Center for Research on Economic and Social Theory

CREST Working Paper

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

Mark Bagnoli
Barton L. Lipman

March, 1987
Number 87--12

DEPARTMENT OF ECONOMICS
University of Michigan
Ann Arbor, Michigan 48109

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

by

Mark Bagnoli
University of Michigan

and

Barton L. Lipman
Carnegie–Mellon University

Revised

March, 1987

* 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, Tom Palfrey, Andreu Mas-Colell, Hal Varian, and participants at the Midwest Mathematical Economics Society Meetings, for helpful comments and suggestions. We would especially like to thank Sanjay Srivastava. As usual we are solely responsible for any errors that we may have made. We are also grateful for the financial support of the National Science Foundation through NSF Grant SES-8520296.

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. Warr [1983], Palfrey and Rosenthal [1984, 1985], Bergstrom, Blume, and Varian [1986], Andreoni [1985], and others have analyzed games in which citizens freely contribute towards the provision of a public good.[1] Confirming the standard view of the free rider problem, the equilibria of these contributions games typically have inefficient outcomes—specifically, not enough of the public good is provided.

By contrast, the literature on full implementation has demonstrated the existence of games for which all equilibria are efficient.[2] Given this fact, why is it supposed that the agents in the economy play a game which has an inefficient equilibrium? Presumably, the games used to model private provision are taken to be "natural" representations of private provision, while the games used in the implementation literature are not viewed this way. The games used to prove existence of games with efficient outcomes generally seem to require a social planner to impose and to "mediate" them. That is, they do not seem to be reasonable descriptions of games which would arise "naturally."

Contrary to this conclusion, we will provide an example of an efficient mechanism which appears to be a very natural contribution game. In the simplest version of this example, we consider a complete information economy with a single private good and a discrete public good. The simple game that we consider is to allow 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, the set of equilibrium outcomes of this game (where the equilibrium notion is a slight modification of perfection) is exactly the core of the economy. We extend the game to the case of a public good that can take on finitely many values and prove essentially the same result. Since this extension works for any finite number of values for the public good, we can reinterpret the finite case as approximating the case where the public good can take on any value in a continuum. It is straightforward to show that as the increment between values (the "step size") goes to zero, the outcome converges to the core of the economy where the public good can take on any value in the continuum. Interestingly, as the step size goes to zero, the game form converges (in an appropriate sense) to a repeated version of the game considered by Bergstrom, Blume, and Varian. We analyze the set of equilibrium outcomes of

[1] Related papers include Bliss and Nalebuff [1985], Cornes and Sandler [1985a, 1985b], Sudgen [1985], Bernheim [1986], and Bernheim and Bagwell [1985].

[2] See Maskin [1977], Moore and Repullo [1986], and Palfrey and Srivastava [1986].

the game at the limit and demonstrate some interesting differences between the limiting outcomes and the outcomes at the limit. This provides an intriguing view of the role of discontinuities in the game form, a point noted in a different context by Aghion [1985].

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 decisions 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 $\{0, 1, \ldots, M\}$ 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 $\{0, \delta, 2\delta, \ldots, M(\delta)\}$ where $\delta > 0$ and $M(\delta)$ is the largest multiple of δ 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. The properties of our game as $\delta \downarrow 0$ are immediate. In Section V, we formalize this and analyze the game at $\delta = 0$. In Section VI, 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 $i \in \mathcal{I} = \{1, \ldots, I\}$. 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 ω, specifies each agent's utility function. The set of possible states will be denoted Ω. We will write the utility function for the i^{th} agent in state ω as $u_i(d, w_i \mid \omega)$. 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 all ω. We will impose some further conditions on Ω in the subsequent sections.

For convenience, we define the cost function for the social decisions as a function $c: \mathbf{R}_+ \to \mathbf{R}_+$. Of course, c is only relevant on D. We assume that the cost function is strictly increasing and concave 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 in $D \times \mathbf{R}^I_+$. We will use θ to denote a generic outcome and will let Θ denote the set of feasible outcomes—that 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 $C: \Omega \to P(\Theta)$ 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 Θ_S as the set of feasible outcomes for coalition $S \subseteq I$. That is,

$$\Theta_S = \{(d, x) \in D \times \mathbf{R}^I_+ \mid \sum_{i \in S} x_i \leq \sum_{i \in S} w_i - c(d)\}$$

In the usual terminology, we will say that the coalition $S \in P(I)$ can block the outcome $\theta \in \Theta$ in state ω if there exists some $\theta' = (d', x') \in \Theta_S$ such that

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

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

A game form, G, is a pair (S, O) where $S = S_1 \times \ldots \times S_I$ and $O: S \to \Theta$. 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 ω. Thus, for a given G, the set of equilibrium strategies must be defined as a correspondence from Ω 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 ω, 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.[3] 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 strict perfection to the resulting game. We will refer to this as successively undominated strictly perfect equilibrium or SUSPE.[4] (Before defining these concepts precisely, we should note that the reader already familiar with these notions can avoid the notation of what follows and proceed to the next section.)

These definitions are notationally complex for several reasons. First, we need to define, for an arbitrary game form, a correspondence from Ω to S giving the equilibrium strategy tuples. However, equilibrium notions are defined on games, not game forms. Thus we must define this correspondence with reference to the game induced by the game form and the state. Rather than carrying around notation indicating that the game depends on the state and the game form, we will simply define our equilibrium notions for a game and then define the equilibrium correspondence.

The second source of complexity is that we work with action sets which are uncountable. As discussed further below, we deal with these sets through sequences of finite approximations. Thus we have an n in our notation to denote the level of approximation. Rather than carrying around the n in our notation, we will define the approximation technique and then only define our equilibrium notions for finite games directly.

Third, we will consider some sequential (or multi-stage) games and thus wish to define our equilibrium notions for the extensive form, not the normal form. Most equilibrium notions like perfection are modified for the extensive form by simplying applying the definition to the agent-normal form, rather than the normal form.[5] We do the same here, requiring us to define the agent-normal form and, in principle, carry notation distinguishing the agent-normal from the normal form. Instead, we will define the agent-normal form of an extensive form game and then

[3] See the classic example in Myerson [1978].

[4] 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 $\sum_i v_i = c$.

[5] See van Damme [1985], Chapter 6, for a detailed discussion.

define our equilibrium notions for a finite, normal form game.

The fourth and final source of notational discomfort is that our equilibrium notions all begin with an elimination of dominated strategies and then apply a standard equilibrium notion. Thus there are basically two levels of definition in each equilibrium notion for finite games. To deal with this notation efficiently, we will define a mapping which reduces a game to a simpler game by eliminating dominated strategies. We will then, separately, define the equilibrium notions applied to these reduced games for general finite games. At the end, we will assemble these definitions to make the linkages as clear as possible.

So consider a normal form game $\Gamma = (S, u)$. In all of the games we consider, S will be an uncountable set. 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.[6] Thus we "discretize" the game and analyze games arbitrarily "close" to the continuous game. To describe the approximation, note first that we can, without loss of generality, define each S_i to be a set of mappings from a set of "histories observed by i," H_i, into a set of "actions," A_i, where there may be some constraint on the set of possible maps (because, for example, some actions may be infeasible on certain histories). That is, we can work in terms of the extensive form of the game, where H_i is the set of information sets for i and A_i is the set of actions i can take at some information set. Clearly, for a one-shot, simultaneous move game, H_i is a singleton, so that, for this case, we are going to more trouble than is necessary. However, we will require this level of care for the sequential games discussed in Sections IV and V. In all of what we consider, each A_i will be an interval of the real line $[\underline{a}^i, \bar{a}^i]$.

Since we will work with the agent–normal form, "making the game finite" simply requires making the strategy set for each agent for any player finite. That is, we only need to discretize the A_i sets, not the S_i sets. Notice that making the A_i sets finite is not sufficient to imply that the strategy sets are finite. For example, if there is an uncountable set of moves by nature which are observable, then the strategy set would be uncountable. In all the games we will consider, there will be no moves by nature so that this consideration need not concern us. Thus in any approximating game, the fact that the set of actions is finite implies that the set of histories a player can observe is, at most, countable. However, the set of maps from a countable set to a finite set is still uncountable, so that the S_i sets can still be uncountable. In fact, as we will see, the

[6] 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.

approximating strategy sets will be uncountable in the game we consider in Section V.

Let $A_i(n)$ denote a finite subset of A_i, which we will denote $\{a_1, \ldots, a_m\}$. We require that $A_i(n)$ have the property that for all $k \in \{1, \ldots, m\}$ and all i,

(1) $$| a_1 - a_k | = (k-1) | a_1 - a_2 | = (k-1)\delta_n$$

Also,

(2) $$| \underline{a}^i - a_1 | \leq \delta_n$$

(3) $$| a_n - \bar{a}^i | \leq \delta_n$$

Notice that δ_n is the same for all i. Thus we require that the elements of the approximating set $A_i(n)$ are equally spaced from each other for all i and must be close to the extreme points of A_i. Let $H_i(n)$ be the set of histories in H_i for which every player has chosen an action in $A_i(n)$. Finally, let $S_i(n)$ denote the set of maps in S_i which only map from points in $H_i(n)$ into points in $A_i(n)$. Then, letting $S(n) = S_1(n) \times \ldots \times S_I(n)$, let $\Gamma(n)$ be the game $(S(n), u)$.

For each equilibrium notion we use, we will say σ is an equilibrium of Γ if it is the limit of a sequence $\{\sigma(n)\}$ of equilibrium points of a sequence of approximating games $\{\Gamma(n)\}$. More precisely, let ϕ be an equilibrium correspondence defined on the set of games with finite strategy sets. We will extend the correspondence ϕ to a game with an uncountable strategy set in the following way. Let N denote some subsequence of the positive integers. We will say that $\sigma \in S$ is an element of $\phi(\Gamma)$ iff there exists a sequence $\{\sigma(n)\}_{n \in N}$ converging to σ such that for each $n \in N$, $\sigma(n) \in \phi(\Gamma(n))$.[7]

Now we are ready to begin defining equilibrium notions. We first define the agent-normal form and then define equilibrium notions to be applied to this form. So consider a game (S, u) where, for each i, S_i is the set of maps from i's possible information sets, H_i, into a finite set of possible actions for i, A_i. The agent-normal form of this game is constructed by replacing the players with "agents" who represent the players, where each player has a different agent working for him at each information set. In other words, we alter the set of players from $\{1, \ldots, I\}$ to the set

$$\{ih \mid i \in I, h \in H_i\}$$

The strategy set of agent ih is simply those actions that i might choose at information set h. That is,

$$S_{ih} = \{a \in A_i \mid a = \sigma_i(h) \text{ for some } \sigma_i \in S_i\}$$

[7] Notice that we are allowing any subsequence of the positive integers and any approximation satisfying (1) through (3) for each n.

Notice that the fact that A_i is finite immediately implies that each agent has a finite strategy set. Given a vector of strategies for all the agents, we can immediately define the implied strategies for the players by

$$\sigma_i(h) = a_{ih}$$

where a_{ih} is the action chosen by ih. Then the payoff to agent ih is $u_i(\sigma)$—that is, ih receives the same payoff as i.

The agent–normal form has several advantages over the normal form when considering a sequential game. Basically, it eliminates correlation in "trembles" across different stages in the game. Since we take the agents to be separate players, if i makes a "mistake" at one stage, there is no reason to believe that he will do so again at a later stage. It is also worth noting that eliminating dominated strategies in the agent–normal form is more difficult than eliminating them in the normal form. Thus performing this removal as we do eliminates fewer possible equilibria than if we worked with the normal form. The reason this is true is quite simple. In the agent–normal form, one considers a strategy for an agent ih to be dominated based on a consideration of what all other agents might do— including i at a different information set. By contrast, when working with the normal form, one can eliminate action combinations over information sets. For example. suppose that i has two successive information sets. Consider a strategy which specifies an action at each information set with the property that his action at the first information set is disastrous given his action at the second information set. This strategy could well be dominated in the normal form. However, in considering the agent–normal form, we do not consider i's choice at the second information set as fixed. Thus we may not be able to eliminate this strategy.

We will now define the equilibrium notions for a game $\Gamma = (S, u)$. We assume that Γ is in agent–normal form so that we can take S_i to be a finite set for each i. (We refer to agents here as i instead of ih for simplicity.) We will say that a strategy for i, $\sigma_i \in S$, is dominated if there exists $\sigma_i' \in S_i$ such that for all $\sigma_{\sim i} \in S_{\sim i}$,

$$u_i(\sigma_i', \sigma_{\sim i}) \geq u_i(\sigma_i, \sigma_{\sim i})$$

with a strict inequality for some $\sigma_{\sim i} \in S_{\sim i}$. Let $R^1(S_i)$ denote the set of strategies for i which are not dominated and let $R^1(S) = R^1(S_1) \times \ldots \times R^1(S_I)$. Then let $R^1(\Gamma)$ denote the game $(R^1(S), u)$. An undominated perfect equilibrium of Γ 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 $R^1(\Gamma)$. In other words, $R^n(S_i) = R^1(R^{n-1}(S_i))$. 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 Γ is a perfect equilibrium of $R^*(\Gamma)$ and similarly for a successively 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. Let Δ_i° be the set of probability mass functions over S_i such that each element of S_i is given strictly positive probability. That is Δ_i° is the set of totally mixed strategies for each player. A typical element of Δ_i° will be written s_i where $s_i(\sigma_i)$ is the probability that i chooses the strategy σ_i. Let $\Delta_{\sim i}^\circ$ and Δ° denote the usual cartesian products. A typical element of $\Delta_{\sim i}^\circ$ will be written $s_{\sim i}$ and a typical element of Δ° will be written s. Finally, let $V_i(\sigma_i, s_{\sim i})$ be the expected payoff to i when he chooses pure strategy σ_i and the other players choose mixed strategies $s_{\sim i}$, where expected utilities are defined in the usual way. Let N denote some sequence of the positive integers. Then σ is a perfect equilibrium for Γ if there exists a sequence $\{s^n\}_{n\in N}$ such that

$$s^n \in \Delta^\circ, \quad \forall n \in N, \tag{4}$$

$$\lim_{n\to\infty} s_i^n(\sigma_i) \to 1, \quad \forall i \in I, \tag{5}$$

and

$$V_i(\sigma_i, s_{\sim i}^n) \geq V_i(\sigma_i', s_{\sim i}^n), \quad \forall i \in I, \ \forall n \in N, \text{ and } \forall \sigma_i' \in S_i. \tag{6}$$

In other words, for each i, σ_i is a best reply to some vector of completely mixed strategies for the other players close to $\sigma_{\sim i}$.

Strict perfection requires much more—essentially that σ_i is a best response to *every* vector of completely mixed strategies for the other players close to $\sigma_{\sim i}$. 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.

More precisely, σ is a strictly perfect equilibrium of Γ iff for *any* sequence $\{s^n\}_{n\in N}$ in Δ° for all $n \in N$ such that σ is the limit as $n \to \infty$ of $\{s^n\}$, we have

$$V_i(\sigma_i, s_{\sim i}^n) \geq V_i(\sigma_i', s_{\sim i}^n)$$

for all i, all $\sigma_i' \in S_i$, and all $n \in N$. 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.

We are finally ready to assemble definitions. Let Γ be any agent–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 player is finite. We will say that σ is an undominated perfect equilibrium (UPE) of Γ iff it is a perfect equilibrium of $R^1(\Gamma)$. Similarly, σ is a successively undominated strictly perfect equilibrium (SUSPE) of Γ iff it is a strictly perfect equilibrium of $R^*(\Gamma)$. If Γ has an uncountable strategy set, we say that σ is a UPE (SUSPE) of Γ if there is a sequence $\{\sigma(n)\}_{n\in N}$ converging to σ such that $\sigma(n)$ is a UPE (SUSPE) of $\Gamma(n)$ for all $n \in N$.

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 σ is a UPE (SUSPE) of the game induced by G at state ω.

III. The One–Streetlight Problem.

In this section, we simplify the above structure to the case where $D = \{0, 1\}$. For simplicity, we let $c = c(1)$. Without loss of generality, we adopt the normalization $u_i(0, w_i \mid \omega) = 0$ for each i and each $\omega \in \Omega$. The valuation of agent i in state ω, $v_i(\omega)$, is defined implicitly by $u_i(1, w_i - v_i \mid \omega) = 0$. Since the valuations define everything about u_i that is relevant for our purposes at a state ω, we will often omit the ω argument and focus on the valuations directly. We also assume that for each state, $w_i > v_i(\omega)$ for all i—that is, $u_i(1, 0 \mid \omega) < 0$ 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. In short, we take the set Ω to be $\{\omega \mid 0 \ll v(\omega) \ll w\}$. 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.[8]

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. A variety of assumptions could be made concerning 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.[9] 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[10] for the same purpose.[11]

More formally, let the strategy set of agent i be $S_i^1 = [0, w_i]$. A strategy choice by i will be denoted σ_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

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.

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 σ such that $0 \leq \sigma_i \leq v_i$ for all i and $\sum_i \sigma_i = c$ must be a Nash equilibrium. Since $\sigma_i \leq v_i$, each agent's equilibrium utility 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 pure strategies. In particular, consider any σ 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 $c = 1$, $I = 3$, and each person's valuation is .5. Then $\sigma_i = 0$ for all i is a perfect equilibrium. In particular, suppose each person puts probability $1 - \epsilon$ on 0, $k\epsilon/(1+k)$ on .5, and the rest of the probability on the remaining strategies. Choose k to be very large. Then it is virtually certain that the sum of the other two players' contributions is either 0, .5, or 1. In any of these cases, i's best strategy is $\sigma_i = 0$ and in the last 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 either he will end up contributing when the good would be provided without his contribution or he would have to contribute v_i to have any effect.

Intuitively, these trembles do not seem plausible. The equilibrium in the example required that the agents were most likely to tremble to contributing their entire valuations. Clearly, though, any smaller contribution weakly dominates this one. Hence if we eliminate such dominated strategies even as trembles. then this possibility is eliminated. This is why we focus on undominated perfect equilibria.

The result 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. Thus, for example, Theorem 1 is trivially extended to undominated proper equilibria or undominated strictly perfect equilibria.

The elimination of dominated strategies before applying one of these two equilibrium notions is not necessary for eliminating the equilibria not in the core when $\sum_i v_i > c$. As we have shown elsewhere,[12] simply focusing on proper equilibrium without first removing dominated strategies has this effect and hence the same must be true for strict perfection.[13] The problem with these concepts is that they remove too many equilibria in the cases where $\sum_i v_i \leq c$. In particular, one can show that the only proper equilibrium outcome when $\sum_i v_i = c$ is $(1, w - v)$ so that the set of equilibrium outcomes does not exactly coincide with the core at such a state. However, one can show that this is the only circumstance in which this is true so that G^1 implements the core of $\mathcal{E}^1$ generically in proper equilibria. One can also show that when $\sum_i v_i \leq c$, there are no strictly perfect equilibria (in pure strategies). Removing dominated strategies first guarantees that enough of the Nash equilibria for the cases where $\sum_i v_i \leq c$ remain when properness or strict perfection is used.

It is also worth noting that the set of successively undominated perfect equilibrium (SUPE) outcomes is identical to the set of UPE outcomes. To see this, note that the strategies removed in the first round of elimination are those where some agent contributes more than v_i or contributions so small that $c - \sigma_i > \sum_{j \neq i} w_j$. Any other contribution can be a best reply to some $\sigma_{\sim i}$ and

[12] See Bagnoli and Lipman [1986].

[13] The fact that properness eliminates the equilibria with outcomes outside the core is intuitively clear from the fact that these equilibria can only be perfect if the trembles are to dominated strategies The problem of trembles to dominated strategies was precisely the motivation for Myerson's [1978] introduction of the properness concept

hence cannot be eliminated. Thus if $\sum_i v_i < c$, this round of elimination already implies that any equilibrium outcome is in the core and no further elimination will change this. Suppose that $\sum_i v_i > c$. Since the equilibria with outcomes in the core all have strict best replies chosen by each player, none of these strategies can be dominated so that no further elimination of dominated strategies will eliminate these equilibria. Since there are no other equilibria, the set of SUPE outcomes would be the same as the set of UPE outcomes. In short, Theorem 1 holds for SUPE and SUSPE as well as UPE.

IV. The Multiple Streetlights Problem.

In this section, we consider a broader social decision set. Here we take D to be $\{0, \ldots, M\}$ for some finite $M \geq 1$—for example, how many streetlights to build. Since we will find it necessary to work with sequential mechanisms, eliminating wealth effects is quite useful. Therefore, we will simplify our assumptions on preferences and assume that $u_i(d, w_i \mid \omega) = U_i(d \mid \omega) + w_i$ 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 $v_i(d \mid \omega) = U_i(d \mid \omega) - U_i(d-1 \mid \omega)$. As before, this will summarize most of what we need and so we will often omit the ω argument.

Recall that we have assumed that $u_i(d, w_i \mid \omega)$ is strictly increasing in d, which implies that $v_i(d \mid \omega) > 0$ for all $d \geq 1$ and all i. We will also assume that utility is strictly concave in d—more precisely, we assume that $v_i(d \cdot \omega)$ 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, $w_i > U_i(M \mid \omega)$ for all i. In other words,

$$\Omega = \{\omega \mid 0 < v_i(d \mid \omega) < v_i(d-1 \mid \omega) \text{ for all } i, d \text{ and } \sum_{d \in D} v_i(d \mid \omega) < w_i \text{ for all } i\}$$

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 $\mathcal{E}^2$.

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

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

Denote this value of d by $d^*(\omega)$. (If there is no $d \in D$ for which this holds, then $d^*(\omega) = 0$.)[14] Clearly, any outcome in the core must have $d = d^*(\omega)$. The distribution of wealth is more complex. To characterize this, we make use of results in Mas–Colell [1980]. We will say that a price system is a vector of functions $(p_1, \ldots, p_I)$ with $p_i : D \times \Omega \to \mathbf{R}_+$. An outcome (d', x') is supported by a price system at ω if

$$\sum_i p_i(d' \mid \omega) = c(d') \tag{7}$$

$$(d', x'_i) \text{ maximizes } u_i(d, x_i \mid \omega) \text{ on } \{(d, x_i) \mid p_i(d \mid \omega) + x_i = w_i\} \tag{8}$$

$$d' \text{ maximizes } \sum_i p_i(d \mid \omega) - c(d) \text{ on } D \tag{9}$$

Mas-Colell proves the following characterization of the core which we will make use of in our proofs.

Proposition. *The outcome (d, x) is a supported by a price system at ω iff $(d, x) \in C(\omega)$.*

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 $c(1)$. Such a game will not implement the core under any refinement of the Nash equilibria. To see this, choose ω such that $d^*(\omega) = 2$, $\sum_i v_i(2 \mid \omega) < c(2)$, and $v_i(2 \mid \omega) < c(2) - c(1)$ for all i. (For a concrete example, suppose $I = 3$, $v_i(1) = 1$ for all i, $v_i(2) = 1/2$ for all i, $c(1) = 1$, $c(2) = 2.3$, and $M = 2$.) Consider any set of contributions summing to $c(2)$ such that each individual contributes less than $v_i(1) + v_i(2)$. This outcome is in the core and these strategies are the only way to achieve this outcome with this game. However, these strategies cannot constitute a Nash equilibrium. Note that the fact that $\sum_i v_i(2) < c(2)$ implies that there must exist some agent i whose contribution strictly exceeds $v_i(2)$. Consider the deviation from the proposed equilibrium whereby this agent reduces his contribution by $v_i(2) + \epsilon$ for some small $\epsilon > 0$. By assumption, $v_i(2)$ is less than $c(2) - c(1)$, so that the outcome with this reduced contribution will have $d = 1$ if ϵ is sufficiently small. But then he loses one streetlight worth $v_i(2)$ to him, but reduces his payments by more than this. Hence by deviating, he increases his expected payoff so that the strategies cannot be an equilibrium.

Thus if we wish to retain the contribution game structure, we must necessarily consider a sequential game of some kind. The one we choose is certainly not the only natural extension of G^1,

[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.

but appears to be quite a reasonable one. We suppose that agents contribute any non-negative amount of wealth they choose. If the amount contributed 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. Intuitively, this structure reproduces the one-streetlight game in successive rounds and thus would seem likely to implement the core. 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.

Defining the game form more precisely is rather notation-intensive. First, we define histories and a function O^* from histories into outcomes. Then we can define the strategy sets, S^2. Since a strategy in our game is a mapping from histories into feasible actions and since actions determine what is feasible at the next round, this order of exposition is necessary. Finally, we define a function from strategies into histories, giving the play of the game for a set of strategies. The outcome function O^2, then, is simply the composition of O^* with this map.

In the game we consider, all past contributions by all agents are common knowledge, so that a history of length m is an mI vector. We define the set of histories analogously to Kalai and Stanford [1986]. The set of histories of length 0, H^0, consists of the single element $\{e\}$ which is the "empty history." The set of histories of length m is $H^m = \mathbf{R}_+^{mI}$. The set of all histories, then, is

$$H = \bigcup_{m=0}^{\infty} H^m$$

Notice that we need not consider histories longer than M as our game must end after $M-1$ rounds. However, it will prove convenient later to allow for arbitrarily long histories. We let h denote a generic element of H and use $\ell(h)$ denote the length of the history h. Not all histories in H are physically possible, not even all of those with length less than M, because of the fact that agents cannot contribute more than their wealth. Once we have a few more definitions, we will identify the set of feasible histories.

To define the mapping from histories into outcomes, we need a bit more notation. Define the projection operators $P_m: H \rightarrow H^1$ for $m = 1, \ldots,$ as giving for any h, the m^{th} I vector in h. If

$m > \ell(h)$, we take $P_m(h)$ to be e. That is, $P_m(h)$ tells what the contributions were at the m^{th} round along the history h. We can use the projection operators to define the concatenation of two histories h and h' as follows. The concatenation, which we will denote $h \cdot h'$, is defined as that history h'' of length $\ell(h) + \ell(h')$ such that

$$P_m(h'') = \begin{cases} P_m(h), & \text{for } m \leq \ell(h); \\ P_{m-\ell(h)}(h'), & \text{otherwise.} \end{cases}$$

Notice that our definition implies $e \cdot h = h \cdot e = h$ for any history h.

We recursively define the outcomes at a round as follows. Suppose h is a history of length $\ell \geq 1$. Then

$$o^1(h) = (k^1(h), r^1(h), t^1(h))$$

where $k^1(h)$ is the largest integer k such that

$$\iota P_1(h) \geq c(k)$$

where ι is an I vector of all ones. The function $r^1(h)$ is given by

$$r^1(h) = \begin{cases} 1, & \text{if } k^1(h) \geq 1; \\ 0, & \text{otherwise.} \end{cases}$$

and $t^1(h)$ gives actual contributions at round 1. Recall that some contributions may be refunded. The actual contribution by the i^{th} agent is the amount he offered minus his refund. His refund is $\iota P_1(h) - c(k^1(h))$ times the proportion his contribution was of the total. Hence

$$t^1(h) = P_1(h) - \frac{[\iota P_1(h) - c(k^1(h))]}{\iota P_1(h)} P_1(h)$$

or

$$t^1(h) = \frac{c(k^1(h))}{\iota P_1(h)} P_1(h)$$

(If $P_1(h) = 0$, then $t^1(h) = 0$.) Then for $m \leq \ell$, we can calculate $o^m(h) = (k^m(h), r^m(h), t^m(h))$ by $o^m(h) = o^{m-1}(h)$ if $r^{m-1}(h) = 0$ and otherwise is given as follows. Analogously to the above, $k^m(h)$ is the largest integer k such that

$$\iota P_m(h) \geq c(k) - c(k^{m-1}(h))$$

$$r^m(h) = \begin{cases} 1, & \text{if } k^m(h) \geq k^{m-1}(h) + 1; \\ 0, & \text{otherwise.} \end{cases}$$

$$t^m(h) = t^{m-1}(h) + \frac{c(k^m(h)) - c(k^{m-1}(h))}{\iota P_m(h)} P_m(h)$$

(Again, if $P_m(h) = 0$, we have $t^m(h) = t^{m-1}(h)$.)

Define $m^*(h)$ as the smallest value of $m \leq \ell(h)$ such that $r^m(h) = 0$ if such an m exists. If no such m exists, the game has not ended yet. In this case, we leave $m^*(h)$ undefined. For h such that $m^*(h)$ is defined, we can define a mapping from histories into outcomes by

$$O^*(h) = (k^{m^*(h)}, w - t^{m^*(h)}(h))$$

We can use the o^m maps to identify the set of feasible histories, H^*. For convenience, we will let H^* be the set of histories in H such that $t^m(h) \leq w$ for all m and such that $P_m(h) = \epsilon$ for all m such that $r^{m-1}(h) = 0$. In other words, agents need not consider histories which are either infeasible or for which contributions are no longer being solicited.[15] We are finally ready to define S^2. A strategy for agent i is a mapping, σ_i, from H^* into $\mathbf{R}_+$ which is feasible. Feasibility requires that

$$\sigma_i(h) + t_i^{\ell(h)}(h) \leq w_i$$

for all $h \in H^*$. Let S_i^2 be the set of such maps and, of course, let $S^2 = S_1^2 \times \ldots \times S_I^2$.

It is easy to define the history induced by a vector of strategies in S^2. For any $\sigma \in S^2$, recursively define the functions ϕ_σ^m for $m = 0, \ldots,$ by

$$\phi_\sigma^0(h) = h$$

and for $m \geq 1$,

$$\phi_\sigma^m(h) = \phi_\sigma^{m-1}(h) \cdot \sigma(\phi_\sigma^{m-1}(h))$$

where, for any $h \notin H^*$, $\sigma(h) = e$. In other words, $\phi_\sigma^0(h)$ is an identity function. Then $\phi_\sigma^1(h)$ gives the history h followed by the actions of the agents on the history h. Similarly, ϕ_σ^2 is the history $h \cdot \phi_\sigma^1(h)$ followed by the actions of the agents on this history and so on. Thus $\phi_\sigma^m(e)$ gives the first m actions of the agents given that they follow the strategies σ. Let $h^*(\sigma) = \phi_\sigma^\infty(e)$. This gives the entire history induced by σ. Notice that $h^*(\sigma)$ will necessarily have length less than or equal to $M + 1$. The outcome function, O^2, then, is defined by $O^2(\sigma) = O^*(h^*(\sigma))$. Let G^2 denote the game form (S^2, O^2).

Unfortunately, this game will not fully implement the core in UPE. First, eliminating dominated strategies once is not sufficient. To see the point, suppose that $d^*(\omega) = 2$ and $M = 4$. If 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

[15] The usefulness of this definition is that it guarantees that H^* is finite in our sequence of approximating games.

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, a strategy specifying such a contribution 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. Recall that successive removal of dominated strategies before applying perfection would not alter the set of equilibrium outcomes in G^1.

Even once we have done this, applying perfection to the resulting game is not sufficient. To see this, suppose we have $I = 10$. For agent 1, we have $v_1(1) = 1$, $v_1(2) = .9$, and $v_1(3) = .89$. For all other agents, $v_i(1) = 4.7/9$, $v_i(2) = 4.6/9$, and $v_i(3) = .01$. Suppose $M = 3$, $c(1) = 2.1$, $c(2) = 4.3$, and $c(3) = 6.6$. Then the strategy for agent 1 of contributing 2.1 in the first round cannot be eliminated by successive removal of dominated strategies. To see this, consider the strategies for the agents other than 1 of contributing nothing in the first round and .5 in the second round. The best reply for agent 1 is to contribute 2.1 in the first round. If he gives any less, his utility will be zero. However, if he gives 2.1, the other agents will contribute a total of 4.5 in the second round, bringing us up to three streetlights. The game will necessarily end there and agent 1's payoff will be $2.79 - 2.1 = .69$. Thus he should contribute 2.1. These strategies, of course, do not constitute a Nash equilibrium. However, none of them can be eliminated by successive removal of dominated strategies. If we try to apply perfection to the game after successive removal, then, this strategy for player 1 will still be possible. Hence we could consider having player 1 tremble to it with high probability to support a strategy by the other players of contributing 0 in the first round. It is easy to see that a slightly more complex example with two "large" agents each of whom have a high probability of trembling to $c(1)$ in the first round can support a perfect equilibrium where no one contributes anything in the first round. Thus we must consider a stronger equilibrium concept.

For these reasons, we are led to focus on successively undominated strictly perfect equilibria or 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 ω 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.

Since any agent can unilaterally deviate to a smaller contribution and lower both his payment and the social decision, if equilibrium outcome has $d = d^*(\omega)$, then no coalition of one could block it. It is not obvious that this implies that the outcome is in the core. However, suppose some coalition of two could block the equilibrium outcome so that it is not in the core. The coalition of two would choose some social decision smaller than $d^*(\omega)$ and pay for it themselves. Yet in equilibrium, either of them could reduce his contribution and cause the social decision to fall to no less than $d^*(\omega)$. The fact that they do not do so indicates that, in fact, the coalition could not block. Hence we see that if an equilibrium outcome is not in the core, it must be true that the social decision is strictly less than the efficient one.

So suppose that this is the case. Then there is some round at which contributions do not add up to enough to build an additional streetlight even though the valuations for that streetlight do sum to more than the cost. Strict perfection requires that strategies be robust with respect to *all* small probabilities of mistakes by other players. So to obtain a contradiction, we can choose any trembles we wish. It is not surprising that this freedom coupled with the fact that we can find Pareto dominating strategies means that we can always induce someone to change their strategy in such a situation. Thus no equilibrium with an outcome outside the core can exist.

Showing that each outcome in the core is achieved by some equilibrium is quite tedious, 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.

V. Approximating the Continuum: Limit Properties.

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

In this section, we will formalize this intuition and also demonstrate some surprising properties of our game at the limit where $\delta = 0$. As we will show, the set of equilibrium outcomes does not intersect the core and is identical to the set of equilibrium outcomes in the game analyzed by Bergstrom, Blume, and Varian (henceforth BBV). As we argue, our game at the limit is essentially a repeated version of their game and so this equivalence is not entirely surprising.

To begin, define $o_\delta^m(h) = (k_\delta^m(h), r_\delta^m(h), t_\delta^m(h))$ analogously to the above definition for $o^m(h)$. That is, $k_\delta^m(h)$ is defined to be the largest integer multiple of δ such that

$$\iota P_m(h) \geq c(k) - c(k_\delta^{m-1}(h))$$

if $r_\delta^{m-1}(h) = 1$ and $m \leq \ell(h)$. Similarly,

$$r_\delta^m(h) = \begin{cases} 1, & \text{if } k_\delta^m(h) \geq k_\delta^{m-1}(h) + \delta; \\ 0, & \text{otherwise.} \end{cases}$$

for h such that $r_\delta^{m-1}(h) = 1$ and $m \leq \ell(h)$. We can define $m_\delta^*(h)$ precisely as above and use this to define O_δ^* as above. We can use this to define the set of feasible histories, H_δ^*, analogously to the above and the strategy set S_δ. Finally, the functions generating the sequence of actions for a given set of strategies are identical to those defined above and the outcome function, O_δ is, again, the composition of O_δ^* with this function. Let $G_\delta = (S_\delta, O_\delta)$.

We can also define the analogous mappings for the game "at the limit." We will denote these by $o_0^m(h) = (k_0^m(h), r_o^m(h), t_0^m(h))$. These are defined recursively by defining $k_0^1(h)$ to be that value of k solving

$$\iota P_1(h) = c(k)$$

Then

$$t_0^1(h) = \iota P_1(h)$$

and

$$r_0^1(h) = \begin{cases} 1, & \text{if } k_0^1(h) > 0; \\ 0, & \text{otherwise.} \end{cases}$$

Then for $m \geq 1$ and h such that $\ell(h) \geq m$, if $r_0^{m-1}(h) = 0$, $o_0^m(h) = o_0^{m-1}(h)$. Otherwise, $k_0^m(h)$ is that value of k solving

$$\iota P_m(h) = c(k) - c(k_0^{m-1}(h))$$

$$t_0^m(h) = t_0^{m-1}(h) + \iota P_m(h)$$

and

$$r_0^m(h) = \begin{cases} 1, & \text{if } k_0^m(h) > k_0^{m-1}(h); \\ 0, & \text{otherwise.} \end{cases}$$

We can define $m_0^\cdot(h)$, $O_0^\cdot(h)$, $H_0^\cdot$, and S_0 exactly analogously to what we did before. Again, the outcome function O_0 is the composition of $O_0^\cdot$ with $h^\cdot$. Let $G_0 = (S_0, O_0)$.

To formalize the notion that the game form G_0 is the limit of the sequence of game forms $\{G_\delta\}_{\delta \in \Delta}$ for some sequence $\delta \in \Delta$ requires some additional terminology. Since G_δ is a pair, we must require convergence of both S_δ to S_0 and O_δ to O_0. The latter is obviously most simply defined in terms of pointwise convergence. However, the fact that O_δ and O_0 are functions from different spaces makes this more complex. Hence we will say that the sequence $\{O_\delta\}_{\delta \in \Delta}$ converges to O_0 if the sequence converges pointwise on $S(\Delta)$ where

$$S(\Delta) = \bigcap_{\delta \in \Delta \cup \{0\}} S_\delta$$

Recall that

$$S_\delta = \{\sigma : H_\delta^\cdot \to \mathbf{R}_-^I \mid \sigma(h) + t_\delta^{\ell(h)} \leq w, \ \forall h \in H_\delta^\cdot\}$$

Hence a natural way to define the convergence of the sets S_δ to S_0 is to require the domains to converge and to require a certain convergence of feasibility. We do this by the following definition. We will say that the sequence $\{S_\delta\}_{\delta \in \Delta}$ converges to S_0 iff each of the following holds:

(10) $\forall h \in H_0^\cdot$, there exists $\bar{\delta} > 0$ such that $\forall \delta \in \Delta$ with $\delta \leq \bar{\delta}$, $h \in H_\delta^\cdot$

(11) $\forall h \notin H_0^\cdot$, there exists $\bar{\delta} > 0$ such that $\forall \delta \in \Delta$ with $\delta \leq \bar{\delta}$, $h \notin H_\delta^\cdot$

(12) $\forall \sigma \in S_0$, there exists $\bar{\delta} > 0$ such that $\forall \delta \in \Delta$ with $\delta \leq \bar{\delta}$,

$$\sigma \mid H_0^\cdot \cap H_\delta^\cdot \in \{\sigma \mid \sigma = \sigma' \mid H_0^\cdot \cap H_\delta^\cdot, \text{ for some } \sigma' \in S_\delta\}$$

In (12), the notation $\sigma \mid A$ for any set A denotes the restriction of σ to A. As we will show in the proof of the lemma, $H_\delta^* \subseteq H_0^*$ for all δ sufficiently small. Hence we can rewrite (12) as

$$\forall \sigma \in S_0, \text{ there exists } \bar{\delta} > 0 \text{ such that } \forall \delta \in \Delta \text{ with } \delta \leq \bar{\delta},\ \sigma \mid H_\delta^* \in S_\delta$$

Now we are ready to state our lemma asserting convergence of G_δ to G_0.

Lemma. $G_\delta \rightarrow G_0$ *as* $\delta \downarrow 0$.

BBV 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.

Yet this is precisely G_0. When $\delta = 0$, it is impossible for contributions to fall strictly between $c(k)$ and $c(k+\delta)$. 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.

Thus in this sense, our game has a repeated version of the BBV game as its limit. As noted above, it is not difficult to prove an analogue of Theorem 2. To state this analogue, let $\mathcal{E}_\delta$ denote the set of economies described above with social decision set D_δ. (To maintain our assumptions on endowments, we will assume that $\sum_i w_i \geq c(M)$ throughout.) Let $C_\delta(\omega)$ be the core of the economy at state ω. Then it is straightforward to see that the game form G_δ fully implements the core of $\mathcal{E}_\delta$ in SUSPE for all $\delta > 0$. Let $\mathcal{E}_0$ denote the set of economies with decision set D_0 and let $C_0(\omega)$ denote the core of $\mathcal{E}_0$ at state ω. Since it is straightforward to show that $C_\delta(\omega)$ converges to $C_0(\omega)$ as $\delta \downarrow 0$,[16] we see that for all ω, the set of SUSPE outcomes of G_δ converges to $C_0(\omega)$ as $\delta \downarrow 0$. Putting it differently,

Theorem 3. *The sequence of game forms G_δ approximately fully implements the core of $\mathcal{E}_0$ in successively undominated strictly perfect equilibrium.*

As mentioned, the set of equilibria at the limit is rather different. In general, as we will show

[16] This convergence is defined identically to the convergence of the sets H_δ^* to H_0^*.

below, there are no SUSPE's. As we will see, this is, in part, due to the approximation technique. We characterize the set of Nash equilibria that are limit points of Nash equilibria of approximating games and show that the set of outcomes so defined does not intersect the core. We conclude by offering some intuition for these results.

It is useful to first describe the set of Nash equilibria in the one–shot version of the BBV game. To do so, let $\hat{d}_i(\omega)$ solve

$$\max_d U_i(d \mid \omega) - c(d)$$

and let $\bar{d}(\omega) = \max_{i \in I} \hat{d}_i(\omega)$. Let $I_1(\omega)$ be the set of i such that $\hat{d}_i(\omega) = \bar{d}(\omega)$ and let $I_2(\omega)$ denote $I \setminus I_1(\omega)$. Then we can define

$$\Theta_1(\omega) = \{(d, x) \in \Theta \mid d = \bar{d}(\omega), \sum_{i \in I_1(\omega)} x_i = \sum_{i \in I_1(\omega)} w_i - c(d),$$
$$x_i \in [0, w_i] \; \forall i, \text{ and } x_j = w_j, \; \forall j \in I_2(\omega)\}$$

In other words, $\Theta_1(\omega)$ is the set of outcomes where d is equal to the largest number anyone would buy for themselves if they could afford it. Those who want this maximum amount pay for it in any way that does not require any agent to pay more than his wealth. The remaining agents pay nothing.[17]

In any Nash equilibrium in which the game ends after a finite number of rounds, the outcome must have $d = \bar{d}(\omega)$. To see this, note that if the game is going to end with a smaller amount of the public good, any individual desiring $\bar{d}(\omega)$ would contribute whatever was necessary to achieve this amount. If the game would end with a larger amount, then, in the last round of contributions, there must be some individual who would prefer to reduce his contribution. It is not difficult to use this to show that the set of Nash equilibrium outcomes of the one–shot BBV game is $\Theta_1(\omega)$.

The above reasoning does not apply to equilibria where contributions are given forever. Notice, though, that in any approximating game, there is some minimum strictly positive contribution. Hence in any approximating game, there must be only a finite number of rounds of contributions. Therefore, any equilibrium outcome of an approximating game must have $d = \bar{d}(\omega)$, implying that the equilibrium outcomes of G_0 must as well. It is not true in general, though, that the only outcomes that can be achieved by Nash equilibria with a finite number of rounds are outcomes in $\Theta_1(\omega)$. A counterexample is given in the Appendix along with the proof of Theorem 4. We will refer

[17] The discussion in the text ignores the possibility that wealth constraints are binding. It is straightforward to show that $w_i \geq U_i(M \mid \omega)$ for all ω implies that wealth constraints will not bind.

to a Nash equilibrium which is the limit point of a sequence of Nash equilibria of approximating games as an approachable equilibrium.

Theorem 4. *For each ω, every approachable equilibrium outcome of G_0 has $d = \bar{d}(\omega)$. Furthermore, $\Theta_1(\omega)$ is a subset of the set of the approachable equilibrium outcomes at ω.*

It is not difficult to show that the set of approachable outcomes of G_0 does not intersect $C_0(\omega)$. To see this intuitively, suppose that $c(d) = d$ and that the utility functions are differentiable. Then $\hat{d}_i(\omega)$ is where agent i's marginal utility is 1. Hence the social decision in an approachable outcome of G_0 is necessarily where some agent's marginal utility is 1 (and the others' marginal utilities are all strictly positive) while the efficient d is where the sum of the agent's marginal utilities is 1. This fact, then, indicates a strong difference between the limiting game and the game at the limit.

This result would suggest that the SUSPE outcomes are disjoint from the core. However, this is not true because the set of SUSPE's is generally empty. To see this, note that in an approximating game, there must be a finite number of rounds in any Nash equilibrium and the outcome must have $d = \bar{d}(\omega)$. Consider the last stage at which contributions are received in such an equilibrium and suppose some person deviates downward in his contribution. As long as positive contributions are received, the game continues. Obviously, for the original strategies to constitute an equilibrium, it must be true that no one makes up the shortfall in contributions for the deviator, as otherwise he would certainly desire to deviate in this fashion. Hence either the shortfall is not made up or the deviator himself makes it up. It is not hard to see that subgame perfection (which is, of course, implied by SUSPE) requires the latter because any contributor in the last round must be some i such that $\hat{d}_i(\omega) = \bar{d}(\omega)$.

It is also not hard to see that the successive elimination of dominated strategies does not pin down very precisely the strategies that may be employed by the deviator at the subsequent history. Thus the deviator may make up the shortfall all in one round or over several rounds and he is certainly indifferent as to how he does so. So consider his optimal strategy in the face of some completely mixed strategies for the other agents. Suppose these completely mixed strategies have a high probability that the deviator himself will "accidentally" give zero tomorrow even if contributions are not yet $c(\bar{d})$ and (comparatively) almost no probability of any other "mistake." It is not hard to see that this implies that the best strategy for the deviator at this history is to make up the entire shortfall at once. However, consider instead completely mixed strategies where the most likely mistake is that some player other than the deviator accidentally gives some money this period and any other mistake is almost impossible by comparison. Then the best strategy for the deviator is to not make up the entire shortfall this period. In short, there is no strategy on this history for which the deviator has a best reply to ***any*** vector of completely mixed strategies close to the equilibrium strategies. Hence there is no SUSPE.

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. The reason is that there are Nash equilibria with strong robustness properties which are not approachable but do have outcomes in the core. For example, consider some outcome Pareto preferred to some outcome in $\Theta_1(\omega)$. For each agent, divide the amount of wealth he is to contribute beyond the amount at the outcome in $\Theta_1(\omega)$ into an infinite sequence. Construct the equilibrium strategies by supposing that all agents contribute to reach the outcome in $\Theta_1(\omega)$ first and then agents alternate with agent i at his n^{th} "turn" contributing 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 there is always some amount of contributions being given in the future means that any agent who cuts his contribution will discontinously 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. In the Appendix, we show that if $I = 2$, there are very robust Nash equilibria[18] which have outcomes in the core.

To understand these results intuitively, suppose $c(d) = d$ and that $U_i(d \mid \omega)$ is differentiable for all v. Note that an approachable equilibrium of G_δ for any $\delta > 0$ can have exactly δ be contributed at each round up to the efficient d. For any positive δ, this process allows us to (eventually) hit the efficient d. However, if $\delta = 0$, we can no longer take this sequence of minimal size steps up to the efficient d. We must necessarily step from one level to another with physically possible outcomes in between. Notice that at the efficient d, it will necessarily be true that each agent has $U_i'(d \mid \omega) < 1$ as the sum of the MRS's must equal 1. Recall that the equilibria must have a finite number of rounds. Therefore, if there is some proposed core equilibrium, any agent can reduce his contribution at this last round by, say, k, at worst cause the level of the public good to fall to the efficient level minus k, and be strictly better off. Thus we cannot reach the efficient outcome. By contrast, when $\delta > 0$, along the equilibrium path, agents cannot decrease their contributions by ***any*** nonzero amount without a discontinuous effect on the public good. The fact that $U_i'(d \mid \omega) < 1$ does not imply that they can reduce their contribution and be better off because a reduction leads to a more than one–for–one decrease in the level of the public good. Hence, each round is essentially

[18] To be more precise, we construct strong Nash equilibria (in the sense of van Damme [1983]) in the agent–normal form.

a unanimity vote between the status quo and some Pareto preferred outcome. Naturally the only robust equilibrium has all agents voting for the Pareto preferred option.

The effect of the approachability requirement is most easily seen by fixing the approximation so that contributions are in multiples of $1/n$ and asking how the outcome is affected by taking δ toward zero. When δ falls below $1/n$, 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. Hence we cannot reach the core with $\delta \leq 1/n$. Clearly, for any fixed $\delta > 0$, we can make n large 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 indivisibilites in the private good, then core outcomes are achieved by this contributions 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. For example, this role is noted by Benassy [1985] and explored in some depth by Aghion [1985] in the context of market games. While this role may seem surprising at first glance, this is the same role discontinuity plays in the efficiency of perfect competition.[19] 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, in the game G_0, the continuity of the outcome function with respect to contributions allows agents to shade their contributions a small amount without consequences as disastrous for them as in G_ϵ.

[19] We are grateful to Andreu Mas-Colell for pointing out this analogy to us.

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 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.

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 implementing 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? It is clear that knowing that a game form fully implements a choice correspondence in one equilibrium notion does not tell us that it will fully implement in another. Furthermore, in general, one needs to know more than which equilibrium notion is the stronger to make the determination. Thus 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, we define the sequence of approximating games. For G^1, the "maps" that comprise the strategy sets are trivial ones, so that approximating S_i is the same as approximating A_i. Thus we can approximate S_i by any set of the form

$$\{\underline{a}^i(n), \underline{a}^i(n) + \delta_n, \underline{a}^i(n) + 2\delta_n, \ldots, w_i(n)\}$$

where $\underline{a}^i(n) \leq \delta_n$, $w_i(n)$ is the largest number less than or equal to w_i that can be written in the form $\underline{a}^i(n) + k\delta_n$ for some integer k, and

$$\lim_{n \to \infty} \delta_n = 0$$

To begin the proof, then, 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.[20] The limit of any such tuple must have a sum strictly less than c. Thus we see that for such ω, 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 ω with an outcome of $(0, w)$, one of the points in the core.

The other point in the core, $(1, w - v(\omega))$, is a bit trickier. To see how to deal with this case, simply note that in the reduced game, one perfect equilibrium has each agent choosing the largest contribution in $R^1(S_i(n))$. These elements must sum to less than c for any n. However, as $n \to \infty$, 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 ω.

[20] For simplicity, our notation does not reflect the fact that which strategies are dominated depends on ω.

The last case, when $\sum_i v_i > c$, is more complex. First, we will show that no UPE outcome can have $d = 0$. Then we will verify that the all points in the core are UPE outcomes. To begin the first task, note that a Nash equilibrium in pure strategies cannot have the contributions add to strictly more than c as any contributor would then prefer a smaller contribution. Hence all UPE's certainly have $\sum_i \sigma_i \leq c$. Thus we only need to eliminate the possibility that contributions are strictly less than c.

So suppose that this does occur in a UPE. This requires that for some large n, we have a UPE, $\sigma^*(n) = (\sigma_1^*(n), \ldots, \sigma_I^*(n))$, where $\sum_i \sigma_i^*(n) < c$ even though $\sum_i v_i > c$. For each i, define $v_i(n)$ as the largest element of $S_i(n)$ strictly less than v_i. Since we have eliminated dominated strategies, this will be the largest element of $S_i(n)$. 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[21]

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

Notice that we must have $\sigma_1' < v_1$. To see this, note that it certainly holds if

$$\sigma_1^*(n) + v_2(n) - \sigma_2^*(n) \leq v_1(n)$$

or

$$v_1(n) - \sigma_1^*(n) \geq v_2(n) - \sigma_2^*(n)$$

which is implied by the way we have chosen to number the players. Since this is a UPE, we must also have

$$V_1(\sigma_1^*(n), s_{\sim 1}^{\epsilon *}(n)) \geq V_1(\sigma_1', s_{\sim 1}^{\epsilon *}(n))$$

where $s^{\epsilon *}(n)$ is a sequence of completely mixed strategies in the reduced game converging to $\sigma^*(n)$ as $\epsilon \downarrow 0$.[22] We can rewrite this equation as

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

where the notation "$\Pr^\epsilon$" indicates that the probability is calculated given the distribution induced by $s^{\epsilon *}(n)$. Rearranging yields:

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

21 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.

22 For convenience, we suppress the dependence of the utility functions and the game on the state ω.

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

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

(For notational simplicity, we will not denote the fact that this correspondence depends on n. Notice, though, that it does not depend on ϵ.) Thus $A(\sigma_1)$ is the set of contributions for the other agents such that $d = 1$ given that player 1 contributes σ_1. Notice that

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

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

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

where $\xi^\epsilon(\hat{\sigma}_{\sim 1}) \downarrow 0$ as $\epsilon \downarrow 0$. To see this, note that any $\hat{\sigma}_{\sim 1} \in A^*$ must contain some components which differ from the corresponding component of $\sigma_{\sim 1}^*(n)$. For each $\sigma_{\sim 1} \in A^*$, choose any i such that $\hat{\sigma}_i \neq \sigma_i^*(n)$ and construct $\sigma_{\sim 1}'$ by replacing $\hat{\sigma}_i$ with $\sigma_i^*(n)$. Note that replacing $\hat{\sigma}_{\sim 1}$ with $\sigma_{\sim 1}'$ 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 $\hat{\sigma}_i \leq v_i(n)$. 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 σ_1' guarantees that total contributions are still at least c. Note also that

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

Let

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

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

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

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

The fact that A^* is a finite set implies that ξ^ϵ exists and that $\xi^\epsilon \downarrow 0$ as $\epsilon \downarrow 0$.

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

$$\xi^{\epsilon} \; \frac{\sum_{\hat{\sigma}_{\sim 1} \in A^{*}} \Pr^{\epsilon}[\sigma_{\sim 1} = g(\hat{\sigma}_{\sim 1})]}{\sum_{\sigma'_{\sim 1} \in A'} \Pr^{\epsilon}[\sigma_{\sim 1} = \sigma'_{\sim 1}]} \geq \frac{u_1(1, w_1 - \sigma'_1)}{u_1(1, w_1 - \sigma^{*}_1(n))} \tag{A.2}$$

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 g^{*} be the vector that $g(\hat{\sigma}_{\sim 1})$ maps into most often for $\hat{\sigma}_{\sim 1} \in A^{*}$ and let

$$\ell = \#\{\hat{\sigma}_{\sim 1} \in A^{*} \mid g(\hat{\sigma}_{\sim 1}) = g^{*}\}$$

(Notice that ℓ is implicitly a function of n and ϵ. However, for any n, it is necessarily finite for all ϵ.) 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

$$\ell \xi^{\epsilon} \geq \frac{u_1(1, w_1 - \sigma'_1)}{u_1(1, w_1 - \sigma^{*}_1(n))}. \tag{A.3}$$

Note that the right–hand side is strictly positive as $\sigma'_1 < v_1$ and is independent of ϵ. Hence if we choose ϵ sufficiently small, we contradict (A.3). 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 Γ with $d = 0$.

Now we only need to establish that for any ω 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 σ, we can always find a sequence of approximating games with $\sigma \in S(n)$ for all n. To see this, simply choose $\delta_n = x/n$ for some x and let $\underline{a}^i(n) = \sigma_i - k\delta_n$ for the largest integer k such that this is positive. For any such σ and choice of $S(n)$, σ is a strong equilibrium of $\Gamma(n)$ (see van Damme [1983]) and hence is a UPE. Since this is true for all n, σ is a UPE of Γ.

Now suppose that $\sigma < v$, but that $\sigma_i = v_i$ for some i. Let B denote the set of $i \in I$ such that $\sigma_i = v_i$ and let $D = I \setminus B$. Choose any allowable approximation for S_i for $i \in B$. We wish to construct a sequence of equilibria in which each of these agents gives $v_i(n)$. Let $t(n) = \sum_{i \in B}(v_i - v_i(n))$. Let k be the number of agents in D and for $i \in D$, let $\sigma_i(n) = \sigma_i + t(n)/k$. For n sufficiently large, this will necessarily have $\sigma_i(n) < w_i$ for all $i \in D$. For $i \in D$, construct $S_i(n)$ so that $\sigma_i(n) \in S_i(n)$ for all n. (This is easily done by adapting the construction above.) For $i \in B$, let $\sigma_i(n) = v_i(n)$. By definition, then,

$$\sum_i \sigma_i(n) = \sum_i \sigma_i = c$$

for all n. It is easy to see that $\sigma(n)$ is a strong equilibrium of $\Gamma(n)$ for all n and hence is a UPE. Since $\sigma(n) \to \sigma$, we see that σ is a UPE of Γ. ▌

Proof of Theorem 2

We prove the theorem through a series of lemmas. Most of the lemmas apply directly to the game induced by G^2 at state ω as well as to any approximating game. When the distinction between the game and approximating game is unnecessary, we will avoid the extra notation involved in describing the lemma for the latter. The lemmas will show that if (d, x) is a SUSPE outcome at state ω, then $(d, x) \in C(\omega)$. They also demonstrate some useful facts which will enable us to prove the converse by construction. A useful definition is the following. We will say that h' is a subhistory of h if there exists a history h'' such that $h' \cdot h'' = h$.

Lemma 2.1. *If (d, x) is a SUSPE outcome at state ω, then $d \leq d^*(\omega)$.*

Proof: Since the SUSPE strategy tuples are a subset of the Nash, it is clearly sufficient to demonstrate the result for Nash outcomes. So suppose we have a Nash equilibrium σ at a state ω with $O^2(\sigma) = (d, x)$ where $d > d^*(\omega)$. Let $h = h^*(\sigma)$—that is, h gives the equilibrium path. Let h' be the subhistory consisting of the first $m^*(h) - 2$ rounds. Thus h' is the history leading up to the last round of contributions that added up to enough to alter the status quo. Let $k^{\ell(h')}(h') = d' < d$. Then we must have

$$\sum_i \sigma_i(h') = c(d) - c(d')$$

To see why we must have equality (as opposed to a weak inequality), note that if contributions add to more than $c(d) - c(d')$, each agent will receive a refund which is some fraction of this excess. Clearly, cutting one's contribution by the amount of this excess instead gives one the entire amount rather than a fraction. Since no further contributions were to be made, the deviator can not contribute anything more subsequently and thus can avoid being "punished" for this deviation.

The fact that $d > d^*(\omega)$ implies

$$\sum_i v_i(d \mid \omega) < c(d) - c(d-1)$$

so that

$$\sum_i v_i(d \mid \omega) < c(d) - c(d')$$

Hence there must be some i such that $\sigma_i(h') > v_i(d \mid \omega)$.

Choose any such i and consider the strategy $\tilde{\sigma}_i$ constructed as follows. For any history h'' which is not equal to h' nor has h' as a subhistory, $\tilde{\sigma}_i(h'') = \sigma_i(h'')$. For any history h'' other than h' with h' as a subhistory, $\tilde{\sigma}_i(h'') = 0$. Finally,

$$\tilde{\sigma}_i(h') = \sigma_i(h') - v_i(d \mid \omega) - \epsilon$$

for some small $\epsilon > 0$. Clearly,

$$\sum_{j \neq i} \sigma_j(h') + \tilde{\sigma}_i(h') = \sum_j \sigma_j(h') - v_i(d \mid \omega) - \epsilon$$

Since $d > d^*(\omega)$, we must have

$$c(d) - c(d-1) > \sum_j v_j(d \mid \omega) > v_i(d \mid \omega)$$

so that for ϵ sufficiently small,

$$\sum_{j \neq i} \sigma_j(h') + \tilde{\sigma}_i(h') \geq c(d-1) - c(d')$$

Hence if i switches to this strategy, the outcome will have at least $d-1$ as the social decision and i will pay at least $v_i(d \mid \omega) + \epsilon$ less than he would have. It is possible that he will receive a refund and thus ends up paying still less than this. Therefore, i's expected payoff with $\tilde{\sigma}_i$ minus his expected payoff with σ_i is at least $\epsilon > 0$. Hence this could not be a Nash equilibrium. ▌

The next lemma shows that we can focus attention on equilibrium social decision. Let $k(h) = k^{\ell(h)}(h)$, $t(h) = t^{\ell(h)}(h)$, and $r(h) = r^{\ell(h)}(h)$.

Lemma 2.2. *Suppose (d', x') is a SUSPE outcome at state ω. If $d' = d^*(\omega)$ or if*

(A.4) $$\sum_i v_i(d^*(\omega) \mid \omega) = c(d^*(\omega)) - c(d^*(\omega) - 1)$$

and $d' = d^*(\omega) - 1$, then $(d', x') \in C(\omega)$.

Proof: Suppose not. Then the Proposition stated in the text implies that there is no price system which supports the equilibrium outcome at state ω. In other words, for any nonnegative set of numbers, $p_i(d)$ for $i \in I$ and $d \in D$, one of the following fails to hold:

(A.5) $$\sum_i p_i(d') = c(d')$$

(A.6) $$(d', x_i') \text{ maximizes } u_i(d, x_i \mid \omega) \text{ on } \{(d, x_i) \mid p_i(d) + x_i = w_i\}$$

(A.7) $$d' \text{ maximizes } \sum_i p_i(d) - c(d) \text{ on } D$$

Let h be the history generated by the equilibrium strategies and let $\mathcal{H}$ be the set of histories that are subhistories of h. We will say that d is reached in equilibrium if there exists $h' \in \mathcal{H}$ such that $k(h') = d$. For each i and each d that is reached in equilibrium, let

$$p_i(d) = t_i(h'(d))$$

where $h'(d)$ is the history in $\mathcal{H}$ such that $k(h') = d$. For any $d < d'$ which is not reached in equilibrium, any player could have caused d to be reached by some deviation. In particular, for any $d < d'$ such that d is not reached, let $h'(d)$ denote the longest history in $\mathcal{H}$ such that $k(h) < d$. Then for all i and each such d, let

$$p_i(d) = t_i(h'(d)) + \max[c(d) - \sum_j t_j(h'(d)) - \sum_{j \neq i} \sigma_j(h'(d)),\ 0]$$

Finally, for $d > d'$, let

$$p_i(d) = t_i(h) + \sum_{m=d'+1}^{d} v_i(m \mid \omega)$$

Obviously, all of these numbers are nonnegative. Furthermore,

$$\sum_i p_i(d') = \sum_i t_i(h) = c(d')$$

so that (A.5) holds. Similarly, by construction, for any d which is reached in equilibrium,

$$\sum_i p_i(d) = c(d)$$

For any $d < d'$ such that d is not reached,

$$\sum_i p_i(d) = c\left(k\left(h'(d)\right)\right) + \sum_i \max[c(d) - \sum_j t_j(h'(d)) - \sum_{j \neq i} \sigma_j(h'(d)),\ 0]$$

Let $k = k(h'(d))$ and let I^+ denote the set of agents for whom the maximum is the first term. Then we have

$$\sum_i p_i(d) = c(k) + \#I^+[c(d) - c(k)] - \sum_{i \in I^+} \sum_{j \neq i} \sigma_j(h'(d))$$

But

$$\sum_{i\in I^+}\sum_{j\neq i}\sigma_j(h'(d)) = \#I^+ \sum_{j\notin I^+}\sigma_j(h'(d)) + (\#I^+ - 1)\sum_{j\in I^+}\sigma_j(h'(d))$$

or

$$\sum_{i\in I^+}\sum_{j\neq i}\sigma_j(h'(d)) = \#I^+ \sum_{j}\sigma_j(h'(d)) - \sum_{j\in I^+}\sigma_j(h'(d))$$

Hence

$$\sum_{i\in I^+}\sum_{j\neq i}\sigma_j(h'(d)) \geq (\#I^+ - 1)\sum_{j}\sigma_j(h'(d))$$

By assumption, though,

$$c(k) + \sum_{j}\sigma_j(h'(d)) > c(d)$$

so

$$\sum_{i\in I^+}\sum_{j\neq i}\sigma_j(h'(d)) > (\#I^+ - 1)(c(d) - c(k))$$

Substituting, we see that

$$\sum_{i} p_i(d) < c(d)$$

For $d > d^*$, the fact that

$$\sum_{i} v_i(m \mid \omega) < c(m) - c(m-1)$$

for all $m > d^*(\omega)$ immediately implies

$$\sum_{i} p_i(d) \leq c(d)$$

Hence (A.7) is satisfied. Thus if we can show that (A.6) must be satisfied, we are done.

To see this, notice first that no agent would prefer any $d > d^*$ at these "prices." These prices have each agent paying for extra streetlights exactly what they are worth to him and hence he cannot strictly prefer $d > d^*$. Furthermore, no agent could strictly prefer any $d < d^*$ such that d is reached in equilibrium. The price he pays for such a social decision is precisely his total contributions up to the point where d is reached. If he strictly preferred staying at this point to following the equilibrium path, he could have simply chosen $\sigma_i(h) = 0$ on all subsequent histories. Since strictly preferring d must imply that this strategy is different from his equilibrium strategy, we see that he cannot strictly prefer any such d. Similarly, he could not prefer $d < d^*$ such that d is not reached in equilibrium. The fact that $d < d^*$ implies that, at some point, total contributions "hop over" $c(d)$. Hence, at this point, he could cut his contribution and either guarantee that d is reached or that some larger amount is reached. He could then refuse to contribute any more than this. It is easy to see that his price for this d guarantees that this choice yields utility at least as large as $u_i(d, w_i - p_i(d) \mid \omega)$ and hence the fact that he does not deviate to this choice implies that he prefers d'. Hence (A.6) holds. Therefore, $(d', x') \in C(\omega)$. ∎

Thus we have shown that if (d, x) is a SUSPE outcome at ω, then, either $(d, x) \in C(\omega)$ or the social decision is strictly smaller than the Pareto efficient decision. Lemma 2.5 will show that the latter cannot be the case, thus establishing that any SUSPE outcome is in the core. The proof of this lemma, however, requires some information about strategies which are removed by successive elimination of dominated strategies. The following lemmas provide this information. First, we define some new terminology. We will say that a history $h \in H^*$ is reached with $R^*(S)$ if there is a $\sigma \in R^*(S)$ such that $\phi_\sigma^{\ell(h)}(e) = h$. That is, there is a $\sigma \in R^*(S)$ such that h is a history generated by σ. We will say that a history $h \in H^*$ can be reached from $h' \in H^*$ with $R^*(S)$ if h' is a subhistory of h and there is $\sigma \in R^*(S)$ such that $\phi_\sigma^{\ell(h)}(e) = h$ and $\phi_\sigma^{\ell(h')}(e) = h'$. That is, both h and h' are generated by σ. Finally, we will say that the outcome can be changed from $h \in H^*$ with $R^*(S)$ if there is a history h' that can be reached from h with $R^*(S)$ such that $k(h') > k(h)$.

Lemma 2.3. *Consider any ω and $h \in H^*$ such that,*

$$(a) \qquad \sum_i \min[w_i - t_i(h),\ v_i(k(h) + 1 \mid \omega)] < c(k(h) + 1) - c(k(h))$$

and, (b), h can be reached with $R^*(S)$. Then the outcome cannot be changed from h with $R^*(S)$.

Proof: First, we offer some further explanation of terminology. Recall that we have defined SUSPE's relative to the agent–normal form. This means that the "players" we consider are ih pairs, not simply i's. Thus we will say a strategy for ih of σ_{ih} is dominated and hence is not in $R^*(S_{ih})$.

The proof is by induction on d. So suppose that we have ω and h satisfying (a) and (b) where $k(h) = M - 1$. Then

$$\sum_i \min[w_i - t_i(h),\ v_i(M \mid \omega)] < c(M) - c(M - 1)$$

Clearly, for any ih such that $k(h) = M - 1$, any σ_{ih} such that

$$\sigma_{ih} > v_i(M \mid \omega)$$

is not in $R^1(S_{ih})$ and hence is not in $R^*(S_{ih})$. Obviously, any σ_{ih} such that

$$\sigma_{ih} > w_i - t_i(h)$$

is not feasible and hence is also not in $R^*(S_{ih})$. Therefore, for any $\sigma \in R^*(S)$, it must be true that

$$\sum_i \sigma_{ih} < c(M) - c(M-1)$$

so that the outcome cannot be changed from h with $R^*(S)$.

To complete the proof, suppose that for some $d' \geq 2$, we have shown the following induction hypothesis for $d \geq d'$:

(H) If h can be reached with $R^*(S)$, $k(h) = d$, and

$$\sum_i \min[w_i - t_i(h),\ v_i(d+1 \mid \omega)] < c(d+1) - c(d), \tag{A.8}$$

then the outcome cannot be changed from h with $R^*(S)$.

We wish to show that (H) holding for all $d \geq d'$ implies (H) for $d = d' - 1$. This will complete the proof. So suppose (H) does not hold for $d = d' - 1$. Then there is an $h' \in H^*$ such that h' can be reached from h with $R^*(S)$ with $k(h') \geq d'$. We will derive a contradiction to this by induction.

To begin this induction, notice that we cannot have $k(h') = M$. This would require

$$\sigma_{ih} > \sum_{m=d'}^{M} v_i(m \mid \omega)$$

for some ih. To see this, note that (A.8) implies

$$\sum_{m=d+1}^{M} \sum_i \min[w_i - t_i(h),\ v_i(m \mid \omega)] < c(M) - c(d)$$

because $v_i(m \mid \omega)$ is strictly decreasing in m and $c(m+1) - c(m)$ is strictly increasing in m. But the left-hand side is weakly larger[23] than

$$\sum_i \min\left[w_i - t_i(h),\ \sum_{m=d}^{M} v_i(m \mid \omega)\right]$$

so that

$$\sum_i \min\left[w_i - t_i(h),\ \sum_{m=d'}^{M} v_i(m \mid \omega)\right] < c(M) - c(d')$$

[23] Recall that all terms in the minimum are positive. It is not difficult to show that for any a, b, and c, all positive, $\min(a, b) + \min(a, c) \geq \min(a, b + c)$.

Obviously, any such σ_{ih} is dominated so $\sigma_{ih} \notin R^*(S_{ih})$. This provides us with the basis for an induction on $k(h')$. So suppose that we have shown that we cannot have $k(h') \geq m'$ for some $m' \geq d' + 1$. We wish to show that this implies that we cannot have $k(h') = m' - 1$, which will provide the contradiction needed to complete the proof.

First, notice that the fact that $k(h')$ cannot be greater than or equal to m' and the induction hypothesis (H) for $d \geq d'$ implies that any σ_{ih} such that

$$\sigma_{ih} > \sum_{m=d'}^{m'-1} v_i(m \mid \omega)$$

cannot be in $R^*(S)$. This is because it is impossible to increase d beyond $m' - 1$ and so no contribution larger than the right-hand side could possibly be worthwhile. But then

$$\sigma_{ih} \leq \sum_{m=d'}^{m'-1} v_i(m \mid \omega)$$

for all $\sigma_{ih} \in R^*(S_{ih})$. Together with the fact that feasibility requires $\sigma_{ih} \leq w_i - t_i(h)$, this implies

$$\sum_i \sigma_{ih} \leq \sum_i \min \left[w_i - t_i, \sum_{m=d'}^{m'-1} v_i(m \mid \omega) \right]$$

Hence (A.8) coupled with an analogous argument to the above implies

$$\sum_i \sigma_{ih} < c(m' - 1) - c(d')$$

so that we cannot have $k(h') = m' - 1$. ∎

The next lemma uses this one to show what some strategies in $R^*(S)$ must be. To define these strategies, let

$$A(h) = \{a \in S_h \mid a_i < v_i(k(h) + 1 \mid \omega) \text{ and } \sum_i a_i = c(k(h) + 1) - c(k(h))\}$$

$$S_{ih}^* = \{a_i \in S_{ih} \mid \exists a_{\sim i} \text{ with } (a_i, a_{\sim i}) \in A(h)\}$$

$$\tilde{H} = \{h \in H^* \mid t_i(h) \leq \sum_{d=1}^{k(h)} v_i(d \mid \omega) \text{ and } h \text{ can be reached with } R^*(S)\}$$

Then we have

Lemma 2.4. *For all i, all ω, and all $h \in \tilde{H}$,*

$$S_{ih}^* \subseteq R^*(S_{ih}).$$

Proof: If $A(h) = \emptyset$, then $S^*_{ih} = \emptyset$ and so it is trivially true that $S^*_{ih} \subseteq R^*(S_{ih})$. Let

$$\hat{H} = \{h \in \tilde{H} \mid A(h) \neq \emptyset\}$$

Let $\hat{d}(\omega) = d^*(\omega) - 2$ if (A.4) holds and let it equal $d^*(\omega) - 1$ otherwise. Thus $\hat{d}(\omega)$ is one smaller than the smallest possible Pareto efficient decision. For h such that $k(h) > \hat{d}(\omega)$, we must have $A(h) = \emptyset$. Hence $h \in \hat{H}$ implies $k(h) \leq \hat{d}(\omega)$.

The proof is by induction on $k(h)$. To begin, consider any $h \in \hat{H}$ such that $k(h) = \hat{d}(\omega)$. Suppose first that $\hat{d}(\omega) = d^*(\omega) - 1$. Consider the following strategy tuple. For all ih' such that $h' \neq h$ and h' is a subhistory of h, choose strategies in $R^*(S_{ih'})$ which lead to h. Since $h \in \hat{H}$, such strategies must exist. For all ih' such that $h' \neq h$ and h is a subhistory of h', choose any strategy in $R^*(S_{ih'})$. By Lemma 2.3, these strategies must not have the contributions adding up to enough to change the outcome. For all ih' such that $h' \neq h$ and neither history is a subhistory of the other, choose any strategy in $R^*(S_{ih'})$. Finally, for history h, choose σ_{ih} so that the vector $\sigma_h \in A(h)$. The fact that $h \in \tilde{H}$ together with our assumption that $w_i > U_i(M \mid \omega)$ implies that these strategies are feasible. Notice that for each i, the strategy σ_{ih} is the unique best reply for ih to the strategies of $\sim (ih)$. This is true because any reduction in the contribution by ih causes contributions not to add up at h and hence i is worse off. Similarly, since contributions will not sum up at the next stage, there is no possible advantage to increasing one's contribution since refunds are only a fraction of the extra contributed.[24] Hence for each ih, $\sigma_{ih} \in R^1(S_{ih})$ as $\sigma_{\sim(ih)} \in S_{\sim(ih)}$. But then for all ih, $\sigma_{ih} \in R^2(S_{ih})$ as $\sigma_{\sim(ih)} \in R^1(S_{\sim(ih)})$ and so on. Therefore, $\sigma_{ih} \in R^*(S_{ih})$ for all ih. Hence if $\hat{d}(\omega) = d^*(\omega) - 1$, then for all h such that $k(h) = \hat{d}(\omega)$, we have the desired result.

Suppose instead that $\hat{d}(\omega) = d^*(\omega) - 2$. Again, consider any $h \in \hat{H}$ such that $k(h) = \hat{d}(\omega)$. Notice that for any h' such that $k(h') = d^*(\omega) - 1$, it is certainly dominated to give more than $v_i(d^*(\omega) \mid \omega)$. Hence the only way contributions at h' could possibly sum to $c(d^*(\omega)) - c(d^*(\omega) - 1)$ is for each agent to receive a zero increment to their payoff at this stage. Hence the issue of whether or not contributions add up at h' is irrelevant to payoffs at the preceding stage. In other words, we can choose $\sigma_{ih'}$ for each ih' exactly as above and the same argument will apply. Hence we see that for any $h \in \hat{H}$ such that $k(h) = \hat{d}(\omega)$, we have $S^*_{ih} \subseteq R^*(S_{ih})$.

This provides us with the basis for the induction. So consider the following induction hypoth-

[24] To be more precise, unless ih is the only contributor at h, his refund will be a fraction of the excess so that he would be strictly worse off increasing his contribution. If ih is the only contributor at h, then he is indifferent between this contribution and any larger one leading to the same outcome. However, the smaller contribution cannot be dominated by a larger one. To see this, simply note that the two contributions do have equal payoffs against these strategies. However, the smaller one is a better reply against any positive contributions by the other players at this round. In other words, comparing this contribution and any other, we see that this one is either strictly better against these strategies or is strictly better against some other strategies. This fact is sufficient for the subsequent arguments.

esis:

(H) For all $h \in \hat{H}$ such that $k(h) = d$, $S^*_{ih} \subseteq R^*(S_{ih})$.

Suppose we have demonstrated (H) for all $d \geq d'$ where d' is between 1 and $\hat{d}(\omega)$. We wish to show that this implies (H) for $d = d' - 1$.

First, notice that for any $h \in \hat{H}$, the set $A(h)$ must have at least two elements. To see this, notice that $A(h)$ is only nonempty when

$$\sum_i v_i(k(h) + 1 \mid \omega) > c(k(h) + 1) - c(k(h))$$

Obviously there will exist many choices of a such that $a_i < v_i(k(h) + 1 \mid \omega)$ and such that the a's sum to $c(k(h) + 1) - c(k(h))$. Importantly, notice that this will also be true for the approximating games for n sufficiently large. Hence we see immediately that there are strategies in $R^*(S)$ such that contributions will not sum to $c(k(h) + 1) - c(k(h))$ for any h such that $k(h) \geq d'$. Simply choose the smallest element of $S^*_{ih'}$ for each ih'. Then we can use this fact to construct $\sigma_{ih'}$ analogously to the above and the same arguments will establish the claim. ∎

We are finally ready to complete the proof that the set of SUSPE outcomes is a subset of the core.

Lemma 2.5. *If (A.4) holds and (d, x) is a SUSPE outcome at ω, then $d \geq d^*(\omega) - 1$. If (A.4) does not hold and (d, x) is a SUSPE outcome at ω, then $d = d^*(\omega)$.*

Proof: Suppose not. Then there is a state ω and a SUSPE at ω, σ, such that $O^2(\sigma) = (d, x)$ and d violates the statement of the lemma. Since σ is a SUSPE, there must be a sequence $\{\sigma(n)\}$ converging to σ such that $\sigma(n)$ is a SUSPE of the n^{th} approximating game for each n. Let $h^*_n = h^*(\sigma(n))$—that is, h^*_n is the equilibrium path. Let h'_n be the subhistory of h^*_n consisting of the first $m^*(h^*_n) - 1$ components. Thus h'_n gives h^*_n up to the round at which contributions do not add up to enough to change the status quo. So, for n sufficiently large, $k(h'_n) = d$ and $\sum_i \sigma_{ih'_n}(n) < c(d + 1) - c(d)$. In what follows, we omit the n subscript on h'_n.

Recall that, by definition, if $\sigma(n)$ is a SUSPE, then, for all ih, $\sigma_{ih}(n)$ is a best response to any vector of completely mixed strategies for the other agents close to $\sigma_{\sim(ih)}(n)$. We will find an ih and a vector of completely mixed strategies which provides a contradiction. For obvious reasons, the history we choose is h'. So choose any i such that $\sigma_{ih'} < v_i(d + 1 \mid \omega)$. The fact that the valuations sum to more than $c(d+1) - c(d)$ and the contributions do not implies that such an i must exist. The construction of the completely mixed strategies is quite simple. For $h \neq h'$, let the probability that jh chooses $\sigma_{jh}(n)$ be $1 - \epsilon^k$ for some large k. The remaining ϵ^k of probability is spread over the strategies in $R^*(S_{jh}(n))$ in any way. For $h = h'$, let the probability that jh chooses $\sigma_{jh}(n)$ be $1 - \epsilon$. Consider the remaining ϵ of probability to be put on jh's other possible strategies. We put probability $y\epsilon/(1+y)$ (for some large y) on a particular strategy which we will denote $\sigma'_{jh}(n)$. For now, we will let $\sigma'_{jh}(n) = x_j(n)$. (The argument is not affected if $x_j(n) = \sigma_{jh}(n)$.) The remaining probability for jh for $h = h'$ is spread over the other strategies in $R^*(S_{jh}(n))$ in any fashion. We will show that there must exist an $x_j(n) \in R^*(S_{jh'}(n))$ for $j \neq i$ such that $\sigma_{ih'}(n)$ is not a best reply to this vector of completely mixed strategies for $\sim (ih)$.

First, note that the fact that d violates the condition stated in the lemma means that

$$\sum_i v_i(d+1 \mid \omega) > c(d+1) - c(d) \tag{A.9}$$

Let

$$J = \{j \in I \setminus \{i\} \mid \sigma_{jh'} < v_j(d+1 \mid \omega)\}$$

It is easy to see that (A.9) implies that J is nonempty. (A.9) also implies that, for all n sufficiently large, there exists a set $B \subset J$ satisfying

$$\sum_{j \in B} v_j(d+1 \mid \omega) + \sum_{j \notin B, j \neq i} \sigma_{jh'}(n) + v_i(d+1 \mid \omega) > c(d+1) - c(d) \tag{A.10}$$

as this is true for $B = J$. The fact that the power set of J is finite implies that there is a finite number of sets B satisfying (A.10). Hence we can certainly find a set with the property that $B \subseteq J$, B satisfies (A.10), and no subset of B other than B itself satisfies (A.10). Choose any such B and call this set B^*.

We will now show that for n sufficiently large, we can choose $x_j(n)$ for $j \in B^* \cup \{i\}$ so that

$$\sigma_{jh'}(n) < x_j(n) < \min[v_j(d+1 \mid \omega),\ w_j - t_j(h')], \tag{A.11}$$

$$\sum_{j \in B^*} x_j(n) + x_i(n) + \sum_{j \notin B^*, j \neq i} \sigma_{jh'}(n) =_n c(d+1) - c(d), \tag{A.12}$$

and

$$\sum_{j \in B^*} x_j(n) + \sigma_{ih'}(n) + \sum_{j \notin B^*, j \neq i} \sigma_{jh'}(n) < c(d+1) - c(d) \tag{A.13}$$

where the notation $=_n$ indicates that contributions are at least $c(d+1) - c(d)$ and are strictly smaller than $c(d+1) - c(d) + \delta_n$.

To see this, first note that the fact that the original strategies were an equilibrium with contributions not adding up at h' implies that $v_j(d+1 \mid \omega) < w_j - t_j(h')$. This is true because

$$w_j \geq \sum_{m=1}^{M} v_j(m \mid \omega)$$

by assumption, so $v_j(d+1 \mid \omega) \geq w_j - t_j(h')$ would imply

$$t_j(h) > \sum_{m=1}^{d} v_j(m \mid \omega)$$

But this means that j gets a payoff strictly less than w_j. Since j can always choose $\sigma_j \equiv 0$ and guarantee himself a payoff at least equal to w_j, this could not be an equilibrium. Hence we can rewrite (A.11) as

$$\sigma_{jh'}(n) < x_j(n) < v_j(d-1 \mid \omega)$$

Since for each $j \in J \cup \{i\}$, $\sigma_{jh}(n)$ is bounded away from $v_j(d-1 \mid \omega)$, we can obviously find $x_j(n)$ in this range for n sufficiently large. Therefore, (A.11) can certainly be satisfied.

To see that (A.11) and (A.12) can be satisfied simultaneously, simply note that the strict inequality in (A.10) means that for n sufficiently large, we can choose $x_j(n)$ close enough to $v_j(d-1 \mid \omega)$ to guarantee an inequality in (A.12) while satisfying (A.11). We can then lower the x_j's gradually to achieve the (approximate) equality required by (A.12). Since this is an equilibrium in which contributions do not add up, we must have

$$\sum_{j \neq i} \sigma_{jh'}(n) + v_i(d+1 \mid \omega) < c(d+1) - c(d)$$

so that we can reach this equality without violating $\sigma_{jh'}(n) < x_j(n)$. Clearly, the equality in (A.12) combined with $\sigma_{ih'}(n) < x_i(n)$ implies (A.13) immediately.

Now we only need to choose $x_j(n)$ for $j \notin B^*$. We will let $x_j(n)$ equal the smallest contribution in $S_{jh'}(n)$ for $j \neq i$, $j \notin B^*$. We have two facts left to establish. First, we will demonstrate that $\sigma_{ih'}(n)$ is not a best response for ih' to the vector of completely mixed strategies constructed in this fashion. Then we will show that $x_j(n) \in R^*(S_{jh'}(n))$ for all j so that this is an allowable vector of completely mixed strategies for the reduced game. This will imply that $\sigma(n)$ is not a SUSPE for n large and hence that σ cannot be a SUSPE.

To complete the first task, we will show that $x_i(n)$ is a strictly better response for ih' to this vector than is $\sigma_{ih'}(n)$ for all ϵ sufficiently small. To see this, notice that it is sufficient for our purposes to suppose that the outcome of randomization by the agents other than ih' is that jh' chooses either $\sigma_{jh'}(n)$ or $x_j(n)$ and, for all $h \neq h'$, jh chooses $\sigma_{jh}(n)$. If we can demonstrate that $x_i(n)$ yields a strictly higher expected payoff for ih' under this assumption, then we can choose y and k large enough that $x_i(n)$ will be a strictly better response to the vector of completely mixed strategies. Under this supposition, the expected payoff to ih' from $x_i(n)$ minus the expected payoff from $\sigma_{ih'}(n)$ is at least the probability that contributions from $jh' \neq ih'$ is at least $c(d+1) - c(d) - x_i(n)$ times a positive number minus the probability that contributions from $jh' \neq ih'$ is at least $c(d+1) - c(d) - \sigma_{ih'}(n)$ times a positive number.[25] Notice that the former probability is a sum of the form:

$$\left(\frac{y\epsilon}{1+y}\right)^{\#B} (1-\epsilon)^{I-1-\#B}$$

where $B \subset I \setminus \{i\}$ such that:

$$\sum_{j\in B} x_j(n) + \sum_{j\notin B, j\neq i} \sigma_{jh'}(n) + x_i(n) \geq c(d+1) - c(d) \tag{A.14}$$

The latter probability is a sum of the same form except that B must satisfy

$$\sum_{j\in B} x_j(n) + \sum_{j\notin B, j\neq i} \sigma_{jh'}(n) + \sigma_{ih'}(n) \geq c(d+1) - c(d) \tag{A.15}$$

We wish to show that this difference is strictly positive for ϵ sufficiently small. By construction, we know that there is no set $B \subset B^*$ which satisfies (A.14). Furthermore, for all $j \neq i$ such that $j \notin B^*$, $\sigma_{jh'}(n) \geq x_j(n)$ so that no set B which does not contain B^* can satisfy (A.14). However, B^* certainly satisfies (A.14) and there may be sets B with B^* as a subset which satisfy (A.14). However, notice that (A.12) and the fact that $j \neq i$ and $j \notin B^*$ implies $\sigma_{jh'}(n) \geq x_j(n)$ means that there is no B satisfying (A.15). Hence this difference must be strictly positive.

Thus if the $\sigma'_{jh'}(n)$ strategies we have constructed are elements of $R^*(S_{jh'}(n))$ for each jh', we are done. But notice that the $x_j(n)$ for all j are in $S^*_{jh'}(n)$ and hence, by Lemma 2.4, must be in $R^*(S_{jh'}(n))$. ∎

We have now established that for any ω, the set of SUSPE outcomes at ω is a subset of $C(\omega)$. Now we only need to show that any outcome in $C(\omega)$ is a SUSPE outcome at ω. This is not as straightforward as the analogue was for Theorem 1 because of the many unreached information

[25] Notice that this implicitly assumes that i will not subsequently contribute more than his valuation for whatever further streetlights are built. Since the SUSPE's are a subset of the subgame perfect equilibria and this is implied by subgame perfection, we see that this must hold.

sets in equilibrium. Strict perfection of the equilibrium essentially requires strict perfection at every information set, often difficult to verify (or satisfy). This is where successive elimination of dominated strategies proves especially useful. Recall that Lemma 2.3 showed that this elimination implies that information sets where adding streetlights is not possible as a Pareto improvement need not concern us. The lemma showed that there are no strategies available to the players that can affect the outcome at that point and hence any specification of strategies in $R^{*}(S)$ for these information sets will not affect whether or not the strategy tuple is a SUSPE.

We now show how to construct a SUSPE for a state ω and a given element of $C(\omega)$. Obviously, if $d^{*}(\omega) = 0$, Lemma 2.3 implies that our task is trivial. So consider any ω for which $d^{*}(\omega) \geq 1$ and any $(d, x) \in C(\omega)$. Since we can choose any approximating action sets we like, we will suppose that the n^{th} approximating game has all feasible contributions being integer multiples of $1/n$. That is, for all i and n,

$$A_i(n) = \{0, 1/n, 2/n, \ldots, \langle w_i \rangle_n\}$$

where for any real number a, $\langle a \rangle_n$ is the largest integer multiple of $1/n$ less than or equal to a. First, suppose that (A.4) does not hold so that $d = d^{*}(\omega)$ or let $d = d^{*}(\omega) - 1$. Furthermore, suppose that for each i,

$$x_i > w_i - \sum_{d \leq d^{*}(\omega)} v_i(d \mid \omega)$$

For each i and n, choose d terms, which we will denote $p_i(m \mid n)$, in such a way as to guarantee that

$$\text{(A.16)} \qquad \sum_i p_i(m \mid n) =_n c(m) - c(m-1), \; \forall \; m \leq d, \; n \in N,$$

$$\text{(A.17)} \qquad 0 < p_i(m \mid n) < v_i(m \mid \omega), \; \forall \; i \in I, \; m \leq d, \; n \in N,$$

$$\text{(A.18)} \qquad p_i(m \mid n) = \langle p_i(m \mid n) \rangle_n, \; \forall i \in I, n \in N, m \leq d.$$

Let $\tilde{H}(n)$ denote the set of histories of length less than or equal to d on which

$$P_m(h) = p(m \mid n), \;\; \forall m \leq \ell(h)$$

Note that $\tilde{H}(n)$ contains exactly one history of each length less than or equal to d, including the "empty history," e. For each $m \leq d$, let $\tilde{h}(m \mid n)$ denote the history of length m in $\tilde{H}(n)$. The final restriction on the numbers to be chosen, then, is

$$\text{(A.19)} \qquad w_i - t_i(\tilde{h}(d \mid n)) \in (x_i - 1/n, x_i + 1/n)$$

This is clearly possible.

Our objective is to choose strategies so that these numbers are the contributions along the equilibrium path up to $\tilde{h}(d \mid n)$. By Lemma 2.3, we will not need to consider histories such that $k(h) \geq d$. To construct the equilibrium strategies, we begin by partitioning the set of histories such that $r(h) = 1$ into four sets. The first set is those histories which Lemma 2.3 tells us we can ignore. That is,

$$H_0^a(n \mid \omega) = \{h \in H_1^*(n) \mid k(h) \geq d \text{ or } k(h) < d \text{ but}$$

$$\sum_i \langle \min[v_i(k(h)+1 \mid \omega),\ w_i - t_i(h)] \rangle_n < c(k(h)+1) - c(k(h))\}$$

where $H_1^*(n)$ is the set of possible histories with $r(h) = 1$. The second set is also one we can essentially ignore. Let

$$H_0^b(n \mid \omega) = \{h \in H_1^*(n) \mid w_i > t_i(h) \text{ for exactly one } i \text{ and } h \in H_0^a(n \mid \omega)\}$$

That is. this is the set of histories for which exactly one person has wealth left and this person wishes to change the outcome. Once we reach such a history, clearly, this person's optimal strategy is rather trivial. Notice that considerations of how he "trembles" on subsequent histories is also essentially irrelevant as failing to contribute enough to change the outcome when he has not yet reached his favorite d is dominated as is changing the outcome to a d larger than his favorite one. Since an individual cannot unilaterally cause others to lose their entire wealth, this histories will not concern us very much.

The third set is an important one. This is the set of histories consistent with our chosen equilibrium path which have not yet reached $k(h) = d$. That is, the third set is

$$H_0(n \mid \omega) = \tilde{H}(n) \setminus \{\tilde{h}(d \mid n)\}$$

That is. H_0 is the set of histories with no deviations from the equilibrium path. The fourth set, which we will partition further in a moment, contains all other histories in $H_1^*(n)$.

We will let $\sigma_{ih}(n)$ be any strategy in $R^*(S_{ih}(n))$ for any $h \in H_0^a(n \mid \omega)$ or $H_0^b(n \mid \omega)$. For any $h \in H_0(n \mid \omega)$, $\sigma_{ih}(n) = p_i(k(h)+1 \mid n)$. It is easy to see that no agent at any history in H_0 wishes to deviate downward. Reducing one's contribution at such a history is strictly worse than the proposed equilibrium strategy as this causes contributions to fail to add up to $c(k(h)+1) - c(k(h))$. Since each person's contribution is strictly less than his valuation for $k(h)+1$, this change must make him strictly worse off. What is more difficult is guaranteeing that no agent wishes to deviate upward in his contribution. Since such a deviation will not end the game, this can only be established by assigning strategies off the equilibrium path properly.

To do so, consider the histories not in H_0^a, H_0^b, or H_0. Without loss of generality, we can write any such history, say h, as $h' \cdot h''$, where h' is the longest subhistory of h which is an element of H_0 and $\ell(h'') \in [1, d - \ell(h')]$. This is true because e, the empty history, is in H_0. We will write this decomposition as $h_0'(h) \cdot h_0''(h)$ whenever there may be some ambiguity about which history we are decomposing in this fashion or what set h' belongs to. Let $H_1^a(n \mid \omega)$ denote the set of histories in $H_1^-(n)$ not in H_0^a, H_0^b, or H_0 such that $\ell(h''(h)) = 1$. For these histories, the first deviation from the equilibrium path has just occurred. Note that every history not in H_0^a, H_0^b, or H_0 must have a subhistory in H_1^a.

Recall that we only need to guarantee that no agent wishes to deviate upward from a history in H_0. However, we do also need to guarantee that no agent wishes to deviate from the prescribed strategy at a history in H_1^a. Clearly, if we choose strategies at H_1^a to guarantee that contributions just sum to the costs and each person's contribution is strictly less than his valuation, then, as above, we will guarantee that no individual wishes to deviate downward at a history in H_1^a. At the same time, we should choose the contributions to punish anyone who deviated upward from a history in H_0 to cause the succeeding history to be in H_1^a. If we can do this, then we will simply need to carry out an analogous construction for a set of histories which have a deviation from the equilibrium path from H_1^a to make sure that no one wishes to deviate upwards from their proposed equilibrium strategy in H_1^a. Obviously, we will do this in an analogous manner. Since we can only have a finite number of periods of deviations, this procedure must end and we will be done.

To see how to do this, consider a history $h \in H_1^a(n \mid \omega)$. Suppose $t(h) = t(\tilde{h}(k(h) \mid n))$. That is, total contributions net of refunds by each person is exactly what it would have been on the equilibrium path. This is only possible if every person deviates upward by exactly the same percentage and thus requires multi-lateral deviations. Since such histories cannot be reached from the equilibrium path by a unilateral deviation, they will not concern us very much. For such an h, we set $\sigma_{ih}(n) = p(k(h) + 1 \mid n)$.

Now suppose $t(h) \neq t(\tilde{h}(k(h) \mid n))$. Since

$$\sum_i t_i(h) = c(k(h)),$$

this implies that some agents have larger net wealth than they would have had on the equilibrium path and others have smaller net wealth. Since a unilateral deviation cannot keep $r = 1$ and achieve larger net wealth, it is those agents who have smaller net wealth that we need to "punish." Accordingly, we will choose an equilibrium path from this point which ends up giving those who now have a smaller net wealth than they should have a smaller net wealth at the end than they would have had. One complication in doing so is that some agents may have lost so much net wealth that it is no longer possible to reach d without some agent contributing more than his valuation or more than his remaining wealth. Notice, though, that if we ensure that at each status quo decision, a deviator has lower net wealth than he would have had at that decision along the equilibrium path, then he is unambiguously worse off deviating even if the social decision changes. This is true because the difference in his ultimate net wealth along the two paths is less than his willingness to pay for the larger social decision. (Clearly, the loss of wealth cannot make it possible to reach a social decision larger than d.) The most straightforward way to punish deviators and to delineate the cases where the loss of wealth must affect the path through the game tree from h is to suppose that the original equilibrium contributions are continued from here if possible. We now construct the histories that would result if this were done. If these are feasible, they will constitute the equilibrium path from h. So let $\tilde{H}(h,n)$ denote the set of h' with h as a subhistory such that for all $m \leq \ell(h')$, $m > \ell(h)$,

$$P_m(h') = p(m + k(h) - \ell(h) \mid n)$$

That is, $\tilde{H}(h,n)$ is the set of histories consistent with these contributions being given once we reach h. As before, for any $m \geq \ell(h)$, $m \leq d - k(h) + \ell(h)$, let $\tilde{h}(m \mid h,n)$ denote the (unique) history of length m in $\tilde{H}(h,n)$. First, suppose that $\tilde{H}(h,n) \subseteq H^*(n)$. That is, it is feasible to continue with the original equilibrium contributions. For such a history, we suppose $\sigma_{ih}(n) = p_i(k(h) - 1 \mid n)$ as before.

Suppose instead that $\tilde{H}(h,n) \not\subseteq H^*(n)$. Let $h^*(h,n)$ denote the shortest history in $\tilde{H}(h,n)$ not in $H^*(n)$ and let $k^* = k(h^*(h))$ and $t^* = t(h^*(h))$. There are three possible cases to consider. First, it may be true that

$$\sum_i \langle \min[v_i(k^* + 1 \mid \omega),\ w_i - t_i^*] \rangle_n < c(k^* + 1) - c(k^*)$$

In this case, Lemma 2.3 implies that contributions cannot add up at $h^*(h)$. In this case, our specification of strategies is irrelevant.

The second possibility is that

$$\text{(A.20)} \quad \sum_i \langle \min[v_i(k^* + 1 \mid \omega),\ w_i - t_i^*] \rangle_n = \sum_i \min[v_i(k^* + 1 \mid \omega),\ w_i - t_i^*] = c(k^* + 1) - c(k^*)$$

For these histories, it is possible that contributions add up but only if everyone gives the most that they are able/willing to pay. It is not immediately clear whether or not there are contributions in

$R^*(S_{ih})$ that will add up at such a stage. If so, then those who have $w_i - t_i^* < v_i(k^* + 1 \mid \omega)$ must contribute $w_i - t_i^*$. This is true because with any small probability that contributions will add up at this stage, those who gain a strictly positive amount from having this outcome will necessarily give the unique contribution for them which makes this possible. If there are no contributions in $R^*(S_{ih})$ such that the outcome can be changed at this history, then, clearly, the contributions are irrelevant. We will suppose that each person gives the largest element of $R^*(S_{ih}))$ at such a history. Note that contributions cannot possibly add up at the stage following such a history.

Finally, the last case is where

$$\sum_i \langle \min[v_i(k^* + 1 \mid \omega),\ w_i - t_i^*] \rangle_n > c(k^* + 1) - c(k^*)$$

To construct the strategies from this point on, first, define I numbers, $p_i(k(h^*) + 1 \mid h, n)$. These numbers should satisfy:

$$0 \leq p_i(k(h^*) + 1 \mid h, n) < \min[v_i(k(h^*) + 1 \mid \omega),\ w_i - t_i(h^*)]$$

$$\sum_i p_i(k(h^*) + 1 \mid n) =_n c(k(h^*) + 1) - c(k(h^*))$$

$$p_i(k(h^*) + 1 \mid n) = \langle p(k(h^*) + 1 \mid n) \rangle_n$$

Let $\tilde{h}^*(k^*, h, n)$ denote that history of length $\ell(h^*(h)) + 1$ which has $h^*(h)$ as a subhistory and

$$P_{\ell(h^*(h))+1}(\tilde{h}^*(k^*, h, n)) = p(k(h^*) + 1 \mid n)$$

Then we also want $p_i(k^* + 1 \mid n)$ to satisfy

$$t_i(\tilde{h}^*(k^*, h, n)) > t_i(\tilde{h}(k^*, n)), \text{ for all } i \text{ such that } t_i(h) > t_i(\tilde{h}(k(h), n) \tag{A.21}$$

Again, it is not hard to see that this is possible. For the next stage along the equilibrium path from $h^*(h)$, repeat the same procedure described above.

This procedure defines the equilibrium path from h for any $h \in H_1^a$. Let the set of histories which follow this σ from some $h \in H_1^a$ be $H_1(n \mid \omega)$. This set includes every history in H_1^a plus some histories for which $\ell(h_0''(h)) \geq 2$. Thus this is the set of histories with one deviation followed by equilibrium play thereafter. Consider the set of histories, h, in $H_1^*(n)$ but not in H_0^a, H_0^b, H_0, or H_1. Analogously to the above, we can decompose any such h into $h' \cdot h''$ where h' is the longest subhistory of h such that $h \in H_1$. Let $h_1'(h) \cdot h_1''(h)$ denote this decomposition and let $H_2^a(n \mid \omega)$ denote the set of such histories for which $\ell(h_1''(h)) = 1$. These are histories which contain a deviation from the equilibrium path at some point and a second deviation in the previous period. Simply imitate the construction given above to construct an equilibrium path from a history in H_2^a with the following alteration. Let h denote a history in H_2^a. Then we rewrite (A.21) to insure that any agent with lower net wealth at h than at the equilibrium path from $h_1'(h)$ has lower net wealth at each step of the path from h compared to the path from $h_1'(h)$. Clearly, we can define H_2 as the set of histories with exactly two deviations, define H_3^a analogously to the above, and so on.

Since the number of rounds of contributions is necessarily finite, this procedure must eventually terminate. It is not hard to see that the resulting strategies will have the property that for any ih, if (A.20) does not hold, $h \in H_1^*(n)$, and h is not in H_0^a or H_0^b, then the prescribed strategy for ih is a strict best reply to the strategies of $\sim (ih)$. Hence this strategy is a best reply to any vector of completely mixed strategies for $\sim (ih)$ close to $\sigma_{\sim(ih)}(n)$. Lemma 2.3 covers the strategies for histories in H_0^a and, as we argued above, histories in H_0^b are straightforward. Finally, as argued above, histories for which (A.20) holds do not affect the analysis. Hence $\sigma(n)$ is a SUSPE for all n. Notice, of course, that as $n \to \infty$, the outcome approaches the point in the core that we sought.

Now consider the case where (A.4) holds and we wish to achieve an outcome with $d = d^*(\omega)$. Here we can certainly follow a path to an outcome with $d = d^*(\omega) - 1$ and then have contributions at the last stage equal to the largest elements of $R^*(S_{ih}))$ for each i and that history. Since

$$\sum_i v_i(d^*(\omega) \mid \omega) = c(d^*(\omega)) - c(d^*(\omega) - 1)$$

and no one will contribute more than their valuation at this stage by Lemma 2.3, we see that this will either have contributions fall short or will have add up exactly. To see that (in the limit) the former will hold, consider the way dominated strategies are eliminated. Contributing more than one's valuation is not a dominated strategy until the possibility of getting past $d^*(\omega)$ has been eliminated by removal of other dominated strategies. Up till this point, it is certainly not dominated to contribute any amount less than one's valuation. Hence at the step in the successive removal where contributing more than one's valuation at this history is eliminated, strategies where one contributes strictly less must still be possible. But then contributing less will not be eliminated as no strategy for ih can affect the outcome and hence no strategy left for ih cannot possibly be dominated. But this implies that the largest element in $R^*(S_{ih} \mid n))$ must converge to $v_i(d^*(\omega) \mid \omega)$.

Finally, consider a core outcome with

$$x_i = w_i - \sum_{m=1}^{d} v_i(m \mid \omega)$$

It is easy to see that we can allow i to contribute slightly less than his valuation at each stage and let the difference between his contribution and valuation go to zero as $n \to \infty$.

Proof of Lemma

First, we show that $o_\delta \to o_0$ by induction. To begin the induction, recall that $k_\delta^1(h)$ is defined as the largest multiple of δ such that

$$\iota P_1(h) \geq c(k)$$

Obviously, as $\delta \downarrow 0$, the continuity of $c(\cdot)$ implies that $k_\delta^1 \to k_0^1$. Similarly, it is clear from this that $t_\delta^1 \to t_0^1$. So consider r_δ^1. Recall that

$$r_\delta^1(h) = \begin{cases} 1, & \text{if } k_\delta^1(h) \geq \delta; \\ 0, & \text{otherwise.} \end{cases}$$

and

$$r_0^1(h) = \begin{cases} 1, & \text{if } k_0^1(h) > 0; \\ 0, & \text{otherwise.} \end{cases}$$

Notice that $r_\delta^1 \to r_0^1$, iff for all δ sufficiently small, $r_\delta^1(h) = r_0^1(h)$. If $k_0^1(h) > 0$, then for δ sufficiently small, k_δ^1 must be strictly larger than δ as k_δ^1 is converging to $k_0^1 > 0$ and δ is converging to 0. Hence for all h such that $k_0^1(h) > 0$, $r_\delta^1(h) = r_0^1(h)$ for all δ sufficiently small. So suppose that $k_0^1(h) = 0$. But this is only possible if $P_1(h) = 0$ in which case $k_\delta^1(h) = 0$ also. Hence, obviously, $r_\delta^1 \to r_0^1$. Hence $o_\delta^1 \to o_0^1$.

So suppose that for some m', we have shown that $o_\delta^m \to o_0^m$ for all $m \leq m'-1$. We wish to show that this implies $o_\delta^{m'} \to o_0^{m'}$. It is easy to see that the induction hypothesis implies $k_\delta^{m'} \to k_0^{m'}$ and similarly for $t_\delta^{m'}$. An argument analogous to the one above establishes the result for $r_\delta^{m'}$, completing the proof. Hence o_δ converges pointwise to o_0 on H.

It is easy to see that this result implies that m_δ^* converges to m_0^* pointwise on H and hence that O_δ^* converges to O_0^* pointwise as well.

We now wish to show that this implies that the set of possible histories for δ converges to H_0^* as $\delta \downarrow 0$ in the sense defined in (10) and (11). Recall that H_δ^* is that subset of H such that

(A.23) $$t_\delta^m(h) \leq w, \text{ for all } m$$

and

(A.24) $$P_m(h) = e, \text{ for all } m \text{ such that } r_\delta^{m-1}(h) = 0.$$

So consider any $h \in H_0^*$. For all δ sufficiently small, the fact that we will have $r_\delta^m(h) = r_0^m(h)$ immediately implies (A.24). Furthermore, notice that $t_0^m(h) = \sum_{m' \leq m} \iota P_{m'}(h)$ while $t_\delta^m(h)$ is necessarily weakly smaller than the same quantity. Hence (A.23) must hold. Therefore, for all δ sufficiently small, $h \in H_\delta^*$. Now consider any $h \notin H_0^*$. If (A.23) is violated at $\delta = 0$, then the fact that t_δ^m converges pointwise to t_0^m implies that (A.23) will be violated for all δ sufficiently small. Similarly, if (A.24) is violated for $\delta = 0$, then convergence will again imply that (A.24) is violated for all δ sufficiently close to zero. Therefore, $h \notin H_\delta^*$ for δ sufficiently small. Hence (10) and (11) hold, so $H_\delta^* \to H_0^*$. In fact, for δ sufficiently small, $H_\delta^* \subseteq H_0^*$.

It is easy to show that (12) holds. The fact that $t_\delta^m(h) \le t_0^m(h)$ implies that for any $h \in H_\delta^*$, if $\sigma_i(h)$ satisfies this feasibility condition for $\delta = 0$, it must do so for δ close to zero. Therefore, (12) holds and $S_\delta \to S_0$.

Obviously, the function translating strategies into histories is entirely unaffected by δ. Therefore, the fact that $O_\delta^* \to O_0^*$ implies immediately that $O_\delta \to O_0$. ▌

Proof of Theorem 3

We wish to show that $C_\delta(\omega) \to C_0(\omega)$ as $\delta \downarrow 0$. Since the set of SUSPE outcomes at state ω in G_δ is the former, this implies that for δ sufficiently small, the set of SUSPE outcomes of G_δ will be approximately $C_0(\omega)$. Recall that we say that $C_\delta(\omega)$ converges to $C_0(\omega)$ if for all δ sufficiently small, $(d', x') \in C_0(\omega)$ implies $(d', x') \in C_\delta(\omega)$ and $(d', x') \notin C_0(\omega)$ implies $(d', x') \notin C_\delta(\omega)$.

To see that this holds, consider any $(d', x') \in C_0(\omega)$. Clearly, if $d' \in D_\delta$, then $(d', x') \in C_\delta(\omega)$. Recall from the Proposition that an outcome is in the core iff there is a price system supporting it. Any price system supporting the allocation for D_0 will work when restricted to D_δ. Hence, choosing the sequence of δ's to be the set of δ such that $\delta_n = d'/n$, we see that for all $(d', x') \in C_0(\omega)$, there is a sequence $\{C_\delta(\omega)\}_{\delta\in\Delta}$ such that $(d', x') \in C_\delta(\omega)$ for all $\delta \in \Delta$.

So consider any $(d', x') \notin C_0(\omega)$. We wish to show that for δ sufficiently small, $(d', x') \notin C_\delta(\omega)$. To see this, notice that there is no price system supporting (d, x). This implies that for any set of non–negative, upper semicontinuous functions, at least one of the following must fail:

(A.25)
$$\sum_i p_i(d') = c(d')$$

(A.26)
$$(d', x_i') \text{ maximizes } u_i(d, x_i \mid \omega) \text{ on } \{(d, x_i) \mid p_i(d) + x_i = w_i\}, \quad \forall i$$

(A.27)
$$d' \text{ maximizes } \sum_i p_i(d) - c(d)$$

Consider, then, any set of non-negative numbers $p_i(d)$ for $i \in I$ and $d \in D_\delta$ for which (A.25) holds. It will necessarily be true that (A.26) or (A.27) will fail to hold if the d which maximizes the relevant function with respect to D_0 is in D_δ. Again, for δ sufficiently small, either this d or one very close to it will be in D_δ. Hence a price system supporting (d', x') for D_δ will not exist so that $(d', x') \notin C_\delta(\omega)$. ▌

Proof of Theorem 4

First, we show that any approachable outcome must have $d = \bar{d}(\omega)$. Then we will show that any outcome in $\Theta_1(\omega)$ is an approachable outcome. Finally, we will give an example in which there is an approachable outcome which is not in $\Theta_1(\omega)$, indicating that the set of approachable outcomes in general includes more than $\Theta_1(\omega)$.

So suppose we have an approachable outcome (d', x') with $d' \neq \bar{d}(\omega)$. Suppose, first, that $d' < \bar{d}(\omega)$. This cannot be a Nash equilibrium, then, as any person for whom $d' < \hat{d}_i(\omega)$ can do better by contributing $c(\hat{d}_i(\omega)) - c(d')$. Hence we must have $d' \geq \bar{d}(\omega)$. Suppose, then, that $d' > \bar{d}(\omega)$. Consider the sequence of approximating games. In any game in the sequence, the number of rounds of contributions is finite. Hence there is a last round in any equilibrium in such an approximating game. But then anyone contributing in that round can cut their contribution by as much as $c(d') - c(\bar{d}(\omega))$ and be strictly better off. Hence the outcome in any approximating game must have $d \leq \bar{d}(\omega)$ and so $d' = \bar{d}(\omega)$.

To see that any outcome in $\Theta_1(\omega)$ is an appraochable outcome, simply consider the following strategies. Let (d', x') be some outcome in $\Theta_1(\omega)$. For the empty history, each agent except one contributes $\langle w_i - x'_i \rangle_n$. Some agent with $\hat{d}_i(\omega) = \bar{d}(\omega)$ makes up the extra contribution required by the approximation. On subsequent histories, if one's past contributions (not netting out refunds) totals at least $\sigma_{ie}(n)$, $\sigma_{ih}(n) = 0$. Otherwise, $\sigma_{ih}(n)$ is the difference between one's past contributions and $\sigma_{ie}(n)$. It is easy to see that for all ih, $\sigma_{ih}(n)$ is a best reply to $\sigma_{\sim(ih)}(n)$. Hence $\sigma(n)$ is a Nash equilibrium (in fact, a subgame perfect Nash equilibrium) and hence its limit as $n \to \infty$ is an approachable equilibrium. Clearly, the limiting outcome is (d', x').

To see that there can be approachable outcomes not in $\Theta_1(\omega)$, suppose $I = 3$. Suppose that for some ω, two agents, 1 and 2, have $U_i(d \mid \omega) = 3\log(1+d)$ and the third has $U_3(d \mid \omega) = 2\log(1+d)$. Suppose $c(d) = d$. It is easy to see that

$$\hat{d}_1(\omega) = \hat{d}_2(\omega) = 2$$

and

$$\hat{d}_3(\omega) = 1$$

Hence every outcome in $\Theta_1(\omega)$ has $d = 2$ and $x_3 = w_3$. As we have just shown, every approachable outcome must have $d = 2$ as well. Hence if there is an approachable outcome which is not in $\Theta_1(\omega)$, it must have $x_3 < w_3$. Consider, then, the following strategies.

$$\sigma_3(h) = \begin{cases} 1 - t_3(h), & \text{if } t_3(h) \leq 1 \text{ and } t_1(h) = t_2(h) = 0; \\ 0, & \text{otherwise.} \end{cases}$$

To define the strategies for 1 and 2, let

$$A_i^1 = \{h \in H^* \mid t_3(h) < 1 \text{ or there exists a subhistory of } h,\ h', \text{ such that } t_{3-i}(h') > 0 \text{ and } t_i(h') = 0\}$$

$$A_i^2 = \{h \in H^* \mid t_3(h) \geq 1 \text{ and } t_1(h) = t_2(h) = 0\}$$

$$A_i^3 = \{h \in H^* \mid \text{there exists a subhistory of } h,\ h', \text{ such that } t_i(h') > 0 \text{ and } t_{3-i}(h') = 0\}$$

Then for $i = 1, 2$, let

$$\sigma_i(h) = \begin{cases} 0, & \text{if } h \in A_i^1; \\ (2 - t_3(h))/2, & \text{if } h \in A_i^2; \\ 2 - t_3(h), & \text{otherwise.} \end{cases}$$

In other words, neither 1 nor 2 contributes any money until 3 has contributed at least 1. Once 3 has contributed this, 1 and 2 each contribute half the difference between 3's contributions and 2. If either 1 or 2 deviates from this by contributing early, the other two players do not contribute anything and the deviator contributes the difference between contributions to date and 2. Notice that for each ih, σ_{ih} is a best reply to $\sigma_{\sim(ih)}$. Hence these strategies constitute a subgame perfect Nash equilibrium.

To guarantee that the same is true of a sequence of approximating games, simply choose the approximation games so that $1/2$ is a multiple of $1/n$. Then all contributions on the equilibrium path will be multiples of $1/n$. Since any deviation in an approximating game will have to be a multiple of $1/n$, the subsequent contribution required for the deviator will also be a multiple of $1/n$ and so will be feasible. Hence this equilibrium is the limit of a sequence of subgame perfect equilibria of approximating games. ▌

Equilibria that are not Approachable

In this section of the Appendix, we consider equilibria that are not approachable for the case of $I = 2$. We show that there are very robust Nash equilibria in this set which do have outcomes in the core. In particular, many core outcomes are outcomes of a strong Nash equilibrium in the agent–normal form. A strong Nash equilibrium (using the definition of van Damme [1983]) is a strategy tuple for which each agent is choosing his unique best reply. We will also see that there are many outcomes not in the core which can be so reached. Our contruction does not yield all core outcomes, but this does not imply that these other core outcomes cannot be strong Nash outcomes. For simplicity, we assume $c(d) = d$, but it is straightforward to eliminate this assumption.

We begin by demonstrating the following result. Consider any feasible outcome (d_0, x_0) with $d_0 \geq \bar{d}(\omega)$ and $\sum_i (w_i - x_i) = d_0$. Let $\Theta^*(d_0, x_0, \omega)$ denote the set of outcomes, (d^*, x^*), with $d^* > d_0$, $d^* \leq d^*(\omega)$, $0 \leq x^* < x_0$, $\sum_i (w_i - x_i^*) = d^*$, and

$$u_i(d^*, x^* \mid \omega) > u_i(d_0, x_0 \mid \omega)$$

for $i = 1, 2$. Consider the subgame reached when contributions to date have been $w - x_0$. Then for each $(d^*, x^*) \in \Theta^*(d_0, x_0, \omega)$, there exist strategies in the subgame which constitute a strong Nash equilibrium of the agent–normal form with the outcome (d^*, x^*).

We construct these strategies as follows. Choose any $\lambda \in (0, 1)$ and define $z_i(d, x)$ by

$$u_i(d + z_i, x_i \mid \omega) = u_i(d^*, x^* \mid \omega)$$

For any history in the subgame—that is, for any history since (d_0, x_0) was reached—let

$$p_1(h) = \begin{cases} \lambda \min\{w_i - x_i^* - t_i(h),\ z_2(k(h), t(h))\}, & \text{if } \ell(h) \text{ is odd;} \\ 0, & \text{otherwise.} \end{cases}$$

and define $p_2(h)$ analogously except that contributions are not zero for even length histories. (We take 0 to be an even number.) Let $\tilde{H}$ denote the set of subgame histories consistent with these contributions. That is, $\tilde{H}$ is the set of subgame histories such that

$$P_m(h) = p(h'), \quad \forall m \leq \ell(h)$$

where h' is the first $m - 1$ components of h. Then let the strategies in the subgame be

$$\sigma_i(h) = \begin{cases} p_i(h), & \text{if } h \in \hat{H}; \\ 0, & \text{otherwise.} \end{cases}$$

To show that these strategies in the subgame form a strong Nash equilibrium in the agent–normal form, notice first that the strategies, if followed, define a sequence of outcomes which we will denote $\{(d_t, x_t)\}$ for $t = 0, 1, \ldots$. We claim the following statements are true about this sequence. First, for all t, for $i = 1, 2$,

(A.26) $$u_i(d_t, x_t \mid \omega) < u_i(d^*, x^* \mid \omega)$$

Second,

$$\text{(A.27)} \qquad \lim_{t\to\infty}(d_t, x_t) = (d^*, x^*)$$

The first statement is obviously true for $t = 0$. For all subsequent t, the utility of the agent who did not contribute at t is clearly less than $u_i(d^*, x^* \mid \omega)$ because the other agent's contribution was chosen to insure this. Furthermore, the same is true of the agent who did contribute. Recall that we started where $d = \bar{d}(\omega)$. Hence for all t, the contributing agent's marginal utility is strictly less than 1. Therefore, his utility, which started below $u_i(d^*, x^* \mid \omega)$, fell as a result of his contribution and hence must satisfy (A.26).

We now demonstrate that (A.27) holds. Suppose it does not. The limit must exist since the sequence of d's is an increasing, bounded sequence and the sequence of x's is a decreasing, bounded sequence. Let the limit be denoted by $(\tilde{d}, \tilde{x})$ where we suppose $(d^*, x^*) \neq (\tilde{d}, \tilde{x})$. By construction, though, $\tilde{x} \geq x^*$ and $\tilde{d} \leq d^*$. Hence if $\tilde{d} = d^*$, the fact that d must equal $\sum_i (w_i - x_i)$ yields a contradiction. Therefore, we must have $\tilde{d} < d^*$ and, for at least one i, $\tilde{x}_i > x_i^*$. But if this holds for exactly one i, for that i, we will have $u_i(\tilde{d}, \tilde{x} \mid \omega) > u_i(d^*, x^* \mid \omega)$ which contradicts (A.26). Therefore, we must have $\tilde{x}_i > x_i^*$ for $i = 1, 2$. But the only way this can hold is if $z_i(\tilde{d}, \tilde{x}) = 0$ for $i = 1, 2$. Otherwise, the contributions would not be converging to zero. Hence we must have

$$U_i(\tilde{d} \mid \omega) + \tilde{x}_i = U_i(d^* \mid \omega) + x_i^*$$

for $i = 1, 2$, so that

$$\sum_i U_i(d^* \mid \omega) + \sum_i w_i - d^* = \sum_i U_i(\tilde{d} \mid \omega) + \sum_i w_i - \tilde{d}$$

Recall, though, that $d^* \leq d^*(\omega)$ by assumption and that $d^*(\omega)$ maximizes $\sum_i U_i(d \mid \omega) - \sum_i w_i - d$. The strict concavity of the utility functions implies that the function being maximized is strictly larger at d^* than at $\tilde{d}$ if $d^* > \tilde{d}$. Hence we have a contradiction. Therefore (A.27) holds.

The fact that (A.26) and (A.27) hold implies that these strategies form a strong Nash equilibrium in the agent–normal form of the subgame. To see this, note simply that any deviation from the equilibrium path by any ih leads to an outcome strictly inferior for either agent than (d^*, x^*). Hence any history on the equilibrium path has ih choosing a strict best reply to the strategies of $\sim (ih)$. Consider any history off the equilibrium path. If the game has not yet ended, the strategies prescribe 0 for each agent. Notice that these strategies are again a strict best replies as the fact that the current d must be at least $\bar{d}(\omega)$ implies that the best unilateral contribution is 0.

We now see that from any outcome (d_0, x_0) with $d_0 \geq \bar{d}(\omega)$, we can reach any Pareto preferred outcome without overprovision of the public good which has $x^* < x_0$. Choose some (d_0, x_0) in $\Theta_1(\omega)$, fix some d^* such that there exists x^* for which $(d^*, x^*) \in \Theta^*(d_0, x_0, \omega)$. Consider the set of x^* such that $(d^*, x^*) \in \Theta^*(d^0, x^0, \omega)$. Since $\Theta^*(d^0, x^0, \omega)$ is an open set, the set of such x^* is open. Hence no matter which x^* in this set we choose, there is another one which is worse for i. This fact will prove convenient below. So choose any $(d^*, x^*) \in \Theta^*(d_0, x_0, \omega)$. Construct equilibrium strategies reaching (d^*, x^*) as follows. On the empty history, each person contributes $w_i - x_{i0}$. If these strategies are followed in the first round, then assign the strategies constructed above which lead to (d^*, x^*) for all subsequent histories.

To handle deviations from the strategies assigned for a history such that $k(h) < d_0$, we adopt the following rules. First, whenever the deviation leads to h' such that $k(h') \geq \bar{d}(\omega)$, then set $\sigma_{ih''} = 0$ for all i and all h'' with h' as a subhistory. Multilateral deviations, of course, must be considered even though they will not affect an agent's decision. We consider all multilateral deviations together at the end. First, then, suppose that one agent deviated downward in the first round and let i^* denote this agent. Choose a new equilibrium path from this history as follows. Let $\sigma_{i^* h}$ equal to $\sigma_{i^* e}$ minus whatever i^* contributed so far. Let $\sigma_{2-i^*, h} = 0$. If followed, these strategies still lead to (d_0, x_0). Let

$$\underline{x}_i = \inf_{x_i}\{x \mid (d^*, x) \in \Theta^*(d_0, x_0, \omega)\}$$

Choose a different outcome in Θ^* to end up at by leaving d unchanged and setting i^*'s wealth at the final outcome to be $x^*_{i^*} - \beta[x_{i^*} - \underline{x}_0]$ for some $\beta \in (0, 1/2)$. By the definition of $\underline{x}_{i^*}$, this outcome must be in Θ^*. Notice, also, that $\beta < 1/2$ implies that the outcome defined by setting x_{i^*} equal to

$$x^*_{i^*} - \sum_{k=1}^{\infty} \beta^k [x^*_{i^*} - \underline{x}_i]$$

is also in Θ^*. If both players follow the strategies assigned from h to (d_0, x_0), then choose the outcome prescribed above and choose the strategies which lead to it. To cover deviations from this path prior to reaching (d_0, x_0), simply adopt the following rules. As above, if the deviation leads to some h such that $k(h) \geq \bar{d}(\omega)$, assign the strategies $\sigma_{ih} = 0$ forever after. Notice that if there is a unilateral deviation by $2 - i^*$, this must be true. Since $2 - i^*$ is supposed to give 0 and i^* is supposed to contribute so we reach (d_0, x_0), if only $2 - i^*$ deviates, we must have $k(h) > d_0 = \bar{d}$. Hence the only other cases to consider are deviations downward by i^* and deviations by both players. If only i^* deviated downward, adjust the strategies so that contributions at the next round take us to (d_0, x_0) and alter the ultimate outcome by subtracting $\beta^k [x^*_{i^*} - \underline{x}_{i^*}]$, where k is the number of times i^* has deviated. The case of multilateral deviations is handled below.

Since what happens after multilateral deviations do not affect the incentives of the agents to deviate unilaterally, we have more flexibility here. Suppose we have a multilateral deviation from the strategies specified so far at some history h' satisfying $k(h') < \bar{d}(\omega)$. Again, if this deviation leads to a history h with $k(h) \geq \bar{d}(\omega)$, we set all contributions on any succeeding history to zero. Suppose, then, that the deviations lead to a history h with $k(h) < \bar{d}(\omega)$. Choose any x_0' such that $0 < x_0' < w - t(h)$ and $\sum_i (w_i - x_{i0}') = \bar{d}(\omega)$. Let the outcome $(\bar{d}(\omega), x_0')$ replace $(\bar{d}(\omega), x_0)$ to imitate the construction given above for the strategies following this history.

Consider the strategies so constructed. Clearly, after a history of length one consistent with the equilibrium strategies, the equilibrium in the subgame is a strong equilibrium in the agent–normal form. In fact, from any history h with $k(h) \geq \bar{d}$, the same is true. This is because such a history either has a deviation at the previous round or else it does not. In the former case, the strategies have zero being contributed and we know that this is a strict best reply to the other player contributing zero. In the latter case, we are following a path like those constructed at the beginning of this section and, again, these strategies are strict best replies to one another.

So consider the strategies assigned for any h for which $k(h) < \bar{d}$. At this point, there is some (d_0, x_0) that the strategies lead to and some (d^*, x^*) that the strategies lead to from there. Notice that any unilateral deviation upward leads to a history with $k(h') \geq \bar{d}$. Hence such a deviation leads to an outcome strictly worse than (d_0, x_0) for the deviator and hence strictly worse than (d^*, x^*). Any unilateral deviation downward leads to an outcome better than (d_0, x_0) but strictly worse for the deviator than (d^*, x^*). Hence the strategies proposed are again strict best replies. Therefore. we have constructed strong Nash equilibria in the agent–normal form.

Hence we have shown that any outcome in $\Theta^*(\bar{d}(\omega), x, \omega)$ for x satisfying $x \geq 0$ and $\sum_i (w_i - x_i) = \bar{d}(\omega)$ can be achieved by a strong Nash equilibrium in the agent–normal form. The fact that this set contains outcomes with $d < d^*(\omega)$ immediately implies that there are outcomes not in $C_0(\omega)$ that can be reached. However, notice that this set also contains all Pareto optimal points which are strictly Pareto preferred to some point which must be strictly individually rational for at least one agent and individually rational for both. With only two agents, of course, the core is simply the set of Pareto optimal, individually rational outcomes. Hence any point in the core except for one in which some agent receives "too little" can be reached by such an equilibrium. In particular, if for one i, $\hat{d}_i(\omega) < \bar{d}(\omega)$, we cannot achieve any point in the core too close to one yielding him only his individually rational payoff. To be more precise, this construction does not yield such outcomes. Whether or not they are possible is an open question.

BIBLIOGRAPHY

Aghion, P., "On the Generic Inefficiency of Differentiable Market Games," *Journal of Economic Theory*, 37, 1985.

Andreoni, J., "Impure Altruism and Donations to Public Goods," University of Michigan working paper, 1985.

Bagnoli, M., and B. Lipman, "Can Private Provision of Public Goods be Efficient?", University of Michigan working paper, 1986.

Benassy, J.-P., "On Competitive Market Mechanisms," *Econometrica*, January, 1986.

Bergstrom, T., L. Blume, and H. Varian, "On the Private Provision of Public Goods," *Journal of Public Economics*, February, 1986.

Bernheim, D., "On the Voluntary and Involuntary Provision of Public Goods," *American Economic Review*, September, 1986.

Bernheim, D., and K. Bagwell, "Is Everything Neutral?," Stanford University working paper, 1985.

Bliss, C., and B. Nalebuff, "Dragon-Slaying and Ballroom Dancing: The Private Supply of a Public Good," *Journal of Public Economics*, November, 1984.

Chatterjee, K., and L. Samuelson, "Perfect Equilibria in Simultaneous-Offers Bargaining," Pennsylvania State University working paper, October, 1986.

Cornes, R., and T. Sandler, "The Simple Analytics of Pure Public Goods Provision," *Economica*, 1985a.

Cornes, R., and T. Sandler, "On the Consistency of Conjectures with Public Goods," *Journal of Public Economics*, June, 1985b.

Ferejohn, J., R. Forsythe, R. Noll, and T. Palfrey, "An Experimental Examination of Auction Mechanisms for Discrete Public Goods," in *Research in Experimental Economics*, vol. 2, ed. by V. Smith, JAI Press, Greenwich, Conn., 1982.

Isaac, R., J. Walker, and S. Thomas, "Divergent Evidence on Free-Riding: An Experimental Examination of Possible Explanations," *Public Choice*, vol. 2, 1984.

Kalai, E., and W. Stanford, "Finite Rationality and Interpersonal Complexity in Repeated Games," Northwestern University working paper, April, 1986.

Mas–Colell, A., "Efficiency and Decentralization in the Pure Theory of Public Goods," *Quarterly Journal of Economics*, June, 1980.

Maskin, E., "Nash Equilibrium and Welfare Optimality," unpublished mimeo, 1977.

Moore, J., and R. Repullo, "Subgame Perfect Implementation," London School of Economics working paper, 1986.

Myerson, R., "Refinements of the Nash Equilibrium Concept," *International Journal of Game Theory*, vol. 7, 1978.

Nash, J., "Two–Person Cooperative Games," *Econometrica*, January, 1953.

Palfrey, T., and H. Rosenthal, "Private Incentives in Social Dilemmas," Carnegie–Mellon working paper no. 35–84–85, 1985.

Palfrey, T., and H. Rosenthal, "Participation and the Provision of Discrete Public Goods," *Journal of Public Economics*, January, 1984.

Palfrey, T., and S. Srivastava, "Nash Implementation using Undominated Strategies," Carnegie–Mellon University working paper, 1986.

Schneider, F., and W. Pommerehne, "Free Riding and Collective Action: An Experiment in Public Microeconomics," *Quarterly Journal of Economics*, November, 1981.

Sugden, R., "Voluntary Contributions to Public Goods," *Journal of Public Economics*, June, 1985.

van Damme, E., *Refinements of the Nash Equilibrium Concept*, Springer–Verlag, Berlin, 1983.

Warr, P., "The Private Provision of a Public Good is Independent of the Distribution of Income," *Economic Letters*, 13, 1983.

Recent CREST Working Papers

87–1: Jeffrey K. MacKie-Mason, "Nonlinear Taxation of Risky Assets and Investment, With Application to Mining" September, 1984.

87–2: Jeffrey K. MacKie-Mason, "Sequential Decision Problems and Asymmetric Information" September, 1985.

87–3: Michelle J. White, "Contract Breach and Contract Discharge due to Impossibility: A Unified Theory" July 29, 1987.

87–4: Ted Bergstrom, "Systems of Benevolent Utility Interdependence" May 19, 1987.

87–5: Ted Bergstrom, "A Fresh Look at the Rotten Kid Theorem—and Other Household Mysteries" November, 1986.

87–6: Michelle J. White, "The Corporate Bankruptcy Decision" July, 1987.

87–7: Michelle J. White, "Location Choice and Commuting Behavior in Cities with Decentralized Employment" July, 1987.

87–8: Lawrence E. Blume and David Easley, "Implementation of Walrasian Expectations Equilibria" December, 1985.

87–9: Lawrence E. Blume, "Lexiocographic Refinements of Nash Equilibrium" April, 1986.

87–10: David Lam, "Lorenz Curves, Inequality, and Social Welfare Under Changing Population Composition" June 16, 1987.

87–11: Mark Bagnoli and Naveen Khanna, "Equilibrium with Debt and Equity Financing of New Projects: Why More Equity Financing Occurs When Stock Prices are High" June, 1987.

87–12: Mark Bagnoli and Barton L. Lipman, "Provision of Public Goods: Fully Implementing the Core through Private Contributions" March, 1987.

87–13: Mark Bagnoli and Barton L. Lipman, "Successful Takeovers without Exclusion" August, 1987.

87–14: Mark Bagnoli and Michael McKee, "Controlling the Game: Political Sponsors and Bureaus" May, 1987.

87–15: Mark Bagnoli and Michael McKee, "Can the Private Provision of Public Goods be Efficient?—Some Experimental Evidence" March, 1987.

87–16: Mark Bagnoli, "Non-Market Clearing Prices in a Dynamic Oligopoly with Incomplete Information" January, 1986.

87–17: John Laitner, "Bequests, Gifts, and Social Security" February 28, 1986.

87–18: John Laitner, "Dynamic Determinacy and the Existence of Sunspot Equilibria" May 12, 1986.

87–19: David Lam, "Does a Uniform Age Distribution Minimize Lifetime Wages?" August 12, 1987.

87–20: David Lam, "Assortative Mating with Household Public Goods" April, 1987.

87–21: Jeffrey A. Miron and Stephen P. Zeldes, "Production, Sales, and the Change in Inventories: An Identity that Doesn't Add Up" June 1987.

87–22: Jeffrey A. Miron and Stephen P. Zeldes, "Seasonality, Cost Shocks, and the Production Smoothing Model of Inventories" December, 1986.

87–23: Hal R. Varian, "Differences of Opinion in Financial Markets" March, 1985.

87–24: Roger H. Gordon and Hal R. Varian, "Taxation of Asset Income in the Presence of a World Securities Market" August, 1986.

87–25: Hal R. Varian, "Measuring the Deadweight Costs of DUP and Rent Seeking Activities" November, 1982.

87–26: Hal R. Varian, "Price Discrimination" January, 1987.

87–27: Roger H. Gordon and Hal R. Varian, "Intergenerational Risk Sharing" October, 1985.

87–28: Hal R. Varian, "Three Papers on Revealed Preference" August, 1987.

87–29: Hal R. Varian, "Optimal Tariffs and Financial Assets" April, 1987.

87–30: Jonathan Cave and Stephen W. Salant, "Cartels That Vote: Agricultural Marketing Boards and Induced Voting Behavior" August, 1987.

87–31: Stephen W. Salant and Donald H. Negri, "Pastures of Plenty: When is the Standard Analysis of Common Property Extraction Under Free Access Incorrect?" July 10, 1987.

87–32: Stephen W. Salant, "When is Inducing Self-Selection Sub- optimal for a Monopolist?" February, 1987.

87–33: Stephen W. Salant, "Treble Damage Awards in Private Lawsuits for Price-Fixing" August, 1987.

87–34: Stephen W. Salant and Roy Danchick, "Air Force Academy Attrition: A New Perspective on the College Dropout Problem" August, 1987.

87–35: Stephen W. Salant and Eban Goodstein, "Committee Voting Under Alternative Procedures and Preferences: An Experimental Analysis" April 20, 1987.

87–36: Robert B. Barsky and Jeffrey A. Miron, "The Seasonal Cycle and the Business Cycle" June, 1987.

87–37: Robert B. Barsky, N. Gregory Mankiw, Jeffrey A. Miron and David N. Weil, "The Worldwide Change in the Behavior of Interest Rates and Prices in 1914" July, 1987.

87–38: Jeffrey K. MacKie-Mason, "Taxes, Information and Corporate Financing Choices" April 1986.