For  $n \geq 2$ , recursively define  $R^n(S_i)$  as the set of strategies for  $i$  which are not dominated in  $R^{n-1}(\Gamma)$  where  $R^n(\Gamma)$  is defined analogously to  $R^1(\Gamma)$ . Finally, define  $R^*(S_i)$  as the set of strategies in  $R^n(S_i)$  for all  $n$  and  $R^*(\Gamma)$  analogously to the above. (Notice that  $R^*(S_i)$  must be nonempty for all  $i$ .) A successively undominated perfect equilibrium of  $\Gamma$  is a perfect equilibrium of  $R^*(\Gamma)$  and similarly for a successively

5. It may be useful to provide some relationships among these concepts and other familiar ones. The UPE's are a subset of the perfect equilibria. In turn, the proper equilibria are a subset of the UPE's. The SUSPE's are also a subset of the UPE's and the strictly perfect equilibria are a subset of the SUSPE's. While one may suspect that the SUSPE's are a subset of the proper equilibria, this is not true. In fact, one can show that the proper equilibria are a subset of the SUSPE's in the game we consider in Section III for the case where  $\sum_i v_i = c$ .

6. The reader interested in more detail is referred to our 1987 working paper.

7. For an example of the latter approach, see Chatterjee and Samuelson (1986). Their analysis is quite related to ours in that they consider essentially a two-agent economy where the decision set is  $\{0, 1\}$  and a provision date.

undominated proper equilibrium or a successively undominated strictly perfect equilibrium.

We are now ready to define the perfect and strictly perfect equilibria of an arbitrary game with a finite strategy set. So, again, consider an arbitrary game  $\Gamma = (S, u)$  where each  $S_i$  is finite. A completely mixed strategy for  $i$  is a mixed strategy which places strictly positive probability on each of  $i$ 's pure strategies. A perfect equilibrium is a vector  $\sigma \in S$  such that  $\sigma_i$  is best reply to some vector of completely mixed strategies for the other players which is close to  $\sigma_{-i}$ . By close to  $\sigma_{-i}$ , we mean that  $\sigma_i$  is a best reply to each point in a sequence of completely mixed strategies converging to  $\sigma_{-i}$ .

Strict perfection requires much more—that  $\sigma_i$  is a best response to *every* vector of completely mixed strategies for the other players close to  $\sigma_{-i}$ . More precisely,  $\sigma_i$  is required to be a best response to each point in any sequence of completely mixed strategies converging to  $\sigma_{-i}$ . Thus perfection requires robustness with respect to *some* small probabilities of mistakes by the other players, while strict perfection requires robustness with respect to *all* small probabilities of mistakes by the other players. It is important to note that not every normal form game possesses a strictly perfect equilibrium. In fact, the games we consider, while they do possess undominated strictly perfect equilibria or successively undominated strictly perfect equilibria, do not in fact possess strictly perfect equilibria.

To summarize, then, let  $\Gamma$  be any normal form game with a finite strategy set or, equivalently, the agent-normal form of an extensive form game where the action set for each agent is finite. We will say that  $\sigma$  is an undominated perfect equilibrium (UPE) of  $\Gamma$  iff it is a perfect equilibrium of  $R^1(\Gamma)$ . Similarly,  $\sigma$  is a successively undominated strictly perfect equilibrium (SUSPE) of  $\Gamma$  iff it is a strictly perfect equilibrium of  $R^*(\Gamma)$ . If  $\Gamma$  has an uncountable strategy set, we say that  $\sigma$  is a UPE (SUSPE) of  $\Gamma$  if there is a sequence  $\{\sigma(n)\}$  converging to  $\sigma$  such that  $\sigma(n)$  is a UPE (SUSPE) of  $\Gamma(n)$ .

Finally, we will say that a game form  $G = (S, O)$  fully implements the core in UPE (SUSPE) if  $E_G(\omega) = C(\omega)$  for all  $\omega \in \Omega$  where  $E_G^*(\omega)$  is the set of  $\sigma \in S$  such that  $\sigma$  is a UPE (SUSPE) of the game induced by  $G$  at state  $\omega$ .

### III. THE ONE-STREETLIGHT PROBLEM.

In this section, we simplify the above structure to the case where  $D = \{0, 1\}$ . We will interpret  $d = 1$  as the decision to provide a streetlight. For simplicity, we let  $c = c(1)$ . Without loss of generality, we adopt the normalization  $u_i(0, w_i | \omega) = 0$  for each  $i$  and each  $\omega \in \Omega$ . The valuation of agent  $i$  in state  $\omega$ ,  $v_i(\omega)$ , is defined implicitly by  $u_i(1, w_i - v_i | \omega) = 0$ . Since the valuations define everything about  $u_i$  that is relevant for our purposes at a state  $\omega$ , we will often omit the  $\omega$  argument and focus on the valuations directly. We also assume that for each state,  $w_i > v_i(\omega)$  for all  $i$ —that is,  $u_i(1, 0 | \omega) < 0$  for all  $\omega \in \Omega$ . This assumption is made so that we do not need to consider what happens when some agents would like to contribute more than their wealth. To guarantee that the problem is interesting, we also assume that  $\sum_i w_i > c$ . We will refer to this class of economies as  $\mathcal{E}^1$ .

Characterizing the core of an economy in  $\mathcal{E}^1$  is quite straightforward. If  $\sum_i v_i(\omega) < c$ , then the only point in the core is  $(0, w)$ . This is true because any other distribution of wealth along with  $d = 0$  clearly cannot be both feasible and in the core. Similarly, any distribution of wealth with  $d = 1$  leaves some agent worse off than if he refused to participate and hence cannot be in the core. If  $\sum_i v_i(\omega) = c$ , the core consists of the points  $(0, w)$  and  $(1, w - v)$ . Again, it is clear that any other distribution of wealth could be blocked by some coalition. Finally, if  $\sum_i v_i(\omega) > c$ , then any outcome in the core must

have  $d = 1$  as this condition is necessary for Pareto optimality. Clearly, the wealth distribution at a core outcome must have each individual with at least  $w_i - v_i$  or else some individual agent would block. If any agent receives more than  $w_i$ , then the coalition of all agents other than this one can block as their loss of wealth is in part received by this individual. Thus the core is certainly no larger than the set of  $(1, x)$  such that  $\sum_i x_i = \sum_i w_i - c$  and  $w_i - v_i \leq x_i \leq w_i$  for all  $i$ . In fact, it is easy to see that no coalition can block an outcome in this set and so this set is precisely the core.<sup>8</sup>

A very natural way to consider the problem of how the agents get together to jointly provide the good is to suppose that they contribute money toward the building of the streetlight. If the contributions add to  $c$  or more, the streetlight is provided, with any excess contributions either kept by the collector or refunded in some fashion. For notational simplicity, we assume that if contributions exceed  $c$ , the collector keeps the excess. However, as we discuss below, our results also hold when he refunds this money, so a requirement of budget-balance can easily be satisfied. A variety of assumptions could be made about what happens when not enough money is contributed, not all of which would lead to efficient outcomes. For example, we might assume that contributions are not refunded regardless of the total.<sup>9</sup> However, with such a structure the possibility of insufficient contributions may deter agents from contributing. An obvious way to avoid this problem is to assume that if contributions add to less than  $c$ , all contributions are refunded. This contribution game is a simple generalization of one used by Palfrey and Rosenthal (1984) to model private provision and similar games have been used in experimental work<sup>10</sup> for the same purpose.<sup>11</sup>

More formally, let the strategy set of agent  $i$  be  $S_i^1 = [0, w_i]$ . A strategy choice by  $i$  will be denoted  $\sigma_i$  and will be referred to as a contribution. Define  $O^1(\sigma)$  by

$$O^1(\sigma) = \begin{cases} (0, w), & \text{if } \sum_i \sigma_i < c; \\ (1, w - \sigma), & \text{otherwise.} \end{cases}$$

We will refer to this particular game form as  $G^1$ .

As we will discuss in more detail below, there are many Nash equilibria of this game, some of which are not in the core. However, we have the following theorem.

**Theorem 1.**  $G^1$  fully implements the core of  $\mathcal{E}^1$  in undominated perfect equilibrium.

To understand this result, first consider the case where  $\sum_i v_i(\omega) < c$ . As noted, the only core outcome at such a state is  $(0, w)$  and it is easy to see that all Nash equilibria must have this outcome. Simply note that no one will contribute more than  $v_i$  if this will cause the streetlight to be built. Hence contributions cannot possibly add to  $c$  in equilibrium.

Now suppose  $\sum_i v_i > c$  and consider the set of (pure strategy) Nash equilibria of this game. It is easy to see that any vector of contributions  $\sigma$  such that  $0 \leq \sigma_i \leq v_i$  for all  $i$  and  $\sum_i \sigma_i = c$  must be a Nash equilibrium. Since  $\sigma_i \geq v_i$ , each agent's equilibrium utility

8. See Mas-Colell (1980) for a more detailed characterization of the core in a setting which has  $\mathcal{E}^1$  as a special case.

9. Palfrey and Rosenthal (1984) considered this possibility.

10. See, for example, Ferejohn, Forsythe, Noll and Palfrey (1982), Schneider and Pommerehne (1981) or Isaac, Walker and Thomas (1984).

11. It is also worth noting that this game is essentially a simplification of Nash's (1953) demand game. While our results had been known and in fact are straightforward to prove for the case of  $I = 2$ , the generalization is new. We know of no proof for the general Nash demand game. Our results do generalize to the original Nash demand game under certain simplifications.



is at least 0 so that no agent can increase his payoff by contributing less. Such a deviation will cause the decision to change and he will get utility of 0. Similarly, an increase in agent  $i$ 's contribution can only make him worse off because the streetlight will be provided at the lower contribution. Thus we see that there is a Nash equilibrium outcome for each point in the core.

However, there are other Nash equilibria. In particular, consider any  $\sigma$  such that  $\sigma_i \geq 0$  for all  $i$ ,  $\sum_i \sigma_i < c$ , and  $\sum_{j \neq i} \sigma_j + v_i \leq c$  for all  $i$ . In this equilibrium, the streetlight is not built, so each agent's utility is 0. This is an equilibrium because for each agent, any contribution that changes the decision exceeds  $v_i$ . Given that no contribution below  $v_i$  will cause the streetlight to be built, agent  $i$  is indifferent among all contributions which do not lead to the streetlight being built. These equilibrium outcomes are not in the core.

All of the equilibria which lead to core outcomes are strong Nash equilibria (see van Damme (1983)) and thus satisfy virtually all robustness requirements ever proposed in game theory. However, the other equilibria are not so robust. Many of them are not perfect, for example. In a perfect equilibrium, each agent's strategy must be robust to small probabilities of "mistakes." It is easy to see that no Nash equilibrium in which  $\sigma_i > v_i$  can be perfect. If the streetlight will not be provided, then contributing more than  $v_i$  is costless since this contribution will be refunded. Hence this can occur in a Nash equilibrium. On the other hand, if there is even a tiny probability that some other agent(s) will "accidentally" contribute enough so that the streetlight is built, then  $i$  strictly prefers contributing less than  $v_i$ .

There are equilibria which do not have outcomes in the core and which are perfect. To see this, suppose that  $c = 1$ ,  $I = 2$ , and each person's valuation is 0.6. Then  $\sigma_i = 0$  for both players is a perfect equilibrium. To construct trembles supporting this, suppose each person puts probability  $1 - \varepsilon$  on 0,  $k\varepsilon/(1+k)$  on 1, and the rest of the probability on the remaining strategies. Choose  $k$  to be very large. Then it is virtually certain that the other player's contribution is either 0 or 1. In either case,  $i$ 's best strategy is  $\sigma_i = 0$  and in the latter case, this is his unique best strategy. Hence any other strategy must yield a strictly lower expected payoff. Intuitively, certain trembles will not induce an agent to increase his contribution because he will end up contributing when the good would be provided without his contribution.

However, these trembles do not seem plausible. The equilibrium in the example required that the agents were most likely to tremble to contributing more than their valuations. Clearly, though, any smaller contribution strictly dominates this one. Hence if we eliminate dominated strategies even as trembles, then this possibility is eliminated. This is why we focus on undominated perfect equilibria. In the example above, this elimination implies that the other player necessarily contributes less than 0.6, so that in any UPE, each player must give more than 0.4 as any smaller contribution is weakly dominated. But this eliminates all the inefficient Nash equilibria.

Theorem 1 holds under a large variety of other equilibrium notions. The robustness of the equilibria with outcomes in the core when  $\sum_i v_i > c$  means that we can use any stronger equilibrium notion given that we first eliminate dominated strategies. It is also not hard to show that extra rounds of elimination of dominated strategies will not change the set of equilibria either. Thus, for example, Theorem 1 is trivially extended to undominated proper equilibria, undominated strictly perfect equilibria, SUPE, or SUSPE.<sup>12</sup>

12. In Bagnoli and Lipman (1986), we showed that all proper equilibrium outcomes are in the core. However, when  $\sum_i v_i = c$ , the only proper equilibrium outcome is  $(1, w - v)$ , so the other outcome in the core,  $(0, w)$ , cannot be obtained. Hence  $G^1$  only generally implements the core in proper equilibria.

Theorem 1 also holds for a large class of variations on  $G^1$  which satisfy budget-balance. To ensure budget-balance, when contributions exceed  $c$ , the excess must be refunded. If we amend  $G^1$  to incorporate any refund scheme with the property that increasing one's contribution by \$1 never increases one's refund by more than \$1, the resulting game fully implements the core in UPE. Intuitively, such a refund scheme guarantees that an agent always strictly prefers a contribution making the total exactly equal  $c$  to any larger contribution. This property implies that all the efficient equilibria are strong Nash equilibria and so are UPE's. Furthermore, since the refunding increases the payoff to a deviation from an inefficient Nash equilibrium, it certainly does not interfere with eliminating these equilibria.

#### IV. THE MULTIPLE STREETLIGHTS PROBLEM.

In this section, we consider a broader social decision set. Here we take  $D$  to be  $\{0, \dots, M\}$  for some finite  $M \geq 1$ —for example, how many streetlights to build. Since we will work with sequential mechanisms, eliminating wealth effects is notationally very convenient. Therefore, we will simplify our assumptions and assume that preferences are quasi-linear; that is, we assume that  $u_i(d, w_i | \omega) = U_i(d | \omega) + w_i$  for all  $i$  and all  $\omega \in \Omega$ . Analogously to the previous section, define the valuation of agent  $i$  for the  $d$ th streetlight as  $v_i(d | \omega) = U_i(d | \omega) - U_i(d-1 | \omega)$ . As before, this will summarize most of what we need and so we will often omit the  $\omega$  argument.

Recall that we have assumed that  $u_i(d, w_i | \omega)$  is strictly increasing in  $d$ , which implies that  $v_i(d | \omega) > 0$  for all  $d \geq 1$  and all  $i$ . We will also assume that  $v_i(d | \omega)$  is strictly decreasing in  $d$ . Finally, as in the last section, it is convenient to eliminate the possibility that some agent wishes to contribute more than his wealth. Hence we will assume that  $w_i$  is greater than  $i$ 's total valuation. That is,  $w_i > U_i(M | \omega)$  for all  $i$ . Analogously to the previous section, we assume  $\sum_i w_i > c(M)$  and for simplicity we also assume  $\sum_i w_i < c(M+1)$ . We will refer to this set of economies as  $\mathcal{E}^2$ .<sup>13</sup>

The core of an economy in  $\mathcal{E}^2$  is not quite as easy to characterize as the core of  $\mathcal{E}^1$ . It is straightforward to define the Pareto optimal decision as the largest  $d \in D$  such that

$$\sum_i v_i(d | \omega) \geq c(d) - c(d-1)$$

Denote this value of  $d$  by  $d^*(\omega)$ . (If there is no  $d \in D$  for which this holds, then  $d^*(\omega) = 0$ .)<sup>14</sup> Clearly, any outcome in the core must have  $d = d^*(\omega)$ . The distribution of wealth is more complex and is discussed in the Appendix where we give the proof of Theorem 2.

There are many ways one could generalize the game of the previous section to the situation considered here. The most obvious generalization would be to suppose that agents contribute any amount they choose and the largest value of  $d$  such that the contributions cover  $c(d)$  is chosen, with some rule to cover the possibility that contributions are less than  $c(1)$ . Unfortunately, in general, none of the Nash equilibria of such a game will be efficient. To see this, suppose we have two players, with  $v_i(1) = 1$  and  $v_i(2) = 0.7$  for both. Suppose  $c(1) = 1$ ,  $c(2) = 2.3$ , and  $M = 2$ . The Pareto optimal  $d$  is 2. Clearly, if we have contributions summing to 2.3, at least one player contributes at least 1. Suppose this player reduces his contribution by 1. Now the second streetlight will not

13. We conjecture that our result also holds if we extend  $\mathcal{E}^2$  to include economies with preferences that are not quasi-linear, so long as the valuation functions are decreasing in  $d$  and increasing in net wealth.

14. If  $\sum_i v_i(d^*(\omega)) = c(d) - c(d-1)$ , then outcomes with  $d = d^*(\omega) - 1$  are also in the core. The proofs do take account of this fact, though, for simplicity, the discussion in the text does not.

be built, but the first one will. This player's utility would have been  $v_i(1) + v_i(2) + w_i - \sigma_i = 1.7 + w_i - \sigma_i$ . Now it will be (at least<sup>15</sup>)  $v_i(1) + w_i - \sigma_i + 1 = 2 + w_i - \sigma_i$ , which is clearly larger. As a result, this simple generalization of the game of the previous section will not guarantee efficiency.

The intuition behind this problem is very simple. Suppose that there are many possible levels for the public good and that  $c(d) = d$ . The Pareto optimal decision will be approximately characterized by the Samuelsonian condition that the sum of the MRS's equals marginal cost, or 1. However, this implies that some agent's MRS is strictly less than 1. Hence if he could cut his contribution and, as a result, increase his wealth and decrease the amount of the public good by exactly this amount, he would be better off. To avoid this problem, the game must have the property that a small decrease in one's contribution causes a large decrease in the amount of the public good provided. Intuitively, one way to do this is to sequentially step from one level to the next all the way up to the efficient decision where at each step, insufficient contributions implies that no further steps are taken. This structure reproduces the one-streetlight game in successive rounds and thus would seem likely to achieve an efficient outcome and even implement the core.<sup>16</sup>

More concretely, we consider a multi-stage game in which agents contribute any non-negative amount of wealth they choose at each stage. If the amount contributed in the first stage falls short of  $c(1)$ , then, as before, the contributions are refunded and no streetlight is built. If the contributions are exactly equal to  $c(k)$  for some  $k \geq 1$ , then we continue with another round of contributions where  $k$  becomes the "status quo" instead of 0. The more difficult part of the game to specify is what happens if contributions fall strictly between  $c(k)$  and  $c(k+1)$  for some  $k$  between 1 and  $M-1$ . Such a situation is "falling short" of the necessary contributions in one sense and "having enough" in another. Hence it is not obvious what the appropriate incentives should be at this point. We assume that in such a situation, the difference between the amount contributed and  $c(k)$  is refunded to the agents in proportion to their contributions. Then we proceed as if exactly  $c(k)$  had been contributed.<sup>17</sup> Note that this specification has the advantage of guaranteeing budget balance. After each round, each player observes all of the contributions in that round.

Defining the game form more precisely is rather notation-intensive.<sup>18</sup> However, the idea of the game should be clear from the description above. We will refer to this game form as  $G^2 = (S^2, O^2)$ .

15. There will be some extra contributions that could be refunded so his final wealth could be larger.

16. There are certainly other ways to enrich the strategy spaces to create the sort of discontinuity efficiency requires. For example, following Mas-Colell (1980) and Bernheim and Whinston (1986), we could allow each agent to offer a contribution vector, where the  $d$ th component is the amount he pledges for  $d$  streetlights. The outcome would be the largest  $d$  such that contributions cover costs. We conjecture that the set of SUSPE outcomes of this game equals the core. However, this game seems to give a great deal of responsibility to the collector and so does not seem as natural a model of private provision as the sequential game we consider. In addition, a sequential structure is not unlike the kind of solicitation for public goods contributions we often see, while the use of alternative techniques does not seem common.

17. The particular refunding scheme used here could be replaced by any method of refunding the excess above  $c(k)$  such that increasing one's contribution by \$1 cannot increase one's refund by more than \$1. However, it is important for our results that excess contributions are always refunded. For example, consider the following variation of our game. If the status quo is unchanged at some round, the contributions at that round are refunded. If the status quo is changed, any excess is not refunded but is instead applied toward the provision of higher levels of the public good. If, for example, the status quo is 10 and more than  $c(10)$  has been contributed, then if contributions at the next round do not bring us up to  $c(11)$ , this round's contributions are refunded but no other funds are. With this scheme, one can show that every SUSPE outcome is in the core. However, in general, there will be no SUSPE's. For the reasons given in the text, weaker refinements would leave us with the possibility of inefficiency.

18. The details are contained in our 1987 working paper.

Unfortunately, this game will not fully implement the core in UPE. For example, suppose we have two identical agents and can produce at most two streetlights. Each agent values the first at 0.8 and the second at 0.7. The cost of the first is 1 and the marginal cost of the second is 1.1. It is easy to see that the following equilibrium has each agent choosing his unique best reply after every history. In the first stage, agent 1 contributes 1 and agent 2 gives nothing. In the second stage (for any history), agent 1 gives 0.45 and agent 2 gives 0.65. This implies that giving 1 in the first stage cannot be a dominated strategy or even successively dominated. Therefore, this can be used as a tremble.

We can use this tremble to support a UPE in which no streetlights are provided. To do this, suppose each player gives zero in the first stage and the most likely tremble for either player at this stage is to contribute 1. If the second stage is reached, no matter how, each player gives 0.55. Trembles at this stage are taken to be an order of magnitude less likely than the tremble to 1 in the first stage. Clearly, if the other player does not tremble to 1 in the first stage, giving zero is a best reply. If he does tremble to 1, giving zero is the unique best reply. Therefore, if any other tremble is sufficiently unlikely, zero is the best strategy. Hence this is both a UPE and a SUPE.

The reason that UPE's are not all efficient is that giving more than  $v_i(1)$  is not dominated, or even successively dominated, because we can end up with more than one streetlight. Thus as in the example of the previous section, this can be used as a tremble to support zero contributions. Hence efficiency requires that strategies be more robust. This is why we use the strict perfection part of our equilibrium definition.

The successive elimination of dominated strategies becomes important because of the strict perfection requirement. Strict perfection alone ensures that contributions must add up to at least the cost of one more streetlight at any stage where adding a streetlight is optimal. Unfortunately, though, there are no strictly perfect equilibria in subgames where we already have the efficient number of streetlights. To see this, consider some player in such a subgame. In a subgame-perfect equilibrium, contributions cannot add up to the cost of another streetlight at this point. Thus the player is necessarily indifferent over a wide range of contributions given the equilibrium strategies of the other players. However, suppose we pick trembles for the other players which makes their total contributions almost certainly either 0 or \$1 less than the cost of the additional streetlight. The player's best response to these trembles is to contribute \$1 if this is less than his valuation. But if we vary the trembles so that the others almost certainly give either 0 or \$0.50 less than the cost of another streetlight, the player's best response is to give \$0.50. Hence no strategy will be optimal for every set of trembles, as strict perfection requires.

To eliminate this problem, we need to reduce the set of possible trembles in such subgames. As the discussion above indicates, the trembles that we need to eliminate have some players contributing more than their valuations and so, again, these trembles seem quite unreasonable. Unfortunately, simply eliminating dominated strategies once does not rule them out. To see the point, suppose that  $d^*(\omega) = 2$  and  $M = 4$ . Suppose we reach a point in the game at which two streetlights have been paid for and contributions are still being collected. Contributing more than  $v_i(3)$  would seem not to be a particularly good strategy for the following reasons. First, notice that no one will contribute more than  $v_i(4)$  at the last round so that even if  $c(3) - c(2)$  is raised at this round, we will never get a fourth streetlight as contributions at the next round could not possibly sum to  $c(4) - c(3)$ . Furthermore, no one will contribute more than  $v_i(3) + v_i(4)$  at this round so that contributions at this round cannot possibly sum to  $c(4) - c(2)$ . Thus contributing more than  $v_i(3)$  is accepting a loss which cannot be made up. However, this strategy is

not dominated. The reason is that the argument suggesting that this strategy is a rather poor one relies on the argument that no one else will use a dominated strategy. Thus this strategy is not dominated, though if we eliminate dominated strategies, this strategy will be dominated in the resulting game. For this reason, we successively eliminate dominated strategies as our first step.

For these reasons, we are led to focus on successively undominated strictly perfect equilibria of SUSPE. Recall that Theorem 1 holds under this equilibrium notion as well as UPE.

**Theorem 2.**  *$G^2$  fully implements the core of  $\mathcal{E}^2$  in successively undominated strictly perfect equilibrium.*

To see the intuition behind the result, fix some  $\omega$  and ask how an equilibrium with an outcome outside  $C(\omega)$  could come about. First, it is clear that the equilibrium outcome must have  $d \leq d^*(\omega)$ . Overprovision cannot be a Nash equilibrium, much less SUSPE. It is not hard to show that strict perfection implies that contributions must add up to at least the cost of one additional streetlight whenever the status quo is less than the efficient level. Strict perfection requires optimality against all possible trembles, including the possibility that all other agents tremble to a contribution which makes it worthwhile to contribute. Since such trembles always exist when the status quo is less than the efficient level, contributions short of marginal cost can never satisfy this requirement. Thus underprovision cannot be a SUSPE either.

Hence if we have an outcome outside the core, it is because some coalition of agents is paying too much. So suppose some coalition could block the allocation. This coalition would choose some social decision less than or equal to  $d^*(\omega)$ , say  $d'$ , and pay for it themselves. Yet in our game, any of them could reduce his contribution enough at the right point to cause the social decision to fall to  $d'$  and still have some of the costs borne by others.<sup>19</sup> The fact that they do not do so indicates that the coalition could not block the allocation.

Showing that each outcome in the core is achieved by some equilibrium is quite tedious (because we have to consider all nonequilibrium histories), but intuitively clear. Notice that we can simply have a succession of rounds with one additional streetlight purchased at each round with the contributions adding to exactly marginal cost at each round. Then each player is choosing a strict best response and no sufficiently small probability of error by another player will induce a deviation.

One of the most intriguing aspects of this game is that when  $D$  is uncountable, the game can be used to approximately fully implement the core in the following sense. Let  $D_\delta = \{0, \delta, 2\delta, \dots, M(\delta)\}$ , where  $M(\delta)$  is the largest integer multiple of  $\delta$  less than or equal to  $M$ . It is simply an alteration of our notation to show that our game fully implements the core with this decision set for any  $\delta > 0$ . Intuitively, as  $\delta \downarrow 0$ , the core of the economy with  $D_\delta$  converges to the core of the economy with  $D_0 = [0, M]$ . Thus, the set of equilibrium outcomes is converging to the core of the economy with  $D_0$ .<sup>20</sup>

There are also some surprising properties of our game at the limit where  $\delta = 0$ . Our game at the limit, which we will denote  $G_0$ , is essentially a repeated version of the game considered by Bergstrom, Blume, and Varian. As we discuss below, the set of equilibrium outcomes of  $G_0$  does not even intersect the core.

19. This statement is a bit loose, but the Appendix shows that the basic reasoning is sound.

20. This statement is made more rigorous and proven in our working paper. Similarly, the issues discussed in the rest of this section are treated more formally and in more detail in the working paper.

Bergstrom, Blume, and Varian consider a game in which each agent can contribute any non-negative amount of the private good he chooses. Letting these contributions be denoted  $g_i$ , the amount of the public good provided is taken to be  $\sum_i g_i$ . A simple alteration of their game would be to allow agents to contribute in successive rounds. If the amount of contributions is strictly positive at the first round, further contributions are solicited. Once the amount contributed at a round is zero, collections cease and the public good is provided in an amount whose cost equals the total amount of money collected over the course of the game. This game is precisely  $G_0$  where  $c(d) = d$ . When  $\delta = 0$ , it is impossible for contributions to fall strictly between two levels of the public good. The determination of when another round of contributions are solicited would be precisely that the level of the public good increased—that is, that nonzero contributions were offered.

As we have seen, the limit as  $\delta \downarrow 0$  of the set of equilibria is the core of the economy with  $D_0$ . However, as we claimed, that is not the set of equilibria of  $G_0$ . Characterizing the set of equilibria in  $G_0$  poses some problems, in part because there are no SUSPE's in  $G_0$ . As discussed above, we used successive elimination of dominated strategies to ensure that we could guarantee robustness of strategies with respect to all trembles in certain subgames. One can show that this is not enough in  $G_0$  and so, in general, there are no strictly perfect equilibria even after the successive elimination of dominated strategies.<sup>21</sup>

However, it is straightforward to show that any refinement of the Nash equilibria of  $G_0$  which is based on our approximation technique and does not eliminate all equilibria will produce a set of equilibrium outcomes disjoint from the core. To see this, consider any approximating game where  $\delta = 0$  and all contributions must be in multiples of  $\mu_n > 0$ . Clearly, since wealth is finite and  $\mu_n$  is strictly positive, the number of rounds of contributions is necessarily finite. This implies that the number of streetlights in any such equilibrium must be the largest number that some agent would purchase for himself. In other words, let  $\hat{d}_i(\omega)$  be the number of streetlights which maximizes  $U_i(d|\omega) - c(d)$  and let  $\bar{d}(\omega)$  be the largest of the  $\hat{d}_i$ 's. If the number of rounds of contributions in equilibrium is finite, then we must end up with  $\bar{d}(\omega)$  streetlights. It is clear that we can't end up with fewer—any agent with  $\hat{d}_i = \bar{d}$  will contribute more. Similarly, we can't end up with a larger number of streetlights as any contributor in the the final round prefers cutting his contribution in this round. Since this happens for every  $\mu_n > 0$ , it must happen as  $\mu_n \downarrow 0$ .

It is easy to see that  $\bar{d}$  cannot be Pareto efficient since  $\bar{d}$  is where some agent's MRS equals marginal cost, not where the sum of the MRS's equals marginal cost. This is precisely the same reason why a one-shot contribution game would fail to achieve efficient provision when  $M \geq 2$ —agents must face a discontinuous drop in the level of the public good from a small decrease in their own contribution. In  $G_0$ , this cannot happen because  $\delta = 0$ .

Interestingly, the above reasoning does not apply to equilibria where contributions are given forever. In fact, there can be Nash equilibria with strong robustness properties whose outcome is in the core. To see the intuition, consider some outcome in the core. For each agent, divide the amount of wealth he is to contribute into an infinite sequence. Construct the equilibrium strategies by supposing the agents alternately contribute and agent  $i$  at his  $n$ -th "turn" contributes the  $n$ -th term in his sequence. If any agent deviates from his sequence, all agents cut their subsequent contributions to zero. The fact that

21. We conjecture that there would be SUSPE's if we defined completely mixed strategies in a manner analogous to Chatterjee and Samuelson rather than in terms of approximating games.

there are always future contributions means that any agent who cuts his contribution will discontinuously reduce provision of the public good. If we choose the sequence correctly, we can guarantee that the "status quo" at any point in any agent's sequence is Pareto-dominated by the proposed equilibrium. Thus no agent will wish to deviate since he knows this will cause the status quo at the time he deviates to be the outcome.<sup>22</sup>

The effect of our approximation technique in  $G_0$  is most easily seen by fixing the approximation so that contributions are in multiples of  $\mu_n$  and asking how the outcome is affected by taking  $\delta$  toward zero. For simplicity, suppose that  $c(d) = d$ . When  $\delta$  falls below  $\mu_n$ , we cannot have the status quo changed by one unit in a given round unless only one person makes a contribution in that round. This cannot happen in equilibrium in the last round, so that the last round of contributions must yield two units. But then either of the two contributors in the last round could deviate to zero and be strictly better off because each must have their MRS strictly less than 1 at this point. Hence we cannot reach the core with  $\delta \leq \mu_n$ . Clearly, for any fixed  $\delta > 0$ , we can make  $\mu_n$  small enough to ensure that this does not happen. But at the limit where  $\delta = 0$ , this problem is unavoidable. Generally, we think of perfect divisibility as an approximation of "small" indivisibilities and presume that this approximation does not affect the analysis. Here we see that indivisibilities in both the public and private goods crucially affect the analysis, particularly the relative magnitudes of the indivisibilities. Loosely speaking, if the indivisibilities in the public good are large relative to the indivisibilities in the private good, then core outcomes are achieved by this contribution game. Certainly it would seem plausible to argue that this is the usual case.

The role of discreteness or discontinuity in generating efficient outcomes has been seen other areas of economics such as in Aghion (1985). While this role may seem surprising at first glance, this is the same role discontinuity plays in the efficiency of perfect competition.<sup>23</sup> Perfect competition yields efficient outcomes because each firm's demand curve as a function of its price is discontinuous. If a firm's demand curve is continuous, it will in general set a price different from its marginal cost because it can exploit the fact that the outcome (its demand) varies continuously with its strategy choice to its advantage. Similarly, when  $\delta = 0$ , the continuity of the outcome function with respect to contributions allows agents to shade their contributions a small amount without consequences as severe for them as when  $\delta > 0$ .

## V. CONCLUSION.

The literature on full implementation has primarily focused on necessary and sufficient conditions on a choice correspondence for that correspondence to be fully implemented by some game form. Thus the games presented are typically used for sufficiency proofs, rather than being chosen for their plausibility. Not surprisingly, then, many of them do not seem plausible as natural games that private agents, absent some social planner, would choose to play. Even the analysis of mechanisms which are put forth as "plausibly useful", such as Groves-Clarke taxes, is focused on mechanisms that a government might actually wish to impose and rarely on mechanisms which private individuals might jointly use. Perhaps for this reason, the literature on private provision of public goods has basically ignored the implementation literature, hypothesized particular games, and demonstrated, among other things, that these games do not have efficient outcomes. We have presented a fairly natural game of private provision of public goods which fully

22. In our working paper, we show how to construct these equilibria for the two-agent case.

23. We are grateful to Andreu Mas-Colell for pointing out this analogy to us.

implements the core, thus suggesting that the literature on full implementation has more to say about private provision than might have been inferred to date.

Our results suggest that, as one might have expected, implementing the core is much more difficult in the multiple streetlight case than in the single streetlight case. To maintain a relatively simple contribution game structure, we were forced to consider a sequential game and to adopt a much stronger refinement notion. This suggests that the complexity of efficient mechanisms increases with the complexity of the environment.

An important part of our analysis has been the consideration of refinements of the Nash equilibrium concept. The fact that we are able to obtain efficient outcomes with such a simple game only by considering such refinement notions is quite suggestive. To what extent are the characteristics of the games implementating various choice correspondences driven by the equilibrium notion? In particular, is there some sense in which the Nash equilibrium concept itself leads to the “unnatural” appearance of the games implementing in Nash? If we wish to use full implementation to study institutions, we will have to learn what the equilibrium notions themselves imply about the implementing game forms.

### APPENDIX

*Proof of Theorem 1.* First, suppose that  $\sum_i v_i(\omega) < c$ . Clearly, the elimination of dominated strategies removes all  $\sigma_i \in S_i(n)$  such that  $\sigma_i \geq v_i(\omega)$ . Hence it is impossible to have contributions add to  $c$  or more in the reduced game. Therefore, all agents are indifferent over all strategies in this game and any strategy tuple in  $R^1(S(n))$  is a perfect equilibrium.<sup>24</sup> The limit of any such tuple must have a sum strictly less than  $c$ , so the set of equilibrium outcomes is  $(0, w)$ , which is the same as the core.

Now suppose that  $\sum_i v_i(\omega) = c$ . Again, once we eliminate dominated strategies, we have eliminated the possibility that the contributions can add to  $c$ . Hence, as above, we can make any strategy tuple in  $R^1(S(n))$  a perfect equilibrium. In particular, we can pick out the smallest element of each  $R^1(S_i(n))$ . This guarantees, then, that there are UPE’s of the game induced by such an  $\omega$  with an outcome of  $(0, w)$ , one of the points in the core. Similarly, we can have each agent choose the largest contribution in  $R^1(S_i(n))$ . As  $n \rightarrow \infty$ , these largest elements necessarily approach  $v_i$ . Therefore,  $\sigma = v(\omega)$  is a UPE of the game induced by such a state, which implies that  $(1, w - v(\omega))$  is a UPE outcome. It is not hard to see that there cannot be any UPE outcome other than  $(0, w)$  and  $(1, w - v)$  so that the set of UPE outcomes is exactly the core for any such  $\omega$ .

The last case, when  $\sum_i v_i > c$ , is more complex. First, we will show that every UPE outcome has  $\sum_i \sigma_i = c$ . Note that a Nash equilibrium cannot have the contributions add to strictly more than  $c$  as any contributor would then prefer a smaller contribution. Hence all UPE’s have  $\sum_i \sigma_i \leq c$ , so that we only need to show that contributions cannot add to strictly less than  $c$ .

So suppose that for some large  $n$ , we have a UPE,  $\sigma^*(n) = (\sigma_1^*(n), \dots, \sigma_n^*(n))$ , were  $\sum_i \sigma_i^*(n) < c$ . For each  $i$ , define  $v_i(n)$  as the largest element of  $S_i(n)$ . Since we have eliminated dominated strategies, this will be strictly less than  $v_i$ . Without loss of generality, number the agents so that  $v_i(n) - \sigma_i^*(n) \geq v_{i+1}(n) - \sigma_{i+1}^*(n)$  for all  $i$ . Consider the following alternative strategy for player 1. Suppose he chooses<sup>25</sup>

$$\sigma'_1 = \sigma_1^*(n) + v_2(n) - \sigma_2^*(n)$$

It is not hard to verify that our numbering of the players implies that  $\sigma'_1 < v_1$ . Since this is a UPE, we must also have

$$Pr^\varepsilon[\sum_{i \neq 1} \sigma_i(n) \geq c - \sigma_1^*(n)]u_1(1, w_1 - \sigma_1^*(n)) \geq Pr^\varepsilon[\sum_{i \neq 1} \sigma_i(n) \geq c - \sigma'_1]u_1(1, w_1 - \sigma'_1)$$

where the notation “ $Pr^\varepsilon$ ” indicates that the probability is calculated given the distribution induced by  $s^{\varepsilon*}(n)$  where this is a sequence of complexity mixed strategies in the reduced game converging to  $\sigma^*(n)$  as  $\varepsilon \downarrow 0$ .<sup>26</sup>

24. For simplicity, our notation does not reflect the fact that which strategies are dominated depends on  $\omega$ .

25. Notice that the way we have done the approximation does not guarantee that this strategy is in  $S_1(n)$ . This is not a problem. For any approximation, for  $n$  sufficiently large, it will have to be true that player 1 can choose a contribution sufficiently close to this one. It is easy to see that this fact is sufficient for our proof.

26. For convenience, we suppress the dependence of the utility functions and the game on the state  $\omega$ .



Rearranging yields:

$$\frac{\Pr^\varepsilon[\sum_{i \neq 1} \sigma_i(n) \geq c - \sigma_1^*(n)]}{\Pr^\varepsilon[\sum_{i \neq 1} \sigma_i(n) \geq c - \sigma_1']} \geq \frac{u_1(1, w_1 - \sigma_1')}{u_1(1, w_1 - \sigma_1^*(n))} \quad (\text{A.1})$$

Notice that  $u_1(1, w_1 - \sigma_1') > 0$  as  $\sigma_1' < v_1$ . Also, since  $\sigma^*(n)$  is a UPE, we must have  $\sigma_1^*(n) < v_1$  so that  $u_1(1, w_1 - \sigma_1^*(n)) > 0$ .

Let

$$A(\sigma_1) \equiv \{\sigma_{-1} \in S_{-1}(n) \mid \sum_{i \neq 1} \sigma_i \geq c - \sigma_1\}$$

be the set of contributions for the other agents such that  $d = 1$  given that player 1 contributes  $\sigma_1$ . Notice that

$$\Pr^\varepsilon[\sum_{i \neq 1} \sigma_i \geq c - \sigma_1] = \sum_{\sigma_{-1} \in A(\sigma_1)} \Pr^\varepsilon[\sigma_{-1} = \sigma_{-1}]$$

For ease of exposition, let  $A^* = A(\sigma_1^*(n))$  and  $A' = A(\sigma_1')$ . We will now show that for any  $\hat{\sigma}_{-1} \in A^*$ , there exists a vector  $\sigma'_{-1} \in A'$  such that

$$\Pr^\varepsilon[\sigma_{-1} = \hat{\sigma}_{-1}] = \xi^\varepsilon(\hat{\sigma}_{-1}) \Pr^\varepsilon[\sigma_{-1} = \sigma'_{-1}]$$

where  $\xi^\varepsilon(\hat{\sigma}_{-1}) \downarrow 0$  as  $\varepsilon \downarrow 0$ . To see this, note that any  $\hat{\sigma}_{-1} \in A^*$  must contain some components which differ from the corresponding component of  $\sigma_{-1}^*(n)$ . For each  $\sigma_{-1} \in A^*$ , choose any  $i$  such that  $\hat{\sigma}_i \neq \sigma_i^*(n)$  and construct  $\sigma'_{-1}$  by replacing  $\hat{\sigma}_i$  with  $\sigma_i^*(n)$ . Note that replacing  $\hat{\sigma}_{-1}$  with  $\sigma'_{-1}$  reduces total contributions from the agents other than 1 by  $\hat{\sigma}_i - \sigma_i^*(n)$ . However, since this is a UPE,  $\hat{\sigma}_i \in R^1(S_i(n))$  so that  $\hat{\sigma}_i \geq v_i(n)$ . Hence

$$\hat{\sigma}_i - \sigma_i^*(n) \leq v_i(n) - \sigma_i^*(n) \leq v_2(n) - \sigma_2^*(n)$$

Thus agent 1's additional contribution at  $\sigma_1'$  guarantees that total contributions are still at least  $c$ . Note also that

$$\Pr^\varepsilon[\sigma_{-1} = \hat{\sigma}_{-1}] = \frac{s_i^{\varepsilon*}(\hat{\sigma}_i; n)}{s_i^{\varepsilon*}(\sigma_i^*(n); n)} \Pr^\varepsilon[\sigma_{-1} = \sigma'_{-1}].$$

Let

$$\xi^\varepsilon(\hat{\sigma}_{-1}) = \frac{s_i^{\varepsilon*}(\hat{\sigma}_i; n)}{s_i^{\varepsilon*}(\sigma_i^*(n); n)}.$$

Note that  $\xi^\varepsilon(\hat{\sigma}_{-1}) \downarrow 0$  as  $\varepsilon \downarrow 0$  by the assumption that  $\sigma^*(n)$  is a UPE. Hence the assertion made above is true.

Let  $\xi^{\varepsilon*}(\hat{\sigma}_{-1})$  be constructed as above for each  $\hat{\sigma}_{-1} \in A^*$ . Let

$$\xi^{\varepsilon*} = \max \{ \xi^{\varepsilon*}(\hat{\sigma}_{-1}) \mid \hat{\sigma}_{-1} \in A^* \}.$$

The fact that  $A^*$  is a finite set implies that  $\xi^{\varepsilon*}$  exists and that  $\xi^{\varepsilon*} \downarrow 0$  as  $\varepsilon \downarrow 0$ .

Now the proof is virtually complete. Let  $g: A^* \rightarrow A'$  be the mapping described above. Then we see that (A.1) implies

$$\xi^{\varepsilon*} \frac{\sum_{\hat{\sigma}_{-1} \in A^*} \Pr^\varepsilon[\sigma_{-1} = g(\hat{\sigma}_{-1})]}{\sum_{\sigma'_{-1} \in A'} \Pr^\varepsilon[\sigma_{-1} = \sigma'_{-1}]} \geq \frac{u_1(1, w_1 - \sigma_1')}{u_1(1, w_1 - \sigma_1^*(n))} \quad (\text{A.2})$$

The numerator of the fraction on the left-hand side is a sum of terms all of which also appear in the denominator. Of course, the sum in the numerator may not include every term in the denominator and may include some terms several times. Let  $l$  be the largest number of times that any term appears in the numerator. Consider any term in the sum in the denominator which appears fewer than  $l$  times in the numerator. If we add this term to the numerator so that it does appear  $l$  times, we will have increased the left-hand side of (A.2). Hence we see that

$$l \xi^{\varepsilon*} \geq \frac{u_1(1, w_1 - \sigma_1')}{u_1(1, w_1 - \sigma_1^*(n))}. \quad (\text{A.3})$$

Note that the right-hand side is strictly positive as  $\sigma_1' < v_1$  and is independent of  $\varepsilon$ . Hence if we choose  $\varepsilon$  sufficiently small, we contradict (A.3). (While  $l$  is a function of  $\varepsilon$ , it is bounded from above and hence cannot go to infinity as  $\varepsilon \downarrow 0$ .) Therefore if  $n$  is sufficiently large and  $\sum_i v_i > c$ , then any UPE of  $\Gamma(n)$  must have  $\sum_i \sigma_i(n) = c$ . This contradicts the existence of a UPE outcome of  $\Gamma$  with  $d = 0$ .

Now we only need to establish that for any  $\omega$  such that  $\sum_i v_i(\omega) > c$ , every outcome in  $C(\omega)$  is a UPE outcome. Recall that  $C(\omega)$  is the set of  $\theta = (1, w - \sigma)$  with  $0 \leq \sigma \leq v(\omega)$  and  $\sum_i \sigma_i = c$ . Consider first the  $\sigma \ll v(\omega)$ . For any such  $\sigma$ , we can always find a sequence of approximating games with  $\sigma \in S(n)$  for all  $n$ . To see this, simply choose  $\mu_n = 1/n$  and let the smallest element of  $S_i(n)$  be  $\sigma_i - k\mu_n$  for the largest integer  $k$  such that this is positive. For any such  $\sigma$  and choice of  $S(n)$ ,  $\sigma$  is a strong equilibrium of  $\Gamma(n)$  (see van Damme

(1983)) and hence is a UPE. Since this is true for all  $n$ ,  $\sigma$  is a UPE of  $\Gamma$ . For the case where some  $\sigma_i = v_i$ , one can easily construct a sequence such that each person pays strictly less than  $v_i$  for each  $n$ , but the limit as  $n \rightarrow \infty$  has the appropriate people paying  $v_i$ . ||

*Proof of Theorem 2.* We provide a sketch of the proof. The interested reader is referred to our 1987 working paper for details.

First, it is easy to see that over-provision cannot even be a Nash equilibrium, much less a SUSPE. This is true because in a proposed equilibrium with over-provision, each agent has an incentive to reduce his contribution and reduce the number of streetlights provided.

It is more difficult to rule out under-provision. So suppose we have a proposed equilibrium with under-provision and consider the last stage at which contributions are solicited. Since we are below the efficient level of the public good, we can find a contribution,  $\sigma'_i$ , for each agent at this stage strictly less than his valuation for an additional streetlight but (weakly) larger than his equilibrium contribution, say  $\sigma_i^*$ , such that  $\sum_i \sigma'_i$  equals the cost of the additional streetlight. It is not difficult to show that contributing  $\sigma'_i$  is not dominated or even successively dominated for any  $i$ . Choose any agent for whom  $\sigma'_i > \sigma_i^*$  and make the  $\sigma'_j$ 's the most likely trembles for the other agents. Suppose also that all other possible trembles are orders of magnitude less likely. The strict perfection part of our equilibrium notion requires that strategies be robust with respect to these trembles. Clearly, though, if any strict subset of agents tremble to their  $\sigma'_j$ 's, the additional streetlight is not provided if  $i$  contributes  $\sigma_i^*$  or  $\sigma'_i$ . Hence in these events,  $i$  is indifferent between these two strategies. However, if all tremble, he is strictly better off if he contributes  $\sigma'_i$ .<sup>27</sup> Hence this strategy is better and so the original strategies are not a SUSPE.

So every SUSPE outcome has the efficient level of the public good. Hence if we have a SUSPE outcome not in the core, it is because the payments of some group of agents are too high. So suppose we have a SUSPE outcome  $(d', x')$  not in the core. Mas-Colell (1980) showed that this means that there is no price system supporting this outcome. That is, for any system of payments assigned to the players for each possible  $d$ , either the total payments at  $d'$  do not sum to  $c(d')$ , or some agent does not demand  $d'$ , or "profits" (the sum of the payments minus  $c(d)$ ) are not maximized at  $d'$ . However, one can construct a price system supporting  $(d', x')$  in the following way. For each person, let the amount he pays for  $d'$  be exactly his total contribution in the equilibrium. For any  $d$  larger than this, let his payment be his total contribution in the equilibrium plus his valuation for the additional streetlights. Ignoring the assignment of prices for  $d < d'$  for a moment, it is clear that this assignment makes the sum of payments equal cost at  $d'$ , guarantees that no agent strictly prefers buying more, and guarantees that profits are smaller at any higher  $d$  (since  $d'$  is efficient).

Assigning payments for  $d < d'$  is slightly trickier. First, consider the case where in equilibrium, we reach a point where  $d$  is the status quo at some round. Let the payments for this  $d$  for  $i$  be the total payments  $i$  has made up to this point. Clearly, profits at such a  $d$  are zero, so that  $d'$  still maximizes profits. Similarly, each agent is willing to continue on and so prefers buying  $d'$  to  $d$ . Second, consider the case where in equilibrium, the status quo "hops over"  $d$ . At the round at which this occurs, each agent could reduce his contribution and cause  $d$  (or a higher level if the constraint that contributions be nonnegative interferes) to become the status quo. Assign  $i$ 's payment to be the amount he has paid up to this round plus whatever payment he would be making when  $d$  becomes the status quo. It is not hard to show that with this assignment of prices,  $d$  cannot yield higher profits than  $d'$ . To see that no agent would prefer to buy  $d$ , notice that in equilibrium, any agent could cut his contributions in the manner specified and refuse to contribute any further. This must leave him at least as well off as at  $d$  with the payment we have assigned. Since he does not choose to do this in equilibrium, he is better off at  $d'$ . Hence the outcome is in the core. Therefore, the SUSPE outcomes are a subset of the core.

We now explain to construct a SUSPE for any given point in the core. First, we construct the equilibrium path by assigning a contribution for each agent for each round up to  $d^*$  so that contributions at a round sum to the cost of one additional streetlight and each agent gives less than his valuation for this streetlight. Clearly, each agent strictly prefers following the equilibrium path to contributing less. If any agent contributes more, the equilibrium path contributions are still followed if possible. More specifically, the contributions for the  $d$ -th streetlight are exactly what they would have been on the equilibrium path unless wealth constraints become binding because of the deviation. If wealth constraints do become binding, we construct a new equilibrium path from here based on the same idea. That is, we choose payments at each round adding to marginal cost with each agent paying less than the minimum of his valuation and his remaining wealth as long as the sum of these minima exceeds marginal cost. We handle deviations from this path in an analogous manner. Since wealth constraints can only become binding for some agent as a consequence of his own actions and since he

27. It is not difficult to see that in subsequent rounds of contributions, he cannot be made worse off since he can always contribute zero for the rest of the game.

must be worse off than in the equilibrium if he does this, no agent will wish to deviate in this way. Furthermore, the fact that increasing one's contributions less drastically leads to a refund to the agent of less than the increase in his contribution and does not change subsequent contributions means that this deviation is also strictly worse than staying on the equilibrium path. The fact that following the assigned strategies is always strictly better than deviating means that this is optimal against any trembles.

Stages at which the sum of the smaller of the agent's valuation for another streetlight and his remaining wealth falls short of marginal cost do not concern us. It is not difficult to show that the successive elimination of dominated strategies means that contributions cannot add to marginal cost at such a stage. This means that any successively undominated strategy at this point is optimal against any trembles. Thus these strategies form a SUSPE that supports the chosen core allocation. ||

*Acknowledgement.* Some of this paper is based on our "Can Private Provision of Public Goods be Efficient?", especially Section III. We would like to thank Ted Bergstrom, Andreu Mas-Colell, Tom Palfrey, Hal Varian, participants at the Midwest Mathematical Economics Society Meetings, John Moore, and two anonymous referees for helpful comments and suggestions. We would especially like to thank Sanjay Srivastava. As usual we are solely responsible for any errors. We are also grateful for the financial support of the National Science Foundation through NSF Grant SES-8520296.

#### REFERENCES

- AGHION, P. (1985), "On the Generic Inefficiency of Differentiable Market Games", *Journal of Economic Theory*, **37**, 126-146.
- ANDREONI, J. (1988), "Privately Provided Public Goods in a Large Economy: The Limits of Altruism", *Journal of Public Economics*, **35**, 57-74.
- BAGNOLI, M., and LIPMAN, B. (1986), "Can Private Provision of Public Goods be Efficient?" (University of Michigan working paper).
- BAGNOLI, M., and LIPMAN, B. (1987) "Provision of Public Goods: Fully Implementing the Core through Private Contributions" (University of Michigan working paper).
- BERGSTROM, T., BLUME, L. and VARIAN, H. (1986) "On the Private Provision of Public Goods", *Journal of Public Economics*, **29**, 25-49.
- BERNHEIM, B. D. (1986), "On the Voluntary and Involuntary Provision of Public Goods", *American Economic Review*, **76**, 789-793.
- BERNHEIM, B. D., and WHINSTON, M. (1986), "Menu Auctions, Resource Allocation, and Economic Influence", *Quarterly Journal of Economic*, **101**, 1-33.
- BLISS, C., and NALEBUFF, B. (1984), "Dragon-Slaying and Ballroom Dancing: The Private Supply of a Public Good", *Journal of Public Economics*, **25**, 1-12.
- GHATTERJEE, K., and SAMUELSON, L. (1986), "Perfect Equilibria in Simultaneous-Offers Bargaining" (Pennsylvania State University working paper).
- CORNES, R., and SANDLER, T. (1985), "The Simple Analytics of Pure Public Goods Provision", *Economica*, **52**, 103-116.
- D'ASPREMONT, C., and GÉRARD-VARET, L. (1979) "Incentives and Incomplete Information", *Journal of Public Economics*, **11**, 25-45.
- FEREJOHN, J., FORSYTHE, R., NOLL, R. and PALFREY, T. (1982), "An Experimental Examination of Auction Mechanisms for Discrete Public Goods", in Smith, V. (ed), *Research in Experimental Economics*, Vol. 2, (Greenwich, Conn.: JAI Press).
- ISAAC, R., WALKER, J. and THOMAS, S. (1984), "Divergent Evidence on Free-Riding: An Experimental Examination of Possible Explanations", *Public Choice*, **43**, 113-150.
- LAFFONT, J.-J., and MASKIN, E. (1982) "The Theory of Incentives: An Overview", in Hildenbrand, W. (ed.), *Advances in Economic Theory* (Cambridge: Cambridge University Press).
- MAILATH, G., and POSTLEWAITE, A. (1988), "Asymmetric Information Bargaining Problems with Many Agents" (University of Pennsylvania working paper).
- MAS-COLELL, A. (1980), "Efficiency and Decentralization in the Pure Theory of Public Goods", *Quarterly Journal of Economics*, **94**, 625-641.
- MASKIN, E. (1977), "Nash Equilibrium and Welfare Optimality" (unpublished manuscript).
- MOORE, J., and REPULLO, R. (1988) "Subgame Perfect Implementation", *Econometrica*, **56**, 1191-1220.
- MYERSON, R. (1978), "Refinements of the Nash Equilibrium Concept", *International Journal of Game Theory*, **7**, 000-000.
- NASH, J. (1953), "Two-Person Cooperative Games", *Econometrica*, **21**, 128-140.
- PALFREY, T., and ROSENTHAL, H. (1984) "Participation and the Provision of Discrete Public Goods", *Journal of Public Economics*, **24**, 171-193.
- PALFREY, T., and ROSENTHAL, H. (1988) "Private Incentives in Social Dilemmas: The Effects of Incomplete Information and Altruism", *Journal of Public Economics*, **35**, 309-332.

- PALFREY, T., and SRIVASTAVA, S. (1986), "Nash Implementation using Undominated Strategies" (Carnegie-Mellon University working paper).
- SCHNEIDER, F. and POMMEREHNE, W. (1981), "Free Riding and Collective Action: An Experiment in Public Microeconomics", *Quarterly Journal of Economics*, **96**, 689-704.
- van DAMME, E. (1983) *Refinements of the Nash Equilibrium Concept* (Berlin: Springer-Verlag).
- WARR, P. (1983), "The Private Provision of a Public Good is Independent of the Distribution of Income", *Economic Letters*, **13**, 207-211.