Available online at www.sciencedirect.com
SCIENCE DIRECT®
Games and Economic Behavior 48 (2004) 111–123
GAMES and Economic Behavior
www.elsevier.com/locate/geb

Bidding for the surplus: realizing efficient outcomes in economic environments

Suresh Mutuswami,[a] David Pérez-Castrillo,[b,*] and David Wettstein[c]

[a] *Department of Economics, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, UK*
[b] *Dep. Economia e Historia Economica and CODE, Universitat Autònoma de Barcelona, Edifici B, 08193 Bellaterra (Barcelona), Spain*
[c] *Department of Economics, Monaster Center for Economic Research, Ben-Gurion University of the Negev, Beersheva 84105, Israel*

Received 7 March 2002

Available online 6 November 2003

Abstract

In this paper, we consider a local public goods environment. The agents are faced with the task of providing local public goods that will benefit some or all of them. We propose a bidding mechanism whereby agents bid for the right to decide upon the organization of the economic activity. The subgame perfect equilibria of the mechanism generate efficient outcomes. We also show how to adapt the mechanism to network economies where the economic activity takes place via the formation of links.

JEL classification: H41; C72; C71

Keywords: Implementation; Local public goods; Bidding; Networks; Shapley value

1. Introduction

A substantial amount of economic activity takes place in settings that are very different from the perfect competition model. A natural concern for such environments is the attainment of efficient outcomes, despite the difficulties stemming from strategic behavior

* Corresponding author.
E-mail addresses: smutus@essex.ac.uk (S. Mutuswami), david.perez@uab.es (D. Pérez-Castrillo), wettstn@bgumail.bgu.ac.il (D. Wettstein).

0899-8256/$ – see front matter
doi:10.1016/j.geb.2003.08.003

and the market power participating agents have. In this paper we consider local public goods environments and resolve the problem of achieving efficiency by constructing a bidding mechanism whose equilibria generate efficient outcomes.

The first stage of the mechanism suggested is a bidding game. The winner of that bidding obtains the right to be the first agent to suggest a coalition and a production plan of local public goods for that coalition. He also offers a vector of transfers to induce all the agents to accept his plan. If the agents accept his proposal, it is carried out. The agents outside the proposed coalition will go on organizing their activity once again using this mechanism. In case the offer is refused, the winner is removed from the game and left on his own. The remaining agents use the same mechanism again, starting with the bidding stage.

The subgame perfect equilibria of the mechanism generate efficient outcomes. The mechanism thus provides a noncooperative framework that allows the agents to sequentially form an efficient coalition structure. Moreover, at equilibrium the coalitions produce and consume efficient levels of local public goods. We also show that the payoffs received by the agents coincide with their Shapley values in an appropriately defined cooperative game. Hence the surplus generated by the agents is shared in a fair and consistent manner.

Further analysis shows that the equilibrium bids satisfy strong stability conditions: they are maxmin strategies and are also robust to deviations by groups of agents. Hence, the equilibrium bids are "very safe" strategies: they guarantee any agent (at least) the equilibrium payoff independently of the bids made by the other agents, furthermore no agent can be hurt by a collusive (coordinated) change of bids made by a group of other agents. These stability properties, which are lacking in most mechanisms suggested to date, would play an important role in the stage of choosing a mechanism and lend further support to the adoption of this mechanism to coordinate the activities of the agents.

The mechanism constructed is related to two mechanisms which have been recently proposed, one to implement the Shapley value in general Transferable Utility games (Pérez-Castrillo and Wettstein, 2001) and the other to achieve efficient outcomes in a general social choice framework (Pérez-Castrillo and Wettstein, 2002). The bidding stage in the Shapley value mechanism, while leading to easily interpretable bids (the bids in equilibrium represented each agent's contribution to the Shapley value of the other agents), was more complex. This, in part, implied that there was no lower bound on an agent's welfare in an out-of-equilibrium outcome. The mechanism constructed to resolve general social choice problems consisted of one stage and while leading to an efficient outcome, as well as being immune to coalitional deviations, failed to provide a unique way of sharing the surplus created, due to an inherent multiplicity of equilibria.

The present mechanism is a hybrid construction of the two, with the bidding stage borrowed from Pérez-Castrillo and Wettstein (2002) (see also Dubins, 1977) and the post-bidding stages taken from Pérez-Castrillo and Wettstein (2001). This "merger" leads to a mechanism with more appealing stability properties that provides, basically, a unique answer to the problem of creating and then sharing a surplus in a specific economic environment, namely a local public goods economy.

Several previous papers were concerned with achieving efficiency in environments similar to ours, see Bagnoli and Lipman (1989), Jackson and Moulin (1992), Bag and Winter (1999), and Mutuswami and Winter (2003). Our mechanism improves upon these constructions in allowing for a more general class of environments, and dispensing with the need for a designer. Furthermore, the agents always receive payoffs corresponding to an appropriately defined Shapley value.

Sequential mechanisms that were constructed by Moore and Repullo (1988) and Maniquet (2003) to realize general social choice functions would work for our environments as well. Due to their large scope of coverage they are however more complex than the other mechanisms mentioned.

In the next section we introduce the environment and the mechanism. We then proceed to analyze the equilibrium outcomes of the mechanism, establishing their efficiency and equity properties as well as the strong stability properties enjoyed by the equilibrium strategies. We compare our proposal with previous mechanisms and discuss its advantages.

Finally we indicate how to adapt the sequential mechanism we have proposed to cope with the issue of forming an efficient network and sharing the surplus generated.

2. A local public goods economy

We consider environments with a set $N = \{1, \dots, n\}$ of agents that consume m *local public goods* and one private good. The preferences of the agents are quasi-linear in the private good. Agent i's preferences are given by $U_i(y, S, x_i) = u_i(y, S) - x_i$, where $S \subseteq N$ denotes the coalition to which the agent belongs, $y \in \Re^m_+$ the level of public goods produced by the members of S, and $x_i \in \Re$ agent i's contribution towards the production of the public goods. The local nature of the public good implies that an agent only enjoys the public good produced by a coalition if he belongs to it. That is, each coalition can bar agents not belonging to the coalition from consuming the public goods produced by it. Also, the utility of an agent depends on both the level of public goods and the identity of the partners in the coalition.

The technology is given by a cost function $c(y, S)$, that describes the cost (in terms of the private good) of producing y by the members of S. Hence it could be that different coalitions have different costs for producing identical amounts of public goods. Differences may stem from size or from the availability of different technologies.

We can model the previous environment as a cooperative game with transferable utility. Denote by $w(T)$ the value associated with any coalition $T \subseteq N$:

$$w(T) = \max\left\{\sum_{i \in T} u_i(y, T) - c(y, T) \;\middle|\; y \in \Re^m_+\right\}.$$

The function $w(T)$ thus measures the maximum total surplus that the members of the coalition T can obtain by producing on their own a vector of public goods.[1] The resulting

[1] We assume that preferences and technology are such that $w(T)$ is well-defined for any coalition T. We could, for example, assume that agents' endowments of the private good are finite, costs are strictly convex in y and utility functions are continuous.

cooperative game (N, w) need not be super-additive. We construct the super-additive cover of this cooperative game and denote its characteristic function by $W(T)$ for $T \subseteq N$:

$$W(T) = \max\left\{ \sum_{S_j \in \pi} w(S_j) \,\middle|\, \pi \text{ is a partition of } T \right\}.$$

Note that $W(T)$ is the maximal total payoff that members of T can generate, possibly splitting into several subcoalitions.

An *efficient partition and production plan* for this economy is a feasible allocation that maximizes the total payoff members in N could obtain, by possibly splitting into coalitions that would produce public goods for their members. Therefore, an efficient partition and production plan for N maximizes the following expression:

$$\left\{ \sum_{S_j \in \pi} \left(\sum_{i \in S_j} u_i(y_j, S_j) - c(y_j, S_j) \right) \,\middle|\, \pi \text{ is a partition of } N \text{ and } y_j \in \Re^m_+ \text{ for all } j \right\}.$$

The *Shapley value* of agent $i \in N$ in the cooperative game with the characteristic function $W(S)$ is denoted by $\phi_i(N, W)$, that is:

$$\phi_i(N, W) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n - |S| - 1)!}{n!} \big[W\big(S \cup \{i\}\big) - W(S) \big].$$

Hence, to reach the outcome prescribed by the Shapley value, it must be the case that an efficient partition is formed and the agents produce the corresponding efficient levels of public goods. Moreover, the sharing of the benefits and costs of production is carried out in an equitable manner.

For the local public goods economy, we suggest a mechanism whose subgame perfect equilibrium (SPE) outcomes generate an efficient partition and production plan for N. We view this as an implementation problem in an environment with complete information. The mechanism constructed will in effect implement the Shapley value $\phi(N, W)$ in SPE.

The *bidding mechanism* that we propose for this environment can be considered as a model of coalition formation in which a bargaining process is modeled explicitly. In this mechanism, coalitions are formed sequentially.

An informal description of the bidding mechanism proceeds as follows: In stage 1 the agents bid to choose the proposer. Each agent bids by submitting an n-tuple of numbers (positive or negative), one number for each agent (including himself). The bids submitted by an agent must sum up to zero, hence they reflect a measure of relative preference. The agent for whom the aggregate bid (sum of bids submitted for him by all agents including himself) is the highest, is chosen as the proposer. Before moving to stage 2, all the agents (including the proposer) pay the bid they submitted for the proposer. In the case where the payments do not balance out, that is the aggregate bid to the proposer is positive,[2] these payments generate a positive surplus, which is then shared equally among the agents.

In stage 2 the proposer offers a vector of payments to all other agents, chooses a coalition he wants to form, and proposes a vector of public goods that will be produced and enjoyed

[2] The highest aggregate bid is always non negative since the bids by each agent must sum up to zero.

by the members of that coalition. The offer is accepted if all the other agents agree. In case of acceptance the coalition is formed, the proposer produces the public goods, and the agents outside the coalition proceed to play the same game again among themselves. In the case of rejection all the agents other than the proposer play the same game again.

Formally, if there is only one agent $\{i\}$, he chooses a vector y of public goods. (If only one agent plays, there is no bidding stage.)

Given the rules of the mechanism for at most $n-1$ agents, the mechanism for $N = \{1, \ldots, n\}$ proceeds as follows:

$t = 1$: Each agent $i \in N$ makes bids $b^i_j \in \Re$, one for every $j \in N$, with $\sum_{j \in N} b^i_j = 0$. Hence, at this stage, a strategy for agent i is a vector $(b^i_j)_{j \in N} \in \mathcal{H}^n$, where

$$\mathcal{H}^n = \left\{ z \in \Re^n \;\Big|\; \sum_{j \in N} z_j = 0 \right\}.$$

For each $i \in N$, define the *aggregate bid* to agent i by $B_i = \sum_{j \in N} b^j_i$. Let $\alpha = \text{argmax}_i(B_i)$ where an arbitrary tie-breaking rule is used in the case of a non-unique maximizer. Once the winner α has been chosen, every agent $i \in N$ pays b^i_α and receives B_α/n.

$t = 2$: Agent α chooses a coalition S_α with $\alpha \in S_\alpha$, a production plan $y_\alpha \in \Re^m_+$ and makes an offer $x^\alpha_i \in \Re$ to every agent $i \neq \alpha$.[3]

$t = 3$: The agents other than α, sequentially, either accept or reject the offer. If an agent rejects it, then the offer is rejected. Otherwise, the offer is accepted.

If the offer is accepted, each agent $i \neq \alpha$ receives x^α_i, agent α forms the coalition S_α and produces y_α bearing the cost $c(y_\alpha, S_\alpha)$. After this, agents in $N \backslash S_\alpha$ proceed to play the game again among themselves. Therefore, the final payoff to agent $i \in S_\alpha \backslash \{\alpha\}$ is $u_i(y_\alpha, S_\alpha) + x^\alpha_i - b^i_\alpha + B_\alpha/n$; agent α receives $u_\alpha(y_\alpha, S_\alpha) - c(y_\alpha, S_\alpha) - \sum_{i \neq \alpha} x^\alpha_i - b^\alpha_\alpha + B_\alpha/n$; and the final payment for an agent $i \in N \backslash S_\alpha$ will be $-b^i_\alpha + B_\alpha/n + x^\alpha_i$ plus the payoff he will obtain in the game played by $N \backslash S_\alpha$.

On the other hand, if the offer is rejected, all agents other than α proceed to play the same game where the set of agents is $N \backslash \{\alpha\}$, and agent α is on his own. The final payoff to agent α is what he can obtain by himself (that is, $\max_{y \in \Re^m_+}[u_\alpha(y, \alpha) - c(y, \alpha)]$) minus the bid b^α_α plus his share of the aggregate bid B_α/n. The final payoff to any agent $i \neq \alpha$ is the payoff he obtains in the game played by $N \backslash \{\alpha\}$ minus the bid b^i_α plus the share B_α/n.

Theorem 1. *The subgame perfect equilibria of the bidding mechanism result in an efficient partition and production plan for the local public goods economy. Moreover, the mechanism implements the Shapley value of (N, W) in subgame perfect equilibria.*

Proof. The proof proceeds by induction on the number of agents n. It is easy to see that the theorem holds for $n = 1$. We assume that it holds for all $k \leqslant n - 1$ and show it is satisfied for n.

[3] For the players in S_α, $-x^\alpha_i$ is the payment made by player i to the player α who bears the whole cost of producing the public goods vector. For players outside S_α, x^α_i is the payment necessary (positive or negative) to induce them to stay outside of S_α.

First we show that the Shapley value is an SPE outcome. Consider the following strategies for the case of n agents:

At $t = 1$, each agent i, $i \in N$, announces $b^i_j = \phi_i(N\backslash\{j\}, W) - \phi_i(N, W)$ for every $j \neq i$ and $b^i_i = W(N) - W(N\backslash\{i\}) - \phi_i(N, W)$.

At $t = 2$, agent i, if he is the proposer, chooses a coalition S_i such that $S_i \in \text{argmax}_{S\subseteq N, S\ni i}\{w(S) + W(N\backslash S)\}$, a production plan y_i efficient for S_i, and offers $x^i_j = \phi_j(N\backslash\{i\}, W) - u_j(y_i, S_i)$ to every $j \in S_i\backslash\{i\}$ and $x^i_j = \phi_j(N\backslash\{i\}, W) - \phi_j(N\backslash S_i, W)$ to every $j \notin S_i$.

At $t = 3$, agent i, if agent $j \neq i$ is the proposer and $i \in S_j$, accepts any offer greater than or equal to $\phi_i(N\backslash\{j\}, W) - u_i(y_j, S_j)$ and rejects it otherwise. If agent $j \neq i$ is the proposer and $i \notin S_j$, agent i accepts any offer greater than or equal to $\phi_i(N\backslash\{j\}, W) - \phi_i(N\backslash S_j, W)$ and rejects it otherwise.

First of all, we notice that the bids at $t = 1$ are acceptable, that is, $\sum_{j\in N} b^i_j = 0$. This is a direct consequence of the well known property of the Shapley value that

$$\phi_i(N, W) = \frac{1}{n}\big[W(N) - W\big(N\backslash\{i\}\big)\big] + \frac{1}{n}\sum_{j\neq i}\phi_i\big(N\backslash\{j\}, W\big). \tag{1}$$

Moreover, if agents make the previous bids, then the aggregate bid to each one is zero. Indeed,

$$B_i = \sum_{j\in N} b^j_i = W(N) - W\big(N\backslash\{i\}\big) - \sum_{j\in N}\phi_j(N, W) - \sum_{j\neq i}\phi_j\big(N\backslash\{i\}, W\big) = 0.$$

The induction argument ensures that agent $i \notin S_\alpha$ will obtain $\phi_i(N\backslash S_\alpha, W)$ if the game is played among the agents in $N\backslash S_\alpha$. Also, agent $i \in S_\alpha$ obtains the utility derived from enjoying the public good, $u_i(y_\alpha, S_\alpha)$ plus the payment x^α_i. Then, it is easy to check that when all the agents follow the previous strategies, the total utility (taking into account the bid) obtained by any agent different from the proposer is $\phi_i(N, W)$. Also, these strategies lead to an efficient partition and to an optimal production plan of public goods for each coalition. Hence, the proposer obtains his Shapley value as well.

We now prove that the previous strategies constitute an SPE. The induction argument makes it clear that the strategy at $t = 3$ is a best response for any agent different from the proposer. At $t = 2$, given the strategies that the other agents will follow at $t = 3$, agent i (if he is the proposer) has to choose a subset S_i (such that $i \in S_i$), a production plan $y_i \in \Re^m_+$ and payments x^i_j to every $j \neq i$. The payments he will offer to the other agents will be the lowest ones that would not be rejected in stage 3.[4] Hence, the payments are given by

$$x^i_j = \begin{cases} \phi_j\big(N\backslash\{i\}, W\big) - u_j(y_i, S_i) & \text{if } j \in S_i\backslash\{i\}, \\ \phi_j\big(N\backslash\{i\}, W\big) - \phi_j(N\backslash S_i, W) & \text{if } j \notin S_i. \end{cases}$$

[4] We are considering proposals that will be accepted. Of course, the proposer always has the possibility of proposing an allocation that will not be accepted by the other players. This possibility is equivalent to proposing $S_i = \{i\}$ and then $x^i_j = \phi_j(N\backslash\{i\}, W) - \phi_j(N\backslash\{i\}, W) = 0$ to every $j \neq i$ and so it is implicitly taken into account in what follows. Notice also that the bids paid by all the players are a sunk cost at this stage and so do not enter the analysis at this point.

In light of that, agent i chooses his strategy so as to maximize

$$\begin{aligned} & u_i(y_i, S_i) - c(y_i, S_i) - \sum_{j \neq i} x_j^i \\ &= \Big[u_i(y_i, S_i) - c(y_i, S_i) - \sum_{j \neq i} \phi_j\big(N \backslash \{i\}, W\big) + \sum_{j \notin S_i} \phi_j(N \backslash S_i, W) \\ &\quad + \sum_{j \in S_i \backslash \{i\}} u_j(y_i, S_i) \Big] \\ &= \Big[\sum_{j \in S_i} u_j(y_i, S_i) - c(y_i, S_i) + W(N \backslash S_i) - W\big(N \backslash \{i\}\big) \Big]. \end{aligned} \tag{2}$$

The coalition S_i and the level of public good y_i that maximize (2) are the ones proposed in the candidate strategy.

Finally, consider the strategies at $t = 1$. Remember that we have shown that $B_i = 0$ for all $i \in N$. Given this, if agent i changes his bid $(b_j^i)_{j \in N}$, the proposer will be the agent (or one of the set of agents) to whom i increases his bid. But then, agent i's payoff will be lower than his Shapley value. Therefore, deviating is not profitable.

To show that any SPE yields the Shapley value outcome, denote the "effective offer" to agent $i \neq \alpha$ in stage 2 when agent α is the proposer by z_i^α:

$$z_i^\alpha = \begin{cases} x_i^\alpha + u_i(y_\alpha, S_\alpha) & \text{if } i \in S_\alpha \backslash \{\alpha\}, \\ x_i^\alpha + \phi_i(N \backslash S_\alpha, W) & \text{if } i \notin S_\alpha. \end{cases}$$

By the induction argument, the effective offer is the total utility (without taking into account the payments generated by the bidding stage) that an agent will obtain (in equilibrium) if the offer is accepted. We proceed by a series of claims. The proofs of claims (i) and (ii) are not difficult and they are similar to the proofs of claims (i) and (ii) in Theorem 1 of Pérez-Castrillo and Wettstein (2001).

Claim (i). In any SPE, any agent i different than the proposer α accepts the offer at $t = 3$ if $z_i^\alpha > \phi_i(N \backslash \{\alpha\}, W)$ for every $i \neq \alpha$. If $z_i^\alpha < \phi_i(N \backslash \{\alpha\}, W)$ for some $i \neq \alpha$, then the offer is rejected.

Claim (ii). In any SPE of the game that starts at $t = 2$, the proposer α chooses a coalition S_α that is part of an efficient partition and announces offers such that $z_i^\alpha = \phi_i(N \backslash \{\alpha\}, W)$ for all $i \neq \alpha$. Finally, at $t = 3$, every agent $i \neq \alpha$ accepts any offer such that $z_i^\alpha \geqslant \phi_i(N \backslash \{\alpha\}, W)$.[5] The final payoffs to the proposer α and to any other agent $i \neq \alpha$ are $W(N) - W(N \backslash \{\alpha\}) - b_\alpha^\alpha + B_\alpha / n$ and $\phi_i(N \backslash \{\alpha\}, W) - b_\alpha^i + B_\alpha / n$, respectively.

Claim (iii). In any SPE, $B_i = 0$ for all $i \in N$.

---

[5] To be rigorous, if $\{\alpha\}$ is part of any efficient partition, then there exist other equilibria in addition to the previous ones. Any strategy profile such that at $t = 2$, α makes offers such that $z_j^\alpha \leqslant \phi_j(N \backslash \{\alpha\}, W)$ to a particular player $j \neq \alpha$ and at $t = 3$, player j rejects any effective offer less than or equal to $\phi_j(N \backslash \{\alpha\}, W)$ also constitutes an SPE.

To prove the claim, let $\Phi \equiv \{i \in N \mid B_i = \max_{j\in N} B_j\}$. If $\Phi \neq N$, then the aggregate bid to all the agents in Φ is positive. In this case, any agent, say agent k, can slightly reduce the bids he made to all the agents in Φ, and increase the bids to the other agents, so as to satisfy that the sum of the bids is still zero, and the set Φ does not change. However, the utility of agent k increases since he pays a lower bid independently of which agent in Φ ends up as the proposer.

Claim (iv). In any SPE, each agent's payoff is the same regardless of who is chosen as the proposer.

This claim is a direct consequence of the previous one: If agent i would strictly prefer that agent k be the proposer, he could slightly increase b^i_k (and decrease b^i_j by the same amount, for some other j) so as to make sure that the aggregate bid to k is the highest at a negligible cost.

Claim (v). In any SPE, the final payment received by agent i is $\phi_i(N, W)$. Moreover, the bids made by the agents at $t = 1$ are $b^i_j = \phi_i(N\backslash\{j\}, W) - \phi_i(N, W)$ for every $j \neq i$ and $b^i_i = W(N) - W(N\backslash\{i\}) - \phi_i(N, W)$.

The payoff of agent i if he is the proposer is $W(N) - W(N/\{i\}) - b^i_i$, while his payoff is $\phi_i(N\backslash\{j\}, W) - b^i_j$ if j is the proposer. Summing over all the agents, using the fact that $\sum_{j\in N} b^i_j = 0$ and the property (1) of the Shapley value highlighted before, we obtain that the sum of the payoffs is $n\phi_i(N, W)$. By claim (iv), this implies that the payoff of agent i is always $\phi_i(N, W)$. The fact that the final payoff to agent i is $\phi_i(N, W)$, whereas the payoffs from stage 2 onwards is given by claim (ii), implies that the agent i's bids must be as described in claim (v). □

Bagnoli and Lipman (1989), Jackson and Moulin (1992), Bag and Winter (1999), and Mutuswami and Winter (2003) also proposed mechanisms that realize efficient outcomes in environments with (local) public goods. We now discuss the advantages and distinguishing features of the bidding mechanism.

The first distinguishing feature of our mechanism is that it does not need a planner. The agents do not submit messages to a planner who then implements the final outcome. In the bidding mechanism, the messages and offers from an agent are made directly to the other agents. The agents can play the mechanism by themselves. The only role that a third party might play is to act as a court in case some agent does not fulfill his commitments. This is an easy task here because the actions by the agents are just contracts (offers) which are easy to verify.

A second advantage of the proposed mechanism is that in contrast to the mechanisms suggested in the previous contributions, it can handle environments with local public goods. It allows the participating agents the possibility of splitting up into smaller coalitions. The outcome generated by the mechanism specifies not only the production/consumption plan the individuals will follow but also the coalition structure that will prevail. Since the outcome is efficient, coalitions other than the grand coalition can form in equilibrium.

Third, the equilibria of the bidding mechanism give rise to efficient outcomes even when the utility of an agent depends on the identity of the partners he is with[6] and the costs of producing the public good differ for different coalitions. The bidding mechanism thus obtains efficiency in a larger class of environments than in previous contributions.

Another distinguishing feature of the bidding mechanism is that it allows for transfers across coalitions. Our starting point is the set of all agents who can, subject to the rules of the mechanism, split up into distinct coalitions. If the efficient coalition structure is not the grand coalition, it is in the agents' own interests to form smaller coalitions. Therefore, it is entirely likely that some agents may have to be compensated so as not to join a coalition S. The rules of our mechanism allow for these compensations to take place via payments that end up transferring resources across coalitions. This feature is necessary to achieve efficient outcomes in those environments where forming the grand coalition is not efficient. This point of view might explain transfers whereby more affluent countries, who form a trading coalition, compensate other countries that are prevented from joining the coalition. This compensation alleviates tensions and pressures that might arise if these other countries try to force their way into the coalition.

The fifth feature is that equilibrium payoffs received by the agents in our mechanism coincide with the Shapley value of the super-additive cover of the corresponding cooperative game. It is important to point out that, in this value, the total contributions made by the agents belonging to a coalition do not necessarily match the cost of the public good produced by this coalition. Due to the possibility of cross subsidies, the value of an agent is a measure of his strategic possibilities not only inside the coalition to which he ends up belonging, but also with agents outside this coalition.

In the two-stage mechanisms proposed by Bag and Winter (1999) and Mutuswami and Winter (2003), the equilibrium payoffs coincide with the Shapley values of the original cooperative game (not its super-additive cover) for some environments.[7] We think that in environments where forming the grand coalition is not efficient, it is the super-additive cover which is the relevant measure of social surplus and consequently, in looking for efficient and equitable outcomes, it is the Shapley value of the super-additive cover which is relevant. Moreover, in these papers, the Shapley value can be obtained as *actual* payoffs by assuming that each agent prefers the game ending in stage 1 (which gives the Shapley value as actual payoffs) to ending in stage 2 (which gives the value as expected payoffs). Indeed, the agents obtain their Shapley value because of their fear in the first stage to be put in a worse position in the subsequent random choice of a new order of play which would come up if some one is unhappy. On the other hand, through the bidding mechanism the agents always obtain the Shapley value as their actual payoffs thanks to the initial bidding

[6] See the papers by Banerjee et al. (2001) and Bogomolnaia and Jackson (2002) for an analysis of "pure hedonic coalitions." In these papers, the players have (ordinal) preferences over the coalitions in which they are members, and the objective is to find conditions on preferences under which there exist a "stable partition" of players.

[7] Of course, the two solution concepts coincide if the game is super-additive. We observe that in the environment analyzed by Bag and Winter (1999), the corresponding cooperative game is *convex*, which is a stronger requirement than super-additivity.

game. No addition assumption is required here. These are two different ways to obtain a similar result, and the two methods may have different applications.

Finally, we would like to emphasize that the equilibrium bidding strategies in the mechanism satisfy very strong properties. We state the properties formally:

Theorem 2. *The bids that agents submit at the subgame perfect equilibria of the bidding mechanism satisfy the following properties*:

(i) *they are unique,*
(ii) *they are maxmin strategies at the bidding stage,*[8] *and*
(iii) *they are robust to deviations by coalitions of agents.*

Proof. (i) See claim (v) in the proof of Theorem 1.

(ii) If an agent bids according to the unique SPE bids, his payoff is, at least, his Shapley value independently of the bids made by the other agents. In fact, he will obtain his Shapley value if the aggregate bid of the proposer is zero (independently of the identity of the proposer) and he will obtain strictly more if the aggregate bid is positive.[9]

(iii) Given that the SPE outcome of the mechanism is efficient, there is no way in which a group of agents can increase their joint payoff without decreasing other agents payoffs. However, decreasing other agents payoffs is not possible since the SPE strategies are also maxmin strategies. □

In the bidding mechanism that we propose, one can argue that the most complex stage to be played is the bidding stage. Theorem 2 states that playing it is in fact relatively easy. The SPE bidding strategy profile is unique, eliminating the problem of coordination that exists when strategies are not unique. Moreover, bidding according to the SPE profile is "simple" and "safe." It is simple since the bid made by agent i to any other agent j is $b^i_j = \phi_i(N\backslash\{j\}, W) - \phi_i(N, W)$ which is the difference between agent i's Shapley value in the game without agent j and his Shapley value in the whole game. The bid b^i_j is the amount agent i pays if j is the proposer. This difference is often negative, in which case $-b^i_j$ represents the compensation paid to agent i if agent j is the proposer. This is the difference between "the value of the game" for agent i, $\phi_i(N, W)$, and the offer he will receive from agent j, that is, "the value of agent i in the game without agent j," $\phi_i(N\backslash\{j\}, W)$. Moreover, the bidding behavior is safe since the strategies are immune to coalitional deviations by groups and are maxmin strategies as well.

[8] A strategy is a maxmin strategy if it maximizes the minimum payoff that a player can possibly obtain for any choice of strategies by the other players.

[9] We are assuming here that the other players can change their bids but that they will stick to their equilibrium behavior from $t = 2$ onwards.

3. Forming networks

A class of economic environments that has recently received a lot of attention is that of network economics. An extensive survey of networks in industrial organization is provided by Bloch (2002), see also Kranton and Minehart (2001) for the analysis of networks as a new model of exchange. In such environments the surplus is created by the formation of links among the agents. A network, or a graph, is a structure of *bilateral relations* among agents. The value of the graph represents the total surplus produced by agents when they form the set of links prescribed by it. The literature restricts attention to externality-free graphs. These are graphs whose value is given by the sum of the values of their components, where a component is a set of agents forming links with one another (directly or indirectly) and having no links with agents outside this set. An allocation rule associates with each graph a sharing of the surplus among the agents comprising the graph. An example of an allocation rule is the one proposed by Jackson and Wolinsky (1996) which associates to each graph, the Shapley value of a transferable utility game associated with the graph.

A network is efficient if it creates the largest possible surplus. Recent literature on social and economic networks, stemming from the paper of Jackson and Wolinsky (1996) has focused on the problem of generating efficient networks. It tries to resolve the tension between efficiency and stability of networks.[10] Jackson and Wolinsky (1996) showed that it is not possible to reconcile this tension if the allocation rule is required to satisfy the properties of anonymity and component balancedness, that is the rule always shares the surplus created by a component among the component members.

In a subsequent paper, Dutta and Mutuswami (1997) showed that this tension can be resolved by using a *mechanism design* approach. They assumed that the planner knows the value function and based on it constructs an allocation rule. The value function and the allocation rule provide a payoff function for a suitably defined link formation game. Their main concern is with the behavior of the allocation rule on the equilibria of the link formation game. Mutuswami and Winter (2002) assumed the planner has just partial information regarding the value function and constructed mechanisms achieving efficient and at times symmetric outcomes.

In another paper, Currarini and Morelli (2000) examined this problem when there is no social planner. They analyzed two sequential-move games. In both games, an agent's strategy consists of two parts. The first part specifying the set of agents with whom the agent wants to form links is common to both games. The second part is a payoff demand and can be either a single absolute payoff demand or a vector of demands, one for each proposed link. Thus, the payoff division in Currarini and Morelli (2000) is endogenous, in contrast to the analysis in Jackson and Wolinsky (1996) and Dutta and Mutuswami (1997). They showed that all subgame perfect equilibria of their games give rise to efficient networks. However, the payoff division is highly asymmetric being sensitive to the order in which agents move.

The *bidding mechanism* we constructed in the previous section can be adapted to handle network economics. The bidding stage stays the same and leads to the choice of a proposer

[10] Stability of a network can be understood as meaning that there does not exist a deviation for some group of agents which makes all deviating members better off.

and a vector of transfers among agents. The proposal in the next stage consists of a component (that is the set of agents with whom the proposer wants to be connected), a set of links for this component and payments to all the agents. If the offer is accepted the component is formed, payments are made and the agents outside the component proceed to play the same game again. In the case of rejection it is all the agents other than the proposer who play the same game again. Hence, the bidding mechanism can be viewed as a model of network formation in which a bargaining process is modeled explicitly. In this mechanism, components are formed sequentially.

The equilibrium outcomes of the bidding mechanism lead to an efficient network. Moreover the payoffs to the agents are uniquely determined and given by the Shapley value of the transferable utility game where the value of the coalition is the maximal surplus its members can generate by acting on their own. That is, the value of a coalition is not calculated using a fixed network, it is determined by the value of the best network that the agents in the coalition can form. This Shapley value is different from the one considered in Jackson and Wolinsky (1996) (see also Myerson, 1977). In fact, it coincides with the one derived from the player-based flexible allocation network rule recently proposed by Jackson (2003), when the network to which the rule is applied is component additive.

Our proposal is in the same line of research as Currarini and Morelli (2000) but with the additional objective of ensuring that the payoffs to the agents are equitable. Thus, the first advantage of our mechanism with respect to that of Currarini and Morelli (2000) lies in the fact that the payoff division corresponds to the Shapley value of a game which takes into account the various network options available to the agents. Second, our result holds for a larger class of value functions while their result holds only for anonymous value functions satisfying size monotonicity. Finally, our mechanism does not require out-of-equilibrium free disposal.

4. Conclusion

In this paper we provided a mechanism through which, in equilibrium, a set of agents forms an efficient coalition structure, and produces efficient levels of local public goods for each of the coalitions. The equilibrium strategies enjoy appealing stability properties. This increases the likelihood that agents would actually agree to interact via such a mechanism. We intend to further support this claim by studying the performance of the mechanism in experimental settings.

We also indicated how to adapt the mechanism to handle network economics. The equilibrium outcomes of the mechanism in such an environment lead to an efficient network and an equitable division of the surplus generated.

Finally, we note that slightly modified versions of the mechanism constructed can be applied in any environment with quasi-linear preferences (transferable utility), where agents can form into subgroups, with each subgroup choosing a collective decision which affects only the agents in the subgroup. Such scenarios include, in addition to those analyzed in the paper, problems of cost sharing, partnership dissolution, bankruptcy issues and so on. The mechanism would generate an efficient partition as well as efficient decisions, and furthermore, the payoffs would be equitable, thus realizing two desirable objectives of efficiency and equity.

Acknowledgments

The authors thank an associate editor and the referee for helpful comments. Pérez-Castrillo acknowledges financial support from BEC2000-0172, BEC2003-01132, and 2001SGR-00162.

References

Bag, P., Winter, E., 1999. Simple subscription mechanisms for excludable public goods. J. Econ. Theory 87, 72–94.

Bagnoli, M., Lipman, B., 1989. Provision of public goods: Fully implementing the core through private contributions. Rev. Econ. Stud. 56, 583–601.

Banerjee, S., Konishi, H., Sönmez, T., 2001. Core in a simple coalition formation game. Soc. Choice Welfare 18, 135–153.

Bogomolnaia, A., Jackson, M., 2002. The stability of hedonic coalition structures. Games Econ. Behav. 38, 201–230.

Bloch, F., 2002. Coalitions and networks in industrial organization. Manchester Sch. 70, 36–55.

Currarini, S., Morelli, M., 2000. Network formation with sequential demands. Rev. Econ. Design 5, 229–249.

Dubins, L.E., 1977. Group decision devices. Amer. Math. Monthly 84 (May), 350–356.

Dutta, B., Mutuswami, S., 1997. Stable networks. J. Econ. Theory 76, 322–344.

Jackson, M., 2003. Allocation rules for network games. Mimeo.

Jackson, M., Moulin, H., 1992. Implementing a public project and distributing its costs. J. Econ. Theory 57, 125–140.

Jackson, M., Wolinsky, A., 1996. A strategic model of social and economic networks. J. Econ. Theory 71, 44–74.

Kranton, R., Minehart, D., 2001. A theory of buyer-seller networks. Amer. Econ. Rev. 91, 485–508.

Maniquet, F., 2003. Implementation under perfect information in economic environments. Soc. Choice Welfare In press.

Moore, J., Repullo, R., 1988. Subgame perfect implementation. Econometrica 46, 1191–1220.

Mutuswami, S., Winter, E., 2002. Subscription mechanisms for network formation. J. Econ. Theory 106, 242–264.

Mutuswami, S., Winter, E., 2003. Efficient mechanisms for multiple public goods. J. Public Econ. In press.

Myerson, R., 1977. Graphs and cooperation in games. Math. Oper. Res. 2, 225–229.

Pérez-Castrillo, D., Wettstein, D., 2001. Bidding for the surplus: A non-cooperative approach to the Shapley value. J. Econ. Theory 100, 274–294.

Pérez-Castrillo, D., Wettstein, D., 2002. Choosing wisely: A multi-bidding approach. Amer. Econ. Rev. 92, 1577–1587.