Formation of Networks and Coalitions

Francis Bloch
Ecole Polytechnique

and

Bhaskar Dutta
University of Warwick

January 2010

1 Introduction

Although much of formal game theory focuses on non-cooperative behavior, sets of individuals often cooperate with one another to promote their own self-interest. For instance, groups of countries form customs unions, political parties form coalitions, firms form cartels and so on. There is also growing awareness that there are a variety of contexts where the particular structure or *network* of interactions has a strong influence on economic and social outcomes. Typically, groups and networks form through the *deliberate* choice of the concerned individuals. Given the influence of groups and networks on the eventual outcome, it is important to have some idea of what kind of coalitions and networks may form. In this survey, we focus on this issue by discussing some recent literature on the endogenous formation of coalitions and networks.

Given the huge literature around this area, we have had to be very selective in what we cover in this survey.[1] For instance, given our focus on the *formation* of networks, we will not discuss the growing and fascinating literature on how the structure of interactions affect outcomes given a *fixed* network. In the area of cooperative game theory, we will eschew completely any discussion of solution concepts like the Shapley value. Neither will we have anything to say on solutions like the various versions of the bargaining set which essentially assume that the coalition of all players will necessarily form. Rather, we focus on coalition formation processes where there is no a priori reason to expect the grand coalition to form.[2]

In much the same spirit, while we do discuss bargaining processes which determine network structures or coalition structures along with individual payoffs, we do not cover the huge literature on bargaining and markets. Perhaps more contentious is our decision to omit the very interesting literature on the so-called Nash program. We decided to leave out this literature because the focus of this literature is *normative* in nature. That is, the objective has often been to describe (for instance) bargaining procedures which sustain some given solution concept rather than to focus on procedures which are *actually used*.[3]

The plan of the paper is the following. In the next section, we first

[1] For more exhaustive surveys of group and network formation, we refer the reader to the collection of surveys in Demange and Wooders (2005), and to the recent books on coalitions by Ray (2008) and on networks by Goyal (2007) and Jackson (2008).

[2] In that respect, we share Maskin (2003)'s view that "Perhaps one reason that cooperative theory has not been more influential on the mainstream is that its two most important solution concepts for games of three or more players, the core and Shapley value, presume that the grand coalition – the coalition of all players– always forms. And thus the possibility of interaction between coalitions – often important in reality – is ruled out from the beginning."

[3] Serrano (2005) is an excellent survey on the Nash program.

describe a general framework which encompasses both coalition structures and network structures. In this section, we also discuss various issues that arise in the formation of groups through one-stage processes or normal form games when players are *myopic*. Section 3 focuses on bargaining or multi-stage group formation procedures. In section 4, we go back to one-stage models but assume that players are *farsighted*. Section 5 describes some recent literature on group formation in a dynamic setting. In section 6, we discuss the tension between efficiency and stability, essentially in the context of network formation. Section 7 concludes and discusses open questions for future research.

2 One-stage Models of Coalition and Network Formation

In this section, we are concerned with issues which arise when agents form networks or coalitions by means of normal form games. That is, agents simultaneously choose which coalitions (or what links to form). Given these simultaneous choices, different "rules of the game" determine exactly what coalitions or networks will actually form.

2.1 A General Framework

We describe a general framework within which we can discuss several issues connected with the one-shot (or simultaneous) formation of both networks and coalitions.

Consider a social environment where $N = \{1, \ldots, n\}$ is a finite set of agents or players, while X is the set of *social states*. Each individual i has a preference relation $\succeq_i$ over X. Let $\succ_i$ denote the strict preference relation corresponding to $\succeq_i$. The power of different groups to change the social state is represented by an *effectivity relation* $\{\rightarrow_S\}_{S \subseteq N}$ on $X \times X$, where for any $x, y \in X$, $x \rightarrow_S y$ means that the coalition or group S has the means to change the state to y if the "status quo" is x. So, a social environment is represented by the collection $\mathbb{E} = (N, X, \{\succeq_i\}_{i \in N}, \{\rightarrow_S\}_{S \subseteq N})$.

Several different examples fit this general description of a social environment.

Normal Form Games : For each $i \in N$, let S_i denote the strategy set of player i, $S^N \equiv \prod_{i \in N} S_i$, while $u_i : S^N \rightarrow \mathbb{R}$ is player i's payoff function. Now, letting $X = S^N$, u_i is simply the numerical representation of the preference relation $\succeq_i$. Finally, the effectivity relation also has a natural interpretation. Consider $x = (s_1, \ldots, s_n)$ and $y = (s'_1, \ldots, s'_n)$. Then, for any group $S \subseteq N$, $x \rightarrow_S y$ iff $s_i = s'_i$ for all $i \notin S$.

Undirected Networks: A rich literature models social and economic interactions by means of networks or graphs. Identify N with the set of nodes. An arc exists between nodes i and j if i and j "interact bilaterally". The network is undirected if bilateral interaction is a symmetric relation. The specific economic or social context being modeled gives meaning to the term bilateral interaction. For instance, the nodes may be a set of firms, and bilateral interaction may refer to firms i and j collaborating in a research joint venture.[4] Alternatively, the graph may represent a friendship (or connections) network where i and j are linked if they are "friends".[5]

Let G be the set of all possible undirected networks with N as the set of all nodes. A *value function* v specifies the aggregate value of each graph, while Y is an *allocation rule* which specifies the payoff corresponding to each value function and each network.

Fix some value function. For simplicity, ignore the dependence of payoffs on the value function. So, $Y_i(g)$ denotes the payoff to i corresponding to a network $g \in G$. Letting $X = G$, we can now identify Y_i as the numerical representation of $\succeq_i$.

The implicit assumption underlying models of the strategic formation of undirected networks is that a link between any pair i and j can form only if both agents decide to form the link. However, an existing link (ij) can be broken unilaterally by either i or j. We will formally define these "rules" of network formation subsequently. However, this informal description is sufficient to describe the relevant effectivity relation. Consider any $S \subseteq N$, and any pair $g, g' \in G$. Then,

$$g \rightarrow_S g' \Leftrightarrow \big(ij \in g' - g \Rightarrow \{i,j\} \subseteq S \text{ and } ij \in g - g' \Rightarrow \{i,j\} \cap S \neq \emptyset\big)$$

Directed Networks: The main difference between directed and undirected networks is that in the former, arcs are directed. So, i can be "connected" to j without j being connected to i. For instance, i can access j's homepage, but j need not access i's webpage. Let G^d be the set of all directed networks on node set N. It is standard to assume that i does not need j's consent to form the directed link to j. So, for any pair $g, g' \in G^d$ and subset S of N,

$$g \rightarrow_S g' \Leftrightarrow (ij \in (g - g') \cup (g' - g) \Rightarrow \{i,j\} \cap S \neq \emptyset)$$

Characteristic Function Games: The cornerstone of cooperative game theory is the *characteristic function.* A (TU) characteristic function

[4] See Goyal and Joshi (2003).

[5] The connections model is due to Jackson and Wolinsky (1996)(henceforth JW). See also Bloch and Dutta (2009) for an analysis which incorporates strength of links in a friendship network.

game is a pair (N, v) where v is the characteristic function describing the "worth" of every coalition. The worth of a coalition is the maximum aggregate utility that a coalition can *guarantee* itself.

For every coalition S, let $A(S) = \{x \in \mathbb{R}^N | \sum_{i \in S} x_i \leq v(S)\}$. So, $A(N)$ is the set of feasible allocations for the grand coalition, while $A(S)$ is the set of allocations which gives members of S what they can get on their own. Identify X with $A(N)$.

Clearly, $x \succeq_i y$ if and only if $x_i \geq y_i$. Also, the effectivity relation is easy to describe. For any x and y,

$$x \to_S y \text{ iff } y_S \in A(S)$$

That is, the coalition S can enforce any social state y in $A(S)$ if the sum of the payoffs to individuals in S does not exceed the worth of S.

A straightforward extension to NTU characteristic function games is readily available. In the NTU version, members of a coalition cannot transfer payoffs amongst themselves on a one-to-one basis. For instance, the situation being modeled may not have any "money" (more generally a private good). Alternatively, even if the model has money, players' utilities may not be linear in money. For instance, consider the familiar exchange economy in which individuals have (ordinal) preferences defined over the commodity space. Individuals also have endowments of goods, and can trade with each other. So, the worth of any coalition is the set of feasible utility vectors that the coalition can get by restricting trade to within the coalition.

Thus, the NTU characteristic function specifies a *set* $V(S)$ of feasible utility vectors for each coalition S. So, $x \to_S y$ if the restriction of y to S is in $V(S)$.

Hedonic games without externalities are "ordinal" versions of characteristic function games in which players are partitioned into groups or communities, and each player's payoff is solely determined by the identity of other members in her coalition.[6] So, each player i has a preference ordering over the set of coalitions to which i belongs. Examples of group interaction which fit this description include the formation of social clubs, local communities which provide local public goods such as roads, etc. Clearly, such games also fit into the general framework outlined here.

Games in partition function form: Characteristic functions (in either the TU or NTU version) cannot adequately describe environments in which there are significant *externalities* across coalitions - the notion of what a coalition can guarantee itself is not always meaningful. For consider situations where the payoff to a coalition S depends on the actions taken by the complementary coalition. Clearly, it may not be in the interest of the

[6]This terminology is due to Dréze and Greenberg (1980).

complementary coalition to take actions which minimize payoffs to S. For instance, in a Cournot oligopoly where each firm has a "large" capacity, payoffs to S are minimized at zero if the firms outside S produce so much output that prices are driven to marginal cost. But, it makes no sense for these firms to do so. So, S has to make some predictions about the behavior of its opponents.

A more general representation-one which incorporates the possibility of externalities- is the *partition function form*. Let Π_S denote the set of all partitions of any coalition $S \subset N$. For any coalition S, S^c denotes the set $N \setminus S$. Call objects of the form $(S; \pi(S^c))$ *embedded coalitions*. Then, (N, w) denotes a game in partition function form, where w specifies a real number for every embedded coalition. We represent this as $w(S; \pi(S^c))$.[7] Notice that this definition incorporates the possibility of externalities since the worth of a coalition depends on how the complementary coalition is organised.[8]

In analogy with the earlier example, one can identify the set of social states with the set of embedded coalitions. Now, suppose $x = (S; (S_1, \ldots S_K))$ where $(S_1, \ldots, S_K)$ is some partition of S^c, and consider any subset T of N. What social state can T induce from x? Suppose members of coalition T believe that once they leave their current coalitions, all others will stay in their original coalitions. That is, there will not be any further reorganization of members. Let $T_0 = S \setminus T$, and $T_k = S_k \setminus T$ for each $j = 1, \ldots, K$, and $y \equiv (T; (T_0, T_1, \ldots, T_K))$. Then, under the *myopic* assumption that players in T^c will "stay put", we can write $x \to_S y$. However, the assumption of myopic agents is typically an assumption of convenience, and we will consider alternatives notions of "farsightedness".

2.2 Models of coalition and network formation

In this subsection, we describe different one-shot noncooperative procedures by which agents form coalitions and networks. These are all procedures which give rise to normal form games. Two classes of models have been discussed in the literature. In the first class, individuals are precluded from transferring money or utility. In these models, strategies are simply an

[7]Hedonic games with externalities are ordinal counterparts of games in partition function form. In such games, a player has a preference ordering over the set of all possible partitions of N.

[8]The derivation of a game in partition function form from a game in normal form is not without problems. One possibility is to treat each coalition as a single entity, and then assume that each such entity plays a non-cooperative game amongst each other. *If* for every partition of N, this non-cooperative game has a *unique* Nash equilibrium, then the unique equilibrium payoff for S corresponding to each $\pi(S^c)$ can be identifies with $w(S; \pi(S^c))$.

announcement of the other players with whom a player wants to form a coalition or link. In the second class of models, individual strategies are *bids* to "buy" resources of other agents or "transfers" or "demands" to set up links with other agents. These bids, transfers and demands are in money or utility. Once again, the rules of the game specify what coalitions or networks form, and *net* payoffs now depend on both the solution concept as well as the bids and transfers. We describe these different models in some detail below.

Models without transfers

The earliest model of coalition formation was proposed by von Neumann and Morgenstern (1944, pp. 243-244). Each player i announces a coalition $S(i)$ to which she wants to belong. The outcome function assigns to any vector of announcements $S(1), ..., S(n)$, a coalition structure π as follows: $S \neq \{i\} \in \pi$ if and only if, for all agents $i \in S$, $S(i) = S$. A singleton i belongs to the coalition structure π if and only if (i) either $S(i) = \{i\}$ or $S(i) = S$ and there exist $j \in S$ such that $S(j) \neq S$. In this procedure, a coalition is formed if and only if *all its members unanimously agree to enter the coalition.*

This procedure was rediscovered by Hart and Kurz (1984), who labeled it 'model γ'. They contrast it with another procedure, labeled 'model δ', where unanimity is not required for a coalition to form. In the δ procedure, the outcome function assigns to any vector of announcements $S(1), ..., S(n)$, a coalition structure π where: $S \in \pi$ if and only if $S(i) = S(j) \supseteq S$ for all $i, j \in S$. In other words, coalitions are formed by any subset S of players who coordinate and announce the same coalition $S(i)$. In this procedure, the announcement serves to coordinate the actions of the players, and indicates what is the largest coalition that players are willing to form.

Myerson (1991) proposes a game of undirected network formation which is very similar to models γ and δ. In particular, agents simultaneously announce the set of agents with whom they want to form links. Hence, a pure strategy in the game is a subset $S(i) \subseteq N \setminus \{i\}$ for every agent i. The formation of a link requires *consent* by both parties. Link ij is formed if and only if $i \in S(j)$ and $j \in S(i)$. Myerson's model is well suited to handle situations where both agents need to agree to form a link (e.g. friendship relations, formal agreements).

In contrast to Myerson, Bala and Goyal (2000) study the formation of directed networks where agents do not need the consent of the other party to form a link. They consider situations (like the formation of communication links) where agents can freely build connections to the existing network. In these situations, one of the two agents initiates the link and incurs its cost. So, every agent announces a subset $S(i)$ of $N \setminus i$, and the directed link $i \to j$ is formed if and only if $j \in S(i)$. Bala and Goyal distinguish between two specifications of payoffs. In the *one-way flow model*, the agent initiating the link is the only one to derive any benefit from the link. In the *two-way flow model*, one agent incurs the cost of forming the link, but both agents benefit from the link formed.

Both models γ and δ are models of *exclusive membership*: players can exclude other players from a coalition by their announcements. Myerson's link formation also has this feature. Other procedures do not give players the ability to exclude other agents from the coalition: these are games of *open membership.* For example, the procedure proposed by d'Aspremont et al. (1983) to study the formation of a cartel is defined as follows: players announce their willingness to participate in the cartel (either 'yes' or 'no'). A cartel is formed by all the players who have announced 'Yes'. Alternatively, the equilibria of the cartel formation game can be characterized by the following two conditions of *internal* and *external stability.* Let $v_i^I(S)$ define the profit of an insider in cartel S and $v_i^O(S)$ the profit of an insider when cartel S forms. A cartel is internally stable if no member of the cartel wants to leave, $v_i^I(S) \geq v_i^O(S \setminus \{i\})$ for all $i \in S$. A cartel is externally stable if no outsider wants to join the cartel, $v_i^O(S) \geq v_i^I(S \cup \{i\})$ for all $i \notin S$. One drawback of the procedure of d'Aspremont et al. (1983) is that it only allows one coalition to form. The procedure can easily be generalized to the following open membership game. Every player announces an address $a(i)$ (taken from a set A of cardinality greater than $n+1$, and with a distinctive element a_0). A coalition S is formed if and only if $a(i) = a(j) \neq a_0$ for all $i, j \in S$. Coalitions are formed by players who announce the same address. Players also have the opportunity to remain singletons by announcing the particular address a_0.

In all procedures defined above, the decision to participate in a group or to form a link was modeled as a discrete $\{0, 1\}$ choice. In reality, agents may choose the amount of resources they spend in different groups and on different links, resulting in a continuous model of participation and link formation. Bloch and Dutta (2009) and Rogers (2005) study this issue in models of link formation. They assume that agents select how to allocate fixed resources X_i on different links. In their models, agents thus choose a vector of investments on every link, $x^i = (x_1^i, x_2^i, ..., x_n^i)$ such that $\sum_j x_j^i = X_i$ for all i. Individual investments are transformed into link quality by a production function, $s_{ij} = f(x_j^i, x_i^j)$, assigning a number between 0 and 1, the quality of the link, as a function of individual investments. The outcome of the link formation game is thus a *weighted network* where links have different values. Similarly, one can consider a model of group participation where agents select the amount of resources they devote to different activities. If there are K different activities or tasks to perform, every agent chooses a vector $x^i = (x_1^i, x_2^i, ..., x_K^i)$ satisfying $\sum_k x_k^i$. These resources are combined in groups to produce surplus, according to a family of production functions, $v(S, k, (x_k^i)_{i \in S})$. In this formulation coalitions are overlapping in the sense that the same player may belong to different coalitions.

Models with transfers

In the games presented in the previous sections, agents were precluded from transferring money or utility. States were defined as coalition structures or networks, and did not include a description of individual payoffs achieved by the players. We now introduce one-stage procedures of coalition or network formation where agents are allowed to transfer utility.

Kamien and Zang (1990)'s model of monopolization in a Cournot industry was originally designed to study mergers in Industrial Organization. However, the first period game of coalition formation that they introduce is quite general and can be applied to any problem of coalition formation. They suppose that every agent i submits a vector of *bids*, b_j^i over all agents j in N. A bid b_j^i for $i \neq j$ is interpreted as the amount of money that agent i is willing to put to acquire the resources of agent j. A bid b_i^i is interpreted as the asking price at which agent i is willing to sell her resources. Given a matrix $B = [b_i^j]$ of nonnegative bids, one can assign the resources of every agent i either to another agent j (or to agent i herself, if she remains a singleton). Formally, let

$$S(i) = \{j \in N, j \neq i, b_i^j \geq b_i^k \forall k \neq j\}$$

denote the set of players other than i such that (i) the bid they offer is no smaller than the bid of any other player and (ii) the bid they offer is higher than the asking price. If $S(i)$ is a singleton, the assignment of the resources of player i to the unique player in $S(j)$ (and hence the formation of a coalition S containing $\{i, j\}$) is immediate. If $S(i)$ is not a singleton, one needs to define an exogenous tie-breaking rule to assign the resources of player i to some member of $S(i)$. As a result of this bidding procedure, resources of some players are bought by other players, resulting both in the formation of a coalition structure π and in transfers across players given by $t_i^j = b_i^j$ and $t_j^i = -b_i^j$ if player j acquires the resources of player i.[9] Multi-bidding games have later been extended by Perez Castrillo and Wettstein (2001) who have also uncovered a connection between bidding mechanisms and the Shapley value.[10]

[9] Perez Castrillo (1994) independently proposed a procedure of coalition formation which bears close a resemblance to Kamien and Zang (1990)'s bidding game. The main difference is that Perez Castrillo introduces competitive outside players (the "coalition developers") who simultaneously bid for the resources of the players.

[10] See also Macho Stadler, Perez Castrillo and Wettstein (2006) for partition function games, and, in the context of networks, Slikker (2007).

Bloch and Jackson (2007) extend Myerson's model of link formation to allow transfers among agents.[11] In their basic setting, every agent announces a vector of bids, $t^i = (t^i_j), j \neq i$. The bid t^i_j may be positive (and then interpreted as an offer to pay t^i_j to player j), or may be negative (and then interpreted as a demand to receive t^i_j) from player j. Given the simultaneous announcement of bids and the matrix $T = [t_{ij}]$, links are formed and transfers made as follows: If $t^i_j + t^j_i \geq 0$, the link between i and j is formed, and players pay (or receive) the transfer that they offered (or demanded). Given this specification, it may be that transfers are wasted out of equilibrium if $t^i_j + t^j_i > 0$. Alternative transfer procedures could be specified, without altering the network formed in any equilibrium of the procedure. Bloch and Jackson then proceed to define richer structures of transfers, where players are not constrained to put money only on the links they form. In one model, players can choose to subsidize links formed by other players, by announcing positive transfers t^i_{jk} on links formed by other players ; in another model, players can announce negative transfers in order to prevent the formation of a link by other players. Finally, in the most general setting, Bloch and Jackson allow players to announce positive or negative transfers contingent on the entire network formed.

2.3 Stability

The processes of group formation described above tell us *how* networks or coalitions form, but not *which* group(s) will actually materialize in any specific context. Since these processes yield well-defined normal form games, it is natural to use game-theoretic notions of equilibrium to predict the network or coalition structure that will be formed. In this section, we describe some of the equilibrium concepts which are relevant when agents are "myopic". We clarify the meaning of this term shortly.

Consider a social environment which is represented by the collection $\mathbb{E} = (N, X, \{\succeq_i\}_{i \in N}, \{\to_S\}_{S \subseteq N})$. Which social states are likely to emerge as social outcomes following strategic interaction amongst the agents?

Definition 1 *A social state $x \in X$ is a k-equilibrium if there is no set $S \subseteq N$ with $|S| \leq k$ such that there is $y \in X$ with $x \to_S y$ and $y \succ_i x$ for all $i \in S$.*

Implicit in this definition of stability is the idea that when a coalition contemplates a deviation from a social state x which is "on the table" to a state y, it compares the utilities associated with x and y. The deviating

[11] Slikker and van den Nouweland (2001) introduced an early model of link formation with transfers where agents, in addition to forming links, submit claims on the value of the network.

coalition does not consider the possibility that y itself may not be stable. That is, it does not take into account the possibility that there may be further round(s) of deviation from y. This is the sense in which the current notion of stability is relevant only when players are "myopic". In a later section, we will consider different definitions of stability for players who "look ahead".

This general definition encompasses several notions of stability that have been used in the literature. For instance, if $k = 1$, then it is analogous to Nash equilibrium. However, in most settings of group formation, Nash equilibrium hardly has any predictive power.[12] Consider for example a setting of hedonic games (with or without externalities) where players prefer to belong to *some* coalition rather than remain single. Then, *any coalition structure* can be sustained as an equilibrium of the γ game of coalition formation. To see this, fix any coalition structure π. Let $S \in \pi$, and suppose individual $i \in S$. If all players in $S \setminus \{i\}$ announce the coalition S, it is a best response of player i to also announce S, even though she may prefer another coalition.[13]

Given the typical indeterminacy of Nash equilibrium particularly in models of undirected networks, it is not surprising that other equilibrium notions have been considered in the literature. Because it takes agreement of both players i and j to form the link ij, it is natural to consider coalitions of size two since this is the minimal departure from a purely non-cooperative equilibrium concept. JW specified a very weak notion of stability for undirected networks.

Definition 2 *A network g is pairwise stable if for all $i, j \in N$,*
(i) $Y_i(g) \geq Y_i(g - ij)$
(ii) $Y_i(g + ij) > Y_i(g)$ implies that $Y_j(g + ij) < Y_j(g)$.

This concept of stability is very weak because it restricts deviations to change only *one* link at a time - either some agent can delete a link or a pair of agents can add the link between them. This notion of stability is not based on any specific procedure of network formation. A stronger concept of stability based on bilateral deviations uses Myerson's network formation game.

Definition 3 *A 2-equilibrium s^* of Myerson's game is a Pairwise Nash equilibrium.*

[12] Somewhat surprisingly, it turns out that pure strategy Nash equilibria do not necessarily exist in Bala and Goyal's model with *heterogeneous* agents. See Galeotti, Goyal and Kamphorst (2006), Billand and Bravard (2005) and Haller and Sarangi (2005).

[13] Similar results are available in the context of both directed and undirected networks. See Bala and Goyal (2000) and Dutta, Tijs and van den Nouweland (1998).

Bloch and Jackson (2007), Calvo-Armengol and Ilkilic (2009) and Gilles, Chakrabarti and Sarangi (2006) analyze the relation between pairwise Nash equilibria and alternative solution concepts. In particular, Bloch and Jackson (2007) observe that the set of pairwise Nash equilibria is the intersection of Nash equilibria of Myerson's game and pairwise stable networks. This intersection may very well be empty even when pairwise stable networks exist. Calvo-Armengol and Ilkilic (2009) characterize the class of network values for which pairwise Nash equilibrium networks and pairwise stable networks coincide.

As an alternative way to select among equilibria involving coordination failures, one may choose to consider only equilibria in *undominated* strategies, as in Dutta, Tijs and van den Nouweland (1998). Selten's trembling-hand perfection may also prove useful, as well as Myerson (1978)'s concept of proper equilibrium. Calvo-Armengol and Ilkilic (2009) focus on proper equilibria, and provide a (complex) condition on network value functions for which pairwise Nash equilibria and proper equilibrium networks coincide. In a different vein, Gilles and Sarangi (2006) propose a refinement based on evolutionary stability (termed 'monadic stability') to select among the equilibria of the linking game. Feri (2007) also applies evolutionary stability arguments to Bala and Goyal's models, and characterizes the set of stochastically stable networks.

Bala and Goyal (2000) follow a different approach. Faced with the multiplicity of equilibria in their one-way and two-way flow models, they propose to concentrate on *strict* Nash equilibria, where every player plays a strict best response to the actions of the other players. Eliminating strategy profiles where players are indifferent results in a drastic reduction of the number of equilibrium networks. Hojman and Szeidl (2008) show that the set of equilibrium networks can also be drastically reduced (to periphery-sponsored stars) when the value function of the network satisfies two conditions (i) strong decreasing returns to scale in the number of links and (ii) decay with network distance.

Of course, the strongest equilibrium notion allows for deviations by any group of players, and so corresponds to n-equilibrium. For normal form games, this is the notion of strong Nash equilibrium. Dutta and Mutuswami (1997) and Jackson and van den Nouweland (2005) study the strong equilibria of Myerson's network formation game.[14] Jackson and van den Nouweland (2005) characterize the set of network value functions for which strong equilibria of the Myerson game exist as follows.

Definition 4 *A network value function v, from the set of all graphs G to $\Re$ is* top-convex *if and only if* $\max_{g \in G^S} \frac{v(g)}{|S|} \leq \max_{g \in G} v(g)/n$.

[14] Jackson and van den Nouweland allow for deviations where some players are indifferent, and so their concept of equilibrium is stronger than the version defined here.

Theorem 1 *The set of strong equilibria in Myerson's game is nonempty if and only if the network value function v is top-convex.*

Jackson and van den Nouweland's characterization theorem shows that strong equilibria only exist when the per capita value of the grand coalition exceeds the per-capita value of any smaller coalition. This very strong convexity property is also the property guaranteeing nonemptiness of the core for symmetric TU games, and as we will see below, also plays a role in Chatterjee et al. (1993)'s study of coalitional bargaining games.

In the context of coalition formation, Hart and Kurz (1983) focus attention on the strong equilibria of the γ and δ games of coalition formation. [15] This concept of strong equilibrium is of course closely related to the familiar concept of the *core* of a characteristic function game.

Definition 5 *An allocation x belongs to the core of the game (N, v) iff x is feasible and*

$$\sum_{i \in S} x_i \geq v(S) \text{ for all } S \subseteq N$$

Of course, not all games have a non-empty core. Bondareva (1963) and Shapley (1967) characterized the class of games which have non-empty cores. Denote by $1_S \in R^n$ the vector such that

$$(1_S)_i = 1 \text{ if } i \in S, (1_S)_i = 0 \text{ if } i \notin S$$

A collection (λ_S) is a *balanced collection of weights* if $\sum \lambda_S 1_S = 1_N$. A game (N, v) is balanced if for all balanced collection of weights (λ_S), $\sum_S \lambda_S v(S) \leq v(N)$.

The classic result of Bonderava-Shapley is the following.

Theorem 2 *A game (N, v) has a non-empty core if and only if it is balanced.*

Notice that since Definition **??** is given in terms of the effectivity relation, the stability of any given social state will depend upon the group formation procedure. Consider, for instance, the γ and δ models of coalition formation applied to the next example.

Example 1 $N = \{1, 2, 3\}$. *Players are symmetric and receive values given by the following partition function $v(123) = (1, 1, 1)$, $v(1|2|3) = (0, 0, 0)$, $v(12|3) = (-1, -1, 2)$.* [16]

---

[15] In a companion paper, Hart and Kurz (1984) provide an example to show that the set of strong equilibria of the procedure of coalition formation may be empty.

[16] We assume that each player in a coalition gets an equal payoff, so that individual values can easily be derived from the partition function.

In game γ, the grand coalition N (giving a payoff of 1 to every player) is formed at a Nash equilibrium. If any player deviates from the announcement N, the coalition structure would collapse into a collection of singletons, resulting in a payoff of 0. By contrast, the grand coalition is not formed at any Nash equilibrium of the game δ. If a player i deviates from the announcement N, the other two players would still form the smaller two-player coalition, and the deviator would receive a payoff of 2 greater than the payoff she received in the grand coalition.

In general, the easier it is for coalitional deviations, the smaller is the set of equilibria. That is, suppose $\{\rightarrow_S^1\}_{S \subseteq N}$ and $\{\rightarrow_S^2\}_{S \subseteq N}$ are two families of effectivity relations with $x \rightarrow_S^1 y$ implying $x \rightarrow_S^2 y$ for all S and all x, y in X. Then, any k-equilibrium corresponding to $\rightarrow_S^2$ must be a k-equilibrium of $x \rightarrow_S^1 y$.

2.4 The degree of consent in group and network formation

The formation of groups and networks is an act which typically involves more than one agent, and may produce externalities on other agents. An important aspect of the procedure of group and network formation is thus the *degree of consent* it requires both from players directly and indirectly affected by the moves. In a model without transfers, this is of paramount importance, as players cannot easily be compensated for the decision taken by other players ; in models with transfers, the issue is somewhat mitigated by the fact that players can propose transfers (e.g. exit and entry prices) in order to internalize the externalities due to the moves of other players. In actuality, the formation of a group may require very little or very strong consent. In international law, agreements are typically open to the signature of all countries without restriction, so that no consent is needed either to enter or to exit the coalition. By contrast, transfers of professional soccer players across European teams require the consent (and the payment of a compensating transfer) both from the team that the player leaves and from the team that the player enters. In the formation of jurisdictions, as discussed in Jehiel and Scotchmer (2001), different constitutional rules on mobility result in very different coalition structures. The following table summarizes the assumptions on the degree of consent in models of group and network formation.

	No consent	Consent to enter	Consent to enter, exit
Coalitions	*Open membership* d'Aspremont (1983)	*Games γ and δ* Hart and Kurz (1984) *Bidding games* Kamien and Zang (1990) Perez-Castrillo (1994) *Individually stable equilibrium* Dréze and Greenberg (1980)	*Individually stable contractual equilibrium* Dréze and Greenberg (1980)
Networks	*Directed networks* Bala and Goyal (2000)	*Linking game* Myerson (1991) *Pairwise stable networks* JW	

For cooperative games without externalities, the introduction of additional constraints on the moves of players (requiring consent to enter and consent to enter and exit) makes deviations harder, and enlarges the set of equilibria. For example, Dréze and Greenberg (1980) note that individually stable contractual equilibria may exist in circumstances where individually stable equilibria fail to exist. When externalities are introduced, the picture becomes less clear, and different rules of consent may yield different predictions on the equilibrium outcomes. Yi (1997) studies this issue by comparing equilibrium outcomes of games with open membership and consent, focussing on the difference between games with *positive externalities* where the formation of a coalition benefits outside players, and games with *negative externalities* where the formation of a coalition harms outside players. The differences between these two types of games can easily be understood considering the following two examples.

Example 2 *A game with positive externalities. $N = \{1, 2, 3\}$. Players are symmetric and receive values given by the following partition function* $v(123) = (1, 1, 1), v(1|2|3) = (0, 0, 0), v(12|3) = (-1, -1, 2)$

As we saw above, the grand coalition is formed at an equilibrium of the γ game. However, if one considers an open membership game, players always want to leave any coalition, and the only equilibrium is one where all players remain as singletons.

Example 3 *A game with negative externalities. $N = \{1, 2, 3\}$. Players are symmetric and receive values given by the following partition function* $v(123) = (1, 1, 1), v(1|2|3) = (0, 0, 0), v(12|3) = (2, 2, -1)$

In this example, the only equilibrium of the γ and δ games results in players forming a coalition of size 2. However, in an open membership game, the third player will always want to join the coalition, and in equilibrium the grand coalition will form. Yi (1997)'s results generalize these two simple examples. He shows that in games with positive externalities, open membership will result in less concentrated coalition structures than games which require consent to enter ; in games with negative externalities, the result is reversed and open membership games yield larger coalitions than games with consent.

In networks, the absence of consent typically results in over-connections. If an agent does not require the consent of her partner to form a link, she might choose to form links which are beneficial to her, at the expense of her partner and all other agents in the network. This is illustrated by the following simple example.

Example 4 $N = 2$. *The payoffs in the graph are as follows:* $Y(\emptyset, v) = (0,0)$*,* $Y_1(\{12\}) = 1, Y_2(\{12\}) = -2$*.*

If consent is needed, the only equilibrium is the (efficient) empty network. If consent is not needed, player 1 can impose the formation of a link to player 2, resulting in the inefficient network $\{12\}$.

2.5 Some Examples

We describe some examples to illustrate how the concepts described earlier have been applied in specific contexts.

The Connections Model

This is due to Jackson and Wolinsky (1996). In this model, a link represents social relationships (e.g. friendship). These offer benefits (favors, information). In addition, individuals also benefit from indirect relationships. However, a "friend of a friend" generates a lower benefit than a friend. In other words, benefits decrease with the (geodesic) distance between any pair of nodes. Note that the benefit available at each node i has the "non-rivalry" characteristic of a pure public good - the benefit does not depend upon how many other nodes are connected to i.

Both i and j pay cost $c > 0$ for setting up link ij.

Hence, the net utility $u_i(g)$ to player i is

$$u_i(g) = \sum_{j \neq i, j \in P_i(g)} \delta^{d(i,j,g)} - c\#\{j | g_{ij} = 1\} \tag{1}$$

where $\delta < 1$, $d(i,j,g)$ is the geodesic distance between i and j, and $P_i(g)$ is the set of j who are path connected to i in g.

In the context of networks, it has become standard to define a network to be efficient if it maximizes the overall value of the network. Notice that this is a stronger notion of efficiency than the more familiar concept of Pareto efficiency.[17]

Definition 6 *Given v, a network g is efficient if $v(g) \geq v(g')$ for all g'.*

The simplicity of the model makes it easy to characterize both the sets of efficient and pairwise stable networks in terms of the two parameters c and δ. For instance, suppose c is smaller than $\delta - \delta^2$. Then, the cost of setting up an additional link ij is $2c$. Individuals i and j each get an additional benefit of at least $\delta - \delta^2$. So, the complete network must be both the unique efficient and pairwise stable network. The complete characterization is described below.

Efficiency: The efficient network is

(i) the complete network if $c < \delta - \delta^2$.

(ii) a star encompassing everyone if $\delta - \delta^2 < c < \delta + \frac{n-2}{2}\delta^2$

(iii) the empty network for $\delta + \frac{n-2}{2}\delta^2 < c$.

Pairwise Stability:

(i) If $c < \delta - \delta^2$, then the complete network is the unique pairwise stable network.

(ii) If $\delta - \delta^2 < c < \delta$, then a star encompassing everyone is one of several pairwise stable networks.

(iii) If $\delta < c < \delta + \frac{n-2}{2}\delta^2$, then all pairwise stable networks are inefficient.

Notice that the last case illustrates the fact that there may be situations in which the efficient network is not pairwise stable.

Bala and Goyal (2000) consider a version of the connections model where each agent i can set up a directed link with j without the consent of j and agents derive utility from directed links (one-way flow) or from undirected links (two-way flow). They make the simplifying assumption that the value of information that i gets from j does not depend upon the distance between i and j, that is there is no decay. In both versions of the model, there are a multiplicity of Nash equilibria. However, Bala and Goyal show that strict Nash equilibrium has a lot of predictive power. In particular, in the one-way flow model, a strict Nash equilibrium is either a wheel or the empty network. In the case of the two-way flow model, a strict Nash equilibrium is either the center-sponsored star or the empty network.

Collaboration amongst Oligopolistic Firms

There is considerable evidence that competing firms in the same industry collaborate with each other in a variety of ways-forming research

[17] See Jackson (2003) for alternative definitions of efficiency for networks.

joint ventures, sharing technology, conducting joint R & D, etc. Goyal and Joshi (2003) analyze research collaboration amongst firms using a two-stage model. In the first stage, each firm simultaneously announces the set of firms with which it wants to set up links. As in the typical two-sided model of link formation, a link forms between firms i and j if each firm has declared that it wants to form a link with the other. A link between firms i and j reduces the cost of production of both firms. Firms i and j also incur a cost $\gamma > 0$ in setting up a link. In the second stage, firms compete in the product market. Assume that firms compete in *quantities*, i.e. they are Cournot oligopolists, although price competition is also easy to analyze.

In its simplest version, the model specifies that n ex-ante identical firms face a linear market demand curve

$$p = a - Q$$

where p denotes the market price and $Q = \sum_{i=1}^{n} q_i$ is the industry output when firms choose the output vector $(q_1, \ldots, q_n)$. Firms have zero fixed cost of production and and initial identical marginal cost c_0. Let $g_{ij} = 1$ if firms i and j set up a collaboration link, and $g_{ij} = 0$ otherwise. Each collaboration link reduces the marginal cost by λ.

Firm i's marginal cost is then

$$c_i(g) = c_0 - \lambda \sum_{j \neq i} g_{ij} \tag{2}$$

The gross profit of a firm is its Cournot profit in the second stage, given a particular network structure, while its net profit is gross profit minus the cost of forming links.

Given any network g, if firms i and j are not linked in g, then by forming the link ij, both firms reduce their marginal cost. This must increase their level of gross profits. Now, suppose link cost, γ, is so low that the change in net profit is always positive for firms i and j whenever the firms set up the additional link ij. Then, the network structure satisfies the general property of *Link Monotonicity*.

Definition 7 *The pair (Y, v) satisfies Link Monotonicity if for all g, for all $ij \notin g$, $Y_i(g + ij, v) > Y_i(g, v)$ and $Y_j(g + ij, v) > Y_j(g + ij, v)$.*

It is obvious that if the network structure satisfies Link Monotonicity, then the complete graph g^N must be the only pairwise stable network. Suppose that we define a collaboration network to be efficient if it maximizes industry profits - that is, we ignore consumer surplus. Then, it is possible to find parameter values such that the complete graph is not the efficient structure.

Risk-Sharing Networks

This is due to Bramoulle and Kranton (2007) and Bloch, Genicot and Ray (2008). It models informal risk sharing across communities. Suppose there are two villages, and the sets of individuals living in villages 1 and 2 are V_1, V_2 respectively. Individual income is a random variable. For agent i living in village v, income is

$$\tilde{y}_i = \bar{y} + \tilde{\epsilon}_i + \tilde{\mu}_v$$

where $\tilde{\epsilon}_i$ is an idiosyncratic shock and $\tilde{\mu}_v$ is a village-level shock. Assume that village shocks are i.i.d. with mean zero and variance σ_μ^2, while idiosyncratic shocks are i.i.d. with mean zero and variance σ_ϵ^2. The village and idiosyncratic shocks are also independently distributed. Individuals have the same preferences with an increasing and strictly concave utility function $u(y)$, so that individuals are risk-averse.

Formal insurance is not available, but pairs of "linked" agents can smooth incomes by transferring money after the realization of shocks. A link between individuals in the same village costs c, while a link between individuals across villages costs $C > c > 0$. Of course, links have to be established ex ante, that is before the realization of the shock. Several interesting questions arise. When will one observe only within-village networks? When will agents also insure against village shocks? Will the latter type of network improve welfare?

A Model of Political Parties

This is due to Levy (2004). She assumes that political parties are composed of factions- groups who differ in their ideological positions. Parties form in order to facilitate commitment policies which represent a compromise between the preferred policies of individual politicians comprising the party.

Assume that a continuum of voters is composed of N finite groups of equal measure. Each group has different preferences over the policy space $Q \subset \mathbb{R}^k$. Voters who belong to group i share single-peaked preferences, represented by a strictly concave utility function $u(q, i)$. The "game" has N politicians, politician i having the preference of group i. Suppose the N politicians are arranged according to some coalition structure π. Interpret each S as a party. For each $S \in \pi$, let Q_S denote the set of Pareto-optimal points for coalition S. Then, each party simultaneously announces a policy platform $q_S \in Q_S \cup \{\emptyset\}$. Let q represent the vector of policy plaforms announced by the different parties. Voters now vote sincerely and the platform with the highest vote wins, ties being broken randomly. Each politician's utility from the game is his expected utility of the electoral outcome. The Nash equilibrium of this game then generates a partition function game.

The Exchange Economy

Each of n individuals have an endowment of L goods. Let $w_i \in \mathbb{R}^L_+$ denote the endowment of individual i. Each individual i has a utility function u_i defined over $\mathbb{R}^L_+$. Each coalition of individuals can trade with each other. This defines a non-transferable utility characteristic function game - notice that there are no externalities across coalitions.

3 Sequential Models of Coalition and Network Formation

3.1 Coalitional bargaining

In this section, we survey sequential models of coalition formation, which are based on Rubinstein (1982)'s model of alternative offers bargaining. As in Rubinstein (1982)'s model, the representative model has an infinite horizon, players discount future payoffs, and at each period in time, one of the players (the proposer) makes an offer to other players (the respondents) who must approve or reject the proposal. Different variants of this scenario have been proposed, each reflecting different assumptions on (i) the type of admissible offers, (ii) the selection of the proposer and (iii) the order of responses.

Coalitional bargaining games extend the two-person bargaining games, by considering general gains for cooperation, which can either be described by a coalitional game with transferable utility, or a partition function game.[18] Chatterjee, Dutta, Ray and Sengupta (1993) propose a model of coalitional bargaining based on an arbitrary game in coalitional form. Players are ordered according to an exogenous protocol. At the initial stage, player 1 chooses a coalition S to which she belongs and a vector of payoffs for all members of $S, \mathbf{x}_S$ satisfying $\sum_{i \in S} x_i = v(S)$. Players in S then respond sequentially to the offer. If all accept the offer, the coalition S is formed, and the payoff vector $\mathbf{x}_S$ is implemented. The first player in $N \backslash S$ is chosen as proposer with no lapse of time. If one of the players in S rejects the offer, one period elapses and the rejector becomes the proposer at the following period.

Chatterjee et al. (1993) look for conditions on the underlying characteristic function v for which efficient equilibria exist. Efficient equilibria must possess two features: (i) agreement must be reached immediately, so that there is no efficiency loss due to delay and (ii) the grand coalition should be formed in equilibrium. Let $m_i(S)$ denote the continuation value of player i

[18] Early extensions of Rubinstein (1982)'s bargaining game to three players were studied by Herrero, Shaked and Sutton – as reported in Sutton (1986) and Binmore (1985).

when she makes an offer and the set of active players is S. The following example shows that for some protocols, all equilibria result in delay.

Example 5 *N=4, $v(\{1,j\})$=50 for j=2,3,4, $v(\{i,j\})$=100, i,j=2,3,4 and $v(S)$=0 for all other coalitions.*

Suppose by contradiction that all equilibria exhibit immediate agreement. Then, all players make acceptable offers, and in particular,

$$\begin{aligned} m_i(N) &= \delta 100/(1+\delta) \text{ for } i = 2,3,4 \\ m_1(N) &= \delta[50 - 100\delta/(1+\delta)] \end{aligned}$$

Clearly, when δ converges to $1, m_1(N)$ converges to 0. Consider then the following deviation for player 1. Player 1 makes an unacceptable offer when the set of players is N, and waits for two players to form a coalition before making an acceptable offer. In that situation, the first coalition will be formed by two of the players $2,3$ and 4 (who will roughly obtain 50 each). Once these two players have left, player 1 and the remaining player will equally share the surplus of 50 and obtain 25 each. Hence, this deviation is profitable for player 1, who has an incentive to make an unacceptable offer (thereby inducing delay) at the beginning of the game.

A careful look at the preceding example shows that delay occurs in equilibrium because one of the players (player 1) is better off waiting for some players to leave before entering negotiations. This suggests that the following condition will be sufficient to rule out delay in equilibrium.

Condition 1 *For all coalitions S and T with $T \subset S$ and all discount factors, $m_i(S) \geq m_i(T)$ for all players i in T.*

Turning now to the second source of inefficiency, the following example shows that even if agreement is reached immediately, the grand coalition may fail to form in equilibrium.

Example 6 $N = 3, v(S) = 0$ *if* $|S| = 1, v(S) = 3$ *if* $|S| = 2, v(N) = 4$.

Suppose by contradiction that the grand coalition forms. Then, $m_i(N) = 4\delta/(2+\delta)$ which converges to $4/3$ as δ converges to 1. But then, any player has an incentive to propose to form a two player coalition, resulting in an expected payoff of $3\delta/(1+\delta)$ which converges to $3/2$ as δ converges to 1.

The preceding example shows that, as long as intermediate coalitions produce a large surplus, players have an incentive to form smaller, inefficient coalitions. A careful look at the example shows that the two-player coalition forms because this is the coalition which maximizes *per capita* payoff. In particular, if the grand coalition were to maximize the per capita payoff of all the players,(condition of top convexity) then it would clearly form in equilibrium. The next Proposition shows that top convexity is a necessary and sufficient condition for the grand coalition to form in all stationary subgame perfect equilibria and for all protocols.

Theorem 3 *The following two statements are equivalent:*

(a) The game v satisfies top convexity

(b) For every protocol, there exists a sequence of discount factors converging to 1 *and a corresponding sequence of efficient stationary subgame perfect equilibria.*

Okada (1996) analyzes a coalitional bargaining game where the proposer is selected at random after every rejection. In that case, no player will strategically make an unacceptable offer, in order to pass the initiative to another player. Hence agreement will be reached at the beginning of the bargaining game. This agreement however may not lead to the formation of the grand coalition – the first proposer may find it optimal to form a smaller coalition. In order to guarantee that the grand coalition forms immediately, the same condition of top convexity identified by Chatterjee et al. (1993), is in fact necessary and sufficient.

If underlying gains from cooperation are represented by a game in partition function form (allowing for externalities across coalitions), players forming a coalition must anticipate which coalitions will be formed by subsequent players. Bloch (1996) proposes a coalitional bargaining game capturing this forward-looking behavior when the division of the surplus across coalition members is fixed.[19] Bloch (1986)'s main result deals with symmetric games where payoffs only depend on the size distribution of coalitions. In that case, the equilibrium coalition structures of the infinite horizon bargaining game can be computed by using the following finite procedure. Let players be ordered exogenously. The first player announces an integer k_1, corresponding to the size of the coalition she wants to form. Player $k_1 + 1$ then announces the size k_2 of the second coalition formed. The game ends when all players have formed coalitions, i.e. $\sum k_t = n$.

[19] In this game, as in the seminal studies of Selten (1981) and Moldovanu and Winter (1995), there is no discounting but all players receive a zero payoff in the case of infinite play.

While Bloch (1996) assumes that the division rule of the surplus is fixed, Ray and Vohra (1999) consider a model of coalitional bargaining with externalities, where the division of coalitional surplus is endogenous, and payoffs are represented by an underlying game in partition function form. Ray and Vohra (1999) first establish the existence of stationary equilibria in mixed strategies, where the only source of mixing is the probabilistic choice of a coalition by each proposer. Their main theorem establishes an equivalence between equilibrium outcomes of the game and the result of a recursive algorithm. This algorithm, in four steps, characterizes equilibrium coalition structures for symmetric games. It can easily be implemented on computers and has been successfully applied in Ray and Vohra (2001) to study the provision of pure public goods. For any vector $\mathbf{n}$ of positive integers, $\mathbf{n} = (n_i)$, let $K(\mathbf{n}) = \sum n_i$.We construct a mapping $t(\mathbf{n})$ for all vectors $\mathbf{n}$ such that $K(\mathbf{n}) \leq n$. This mapping associates a positive integer to any vector $\mathbf{n}$. Applying this mapping repeatedly, starting at the empty set, when no coalition has formed, we obtain a coalition structure, $c(\varnothing) = \mathbf{n}^*$ that will be the outcome of the algorithm.

Step 1: For all $\mathbf{n}$ such that $K(\mathbf{n}) = n - 1$, define $t(\mathbf{n}) = 1$.

Step 2. Recursively, suppose that $t(\mathbf{n})$ has been defined for all $\mathbf{n}$ such that $K(\mathbf{n}) \geq m$ for some m. Suppose moreover that $K(\mathbf{n}) + t(\mathbf{n}) \leq n$. Then define

$$c(\mathbf{n}) = (\mathbf{n}, t(\mathbf{n}), t(\mathbf{n}, t(\mathbf{n})), ...)$$

which is a list of integers, corresponding to the repeated application of the mapping t starting from the initial state $\mathbf{n}$.

Step 3 For any $\mathbf{n}$ such that $K(\mathbf{n}) = m$, define $t(\mathbf{n})$ to be the largest integer in $\{1, ..., n - m\}$ that maximizes the expression

$$\frac{v(t, c(\mathbf{n}, t))}{t}.$$

Step 4 Since the mapping t is now defined recursively for all vectors $\mathbf{n}$, start with the initial state where no coalition has formed, and compute $\mathbf{n}^* = c(\varnothing)$.

Ray and Vohra (1999) then show that, in symmetric games where payoffs are increasing in the order in which coalitions are formed, the preceding algorithm fully characterizes equilibrium coalition structures when the discount factor converges to 1.[20]

Political science is an important area of application of models of coalitional bargaining. Political agents have to build majority coalitions in order

[20]Montero (1999) considers a version of the Ray and Vohra (1999) model where the proposer is chosen at random, as in Okada (1996).

to secure the passing of legislation or the implementation of policies. Majority building occurs in many different political processes: in the formation of coalitional governments in parliamentary democracies, in the passing of legislation in parliament, or in the choice of policies in supranational bodies, like the United Nations Security Council or the European Council. Political scientists have developed a specific analysis of coalition formation, encompassing both theoretical models and empirical estimations. The analysis of coalition formation in political science exhibits two distinctive features: (i) the exact process by which coalitions are formed (the "rules of the game") are often well specified, either through custom or through constitutional provisions, and (ii) the coalitional game is a simple game, where coalitions are either winning or losing.

Baron and Ferejohn (1989) consider an extension of Rubinstein (1982)'s alternating offers model, where members of a legislature bargain over the division of a pie of fixed value (interpreted as the distribution of benefits to different constituencies). In order to be accepted, a proposal must receive the approval of a simple majority of members of the legislature. In the *closed rule* model they consider, a proposer is chosen at random, and his offer is immediately voted upon by the legislature.[21] Baron and Ferejohn (1989)'s first result is an indeterminacy result, showing that any distribution of payoffs can be reached in a subgame perfect equilibrium of the closed-rule game.

Theorem 4 *For an n-member majority rule legislature with a closed rule, if the discount factor satisfies $1 > \delta > \frac{n+2}{2(n-1)}$ and $n \geq 5$, any distribution x of the benefits may be supported as a subgame-perfect equilibrium.*

The intuition underlying this indeterminacy result is easy to grasp. In Baron and Ferejohn (1989)'s game, a deviating player can always be punished (independently of the discount factor) by other players systematically excluding him from any coalition. These equilibrium punishment strategies can be used to deter any deviation from an arbitrary distribution of benefits x. In order to select among equilibria, Baron and Ferejohn (1989) impose a further restriction on equilibrium strategies, by assuming that strategies are stationary – namely cannot depend on the entire history of play but only on the current offer. With stationary strategies, members of the legislature cannot exclude other members in response to a deviation, and the equilibrium distribution of benefits becomes unique, as shown in the following Proposition.

Theorem 5 *For all $\delta \in [0,1]$, a configuration of pure strategies is a stationary subgame perfect equilibrium in a closed rule game (with an odd number*

[21] The second model they consider – the *open rule model* – is more complex and allows for amendments to the status quo.

of legislators) if and only if a member recognized proposes to receive $1-\delta\frac{n-1}{2n}$ *and offers* $\frac{\delta}{n}$ *to* $\frac{n-1}{2}$ *members selected at random, and each member votes for any proposal in which at least* $\frac{\delta}{n}$ *is received.*

By contrast to Rubinstein (1982)'s game, where the shares of the proposer and respondent converge to $\frac{1}{2}$ when δ converges to 1, the proposer in Baron and Ferejohn (1989)'s game retains a large advantage over the respondents, even when all players become perfectly patient. This is due to the fact that a respondent is not sure to be included in the next majority coalition if she rejects the offer, so that the minimal offer she is willing to accept may be quite low. In an application to the formation of coalitional governments in parliamentary systems, Baron and Ferejohn (1989) note that, even if the probability of recognition is proportional to the number of seats in parliament, the proposer's advantage remains very large, and results in the first proposer (even if it is a small party) obtaining a disproportionate number of cabinet posts. In the open rule procedure, the proposer's advantage is mitigated by the fact that a second member of parliament can propose an amendment. This reduces the power of the first proposer, and results in a larger share of the benefits for the respondents. Furthermore, for low values of the discount factor, the first proposer will choose in equilibrium to make offers to a supermajority of members, in order to reduce the probability that a second member of parliament propose an amendment to his offer.[22]

An important extension of Baron and Ferejohn (1989)'s model was proposed by Merlo and Wilson (1995) to take into account uncertainty and random shocks on the surplus from cooperation. In their model of bargaining in a stochastic environment, players share a cake whose size varies from period to period according to a general Markov process. They consider a bargaining problem, where the agreement has to be unanimously accepted by all players to be effective. This model is well suited to analyze the formation and collapse of coalitional governments which face an uncertain, stochastic environment.

While the coalitional bargaining models of Chatterjee et al. (1993) and Baron and Ferejohn (1989) are straightforward extensions of Rubinstein (1982)'s bargaining model, other more complex extensive form procedures have also been studied, often with the objective of providing a noncooperative foundation to a cooperative solution concept. Most papers in this "Nash program" vein aim at supporting solution concepts like the core or the

[22]The coalitional bargaining model of Baron and Ferejohn (1989) has generated a considerable theoretical and empirical literature in political science. See Harrington (1989), Baron and Kalai (1993), Winter (1996), Merlo (1997), Banks and Duggan (2000), Jackson and Moselle (2002), Norman (2002), Eraslan (2002) and Seidmann and Winter (2007) for theoretical contributions and Merlo (1997), Diermeier and Merlo (2004) and Diermeier, Eraslan and Merlo (2003) for empirical tests.

Shapley value, where the grand coalition is *assumed to form*. Hence, these procedures are usually not well suited to analyze the formation of partial coalitions.[23] However, recent work on the Shapley value with externalities discusses procedures which can lead to the formation of partial coalitions. For example, Maskin (2003) proposes a sequential procedure where players enter the game according to an exogenous rule of order, and existing coalitions simultaneously bid for the entering player. Macho-Stadler, Perez-Castrillo and Wettstein (2007) propose a different bidding procedure which implements another type of Shapley value with externalities. De Clippel and Serrano (2008) and Dutta, Ehlers and Kar (2008) propose other extensions of the Shapley value to games with externalities but do not discuss non-cooperative implementation.

3.2 Sequential models of network formation

Sequential models of network formation have been proposed in order to circumvent two difficulties in models of network formation. First, in a sequential procedure, agents do not behave myopically and choose their actions anticipating the reaction of subsequent players. In the words of Aumann and Myerson (1988):

> 'When a player considers forming a link with another one, he does not simply ask himself whether he may expect to be better off with this link than without it, given the previously existing structure. rather, he looks ahead and asks himself: "Suppose we form this new link, will other players be motivated to form further links that were not worthwhile for them before? Where will it all lead? Is the *end result* good or bad for me?"

(Aumann and Myerson (1988, p. 178)).

Second, as a finite sequential game of complete information generically possesses a unique subgame perfect equilibrium, the use of sequential procedures helps to refine the set of Nash equilibria and to resolve the coordination issues involved in link formation. Both the modeling of players as forward-looking agents, and the resolution of coordination problems due to the sequentiality of decisions indicate that sequential models are more likely to produce *efficient networks* than simultaneous procedures. Unfortunately, as shown by Currarini and Morelli (2000) and Mutuswami and Winter (2002)

[23] Some representative papers include Gul (1989) , Hart and Mas Colell (1996), and Perez-Castrillo and Wettstein (2000) for noncooperative implementation of the Shapley value, Perry and Reny (1994), Lagunoff (1988) and Serrano and Vohra (1997) for the core, or Binmore, Rubinstein and Wolinsky (1986) and Krishna and Serrano (1996) for the bargaining solution.

even sequential procedures may not produce efficient networks except under restrictive conditions on the underlying structure of gains from cooperation.

In the first attempt to study sequential procedures of network formation, Aumann and Myerson (1988) consider a finite game where, at any point in time, a pair of players who are not yet linked is called to form a new link by mutual agreement. Links are never destroyed. The game is finite, but the rule of order must be such that, after the last link is formed, all pairs of players who are not linked have a last opportunity to form a new link. Aumann and Myerson (1988)'s primary objective is to emphasize the role of anticipation and forward-looking behavior on the formation of networks. Consider the following example:

Example 7 $N = 3$. *If* $g = \emptyset$, $Y_i(g) = 0$. *If* $g = \{ij\}$, $Y_i(g) = Y_j(g) = 30, Y_k(g) = 0$. *If* $g = \{ij, ik\}$, $Y_i(g) = 44, Y_j(g) = Y_k(g) = 14$. *If* $g = \{ij, ik, jk\}$, $Y_i(g) = Y_j(g) = Y_k(g) = 24$.

In this Example, the complete network is efficient, but, for any rule of order, the subgame perfect equilibrium of Aumann and Myerson (1988)'s game is for two players to form a single link. After this link is formed, if one of the players tries to form another link in order to obtain the payoff of 44, this will be followed by the formation of the last link, resulting in a payoff of 24. Hence, forward-looking players will never choose to form an additional link after the first link is formed. The same line of reasoning can be applied to characterize the subgame perfect equilibrium in an "apex game" where one large player faces four small players.

Example 8 $N = 5$. *Players get payoffs which only depend on the components of the network, and not the way players are linked. There are two types of player: one large player (player* 1*) and four small players,* $(2, 3, 4, 5)$*. Winning coalitions either include the large player, or consist of all four small players. Payoffs are based on the Shapley value of this apex game. If coalition* $\{1, i\}$ *forms, player* 1 *and* i *get* $\frac{1}{2}$*. If coalition* $\{1, i, j\}$ *forms, player* 1 *gets* $\frac{2}{3}$ *and players* i *and* j *get* $\frac{1}{6}$ *each. If coalition* $\{1, i, j, k\}$ *forms, player* 1 *gets* $\frac{3}{4}$ *and players* i, j, k *get* $\frac{1}{12}$ *each. If the four small players form a coalition, they receive* $\frac{1}{4}$ *each. If the grand coalition forms, the large player receives* $\frac{3}{5}$ *and the four small players get* $\frac{1}{10}$ *each.*

In this Example, the unique equilibrium structure is for the four small players to form a coalition. By backward induction, we observe that, if coalition $\{1, i, j, k\}$ forms, small players have an incentive to form a link to the excluded small player, and the grand coalition results. Hence, coalition $\{1, i, j\}$ is stable, because the large player knows that, if she invites another small player to join, the end result will be the grand coalition, with a payoff of $\frac{3}{5}, \frac{2}{3}$. On the other hand, if coalition $\{1, i\}$ forms, the large player has an incentive to bring in another small player, resulting in the stable coalition $\{1, i, j\}$. At the beginning of the game, small players realize that they can either obtain a payoff of $\frac{1}{4}$ (if they form a coalition of small players) or $\frac{1}{6}$ (if they join the large player in a three-player coalition), and prefer to form the coalition of small players.

Aumann and Myerson (1988)'s model makes strong assumptions on the rule of order to ensure that the game is finite. Attempts to construct general, infinite horizon models of network formation based on the same structure as models of coalitional bargaining have so far remained elusive. One exception is Watts (2002)'s construction of a subgame perfect equilibrium in the connections model where agents' utilities can be decomposed as the sum of benefits from communication (discounted by the distance in the network) and costs of direct links. In her model, at each point in time, when a pair of players is selected, it can either choose to form a new link or to destroy the existing link, resulting in a game with infinite horizon – and a large number of equilibria. When the discount factor goes to 1, Watts (2002) exhibits one subgame perfect equilibrium where players form the circle network. To sustain this equilibrium, players employ a grim strategy, where players who fail to cooperate are punished by being ostracized. However, this is only one equilibrium among many, and Watts (2002) does not propose a full characterization of the set of subgame perfect equilibrium outcomes.

In two related contributions, Currarini and Morelli (2000) and Mutuswami and Winter (2002) propose finite procedures to study the relation between efficiency and equilibrium. In both procedures, agents are ordered according to an exogenous rule, and make announcements in sequence. In both models, one needs to impose a monotonicity condition on the value of the network to guarantee that any subgame perfect equilibrium is efficient.

Currarini and Morelli (2000) suppose that the total value of the network is given by a mapping $v : G \rightarrow \Re$, associating a real number $v(g)$ to any graph g. They suppose that the value is monotonic in the following sense:

Definition 8 *A link ij in graph g is* critical *if and only if the number of components of $g - ij$ is strictly greater than the number of components of g. The value function v satisfies* size monotonicity *if and only if for all graphs g and critical links ij, $v(g) > v(g - ij)$.*

They consider a finite procedure where each player makes a single move. At stage i, player i announces a pair (g_i, d_i) where g_i is a set of *links* to agents in $N \setminus i$ and d_i is a real number, expressing the *demand* of agent i. Given these announcements, one constructs a network g by letting link ij be formed if and only if both parties agree to the formation of that link. For any component h of g, one verifies whether the value of the component

$v(g(h))$ can cover the demands of the agents in h. If the answer is positive, network $g(h)$ is formed, and every member of h receives her demand d_i. If the answer is negative, then all members of h are isolated and receive a payoff of zero. In this model, they prove that all subgame perfect equilibria are efficient:

Theorem 6 *Let v satisfy size monotonicity. Then any subgame perfect equilibrium network of Currarini and Morelli (2000)'s sequential game is efficient.*

Mutuswami and Winter (2002) consider instead a model where players have private values over the network, $v_i(g)$, face a known cost function $c(g)$, and can transfer utility only insofar as they share the cost of the network. In their mechanism, at stage i, player i announces a pair (g_i, x_i) where g_i is a set of links that player i wants to see formed and x_i is a positive number, representing the *conditional cost contribution* of player i, that she commits to pay if the network formed is a superset of g_i. Given these announcements, the coalition S is said to be *compatible* if and only if: (i) $g_i \in G^S$ for all $i \in S$ and (ii) $\sum_{i\in S} x_i \geq c(\cup_{i\in S} g_i)$. The mechanism then selects the *largest* compatible coalition S^* among the connected coalitions $1, 12, 123, ..., 123...n$. The network formed is $g = \cup_{i\in S^*} g_i$ and every player in S^* receives a payoff $v_i(g) - x_i$ whereas players in $N \setminus S^*$ receive a payoff of zero. Mutuswami and Winter (2002) consider the following notion of monotonicity:

Definition 9 *The value function v_i is monotonic if and only if, whenever $g \subset g'$, $v_i(g) \leq v_i(g')$.*

They prove that any subgame perfect equilibrium network is efficient.

Theorem 7 *Suppose that $c(\emptyset) = 0$ and that $v_i(\cdot)$ is monotonic for every agent i. Then every subgame perfect equilibrium network in the Mutuswami and Winter (2002) sequential mechanism is efficient. Furthermore, in equilibrium, every agent receives his marginal contribution:*

$$u_i = \max_{g\in G^{\{1,\ldots i\}}} \sum_{j=1}^{i} v_j(g) - c(g) - \max_{g\in G^{\{1,\ldots i-1\}}} \sum_{j=1}^{i-1} v_j(g) - c(g).$$

4 Farsightedness

In section 2.3, we described the notion of a k-equilibrium. Recall that implicit in the definition of a k-equilibrium is the idea that when a group or individual contemplates deviation from the proposal x "on the table" to another social state y, it simply compares the utilities it gets under y and x. But, consider the familiar *voting paradox*.

Example 9 *$N = \{1, 2, 3\}$, $X = \{x, y, z\}$, and $a \to_S b$ for all $a, b \in X$ iff $|S| \geq 2$. The preferences of the three individuals are described below.*

$$x \succ_1 y \succ_1 z, y \succ_2 z \succ_2 x, z \succ_3 x \succ_3 y$$

No social state is a 2-equilibrium. Individuals 1 and 3 prefer to move from y to x and have the power to do so, 2 and 3 want to move away from x to z, and the cycle is completed because 1 and 2 prefer to move from z to y. Given this cycle, why does 1 agree to join 3 in moving from y to z when z itself is not a "stable outcome"? That is, what is the relevance of the utility that she derives from z when there is no guarantee that the group will agree to adopt z as the final outcome.

The answer must be that the concept of k-equilibrium models players who are *myopic* - they do not look ahead to the possibility of further deviations once they themselves have moved away from the status quo. In this section, we discuss concepts of stability in one-stage models of group formation when players are *farsighted* in the sense that they take into account the "final" outcome(s) which can result from their initial deviation.[24] Of course, since players move only once, these notions of farsightedness must involve introspection.

Notice that equilibrium predictions in the bargaining models described in the last section already incorporate this kind of farsighted behavior. For instance, an initial proposer evaluates what will happen if he makes a proposal which will be rejected by some other player. Will the new proposal give him less than he is asking for now? Similarly, a player who contemplates rejecting the current proposal must also look far ahead and anticipate the proposal which will ultimately be approved by the players.

There are different options in modeling players' farsighted behavior when players choose actions simultaneously, depending at least partly on the social environment. First, suppose the social environment is such that there are no externalities across groups. Consider, for instance, games in characteristic function form or hedonic games without externalities, so that "subgames" are well-defined. The standard myopic stability notion then is the core. Clearly, one implication of farsighted behavior is that the act of blocking must be credible. In particular, if a coalition S contemplates breaking away from the grand coalition, then it must also take into account the possibility that a sub-coalition of itself may effect a further deviation.

It is possible to generalize this notion. For each S, let F_S be the set of feasible outcomes for S, and Θ_S be a solution concept. Then, the solution concept should satisfy the requirement that an allocation x is in Θ_S only

[24]In the next section, we will discuss farsightedness in the context of dynamic situations where groups interact over time.

if no sub-coalition T of S can block x with an allocation which is itself a solution for T. More formally,

Definition 10 *The solution Θ_S satisfies Internal Consistency if for all games (N, v) and for all $S \subseteq N$,*

$$\Theta_S(v) = \{x \in F_S(v) | \text{ there is no } T \subset S, y \in \Theta_T(v) s.t.\ y \succ_i x\ \forall i \in T\}$$

Greenberg (1990) and Ray (1989) prove the following.

Theorem 8 *The core satisfies Internal Consistency.*

The underlying intuition is quite simple. For suppose $x \notin C_S(v)$. Then, some $T \subset S$ blocks x with y. If y is not in core of T, then some subset of T, say R blocks y. But, notice that this implies

$$v(R) > \sum_{i \in R} y_i > \sum_{i \in R} x_i$$

So, R blocks x. And so on until some singleton coalition does the blocking.[25]

It is not straightforward to apply this definition of internal consistency to games with externalities since subgames on coalitions are not well-defined. Below, we describe other approaches to farsighted behaviour in general social environments.

Fix some social environment $\Gamma = (N, \{\rightarrow_S\}_{S \subseteq N}, \{\succeq\}_{i \in N})$. For all $x, y \in X$, define the binary relation $>$ as follows

$$x > y \text{ if } \exists S \subseteq N \text{ such that } y \rightarrow_S x \text{ and } x \succ_i y\ \forall i \in S$$

Now, farsighted behavior "could" mean the following:

(i) If x is not "stable", then some coalition should be able to deviate profitably to some y *and* y should itself be stable.

(ii) If x is stable, then no coalition can have a profitable deviation from x to another stable outcome y.

It is worth pointing out why (i) and (ii) incorporate farsightedness. Implicit in (i) is the idea that if a coalition S has the power to deviate from x to some stable outcome y, then the stability of y ensures that there will not

[25] Consistency requires that a coalition can only block with allocations from its own set of unblocked allocations. Suppose "norms" dictate that all unblocked allocations are not available, but only those which pass the norms test. Dutta and Ray (1989) define a recursive notion of *Norm consistency*, which imposes the requirement that each coalition can only block with those feasible allocations which pass the norms test from its one set of unblocked allocations. Their *egalitarian solution* uses the norm of selecting the Lorenz-maximal elements in each set.

be any further deviation from y. So, if members of S all prefer y to x, then they will go ahead with the deviation and so x cannot be stable. The same logic also suggests the criterion that any stable social state must satisfy - if x is stable, then no coalition should want to move to another stable state.

These considerations lead to the vNM *stable set* with respect to the binary relation $>$.

Definition 11 *The vNM stable set for any asymmetric relation $\succ$ over X is a set $V(\succ)$ satisfying:*

- *External Stability: If $y \notin V(\succ)$, then there is some $x \in V(\succ)$ such that $x \succ y$.*
- *Internal Stability: If $y \in V(\succ)$, then there is no $x \in V(\succ)$ such that $x \succ y$.*

Lucas (1969) showed that a stable set need not exist. For instance, it is empty in example **??** for the relation $>$. Even if stable sets do exist, they need not be unique. In fact, as the following example illustrates, there may be a continuum of vNM solutions even in very simple games.

Example 10 *Let $N = \{1, 2, 3\}$, $v(S) = 1$ if $|S| \geq 2$ and $v(S) = 0$ otherwise. Choose any $a \in [0, 1/2)$, and any $i \in N$. Then, the set $V_i(a) = \{x \in R^3_+ | x_i = a, x_j + x_k = 1 - a\}$ constitutes a vNM set for the relation $>$.*

It is easy to check that $V_i(a)$ satisfies internal consistency. To check external stability, take any $y = (y_1, y_2, y_3)$ such that $\sum_{i=1}^{3} y_i = 1$ and $y \notin V_i(a)$. If $y_i > a$, then $y_j + y_k < 1 - a$, and there must exist $x \in V_i(a)$ such that $x_j > y_j$ and $x_k > y_k$. If $y_i < a$, then without loss of generality assume that $y_j \geq y_k$. Then, there exists $x \in V_i(a)$ such that $x_i > y_i$ and $x_k > y_k$.

The non-uniqueness of vNM sets is possibly one reason why this solution concept has not been very popular in applications. Harsanyi (1974) also felt that the vnM set $V(>)$ does not really incorporate farsighted behavior. The example below illustrates the nature of his criticism.

Example 11 *Let $X = \{x, y, z, w\}, N = \{1, 2\}$, and individual preferences be*

(i) $x \succ_1 y \succ_1 z \succ_1 w$.
(ii) $w \succ_2 y \succ_2 x \succ_2 z$.
Finally, the effectivity relation is:

$$x \rightarrow_2 y, z \rightarrow_1 w, w \rightarrow_1 x, y \rightarrow_2 z, z \rightarrow_{\{1,2\}} y.$$

The vnM set is $\{y, w\}$. The underlying logic is the following. Player 1 does not deviate from w to x because x itself is not stable since 2 will deviate from x to y. But, why should this deter 1 from deviating since she prefers y to w? Thus, 1 does not look sufficiently far ahead.

Harsanyi's objection can be interpreted as a criticism not of the stable set, but of the direct domination relation $>$. Notice that the problem unearthed in example **??** arises partly because 1 does not anticipate the *chain* of deviations which might follow from her initial deviation from w. This suggests that the domination relation itself should be modified so as to consider a *chain* of deviations. The *indirect domination* relation, defined below, captures this aspect.

Definition 12 *A state y is indirectly dominated by x if there exist sequences of states $\{y_0, \ldots, y_K\}$ and coalitions $\{S_0, \ldots S_{K-1}\}$ such that $y_0 = y, y_K = x$, and for all $k = 0, \ldots K-1$,*

(i) $y_k \rightarrow_{S_k} y_{k+1}$.

(ii) $y_K \succ_i y_k$ for all $i \in S_k$.

Let $x >> y$ denote the relation that x indirectly dominates y.

The indirect domination relation incorporates the idea that coalitions look beyond their own immediate actions to the "ultimate" consequence-each coalition S_k compares y_k with the "end point" y_K. But how is the end point itself determined? Our previous discussion suggests that a natural candidate for the end point is that it should itself be stable so that there is no further deviation from it. This in turn suggests $(V >>)$ as a candidate for a farsighted consistent solution.[26]

In example **??**, $(V >>) = \{y\}$, and so $V(>>) \subset V(>)$. However, this is not true in general,[27] as the following example illustrates.

Example 12 *Let $X = \{x, y, z, w\}$, $N = \{1, 2, 3\}$. Preferences and the effectivity relation are as follows.*

(i) $y \succ_1 x \succ_1 z$.

(ii) $z \succ_2 w$.

(iii) $z \succ_3 y$.

(iii) $y \rightarrow_3 w, w \rightarrow_2 z, x \rightarrow_1 y$.

[26] Diamantoudi and Xue (2007) advance this as a solution concept in their analysis of hedonic games with externalities. See also Diamantoudi and Xue (2003) and Xue (1998).

[27] What is true however is that $V(>)$ can never be a subset of $V(>>)$. For suppose $V(>) \equiv V \subset V' \equiv V(>>)$. Let $y \in V' \setminus V$. Then, by External Stability of V with respect to $>$, there is $x \in V$ with $x > y$. But, $x > y$ implies $x >> y$, and so V' violates Internal Stability with respect to $>$.

Then, $y > x$, $z >> y$, $z > w$ and hence $z >> w$. So, $V(>) = \{y, z\}$ and $V(>>) = \{x, z\}$.

Chwe (1994) points out that $(V >>)$ may sometimes be too restrictive.[28] Consider the following example.

Example 13 *Let* $X = \{x, y, z, w\}, N = \{1, 2\}$ *and preferences and effectivity functions be as follows.*

(i) $w \succ_1 y \succ_1 x \succ_1 z$.

(ii) $z \succ_2 w \succ_2 y \succ_2 x$.

(iii) $x \rightarrow_1 y,\ y \rightarrow_2 z,\ y \rightarrow_{\{12\}} w$

Then, $y >> x, w >> x, z >> y$, and $(V >>) = \{w, z\}$. The logic ruling out x as a "stable" outcome is that 1 will move to y anticipating that $\{1, 2\}$ will deviate further to w which 1 prefers to x. But, is this anticipation reasonable? After all, once 1 deviates to y, individual 2 has the option of moving to either z or w. Surely, 2 will prefer the move to z which is worse than x as far as 1 is concerned. Thus, if 1 makes the right inference, then 1 should not plan to deviate from x.

Chwe (1994) advances a different solution concept based on the indirect domination relation.

Definition 13 *A set* $Y \subset X$ *is consistent if* $x \in Y$ *iff* $\forall y, S$ *s.t.* $x \rightarrow_S y$, *there is* $z \in Y$ *s.t. either* $z = y$ *or* $z >> y$ *and* $x \succ_i z$ *for some* $i \in S$.

Chwe (1994) shows that there is a *largest consistent set*, denoted LCS, and offers this as a solution concept. He proves the following.

Theorem 9 *Suppose* X *is finite. Then,*

(i) The LCS is nonempty and has the external stability property with respect to $>>$.

(ii) $V(>>) \subset LCS(>>)$.[29]

Suppose a coalition S has the power to move from x to y. When will it decide not to deviate to y? Being a farsighted solution concept, y is relevant if y itself is in the consistent set. If y is indeed in a consistent set, and some individual in S prefers x to y, then S will not move to y. Otherwise, S

[28] Of course, it can sometimes be empty as in the voting paradox.

[29] Since the set of all networks on a finite player set is finite, this existence theorem proves very useful to Kamat, Page and Wooders (2005) and Page and Wooders (2009) in their analysis of network formation. For a different concept of farsighted network formation based on pairwise stability, see Herings et al. (2009).

looks ahead to *some* chain of deviations from y to an element z in Y which indirectly dominates y, and compares z to x.

In example **??**, the LCS=$\{x, z, w\}$, because 1 will be deterred from moving to y apprehending the further move to z by 2. But, while ($V >>$) can be faulted for being too restrictive, the LCS errs in the opposite direction by being too permissive. That is, it labels states as being "stable" when there are good reasons for declaring them to be unstable. Consider a slight modification of example **??**.

Example 14 *Let* $X = \{x, y, z, w\}, N = \{1, 2\}$ *and preferences and effectivity functions be as follows.*

(i) $z \succ_1 x \succ_1 w$.

(ii) $z \succ_2 w \succ_2 y$.

(iii) $x \rightarrow_1 y$, $y \rightarrow_2 z$, $y \rightarrow_2 w$

Then, the LCS contains x. The argument justifying this is the following. Suppose 1 moves to y. Then, 2 could move further to w which is in the LCS. Since $x \succ_1 w$, this should deter the move from x. But, why should 2 move to w from y when she can also move to z which she prefers to both y *and* w? Notice that if 2 actually moves to z, then this justifies 1's deviation from x to y. Hence, in this example, x should not be deemed to be stable. Note that Chwe himself is aware of this problem. He suggests the following interpretation of the LCS - if a state does not belong to the LCS, then it cannot possibly be stable.

These examples suggest that none of these attempts to define a farsighted solution concept is completely satisfactory.

The problem with solution concepts such as $V(>>)$ and LCS is that their arguments for inclusion or exclusion vis-a-vis stability are based on *some* sequence of deviations. For instance, Greenberg (1990) had pointed out that the vnM stable set (with respect to any binary relation) was based on optimistic predictions. These problems are avoided by Ray and Vohra (1997) in their solution concept of *equilibrium binding agreements* (EBA). They study the very general framework of normal form games. Their basic idea is that once a particular structure forms, the players within each coalition will have signed a binding agreement to cooperate with one another. But, prior to joining a coalition and agreeing on any specific course of actions, each player must predict what coalition structure will form in equilibrium. And players will evaluate their payoffs in the "equilibrium" before signing any binding agreement.

In what follows, we restrict attention to normal games where for each coalition structure or partition of the player set N, the game in which each coalition plays as a single "player" has a unique Nash equilibrium. Assuming that players within each coalition also agree on how to divide the coalitional payoffs corresponding to the Nash equilibrium, we then have a hedonic game with externalities. Start from the grand coalition N. Suppose S contemplates a deviation from N. Then, the "temporary" partition will be $\{S, N-S\}$. Of course, since players are farsighted, S will not necessarily believe that this is the end of the process. There are several possible types of deviations.

(i) S itself may break up further.

(ii) $N-S$ may break up

(iii) Some group from S may join with some group from $N-S$.

Ray and Vohra (1997) allow for deviations of types (i) and (ii).[30] So, partitions can only become finer. The coalition structure of singletons, say π^*, is stable by definition since there is no finer partition. Now, consider any partition π whose only finer partition is π^*. Then, π is an EBA if no one one wants to deviate to π^*. Recursively, suppose all the set of all EBAs which are finer than a given partition π have been determined. Then, any coalition contemplating a deviation from π will compare the payoffs in π to what they can get in the next finer EBA. In other words, at each stage potential deviators predict the partition from which there will be no further deviation. The "solution" of the game is the set of coarsest EBAs.

Diamantoudi and Xue (2007) reformulate the concept of EBAs for hedonic games with externalities. For this restricted class of games, the set of EBAs is a vNM stable set of a particular binary relation. Their analysis is somewhat simpler than the original definition, and we describe their reformulation.

Definition 14 *A coalition structure π' is reachable from π via a sequence of coalitions $T_0, \ldots, T_{K-1}$ if $T_k \in \pi'$ and $\pi_k \rightarrow_{T_k} \pi_{k+1}$ for all $k = 0, K-1$, with $\pi_0 \equiv \pi$ and $\pi_K = \pi'$. Moreover, if $\pi' \succ_j \pi_k$ for all $j \in T_k$, then π' sequentially dominates π.*[31]

Suppose π' sequentially dominates π and is an EBA. Can we rule then out π as an EBA? This would mean that members of T_0 have an optimistic prediction about the order of deviations - precisely the criticism we have made about $V(>>)$. Ray and Vohra are careful to avoid this pitfall.

[30]Bernheim, Peleg and Whinston(1987)'s notion of *Coalition-proof Nash equilibrium* assumes that only deviations of type (i) are possible. Yi (1997) studies coalition-proof Nash equilibria of the open membership game of coalition formation. In the context of networks, Dutta and Mutuswami (1997) and Dutta, Tijs and van den Nouweland (1998) apply coalition-proofness to select among equilibria in the Myerson network formation game. When link monotonicity holds, they show that coalition-proof equilibrium networks are equivalent to the complete network.

[31]Notice the similarity with indirect domination.

Definition 15 *π' RV-dominates π, $\pi' RV \pi$, if there exist $T_0, \ldots, T_{K-1}$ s.t. each $T_k \in \pi'$, and*

(i) π' is reachable from π via $T_0, \ldots, T_{K-1}$

(ii) $\pi' \succ_i \pi$ for all $\imath \in T_0$.

(iii) If $Q = \hat{\pi}$ or Q is reachable from $\hat{\pi}$ via a subcollection of $\{T_1, \ldots, T_{K-1}\} - \{T\}$, where $\pi \rightarrow_{T_0} \hat{\pi}$, then $\pi' RV Q$.

It is condition (iii) which avoids the problems which can arise by focusing on specific sequences of deviations. For, let T_0 be the coalition which has initiated the deviation from π. For suppose some coalition(s) in $\{T_1, \ldots, T_{K-1}\}$ deviate from the path prescribed in the move from π to π'. Then, the resulting partition is also RV-dominated by π'. Hence, condition (iii) gives the specific sequence of deviations from π to π' a certain robustness.

Diamantoudi and Xue (2007) prove the following.

Theorem 10 *The set $V(RV)$ is the set of EBAs.*

The concept of Efficient Binding Agreements captures "almost" perfectly the intuitive basis of stability in group formations. The only caveat is the restriction that coalitional deviations can only result in finer partitions. This makes the definition "workable" since π^*, the partition of singletons, is by definition an EBA. And this then allows the recursion to be well-defined. But, of course, π^* may not satisfy one's intuitive sense of stability if individuals across elements in a partition can execute joint deviations. In other words, a fully satisfactory definition of stability needs to allow for coalitional deviations which do not necessarily result in finer deviations. Unfortunately, as pointed out by Ray and Vohra (1997) and Ray (2007), this would then result in possibly infinite chains of deviations because of cycles - the same partition can figure infinitely often in any sequence.[32]

The issue of *existence* of EBAs is not of any interest since π^* is always an EBA. What is of interest is whether efficient outcomes can be sustained as EBAs. This issue goes back to the Coase (1937) who asserted that in a world of complete information, and if there are no restrictions on negotiations, individuals should be able to reach efficient outcomes. However, Ray and Vohra (1997, 2001) and Diamantoudi and Xue (2007) produce examples to show that this optimism is misplaced. The following example from Ray (2007) illustrates.

Example 15 *Consider a public good economy with 3 symmetric agents, where m units of a private good (money) generate m units of public good,*

[32] Diamantoudi and Xue (2007) extend the notion of EBAs by allowing for such deviations. Their solution concept is $V(>>)$. But, as Chwe (1994) pointed out, this can be too restrictive as a solution concept.

but generates utility cost of $(1/3)m^3$. *Individuals have endowment of money. Any coalition of* s *will contribute per capita amount of* $m(s)$ *to maximize*

$$sm(s) - (1/3)m(s)^3$$

If production elsewhere is z*, then payoff to coalition of size* s *is*

$$s\left[z + (2/3)s^{3/2}\right]$$

Hence, the partition function is

$$v(123) = 6\sqrt{3}, v(1|2|3) = (8/3, 8/3, 8/3), v(1|23) = (2\sqrt{2}+2/3, 2[1+2/3\sqrt{8}])$$

Of course, the grand coalition is the efficient partition. But, this cannot be sustained as an EBA. The per capita payoff in the grand coalition is $2\sqrt{3} < 2\sqrt{2} + 2/3$. If i leaves N, jk will stay together. That is, $\{\{i\}, \{jk\}\}$ is an EBA. But this ensures that the grand coalition cannot be an EBA.

Levy (2004) and Ray and Vohra (2001) illustrate the possibility of applying the concept of EBAs in specific contexts.

5 Group Formation in Real Time

In the preceding sections, we have discussed the formation of networks and coalitions in a static setting; that is, one where the set of individuals form either a network or a coalition structure once and for all with payoffs being generated only once. In this section, we describe some recent literature which models group formation in a dynamic context where for instance networks evolve over time or coalitions form and break up as individuals renegotiate for better payoffs. Although there has been relatively little work in this dynamic framework, the approach has some advantages over the more conventional approach. First, there are many contexts where it is perhaps a better description of how groups *actually* interact. For example, relationships typically evolve over time, suggesting that the structure of links in a communication network does change over time. Second, as Konishi and Ray (2003) point out, cycles in sequences of deviations which pose problems for a satisfactory definition of farsightedness can be handled easily in a dynamic setting. Payoffs from cycles can be evaluated just as any other sequence of deviations.

We first describe the dynamic formation of models of networks and coalitions where players do not bargain over payoffs. That is, these are models where in every period, each player's payoff is specified completely by the network or coalition structure which forms. We then go on to describe models where per period payoffs are determined by a dynamic bargaining procedure.

5.1 Dynamic Network Formation

Watts (2001) was the first to study the dynamic evolution of networks for the specific case of Jackson-Wolinsky communication networks. We describe a more general framework by not restricting attention to communication networks. Time is divided into discrete time periods $T = \{1, 2, \ldots, \}$. In any period t, g_t is the historically given graph. A pair i, j meet randomly with probability p_{ij} in period t. The selected pair can decide to

(i) Form the link ij if $ij \notin g_t$. In keeping with the usual assumption in network formation, the link forms if both i and j agree to form the link.

(ii) Either i or j can unilaterally break the link ij if $ij \in g_t$.

Assume that agents are *myopic*, so that each pair of active agents in any period t choose their actions only by looking at their t-period payoff. Not surprisingly, Watts is able to show that in the case of communication networks, the dynamic process does not always converge to the efficient network. For instance, if $\delta < c$, then no link can ever form since the pair forming the first link is better off (in a myopic sense) by not forming the link.[33]

Of course, the specification of myopic behavior is often an extreme assumption. Suppose, for instance, that the formation of the first link results in a payoff of $-\epsilon$ to each agent involved in the link, while two or more links give each linked agent a payoff of 1 million. Myopic agents simply cannot get the process off the ground, and so cannot exploit even very high increasing returns to network formation. There are two ways of avoiding this phenomenon. First, one can stick to the assumption of myopic behavior, but allow for the possibility that there may be exogenous shocks or "mutations" which cause a link to form (that would otherwise not form due to myopia), and thus help the process of network process. Second, one can assume that players are farsighted and so be willing to suffer a small initial loss in order to reap large gains in the future. Both avenues have been explored in the literature, with Jackson and Watts (2002), Feri (2007) and Tercieux and Vannetelbosch (2006) adopting the first approach, while Dutta, Ghosal and Ray (2005) model farsighted behavior in the dynamic network formation. We describe each in turn.

Jackson and Watts (2002) add random perturbations to the basic stochastic process described above, and examine the distribution over networks as the level of random perturbation goes to 0. The basic stochastic process and myopic behavior gives rise to the notion of *improving paths.* An improving path from any network g is a sequence of networks $\{g_0, g_1, \ldots g_K\}$ with $g = g_0$ such that for each $k = 0, \ldots, K-1$

(i) $g_{k+1} = g_k - ij$ for some ij such that $Y_i(g_k - ij) > Y_i(g_k)$

[33] On the other hand, the star network is efficient if $(\delta - \delta^2) < c < \delta + \frac{n-2}{2}\delta^2$.

(ii) $g_{k+1} = g_k + ij$ for some ij such that $Y_l(g_k + ij) > Y_l(g_k)$ for $l = i, j$.

So, an improving path is one where each pair of adjacent networks differ only by the addition or deletion of just one link, and the link addition or deletion is the result of myopic payoff-maximizing behavior by the concerned pair of agents. Improving paths must lead either to pairwise stable networks or cycles where the same set of networks is visited repeatedly over time.

Now consider a process of mutation so that in any period t, after actions are taken by the active pair, there is a probability $\epsilon > 0$ that a tremble takes place, and the link is deleted if present and added if absent. This process defines a Markov chain where the states are the networks existing at the end of every period. The Markov chain has a unique stationary distribution which converges to a unique limiting distribution as ϵ converges to 0. Call a network *stochastically stable* if it is in the support of this limiting distribution. Jackson and Watts (2002) show that the set of stochastically stable networks are those networks which minimize *resistance*.[34] Jackson and Watts provide an interesting application of their analysis to matching models like the marriage market and the college-admissions model. They show that the set of stochastically stable networks coincides with the set of core matchings in these models.

Dutta, Ghosal and Ray (2005) modify the analysis of Konishi and Ray (2003) (to which we turn in the next subsection) to model farsighted behavior when networks evolve over time. Unlike Watts (2001) and Jackson and Watts (2002), they assume that players take into account the discounted value of all future payoffs from any sequence of networks. They also allow the active pair in period t to unilaterally break some existing links in g_t.

Define a state to be a pair (g, ij) where g represents the current network while ij is the pair of active agents in any period. A (mixed) **Markov** strategy for any player is a probability distribution over possible actions at each state s in which player i is an active agent. It is Markovian because the actions of each pair of active agents depends only on the state s, and not on the history of how the network evolved in the past.

A strategy profile precipitates for each state s some probability measure λ_s over the feasible set $F(s)$ of future networks starting from s.

So, a Markov process is induced on the set S of states; while λ_s describes the movement to a new network, the given random choice of active players moves the system to a new active pair.

The process creates values for each player. For any strategy profile μ,

$$V_i(s, \mu) = \sum_{g' \in F(s)} \lambda_s(g')[Y_i(g') + \delta \sum_{i'j'} \pi(i'j')V_i(s', \mu)]$$

[34] The concept of *resistance* is originally due to Freidlin and Wentzell (1984). The resistance of a network measures how many mutations are needed in order to get away from the network to an improving path leading to another network.

where $\delta \in (0,1)$ is the discount factor, λ_s is the probability over $F(s)$ associated with μ, $\pi(i'j')$ is the probability that a pair $i'j'$ will be active "tomorrow", and s' stands for the state $(g', i'j')$.

An *equilibrium process of network formation* is a strategy profile μ with the property that there is no active pair at any state s which can benefit –either unilaterally or bilaterally — by departing from $\mu(s)$. Notice that this is the dynamic counterpart of the static concept of pairwise Nash equilibrium.

Define a state to be *absorbing* if there is no deviation from that state. A state is *strongly absorbing* if it is absorbing, and the process converges to that state from every other state. Strongly absorbing states are the dynamic counterpart of stable networks in the static setting.

DGR show that if the network structure satisfies Link Monotonicity, then for all δ, there will be *some* equilibrium μ^* such that the complete graph g^N will be strongly absorbing. Of course, Link Monotonicity does not ensure that the complete network is efficient. However, they also show that a strong form of increasing returns (which implies that g^N is the unique efficient network) ensures that g^N is a strongly absorbing graph for *some* equilibrium. This positive result cannot be improved to show that the complete network is absorbing for *all* equilibria - static coordination failures can occur even with strong increasing returns to scale. This is illustrated in the next example.

Example 16 *Let* $N = \{1,2,3,4\}$*,* $v(g^N) = 4$*,* $v(\{ij\}) = -100$ *and* $v(g) < 0$ *for all other* g*. Let the allocation rule be equal division within each component, and* $\delta > \frac{6}{151}$*.*

Let μ be such that each pair of active agents breaks all links at all networks, and also refuses to form any link. Then, the empty graph is the strongly absorbing graph.

This also turns out to be an equilibrium. For by following this strategy, $V_i(g, ij, \mu) = 0$. It is enough to check for deviation at g^N. If ij do not break any links at g^N, then

$$V(g^N, ij, \mu') = \frac{6(6 - 151\delta)}{(6-\delta)^2} < 0$$

5.2 Coalition formation over time

Konishi and Ray (2003) (henceforth KR) study coalition formation in a dynamic setting. Let us restrict attention to characteristic function games, where for technical convenience assume that each coalition S has a non-empty and *finite* set of payoff vectors in $\mathbb{R}^S$. Let X be the finite set of possible states, where a state is a pair $x = (\pi, u)$, where π is a coalition structure and u is such that u_S is a feasible payoff vector for each $S \in \pi$.

For each coalition S, let $F_S(x) = \{y \in X | x \rightarrow_S y\}$ be the set of states that S can reach by a coalitional deviation. KR consider a scenario in which at current state x, some coalition S is selected at random, and the selected coalition then chooses (perhaps randomly) some feasible state out of $F_S(x)$. So, there is a ransom transition from one state to another. This process is captured in the following definition.

Definition 16 *A process of coalition formation (PCF) is a transition probability* $p : X \times X \rightarrow [0, 1]$ *so that* $\sum_{y \in X} p(x, y) = 1$.

Player i's payoff from a sequence of probabilities $\{\lambda_t\}$ is

$$\sum_{t=0}^{\infty} \delta_i^t \left(\sum_{x \in X} \lambda_t(x) u_i(x) \right)$$

Hence, each PCF induces a value function for each i.

$$V_i(x, p) = (1 - \delta) u_i(x) + \delta_i \sum_{y \in X} p(x, y) V_i(y, p)$$

Fix x, S, p. Say that S has a *weakly profitable move* from x if there is $y \in F_S(x)$ such that $V_i(y, p) \geq V_i(x, p)$ for all $i \in S$. S has a *strictly profitable move* if the inequality is strict for all $i \in S$.

A move y is *efficient* for S if there is no z such that $V_i(z, p) > V_i(y, p)$ for all i.

Definition 17 *A PCF p is an equilibrium process of coalition process (EPCF) if*

(i) whenever $p(x, y) > 0$ for some $y \neq x$, then there is S such that y is a weakly profitable and efficient move for S.

(ii) if there is some strictly profitable move from x, then $p(x, x) = 0$ and there is a strictly profitable and efficient y such that $p(x, y) > 0$.

A PCF is an "equilibrium" if at all dates and all states, a coalitional move to another state is "justified" only if the move gives the coalition a higher present value. Notice that this incorporates farsighted behavior because individuals evaluate PCFs in terms of the discounted value of future payoffs. KR show that an equilibrium process of coalition formation exists. An equilibrium PCF may not necessarily be *deterministic*, where a deterministic PCF has $p(x, y) \in \{0, 1\}$.

KR establish a striking result which provides a strong justification for the *core* as a solution concept in this dynamic setting. Call a state x to be *absorbing* if $p(x, x) = 1$, and a PCF to be absorbing if for every y, there is an absorbing x s.t. $p^k(y, x) > 0$. A PCF has unique limit if it is absorbing and possesses a single absorbing state. They show that every core allocation can be described as the limit of some deterministic EPCF. Conversely, if a deterministic EPCF has a unique limit, then that limit must be a core allocation.

What happens if deterministic EPCFs do not have a unique absorbing limit? KR show that all absorbing deterministic EPCFs have absorbing states which lie in the LCS. They also construct an interesting example which shows that not all states in the LCS can be supported as absorbing states of EPCFs. This illustrates the point we made in the previous section that the LCS is too "large".

This line of thought also suggests that the set of such absorbing states can itself be interpreted as a farsighted solution set. Clearly, the analysis of group formation in a dynamic setting is a very promising area, where much work needs to be done. Perhaps, it would be more fruitful to focus on specific applications rather than adopt the more abstract framework.

5.2.1 Coalitional Bargaining over Time

The coalitional bargaining models described in Section 3 assume that coalitions leave the game after they are formed. The extensive forms do not allow for renegotiations, and the Coasian intuition that bargaining should allow players to reach efficient outcomes is not supported by the models. Clearly, the assumption that coalitions exit the game after they are formed is a convenient device to solve the coalitional bargaining model recursively, but is hard to justify in real world negotiations. The models we discuss here allow players to renegotiate agreements continuously.

Seidmann and Winter (1998) propose a model with continuous renegotiation, which is a direct extension of Chatterjee et al. (1993).[35]At the initial phase, one player is chosen according to a fixed protocol to be the proposer. This player can either pass the initiative to another player or propose an agreement (S, x) where x is a vector of payoffs for all members of coalition S satisfying $\sum x_i = v(S)$. If one member of S rejects the offer, she becomes the proposer at the next period. If all members agree, the coalition is formed, and (S, x) becomes the current state. Any state of the game can thus be described as a list of "interim agreements" (S_t, x_t) between the players. After the initial period, the game enters a phase of renegotiation. The protocol selects one player who can either pass the initiative to another player or make a proposal (R, y) which must satisfy $\sum y_i = v(R)$ and if $R \cap S_t \neq \emptyset$ then $S_t \subseteq R$. This last condition states that a proposal must include *all the members of interim coalitions* and defines the role of interim coalitions in the model.

[35]This is the model with "reversible" actions in Seidmann and Winter (1998).

Seidmann and Winter (1998)'s main emphasis is on the dynamics of coalition formation. They first show that stationary perfect equilibria of the game will always lead to the formation of the efficient grand coalition. However, this process can either be immediate or gradual, according to the properties of the underlying coalitional game.

Theorem 11 *If the core of the underlying coalitional game is empty, then there exists* $\underline{\delta} < 1$ *such that for all* $\delta \geq \underline{\delta}$*, the game does not possess a stationary perfect equilibrium where the grand coalition is formed immediately.*

The previous theorem clarifies conditions under which the grand coalition forms immediately. It shows that agreement *must be gradual* if the game has an empty core. While this result provides a necessary condition for immediate agreement, it does not give a sufficient condition. Seidmann and Winter (1998) note that in a class of games where the marginal worth of the grand coalition is very large (and hence the core is nonempty), there will always be an equilibrium with immediate agreement.

Okada (2000) proposes a model of bargaining with renegotiation based on Seidmann and Winter (1998) where proposers are chosen at random every period. Based on Okada (1996)'s analysis, he shows that agreements will be reached at every period, so that the grand coalition forms in at most $(n-1)$ steps.

Gomes (2005) extends the analysis to games in partition function form. In his model, proposers are chosen at random (as in Okada (2000)) and probability p_i denotes the recognition probability of player i. A contract consists in a coalition S and a vector of contingent payoffs $x_i(\pi)$ which depend on the entire coalition structure π and verify $\sum_{i \in S} x_i(\pi) = v(S, \pi)$ for all coalition structures such that $S \in \pi$. Gomes (2005)'s first result extends Okada (2000)'s analysis in showing that, when the grand coalition is efficient, it will ultimately form in a finite number of steps. He also proposes a condition on the partition function which guarantees that the grand coalition forms immediately. In order to describe this condition, define, for any collection $\mathcal{S}$ of coalitions in a coalition structure π, the coalition structure obtained by merging all coalitions in $\mathcal{S}$ by $\pi\mathcal{S}$, i.e. $\pi\mathcal{S} = \pi \backslash (\cup_{S \in \mathcal{S}} S) \cup (\cup_{S \in \mathcal{S}} S)$.(By an abuse of notation, we shall also denote $\mathcal{S} = \cup_{S \in \mathcal{S}} S$.)

Theorem 12 *Suppose that for all coalition structures* π *and all collections of coalitions* $\mathcal{S}$ *in* π,

$$v(\mathcal{S}, \pi\mathcal{S}) + \sum_{i \in \mathcal{S}} p_i (v(N) - \sum_{T \in \pi\mathcal{S}} v(T, \pi\mathcal{S})) \leq \sum_{S \in \mathcal{S}} (v(S, \pi) + \sum_{i \in S} p_i (v(N) - \sum_{T \in \pi} v(T, \pi))),$$

then there exists a stationary perfect equilibrium where the grand coalition forms immediately at any state. Conversely, if this condition fails, there exists a state at which the grand coalition does not form in any stationary perfect equilibrium for δ large enough.

Gomes and Jehiel (2005) extend the preceding analysis by considering a general setup, described by a set of states in the economy, and an effectivity function over the states. At every period in time, a player is chosen at random to make an offer, consisting of a transition from the current state to a new state, and a vector of transfers to members of the coalition effective in that transition. Any Markov perfect equilibrium of this game generates a Markov process, and Gomes and Jehiel (2005) first study the convergence properties of this process. Their main result shows that, if players are sufficiently patient, the aggregate value is the same in any recurrent set.[36]

Theorem 13 *The aggregate equilibrium values are approximately the same at all states in the recurrent sets for all economies when δ converges to one.*

As a consequence of this theorem, generically, the Markov process can only admit one recurrent set set. The preceding theorem does not imply that the process always converge to an *efficient* state. In fact, this may not be the case, and one needs the following sufficient condition to guarantee that the process converges to an efficient recurrent set.

Definition 18 *A state a is* negative externality free *if for all $i \in N$, all sequences of moves, $a \to_{S_1} \to \ldots \to_{S_K} b$ involving coalitions $S_k \subseteq N \setminus \{i\}$, $v_i(b) \geq v_i(a)$. A state a is an ENF state if it is efficient and negative externality free.*

Theorem 14 *Any ENF state is in a recurrent set. Hence, if there exists an ENF state, the Markov process always converges to an efficient allocation.*

Hyndman and Ray (2007) build on Gomes and Jehiel (2005) and Konishi and Ray (2003) to propose a general dynamic model of coalition formation, without imposing any stationarity restriction. Like Gomes and Jehiel (2005), they consider an abstract set of states, and an effectivity function, specifying which coalitions are effective in the transition between states. Their first result shows that, in the absence of externalities across coalitions, all equilibria are asymptotically efficient.

[36] A recurrent set of a Markov process is the set of states which are reached in the long run,i.e. a set of states that the Markov process never leaves once it has reached it, and such that the Markov process visits all states in the recurrent set.

Theorem 15 *In characteristic function games with permanently binding agreements, any benign subgame perfect equilibrium (where players always opt for a strategy which makes other players better off when they are indifferent) results in an asymptotically efficient payoff.*

In the presence of externalities, the efficiency result fails, as the following example shows

Example 17 $N = 3$ *and payoff vectors are as follows* $v(1|2|3) = (6,6,6), v(1|23) = (0,10,10), v(12|3) = (5,5,0), v(13|2) = (5,0,5), v(123) = (0,0,0)$

In this example, the two states $\{12|3\}$ and $\{13|2\}$ are Pareto dominated by state $\{1|2|3\}$, but they are absorbing states. Player 1 will never accept to move to state $\{1|2|3\}$, because she knows that players 2 and 3 will subsequently form coalition 23. Hence, any transition out of the two states $\{12|3\}$ and $\{13|2\}$ will be blocked by player 1.

Hyndman and Ray (2007) provide sufficient conditions under which games with externalities result in efficient outcomes (e.g. "grand coalition superadditivity", which states that the grand coalition is efficient). They also provide a four player example where, starting from any state, the outcome is asymptotically inefficient. Hence, the final conclusion of their analysis is mixed: the presence of externalities does indeed lead to inefficiencies, which can only be alleviated by placing stringent restrictions on the value functions.

The previous models assume that players have the ability to continuously renegotiate agreements until the grand coalition forms. Two studies have considered what happens when players choose *endogenously* whether to exit the game. In these contributions, the players' decisions is whether to take an outside option (defined by the current contract or agreement) or to continue negotiating. As the grand coalition is always assumed to be efficient, the issue is whether players ever choose to exit the game inefficiently early.

Seidmann and Winter (1998) argue, through an example, that this may indeed occur in a model of coalitional bargaining. The model they construct follows the same rules as the model of coalitional bargaining with renegotiation described above, with the following differences.[37] At any period, after a contract has been offered and respondents have made their decisions, players are given the opportunity to implement the current agreement. Implementation means that the players commit, in an irreversible fashion, to leave the negotiations with their current agreement. In Seidmann and Winter (1998)'s model, all players simultaneously decide whether to implement, and every

[37] Seidmann and Winter (1998) call this version of their model the model with "irreversible decisions".

player has a veto power over the decision for his coalition. In other words, a contract involving a coalition S of players is implemented as soon as one of the players in S chooses to implement it.[38]

Example 18 $N = 3, v(S) = v > 2/3$ *if* $|S| = 2, v(N) = 1$. *The protocol specifies that, after a two-player coalition forms, the next proposer is the outsider.*

In this example, all stationary perfect equilibria are inefficient, and result in players forming a coalition of size 2 and then exiting with positive probability. Note first that because the core of the game is empty, the grand coalition cannot form immediately. Suppose that the grand coalition indeed forms after a two-player coalition is formed. Without loss of generality, suppose that the coalition $\{1, 2\}$ forms. The interim contract is given by (x_1, x_2) with $x_1 + x_2 = v$. This defines the outside options of players 1 and 2. Now suppose that player 3 makes an offer which is indeed accepted. If the outside options of the other two players are not binding, then the offer will converge to $(\delta/(1+2\delta), \delta/(1+2\delta), 1/(1+2\delta))$. But, because $v > 2/3$ either x_1 or x_2 must be larger than $1/3$ and hence one of the two outside options has to be binding – a contradiction. If now one of the outside options is binding (say x_1) then, as long as $\delta < 1$, player 1 has an incentive to implement the contract early rather than wait to obtain his outside option.

Bloch and Gomes (2006) revisit the issue of endogenous exit in a model with random proposers and externalities across coalitions. They assume that players are engaged in two parallel interactions: they form coalitions, and take part in a game in strategic form which determines the flow payoffs at every period. Every period is separated into two subphases. Players first take part in a contracting phase, where one of the players, chosen at random, proposes to buy out the resources of other players. Every member of the coalition then responds in turn to the offer, and resources are bought only if all coalition members agree. In the second phase, the action phase, all active players choose an action in a set which contains both reversible and irreversible decisions. Reversible decisions (or inside options) determine the flow payoff of the current period, irreversible decisions (or outside options) affect the payoff of all future periods.

Bloch and Gomes (2006) distinguish between two cases: outside options are *pure* if every player is guaranteed to obtain the same payoff by exiting, irrespective of the choices of other players. Otherwise, the model displays

[38]This last hypothesis allows for ”trivial” equilibria, where all players choose to implement the contract. These equilibria of course give rise to inefficient exit, but arise purely because of coordination failures among members of a coalition.

externalities in outside options. Bloch and Gomes (2006) first show that when the choice of proposers is random, in contrast to Seidmann and Winter (1998), there always exists a stationary perfect equilibrium where players make acceptable offers at each period. More precisely, one has to refine the set of equilibria, to take care of coordination failures arising from the fact that all players simultaneously choose whether to exit (and hence, all players exiting may be an equilibrium, even though it is a dominated equilibrium for all the players). Bloch and Gomes (2006) define $\varepsilon - R$ equilibria as equilibria where all players remain in the game with probability ε at every state.

Theorem 16 *For any underlying game with pure outside options and any ε, there exists $\underline{\delta}$ such that for all $\delta \geq \underline{\delta}$ an $\varepsilon - R$ stationary perfect equilibrium exists. Furthermore, as δ converges to 1, the probability of exit in all $\varepsilon - R$ equilibria converges to zero, and hence the outcome of the coalitional bargaining game is approximately efficient.*

However, Bloch and Gomes (2006) note that this result depends crucially on the fact that outside options are pure, and hence the same outside option is available from one period to the next. In games with externalities in outside options, the argument breaks down, and all equilibria may be inefficient.

6 The Tension between Efficiency and Stability

A recurrent theme in the previous sections has been that when coalitions or networks form endogenously, the efficient group structure is often not supportable as a stable outcome. Notice that the failure of the efficient group structure to be supported as a stable outcome occurs in a framework of complete information and in a frictionless world - there are no transaction costs constraining the attainment of the "optimal" structure. This seems to go squarely against the Coasian intuition that rational agents should be able to attain efficiency.

Following the work of JW, several authors have discussed the different ways in which stability can be reconciled with efficiency in the context of the endogenous formation of networks. We review some of this literature in this section.

JW show that if an allocation rule satisfies *component balance* and *anonymity*, [39] then there may be value functions for which no efficient network is pairwise stable.

[39] Component Balance means that there can be no cross-subsidization across components of a network. Anonymity means that individuals who are "alike" in the sense of occupying symmetric positions in the network as well as in terms of contribution to the value must be given equal rewards.

Theorem 17 *Let the allocation rule satisfy anonymity and pairwise stability. Then, there is a value function such that no efficient network is pairwise stable.*

The proof consists of a counterexample. Let $N = \{1, 2, 3\}$ and consider the value function v such that $v(g) = 12$ if g is the complete graph, or if g has exactly one link, and $v(g) = 13$ if g has exactly two links.

Then, g is efficient if and only if it has exactly two links. The proof is completed by showing that no graph with two links can be pairwise stable if the allocation rule Y satisfies anonymity and component balance.

Since v is symmetric, it is sufficient to show that if $g^1 = \{ij, ik\}$, then g^1 is not pairwise stable. Let $g^2 = \{ij\}$. Now, $Y_i(g^2, v) = 6$ since Y satisfies anonymity and component balance. So, $Y_i(g^1, v) \geq 6$ if g^1 is to be pairwise stable. Otherwise, i can sever his link with k.

Hence, from Anonymity we have that $Y_k(g^1, v) = Y_j(g^1, v) \leq 3.5$. But, now both j and k have the incentive to form the link jk since both get 4 at the complete graph.

This shows that g^2 is not pairwise stable.[40]

The tension between efficiency and stability surfaces again even in a dynamic setting when networks evolve over time. DGR (2005) prove a dynamic version of the JW result by constructing an example in which no efficient network is strongly efficient if the allocation rule satisfies efficiency and *limited liability.*

Definition 19 *: Y satisfies limited liability if for all $i \in N$, $i \in N(h)$ and $v(h) \geq 0$ implies $Y_i(g, v) \geq 0$ for every $g \in G$.*

Example 19 *Let $N = \{1, 2, 3\}$. Consider a symmetric value function where $v(g) = 2\alpha$ if g has just one link, $v(g^N) = 3\alpha$ and $v(g) = 0$ otherwise, where α is some positive number.*

The unique efficient graph is the complete graph g^N. However, there is no pure strategy equilibrium μ^* such that g^N is strongly absorbing.

For consider any pure strategy equilibrium μ^*. Notice that since Y is anonymous, $Y_i(\{i, j\}) = Y_i(g^N) = \alpha$. Also, for all other g, $Y_\ell(g) = 0$ for all $\ell \in N$ from limited liability.

If g^N is to be strongly absorbing, then there must be a path from a one-link network to the complete network. In particular, this means that there must be some pair (i, j) and k such i, j form the link ij and then (say) i forms the link with k.

[40]For a similar result for directed networks, see Dutta and Jackson (2000).

Suppose i and j have formed the link ij. Then, i and j both have a payoff of α. Notice that no graph gives them a higher payoff while the two-link graph gives both i, j a strictly lower payoff. Clearly, neither i nor j have an incentive to form the link with k.

This is essentially a heuristic proof of the following theorem.

Theorem 18 *Suppose the allocation rule satisfies anonymity and limited liability. Then, there is a value function such that no efficient network is strongly absorbing.*

The literature on ways of resolving the tension between efficiency and stability can be divided into two broad areas. The first approach assumes that the individual agents who form the nodes can influence their payoffs or rewards only through their decisions on which network to form. In other words, the *allocation rule* itself is specified exogenously to the agents. In the second approach, the agents determine both the network as well as the allocation rule, perhaps through bargaining.

Consider the first approach, and suppose the allocation rule is such that *all* agents receive the *average* payoff $v(g)/n$. The "average rule" fully aligns individual incentives with social goals, so that there can be no conflict between efficiency and stability with this rule. However, the average rule does not satisfy Component Balance, one of the specified restrictions on the allocation rule in the J-W Theorem. Of course, one could ask why it is important to restrict attention to allocation rules satisfying Component Balance. The example used to prove the JW Theorem provides an answer. Consider any of the one link graphs g, say $g = \{ij\}$. The average rule requires that k be given 4. But why should i and j agree to this rule? If they have the option of breaking away from the "society", they will certainly exercise this option.

A rule which satisfies Component Balance but is similar in spirit to the average rule is the "Component Average Rule" which divides payoffs equally within each component. Denoting this rule as Y^a, we have

$$Y_i^a(v, g) = \frac{v(h)}{n_h} \text{ for all } i \in N(h), h \in C(g)$$

It is of some interest to ask the condition under which Y^a will ensure stability of efficient networks. For any network g, call ij a *critical link* in g if $g - ij$ has more components than g. In other words, the severance of a critical link breaks up an existing link into two components.

Definition 20 *A pair (v, g) satisfies Critical Link Monotonicity if for any critical link in g and its associated components h, h_1, h_2, $v(h) \geq v(h_1)+v(h_2)$ implies that $v(h)/n_h \geq \max\left[v(h_1)/n_{h_1}, v(h_2)/n_{h_2}\right]$.*

The following theorem is due to JW.

Theorem 19 *Let g be any efficient network. Then, g is pairwise stable given the allocation rule Y^a iff (g, v) satisfies Critical Link Monotonicity.*

This theorem uses Y^a. An interesting question is to characterize sets of value functions under which efficient networks can be supported as pairwise stable networks for other allocation rules.

The Mechanism Design Approach

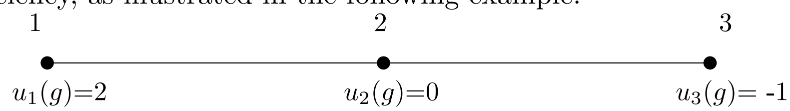
Suppose the implicit assumption or prediction is that only those networks which correspond to Pairwise Nash equilibria or Strong Nash equilibrium of the link formation game will form. Then our interest in the *ethical* properties of the allocation rule should be restricted only to how the

rule behaves *on the class of these networks*. That is, since networks outside this class will not form, why should we bother about how the allocation rule behaves on these networks?

So, suppose the prediction is that only strongly stable networks will form. Then, if we want symmetry of the allocation rule, we should be satisfied if the allocation rule is symmetric on the *subdomain* of strongly stable graphs. This gives some freedom about how to specify the allocation rule. Choose some efficient $g^* \in G$. Suppose s^* induces g^*, and we want to ensure that g^* is strongly stable. Now, consider any g which is different from g^*, and let s induce g. Then, the allocation rule must *punish* at least one agent who has deviated from s^* to s. This is possible only if a deviant can be *identified.* This is trivial if there is some $(ij) \in g|g^*$, because then both i and j must concur in forming the extra link (ij). However, if $g \subset g^*$, say $(ij) \in g^*|g$, then *either i or j* can unilaterally break the link. The only way to ensure that the deviant is punished is to punish *both i* and j.

Several simple punishment schemes can be devised to ensure that at least two agents who have deviated from s^* are punished sufficiently to make the deviation unprofitable. However, since the allocation rule has to be component balanced, these punishment schemes may result in some other agent being given more than the agent gets in g^*. This possibility creates a complication because the punishment scheme has to *satisfy an additional property.* Since we also want to ensure that inefficient graphs are *not* strongly stable, agents have to be provided with an incentive to deviate from any graph which is not strongly efficient. Hence, the punishment scheme has to be relatively more sophisticated. Dutta and Mutuswami (1997) describe conditions under which this approach will reconcile the conflict between efficiency and stability.

An example of the second approach is the paper by Bloch and Jackson (2007) on network formation with transfers. Bloch and Jackson (2007) allow payoffs to be determined endogenously since players are allowed to make transfers. They highlight two difficulties in reaching efficient networks in a model of network formation with transfers. The first difficulty is due to the presence of externalities. If the formation of a link produces positive externalities on other players, direct transfers may not be sufficient to attain efficiency, as illustrated in the following example:



All other networks result in a utilities of 0 for all players.

The efficient network is the line $\{12, 23\}$. For this network to be supported, we must have $t^3_{23} = -1$. But if $t^2_{23} = 1$, player 2 has an incentive to deviate, so network $\{12, 23\}$ cannot be supported in equilibrium. Notice that if player 1 could subsidize the formation of link 23, this difficulty will disappear.

However, there remains another, more subtle difficulty, due to the fact that players' marginal benefits from a set of links may not be equal to the sum of the marginal benefits of every link in the set.

Example 20 *Consider a three-player society and a profile of utility functions described as follows. Any player gets a payoff of 0 if he or she does not have any links. Player 1 gets a payoff of 2 if she has exactly one link, and a payoff of 1 if she has two links. Player 2 gets a payoff of -2 if he has exactly one link, and a payoff of 0 if she has two links. Player 3's payoff function is similar to that of player 2*

It is clear from this specification that all players' payoffs depend only on the configuration of their own links and so there are no externalities in payoffs. However, we claim that the efficient network structure (the complete network) cannot be reached in equilibrium. By setting $t^2_{2i} = 0$ for each i, player 2 gets a payoff of at least 0. The same is true for player 3. Thus, players 2 and 3 must have a payoff of at least 0 in any equilibrium. Now, suppose by contradiction that the complete network were supported in an equilibrium. It would follow that $t^1_{1i} = 0$ for at least one i, or otherwise one of players 2 and 3 would have a negative payoff. Without loss of generality, suppose that $t^1_{12} = 0$. Player 1's payoff would then be $1 - t^1_{12} - t^1_{13}$. Suppose that player 1 deviated so that network $13, 23$ forms. Then, player 1's payoff would become $2 - t^1_{13}$ which is greater than $1 - t^1_{12} - t^1_{13}$. For this inefficiency to disappear, one needs to allow transfers which are contingent on the entire network.

7 Conclusions and Open Problems

This paper has surveyed a large number of models by which agents driven by their self-interest choose to cooperate in coalitions and networks. The models belong to three broad categories. The earlier literature considered static models of group formation, where players simultaneously announce the links or group they want to form. Following Rubinstein (1981)'s analysis of two-player bargaining, the next step has been to represent group and network formation as a sequential bargaining process. Finally, more recent work has focussed on general, abstract dynamic processes of coalition and network formation among farsighted players.

The three categories of models have advantages and disadvantages. Static games are easy to analyze, but typically exhibit multiple equilibria due to coordination failures, and refinements must be used to obtain clear predictions. Sequential bargaining models generically result in a single outcome, but the characterization of equilibrium networks or coalition structures requires complex recursive computations, and the outcome of the game may depend on details of the bargaining protocol. Dynamic processes are immune to that last criticism, as they model the process of network or group formation as an abstract game, but cannot be used to predict the exact outcome of particular models of cooperation.

What are the next steps in the study of coalition and network formation? In our opinion, three basic issues remain unexplored and should stimulate new research in the next few years.

Robustness Analysis: The large variety of models of coalition and network formation results in a large variety of predictions on the coalitions and networks which are likely to form in specific situations. For sequential games, the outcome of the process of coalition and network formation depends on fine details of the bargaining protocol. One needs to get a better understanding of the mapping between the procedure of coalition or network formation and the equilibrium outcomes. What are the features of the procedure of coalition or network formation which guarantees that efficient outcomes arise? Which processes result in single or multiple outcomes? How does the absence or presence of transfers affect the outcome of the procedure and the ability to reach efficient results? Can players manipulate procedures of coalition formation to their advantage? The answer to these important questions requires a meta-analysis of the procedures of coalition and network formation.

Coalitions versus Networks: This survey shows that networks and coalitions are related yet different modes of cooperation among agents. In one sense, networks are more general descriptions of cooperation possibilities, as they include information not only about the identity of groups of agents who cooperate (the components of the network), but also about the details of the structure of bilateral links which favors cooperation. For each network, one can derive a single coalition structure by distinguishing network components, but each coalition structure may be obtained from a large number of different networks. At another level though, networks and coalitions are different but unrelated ways of describing cooperation: some situations are inherently bilateral (e.g. face to face communication) whereas others are multilateral (e.g. large meetings). The choice between describing cooperation through coalitions or networks is then driven by the application. Even when the description of gains from cooperation is the same for bilateral and multilateral interactions (for example when gains from cooperation are described by a game in partition function form), the process of cooperation (bilateral or multilateral) may affect the outcome. Forming bilateral links is easier than agreeing with a group of agents, inducing more cooperation in networks than coalitions. But if agents can unilaterally sever some of their links while keeping others, they may be less likely to cooperate in the first place. Examples can be given to show that network formation models result in more cooperation as in the formation of cost-reducing alliances among firms studied by Bloch (1995) for coalitions and Goyal and Joshi (2000) for networks. But there are also examples pointing in the other direction where cooperation is easier in coalitions than networks as in the models of informal insurance in groups and networks of Genicot and Ray (2003) and Bloch, Genicot and Ray (2008).

Incomplete information: All the models considered in this survey assume that agents possess complete information on gains from cooperation. Relaxing this assumption is a difficult task, as it would introduce a number of new problems essentially revolving around the amount of information which can be credibly shared amongst members of a coalition. compatibility. Wilson(1968) initiated some recent literature on the core of the exchange economy with incomplete information. [41] However, the literature group formation with incomplete information is still in its infancy, and much remains to be done.

8 References

d'Aspremont, C., A. Jacquemin, J. Jaskold Gasbszewicz and J. Weymark (1983) "On the Stability of Price Leadership," *Canadian Journal of Economics* **16**, 17-25.

Aumann, R. and R. Myerson (1988) "Endogenous Formation of Links Between Players and of Coalitions: An Application of the Shapley Value," in *The Shapley Value: Essays in Honor of Lloyd Shapley*, A. Roth ed., 175-191. Cambridge University Press.

[41] Forge et al (2002) is an elegant survey of some of this literature.

Bala, V. and S. Goyal (2000) "A Noncooperative Model of Network Formation," *Econometrica* **68**, 1181-1229.
Banks, J. and J. Duggan (2000) "A Bargaining Model of Collective Choice *American Political Science Review* **94**, 73-88.
Baron, D. and J. Ferejohn (1989) "Bargaining in Legislatures," *American Political Science Review* **83**, 1181-1206.
Baron, D. and E. Kalai (1993) "The Simplest Equilibrium of a Majority-Rule Division Game," *Journal of Economic Theory* **61**, 290-301.
Bernheim, D., Peleg, B. and M. Whinston (1987) "Coalition-Proof Nash Equilibria I. Concepts" *Journal of Economic Theory* **42**, 1-12.
Billand, P. and C. Bravard (2005) "A Note on the Characterization of Nash Networks," *Mathematical Social Sciences* **49**, 355-365.
Binmore, K. (1985) "Bargaining and Coalitions", in *Game-Theoretic Models of Bargaining*, A. Roth ed., Cambridge: Cambridge University Press.
Binmore, K., A. Rubinstein and A. Wolinsky (1986) "The Nash Bargaining Solution in Economic Modelling," *RAND Journal of Economics* **17**, 176-188.
Bloch, F. (1995) "Endogenous Structures of Association in Oligopolies, *RAND Journal of Economics* **26**, 537-556.
Bloch, F. (1996) "Sequential Formation of Coalitions in Games with Externalities and Fixed Payoff Division", *Games and Economic Behavior* **14**, 90-123.
Bloch, F. and B. Dutta (2009)"Communication Networks with Endogenous Link Strength", *Games and Economic Behavior* **66**, 39-56.
Bloch, F. and A. Gomes (2006) "Contracting with Externalities and Outside Options", *Journal of Economic Theory* **127**, 172-201.
Bloch, F. and M.O. Jackson (2006) "Definitions of Equilibrium in Network Formation Games *International Journal of Game Theory* **34**, 305-318.
Bloch, F., G. Genicot and D. Ray (2008) "Informal Insurance in Social Insurance", *Journal of Economic Theory* **143**, 36-58.
Bloch, F. and M.O. Jackson (2007) "The Formation of Networks with Transfers between Players", *Journal of Economic Theory* **133**, 83-110.
Bondareva O. (1963) "Some Applications of Linear Programming Methods to the Theory of Cooperative Games (In Russian)", *Problemy Kybernetiki* **10**, 119-139.
Bramoulle, Y. and R. Kranton (2007) "Risk Sharing Networks, *Journal of Economic Behavior and Organization* **64** 275-294.
Calvo-Armengol, A. and R. Ilkilic (2009) "Pairwise Stability and Nash Equilibria in Network Formation," *International Journal of Game Theory* **38**, 51-79.
Chatterjee, K., B. Dutta, D. Ray and K. Sengupta (1993) "A Noncooperative Theory of Coalitional Bargaining", *Review of Economic Studies*

60, 463-477.
Chwe, M. (1994) "Farsighted Coalitional Stability", *Journal of Economic Theory* **63**, 299-325.
Coase R. (1937) "The Nature of the Firm," *Economica*, 386-405.
Currarini, S. and M. Morelli (2000) "Network formation with Sequential Demands", *Review of Economic Design* **5**, 229-250.
Declippel, G. and R. Serrano (2008) "Marginal Contributions and Externalities in the Value", *Econometrica* **76**, 1413-1436.
Demange, G. and M. Wooders, eds. (2005) *Group Formation in Economics: Networks, Clubs and Coalitions*, Cambridge: Cambridge University Press.
Diamantoudi, E. and L. Xue (2003) "Farsighted Stability in Hedonic Games," *Social Choice and Welfare* **21**, 39-61.
Diamantoudi, E. and L. Xue (2007) "Coalitions, Agreements and Efficiency," *Journal of Economic Theory* **136**, 105-125.
Diermeier, D., H. Eraslan and A. Merlo (2003) "A Structural Model of Government Formation," *Econometrica* **71**, 27-70.
Diermeier, D. and A. Merlo (2004) "An Empirical Investigation of Coalitional Bargaining Procedures," *Journal of Public Economics* **88**, 783-797.
Dréze, J. and J. Greenberg (1980) "Hedonic Coalitions: Optimality and Stability," *Econometrica* **48**, 987-1003.
Dutta, B., L. Ehlers and A. Kar (2009) "Externalities, Potential, Value and Consistency", mimeo., Department of Economics, University of Montreal.
Dutta, B., S. Ghosal and D. Ray (2005) "Farsighted Network Formation", *Journal of Economic Theory* **122**, 143-164.
Dutta, B. and M.O. Jackson (2000) "The Stability and Efficiency of Directed Networks", *Review of Economic Design* **5**, 251-272.
Dutta, B. and S. Mutuswami (1997) "Stable Networks", *Journal of Economic Theory* **76**, 322-344.
Dutta, B. and D. Ray (1989) "A Concept of Egalitarianism under Participation Constraints," *Econometrica* **57**, 615-635.
Dutta, B., S. Tijs and A. van den Nouweland (1998) "Link Formation in Cooperative Situations," *International Journal of Game Theory* **27**, 245-256.
Eraslan, H. (2002) "Uniqueness of Stationary Equilibrium Payoffs in the Baron-Ferejohn Model," *Journal of Economic Theory* **103**, 11-30.
Feri, F. (2007) "Stochastic Stability in Networks with Decay," *Journal of Economic Theory* **135**, 442-457.
Forge, F., E. Minelli and R.Vohra (2007), "Incentives and the Core of an Exchange Economy: a survey", *Journal of Mathematical Economics*,

38, 1-41.
Freidlin, M. and A. Wentzell (1984) *Random Perturbations of Dynamical Systems*, New York: Springer Verlag.
Galeotti, A., S. Goyal and J. Kamphorst (2006) "Network Formation with Heterogeneous Players," *Games and Economic Behavior* **54**, 353-372.
Genicot, G. and D. Ray (2003) "Group Formation in Risk-Sharing Arrangements," *Review of Economic Studies* **70**, 87-113.
Gilles, R. and S. Sarangi (2006) "Building Social Networks," mimeo., Virginia Tech and Louisiana State Universities.
Gilles, R., S. Chakrabarti and S. Sarangi (2006) "Social network Formation with Consent: Nash Equilibria and Pairwise Refinements," mimeo., Virginia Tech, Queen's University Belfast and Louisiana State University.
Gomes A. (2005) "Multilateral Contracting with Externalities," *Econometrica* **73**, 1329-1350.
Gomes, A. and P. Jehiel (2005) "Dynamic Processes of Social and Economic Interactions: On the Persistence of Inefficiencies," *Journal of Political Economy* **113**, 626-667.
Goyal S. (2007) *Connections: An Introduction to the Economics of Networks*, Princeton: Princeton University Press.
Goyal, S. and S. Joshi (2003)"Networks of Collaboration in Oligopoly ", *Games and Economic Behavior*, **43**, 57-85.
Greenberg, J. *The Theory of Social Situations: An Alternative Game-Theoretic Approach*, Cambridge: Cambridge University Press.
Gul F. (1989) "Bargaining Foundations of the Shapley Value," *Econometrica* **57**, 81-95.
Haller, H. and S. Sarangi (2005) "Nash Networks with Heterogeneous Links," *Mathematical Social Sciences* **50**, 181-201.
Harrington, J. (1989) "The Advantageous Nature of Risk Aversion in a Three Player Bargaining Game where Acceptance of a Proposal Requires a Simple Majority," *Economics Letters* **30**, 195-200.
Harsanyi, J. (1974) "An Equilibrium Point Interpretation of Stable Sets and a Proposed Alternative Definition," *Management Science* **20**, 1472-1495.
Hart, S. and M. Kurz (1983) "Endogenous Formation of Coalitions," *Econometrica* **51**, 1047-1064.
Hart, S. and M. Kurz (1984) "Stable Coalition Structures," in *Coalitions and Collective Action* (M. Holler ed.), Physica-Verlag.
Hart, S. and A. Mas-Colell (1996) "Bargaining and Value," *Econometrica* **64**, 357-380.
Hojman, D. and A. Szeidl (2008) "Core and Periphery in Networks," *Journal of Economic Theory*, **139**, 295-309.

Hyndman, K. and D. Ray (2007) "Coalition Formation with Binding Agreements," *Review of Economic Studies* **74**, 1125-1147.
Jackson, M.O. (2003) "The Stability and Efficiency of Economic and Social Networks," in *Networks and groups: Models of Strategic Formation* (B. Dutta and M. Jackson, eds.), Springer Verlag: Heidelberg.
Jackson M. O. (2008) *Social and Economic Networks*, Princeton: Princeton University Press.
Jackson M.O. and B. Moselle (2002) "Coalition and Party Formation in a Legislative Voting Game," *Journal of Economic Theory* **103**, 49-87.
Jackson, M.O. and A. van den Nouweland (2005) "Strongly Stable Networks," *Games and Economic Behavior* **51**, 420-444.
Jackson, M.O. and A. Watts (2002a) "The Evolution of Social and Economic Networks" *Journal of Economic Theory* **106**, 265-295.
Jackson, M.O. and A. Watts (2002b) "The Existence of Pairwise Stable Networks," *Seoul Journal of Economics* **14**, 299-321.
Jackson, M. O. and A. Wolinsky (1996) "A Strategic Model of Economic and Social Networks," *Journal of Economic Theory* **71**, 44-74.
Jehiel, P. and S. Scotchmer (2001) "Constitutional Rules of Exclusion in Jurisdiction Formation," *Review of Economic Studies* **68**, 393-413.
Kamien, M. and I. Zang (1990) "The Limits of Monopolization through Acquisition," *Quarterly Journal of Economics* **90**, 465-499.
Konishi, H. and D. Ray (2003) "Coalition Formation as a Dynamics Process" *Journal of Economic Theory* **110**, 1-41.
Krishna, V. and R. Serrano (1996) "Multilateral Bargaining," *Review of Economic Studies* **63**, 61-80.
Lagunoff R. (1994) "A Simple Noncooperative Core Story," *Games and Economic Behavior* **7**, 54-61.
Levy G. (2004) "A Model of Political Parties," *Journal of Economic Theory* **115**, 250-277.
Macho-Stadler, I., D. Perez Castrillo and D. Wettstein (2007) "Efficient Bidding with Externalities", *Games and Economic Behavior* **57**, 304-320.
Maskin, E. (2003) "Bargaining, Coalitions and Externalities", Presidential Address, Econometric Society.
Merlo, A. and C. Wilson (1995) "A Stochastic Model of Sequential Bargaining with Complete Information," *Econometrica* **63**, 371-399.
Merlo, A. (1997) "Bargaining over Governments in a Stochastic Environment," *Journal of Political Economy* **105**, 101-131.
Moldovanu, B. and E. Winter (1995) "Order Independent Equilibria," *Games and Economic Behavior* **9**, 21-34.
Montero M. (1999) "Coalition Formation in Games with Externalities," mimeo., Tilburg University.

Mutuswami, S. and E. Winter (2002) "Subscription Mechanisms for Network Formation" *Journal of Economic Theory* **106**, 242-264.

Myerson, R. (1978) "Refinements of the Nash Equilibrium Concept," *International Journal of Game Theory* **15**, 133-154.

Myerson, R. (1992), *Game Theory*, Harvard University Press, Cambridge.

Norman, P. (2002) "Legislative Bargaining and Coalition Formation," *Journal of Economic Theory* **102**, 322-353.

Okada, A. (1996) "A Noncooperative Coalitional Bargaining Game with Random Proposers," *Games and Economic Behavior* **16**, 97-108.

Okada, A. (2000) "The Efficiency Principle in Noncooperative Coalitional Bargaining," *Japanese Economic Review* **51**, 34-50.

Page, F., M. Wooders and S. Kamat (2005) "Networks and Farsighted Stability," *Journal of Economic Theory* **120**, 257-269.

Page, F. and M. Wooders (2009), "Strategic Basins of attraction, the Path Dominance Core and Network Formation Games," *Games and Economic Behavior* **66**, 462-487.

Perez Castrillo, D. (1994) "Cooperative Outcomes through Noncooperative Games," *Games and Economic Behavior* **7**, 428-440.

Perez Castrillo, D. and D. Wettstein (2002) "Bidding for the Surplus: A Noncooperative Approach to the Shapley Value," *Journal of Economic Theory* **100**, 274-294.

Perry, M. and P. Reny (1994) "A Noncooperative View of Coalition Formation and the Core," *Econometrica* **62**, 795-817.

Ray, D. (1989) "Credible Coalitions and the Core," *International Journal of Game Theory* **18**, 185-187.

Ray, D. (2008) *A Game-Theoretic Perspective on Coalition Formation*, Oxford: Oxford University Press.

Ray, D. and R. Vohra (1997) "Equilibrium Binding Agreements," *Journal of Economic Theory* **73**, 30-78.

Ray, D. and R. Vohra (1999) "A Theory of Endogenous Coalition Structures," *Games and Economic Behavior* **26**, 286-336.

Ray, D. and R. Vohra (2001) "Coalitional Power and Public Goods," *Journal of Political Economy* **109**, 1355-1384.

Rogers B. (2005) "A Strategic Theory of Network Status," mimeo., Northwestern University.

Rubinstein, A. "Perfect Equilibrium in a Bargaining Game," *Econometrica* **50**, 97-109.

Selten R. (1975) "A Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory,* **4**, 25-55.

Selten, R. (1981) "A Noncooperative Model of Characteristic Function Bargaining," in *Essays in Game Theory and Mathematical Economics in*

Honor of O. Morgenstern, (V. Bohm and H. Nachtkamp eds.), Mannheim: Bibliographisches Institut Mannheim.
SEIDMANN, D. AND E. WINTER (1998) "Gradual Coalition Formation," *Review of Economic Studies* **65**, 793-815.
SEIDMANN, D., E. WINTER AND E. PAVLOV (2007) "The Formateur's Role in Government Formation," *Economic Theory* **31**, 427-445.
SERRANO, R. (2005) "Fifty Years of the Nash Program 1953-2003," *Investigaciones Economicas* **29**, 219-258.
SERRANO, R. AND R. VOHRA (1997) "Noncooperative Implementation of the Core," *Social Choice and Welfare* **14**, 513-525.
SHAPLEY, L. (1967) "On Balanced Sets and Cores", *Naval Research Logistics Quarterly* **14**, 453-460.
SLIKKER, M. (2007) "Bidding for the Surplus in Network Allocation Problems," *Journal of Economic Theory* **137**, 493-511.
SLIKKER, M. AND A. VAN DEN NOUWELAND (2001) "A One-Stage Model of Link Formation and Payoff Division," *Games and Economic Behavior* **34**, 153-175.
SUTTON, J. (1986) "Noncooperative Bargaining Theory: An Introduction," *Review of Economic Studies* **53**, 709-724.
TERCIEUX, O. AND V. VANNETELBOSCH (2006) "A Characterization of Stochastically Stable Networks," *International Journal of Game Theory* **34**, 351-369.
VON NEUMANN, J. AND O. MORGENSTERN (1944) *Theory of Games and Economic Behavior*, Princeton: Princeton University Press.
WATTS, A. (2001) "A Dynamic Model of Network Formation," *Games and Economic Behavior* **34**, 331-341.
WILSON, R. (1978), "Information, Efficiency and the Core of an Economy", *Econometrica* **46**, 807-816.
WINTER, E. (1996) "Voting and Vetoeing," *American Political Science Review* **90**, 813-823.
XUE, L. (1998) "Coalitional Stability under Perfect Foresight," *Economic Theory* **11**, 603-627.
YI, S.S. (1997) "Stable Coalition Structures with Externalities," *Games and Economic Behavior* **20**, 201-237.