Efficient Mechanisms for Multiple Public Goods

1 Introduction

Designing incentive compatible mechanisms that lead to efficient production of public goods has long been an important topic in the theoretical literature on public economics. Within this literature, various mechanisms have been proposed. Among others, Bagnoli and Lipman (1989) proposed a mechanism that fully implements the core in a discrete and non-excludable public good economy and Jackson and Moulin (1992) introduced a mechanism for the single indivisible public good economy to achieve efficient and equitable allocations. Abreu and Sen (1990), Maniquet (1999) and Moore and Repullo (1988) construct mechanisms in general economic environments that under certain conditions can implement general social choice functions in the public goods framework.

Our paper is motivated by the same objectives but we depart from the existing literature in a number of ways. First, our framework is rather general. In contrast to the single public good case to which most of the literature confines itself, we are dealing with multiple public goods on which consumers’ preferences are not necessarily separable. Our model can be interpreted in two different ways. On the one hand, we can think of the multiple public goods as involving intrinsically different goods like public school education, roads, national defense and so on. We can also interpret the model as that of a single public good with multiple attributes. For example, suppose that a common computer laboratory is required to serve various department of a firm or university. This means deciding on the amounts to be spent on each attribute like software, hardware, furniture, programmers and so on. Thus, even when the public good problem involves a single facility, the multiple public goods model might be the appropriate way of describing the problem.

Second, while existing literature distinguishes between the excludable public good case (for which agents’ consumption can be restricted at zero cost, e.g. zoos, museums, cable TV etc.) and the non-excludable case (e.g. clean air, defense and security), our framework

*We thank an anonymous referee for detailed comments on two previous versions of this paper. We also thank Pierre Pestieau and another anonymous referee for their comments. Some parts of this paper were written during Suresh Mutuswami’s visits to CORE, Belgium and the Center for Rationality, Hebrew University, Jerusalem. He thanks these institutions for their support. Eyal Winter wishes to thank the Falk Institute for Economic Research for its support.
allows for both type of goods. It even allows some of the goods to be excludable and others non-excludable within the same problem.

Third, our concern is with simple mechanisms, where simplicity is meant to be in terms of the information that participants are requested to submit to the planner. In contrast to standard mechanisms (like Groves-Clarke or Groves-Ledyard mechanisms) where players are required to submit a full utility function over all possible levels of the public good, in ours players are only required to announce their desirable level of the public good/s, and subscribe to a contribution that will be paid conditional on their demand being met. In addition to their simplicity we believe that the operation of our mechanism resemble the way real life collective decisions on public good production are often made. Given their efficiency, they become natural candidates for application in concrete real life environments.

Consider for instance, the problem of building a common infrastructure for use by cellular phone companies in a country. This trend is already visible in Europe among cellular phone companies using the third generation technology. A common infrastructure is preferable to the alternative of having each company building it’s own infrastructure on two counts. Firstly, it is efficient. Secondly, the potential health hazards are minimized by having a common infrastructure. The problem is one of deciding where to locate the transmission towers and how much to charge each company. Since each company has different preferences concerning location because of the different geographical profiles of its customers, we believe that this problem is best treated as one involving multiple public goods. The government can divide the country into different geographical zones with each zone being treated as a public good. If our mechanism is used, each company would be required to submit a list of the public goods it requires along with a cost contribution.

We will be interested in two properties to be satisfied by the mechanism, efficiency and monotonicity with respect to individual preferences. The first property concerns the public good level while the second concerns the allocation of costs. Efficiency requires that the mechanism overcomes the free riding problem, i.e., the equilibrium behavior within the mechanism should result in a socially optimal vector of public goods. Monotonicity requires that if player $i$’s marginal utility for every good at all points exceeds player $j$’s marginal utility for every good, then player $i$ has a higher net utility in equilibrium. In particular, it requires that if two players share the same preferences regarding the public goods, then they will end up having the same net utility, and thus will also be making the same contribution. In addition to these two properties, we argue that in special cases, our mechanism satisfies a third property of core stability, i.e., the resulting outcome is immune to deviations by coalitions. Indeed this property will hold whenever the problem involves a single excludable public good (as shown by Bag and Winter (1999)). We note that no mechanism satisfies this property on the whole domain as such outcomes may not exist in the general case.

This paper is related to the literature on subgame perfect implementation, specifically the papers of Abreu and Sen (1990), Bag and Winter (1999), Moore and Repullo (1988) and Maniquet (1999). Abreu and Sen (1990) and Moore and Repullo (1988) have provided
necessary and sufficient conditions for a general social choice function to be implementable in subgame perfect equilibrium. Their mechanisms are quite complex relying on constructs such as integer games.

Maniquet (1999) characterizes the set of anonymous and individually rational social choice functions implementable through a “Divide and Challenge” mechanism. In the first stage of this \( n \)-stage mechanism, an agent (the “divider”) proposes an allocation along with the preference profile of all agents. In subsequent stages, the other \( n - 1 \) agents (the “choosers”) can challenge the divider by stating the name of an agent for whom the stated preference is incorrect and proposing an alternate allocation. While this mechanism is applicable in our context as well, it needs at least three agents to be present. In contrast, our mechanisms are valid even with two agents. Maniquet’s mechanism also requires an agent to announce the entire preference profile of all agents. As compared to this, our mechanisms are “simpler” since an agent’s announcement consists only of a desired vector of public goods and a (contingent) cost contribution.

Our paper is also related to that of Bag and Winter (1999) who propose two related mechanisms for the single excludable public good case. While simpler, their mechanisms do not work if the public good is pure or if there are multiple excludable public goods. Our mechanisms, in contrast, are more general in that they work when there are multiple public goods, each of which may be pure or excludable. This greater generality comes at the cost of making our mechanisms somewhat more complex, but it is worth noting that excludability plays no role in our mechanisms. In Bag and Winter (1999), the threat of exclusion is used to induce efficiency, whereas in our mechanisms, the “threat” is through manipulating the amounts of public goods produced. The mapping of bids into outcomes in our mechanisms therefore differs from the mechanisms of Bag and Winter even in the single excludable public good case. Our paper shows that efficiency can be obtained without resorting to threats of exclusion.

We organize the remainder of the paper as follows. In Section 2 we set up the multiple public goods model. In Section 3 we define two sequential mechanisms in which players move in turns by making bids regarding the requested level from each of public goods and conditional contribution. The first auxiliary mechanism (called \( \Gamma_m \)) yields an efficient outcome in equilibrium but the equilibrium outcome may not be equitable as players moving earlier have an advantage. The second mechanism (called \( \Gamma_{m1} \)) corrects for this asymmetry by allowing the mechanism to be played twice with the last period order being chosen randomly. Section 4 contains the formal analysis of the two mechanisms defined in Section 3. We demonstrate the operation of these mechanisms in Section 5 on a prominent 3-agent, 3-good problem proposed by Moulin (1995). In Section 6 we discuss conditions under which the equilibrium outcome of the mechanism is core stable. In particular, we demonstrate that this property holds whenever the public goods are excludable, the cost function is submodular and the utility functions are supermodular. We conclude in Section 7 by discussing the robustness of our mechanisms to the assumptions on technology and preferences. In particular, we discuss
the consequences of relaxing the assumption of quasi-linear preferences.

2 The Model

Let \( N = \{1, 2, \ldots, n\} \) denote the set of agents. There are \( p \) public goods in the economy. The set of feasible vectors of public goods for the economy are located in \( Y \subset \mathbb{R}_+^p \). A generic element of \( Y \) shall be denoted as \( y \). We shall assume that \( Y \) is non-empty, closed, unbounded, contains the origin (denoted by \( 0 \)) and a lattice, i.e., if \( y, y' \in Y \), then \( Y \) also contains \( y \lor y' = \max\{y, y'\} \) and \( y \land y' = \min\{y, y'\} \). Our formulation allows for the fact that some of the public goods may be “lumpy” in that they can be consumed in discrete units only. It also permits a given public good to be pure or excludable. In addition to the public goods, there is also a private, perfectly divisible good in the economy which can be interpreted as “money.” It is assumed that agent \( i \) \((i = 1, \ldots, n)\) has a strictly positive endowment of the private good denoted as \( w_i \).

The preferences of the individuals are quasi-linear, given by

\[
U_i(y, x_i) = v'_i(y) + w_i - x_i
\]

where \( v'_i(y) \) is \( i \)'s valuation (in units of the private good) for the vector \( y \) of public goods and \( x_i \) his contribution toward the cost of production. Since \( w_i \) is just a constant, we shall mostly work with the function \( v_i(y) = v'_i(y) + w_i \) in what follows. The technology in the economy is given by a production function \( f \) with an associated cost function \( c \). The following restrictions are imposed on the preferences and the technology.

**Assumption 1**

1. The function \( v'_i \) is continuous, non-decreasing and normalized such that \( v'_i(0) = 0 \) for all \( i \in N \).

2. The cost function \( c \) is continuous, non-decreasing and satisfies \( c(0) = 0 \). Also, \( c(y) \to \infty \) as \( y \to \infty \).

**Remark 1** Quasi-linear preferences are a standard assumption in the mechanism design literature on public goods. We discuss the consequences of relaxing this assumption in the Conclusion.

**Remark 2** The assumption that the economy has a finite endowment of the private good can be weakened if we assume that each agent’s preferences in the public goods is subject to satiation. In other words, for each agent \( i \), there exists \( y_i^M \) such that \( v'_i(y) \leq v'_i(y_i^M) \) for all \( y \in Y \). The analysis which follows will go through unchanged. As will become clear, we basically need conditions on preferences and cost which ensure that the efficient vector of public goods for any coalition of agents is not unbounded.

---

1 The maximum and minimum are taken coordinate-wise.

2 Quasi-linear preferences are a standard assumption in the mechanism design literature. We discuss the consequences of relaxing this assumption in the Conclusion.
We end our discussion of the model by introducing some notation which will be used in this paper. The subset of players \(\{k, k+1, \ldots, n\}\) is denoted \(S_k\) and \(S_{n+1}\) indicates the empty set. The announcement of agent \(i\) in a mechanism is indicated by the use of a subscript as in \(z_i, z'_i\) and so on while the \(n\)-tuple of all agents’ announcements is denoted as \(z, z'\) and so on. The term “announcement” is used interchangeably to indicate the announcement of a particular player and also to the vector of announcements of all players. This should not cause any confusion as the context will make it clear as to which usage is relevant.

3 The mechanisms \(\Gamma_m\) and \(\Gamma^1_m\)

Implementation Theory distinguishes between two distinct scenarios, complete information environments and incomplete information environments. In both scenarios, the planner who wishes to implement a social choice function meeting her objectives is ignorant about the environment. In the complete information setting, all agents know the environment while in the incomplete information setting, agents may have private information so that an agent may not know the true environment either. In our setting, all agents including the planner know the technology while the preferences of an agent is known to all other agents excluding the planner: thus, the setting is one of complete information. The fact that agents know each other’s preferences is used in a crucial way in our mechanisms.

In the mechanism \(\Gamma_m\), agents move sequentially and agent \(k\), when it is her turn to move, chooses a tuple \(z_k = (y_k, x_k)\) where \(y_k\) is the vector of public goods she wishes to consume and \(x_k\) is her contribution toward the cost of production conditional on her demands being met. We assume that the agents move in the “natural order,” viz., first, agent 1 moves, then agent 2, and so on. The analysis can easily be reinterpreted to cover the case where the agents move in a different order.

Let \(S = \{i_1, \ldots, i_{|S|}\} \subset N\). Given a collection \((y^k)_{k \in S}; y^k \in Y\) for \(k \in S\), we denote by \(\hat{y}_S\) the vector \(y^{i_1} \lor \ldots \lor y^{i_{|S|}}\).

**Definition 1** The coalition \(S\) is compatible with the announcement \(z = \{(y^k, x_k)\}_{k \in N}\) if \(\sum_{k \in S} x_k \geq c(\hat{y}_S)\).

Thus, a coalition \(S\) is compatible if the total contributions of the members of \(S\) is sufficient to produce a vector of public goods such that the demands of all members of \(S\) can be satisfied.

Once all players have announced, the planner selects the largest compatible coalition in the set \(\{S_1, \ldots, S_{n+1}\}\) which is called the maximal compatible coalition and denoted \(S^*(z)\).\(^3\)

He then announces that the vector \(y^*(z) = \hat{y}_{S^*(z)}\) is to be produced (obviously, \(y^*(z) = 0\) if

\(^3\)Since \(S_{k+1} \subset S_k\) for all \(k = 1, \ldots, n\) and \(S_{n+1} = \emptyset\) is compatible by definition, the maximal compatible coalition always exists.
\( S^*(z) = \emptyset \) and charges the agents as follows:

\[
x^*_k(z) = \begin{cases} 
  x_k & \text{if } k \in S^*(z), \\
  0 & \text{otherwise}.
\end{cases}
\]

The mechanism \( \Gamma_m^1 \) is a two-stage game. In the first stage, agents play the game \( \Gamma_m \) according to a randomly selected order of the agents. At the end of the first stage, each agent is asked whether she would like to replay the game. If all agents answer "NO," then the game ends. If at least one agent answers "YES," then an order of the agents is chosen at random from the \( n! \) possibilities with equal probability for each order and the agents replay the game \( \Gamma_m \) according to this randomly chosen order. The game ends after the optional second stage.

4 The Results

In this section we show that every subgame perfect equilibrium (hereafter, SPE) of the mechanisms \( \Gamma_m \) and \( \Gamma_m^1 \) results in the production of an efficient vector of public goods. For the mechanism \( \Gamma_m^1 \) we also show that the equilibrium payoffs correspond to the Shapley value of a cooperative game that summarizes the production opportunities of coalitions in the economy.

We start with the following preliminaries. Consider the optimization problem for any \( S \subset N \):

\[
\begin{align*}
\max & \quad \sum_{k \in S} v_k(y) - c(y) \\
\text{subject to} & \quad c(y) \leq \sum_{i \in S} w_i, \quad y \in Y.
\end{align*}
\]

Assumption 1 ensures that this problem is well-defined. Let \( Y^* = \{ y \in Y | 0 \leq c(y) \leq \sum_{i \in S} w_i \} \). Since \( c \) is continuous and \( c(y) \to \infty \) as \( y \to \infty \), it follows that \( Y^* \) is both closed and bounded and therefore, compact.\(^4\) The optimization problem (1) thus reduces to one where the maximum is taken over all \( y \in Y^* \). Since the functions \( c, v_1, \ldots, v_n \) are continuous and \( Y^* \) is compact, it follows that the maximum exists. We shall use the notation \( y_S^* \) to signify an efficient vector of public goods for \( S \), that is, \( y_S^* \) solves (1). Of course, \( y_S^* \) need not be unique.

Definition 2 The stand alone payoff to a coalition \( S \) is defined as

\[
sa(S) = \begin{cases} 
  \sum_{k \in S} v_k(y_S^*) - c(y_S^*) & \text{if } S \neq \emptyset, \\
  0 & \text{otherwise}.
\end{cases}
\]

\(^4\)Note also that since \( c(0) = 0, \emptyset \in Y^* \).
The stand-alone payoff summarizes the aggregate payoff possibilities for each coalition within the economy. We therefore have a natural associated cooperative game for our economy, given by \((N, sa)\).

**Definition 3** The “marginal contribution” of agent \(k, k = 1, \ldots, n\) is \(u^*_k = sa(S_k) - sa(S_{k+1})\).

**Remark 3** Assumption 1 implies that \(0 \leq sa(S) \leq sa(T)\) if \(S \subset T\). The coalition \(T\) can always produce \(y^*_S\), and thus by Assumption 1, \(sa(T) \geq \sum_{i \in T} v_i(y^*_S) - c(y^*_S) \geq sa(S)\). The first inequality follows because \(0\) is a feasible production plan for any coalition and therefore, using Assumption 1 once more, \(sa(S) \geq \sum_{i \in S} v_i(0) - c(0) \geq 0\).

The main results concerning the mechanisms \(\Gamma_m\) and \(\Gamma^1_m\) can be summarized in the following theorems.

**Theorem 1** In all SPE of \(\Gamma_m\), an efficient vector of public goods is produced. Also, if \(y^*\) is an efficient vector of public goods, then there exists a SPE which supports \(y^*\) as an equilibrium outcome. Furthermore, the net payoffs of the agents in all SPE are uniquely given by \((u^*_1, \ldots, u^*_n)\).

**Theorem 2** In all SPE of \(\Gamma^1_m\), an efficient vector of public goods is produced. Also, if \(y^*\) is an efficient vector of public goods, then there exists a SPE which supports \(y^*\) as an equilibrium outcome. Furthermore, the net payoffs to the agents in all SPE are given uniquely by the Shapley value of \((N, sa)\).

Let us first consider the mechanism \(\Gamma_m\). Suppose that agents in \(\{1, \ldots, j\}\) have already announced \(\{z_k = (y^k, x_k)\}_{k=1}^j\). Let \(i \in \{1, \ldots, j\}\) be such that

\[
\sum_{k=j+1}^n v_k(\tilde{y}^j) - c(\tilde{y}^j) + \sum_{k=i}^j x_k > sa(S_{j+1})
\]  

(2)

for some \(\tilde{y}^j\) such that \(\tilde{y}^j \geq y^k\) for all \(k = i, \ldots, j\).

To understand what’s being said here, note that agents in \(S_{j+1}\) can expect a collective net payoff of at most \(sa(S_{j+1})\) without the cooperation of other agents. If an \(i \leq j\) satisfying (2) exists, then it means that the agents in \(S_{j+1}\) can obtain a collective payoff strictly greater than \(sa(S_{j+1})\) by cooperating with agents in \(\{i, i + 1, \ldots, j\}\).

Let \(i_j\) be the largest integer in \(\{1, \ldots, j\}\) satisfying (2). Note that \(i_j\) need not exist at all. However, if \(i_j\) exists, then \(S_{i_j}\) is the *smallest* compatible coalition ensuring a collective

\[\text{Compatibility requires that if any of the agents preceding } j + 1 \text{ are part of the maximal compatible coalition, then they necessarily must form a “connected” set of agents. Recall also that the only compatible coalitions considered are those in } \{S_1, \ldots, S_n, S_{n+1}\}.\]
payoff greater than \( sa(S_{j+1}) \) to the agents following \( j \). In other words, if the resulting maximal compatible coalition in some SPE of the subgame following \( j \)'s announcement is not a superset of \( S_{ij} \), then the agents in \( S_{j+1} \) must be receiving a collective net payoff less than or equal to \( sa(S_{j+1}) \).

The following lemma says that if \( i_j \) exists, then all SPE of the subgame following \( j \)'s announcement will be such that the resulting maximal compatible coalition is a superset of \( S_{ij} \).

**Lemma 1** Suppose that agents in \( \{1, \ldots, j\} \) have announced \( \{(y^k, x_k)\}_{k=1}^j \) in the mechanism \( \Gamma_m \) and that \( i_j \) exists. Then, all SPE of the subgame following \( j \)'s announcement will be such that the resulting maximal compatible coalition is a superset of \( S_{ij} \).

**Proof:** We proceed by induction on \( j \). We will show that if the maximal compatible coalition in some SPE of the subgame following \( j \)'s announcement is not a superset of \( S_{ij} \), then there is a profitable deviation for some agent in \( S_{j+1} \).

If \( j = n \), then there is no subgame following \( n \)'s announcement. Then, (2) implies that \( S_{in} \) is a compatible coalition Therefore, it follows that the resulting maximal compatible coalition is a superset of \( S_{in} \).

Assume that the lemma is true for all \( j > J \). Suppose \( i_j \) exists but the maximal compatible coalition is not a superset of \( S_{ij} \) in some SPE of the subgame following \( J \)'s announcement. Let \( (u_{j+1}, \ldots, u_n) \) be the corresponding net payoffs to the agents following agent \( J \). Since the maximal compatible coalition is not a superset of \( S_{ij} \), we must have \( \sum_{k=j+1}^n u_k \leq sa(S_{j+1}) \).

We can distinguish between two cases.

**Case 1:** There exists \( k > J \) such that \( u_k < u_k^* \).

Let agent \( k \) deviate by announcing \((y_{S_k}^*, x_k')\) where\(^7\)

\[
sa(S_{k+1}) - \sum_{i=k+1}^n v_i(y_{S_k}^*) + c(y_{S_k}^*) < x_k' < v_k(y_{S_k}^*) - u_k.
\] (3)

The upper bound on \( x_k' \) is strictly greater than the lower bound if and only if

\[
v_k(y_{S_k}^*) - u_k > sa(S_{k+1}) - \sum_{i=k+1}^n v_i(y_{S_k}^*) + c(y_{S_k}^*). \tag{4}
\]

Since \( \sum_{i=k}^n v_i(y_{S_k}^*) - c(y_{S_k}^*) = sa(S_k) \), it follows that (4) is true if and only if \( u_k < sa(S_k) - sa(S_{k+1}) = u_k^* \). This is true by assumption. Thus, a value of \( x_k' \) satisfying (3) exists.

Since \( \sum_{i=k+1}^n v_i(y_{S_k}^*) - c(y_{S_k}^*) + x_k' > sa(S_{k+1}) \), it follows that \( i_k = k \) after \( k \)'s deviation.

Since \( k > J \), the induction hypothesis implies that the resulting maximal compatible coalition in any SPE of the subgame following \( k \)'s deviation will be a superset of \( S_{ik} = S_k \). By

\(^6\)Remember that agents are announcing in the order 1, 2, \ldots, \( n \) so that \( n \) is the last to announce.

\(^7\)Recall that for any \( S \subset N \), \( y_S^* \) denotes an efficient vector of public goods for \( S \).
compatibility, the resulting vector of public goods, say $y$, will be such that $y \geq y^*_k$. By Assumption 1, we have $v_k(y) - x_k \geq v_k(y^*_k) - x'_k > u_k$. Thus, $k$ has a profitable deviation.

**Case 2:** $u_k = u^*_k$ for all $k = J + 1, \ldots, n$.

Let agent $J + 1$ deviate by announcing $(\tilde{y}^J, x'_{J+1})$ where $x'_{J+1}$ satisfies the following conditions.\(^8\)

1. $sa(S_{J+2}) - \sum_{i=J+2}^n v_i(\tilde{y}^J) + c(\tilde{y}^J) - \sum_{i=J}^J x_i < x'_{J+1}$,
2. $x'_{J+1} < v_{J+1}(\tilde{y}^J) - sa(S_{J+1}) + sa(S_{J+2})$.

The upper bound on $x'_{J+1}$ is strictly greater than the lower bound if and only if

$$\sum_{i=J+1}^n v_i(\tilde{y}^J) - c(\tilde{y}^J) + \sum_{i=J}^J x_i \geq sa(S_{J+1})$$

which is true because $\tilde{y}^J$ satisfies (2).

Since $\sum_{i=J+2}^n v_i(\tilde{y}^J) - c(\tilde{y}^J) + \sum_{i=J}^J x_i + x'_{J+1} > sa(S_{J+2})$, it follows that $i_{J+1}$ exists after $J + 1$’s deviation. The induction hypothesis implies that the resulting maximal compatible coalition in any SPE of the subgame following $J + 1$’s deviation will be a superset of $S_{i_{J+1}}$. Since $i_{J+1} \leq J + 1$, it follows that $J + 1$ is a member of the resulting maximal compatible coalition. If $y$ is the resulting vector of public goods, then compatibility implies that $y \geq \tilde{y}^J$.

By Assumption 1, we have $v_{J+1}(y) - x_{J+1}' \geq v_{J+1}(\tilde{y}^J) - x'_{J+1} > sa(S_{J+1}) - sa(S_{J+2}) = u^*_{J+1}$. Thus, $J + 1$ has a profitable deviation. This completes the proof of the lemma. ■

**Corollary 1** Let $(u_1, \ldots, u_n)$ be the net payoffs to the agents in some SPE of $\Gamma_m$. Then, $u_k \geq u^*_k$ for all $k \in N$.

**Proof:** If $u_k < u^*_k$ then agent $k$ can deviate by announcing $(y^*_k, x_k)$ where $x_k$ satisfies

$$sa(S_{k+1}) - \sum_{j=k+1}^n v_j(y^*_k) + c(y^*_k) < x_k \leq v_k(y^*_k) - u_k.$$  \hspace{1cm} (6)

It can be checked that a value of $x_k$ satisfying (6) exists. Since $\sum_{j=k+1}^n v_j(y^*_k) - c(y^*_k) + x_k > sa(S_{k+1})$, it follows that $i_k = k$ after $k$’s announcement. By Lemma 1, the maximal compatible coalition in any SPE of the game following $k$’s deviation must be a superset of $S_k$, and therefore, include $k$. If $y$ is the resulting vector of public goods, then by compatibility, $y \geq y^*_k$. Thus, by Assumption 1, we have $v_k(y) - x_k \geq v_k(y^*_k) - x_k > u_k$. This shows that $k$ has a profitable deviation, a contradiction. ■

We are now in a position to prove Theorems 1 and 2.

---

\(^8\)Recall that $\tilde{y}^J$ is a vector of public goods satisfying (2) and such that $\tilde{y}^J \geq y^k$ for $k = i_J, \ldots, J$. Such a $\tilde{y}^J$ exists because $i_J$ exists by assumption.
Proof of Theorem 1: Let \((u_1, \ldots, u_n)\) be the net payoffs to the agents in some SPE of the game \(\Gamma_m\). Let \(y^*\) be the corresponding vector of public goods produced. By Corollary 1, \(u_i \geq u_i^*\) for all \(i \in N\) in any SPE of \(\Gamma_m\). If \(u_i > u_i^*\) for some \(i\) then we must have \(\sum_{j=1}^n u_i = \sum_{j=1}^n v_j(y^*) - c(y^*) > sa(N)\). However, Remark 3 which follows from Assumption 1 shows that the maximum possible stand-alone payoff is \(sa(N)\). The contradiction shows that \(u_i = u_i^*\) for all \(i \in N\). The fact that the efficient vector of public goods is produced follows trivially from the observation that \(\sum_{i=1}^n u_i^* = sa(N)\).

Suppose now that \(y^*\) is an efficient vector of public goods. Consider the strategy profile \(z\) where \(z_i = (y^*, x_i^*)\), \(x_i^* = v_i(y^*) - u_i^*\) for all \(i = 1, \ldots, n\). We claim that \(z\) is a SPE of \(\Gamma_m\). Suppose not. Then, there exists \(j\) such that any agent \(i\) prior to \(j\) announces \(z_i\) and \(j\) announces \(z_j^* \neq z_j\). Following \(j\)'s announcement, we have a subgame. Let \((v_1, \ldots, v_n)\) be the net payoffs of the agents in the equilibrium resulting from this subgame. We can replicate the arguments in Lemma 1 to conclude that \(v_k \geq u_k^*\) for all \(k > j\). If \(v_j < u_j^*\), then clearly, \(j\) has not benefited from the deviation. So suppose that \(v_j > u_j^*\). Then, we have \(\sum_{j=0}^n v_j > sa(S_j)\). Now, if the maximal compatible coalition is also \(S_j\), then we have a contradiction. So, suppose that the maximal compatible coalition is \(S_t\) for some \(t < j\). Note that \(v_l \geq u_l^*\) for all \(l \leq l < j\). Thus, we have \(\sum_{l=1}^n v_l > sa(S_l)\) which is once again, a contradiction. This completes the proof of the theorem. ■

Proof of Theorem 2: Let \((\phi_1(sa), \ldots, \phi_n(sa))\) denote the Shapley value payoffs to the agents in \((N, sa)\). Suppose the game enters Stage 2 of \(\Gamma_m^1\). We know from Theorem 1 that for any order selected by the planner, the SPE payoffs will be given by the “marginal contribution” vector. Since each order is equally likely, it follows that the expected payoff to any agent at the beginning of the second stage is exactly his Shapley value in the game \((N, sa)\).

Consider now the agents’ decisions at the beginning of Stage 1. If any agent gets less than her Shapley value payoff at the end of Stage 1, then she will force the game into the second stage. It thus follows that the strategy profile \(\{z_i^* = (y_N^*, x_i^*)\}, \text{“NO”}\}_{i=1}^n\) where \(x_i^* = v_i(y_N^*) - \phi_i(sa)\) constitutes a SPE of the game \(\Gamma_m^1\). However, this may not be a unique SPE as it is possible that some agent is indifferent between the game ending in the first stage and getting his Shapley value payoff and getting the same payoff in expected terms in Stage 2.10 If we assume that all agents have a lexicographic preference for the game ending in the first stage, then all SPE of the game \(\Gamma_m^1\) will end in the first stage with the production of an efficient vector of public goods, and the agents getting their Shapley value payoffs. ■

Remark 4 If agents discount the future instead of having a lexicographic preference for the game ending in stage 1, then the game \(\Gamma_m^1\) always ends in the first stage and the payoffs approach the Shapley value as the discount factor approaches one.

---

9 The vector of public goods produced must be \(y \geq y^*\) since all members in \(\{t, \ldots, j - 1\}\) announce \(y^*\) as the desired vector of public goods. Since the utility functions are non-decreasing, \(v_t(y) - x_t \geq v_t(y^*) - x_t = u_t^*\) for all \(t = t, \ldots, j - 1\).

10 Note that the efficient vector of public goods is produced in either case.
Let \( 0 < \beta < 1 \) be the discount factor. If the game \( \Gamma_m^1 \) reaches the second stage, then the expected payoffs to the agents at the beginning of Stage 2 are obviously \( \beta(\phi_1(sa), \ldots, \phi_n(sa)) \). Thus, if \( i \)'s payoff at the end of stage 1 is strictly less than \( \beta \phi_i(sa) \), then she will move the game to the second stage. Therefore, the optimal strategy for player \( i \) in stage 1 is to announce so as to leave exactly the second stage payoffs \( (\sum_{j=i+1}^n \beta \phi_j(sa)) \) to the agents following her. Using the argument recursively, it follows that agent 1 will expropriate the entire surplus that accrues on account of time discounting. Note that the game cannot go to the second stage because agent 1 would prefer to concede a little to the agents following her rather than having the game go to the second stage.\(^{11} \) Therefore, with discounting, the game \( \Gamma_m^1 \) always ends in the first stage with production of an efficient vector of public goods and net payoffs \( (sa(N) - \beta \sum_{j=2}^n \phi_j(sa), \beta \phi_2(sa), \ldots, \beta \phi_n(sa)) \). Thus, as \( \beta \to 1 \), the payoffs converge to the Shapley value.

**Remark 5** The symmetry property of the Shapley value implies that if two agents have the same preferences, then they receive the same net utility from the mechanism \( \Gamma_m^1 \). More generally, the mechanism \( \Gamma_m^1 \) has the following "monotonicity property": if agents \( i \) and \( j \) are such that \( i \)'s marginal utility at all points is at least as high as \( j \)'s marginal utility, then \( i \)'s net payoff must be at least as high as \( j \)'s net payoff. It is easy to check that if \( i \)'s marginal utility is higher than \( j \)'s marginal utility at all points, then this implies that \( sa(S \cup \{i\}) - sa(S) \geq sa(S \cup \{j\}) - sa(S) \) for all \( S \subset N \setminus \{i, j\} \). A simple computation using the formula for the Shapley value then implies that \( i \)'s payoff is at least as large as \( j \)'s payoff.

## 5 An Example

We now consider an example due to Moulin (1995) which illustrates the operation of the two mechanisms considered above. Let \( N = \{1, 2, 3\} \) and let the set of public goods be given by \( \{a, b, c\} \). Following Moulin (1995), we can interpret the public goods as being "street-lights." Let \( a \) be the street-light on the road between 1 and 2, \( b \) be the street-light on the road between 1 and 3, and \( c \) the street-light on the road between 2 and 3. We shall say that the street-light \( x \) is adjacent to agent \( i \) if \( x \) is located on a road connecting \( i \) and some other agent \( j \). The utility function of agent \( i \) is given as follows:

\[
U_i = \begin{cases} 
30 & \text{if there is one street-light adjacent to } i, \\
45 & \text{if there are two street-lights adjacent to } i, \\
0 & \text{otherwise.}
\end{cases}
\]

\(^{11} \)If the game does go to a second stage, then agent 1 can deviate by conceding a fraction \( 0 < \alpha < 1 \) of the surplus \( (S = sa(N) - \beta \sum_{j=2}^n \phi_j(sa)) \) to agents following her, agent 2 can follow by conceding some of \((1 - \alpha)S\) to the subsequent agents and so on. This ensures that the game ends in the first stage and agent 1 is strictly better off from the deviation. This argument has been used before in Bag and Winter (1999).
The cost of production is a uniform $40 per street-light. It is easily seen that in this example \( sa\{i\} = 0, sa\{i, j\} = 20 \) and \( sa(N) = 25 \). This follows because the optimal action for an individual agent is not to construct any street-light; for any two agents to construct the street-light on the road connecting them; and for the grand coalition to construct any two street-lights. Observe that the optimal vector of public goods is not uniquely defined for the grand coalition.

Consider the mechanism \( \Gamma_m \) where the agents move in the order 1, 2, 3. Using Theorem 1, it follows that the equilibrium net payoffs to the agents are \((u^*_1, u^*_2, u^*_3) = (5, 20, 0)\). Similarly, using Theorem 2, it follows that the equilibrium net payoffs in the game \( \Gamma^1_m \) are given by \((v^*_1, v^*_2, v^*_3) = (8 \frac{1}{3}, 8 \frac{2}{3}, 8 \frac{1}{3})\). Note that the equilibrium strategies cannot be uniquely specified in the two mechanisms. For instance, in the mechanism \( \Gamma_m \), the strategy profiles \( \{(a, b, c), 40\}, \{(a, b, c), 10\}, \{(a, b, c), 30\} \) and \( \{(a, c, 25\}, \{(a, c, 25\}, \{(a, c, 30\} \) are both equilibrium strategy profiles. Similarly, in the mechanism \( \Gamma^1_m \), the strategy profiles \( \{(a, b, 36 \frac{2}{3}\}, \{(a, b, 21 \frac{2}{3}\}, \{(a, b, 21 \frac{2}{3}\} \) and \( \{(a, c, 21 \frac{2}{3}\}, \{(a, c, 36 \frac{2}{3}\}, \{(a, c, 21 \frac{2}{3}\} \) are both equilibrium strategy profiles. It is a feature of both the \( \Gamma_m \) and \( \Gamma^1_m \) mechanisms that even though the equilibrium strategies cannot be specified uniquely, the equilibrium payoffs are nonetheless unique. Furthermore, the vector of public goods produced in equilibrium is always optimal.

Note also that the core of the game \((N, sa)\) is empty. This brings forward an important point: the mechanisms that we have proposed may not be immune to coalitional deviations. To see this more clearly, imagine that the public goods in this example are excludable and that the technology is freely available to all coalitions. Observe that the coalition \{1, 3\} collectively gets a payoff of 5 in any SPE of \( \Gamma_m \) (assuming that agents move in the order 1, 2, 3). However, the coalition \{1, 3\} can do better by opting to produce the good \( b \) by themselves and sharing the cost equally. This leaves both 1 and 3 with a surplus of 10 which is more than what they get in any SPE of \( \Gamma_m \). Similarly, in the mechanism \( \Gamma^1_m \), any pair of agents can deviate because the total payoff to \( S = \{i, j\} \) in any SPE of \( \Gamma^1_m \) is \( 16 \frac{2}{3} < sa(S) = 20 \).

In the following section we discuss conditions under which the mechanisms \( \Gamma_m \) and \( \Gamma^1_m \) are also coalition stable.

### 6 Coalition Stability

We confine our discussion of coalition stability of our mechanisms to the case where all public goods are excludable. The reason for this restriction is that it is only in this case that coalition stability is well-defined through the notion of stand alone core. This is defined as follows. First, note that an allocation for an economy with multiple excludable public goods is \((y, x_1, \ldots, x_n)\) (where \( y \) is a vector of public goods and \( x_i \) the contribution of agent \( i \)) satisfying \( c(y) \leq \sum_{i=1}^{n} x_i \).

---

12These two strategy profiles do not exhaust the set of equilibrium strategies in this example.
Definition 4 An allocation \((y, x_1, \ldots, x_n)\) is in the stand alone core of the excludable public goods economy if there does not exist \(\emptyset \neq S \subset N\), \((y', (x'_i)_{i \in S})\) satisfying

1. \(c(y') \leq \sum_{i \in S} x'_i\),

2. \(v_i(y') - x'_i \geq v_i(y) - x_i\) for all \(i \in S\) with a strict inequality for at least one \(i \in S\).

Remark 6 It is easily confirmed that an allocation \((y, x_1, \ldots, x_n)\) is in the stand alone core of the excludable good economy if and only if it is in the core of the TU-game \((N, sa)\). The stand-alone core is equivalent to the well-known notion of the \(\alpha\)-core. In our context, it is also equivalent to the \(\beta\)-core.\(^{13}\)

The definition of the stand-alone core implies that a deviating coalition \(S\) does not have access to the public goods produced by \(N \setminus S\). We argue that this assumption is reasonable because exclusion serves as a threat against free-riding and as assumed, costs nothing to the excluding coalition. Our definition would make less sense if there are pure public goods in the economy because a deviating coalition cannot be excluded from the set of pure public goods produced by others. In the pure public good case, Definition 4 implies that in the event of a deviation by the coalition \(S\), the coalition \(N \setminus S\) will not produce any public goods whatsoever. This is obviously a strong requirement for this framework. Carraro and Siniscalco (1993) and Chander and Tulkens (1997) have proposed coalition stability concepts based on weaker behavioral assumptions in the context of environmental externalities but we choose to restrict ourselves to the excludable public goods case where the notion of the \(\alpha\)-core seems quite intuitive.

In the case of a single excludable public good economy, Bag and Winter (1999) and Moulin (1995) show that the stand alone core is non-empty under weak assumptions. Indeed, it turns out that the TU-game \((N, sa)\) is convex in this case. The mechanisms proposed by Bag and Winter (1999) for the single excludable public good economy are coalition stable because they implement payoff vectors which are in the stand alone core. In the multiple excludable public goods case, the stand alone core may be empty as the example in the previous section demonstrates. Moulin (1995) gives sufficient conditions for the convexity of the TU-game \((N, sa)\) which ensures that the stand alone core is non-empty. We turn to these conditions now.

Definition 5 A function \(f : \mathbb{R}^P \to \mathbb{R}\) is supermodular [submodular] if for all \(y, y'\), \(f(y \lor y') + f(y \land y') \geq \leq \lceil f(y) + f(y')\).

The proof of the following lemma is straightforward and is omitted.

Lemma 2 Suppose that \(v_i\) satisfies the conditions in Assumption 1 and is also supermodular for all \(i = 1, \ldots, n\). Also, let the cost function \(c\) satisfy the conditions in Assumption 1 as well as submodularity. Then, the game \((N, sa)\) is convex.

\(^{13}\)See Aumann (1959) for precise definitions of the \(\alpha\)-core and \(\beta\)-core.
This now leads to the following result.

**Theorem 3** Suppose that \( v_i \) is supermodular for \( i = 1, \ldots, n \) and \( c \) is submodular. Then, the mechanisms \( \Gamma_m \) and \( \Gamma_m^1 \) are coalition stable.

**Proof:** By Remark 6 it suffices to show that the allocations achieved by \( \Gamma_m \) and \( \Gamma_m^1 \) are in the core of \((N, sa)\). By Lemma 2, \((N, sa)\) is convex. The coalition stability of \( \Gamma_m \) thus follows from the observation following the proof of Lemma 2. Since the Shapley value is simply a convex combination of the “marginal contribution” vectors, the coalition stability of \( \Gamma_m^1 \) follows also. ■

**Remark 7** Observe that supermodularity and submodularity impose no restrictions when there is only one public good. Lemma 2 thus confirms the result of Moulin (1995) that in the case of a single excludable public good, the game \((N, sa)\) is convex under weak assumptions.

**Remark 8** Topkis (1978) has shown that if a function is smooth, then supermodularity [submodularity] is equivalent to the condition that all cross-partial are non-negative [non-positive]. Thus, on the preference side, supermodularity implies complementarity between the public goods in the sense that the marginal utility from a public good does not decrease when the amount of some other public good increases. A number of well-known utility functions are supermodular; for instance, the Cobb-Douglas, separable, and Leontief functions are all supermodular.

Submodularity of the cost function suggests a similar complementarity: the marginal cost of a public good does not increase when the quantity produced of some other public good increases. Such complementarity may not be unrealistic: an example of such a scenario is when one (excludable) public good is public education and the other is a direct intervention to reduce poverty. Examples of submodular cost functions include \( c(y) = \sum_{i=1}^{p} q_i y_i \) (where \( q_i \) is the unit cost of good \( i \)) and \( c(y) = \ln(1 + \sum_{i=1}^{p} y_i) \).\(^{14}\)

### 7 Conclusion

In concluding, we discuss the sensitivity of our results to the assumptions on technology and preferences. As can be seen, the assumptions on technology are weak. On the other hand, the assumption of quasi-linear preferences is a fairly strong – though standard – restriction and one can ask how crucial this is to the results obtained in the paper. We conjecture that techniques similar to those used in proving Theorem 1 can be used to construct efficient mechanisms when preferences are not quasi-linear but these will not be order-independent. In the quasi-linear case, the Pareto frontier is affine and hence, a convex combination of points on the Pareto frontier is still Pareto efficient. When preferences are not quasi-linear,

---

\(^{14}\)Note that \( c(y) = \sum_{i=1}^{p} q_i y_i \) is only weakly submodular while \( c(y) = \ln(1 + \sum_{i=1}^{p} y_i) \) is strictly submodular. Also, observe that the former cost function is a “natural” specification in many contexts.
a convex combination of points on the Pareto frontier can be inefficient, and therefore, the
randomization technique used to generate order-independent and efficient allocations in the
mechanism $\Gamma^m$ will not work. The following example illustrates this point.

**Example 1** Let $N = \{1, 2\}$ and $p = 1$. Let $u_1(x, y) = xy, u_2(x, y) = x + y, w_1 = w_2 = 1$
and $c(y) = y$. The stand-alone utilities can be computed to be $u^*_1 = 1/4, u^*_2 = 1$. In the
mechanism $\Gamma_m$, when the agents announce in the order $1, 2$, the SPE involves producing 1
unit of the public good, 1 contributing 0 and 2 contributing 1 and the net payoffs to the
agents are $(u_1, u_2) = (1, 1)$. When the agents announce in the order $2, 1$, the SPE involves
producing $(2 + \sqrt{3})/2$ units of the public good, 2 contributing 1, 1 contributing $\sqrt{3}/2$ with
net payoffs given by $(u_1, u_2) = (1/4, (2 + \sqrt{3})/2)$. The average payoff across the two orders
is thus $(\bar{u}_1, \bar{u}_2) = (5/8, (4 + \sqrt{3})/4)$, which is not efficient.\textsuperscript{15}

If the two agents do play the mechanism $\Gamma^1_m$, then the order in which the agents announce
in stage 1 is still important. Consider what happens when agents announce in the order $2, 1$.
If agent 2 wants the game to end in stage 1 itself, then she has to ensure agent 1 a payoff of
$5/8$. The optimal strategy involves agent 2 announcing $((4 + \sqrt{6})/4, 1)$, agent 1 contributing
$(4 - \sqrt{6})/4$ which gives rise to the payoff vector $(u_1, u_2) = (5/8, (4 + \sqrt{6})/4)$. On the other
hand, when the agents announce in the order $1, 2$, the optimal strategies involve agent 1
announcing $((4 + \sqrt{3})/4, \sqrt{3}/4)$, agent 2 contributing 1 which give corresponding payoffs of
$(u_1, u_2) = (13/16, (4 + \sqrt{3})/4)$.

Example 1 suggests that the randomization technique will typically not give rise to order-
independent outcomes when preferences are not quasi-linear. In such a situation, one can use
more complicated and abstract mechanisms like those proposed by Abreu and Sen (1990),
Moore and Repullo (1988) and Maniquet (1999). Whether there exist simple mechanisms –
similar to those proposed in this paper – which will work when preferences are not quasi-
linear is an open question.

**References**


A., Luce, D. (Eds), Contributions to the Theory of Games IV, Volume 40 of Annals of

\textsuperscript{15}To see this, consider the allocation $(x_1, x_2, y) = ((4 - \sqrt{3})/4, 0, (4 + \sqrt{3})/4)$. The reader can check that
this is a feasible allocation and gives rise to the utility vector $(u_1, u_2) = (13/16, (4 + \sqrt{3})/4)$. This weakly
dominates $(\bar{u}_1, \bar{u}_2)$ but one can easily modify the allocation to make both agents strictly better off relative
to $(\bar{u}_1, \bar{u}_2)$. 15


