



TI 2002-069/1

Tinbergen Institute Discussion Paper

# Network Formation with Heterogeneous Players

*Andrea Galeotti<sup>1,2</sup>*

*Sanjeev Goyal<sup>1,2,3</sup>*

<sup>1</sup> *Department of Econometrics, Faculty of Economics, Erasmus University Rotterdam, and Tinbergen Institute.*

<sup>2</sup> *Department of Economics, Queen Mary, University of London,  
Econometric Institute, Faculty of Economics, Erasmus University Rotterdam, and  
Tinbergen Institute.*

### **Tinbergen Institute**

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam.

### **Tinbergen Institute Amsterdam**

Keizersgracht 482  
1017 EG Amsterdam  
The Netherlands  
Tel.: +31.(0)20.5513500  
Fax: +31.(0)20.5513555

### **Tinbergen Institute Rotterdam**

Burg. Oudlaan 50  
3062 PA Rotterdam  
The Netherlands  
Tel.: +31.(0)10.4088900  
Fax: +31.(0)10.4089031

Most TI discussion papers can be downloaded at  
<http://www.tinbergen.nl>

# Network Formation with Heterogeneous Players

Andrea Galeotti\*    Sanjeev Goyal†

May 2002

## Abstract

This paper studies network formation in settings where players are heterogeneous with respect to benefits as well as the costs of forming links. Our results demonstrate that centrality, center-sponsorship and short network diameter are robust features of equilibrium networks. We find that in a society with many groups, where it is cheaper to connect within groups as compared to across groups, strategic play by individuals leads to a network architecture in which there is a core group which is entirely internally connected while all the other groups are entirely externally linked and hence completely fragmented. Since internal/within group links are cheaper to form, this implies that individual incentives may generate a significant waste of valuable social resources.

---

\*Tinbergen Institute & Econometric Institute, Erasmus University Rotterdam. E-mail: galeotti@few.eur.nl

†Department of Economics, Queen Mary, University of London & Econometric Institute, Erasmus University Rotterdam. E-mail: s.goyal@qmul.ac.uk

# 1 Introduction

This paper studies the nature of network formation when individual players differ in their cost of forming links with other players as well as in the benefits derived from other players. There is a variety of settings in which asymmetries between players are natural. For example, some players may be more productive or more informed as compared to others; likewise, some players may have high costs of linking while others have low cost of forming links. We examine the incentives of individual players to form or sever links and the architecture of the strategic stable and efficient networks in such settings.

Our point of departure is the paper by Bala and Goyal (2000). More precisely, we will be studying a model where links can be formed by individuals independently (they are one-sided), but flow of benefits is frictionless and two-sided.<sup>1</sup> Bala and Goyal paper provides a complete characterization of (strict) Nash networks when players are homogeneous. In particular, they show that when a player's payoffs are increasing in the number of other players accessed and decreasing in the number of links formed, the strict Nash network is either a center-sponsored star or the empty network. Thus partially connected networks cannot arise in equilibrium and only a very specific architecture can be sustained within the class of connected networks.<sup>2</sup> It is worthwhile to briefly sketch the argument underlying these two aspects of the result. *First*, we discuss the absence of *partial connectedness*. Suppose there are two components in an equilibrium network with one of them being non-singleton. In the non-singleton component, it must be the case that there is a player who forms a link and that this individual's returns to forming the link exceed the costs of doing so. However, for a player external to the component, the returns to forming a link with someone in the component are strictly larger (since he has access to all the members of the component and *valuations are identical*) while the *costs are the same* as that for the player internal to the component. Hence the player external to the component would like to form a link with the component and this network cannot be sustained in equilibrium. *Second*, we discuss the argument underlying the *center-sponsored star*. Suppose that a player  $i$  forms a link with a player  $j$  in the network. Then it cannot be the case that player  $j$  is directly connected with any other player. This is because if player  $j$  is linked with a player  $k$ , then player  $i$  is indifferent between linking with player  $j$  or with player  $k$  (*since costs of forming links are identical*). Hence, the network is not a strict Nash network. Thus player  $j$  must not have any links with other players; however, the network is connected, and so player  $j$  must access every player via player  $i$ . Using a variant of the above switching argument it can be shown that player  $i$

---

<sup>1</sup>This model is referred to as the two-way flow model in Bala and Goyal (2000).

<sup>2</sup>A network is connected if there is a path between every pair of agents. A network is partially connected if it is not empty and there exists at least a pair of agents which are not linked by any path. A center-sponsored star network is characterized by one agent (the center) who bears all the links with the remaining players.

must form a link with every other player. Hence player  $i$  must be the centre of a star and he must sponsor all the links in the network.

The above arguments suggest that homogeneous costs and valuations play an important role in obtaining the results in Bala and Goyal (2000). In this paper we examine the implications of relaxing the assumption of homogeneity.

We start with a general model of heterogeneous players: the costs to player  $i$  of a link with player  $j$  as well as the benefits are allowed to depend on both  $i$  and  $j$ . In addition, we assume that the length of the path does not matter in defining the benefits (there is no decay). This assumption implies that any Nash network must be acyclic or minimal. Our first result establishes an equivalence between the set of minimal and equilibrium networks: every equilibrium network is minimal and every minimal network can be sustained as a strict Nash equilibrium for some set of costs and value parameters. This result shows that individual incentives and strategic interaction generate no further restrictions apart from minimality. This result motivates a closer examination of the role of cost and value heterogeneity.

We then study a setting where costs of forming links for a particular player are the same across the links he forms, while the valuations are allowed to vary freely. In this setting, we show that an equilibrium network is either a center-sponsored star or a collection of such architectures. The converse result also holds: any such network can be supported in a Nash equilibrium for some parameters. Moreover, if the benefits of accessing a particular individual are the same for everyone but different players are differently valued by everyone then there is at most one non-singleton component and this is a center-sponsored star. A comparison of this result with the above mentioned result of Bala and Goyal (2000) suggests that heterogeneity in values is important in defining the *level of connectedness* but does not play any role in defining the (center-sponsored star) architecture of individual components.

We next study the case of cost heterogeneity and homogeneous values. We start by showing that the above mentioned correspondence between the set of minimal networks and the set of equilibrium networks still obtains. We therefore need to place some restrictions on the cost parameters of the model in order to obtain any further restrictions on the set of equilibrium networks. This leads us to consider an *insider-outsider model* where the society is formed by several distinct groups which are spatially arranged. The cost of forming a link between two players is non-decreasing in the distance between the groups to which the two players belong. Thus, the spatial distance among groups represents the degree of heterogeneity across players. In this setting, our main result is a complete characterization of strict Nash networks: a connected equilibrium network is a *generalized center-sponsored star*. This architecture has the following features: each path in the network is oriented towards a unique player, say  $i$ . Therefore,

this player plays a central role in the network itself. Furthermore, the group to which player  $i$  belongs, say  $N_i$ , is the only group to be entirely internally linked: any two players belonging to  $N_i$  are either directly linked or indirectly linked via a path that contains only members of their own group. This group represents the *core* of the network. Furthermore, all the remaining groups are entirely externally linked: members of these groups are completely split up in the network structure. From these properties it follows that the diameter of this class of architectures depends only on the number of groups composing the society and not on the number of network participants per se. Figure 1A depicts a generalized center-sponsored star in a society composed by three groups, each of them containing three players, *i.e.*  $1a$  denotes player  $a$  belonging to group 1. Each path in this network is oriented toward player  $1a$ ; hence, player  $1a$  is the center of this graph and group 1 is the core of the network. Furthermore, it is easy to note that each pair of players both belonging either to groups 2 or 3, say  $2a$  and  $2b$ , access each other through intermediary players, in this case  $1b$ , that always belong to a different group from their own. Similarly to the Bala and Goyal (2000) result where in a connected equilibrium network, *i.e.* the center-sponsored star, all players but the center are split up in the structure, we obtain that, when heterogeneity is taken into account, then each group but the core will be split up in the generalized center-sponsored network. It is possible that an equilibrium network is *partially connected*. In this case each of the components consists of members of only one group and has the center-sponsored star architecture. Figure 1B depicts a partially connected strict Nash network. Player  $1a$  bears all the links with members belonging to his own group (as represented by the filled circles on each link adjacent to this player), and the same holds for player  $2a$  and  $3a$  that are the centers of groups 2 and 3, respectively.

Our final set of results pertain to the architecture of *efficient* networks in the insider-outsider model. It is clear that an efficient network must minimize the number of outsider links since they are more costly as compared to insider links. Indeed, an efficient connected network is characterized by having each group entirely internally linked and  $m - 1$  outsider links, where  $m$  is the number of groups in the network formation game (see Figure 2). However, our characterization result above states that a (connected) equilibrium network is a generalized centre-sponsored star, with  $|N| - |N_i|$  outsider links (where  $|N_i|$  is the number of players in the core group). Thus strategic incentives of players potentially generate a significant waste of resources. This conflict disappears when the cost of forming outside links is so high to make connectivity between groups socially inefficient. When this is the case, our model becomes essentially equivalent to a set of homogeneous player models and the general result of Bala and Goyal obtains, sustainable networks are also efficient.

To summarize: our analysis illustrates the scope of the research programme which studies network formation from the perspective of individual incentives.

Our results highlight the robustness of properties such as centrality, center-sponsorship and short network diameter. Moreover, the characterization of Nash networks shows how individual incentives can generate very particular and somewhat unexpected network outcomes, such as the generalized center-sponsored star, where there is a core group which is entirely internally connected while all the other groups are entirely externally linked and hence completely fragmented. This strategic effect has high social costs and can lead to significant waste of resources.

Our paper is a contribution to the theory of network formation. This is an active area of research; earlier work includes Bala and Goyal (2000), Dutta, van den Nouweland and Tijs (1995), Jackson and Wolinsky (1996) and Kranton and Minehart (2001), among others. These papers and indeed most of the existing literature focuses on symmetric settings. We briefly discuss three other papers which examine the role of heterogeneity in network formation. McBride (2001) considers a non-cooperative network formation game in a framework characterized by imperfect information about valuations of players. Haller and Sarangi (2000) study the implications of heterogeneous reliability levels across players. Johnson and Gilles (2001) investigate the effects of heterogeneous cost of linking across players in a two-sided network formation game. The main difference between their paper and the present paper is the nature of link formation. There are also some other differences in terms of the model specification, such as the role of decay and the insider-outsider formulation that we use. These differences lead to very different results. More precisely, the authors show that if cost of linking is low in respect to the potential benefit, locally complete networks where a connected agent is always connected to at least one of his direct neighbors and belong to a complete subnetwork, are more likely to arise in equilibrium. This is in sharp contrast to our findings both in the context of the generalized heterogeneity model as well as in the context of the insider-outsider model.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 presents results on network formation when costs and values are allowed to vary freely. Section 4 analyzes an insider-outsider model of network formation. Section 5 concludes.

## 2 The Model

Let  $N = \{1, \dots, n\}$  be a set of players and let  $i$  and  $j$  be typical members of this set. To avoid trivialities, we shall assume throughout that  $n \geq 3$ . Each player is assumed to possess some information of value to himself and to other players. He can augment his information by communicating with other people; this communication takes resources, time and effort and is made possible via *pair-wise* links.

A strategy of player  $i \in N$  is a (row) vector  $g_i = (g_{i,1}, \dots, g_{i,i-1}, g_{i,i+1}, \dots, g_{i,n})$  where  $g_{i,j} \in \{0, 1\}$  for each  $j \in N \setminus \{i\}$ . We say that player  $i$  has a link with  $j$  if  $g_{i,j} = 1$ . A link between player  $i$  and  $j$  can allow for either one-way (asymmetric) or two-way (symmetric) flow of information. We assume throughout the paper that a link  $g_{i,j} = 1$  allows both players to access each other's information. The set of strategies of player  $i$  is denoted by  $G_i$ . Throughout the paper we restrict our attention to pure strategies. Since player  $i$  has the option of forming or not forming a link with each player of the remaining  $n - 1$  players, the number of strategies of player  $i$  is clearly  $|G_i| = 2^{n-1}$ . The set  $G = G_1 \times \dots \times G_n$  is the space of pure strategies of all the players.

A strategy profile  $g = (g_1, \dots, g_n)$  can be represented as a direct network. Let  $g \in G$ . To describe information flows, it is useful to define the closure of  $g$ : this is a non-directed network denoted  $\bar{g} = \text{cl}(g)$ , and define by  $\bar{g}_{i,j} = \max \{g_{i,j}, g_{j,i}\}$  for each  $i$  and  $j$  in  $N$ .<sup>3</sup> Pictorially, the closure of a network simply means replacing every directed edge of  $g$  by a non-directed one. We say there is a path in  $g$  between  $i$  and  $j$  if either  $\bar{g}_{i,j} = 1$  or there exist players  $j_1, \dots, j_m$  distinct from each other and  $i$  and  $j$  such that  $\{\bar{g}_{i,j_1} = \dots = \bar{g}_{j_m,j} = 1\}$ . We write  $i \xleftrightarrow{\bar{g}} j$  to indicate a path between  $i$  and  $j$  in  $g$ . Furthermore, a path between  $i$  and  $j$  is said to be  $i$ -oriented if either  $g_{i,j} = 1$  or there is a sequence of distinct players  $i_1, i_2, \dots, i_n$  with the property that:  $\{g_{i,i_1} = g_{i_1,i_2} = 1, \dots, g_{i_n,j} = 1\}$ . Define  $N^d(i; g) = \{k \in N \mid g_{i,k} = 1\}$  as the set of players with whom  $i$  maintains a link and let  $\mu_i^d(g) = |N^d(i; g)|$  be the cardinality of the set. The set  $N(i; \bar{g}) = \{k \in N \mid i \xleftrightarrow{\bar{g}} k\} \cup \{i\}$  consists of players that  $i$  observes in  $g$ , while  $\mu_i(g) = |N(i; \bar{g})|$  is its cardinality. To complete the definition of a normal-form game of network formation, we specify the payoffs. Let  $V_{i,j}$  denote the benefits that player  $i$  derives from accessing player  $j$ . Similarly, let  $c_{i,j}$  denote the cost for player  $i$  of forming a link with player  $j$ . The payoff function can now be stated as follows:

$$\Pi_i(g) = \sum_{j \in N(i; \bar{g})} V_{i,j} - \sum_{j \in N^d(i; g)} c_{i,j} \quad (1)$$

Throughout we shall assume that  $c_{i,j} > 0$  and  $V_{i,j} > 0$  for all  $i, j$  in  $N$ .

Given a network  $g \in G$ , let  $g_{-i}$  denote the network obtained when all of player  $i$ 's links are removed. Note that the network  $g_{-i}$  can be regarded as the strategy profile where  $i$  chooses not to form a link with anyone. The network  $g$  can be written as  $g = g_i \otimes g_{-i}$  where the  $\otimes$  indicates that  $g$  is formed as the union of the links in  $g_i$  and  $g_{-i}$ . The strategy  $g_i$  is said to be a *best response* of player  $i$  to  $g_{-i}$  if:

$$\Pi_i(g_i \otimes g_{-i}) \geq \Pi_i(g'_i \otimes g_{-i}) \text{ for all } g'_i \in G_i. \quad (2)$$

---

<sup>3</sup>Note that  $\bar{g}_{i,j} = \bar{g}_{j,i}$  so that the order of players is irrelevant.



The set of all of player  $i$ 's best responses to  $g_{-i}$  is denoted by  $BR_i(g_{-i})$ . Furthermore, a network  $g = (g_1, \dots, g_n)$  is said to be a *Nash network* if  $g_i \in BR_i(g_{-i})$  for each  $i$ , i.e. players are playing a Nash equilibrium. A *strict Nash network* is one where each player gets a strictly higher payoff with his current strategy than he would with any other strategy.

Finally, in order to analyze the efficient architectures we need to introduce a welfare measure. As in Bala and Goyal we define the social welfare of a network  $g$  as the sum of payoffs of all players. Formally, given a network  $g$ , its welfare,  $W : G \rightarrow R$ , can be stated as follows:

$$W(g) = \sum_{i=1}^n \Pi_i(g) \text{ for } g \in G. \quad (3)$$

A network is said to be efficient if  $W(g) \geq W(g')$  for any  $g' \in G$ . Hence, an efficient architecture can be seen as the one that minimize the cost of providing a certain amount of information to the players.

### 3 General Heterogeneity

We begin our analysis with some results that outline the scope of our study. In our analysis, we shall use the idea of minimal networks. Given a network  $g$ , we define a component of  $g$ ,  $C(g)$ , a set  $C(g) \subset N$  such that  $\forall i, j \in C(g)$  there exists a path between them and there does not exist a path between  $\forall i \in C(g)$  and an player  $k \in N \setminus C(g)$ . Given a network  $g$ , let  $\#C(g)$  be the number of components in  $g$ . A network  $g$  is said to be minimal if  $\#C(g) < \#C(g - g_{i,j})$ ,  $\forall i \neq j$ . Moreover a network  $g$  is said to be connected if it is composed by only one component, i.e.  $\#C(g) = 1$ . If this component is minimal, then  $g$  is said to be minimal connected. It follows that each link in a minimal connected network is critical in the way that it is enough to delete it, ceteris paribus, to induce some degree of social isolation in the society.<sup>4</sup>

Our first result shows a correspondence between the set of minimal networks and the set of strict Nash equilibrium networks. Here, we start with assuming value homogeneity, i.e.  $V_{i,j} = V$ ,  $\forall i, j \in N$ .

**Proposition 3.1:** *Let the payoffs satisfy (1). Then a strict Nash network is minimal. Given any minimal network  $g$  there exist costs and benefits,  $\{c_{i,j}, V\}$ , such that this network is a strict Nash network.*

---

<sup>4</sup>Social isolation stands for a situation where the social architecture does not allow each player to observe the whole society.

**Proof:** We first show that an equilibrium network is minimal. Let  $g$  be a Nash network, and suppose that it is not minimal. Then there is a link  $g_{i,j} = 1$  such that  $N(i; g) = N(i; g - g_{i,j})$ , for all  $i \in N$ . Given the specification of payoffs, and the assumption that  $c_{i,j} > 0$ , for all  $i, j$ , player  $i$  can strictly increase his payoff by deleting the link. This violates the assumption that  $g$  is Nash.

We now prove the converse. Fix some minimal network  $g$ . For any link  $g_{i,j} = 1$ , set the corresponding cost  $c_{i,j} = \epsilon < V$ , while for any link  $g_{i,j} = 0$ , set the corresponding  $c_{i,j} > (n - 1)V$ . The optimality of forming the existing links follows from the cost restrictions and the fact that the network  $g$  is minimal. The optimality of not forming the link follows directly from the assumption on the costs.  $\square$

Proposition 3.1 motivates an examination of conditions under which we can derive some restrictions on the strict Nash networks, apart from minimality. We note that the second part of the result shows that any minimal network can be sustained in a setting where values are homogeneous. A comparison of this result with the earlier results of Bala and Goyal for the homogeneous players setting suggests that costs homogeneity plays a crucial role in restricting the architecture of equilibrium networks. To get a clearer idea of this issue, we now analyze the case where the cost of link formation is homogeneous across links for any particular individual i.e.  $c_{i,j} = c_i$  for all  $j \in N \setminus \{i\}$ , but it may vary across individuals,  $c_i \neq c_j$  is allowed. In addition we allow values to vary freely.

**Proposition 3.2:** *Suppose that for each  $i \in N$ ,  $c_{i,j} = c_i$ , for every  $j \in N \setminus \{i\}$ . Let  $g$  be a strict Nash network and suppose that  $C(g)$  is a component in  $g$ , with  $|C(g)| \geq 3$ . Then  $C(g)$  is a center-sponsored star. Let  $g$  be a minimal network in which every component with 3 or more players is a center-sponsored star. Then there exist costs and benefits  $\{c_i, V_{i,j}\}$ , such that this network is strict Nash.*

**Proof:** We start with the first part. Let players  $i, j$  and  $k$  belong to  $C(g)$ . Suppose  $g_{i,j} = 1$ . We claim that player  $j$  cannot have any other link. Suppose not and let  $\bar{g}_{j,k} = 1$ . Since  $c_{i,j} = c_{i,k}$  it then follows from the payoffs (1) that player  $i$  is indifferent between forming a link with players  $j$  or  $k$ . This contradicts strictness of equilibrium. Since  $C(g)$  is a component, it follows that player  $j$  accesses everyone in  $C(g)$  via the link  $g_{i,j}$ . By analogous reasoning we can infer that no player  $k$  forms a link with player  $i$ . Hence, player  $i$  must form all links and must be the center of a star. This implies that the component must be a center-sponsored star.

We prove the second part now. Fix some minimal network  $g$  with the said properties. Let there be  $m$  components in this network,  $C_1(g), \dots, C_m(g)$ . Fix some player  $i$  and without loss of generality, let  $i \in C_1(g)$ . For any link  $g_{i,j} = 1$ ,

set the corresponding returns  $c_{i,j} = c_i < V_{i,j}$ , while for every component  $C_k(g)$ ,  $k = 2, \dots, m$ , and any player  $j \in C_k(g)$ , let  $\sum_{j \in C_k(g)} V_{i,j} < c_i$ . The optimality of forming the existing links follows from the cost restrictions and the fact that the network  $g$  is minimal. The optimality of not forming the link follows directly from the assumption on the costs and benefits. Since  $i$  was arbitrary, the proof follows.  $\square$

This result shows that the center-sponsored star architecture plays a prominent role even in the presence of heterogeneous values and differences in cost of forming links across players. A comparison of this result with the earlier result, Proposition 3.1, also suggests that the assumption  $c_{i,j} = c_i, \forall j$ , plays a critical role in the analysis. In section 4 we shall explore further the role of heterogeneity in costs of forming links.

The above proposition shows that value and cost heterogeneity permit the existence of more than one non-singleton component in an equilibrium. We now show that a slightly stronger restriction on valuations,  $V_{j,i} = V_i$ , for all  $j$ , implies that this is no longer possible in equilibrium.

**Proposition 3.3:** *Suppose that for any player  $i$ ,  $c_{i,j} = c_i$ , and  $V_{j,i} = V_i$  for all  $j \in N$ . Then every strict Nash network is minimal and has at most one non-singleton component (which is a center sponsored star). Moreover, any minimal network with these properties is sustainable as a strict Nash network for some value of  $\{c_i, V_i\}$ .*

**Proof:** Suppose that  $g$  is a strict Nash network and  $C_1(g)$  and  $C_2(g)$  are two non-singleton components. From earlier results we know that the network is minimal and that in each of the two components only one player forms all the links. Let  $i \in C_1(g)$  and let  $j \in C_2(g)$  be these players. Since  $g$  is strict Nash network, it follows that  $\Pi_i(g) > \Pi_i(g - g_{i,k})$ , with player  $k \in C_1(g)$ . In other words,

$$V_s - c_i > 0 \quad \forall s \in C_1(g). \quad (4)$$

However, the net payoff to player  $j$  from forming a link with player  $i$  is given by:

$$\sum_{s \in C_1(g)} V_s - c_j < 0 \quad (5)$$

Since player  $j$  does not form this link. It must be the case that  $c_i < c_j$ . Likewise, we can now reason that:

$$V_t - c_j > 0, \quad \forall t \in C_2(g) \quad (6)$$

However, the net payoff to player  $i$  from forming a link with player  $j$  is given by:

$$\sum_{t \in C_2(g)} V_t - c_i < 0 \quad (7)$$

Since player  $i$  does not form this link, it must be the case that  $c_i > c_j$ . This leads to a contradiction. Hence, there can be at most one non-singleton component in a strict Nash network.

We now prove that every such network can be sustained as a strict Nash network. The proof is by construction. Take any such network: minimal, with one non-singleton component, where only one player forms all the links. Suppose that  $C_1(g)$  is the non-singleton component, and that player  $i \in C_1(g)$  forms all the links. Then this network is strict Nash for the following cost/value parameters: (i)  $c_i < V_j, \forall j \in C_1(g)$ ;  $c_i > V_k, \forall k \notin C_1(g)$ ; (ii)  $c_k > \sum_{l \neq k} V_l$ .  $\square$

The results in this section demonstrate several points. *First*, that if we allow for values and costs to vary freely then the only restriction imposed by the equilibrium requirement is minimality. Minimality is a direct consequence of the assumption that the distance between players does not affect the transmission of value. *Second*, if we allow for value heterogeneity but only a moderate amount of cost heterogeneity, then equilibrium has considerable bite. In particular, if any individual  $i$ 's costs of forming links are the same, i.e. if  $c_{i,j} = c_i$ , for all  $j \in N \setminus \{i\}$ , then any equilibrium network is either a center-sponsored star or comprises of smaller center-sponsored stars. Thus, the results demonstrate that the center-sponsored stars continue to be prominent even in settings with considerable value and cost heterogeneity across individuals as long as a player's costs of forming links is independent of the identity of the player being connected to. However, little can be said about equilibrium networks if costs of link formation differ across links for the same individual. The next section presents a model with heterogeneous costs where individual incentives do have strong implications.

## 4 An insider-outsider model

We consider a society composed by  $m$  groups. We define the set of players belonging to group  $l$  as  $N_l, l = 1, \dots, m$ . Let  $|N_l|$ , be the size of group  $l$ , with  $l = 1, 2, 3, \dots, m$ . The set of players is then  $N \equiv \cup_{l=1}^m N_l$ . We assume perfect symmetry in value across individuals and we normalize it to one, i.e.  $V_{i,j} = 1$  for all  $i, j \in N$ . To allow for cost heterogeneity we consider a spatial cost structure: groups can be ordered in a line according to some well defined characteristics. Hence, the distance between two groups can be interpreted as a measure of the heterogeneity that distinguishes them. Given two players  $i \in N_l$  and  $j \in N_k$ ,

the cost of forming a link  $g_{i,j}$ , is:

$$c_{i,j} = c_{j,i} = f(|l - k|) \quad (8)$$

If  $i$  and  $j$  belong to the same group we let:

$$c_{i,j} = c_{j,i} = f(0) = c_L \quad (9)$$

We shall assume that  $f(\cdot)$  is non-decreasing in its argument and  $c_L > 0$ . Let  $N^{d,k}(i;g) = \{j \in N_k | g_{i,j} = 1\}$ , for  $k = 1, \dots, m$ ; then define  $N^d(i;g) \equiv \cup_{k=1}^m N^{d,k}(i;g)$ . Furthermore, let  $\mu_i^{d,k}(g)$  be the cardinality of  $N^{d,k}(i;g)$ . In other words,  $\mu_i^{d,k}(g)$  represents the number of links initiated by  $i$  with members of group  $k$ . Hence, given a network  $g$  and an player  $i \in N_l$ , the payoff function described by (1) can be rewritten as follows:

$$\Pi_i(g) = \mu_i(g) - \sum_{k=1}^m \mu_i^{d,k}(g) f(|l - k|) \quad (10)$$

We note two interesting special cases of our specification.

**1. Homogeneous Players:** This case arises when  $f(0) = f(1) = \dots = f(m-1) = c$ . This implies that player  $i$ 's payoff is the number of players he observes less the total cost of link formation. Clearly, the distinction between inside and outside links becomes irrelevant and we can consider that the whole society is composed of one group. In this case, the payoff is given by:

$$\Pi_i(g) = \mu_i(g) - \mu_i^d(g) c \quad (11)$$

**2. Two-cost levels:** The case of two-cost levels arises when we assume that  $f(d) = c_H, \forall d \geq 2$ , and  $f(0) = c_L < c_H$ . We can then write the cost structure as follows:

$$c_{i,j} = \begin{cases} c_L, & \text{if } i, j \in N_l \\ c_H, & \text{if } i \in N_l \text{ and } j \in N_k, l \neq k \end{cases} \quad (12)$$

In words, the cost of creating a link across groups (outside link),  $c_H$ , is equal or higher than the cost of creating a link within a group,  $c_L$  (inside link). However, links formed with different external groups are equally costly. The two-cost levels case will be discussed below to illustrate some of our results.

We now develop some additional notation. We say that  $i, j \in N_l$  are entirely internally linked if either  $\bar{g}_{i,j} = 1$  or there is a path between  $i$  and  $j$  in which all players belong to  $N_l$ . We say that  $i, j \in N_l$  are externally linked if they are linked but not entirely internally linked. Moreover, let the diameter of a non-singleton component,  $C(g)$ , be defined as the length of the largest geodesic distance between any pair of players belonging to it, *i.e.*

$D(C(g)) = \max_{i,j \in C(g)} d(i, j; C(g))$ .<sup>5</sup> To complete the notation, in the characterization of strict Nash networks we will deal with the class of generalized center-sponsored star architectures. More specifically, a generalized center-sponsored star architecture is a minimally connected network which satisfies the following conditions: (A)  $\exists l$  and  $\exists i \in N_l$  such that  $g_{i,j} = 1, \forall j \in N_l \setminus \{i\}$ ; (B) Let  $i \in N_l$  be the player identified in (A), then for any  $j \in N, i \xrightarrow{\bar{g}} j$ , is an  $i$ -oriented path; (C) consider an  $i$ -oriented path. Without loss of generality, numbered the players as  $i, i_1, i_2, \dots, i_n$  with  $\{g_{i,i_1} = g_{i_1,i_2} = \dots = g_{i_{n-1},i_n} = 1\}$ . Let  $i_k \in N_{l_k}$ , then  $f(|l_k - l_{k+1}|) < f(|l_k - l_x|)$  for  $x \in \{k+2, k+3, \dots, n\}$ ; (D)  $D(g) \leq 2m$ .

## 4.1 Equilibrium Networks

Our first result describes Nash equilibrium networks.

**Proposition 4.1:** *Let the cost structure be given by (8) and let the payoffs be given by (10). A Nash network is minimal. In particular, depending on the cost levels, a Nash network can be empty, connected or partially connected.*

The proof is given in the appendix; we briefly outline the intuition here. *First*, we note that a Nash network is minimal; this follows from the no-decay assumption. *Next*, we note that in the presence of heterogeneity in cost levels, partially connected networks (with each component being composed of members belonging to the same group) can be sustained in equilibrium. There are two reasons for this: a coordination problem and an incentive problem. The coordination problem arises as follows: suppose the inside cost is slightly higher than 1 and the outside cost,  $f(1)$ , is a bit higher than the size of the smallest group, say  $N_l$ . Consider a network where the smallest group is internally linked and all other players are singletons. It is easy to see that such a network can be sustained as a Nash equilibrium. However, if the largest group is able to coordinate by generating a minimal connected component, it will create the right incentives to achieve a connected network, and therefore a partially connected structure is not sustainable anymore. The incentive problem arises in the following way: suppose that the inside cost is low,  $c_L \in (0, 1)$ . Then members of each group will have an incentive to be internally linked. If  $f(1)$  is sufficiently high then a player has no incentive to link up with a component comprising of members of any other single group. In this situation a partially connected network with  $m$  minimal components is a Nash equilibrium. The above result suggests that a wide range of networks can arise in a Nash equilibrium. Earlier work suggests

---

<sup>5</sup>Given two players  $i$  and  $j$  in  $g$ , the geodesic distance,  $d(i, j; g)$ , is defined as the length of the shortest path between them.

that strictness has considerable bite in homogeneous player settings. Is this also true in a setting with heterogeneous players? Our next result provides a complete response to this question.

**Proposition 4.2:** *Assume that  $|N_l| \geq 2, \forall l = 1, \dots, m$ . Let the cost structure be given by (8) and let the payoffs be given by (10):*

- 1) *If  $c_L > 1$  then the only strict Nash network is the empty network.*
- 2) *Suppose  $c_L \in (0, 1)$ , then there are three cases: 2a) if  $f(1) \in (c_L, 1)$ , then a strict Nash network is a generalized center-sponsored star. 2b) If  $f(1) \in (1, \max[N_1, \dots, N_m])$ , then a strict Nash network does not exist. 2c) If  $f(1) > \max[N_1, \dots, N_m]$ , then the only strict Nash network is partially connected with each group constituting a center-sponsored star.*

The *first* step of the proof consists of showing that in each non-singleton component there exists one group that is entirely internally linked. We start proving that in each non-singleton component there exists at least one inside link. Suppose  $g$  is a strict Nash network. For simplicity assume that it is connected; then there is a path between any two players belonging to the same group, say  $i, i' \in N_l$ . There are two possible path configurations. First, the two players are directly linked and if this is the case the claim follows. Second, the two players access each other indirectly, through other players. In this case, it can be shown by an application of the switching argument that this path has to have the following pattern of links:  $\{g_{j,i} = 1, \dots, g_{j',i'} = 1\}$ . Next we note that the same property must also hold for any other pair of players belonging to the same group. Since the number of groups is finite and each group is composed by at least two players, we can iterate this argument to show that there always exist two players belonging to the same group who access each other via a direct connection. We then use network externality effects to argue that if two players of a group are directly linked then all members of this group must belong to the same component. Finally, we use the switching argument to show that given an inside link, i.e.  $g_{i,i'} = 1$  with  $i, i' \in N_l$ ,  $i$  will bear all the links with members of his own group. Hence, group  $N_l$  is entirely internally linked.

The *second* step in the proof shows that if a group is not entirely internally linked then it is entirely externally linked. Consider a connected strict Nash network. Let  $N_l$  be the group highlighted in the previous step and let  $i$  be the center of this group. Consider a path between  $i$  and an arbitrary player  $j$ . Using a variant of the argument sketched above, we show that the path must be  *$i$ -oriented*. Now it is easy to see that along any path leading away from player  $i$ , there can be at most one player of any specific group. Hence it follows that if we take a pair of players in a group  $l' \neq l$  there exists a path (since  $g$  is connected) and along this path there is no player of group  $l'$ . Thus all groups apart from  $l$  are

entirely externally linked. The *final* step in the proof consists of combining the above observations for different cost parameters.

We discuss some aspects of this characterization result. The *first* remark is about insider and outsider links. Our result shows that there is one group which is entirely internally linked in the connected strict Nash network, while all other groups are entirely externally linked. In other words, the formation of local connections is not allowed in equilibrium (except for one group). This is an unexpected result and it suggests that incentives for link formation completely undermine the structure that one might have expected: a set of local center-sponsored stars (corresponding to individual groups) linked with each other. The *second* observation concerns the centrality and center-sponsorship properties. If the strict Nash network is connected, there is a player  $i$  such that all paths are oriented toward him. Hence, this player plays a particularly central role in the network. Furthermore, if the strict Nash network is non-empty but unconnected, then each component consists of members of one group and it has the center-sponsored star structure. *Third*, it is worth noting that the diameter of connected strict Nash network is independent of the number of players, while it only depends on the number of groups. *Fourth*, we consider the two special cases introduced in the specification of the insider-outsider model. When applying Proposition 4.2 to the homogeneous case we obtain the result provided by Bala and Goyal (2000): if  $c > 1$  the only strict Nash network is the empty one, while if  $c \in (0, 1)$  then the only strict Nash network is a center-sponsored star. Let's now turn to the two-cost levels case. When  $c_H \in (c_L, 1)$ , a strict Nash network has a generalized center-sponsored star architecture. More formally, there is an individual, say  $i \in N_l$  which is the center of the whole network: each path in the network is oriented to him. Furthermore, group  $N_l$  is the only group to be entirely internally linked. Moreover, the members of all the remaining groups are passively linked with some members belonging to group  $N_l$ .<sup>6</sup> In particular, if all the remaining players are passively linked with player  $i$ , then the network is a center-sponsored star (see Figure 3).

Finally, we remark on the assumption that there are at least two members in each group. If we relax this assumption and allow for some groups to have only one member then two substantial changes occur. The first change is that there may exist more than one entirely internally linked group while the second change is that the non-existence result may be ruled out. The following example illustrates these points. Consider a society composed by three groups, where group  $N_1$  and  $N_3$  consist of two players and group  $N_2$  has only one player. Let  $g$  be a connected network depicted in Figure 4. When  $f(1) \in (c_L, 1)$ ,  $g$  is strict Nash. We note that in  $g$  all groups are entirely internally linked. Now, suppose that  $f(1) = 1 + \epsilon$ , where  $\epsilon$  is positive and small enough. Again, the network  $g$

---

<sup>6</sup>We say that an agent  $i$  is passively linked with an agent  $j$  if  $g_{j,i} = 1$  and  $g_{i,j} = 0$ .



is strict Nash. However, if we assume that group  $N_2$  consists of more than one player, a standard switching argument leads to the non-existence result.

## 4.2 Efficient Networks

We now turn into the issue of efficiency. We first introduce some new terminology that will be used in the proposition below. Let  $g^{mc}$  refer to a minimal connected network with each group  $N_l$  forming a minimal connected component with  $|N_l| - 1$  inside links respectively and with  $(m - 1)$  outside links of distance one (see Figure 2). A network  $g$  is said to be partially connected,  $g^{pc}$ , if there exists at least one non-singleton component but the network is not connected. Finally, a partially connected network with each group generating a minimal connected component will be denoted as  $g_m^{pc}$  (see Figure 1B).

In the further analysis we assume that groups are of equal size. We develop a simple example in order to illustrate how relaxing this assumption can lead to a variety of efficient networks. For sake of clearness, we consider the two-cost levels case. Let the society be composed by three groups where group  $N_1$  is small while groups  $N_2$  and  $N_3$  are large. Suppose now that  $c_L \in (0, 1)$  and  $c_H < 2|N_2||N_3|$ , then an efficient network must have the three groups internally linked and group  $N_2$  and  $N_3$  connected by one outside link. However, if  $c_H \in (2|N_1|(|N_2| + |N_3|), 2|N_2||N_3|)$  then it is socially efficient to leave group  $N_1$  isolated. Therefore, the efficient network is one in which the three groups are linked internally and where group  $N_2$  and  $N_3$  are connected by one outside link while group  $N_1$  is left out. Clearly, if  $c_L \in (0, 1)$  and  $c_H < 2|N_1|(|N_2| + |N_3|)$  the efficient network is minimal connected with  $m - 1$  outside links, while if  $c_H > 2|N_2||N_3|$  then only a partially connected network where the three groups are linked internally is efficient. Finally for  $c_L > \max\{|N_1|, |N_2|, |N_3|\}$  and  $c_H$  sufficiently high the only efficient network is the empty one.

The following result provides a complete characterization of efficient networks for the case of equal group sizes. Let  $|N_l| = |\bar{N}|$  for all  $l = 1, 2, \dots, m$ ; moreover, we define  $c_1 = m|\bar{N}|^2$  and  $c_2 = [m|\bar{N}|(m|\bar{N}| - 1) - (m|\bar{N}| - m)c_L]/(m - 1)$ .

**Proposition 4.3:** *Let the cost structure be given by (8) and let the payoffs be given by (10). In addition suppose that  $N_l = \bar{N}$  for all  $l = 1, 2, \dots, m$ .*

- 1) *Suppose  $c_L \in (0, |\bar{N}|)$ . If  $f(1) \in (c_L, c_1)$  the network  $g^{mc}$  is uniquely efficient, while if  $f(1) > c_1$  then the network  $g_m^{pc}$  is uniquely efficient.*
- 2) *Suppose  $c_L \in (|\bar{N}|, m|\bar{N}|)$ . If  $f(1) \in (c_L, c_2)$  then the network  $g^{mc}$  is uniquely efficient, while if  $f(1) > c_2$  then the empty network is uniquely efficient.*
- 3) *If  $c_L > m|\bar{N}|$  then the empty network is uniquely efficient.*

The proof is presented in the appendix. We briefly sketch the intuition here. An efficient network is minimal; this follows from the no-decay assumption. When  $c_L$  is high enough the empty network is efficient, while if  $c_L$  is relatively low it is beneficial for the society to have each group internally linked. Considerations on  $f(1)$  allow us to divide this cost space into two sub-spaces: for  $f(1)$  high enough the society is better-off leaving each group isolated by the others, yielding the network  $g_m^{pc}$ , while if  $f(1)$  is not so high then the connected network arises. However, only a connected network with a minimal number of outside links ( $m - 1$ ) and all of ‘length’ one, is efficient. This yields us  $g^{mc}$ .<sup>7</sup>

The above characterization of efficient networks allows us to make some remarks on the trade-off between efficient and sustainable architectures.<sup>8</sup> We have showed that if  $g^{mc}$  is efficient, the corresponding set of strict Nash networks does not contain any architectures compatible with the efficient one. This conflict persists until the level of  $f(1)$  is such that any outside link is not beneficial both from an individual and social point of view. When this is the case, the heterogeneity introduced in the model becomes irrelevant and our problem degenerates in a sum of independent homogeneous problems leading to partially connected strict Nash network with each group generating a center-sponsored star component. It follows that the trade-off between efficiency and stability fades in this case.

The conflict between efficient and sustainable connected architectures arises out of a misallocation of links: too many outside links are set-up in order to obtain connectedness. Consider a connected network  $g$  and pick two players belonging to a group different from the core group, then if  $g$  is strict Nash, they will access each other via a sequence of outside links. This does not allow network participants to minimize the costs of connecting with each other and this lowers social welfare.

This result is altered if we relax the assumption, used in the characterization of strict Nash networks, that each group is composed of at least two players. Consider a society composed by three groups where groups  $N_1$  and  $N_3$  consist of two individuals each while group  $N_2$  consists of a single individual. Suppose  $f(1) \in (c_L, 1)$ . The network depicted in Figure 4 is strict Nash. Moreover, this network satisfies all the necessary conditions for a connected network to be efficient: the allocation of links is optimal from a societal point of view. In general, the presence of a single player between two heterogeneous groups composed by at least two individuals mitigates substantially the conflict between the notion of efficiency and strategic stability.

---

<sup>7</sup>Each minimal connected network produces the same gross social welfare but different minimal connected networks will be characterized by a different total cost depending on the allocation of links.

<sup>8</sup>It is worth noting that the characterization of strict Nash networks presented in Proposition 4.2 is valid if  $|N_l| \geq 2$ , for all  $l = 1, 2, \dots, m$ . This allows us to compare the set of strict Nash networks with the set of efficient ones.

## 5 Conclusion

In this paper we have investigated the implications of heterogeneity for network formation. We have used an extension of the Bala and Goyal (2000) model of one-sided link formation to study this issue. Our analysis suggests two general observations: *first*, by showing the robustness of equilibrium properties such as centrality, center-sponsorship and short network diameter this paper illustrates the scope of the research programme which seeks to understand social and economic structure in terms of individual incentives. Moreover, our characterization of equilibrium networks shows that individual incentives lead to very particular and somewhat unexpected networks, in which there is a core group which is entirely internally connected while all the other groups are entirely externally linked and hence completely fragmented. This finding leads to our *second* point: individual incentives can generate a significant waste of valuable social resources.

## 6 Appendix:

**Proof of Proposition 4.1** Minimality follows as a direct consequence of the no-decay assumption. The proof that the empty network is Nash if and only if  $c_L \geq 1$  is straightforward and omitted. Next, we argue that if  $c_L < 1$  and  $f(1) < \max\{|N_1|, \dots, |N_m|\}$ , then a Nash network is connected. Since  $c_L < 1$  there must exist a path between any pair of players in the same group  $N_l$ . Let, without loss of generality,  $m$  be the largest group. Since  $f(1) < |N_m|$ , it follows that players in all the groups other than  $N_m$  have an incentive to form a link with group  $N_m$ . Hence, a Nash network must involve a path between any two players  $i \in N_l$  and  $j \in N_m$ , i.e. the network is connected.

Finally we note that if  $c_L < 1$  but  $f(1) > |N_m|$  then no player  $i \in N_l$  has an incentive to form a link with  $j \in N_{l'}$ ,  $l \neq l'$ , so long as  $N_{l'}$  constitutes a component by itself. Hence, it is possible to sustain a network with  $m$  components, each component consisting of members of one and only one group, respectively, in equilibrium. Next, we argue that if  $g$  is Nash and there exists a link  $g_{i,j} = 1$ , where  $i \in N_l$  and  $j \in N_{l'}$ , with  $l \neq l'$ , then  $g$  is connected. If  $g_{i,j} = 1$ , then it must be the case that the returns to player  $i$  from the link  $g_{i,j}$ , (say)  $|C(g - g_{i,j})|$ , exceed the cost of the link,  $c_{i,j}$ . Suppose that  $g$  is not connected. Let  $j \in N \setminus C(g)$  and suppose that  $j$  belongs to a group that is closest to the groups represented in  $C(g)$ . It then follows that if  $j \in N_x$ , then there is a player  $j' \in N_y$  with  $j' \in C(g)$  and  $|x - y| \leq 1$ . It is easy to see that the payoff to player  $j$  from linking with player  $j'$  is  $|C(g)| > |C(g - g_{i,j})|$ , while the cost is (weakly) smaller than the cost to player  $i$  of the link  $g_{i,j}$ . Hence, player  $j$  has a strict incentive to form a link with some player in  $C(g)$ , contradicting the hypothesis that  $g$  is a Nash network.  $\square$

**Proof of Proposition 4.2:** We recall some definitions that will be used in the proof. In a network  $g$ , a path between  $i$  and  $j$  is said to be  $i$ -oriented if either  $g_{i,j} = 1$  or there is a sequence of distinct players  $\{i_1, i_2, \dots, i_n\}$  with the property that:  $\{g_{i,i_1} = g_{i_1,i_2} = 1, \dots, g_{i_n,j} = 1\}$ . The proof consists of a sequence of steps, which are covered in the following lemmas.

**Lemma 1** *Suppose  $g$  is a strict Nash network. If  $g_{i,j} = 1$ , where  $i \in N_l$  and  $j \in N_{l'}$ ,  $l \neq l'$ , then  $\bar{g}_{j,j'} = 0$ ,  $\forall j' \in N_k$  where  $k$  is such that  $|l - k| \leq |l - l'|$ .*

**Proof:** Consider a strict Nash network  $g$ . Choose  $i \in N_l$  and  $j \in N_{l'}$ ,  $l \neq l'$ , such that  $g_{i,j} = 1$ . Let  $j' \in N_k$  where  $k$  is such that  $|l - k| \leq |l - l'|$ . Suppose that  $\bar{g}_{j,j'} = 1$ . The spatial cost structure implies that  $i$  can do at least as well by deleting his link with  $j$  and forming a link with  $j'$ . This contradicts strict Nash.  $\square$

**Lemma 2** *Suppose  $g$  is a strict Nash network. If  $i \in N_l$  and  $j \in N_{l'}$ ,  $l \neq l'$ , and  $g_{i,j} = 1$ , then  $g_{j',i} = 0$ ,  $\forall j' \in N_k$  such that  $|k - l| \geq |k - l'|$ .*

**Proof:** Suppose  $g_{j',i} = 1$ . Since the cost of forming links is non-decreasing in the distance between players' groups,  $j'$  can do at least as well by deleting his link with  $i$  and forming a link with  $j$ . This contradicts strict Nash.  $\square$

**Lemma 3** *Assume  $|N_l| \geq 2$ ,  $\forall l = 1, \dots, m$ . Suppose  $g$  is a strict Nash network, then in any non-singleton component there exists a pair of players who belong to the same group (this group will differ across components) and have a direct link.*

**Proof(Sketch):** Consider a non-singleton component  $C(g)$ . There exists  $g_{i,j} = 1$ ,  $i \in N_l$  and  $j \in N \setminus \{i\}$ . Suppose that  $j \in N_{l'}$ ,  $l \neq l'$ . We first note that, given  $g_{i,j} = 1$ , it must be true that  $N_l \subset C(g)$ . This follows by noting that the returns to a player  $k \in N_l$  from linking with component  $C(g)$  are strictly greater than the returns to player  $i$ , while the costs are strictly smaller (since  $k$  forms a link with  $i$ ). Hence every player  $k \in N_l$  must belong to  $C(g)$ . Therefore  $i \in N_l$  must access every  $i' \in N_l$  in  $g$ . There are two possibilities. One,  $i$  accesses  $i'$  via  $j$ . This violates Lemma 1. Two,  $i$  accesses  $i'$  via a player  $j'$ , where  $g_{j',i} = 1$ . Given  $g_{i,j} = 1$ , Lemma 2 implies that the link  $g_{j',i} = 1$  is sustainable in a strict Nash network, only if  $j'$  belongs to a group that is not accessed by  $i$  before the link  $g_{j',i} = 1$  has been formed. Next note that, using the above argument, it follows that all members of  $j'$ 's group must belong to  $C(g)$ . Suppose  $j' \in N_x$ . Then Lemmas 1 and 2 imply that  $j'$  accesses any  $j'' \in N_x$  either by being directly linked, and if this is the case the proof trivially follows, or by being passively linked with some player  $j''' \in N_y$ , belonging to a group other than  $l$ . We can then repeat the same argument with respect to  $j''$  and  $j'''$ . Since the number of groups is finite, we will eventually arrive at a point where two members of the same group are directly linked. The proof follows.  $\square$

**Lemma 4** Assume  $|N_l| \geq 2, \forall l = 1, \dots, m$ . Suppose  $g$  is a non-empty strict Nash network. If  $g_{i,i'} = 1, i, i' \in N_l$ , then  $g_{i,i''} = 1, \forall i'' \in N_l \setminus \{i\}$ .

**Proof:** Consider a non-singleton component,  $C(g)$ . Given the argument in Lemma 3, if  $g_{i,i'} = 1$ , for  $i, i' \in N_l$ , then  $N_l \subset C(g)$ . We first note that, if  $g_{i,i'} = 1$ , then  $g_{i'',i} = 0, \forall i'' \in N_l \setminus \{i\}$ . This follows from the standard switching argument. We have two possible configurations. First, suppose that  $N_l \equiv C(g)$ . Then an application of the switching argument immediately implies that  $g_{i,i''} = 1$ , for all  $i'' \in N_l$ . Second, suppose  $N_l \subsetneq C(g)$ . Since  $C(g)$  is connected, there is a path between  $i$  and  $i''$ , and  $d(i, i'') \geq 2$ . Then there is some player  $j \neq i''$  such that  $\bar{g}_{i,j} = 1$ . Suppose that  $j \in N_l$ . If  $g_{i,j} = 1$  then a simple switching argument applies with regard to player  $i$  and this contradicts the hypothesis that  $g$  is strict Nash. If  $g_{j,i} = 1$  then the switching argument applies to player  $j$ , who is indifferent between the link with  $i$  and the link with  $i'$ . This contradicts the hypothesis that  $g$  is strict Nash. Similar arguments can be used in the case that  $j \notin N_l$  to complete the proof of this lemma.  $\square$

**Lemma 5** Assume  $|N_l| \geq 2, \forall l = 1, \dots, m$ . Suppose  $g$  is a connected strict Nash network and let  $i \in N_l$  be the player identified by Lemma 4. Then any path  $i \xrightarrow{\bar{g}} j, \forall j \in N \setminus \{i\}$ , is  $i$ -oriented.

**Proof:** Let  $g$  be a strict Nash network which is connected. Since  $g$  is minimal, every path starting at  $i$  ends with a well defined end-player. The proof proceeds by contradiction. Suppose there is a path ending with player  $j$ , which is not  $i$ -oriented. Suppose  $\bar{g}_{i,j} = 1$ ; in this case  $g_{j,i} = 1$ . Since  $|N_l| \geq 2$ , a switching argument can be applied for player  $j$  with respect to some other member of his group which leads to a contradiction with the hypothesis that  $g$  is a strict Nash network.

Suppose next that  $\bar{g}_{i,j} = 0$ . Let  $\{i_1, i_2, i_3, \dots, i_n\}$ , be the players on the path between  $i$  and  $j$ , with  $\bar{g}_{i,i_1} = \dots = \bar{g}_{i_n,j} = 1$ . We first take up the case  $g_{j,i_n} = 1$ . Let  $j \in N_x$ ; if  $i_n \notin N_x$  then a simple switching argument with regard to player  $j$  and some member of his group implies that  $g$  is not a strict Nash network. If  $i_n \in N_x$ , there are two possibilities: (i)  $g_{i_{n-1},i_n} = 1$  and (ii)  $g_{i_n,i_{n-1}} = 1$ . In the first case, player  $i_{n-1}$  is indifferent between a link with player  $i_n$  and a link with player  $j$ . This contradicts the hypothesis that  $g$  is a strict Nash network. In the second case, there are two sub-cases: suppose  $i_n$  and  $i_{n-1}$  belong to the same group; then a switching argument applies to player  $j$ , with respect to players  $i_n$  and  $i_{n-1}$ . If  $i_n$  and  $i_{n-1}$  belong to different groups then a switching argument applies to player  $i_n$  with regard to members of the group of  $i_{n-1}$  (given that  $|N_l| \geq 2$ , for all  $l = 1, 2, \dots, m$ ).

Consider finally the case  $g_{i_n,j} = 1$ . Let  $k$  be the first player along the path  $\{i_1, i_2, \dots, i_n\}$ , such that  $g_{k,k-1} = 1$ . Let  $i_{k-1} \in N_y$ . Since  $g_{k-2,k-1} = 1$ , Lemma

1 implies that  $i_k, i_{k+1}, \dots, i_n \notin N_y$ . By hypothesis,  $|N_y| \geq 2$ , and so there is a player  $p \in N_y$ ,  $p \neq i_{k-1}$ , and we know that  $p \notin \{i_k, i_{k+1}, \dots, i_n, j\}$ . Thus,  $p \in N \setminus \{i_{k-1}, i_k, \dots, i_n, j\}$ . In this case however, a switching argument would apply to player  $i_k$  with regard to  $p$ . Hence  $g$  is not a strict Nash network. This contradiction completes the proof of the lemma.  $\square$

**Lemma 6** *Assume  $|N_l| \geq 2, \forall l = 1, \dots, m$ . Suppose  $g$  is a connected strict Nash network. Then  $D(g) \leq 2m$ .*

**Proof:** This follows directly by Lemma 1, 3, 4 and 5  $\square$

We now complete the proof of Proposition 4.2.

1. Consider a strict Nash network  $g$  and suppose  $c_L > 1$ . We claim that the only strict Nash network is the empty one. Suppose that there exists a non-singleton component  $C(g)$ . Using arguments from Lemma 3 it follows that if  $i \in N_l$ , and  $g_{i,j} = 1$ , then  $N_l \subset C(g)$ . If  $N_l \equiv C(g)$ , then it is easy to show by applying the switching argument that  $C(g)$  is a center-sponsored star. However, this is impossible given the hypothesis that  $c_L > 1$ . If on the other hand,  $C(g)$  contains players from more than one group then it follows that  $g$  is a connected network. Lemma 5 now implies that there is central player and that all paths are oriented towards this player. However, given that  $f(1) \geq c_l > 1$ , this is not sustainable in equilibrium. This contradicts the hypothesis that  $g$  is a strict Nash equilibrium. Hence the empty network is the only possible strict Nash network.
- 2a. Suppose  $c_L \in (0, 1)$  and  $f(1) \in (c_L, 1)$ . Suppose  $g$  is a strict Nash network; given the parameter restrictions, it is immediate that  $g$  must be connected. Lemma 3 and Lemma 4 imply that  $g$  satisfies property (A). Since  $g$  is connected, Lemma 5 holds and that implies property (B). Considering the restrictions imposed by Lemma 1, Property (C) follows by verification. Finally, Lemma 6 implies that the diameter of  $g$  satisfies properties (D)
- 2b. Suppose  $c_L \in (0, 1)$  and  $f(1) \in (1, \max[N_1, \dots, N_m])$ . Suppose  $g$  is a strict Nash network; it follows from Proposition 4.1 that it is connected. Lemma 5 implies that  $g$  has a central player  $i$ , and that all paths are  $i$ -oriented. However,  $f(1) > 1$ ,  $g$  cannot be sustained in equilibrium, leading to a contradiction. Hence, there does not exist a strict Nash network.
- 2c. Suppose  $c_L \in (0, 1)$  and  $f(1) > \max[N_1, \dots, N_m]$ . Consider a strict Nash network  $g$ . From Lemmas 3 and 4 it follows that either  $g$  has  $m$  components corresponding to each of the groups or it is connected. In the former case, Lemmas 3 and 4 imply that each of the components is a center-sponsored star. In the latter case, Lemma 5 implies that  $g$  has a central player and all the paths are oriented towards this player. But then the argument from Part 2b applies and such a network cannot arise in equilibrium given that  $f(1) > \max[N_1, N_2, N_3, \dots, N_m]$ .

□

**Proof of Proposition 4.3:** In this proposition we assume equal group size, *i.e.*  $|N_l| = |\bar{N}|$  for any  $l = 1, \dots, m$ . We first start with two observations: (a) The no-decay assumption implies that each non-singleton component part of an efficient architecture is minimal; (b) If  $g$  is efficient and non-empty then it is either minimally connected with  $m - 1$  outside links of ‘length’ one and  $m|\bar{N}| - m$  inside links, or partially connected with each group generating a minimal connected component. This observation follows by the assumption of equal group size and by the definition of efficiency concept. If a link between two members of the same group is socially efficient, then, from a societal point of view, each group should be internally linked. Furthermore, the assumption of equal group sizes implies that each group internally linked contributes equally to the total social welfare produced by the network. It follows that if an outside link is social enhancing, then an efficient network should be minimally connected. Moreover, since the definition of efficiency requires the minimization of the total cost of information flow, a connected efficient network should have  $m - 1$  outside links of length one. Using these observations we compare three different architectures:

- 1) The social welfare from  $g^{mc}$ , is given by:

$$W(g^{mc}) = (m|\bar{N}|)^2 - m(|\bar{N}| - 1)c_L - (m - 1)f(1) \quad (13)$$

- 2) The social welfare from  $g_m^{pc}$ , is given by:

$$W(g_m^{pc}) = m(|\bar{N}|)^2 - m(|\bar{N}| - 1)c_L \quad (14)$$

- 3) The social welfare from  $g^e$  is given by:

$$W(g^e) = m|\bar{N}| \quad (15)$$

First, we compare  $g_m^{pc}$  with  $g^e$ . It is easily checked that  $W(g_m^{pc}) \geq W(g^e)$  if and only if  $c_L \leq |\bar{N}|$ .

Second, suppose  $c_L \in (0, |\bar{N}|]$  and compare  $g^{mc}$  with  $g_m^{pc}$ . Simple computations show that  $W(g^{mc}) \geq W(g_m^{pc})$  if and only if  $f(1) \leq m|\bar{N}|^2 = c_1$ . It follows that given  $c_L \in (0, |\bar{N}|]$  if  $f(1) \in (c_L, c_1]$  the only efficient network is  $g^{mc}$ , while if  $f(1) > c_1$  the only efficient network is  $g_m^{pc}$ . This proves part (1).

Third, suppose  $c_L > |\bar{N}|$  and compare  $g^{mc}$  with  $g^e$ . Again, simple computations show that  $W(g^{mc}) \geq W(g^e)$  if and only if  $f(1) \leq \frac{m|\bar{N}|(m|\bar{N}|-1) - (m|\bar{N}|-m)c_L}{m-1} = c_2$ . We note that  $c_2$  is a decreasing function of  $c_L$  and attains the value  $m|\bar{N}|$  when  $c_L = m|\bar{N}|$ . Suppose therefore that  $c_L \in (|\bar{N}|, m|\bar{N}|)$ . If  $f(1) \in (c_L, c_2]$  then  $g^{mc}$  is uniquely efficient, while if  $f(1) > c_2$  then  $g^e$  is uniquely efficient. Finally, if  $c_L \geq m|\bar{N}|$  then  $c_2 \leq c_L$ . Given our hypothesis that  $f(1) > c_L$  it follows that empty network is uniquely efficient. This proves parts (2) and (3). □

## References

- [1] Bala, V. and S. Goyal (2000): A Non-cooperative Model of Network Formation, *Econometrica*, 68, 1181-1229.
- [2] Dutta, B., van den Nouweland, A. and S. Tijs (1995), *International Journal of Game theory*, 5, 205-228.
- [3] Jackson, M. and A. Wolinsky (1996): A Strategic Model of Social and Economic Networks, *Journal of Economic Theory*, 71, 44-74.
- [4] Johnson, C. and R.P. Gilles (2001): Spatial Social Networks, *Review of Economic Design*, 5, 3, 273-301.
- [5] Haller, H. and S. Sarangi (2000): Nash Networks with Heterogenous Players, *Working Paper VPI*.
- [6] Kranton, R. E. and D. F. Minehart (2001): A Theory of Buyer-Seller Networks, *The American Economic Review*, 91, 485-508.
- [7] McBride, M. (2001): Non-cooperative Network Equilibria under Heterogeneity and Uncertainty, mimeo, Yale University.
- [8] Wasserman, S. and K. Faust (1994): *Social Network Analysis: Methods and Applications*. Cambridge University Press.



Figures:

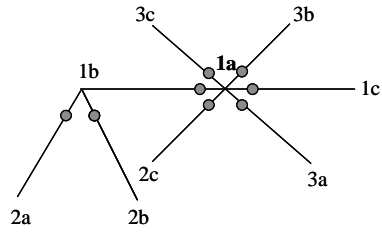


Figure 1A: Generalized Center-Sponsored Star

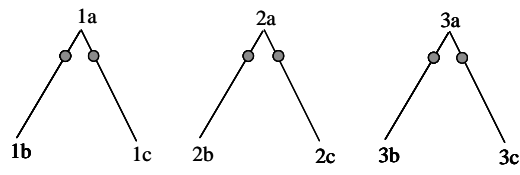


Figure 1B

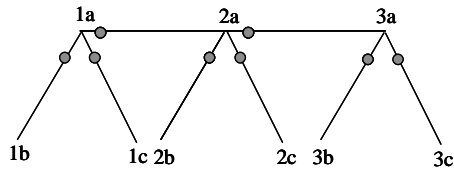


Figure 2

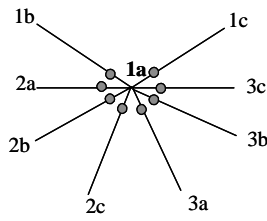


Figure 3: Center-Sponsored Star

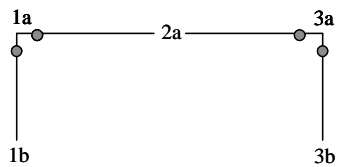


Figure 4