Journal of Economic Theory **106**, 242–264 (2002)
doi:10.1006/jeth.2001.2920

Subscription Mechanisms for Network Formation[1]

Suresh Mutuswami

Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom
smutus@essex.ac.uk

and

Eyal Winter

Department of Economics, European University Institute, Badia Fiesolana, Via dei Roccettini, 9, 50016 San Domenico di Fiesole (FI), Italy; and Center for Rationality and Interactive Decision Theory, Hebrew University of Jerusalem, Feldman Building, Givat Ram, 91904 Jerusalem, Israel
mseyal@mscc.huji.ac.il

Received February 26, 2000; final version received April 13, 2001

We analyze a model of network formation where the costs of link formation are publicly known but individual benefits are not known to the social planner. The objective is to design a simple mechanism ensuring efficiency, budget balance, and equity. We propose two mechanisms towards this end; the first ensures efficiency and budget balance but not equity. The second mechanism corrects the asymmetry in payoffs through a two-stage variant of the first mechanism. *Journal of Economic Literature* Classification Numbers: C71, C72, D20.

Key Words: networks; mechanism design; shapley value.

1. INTRODUCTION

The literature on network formation has mainly been concerned with finding ways of distributing the surplus from a network so as to reconcile

[1] We thank Francis Bloch, an associate editor and an anonymous referee for their comments on this paper. Some parts of this paper were written during the first author's visits to the Center for Rationality and Interactive Decision Theory at The Hebrew University of Jerusalem and CORE, Belgium. He thanks these institutions for their hospitality and financial support. The second author thanks the European Commission for financial support through its TMR grant.

0022-0531/02 $35.00

the tension between efficiency and stability. In Operations Research, this is exemplified by the vast literature on the minimum spanning tree problem.[2] There is also a recent literature which focuses on the economic and social aspect of networks. This literature owes its genesis to the seminal paper by Jackson and Wolinsky [12] and has been explored by a number of other authors.[3] It differs from the Operations Research literature in that a network is regarded as a set of bilateral relationships amongst agents rather than something material. However, the theme here is similar to that of the literature on the minimum spanning tree problem—to develop ways of sharing the surplus such that the surplus-maximizing network is also stable.[4] In both these literatures it is assumed that the benefits to individual agents from network formation are publicly known. This is a strong assumption and it is thus important to understand network formation when the benefits to individual agents are not known publicly.

Our paper makes a first attempt in this direction by examining how a social planner can ensure the formation of an efficient network in a scenario where the costs of network formation are publicly known but an individual player's benefits from network formation are not known to him. As an example of where our approach may be valid, consider the minimum spanning tree problem. Assume, as in the traditional formulation of this problem, that the costs of network formation are common knowledge. Suppose that player 1's benefits from being connected to the source can take two possible values: a high value and a low value and that if player 1's benefits are low, then it is not socially optimal to provide her the input. Thus, the optimal network should not involve player 1 when her benefits from the input are low. The classical formulation of the problem does not take this aspect into consideration and proceeds as if player 1's benefit from the input is always high.

The fact that the social planner may not know the exact benefits of each agent from a given network forms the cornerstone of our analysis. The problem is one of mechanism design because there is asymmetric informa-

[2] The minimum spanning tree problem involves a source who supplies an input to a number of agents. There are two parts to this problem: (i) to find the network minimizing the aggregate cost of connecting all agents to the source; and (ii) to find a cost sharing method such that no subset of agents have an incentive to deviate and build their own network to the source. See Graham and Hall [8] for a historical survey of this problem.

[3] See the papers of Currarini and Morelli [4], Dutta and Mutuswami [6], Johnson and Gilles [14] and Slikker and van den Nouweland [19]. Bala and Goyal [3] are also concerned with the efficiency-stability problem but they use a dynamic framework to examine this issue.

[4] Jackson and Wolinsky [12] and Dutta and Mutuswami [6] use different ways of evaluating the stability of a network. The reader is referred to their papers for details.

tion between the planner and the agents.[5] We are interested in finding mechanisms through which an uninformed planner can nonetheless ensure three different objectives: (i) efficiency, which means ensuring the formation of a network maximizing net social surplus at all preference profiles, (ii) budget balance, which means that the costs of network formation are met by the contributions of the agents themselves, and (iii) equity. We shall be precise as to what we mean by equity later on but at this point it suffices to say that we would like the net payoffs of the agents to correspond to the benefits they derive from network formation.

Our formulation is general in that it can accomodate the Operations Research literature where the emphasis is on physical networks as well as the economic networks literature where a network is to be interpreted as a set of bilateral relationships amongst agents. Our main result shows that a mechanism can be designed meeting all of the above three objectives under weak restrictions on preferences and technology.

We discuss two mechanisms in this paper. The first mechanism involves agents announcing sequentially. Each agent, when it is her turn to move, announces the set of links that she would like to see formed and a conditional cost contribution. Once all agents have announced, the planner selects the network to be formed and the cost shares of the agents. This mechanism ensures the formation of an efficient network but the net payoffs to the agents are asymmetric, being sensitive to the order in which agents announce. We then modify this mechanism in a way which makes the agents symmetric. This modified mechanism ensures the formation of an efficient network along with equitable net payoffs.

The mechanisms proposed in this paper are not intended to be descriptions of actual network formation situations. Our objective—as mentioned before—is to design a mechanism which is simple, efficient and which leads to a reasonable payoff for the agents. The approach adopted here thus differs from a recent literature on network formation which analyze link formation games with endogenous payoff determination. Our understanding is that the games analyzed there are meant to be stylized descriptions of actual network formation situations. We discuss the relationship of our work with this literature in detail in Section 8.

In what follows, we set up the basic model in Section 2, the two mechanisms are introduced in Section 3 and analyzed in Section 4. Section 5 shows that the results obtained in Section 3 depend on the assumption that the utility functions are monotonic and shows how the mechanisms introduced in Section 3 can be modified to obtain identical results to those in

[5] Even though Dutta and Mutuswami [6] "take a mechanism design approach," theirs is not a mechanism design problem because there is no informational asymmetry between the planner and the players in their model.

Section 4. Section 6 discusses the extension to directed graphs and Section 7 discusses the coalitional stability of the mechanisms. Section 8 discusses the relationship of our work with the literature on network formation with endogenous payoff determination and also a smaller literature on cost sharing in multicast trees. We conclude in Section 9.

2. THE MODEL

Let $N=\{1, 2, ..., n\}$ be the set of agents. For any $S \subset N$, let $g^S=\{T \mid T \subset S, |T|=2\}$ be the set of all subsets of S of size 2. A graph or network, denoted generically by g is some subset of g^N. In other words, a graph is a structure of bilateral relationships between the agents with i and j having a bilateral relationship if and only if $\{i, j\} \in g$. The element $\{i, j\}$ of g is also called the *link* between i and j and denoted as (ij). For every $S \subset N$, $G_S=\{g \mid g \subset g^S\}$ denotes the set of graphs involving links only between members of S.

The following graph-theoretic terminology is needed for what follows. Players i and j are *connected in the graph* g if there exists a sequence of agents $i=i_0, i_1, ..., i_K=j$ such that $(i_k i_{k+1}) \in g$ for all $k=0, ..., K-1$. Let $N(g)=\{i \mid \exists j \text{ such that } (ij) \in g\}$ denote the set of agents who are part of at least one link. All players in $N \setminus N(g)$ are said to be *isolated*. The graph $h \subset g$ is a *connected component* of g if all agents in $N(h)$ are connected to each other in h, and for all $i \in N(h)$, $j \notin N(h)$, $(ij) \notin g$. The set of all connected components of g is denoted as $C(g)$.

An agent's benefits from being part of a network is given by a quasi-linear utility function $U_i(g, x_i)=v_i(g)-x_i$. Here, x_i is interpreted as the cost share imputed to agent i. We impose the following restrictions on the function v_i.

Assumption 1. The functions v_i, $i \in N$ are non-negative and monotonic. Formally, for all $i \in N$ and all $g, g' \in G_N$, $g \subset g'$, $v_i(g') \geqslant v_i(g) \geqslant 0$.

Remark 1. Assumption 2 says that the gross benefit to an agent is non-decreasing in the set of links. This is a weak assumption because it only applies to situations where one graph is a subgraph of the other graph. This assumption is satisfied in a number of models. For instance, the minimum spanning tree problem where an agent gets a utility of V_i if he is connected to the source and zero otherwise, satisfies Assumption 1.

There are, however, interesting models where Assumption 1 is not satisfied. In such cases, our mechanisms can be modified in an appropriate way. We shall show how this can be done later.

Finally, note that Assumption 1 allows for the presence of *externalities* across components. Thus, if $i \in N(h), h \in C(g)$, then her utility can be affected by the formation of a link between two players in $N \setminus N(h)$.

The cost of establishing a network is given by a cost function c. We impose the following restrictions on the cost function.

Assumption 2. The cost function is non-negative and satisfies $c(\emptyset) = 0$.

Assumption 2 amounts to saying that there are no fixed costs. This is a weak restriction.

Remark 2. Jackson and Wolinsky [12] use a framework where the object of analysis is a *value function* which is a real-valued mapping on the set of all possible networks, G_N. The value function is the direct analogue of the characteristic function in cooperative game theory which is defined on the set of all coalitions. In particular cases—for instance, in their discussion of the *symmetric connections* and the *co-author* models—the value function is defined as the sum of all agents' net utilities where i's net utility is the difference between the gross benefits he receives from the graph and the cost of forming the links with which he is directly associated. Since we also work in a transferable utility framework, our setup can be translated into the value function framework by specifying the corresponding value function as $v(g) = \sum_{i \in N} v_i(g) - c(g)$ for every $g \in G_N$.[6]

In what follows, we let S_k, $k = 1, ..., n$, denote the subset $\{k, k+1, ..., n\}$ of agents with S_{n+1} being the empty set. Player i's announcement in a mechanism is indicated by the use of a subscript as in v_i, w_i, etc. The vector of announcement of all players is denoted as v, w, etc. The term "announcement" shall be used to refer both to a player's announcement and the vector of announcements of all players and this will not lead to confusion because the context will make it clear as to which usage is relevant. Finally, the term "subgame perfect equilibria" will be abbreviated to SPE.

3. THE MECHANISMS Γ_n AND Γ_n^1

Our formulation assumes that the cost function is commonly known in the society. On the other hand, the benefit functions of the agents are not known to the planner even while they are common knowledge in the rest of society. Undoubtedly, this is a strong assumption, corresponding to the "complete information" framework of implementation theory.

[6] Jackson and Wolinsky [12] confine attention to value functions satisfying *component additivity*: For all $g \in G_N$, $v(g) = \sum_{h \in C(g)} v(h)$. This formulation explicitly rules out externalities among components. It is easy to see that the value functions derived from the utility and cost functions of our framework may not satisfy component additivity.

The mechanisms considered here require agents to move sequentially. We assume throughout that they move in the "natural order," viz. first, agent 1 moves, then 2, and so on. As will become clear, this is without loss of generality.

In the mechanism Γ_n, players move sequentially and player i when it is her turn to move, announces a tuple $w_i = (g_i, x_i)$ where $g_i \subset g^N$ is the set of links that i wants to see formed and x_i is her conditional cost contribution. The cost contribution is conditional in the sense that it is paid only if the resulting network g is a supergraph of g_i.

DEFINITION 1. The coalition S is *compatible* with the announcement $w = \{(g_i, x_i)\}_{i=1}^{n}$ if

1. $g_i \in G_S$ for all $i \in S$,
2. $\sum_{i \in S} x_i \geqslant c(g)$ where $g = \bigcup_{i \in S} g_i$.

The first part of the definition says that each agent in S proposes links which only involve members of S. The second part says that the costs of forming the resulting graph are covered by the contributions of members of S.

Given the announcement w, the planner selects the largest compatible coalition in the set $\{\emptyset, S_1, ..., S_n\}$ which is called the *maximal compatible coalition* and denoted $S^*(w)$. He then forms the graph $g^* = \bigcup_{i \in S^*(w)} g_i$, and charges the players as follows:

$$x_i^*(w) = \begin{cases} x_i & \text{if} \quad i \in S^*(w), \\ 0 & \text{otherwise.} \end{cases}$$

This completes the description of the mechanism Γ_n.

The mechanism Γ_n^1 operates in two stages. In the first stage, players play the game Γ_n according to any (arbitrarily chosen) order of the agents. At the end of this stage, each player is asked whether she wants to replay the game. The game ends if all players answer "NO." Otherwise, the game Γ_n is replayed according to a randomly chosen order of the players with equal probability for each order. The game Γ_n^1 ends after this optional second stage.

4. ANALYSIS OF THE MECHANISMS Γ_n AND Γ_n^1

We start with the following definitions.

DEFINITION 2. The *stand alone* payoff for a coalition S is defined as

$$sa(S) = \begin{cases} \max_{g \in G_S} \sum_{i \in S} v_i(g) - c(g) & \text{if} \quad S \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

DEFINITION 3. A graph $g \in G_S$ is *efficient for* S if $\sum_{i \in S} v_i(g) - c(g) = sa(S)$. If g is efficient for N, then g will simply be called an efficient graph.

DEFINITION 4. The *marginal contribution* of player $k, k = 1, ..., n,$ is $u_k^* = sa(S_k) - sa(S_{k+1})$.

Remark 3. It can be seen from Definition 2 that the stand alone payoff for S is computed by considering only graphs in G_S. In other words, all agents in $N \setminus S$ are assumed to be isolated. This is unproblematic so long as there are no consumption or cost externalities across components. Our formulation, though, does not rule out these possibilities. If there are externalities across components, then there is no unambiguous way to measure the worth of a coalition. One faces a similar problem in defining the characteristic function game of a pure public good economy.

Remark 4. Assumption 1 implies that if $S \subset T$ then $sa(S) \leqslant sa(T)$. This follows from the simple observation that a graph g_S^* which is efficient for S can always be formed by the larger coalition T. We therefore have $sa(T) \geqslant \sum_{i \in T} v_i(g_S^*) - c(g_S^*) \geqslant sa(S)$. The first inequality follows because g_S^* typically will not achieve the stand alone payoff for T. The second inequality follows because $v_i(g_S^*) \geqslant 0$ for all $i \in T \setminus S$ by Assumption 1.

Our main results regarding the mechanisms Γ_n and Γ_n^1 can be summarized in the following theorems.

THEOREM 1. *Suppose that Assumptions* 1 *and* 2 *are satisfied. Then, all SPE of Γ_n result in the formation of an efficient graph. The net payoffs to the agents in all SPE are given by the marginal contribution vector* $(u_1^*, ..., u_n^*)$.

THEOREM 2. *Suppose that Assumptions* 1 *and* 2 *are satisfied. Then, all SPE of Γ_n^1 result in the formation of an efficient graph. The net payoffs to the agents in all SPE are given by the Shapley value of the TU-game* (N, sa).

Remark 5. Note that Theorems 1 and 2 do not specify the equilibrium strategies uniquely. If there are a number of possible efficient graphs, then they do not say anything about which efficient graph forms in equilibrium either. However, they are able to pin down the equilibrium net payoffs uniquely.

Consider first, the mechanism Γ_n. Suppose that agents in $\{1, ..., k\}$ have already announced $\{(g_i, x_i)\}_{i=1}^k$. Let $i_k \in \{1, ..., k\}$ be the *largest* integer such that there exists a graph $g \in G_{S_{i_k}}$ satisfying

$$g_i \subset g \qquad \text{for all} \quad i \in \{i_k, ..., k\}, \tag{1}$$

$$\sum_{j \in S_{k+1}} v_j(g) - c(g) + \sum_{j=i_k}^{k} x_j > sa(S_{k+1}). \tag{2}$$

Note that (2) says that by utilizing the contributions of agents in $\{i_k, ..., k\}$, the agents in S_{k+1} can collectively obtain a payoff greater than their stand-alone payoff. The fact that i_k is the largest integer satisfying (1) and (2) means that S_{i_k} is the *smallest* compatible coalition which can ensure a collective payoff greater than $sa(S_{k+1})$ to the agents in S_{k+1}.[7] In Lemma 1 we show that if i_k exists, then the resulting maximal compatible coalition must be a superset of S_{i_k}.

We need a last piece of notation before stating and proving Lemma 1. Suppose that i_k exists; then $G_k^* = \{g \in G_{S_{i_k}} \mid g \text{ satisfies (1) and (2)}\}$.

LEMMA 1. *Suppose that agents in* $\{1, ..., k\}$ *have announced* $\{(g_i, x_i)\}_{i=1}^k$ *in the game* Γ_n. *Suppose also that* i_k *exists. Then, all SPE of the subgame following k's announcement will be such that the resulting maximal compatible coalition is a superset of* S_{i_k}.

Proof. The proof is by induction. If $k = n$, then there is no subgame following n's announcement.[8] It is easy to check that in this case, (1) and (2) imply that S_{i_n} is a compatible coalition and therefore, the resulting maximal compatible coalition must be a superset of S_{i_n}.

Suppose that the lemma is true for all $k > K$ but that the maximal compatible coalition is not a superset of S_{i_K} in some SPE of the subgame following K's announcement. Let $(u_{K+1}, ..., u_n)$ be the resulting net utilities of the agents following K. Since the maximal compatible coalition must be a superset of S_{i_K} if agents in S_{K+1} are to receive a collective payoff greater than $sa(S_{K+1})$, we must have $\sum_{j=K+1}^{n} u_j \leqslant sa(S_{K+1})$. We can distinguish between two cases here.

Case 1. $u_j < u_j^*$ for some $j \in \{K+1, ..., n\}$.

Let agent j deviate by announcing (g'_j, x'_j) where g'_j is an efficient graph for S_j and x'_j such that

$$sa(S_{j+1}) - \sum_{k=j+1}^{n} v_k(g'_j) + c(g'_j) < x'_j < v_j(g'_j) - u_j. \tag{3}$$

[7] Recall that the only relevant compatible coalitions are $\varnothing, S_1, ..., S_n$.
[8] Recall that agents announce in the order 1, 2, ..., n so that n is the last to announce.

Observe that the upper bound on x'_j is strictly greater than the lower bound if and only if $u_j < sa(S_j) - sa(S_{j+1}) = u^*_j$ which is true by assumption.[9] Thus, a value of x'_j satisfying (3) exists. Since $\sum_{k=j+1}^{n} v_k(g'_j) - c(g'_j) + x'_j > sa(S_{j+1})$ and $g'_j \in G_{S_j}$, it follows that $i_j = j$ after j's deviation. The induction hypothesis implies that the maximal compatible coalition resulting from any SPE of the subgame following j's deviation will be a superset of S_j and will therefore, always contain j. Compatibility implies that the resulting graph, say $\tilde{g}$, will be such that $g'_j \subset \tilde{g}$. Assumption 1 implies that $v_j(\tilde{g}) - x'_j \geqslant v_j(g'_j) - x'_j > u_j$. This shows that j has a profitable deviation, a contradiction.

Case 2. $u_j = u^*_j$ for all $j \in \{K+1, ..., n\}$.

Choose some $g^*_K \in G^*_K$. Let agent $K+1$ deviate by announcing (g^*_K, x'_{K+1}) where

$$sa(S_{K+2}) - \sum_{i=K+2}^{n} v_i(g^*_K) + c(g^*_K) - \sum_{i=i_K}^{K} x_i < x'_{K+1} < v_{K+1}(g^*_K) - u^*_{K+1}. \quad (4)$$

To see that a value of x'_{K+1} satisfying (4) exists, observe that the upper bound on x'_{K+1} is strictly greater than the lower bound if and only if $\sum_{i=K+1}^{n} v_i(g^*_K) - c(g^*_K) + \sum_{i=i_K}^{K} x_i > sa(S_{K+1})$. This is true because $g^*_K \in G^*_K$.

Since $\sum_{i=K+2}^{n} v_i(g^*_K) - c(g^*_K) + \sum_{i=i_K}^{K} x_i + x'_{K+1} > sa(S_{K+2})$, $g_i \subset g^*_K \in G_{S_{i_K}}$ for all $i \in \{i_K, ..., K+1\}$, it follows that i_{K+1} exists after $K+1$'s deviation.[10] The induction hypothesis implies that the maximal compatible coalition resulting from any SPE of the subgame following $K+1$'s deviation must be a superset of $S_{i_{K+1}}$. Since $i_{K+1} \leqslant K+1$, it follows that $K+1$ is a member of the maximal compatible coalition. By compatibility, the resulting graph, say $\tilde{g}$, must be a supergraph of g^*_K. By Assumption 1, we have $v_{K+1}(\tilde{g}) - x'_{K+1} \geqslant v_{K+1}(g^*_K) - x'_{K+1} > u^*_{K+1}$ which shows that $K+1$ has a profitable deviation. This contradiction completes the proof of the lemma. ∎

COROLLARY 1. *Let $(u_1, ..., u_n)$ be the net payoffs to the agents in some SPE of Γ_n. Then, $u_i \geqslant u^*_i$ for all $i \in N$.*

Proof. Suppose $u_k < u^*_k$ for some k in some SPE of Γ_n. Let k deviate by announcing (g'_k, x'_k) where g'_k is efficient for S_k and x'_k such that

$$sa(S_{k+1}) - \sum_{j=k+1}^{n} v_j(g'_k) + c(g'_k) < x'_k < v_k(g'_k) - u_k.$$

[9] Since g'_j is efficient for S_j, we have $\sum_{k=j}^{n} v_k(g'_j) - c(g'_j) = sa(S_j)$.

[10] Clearly, we must have $i_K \leqslant i_{K+1} \leqslant K+1$.

Such an announcement is possible since $u_k < u_k^*$. Since $\sum_{j=k+1}^{n} v_j(g'_k) - c(g'_k) + x'_k > sa(S_{k+1})$ and $g'_k \in G_{S_k}$, it follows that $i_k = k$ after k's deviation. By Lemma 1, it follows that the maximal compatible coalition from any SPE of the subgame resulting after k's deviation is a superset of S_k, and will, therefore, always contain k. By compatibility, the resulting graph, say g, must be a supergraph of g'_k and by Assumption 1, it follows that k has a profitable deviation, a contradiction. ∎

Proof of Theorem 1. Let $(u_1, ..., u_n)$ be the net payoffs to the agents in an SPE of Γ_n. By Corollary 1, $u_i \geqslant u_i^*$ for all $i \in N$. If $u_i > u_i^*$ for some i, then $\sum_{j=1}^{n} u_j > sa(N)$ which is a contradiction since Remark 4 (which uses Assumptions 1 and 2) shows that the maximum surplus attainable is $sa(N)$. Thus, $u_i = u_i^*$ for all $i \in N$. The fact that an efficient network forms follows trivially from the observation that $\sum_{i=1}^{n} u_i = sa(N)$. ∎

Proof of Theorem 2. Let $(\phi_1(sa), ..., \phi_n(sa))$ denote the Shapley value payoffs in the game (N, sa). Suppose the game enters Stage 2 of Γ_n^1. We know from Theorem 1, that for any order selected by the planner, the SPE payoffs will be given by the corresponding "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. Thus, if g_N^* is an efficient graph, then the strategy profile $(\{(g_N^*, x_i^*)\}, \text{"NO"})_{i=1}^{n}$ where $x_i^* = v_i(g_N^*) - \phi_i(sa)$ constitutes a SPE of Γ_n^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. (Note that an efficient network forms in either case.) If we assume that all agents have a lexicographic preference for the game ending in the first stage, then all SPE of the game Γ_n^1 will end in the first stage with the formation of an efficient network and the agents getting their Shapley value payoffs.[11] ∎

Remark 6. If agents discount the future instead of having a lexicographic preference for the game ending in stage 1, then the game Γ_n^1 always ends in the first stage and the payoffs approach the Shapley value as the discount factor approaches one.

Let $0 < \beta < 1$ be the discount factor. If the game Γ_n^1 reaches the second stage, then the expected payoffs to the agents at the beginning of Stage 2

[11] The assumption that all agents have a lexicographic preference for the game ending in the first stage is also used by Bag and Winter [2] in their analysis of production and cost sharing of an excludable public good.

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.[12] In other words, with discounting, the game Γ_n^1 always ends in the first stage and the net payoffs are given by $(sa(N) - \beta\sum_{j=2}^{n} \phi_j(sa), \beta\phi_2(sa), \ldots, \beta\phi_n(sa))$. Thus, as $\beta \to 1$, the payoffs converge to the Shapley value.

Remark 7. Jackson and Wolinsky [12] propose a way of allocating the surplus from a network which also uses the Shapley value. Their rule is defined as follows. Given g and $\varnothing \neq S \subseteq N$, let $g \mid S = \{\{i, j\} \mid \{i, j\} \in g$ and $\{i, j\} \subset S\}$. Define the game (N, v^g) by $v^g(S) = \sum_{i \in S} v_i(g \mid S) - c(g \mid S)$. The Jackson–Wolinsky allocation rule is then the Shapley value of (N, v^g).[13] Note that while the Jackson–Wolinsky procedure specifies the sharing rule for every possible graph, our procedure is valid only for *efficient* graphs. However, even on the domain of efficient graphs, the two rules differ. This is illustrated by the following example.

EXAMPLE 1. Let $N = \{1, 2, 3\}$ and for $j = 1, 2, 3$, let $g^j = \{\{i, j\}, \{j, k\}\}$ where $k \neq i, j \neq i, k \neq j$. The gross payoffs of the agents are as follows: $v_1(\{1, j\}) = 30$ for $j = 2, 3$, $v_1(g^j) = 36$ for all j, $v_1(g^N) = 40$; $v_2(\{2, j\}) = 10$ for $j = 1, 3$, $v_2(g^j) = 26$ for all j, $v_2(g^N) = 30$; $v_3(\{3, j\}) = 10$ for $j = 1, 2$, $v_3(g^j) = 16$ for all j and $v_3(g^N) = 20$. An isolated agent gets zero. The cost of establishing any link is 15. It is easy to check that the utility and cost functions satisfy Assumptions 1 and 2. Some simple calculations reveal that (i) the efficient graphs for N are g^i, $i = 1, 2, 3$, and (ii) the efficient graph for $\{i, j\}$ is $\{\{i, j\}\}$. The stand alone payoffs are as follows: $sa(\{i\}) = 0$ for all i, $sa(\{1, 2\}) = sa(\{1, 3\} = 25$, $sa(\{2, 3\}) = 5$ and $sa(N) = 48$. Using

[12] 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 the agents following her, agent 2 can follow by conceding some fraction of αS 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 [2] and is implicitly used in Lemma 1 and Corollary 1.

[13] As observed before, Jackson and Wolinsky [12] use the value function framework which is different from the framework of this paper. The definition given here corresponds "naturally" to the Jackson–Wolinsky rule in the context of our model. The Jackson–Wolinsky rule itself is an extension of the value proposed by Myerson [16] for "graph-restricted" games.

Theorem 2, it follows that the net payoffs to the agents in any SPE of Γ_n^1 are given by the Shapley value of (N, sa) which works out to $(\phi_1, \phi_2, \phi_3) = (136/6, 76/6, 76/6)$.

The Jackson–Wolinsky procedure applied to g^1 gives $(146/6, 71/6, 71/6)$, to g^2 gives $(111/6, 126/6, 51/6)$, and to g^3 gives $(111/6, 51/6, 126/6)$. Thus, the equilibrium payoffs in the mechanism Γ_n^1 are always different from those obtained through the Jackson–Wolinsky procedure. The difference is on account of the fact that the computation of a player's "marginal contribution" differs in the two procedures. While we compute the "marginal contribution" for a player by looking at the difference in stand-alone payoffs $(sa(S \cup i) - sa(S))$, the Jackson–Wolinsky procedure for a given graph g looks at the difference in the values of the graph restricted to the coalitions $S \cup i$ and S, that is, $v^g(S \cup i) - v^g(S)$. The two procedures are bound to give different answers since $g \mid S$ will typically not be the graph that achieves the stand-alone payoff for S even when g is efficient for the grand coalition.

Remark 8. The allocation rules specified in Theorems 1 and 2 do not necessarily satisfy component balancedness.[14] We have two observations in this regard. Firstly, the literature (which mostly uses the value function setup) typically assumes that the value function is component additive.[15] While requiring the allocation rule to be component balanced makes sense if the value function is component additive, this restriction makes less sense when the value function is not component additive. In our context, given the weak restrictions on the utility and cost functions, there is no guarantee that the induced value function is going to be component additive. Secondly, component balancedness as an axiomatic restriction on the allocation rule makes sense in a setup where the primitive is the value function while our primitives are the utility and cost functions. We are not sure what component balancedness means in our context and therefore do not impose this restriction.

5. NONMONOTONIC UTILITY FUNCTIONS

Theorems 1 and 2 depend crucially on Assumption 2 which requires the gross benefit function of an agent to be monotonic. The following example based on the co-author model of Jackson and Wolinsky [12] illustrates this point.

[14] Component balancedness says that if h is a connected component of a graph g, then the value $v(h)$ must be distributed only to the members making up that component. It rules out transfers across components.

[15] See footnote 2. This assumption has been used in a number of papers, for example, Currarini and Morelli [4], Dutta and Mutuswami [6], Jackson and Wolinsky [12], etc.

EXAMPLE 2. Let $N=\{1,2,3\}$. For any $g\subset g^N$, let n_k denote the number of links involving player k. Denote by g^i the graph $\{\{i,j\},\{i,k\}\}$. The utility of agent k is given by

$$u_k(g)=\begin{cases}\sum_{j:\,\{k,j\}\in g}\dfrac{1}{n_k}+\dfrac{1}{n_j}+\dfrac{1}{n_k n_j} & \text{if } k \text{ is not isolated and } g\neq g^2 \text{ or } g^3,\\ 0 & \text{otherwise.}\end{cases} \tag{5}$$

Let $c(g)=0$ for all g. It is easy to verify that u_k is not monotonic and that the unique efficient graph is g^1. The stand alone payoffs for the various coalitions are as follows: $sa(N)=8$, $sa(\{i,j\})=6$, $sa(\{i\})=0$. Let the agents move in the order 1, 2, 3. Suppose that 1 announces (g_1,x) so as to be compatible with the formation of g^1. Observe that (g_1,x) is compatible with the formation of g^1 so long as $\{2,3\}\notin g_1$ and $x\geqslant 0$.[16] Now, note that $u_2(g^1)=u_3(g^1)=2$ while $u_2(g^N)=u_3(g^N)=2.5$. Therefore, if 2 announces so as to be compatible with the formation of g^1, then she can get a net payoff of at most $x+2+2=x+4$ which is the sum of her own gross payoff in g^1 and the amounts that can be extracted from 1 and 3. However, she can do better by announcing $\{g^N,-(2+x)\}$. It is easy to verify that (one of) 3's optimal strategy now is to announce $\{g^N,2\}$ which leads to the graph g^N and a net payoff vector of $(2.5-x, 4.5+x, 0.5)$. Thus, all SPE (conditional on the agents moving in the order 1,2,3 or 1,3,2) involve inefficient graphs.

The problem here is that agent 1 has no say regarding the links that she does not want to see formed. Agent 1 would like 2 and 3 to not form the link $\{2,3\}$ but there is no way to accomplish this objective. No matter how much agent 1 offers to contribute, agents 2 and 3 will always form the link $\{2,3\}$ since their gross payoff in the complete graph is higher. This problem does not arise when the utility function is monotonic because an agent would never have to compensate agents moving later to not form a link.

The obvious "fix" then is to allow each agent to specify the *entire graph* which she would like to see formed, not just the links that she wants to see formed. An announcement of player i in the modified game Γ_m is then $w_i=(g_i,x_i)$ where g_i is the graph that i wants to be formed and x_i her cost contribution if g_i does get formed. A coalition S is now compatible if (i) $g_i=g$ for all $i\in S$ where $g\in G_S$, and (ii) $\sum_{i\in S}x_i\geqslant c(g)$. The maximal compatible coalition is defined as before. If Γ_m^1 is the two-stage version of

[16] If $x<0$, then agents 2 and 3 will simply form the graph $\{2,3\}$ and divide the surplus of that graph between themselves.

Γ_m (in the same way that Γ_n^1 was a two-stage version of Γ_n), then we have the following result.

THEOREM 3. *Suppose that the functions v_i are non-negative for all $i \in N$ and the cost function satisfies Assumption 2. Then, all SPE of Γ_m and Γ_m^1 give rise to efficient graphs. The net payoffs to the agents in Γ_m are given by the "marginal contribution" vector $(u_1^*, \ldots, u_n^*)$ and in Γ_m^1 by the Shapley value of the TU-game (N, sa).*

Proof. The proof is omitted since it is almost the same as the proofs of Theorems 1 and 2. ∎

6. DIRECTED GRAPHS

So far, our analysis has focused on situations where the links are non-directional. However, some contexts, for example, information acquisition, can be modelled as situations involving directed links. Directed graphs have been analyzed in a couple of recent papers, notably those of Bala and Goyal [3] and Dutta and Jackson [5]. We now show that the mechanisms Γ_n and Γ_n^1 work equally well when the links are directed.

For any $S \subset N$, let $\bar{g}^S = \{(i, j) \in S \times S \mid i \neq j\}$ be the collection of all elements in $S \times S$ excluding those of the form (i, i). We can think of $\bar{g}^N$ as the *complete directed graph* on N. A *directed graph* g is some subset of $\bar{g}^N$. The element (i, j) of g is referred to as the *link* from i to j and denoted by (ij). (Note that in this notation, (ij) and (ji) are different.) Finally, let $\bar{G}_S = \{g \mid g \subset \bar{g}^S\}$ be the set of all graphs involving links only between members of S.

A *path* from i to j in g is a sequence of distinct agents $\{i_0, \ldots, i_K\}$ such that $i_0 = i$, $i_K = j$ and $(i_k i_{k+1}) \in g$ for all $k = 0, \ldots, K-1$. Let $\mu_i(g) = \{j \neq i \mid \exists$ a path from i to $j\}$. As before, the utility function of i is given by $U_i(g) = v_i(g) - x_i$ where $v_i(g)$ is i's gross benefit from g and x_i is his cost share. The following assumption, which is a direct analogue of Assumption 1, is imposed on v_i.

Assumption 3. The function v_i is non-negative and monotonic for all $i \in N$: $v_i(g) \geqslant 0$ for all g and $[\mu_i(g) \subseteq \mu_i(g')] \Rightarrow v_i(g) \leqslant v_i(g')$.

Remark 9. Assumption 3 implies that i is not made worse-off by links that other agents establish to her. Note that if monotonicity is violated, then we could have a problem similar to that illustrated in Section 5. The "fix" in this case would be a similar one as well.

The cost function is assumed to satisfy Assumption 2. An announcement of i is now a tuple (g_i, x_i) where $g_i \subset \bar{g}^N$ is the set of links that i wants to see formed and x_i is his contribution towards network formation.

DEFINITION 5. The coalition S is *compatible* with the announcement $w = \{(g_i, x_i)\}_{i=1}^n$ if

1. For all $i \in S$, $g_i \in \bar{G}_S$,
2. $\sum_{i \in S} x_i \geqslant c(g)$ where $g = \bigcup_{i \in S} g_i$.

The maximal compatible coalition is defined as before: it is the largest compatible coalition in the set $\{\varnothing, S_1, \ldots, S_n\}$.

It is straightforward, with these modified definitions, to prove the equivalents of Lemma 1 and Corollary 1 in this context. The proofs are omitted. We can now state the following combined result for the mechanisms Γ_n and Γ_n^1 as they apply to directed graphs. The proof is omitted as it involves similar arguments to Theorems 1 and 2.

THEOREM 4. *Suppose that Assumptions* 2 *and* 3 *are satisfied. Then, all SPE of Γ_n and Γ_n^1 result in the formation of efficient graphs. The net payoffs of the agents in all SPE of Γ_n is given by the "marginal contribution" vector* $(u_1^*, \ldots, u_n^*)$ *and in Γ_n^1 by the Shapley value of the TU-game* (N, sa).

7. COALITION STABILITY

Immunity to deviations by coalition is a desirable property of any mechanism. Unfortunately, our mechanisms do not possess this property, as the following example illustrates.

EXAMPLE 3. Consider a minimum spanning tree problem with four agents (1, 2, and 3) and a source (agent 0).[17] The cost function is additive in the set of links and the cost of establishing links are as follows: $c(\{0,1\}) = 2$, $c(\{0,2\}) = 50$, $c(\{0,3\}) = 4$, $c(\{1,2\}) = 3$, $c(\{1,3\}) = 50$, $c(\{2,3\}) = 3.5$. It is easily confirmed that the unique efficient network is $g^* = \{\{0,1\}, \{1,2\}, \{2,3\}\}$. Assume that each agent gets a utility of $V > 50$ if she is connected to the source and zero otherwise. A simple computation shows that $sa(1) = V - 2$, $sa(2) = V - 50$, $sa(3) = V - 4$, $sa(12) = 2V - 5$, $sa(13) = 2V - 6$, $sa(23) = 2V - 7.5$ and $sa(N) = 3V - 8.5$. Suppose that the agents move in the order 1, 3, 2. The reader can confirm that the following strategy profile constitutes a SPE of Γ_n: $\{(g^*, 1), (g^*, -42.5), (g^*, 50)\}$. However, in this SPE, agents 1 and 2 collectively pay \$51 even though they can connect themselves to the source at a cost of only \$5! This illustrates the fact that the outcomes resulting from Γ_n and Γ_n^1 may be vulnerable to coalitional deviations.

[17] The source is a nonstrategic agent, as in the classic formulation of the problem.

Thus, in general, our mechanisms will not be immune to coalitional deviations. We now identify sufficient conditions which will guarantee coalition stability. Before doing so, we impose minimal requirement on the utility and cost functions which rule out externalities between components. The reason for doing this is that coalitional stability is difficult to define in the presence of externalities. As mentioned before, this is similar to the problem of defining the core for a pure public good economy.

Assumption 4. 1. For all $i \in N$, for all $g, g' \in G_N$, if $h \in C(g) \cap C(g')$ and $i \in N(h)$, then $v_i(g) = v_i(g')$.

2. For all $g \in G_N$, $c(g) = \sum_{h \in C(g)} c(h)$.

These conditions are easy to understand. The first condition says that an agent's gross utility depends only on the connected component to which she belongs and is not affected by link formation outside this component. The second condition rules out cost externalities among connected components.

Remark 10. Assumption 4, by ruling out externalities, ensures that the stand-alone value of a coalition is a true reflection of the power of that coalition. In such a situation, the core of the game (N, sa) represents allocations that cannot be improved upon by any coalition.

Remark 11. Observe that the utility and cost functions of Example 3 satisfy Assumption 4. Thus, Assumption 4 by itself is not sufficient to guarantee coalitional stability.

The sufficiency conditions identified by us are based on the well-known result that in a convex game, the "marginal contribution" vector and the Shapley value are both core points. Therefore, if the TU-game (N, sa) is convex, then our mechanisms will be coalitionally stable. The following conditions, which are analogues of corresponding conditions identified by Moulin [15] for excludable public good economies, guarantee that the game (N, sa) is convex.

DEFINITION 6. A function $f: G_N \to R$ is *supermodular [submodular]* if $f(g_1) + f(g_2) \leqslant [\geqslant]\, f(g_1 \cup g_2) + f(g_1 \cap g_2)$ for all $g_1, g_2 \in G_N$.

The following lemma follows directly from the definitions of supermodularity and submodularity and its proof is hence omitted.

LEMMA 2. *Suppose that v_i is supermodular, non-negative and satisfies $v_i(\emptyset) = 0$ for all $i \in N$ and that the cost function c is submodular. Then the TU-game (N, sa) is convex.*

Remark 12. The conditions imposed in Lemma 2 while strong, are not trivial. For instance, *additive* cost functions (of the type used in Example 3) satisfy submodularity. An example of a utility function satisfying supermodularity is one where an agent's utility is a convex, increasing function of the number of people to whom he is connected.

We end this section with the following result.

THEOREM 5. *Suppose that the utility and cost functions satisfy Assumption 4. In addition, let v_i be non-negative, supermodular and satisfy $v_i(\emptyset) = 0$ for all $i \in N$ and c be non-negative and submodular. Then, all SPE of Γ_n and Γ_n^1 result in the formation of an efficient network. The net payoffs in all SPE of Γ_n are given by the "marginal contribution" vector and in Γ_n^1 by the Shapley value of the TU-game (N, sa). Furthermore, all SPE of Γ_n and Γ_n^1 are coalitionally stable.*

Proof. The proofs of the first two parts follow from Theorems 1 and 2. Coalition stability follows from Lemma 2. ∎

8. RELATIONSHIP WITH THE LITERATURE

The influential paper of Jackson and Wolinsky [12] used an extension of the cooperative game framework where the "value function" was the natural analogue of the characteristic function. The research question posed by Jackson and Wolinsky [12] was whether an allocation rule could be designed (for any arbitrary value function) so that agents acting in their own interest would form an efficient network. Recent research has relaxed the Jackson–Wolinsky framework by explicitly introducing costs of link formation and also endogenizing the payoffs but the approach for the most part has been to analyze games which are regarded as stylized descriptions of actual network formation processes. Our paper is related to this strand of research but it differs from the existing literature in looking at the problem as one of mechanism design.

The fact that we look at the problem as one of mechanism design and not as a stylized description of actual network formation situations means that certain issues which arise naturally in the latter context are of less importance here. For instance, issues like robustness of our results with respect to the game form, or whether our game is descriptive of real life network formation situations are of less importance in a mechanism design approach.[18] On the other hand, the simplicity of the mechanism and the

[18] The feature of sequential bids is, however, typical of many public good production situations.

fact that it can be implemented within a finite number of periods are of major importance in the mechanism design approach.

As an alternative to our finite horizon mechanism one might be interested in infinite horizon (bargaining type) models of the sort used by some authors to address coalition formation. We have not chosen this avenue here for two reasons. Firstly, our objective is not to model how network are formed but rather to propose a simple mechanism for network formation that guarantees efficiency. We view the finiteness of the mechanism as an advantage rather than a drawback. Secondly and on a more technical level it has been shown by several authors that multilateral (infinite horizon) bargaining models often give rise to inefficient equilibria. This was demonstrated among others by Okada [17], Seidmann and Winter [18], and in the context of a model supporting the Shapley value by Gul [10]. Indeed, as a descriptive model of network formation infinite horizon bargaining model may turn out to be a very useful tool. A challenging task would then be to characterize the environments under which the bargaining give rise to efficient networks. However this approach is outside the scope of the current paper.

Currarini and Morelli [4] have independently developed a sequential link formation game with endogenous payoff determination in the Jackson–Wolinsky framework.[19] In their model, agents announce sequentially and agent i's announcement is a tuple specifying the subset of agents with whom she would like to establish links and a payoff demand. The payoff demand can be either an absolute claim on the total surplus or a vector specifying a claim for each proposed link or a number specifying a share of the total surplus. Currarini and Morelli [4] show that if the value function is *size monotonic* and the payoff demand is either an absolute claim on total surplus or a vector of "bilateral" claims, then all SPE lead to the formation of efficient networks.

Our model is different from that of Currarini and Morelli in a number of respects. First, the mechanisms differ in terms of the agents announcements and in the outcomes they implement. While Currarini and Morelli only require that an agent announce the set of links that she wants to form, we allow an agent to announce the complete set of links that she wants to see formed. We thus allow an agent to have some say about links not involving the agent herself. This may appear strange, but it should be noted that the formation of links not involving the agent herself *can* impact the agent in a negative way. It should also be noted that monotonicity is not a strong assumption in our context because the subgraph relation is only a partial

[19] Currarini and Morelli [4] motivate their paper by saying that they are dealing with situations where there is no planner available to enforce an allocation rule. Their games are thus to be understood as stylized descriptions of actual network formation situations.

order. One can easily construct an example with monotonic utility functions in which announcing just the set of links that an agent wants to form along with cost contributions results in inefficiency. In terms of outcomes, the two mechanisms analyzed here yield the "marginal contributions" vector and the Shapley value (with respect to an appropriate cooperative game) respectively. The games analyzed by Currarini and Morelli do not yield these outcomes.

Secondly, Currarini and Morelli [4] require that the value function be both *anonymous* and *component additive*. These assumptions rule out externalities between components, but as discussed in Section 2, we do allow for such externalities in the utility and cost functions. Neither do we impose any sort of symmetry *a priori* on the agents.

More importantly, the game Γ_n—unlike the Currarini–Morelli game $\Gamma_1(v)$—is *not* invariant to the "net payoff" function.[20] This is illustrated below.

EXAMPLE 4. Consider the same setup as that in Example 2. Since $c(g) = 0$ for all g, we have the following induced value function[21]

$$v(g) = \begin{cases} 0 & \text{if } \; g = \varnothing, g^2 \;\text{ or }\; g^3, \\ 6 & \text{if } \; g = \{\{i, j\}\}, \\ 8 & \text{if } \; g = g^1 \\ 7.5 & \text{if } \; g = g^N. \end{cases} \tag{6}$$

Consider now a different model with the specifications

$$\bar{u}_i(g) = \begin{cases} 10 & \text{if } i \text{ is not isolated in } g, \\ 0 & \text{otherwise.} \end{cases}$$

$$\bar{c}(g) = \begin{cases} 0 & \text{if } \; g = \varnothing, \\ 14 & \text{if } \; g = \{\{i, j\}\}, \\ 22 & \text{if } \; g = g^1, \\ 30 & \text{if } \; g = g^2 \text{ or } g^3, \\ 22.5 & \text{if } \; g = g^N. \end{cases}$$

[20] We are grateful to the associate editor for pointing this out to us.

[21] Recall that the induced value function is given by $v(g) = \sum_{i \in N} u_i(g) - c(g)$ for all $g \in G_N$.

It can be checked that $\{\bar{u}_i\}_{i \in N}$ and $\bar{c}$ give rise to the same value function as that in (6). However, $\{\bar{u}_i\}_{i \in N}$ and $\bar{c}$ satisfy Assumptions 1 and 2 and therefore Theorems 1 and 2 imply that both Γ_n and Γ_n^1 yield g^1 in all SPE. In contrast, Example 2 yields different networks in equilibrium.

In general, given *any* value function v satisfying $v(\varnothing) \geqslant 0$, we can always find utility and cost functions satisfying Assumptions 1 and 2 which give rise to v. Indeed, letting $M > \max_{g \in G_N} v(g)$, we can select the utility and cost functions as follows:

$$u_i(g) = \begin{cases} v(g)/n & \text{if } \quad g = \varnothing, \\ M & \text{otherwise.} \end{cases}$$

$$c(g) = \begin{cases} 0 & \text{if } \quad g = \varnothing, \\ nM - v(g) & \text{otherwise.} \end{cases}$$

On the other hand, there are circumstances where the Currarini–Morelli game $\Gamma_1(v)$ gives rise to efficient graphs while our mechanisms do not do so. An example is the Jackson–Wolinsky co-author model with $|N| = 3$, a modified version of which was discussed in Example 2.[22] The value function of this example satisfies *size monotonicity* and hence by Theorem 2 of Currarini and Morelli [4], all SPE of $\Gamma_1(v)$ give rise to efficient graphs.[23] The important point here is that "claims on net payoffs" is different from "contribution towards cost" and *size monotonicity* neither implies nor is implied by our Assumptions 1 and 2.

Johnson and Gilles [14] examine a model of costly network formation based on the *symmetric connections* model of Jackson and Wolinsky [12]. In their model, an individual's gross benefits from a graph is given by a utility function whose specification is the same as the *symmetric connections* model of Jackson and Wolinsky [12]. The agents are assumed to be located on a line and the cost of linking agents i and j is given by a *cost topology* which has the feature that the cost of establishing a link is less if the two agents are "near" to each other. As in Jackson and Wolinsky [12], the net benefit of an agent is the difference between his gross benefits and the total cost associated with those links in which she is involved. Johnson and Gilles [14] characterize the set of *pairwise stable* graphs as well as the set of efficient graphs in their model and show that they may not coincide.

[22] Example 2 differs slightly from the Jackson–Wolinsky co-author model. The Jackson–Wolinsky specification is obtained by using the formula on the first line of (5) for *all* situations where k is not isolated.

[23] A link $\{i, j\}$ in a graph is said to be *critical* if deleting that link increases the number of components in the graph. A value function satisfies *size monotonicity* if for all critical links $\{i, j\} \in g$, $v(g) > v(g \setminus \{i, j\})$.

They then analyze a sequential move game and show that under certain conditions on parameters, one can find an order of moves such that the corresponding SPE involves a pairwise stable graphs. In the Johnson–Gilles game, a pair of agents move at each stage, in a manner similar to the game analyzed by Aumann and Myerson [1]. However, in contrast to the analysis in this paper as well as that in Currarini and Morelli [4], the Johnson–Gilles game is a pure linking game where the only decision for an agent at any stage is whether to form the link with the other agent with whom he is paired. Since no transfers are allowed between agents, it is not surprising that the SPE of the Johnson–Gilles game need not give rise to efficient graphs which was one of our prime objectives.[24]

Slikker and van den Nouweland [20] have also analyzed network formation with endogenous payoff determination but they have used a simultaneous move game to do so. In their game, agents simultaneously announce a subset of agents with whom they wish to link as well as a payoff claim. Slikker and van den Nouweland [20] confine themselves to value functions which can be derived from cooperative games.[25] Their main result is that the equilibrium graph will not have cycles and that a player may not be able to profit from a "central" position in the graph. In another paper, Slikker and van den Nouweland [19], again confining themselves to value functions derived from cooperative games, examine a game similar to the Johnson–Gilles one when there is a fixed cost $c > 0$ of forming a link and the allocation rule is given by the *Myerson value*. They show that as costs increase, the pattern of the resulting equilibrium graph depends on whether the underlying cooperative game is superadditive and/or convex. These papers have a somewhat different focus than ours and the results obtained there are not really comparable to our results.

Finally, our paper is related to the tiny literature in computer science on cost sharing in multicast trees examined in the works of Feigenbaum *et al.* [7], Herzog *et al.* [11], and Jain and Vazirani [13]. The problem of multicast transmission involves a given computer network and a set of users who each desire to receive a message (like a movie). The willingness to pay for the message of any agent is private information. The planner's problem is to design a mechanism (typically, the literature focuses on dominant strategy mechanisms) which elicits the willingness to pay of each agent,

[24] Johnson and Gilles [14] refer to their result as an 'implementation' result. However, it is not clear why they rule out transfers amongst agents. Our results indicate that the planner can do better by allowing for such transfers.

[25] A value function v is said to be derived from a cooperative game if it is component additive and the following holds: Let g, g' be graphs such that for every $h \in C(g)$, there exists $h' \in C(g')$ such that $N(h) = N(h')$. Then, $v(g) = v(g')$. Thus, two connected components with the same player set have the same value.

then selects the subset of agents to receive the message, a way of routing the message to this subset of agents and an assignment of costs (the sum of the costs of the links used in the selected route) across the agents. The works cited above show that there are no dominant strategy and efficient mechanisms (a result similar to that obtained by Green and Laffont [9]) and focus on dominant strategy mechanisms which are budget balanced but inefficient. One important component of these works is the attention given to *computability*: the papers of Feigenbaum *et al.* [7] and Jain and Vazirani [13] both contain explicit algorithms for implementing the mechanisms. This is important because determining the optimal route for any subset of agents is known to be computationally NP-hard. Economists have not paid much attention to computability but the papers of Bala and Goyal [3] and Johnson and Gilles [14] which contain simulation results illustrate the complex computations involved in even the simplest examples. We have two observations with regard to this literature. First, given our "complete information" assumption, it is not surprising that we obtain more positive results: in particular, we obtain efficiency and to an extent, also equity. Second, regarding computability, it is difficult to say anything given the generality of our model, although it is obviously an important issue.

9. CONCLUSION

In this paper, we have examined a model of network formation with costs where individual benefits from network formation are not known to the planner. The cost function is assumed to be common knowledge in the society. We have proposed two mechanisms for this problem with quasi-linear and monotonic preferences. The only restrictions imposed on the cost function are that there are no fixed costs. In both mechanisms, players move sequentially. Each agent's announcement is a tuple consisting of the set of people with whom he wants to form links, and a monetary contribution, interpreted as the player's contribution towards the cost of network formation. Our first mechanism ensures the formation of an efficient network in all subgame perfect Nash equilibria; however the net payoffs to the agents are asymmetric. In general, agents moving earlier are better off than agents moving later. The second mechanism corrects for this asymmetry and ensures not only the formation of efficient networks but also equitable net payoffs. We also discuss the extension of the basic model to cover the case of nonmonotonic preferences and directed graphs. Finally, we discuss conditions under which the mechanisms we propose are immune to coalitional deviations.

REFERENCES

1. R. Aumann and R. Myerson, Endogenous formation of links between players and coalitions: An application of the Shapley value, *in* "The Shapley Value" (A. Roth, Ed.), Cambridge Univ. Press, Cambridge, UK, 1988.
2. P. Bag and E. Winter, Simple subscription mechanisms for excludable public goods, *J. Econ. Theory* **87** (1999), 72–94.
3. V. Bala and S. Goyal, A non-cooperative model of network formation, *Econometrica* **68** (2000), 1181–1229.
4. S. Currarini and M. Morelli, Network formation with sequential demands, *Rev. Econ. Design* **5** (2000), 229–249.
5. B. Dutta and M. Jackson, The efficiency and stability of directed economic networks, *Rev. Econ. Design* **5** (2000), 251–272.
6. B. Dutta and S. Mutuswami, Stable networks, *J. Econ. Theory* **76** (1997), 322–344.
7. J. Feigenbaum, C. Papadimtriou, and S. Shenker, Sharing the cost of multicast transmission, mimeo, 2000.
8. R. Graham and P. Hall, On the history of the minimum spanning tree problem, *Ann. History Comput.* **21** (1981), 1–18.
9. J. Green and J.-J. Laffont, Incentives in public decision making, *in* "Studies in Public Economics," Vol. 1, North-Holland, Amsterdam, 1979.
10. F. Gul, Bargaining foundations of the Shapley value, *Econometrica* **57** (1989), 81–95.
11. S. Herzog, S. Shenker, and D. Estrin, Sharing the "cost" of multicast trees: an axiomatic analysis, mimeo, 1996.
12. M. Jackson and A. Wolinsky, A strategic model of social and economic networks, *J. Econ. Theory* **71** (1996), 44–74.
13. K. Jain and V. Vazirani, A budget balanced, group strategyproof cost sharing method for a non-submodular function, mimeo, 1999.
14. C. Johnson and R. Gilles, Spatial social networks, *Rev. Econ. Design* **5** (2000), 273–299.
15. H. Moulin, "Cooperative Microeconomics: A Game-Theoretic Introduction," Princeton Univ. Press, Princeton, NJ, 1995.
16. R. Myerson, Graphs and cooperation in games, *Math. Oper. Res.* **2** (1977), 225–229.
17. A. Okada, A non-cooperative coalitional bargaining game with a random proposer, *Games Econ. Behav.* **16** (1996), 97–108.
18. D. Seidmann and E. Winter, A theory of gradual coalition formation, *Rev. Econ. Stud.* **65** (1998), 793–815.
19. M. Slikker and A. van den Nouweland, Network formation models with costs for establishing links, *Rev. Econ. Design* **5** (2000), 333–362.
20. M. Slikker and A. van den Nouweland, A one-stage model of link formation and payoff division, *Games Econ. Behav.* **34** (2001), 153–175.