



No. 2004–70

**SOCIAL NETWORK FORMATION WITH CONSENT**

By R.P. Gilles, S. Sarangi

August 2004

ISSN 0924-7815

# Social Network Formation with Consent\*

Robert P. Gilles<sup>†</sup>

Sudipta Sarangi<sup>‡</sup>

July 2004

## Abstract

We investigate the equilibria of game theoretic models of network formation that are based on individual actions only. Our approach is grounded in three simple and realistic principles: (1) Link formation should be a binary process of consent. (2) Link formation should be costly. (3) The class of network payoff functions should be as general as possible.

It is accepted that these consent models have a very large number of equilibria. However, until now no characterization of these equilibria has been established in the literature. We aim to fill this void and provide characterizations of stable networks or the cases of two-sided and one-sided link formation costs. Furthermore, we provide a comparison of Nash equilibria with potential maximizers for a certain specification.

**Keywords:** Social networks; network formation; individual stability.

---

\*We would like to thank Matt Jackson, John Conlon, Dimitrios Diamantaras and Hans Haller for elaborate discussions on the subject of this paper and related work. We also thank seminar participants at LGS 3 (Siena, Italy), GTS 2004 (Marseille, France), Virginia Tech, and Louisiana State University for their comments and suggestions.

<sup>†</sup>**Corresponding author.** *Address:* Department of Economics, Virginia Tech (0316), Blacksburg, VA 24061, USA. *Email:* rgilles@vt.edu. Part of this research was done at the Center for Economic Research at Tilburg University, Tilburg, the Netherlands. Their hospitality is gratefully acknowledged.

<sup>‡</sup>*Address:* Department of Economics, Louisiana State University, Baton Rouge, LA 70803, USA. *Email:* sarangi@lsu.edu

# 1 Modeling consent in link formation

Networks impact the way we behave, the information we receive, the communities we are part of, and the opportunities that we pursue. They affect the machinations of corporations, the benevolence of non-profit organizations, and the workings of the state. Two recent overviews of the literature on statistical properties of large scale networks, Watts (2003) and Newman (2003), discuss the relevance of networks for fields as diverse as physics, social psychology, sociology, and biology. There has been a similar resurgence of interest in economics to understand the phenomenon of network formation. A number of recent contributions to the literature have recognized that networks play an important role in the generation of economic gains for groups of decision makers. Different network structures usually lead to different levels of generated gains, and network relationships between individuals have been interpreted in different ways. Among others, for example, such relationships could represent communication possibilities (Bala and Goyal 2000), trade relations (Kranton and Minehart 2001), or authority relationships between superiors and subordinates (van den Brink and Gilles 2003, Slikker, Gilles, Norde, and Tijs 2004).

In this paper we study two game-theoretic models of social network formation based on individual actions only. Players in our framework are represented by nodes and their social ties with others by links between these nodes. Nodes and links form together a representation of a social network. Our theory of social network formation is based on *three simple and realistic principles* that govern most real-world networks: (1) Link formation should be based on a binary process of consent. (2) Link formation should in principle be costly. (3) The payoff structure of network formation should be as general as possible.

We develop our approach from the hypothesis that creation of social ties requires some prior interaction and, therefore, the process of link formation under mutual consent principally occurs between social acquaintances. In the sociology literature it has been established that social networks are indeed primarily formed between acquaintances. This literature is founded on Granovetter (1973) and confirmed empirically by Friedkin (1980), Wellman, Carrington, and Hall (1988), and Tyler, Wilkinson, and Huberman (2003).<sup>1</sup>

Here we follow this line of reasoning and differentiate between familiarity among individuals, who can at best only be acquaintances, and the possibility of explicitly creating a mutually beneficial but costly relationship between the same individuals. This is in line with Brueckner (2003), who categorically distinguishes the set of acquaintances a player has, from the friendship links she establishes between them. This also places our approach

---

<sup>1</sup>More recently new methodologies have been developed to detect community structures in social networks for testing such hypotheses. We refer to Newman and Girvan (2004) and Newman (2004) for the details of this methodology.

within the context of Granovetter’s notion of strong social ties.

In our theory, the creation of a social tie or “link” requires the consent of both players involved; the link between players  $i$  and  $j$  is only established when player  $j$  is willing to accept the link initiated by player  $i$  or vice versa. As suggested by our second principle, we emphasize that link formation is costly. Costs depend on the strategies chosen by the player in the link formation process and are incurred independent of the outcome, i.e., even if a link is not established the initiating player still has to pay for the act of trying to form that link. We consider both two-sided and one-sided costs of link formation. In the first model both players bear an individually determined cost of link formation, while in the latter model we distinguish between an “initiator” and a “respondent” in the link formation process with only the initiator incurring the link formation cost. To meet our third requirement, we consider a very general payoff structure that has two components — an arbitrary benefit function corrected for additive link formation costs.<sup>2</sup> We emphasize that benefits depend on the resulting network, and the costs on the link formation strategies chosen by the actors.

The process of network formation studied here is a generalization of the simple network formation model developed by Myerson (1991, page 448). Following Myerson, we model the link formation process as a normal form non-cooperative game. This model incorporates the fundamental idea that networks are the result of costly, consensual link formation between pairs of players. We enhance this model by taking into account the three requirements discussed above. Since this model is rather well known in the literature, we call this generalization of Myerson’s model the *standard model of network formation*.<sup>3</sup>

In the literature, the standard model often features in discussions on social network formation but has been portrayed as being problematic since it is believed to have “too many” Nash equilibria. (Jackson 2003, for example) However, until now there has been made no attempt to provide a complete characterization of the set of these Nash equilibria and our paper aims to fill this void in the literature. Our characterization reveals that the resulting networks have some appealing properties. Also, to abandon a realistic and elegant model because it is not discerning enough in terms of its permissible equilibria seems hardly justifiable.

In order to understand the importance of the ability to break (or deny) links in the process of network formation we introduce a stability concept called *link deletion proofness*: a

---

<sup>2</sup>An arbitrary cost structure would require costs to be dependent on the outcome. The payoff specification then would become game dependent forcing us to give up generality in the results. We believe that the chosen payoff structure based on arbitrary benefits and *additive* link formation costs has the added advantage of capturing what genuinely occurs in a realistic process of link formation.

<sup>3</sup>For other sources on the standard model we refer to Belleflamme and Bloch (2004), Bloch and Jackson (2004), Calvó-Armengol and Ilkiliç (2004), and Gilles, Chakrabarti, Sarangi, and Badasyan (2004).

network is link deletion proof when players get a lower payoff by deleting *exactly one* of their established links. A variation called *strong link deletion proofness* allows players to consider the simultaneous deletion of *multiple* links.

Subsequently we examine the relationship between the classes of networks satisfying these stability concepts, and the set of networks resulting from the Nash equilibria of the network formation game. The latter class is denoted as the set of *individually stable networks*. In general we find that link deletion proofness and strong link deletion proofness are equivalent if and only if network payoffs satisfy a convexity property. This *network convexity* condition is weaker than the  $\alpha$ -convexity condition introduced by Calvó-Armengol and Ilkiliç (2004).

Next we turn to the characterization of individually stable networks. For the case with two-sided link formation costs we find that a network is individually stable if and only if it is strong link deletion proof. This result confirms the well-accepted conjecture that there are a multitude of Nash equilibria in network formation models under consent.

Finally, we study the one-sided cost model where only the link initiating player incurs a cost. We find that if a network is individually stable under two-sided link formation costs, then it is also individually stable under one-sided link formation costs. The reverse does not hold. Moreover, we find that all strong link deletion proof networks are individually stable while the converse does not hold. On the other hand, we provide a (partial) characterization that shows that individually stable networks can be captured in a very large class of partially stable networks. Again these insights confirm the well-accepted conjecture that the class of individually stable networks is extremely large and non-discerning.

We conclude our investigations with the analysis of a simple payoff specification based on link-based network benefits. This setting is used to investigate the relationship between potential maximizers and Nash equilibria. We find that the potential maximizer concept is an useful refinement of Nash equilibrium for the model with one-sided link formation costs, contrary to the case of two-sided link formation costs.

Since the standard model of network formation is sufficiently general it can incorporate a number of existing network models. We first point to the existence of individually stable networks. Under two-sided link formation costs, it is possible to find parallels in the literature on pairwise stability. This implies that the existence of individually stable networks for the two-sided cost model is guaranteed for a large class of specifications. (Jackson and Watts 2002) For the case of one-sided link formation costs, similar parallels can be drawn with the Nash network formulation developed by Bala and Goyal (2000). In our framework the flow of benefits is two-way, while only the initiating player incurs the cost of the link in the one-sided case. Since giving consent to link formation under one-sided costs

is costless, and under the Bala-Goyal type of specification always yields positive benefits, the responding player would immediately consent to the link. Hence, existence of individually stable networks under one-sided link formation costs is guaranteed for a large class of specifications.

The remainder of this paper is organized as follows. Section 2 of the paper provides notation and the model setup. Section 3 introduces the standard model of link formation under consent and two-sided link formation costs. Section 4 discusses the case of one-sided link formation costs. Section 5 elaborates on the interesting case of link-based network payoffs. The proofs of the main results are relegated to Section 6.

## 2 Preliminaries and notation

In this section we introduce the basic concepts and notation pertaining to non-cooperative games and networks.

### 2.1 Non-cooperative games

A *non-cooperative game* on the fixed, finite player set  $N = \{1, \dots, n\}$  is given by a list  $(A_i, \pi_i)_{i \in N}$  where for every player  $i \in N$ ,  $A_i$  denotes an action set and  $\pi_i: A \rightarrow \mathbb{R}$  denotes player  $i$ 's payoff function, where  $A = A_1 \times \dots \times A_n$  is the set of *action tuples*. An individual action of player  $i \in N$  is denoted by  $a_i \in A_i$  and an action tuple is written as  $a = (a_1, \dots, a_n) \in A$ . For every action tuple  $a \in A$  and player  $i \in N$ , we denote by  $a_{-i} = (a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n) \in A_{-i} = \prod_{j \neq i} A_j$  the actions selected by the players other than  $i$ . In the rest of the paper we also denote a non-cooperative game on  $N$  for short by the pair  $(A, \pi)$ , where  $\pi = (\pi_1, \dots, \pi_n): A \rightarrow \mathbb{R}^N$  is the composite payoff function. In this paper we only discuss *finite* non-cooperative games in the sense that for every  $i \in N$  the action set  $A_i$  is finite.

An action  $a_i \in A_i$  for player  $i \in N$  is called a *best response* to  $a_{-i} \in A_{-i}$  if for every action  $b_i \in A_i$  we have that  $\pi_i(a_i, a_{-i}) \geq \pi_i(b_i, a_{-i})$ . A best response  $a_i$  to  $a_{-i}$  is *strict* if for every  $b_i \neq a_i$  we have that  $\pi_i(a_i, a_{-i}) > \pi_i(b_i, a_{-i})$ . An action tuple  $\hat{a} \in A$  is a *Nash equilibrium* of the game  $(A, \pi)$  if for every player  $i \in N$

$$\pi_i(\hat{a}) \geq \pi_i(b_i, \hat{a}_{-i}) \text{ for every action } b_i \in A_i.$$

Hence, a Nash equilibrium  $\hat{a} \in A$  satisfies the property that for every player  $i \in N$  the action  $\hat{a}_i$  is a best response to  $\hat{a}_{-i}$ . A Nash equilibrium  $\hat{a} \in A$  is called *strict* if for every player  $i \in N$  the action  $\hat{a}_i$  is a strict best response to  $\hat{a}_{-i}$ .

A function  $Q: A \rightarrow \mathbb{R}$  is a *potential* of the non-cooperative game  $(A, \pi)$  on the player set  $N$  if for every player  $i \in N$ , action tuple  $a \in A$  and action  $b_i \in A_i$ :

$$\pi_i(a) - \pi_i(b_i, a_{-i}) = Q(a) - Q(b_i, a_{-i}).$$

The notion of a potential game was introduced by Monderer and Shapley (1996) based on the seminal work of Hart and Mas-Colell (1989). Monderer and Shapley (1996) proposed the notion of a *potential maximizer* being an action tuple  $a \in A$  such that  $Q(a) \geq Q(b)$  for every  $b \in A$ . The set of potential maximizers is denoted by  $PM(A, \pi) \subset A$ . It is obvious that each potential maximizer is a Nash equilibrium and, hence, this notion is a refinement of the Nash equilibrium concept. Monderer and Shapley (1996) show that  $PM(A, \pi) \neq \emptyset$  for every finite potential game  $(A, \pi)$  on  $N$ .

An alternative description of a potential game has been introduced by Ui (2000) as follows. A *coalition* is any subset of players  $S \subset N$  and for a coalition  $S$  we denote by  $A_S = \prod_{i \in S} A_i$  its restricted action tuple set. A set of functions  $\{\Phi_S: A_S \rightarrow \mathbb{R} \mid S \subset N\}$  is an *interaction potential* of the game  $(A, \pi)$  if for every  $i \in N$  and every  $a \in A$  it holds that

$$\pi_i(a) = \sum_{S \subset N: i \in S} \Phi_S(a_S).$$

Ui showed that potentials and interaction potentials are essentially the same:

**Lemma 2.1 (Ui 2000, Theorem 3)** *The game  $(A, \pi)$  has a potential  $Q: A \rightarrow \mathbb{R}$  if and only if  $(A, \pi)$  possesses an interaction potential  $\{\Phi_S \mid S \subset N\}$ . Furthermore, for the latter case a potential  $Q$  of the game  $(A, \pi)$  is given by  $Q(a) = \sum_{S \subset N} \Phi_S(a_S)$ .*

We will use these insights to analyze properties of certain specifications of network payoffs in Section 5.

## 2.2 Networks

In our discussion of the foundations of the theory of networks we use established notation from Jackson and Wolinsky (1996), Dutta and Jackson (2003), and Jackson (2003). The reader may refer to these sources for a more elaborated discussion.

We limit our discussion to *non-directed networks* on the player set  $N$ . In these networks the two players making up a single link are essentially equal. Formally, if two players  $i, j \in N$  with  $i \neq j$  are related we say that there exists a *link* between players  $i$  and  $j$ . We use the notation  $ij$  to describe the binary link  $\{i, j\}$ .<sup>4</sup> We define  $g_N = \{ij \mid i, j \in N, i \neq j\}$  as the set of all potential links.

---

<sup>4</sup>We reiterate that network relationships are non-directed, i.e., in this context  $ij = ji$ . However, for the costs of establishing a link one may distinguish between the costs related to  $ij$  and the costs related to  $ji$ , i.e., possibly it holds that  $c_{ij} \neq c_{ji}$ .

Any set of links  $g \subset g_N$  now defines a *network* on  $N$ . We apply the convention that  $g = g_N$  is called the *complete network* and that  $g = g_0 = \emptyset$  is indicated as the *empty network*. We denote by  $\mathbb{G}^N = \{g \mid g \subset g_N\}$  the class of all networks on  $N$ .

The set of (direct) *neighbors* of a player  $i \in N$  in the network  $g$  is given by

$$N_i(g) = \{j \in N \mid ij \in g\}.$$

Similarly we introduce

$$L_i(g) = \{ij \in g_N \mid j \in N_i(g)\}$$

as the *link set* of player  $i$  in the network  $g$ . These are exactly the links with  $i$ 's direct neighbors in  $g$ . We apply the convention that for every player  $i \in N$  we denote by  $L_i = L_i(g_N) = \{ij \mid i \neq j\}$  the set of all potential links involving player  $i$ .

For every pair of players  $i, j \in N$  with  $i \neq j$  we denote by  $g + ij = g \cup \{ij\}$  the network that results from adding the link  $ij$  to the network  $g$ . Similarly,  $g - ij = g \setminus \{ij\}$  denotes the network resulting from removing the link  $ij$  from network  $g$ . More generally for any  $h \subset g$  we let  $g - h = g \setminus h$  and for any  $h \subset g_N$  with  $h \cap g = \emptyset$  we let  $g + h = g \cup h$ .

Within a network, payoffs for the players are generated depending on how they are connected to each other. This is represented by a “network payoff function” for every player. For player  $i \in N$  the function  $\varphi_i: \mathbb{G}^N \rightarrow \mathbb{R}$  denotes her *network payoff function* which assigns to every network  $g \subset g_N$  a value  $\varphi_i(g)$  that is obtained by player  $i$  when she participates in network  $g$ . The composite network payoff function is now given by  $\varphi = (\varphi_1, \dots, \varphi_n): \mathbb{G}^N \rightarrow \mathbb{R}^N$ . We emphasize that these payoffs can be zero, positive, or negative and that the empty network  $g_0 = \emptyset$  generates (reservation) values  $\varphi(g_0) \in \mathbb{R}^N$  that might be non-zero as well.

Several examples of standard network payoff functions for both noncooperative and cooperative games are reviewed in Jackson (2003).<sup>5</sup>

## 2.3 Link-based stability concepts

We conclude the preliminaries on network theory with the definition and discussion of several stability conditions. Note that the stability notions introduced below are based on the properties of the network itself rather than strategic considerations of the players. This latter viewpoint originates from Jackson and Wolinsky (1996).

---

<sup>5</sup> We mention a specific class of network payoff functions, which is investigated in van den Nouweland (1993), Dutta, van den Nouweland, and Tijs (1998), Slikker (2000), Slikker and van den Nouweland (2000), and Garratt and Qin (2003). There these network payoff functions are defined as allocation rules based on underlying cooperative games. These papers extend the seminal contribution Myerson (1977) that set this game-theoretic literature on network formation into motion.



First we consider some auxiliary notation: Let  $\varphi: \mathbb{G}^N \rightarrow \mathbb{R}^N$  be some network payoff function. For a given network  $g \in \mathbb{G}^N$  we now define the following concepts:

- (a) For every  $ij \in G$  the *marginal benefit* of the link  $ij$  in  $g$  is given by

$$D(g, ij) = \varphi(g) - \varphi(g - ij) \in \mathbb{R}^N \quad (1)$$

and for every player  $i \in N$  the marginal benefit of  $ij \in L_i(g)$  is thus given by

$$D_i(g, ij) = \varphi_i(g) - \varphi_i(g - ij) \in \mathbb{R}.$$

- (b) For every player  $i \in N$  and link set  $h \subset L_i(g)$  the *marginal benefit* to player  $i$  of link set  $h$  in  $g$  is given by

$$D_i(g, h) = \varphi_i(g) - \varphi_i(g - h) \in \mathbb{R} \quad (2)$$

Using these additional tools we can give a precise description of the various link-based stability concepts.

**Definition 2.2** Let  $\varphi$  be a network payoff function on the player set  $N$ .

- (a) A network  $g \in g_N$  is **link deletion proof** if for every player  $i \in N$  and every  $j \in N_i(g)$  it holds that  $D_i(g, ij) \geq 0$ .
- (b) A network  $g \in g_N$  is **strong link deletion proof** if for every player  $i \in N$  and every  $h \subset L_i(g)$  it holds that  $D_i(g, h) \geq 0$ .
- (c) A network  $g \in g_N$  is **link addition proof** if for all players  $i, j \in N$ :  $\varphi_i(g + ij) > \varphi_i(g)$  implies  $\varphi_j(g + ij) < \varphi_j(g)$ .

The two link deletion proofness notions are based on the severance of links in a network by individual players. In particular, the notion of link deletion proofness considers the stability of a network with regard to the deletion of a *single* link. Strong deletion proofness considers the possibility that a player deletes any subset of her existing links. Clearly, strong link deletion proofness implies link deletion proofness.

Similarly, link addition proofness considers the addition of a single link by two consenting players to an existing network. A network is link addition proof if for every pair of non-linked players at least one of these two players has negative benefits from the addition of a link between them. Hence, there are no incentives to add any additional links to the existing network.<sup>6</sup>

---

<sup>6</sup>Closely related to these basic stability concepts is the notion of *pairwise stability* seminally introduced by Jackson and Wolinsky (1996). Formally, a network is pairwise stable if it is link deletion proof as well as link addition proof.

We denote by  $\mathcal{L}(\varphi) \subset \mathbb{G}^N$  the family of link deletion proof networks for  $\varphi$ . Similarly, we let  $\mathcal{L}_s(\varphi) \subset \mathbb{G}^N$  be the family of strong link deletion proof networks for  $\varphi$ .

Next we state the precise conditions under which link deletion proofness and strong link deletion proofness are equivalent.

**Definition 2.3** For a player  $i \in N$  the network payoff function  $\varphi_i: \mathbb{G}^N \rightarrow \mathbb{R}$  is **network convex** on the network  $g \in \mathbb{G}^N$  if for every link set  $h \subset L_i(g)$  we have that

$$\sum_{ij \in h} D_i(g, ij) \geq 0 \text{ implies } D_i(g, h) \geq 0.$$

The following result justifies the introduction of this network convexity property. It corrects the assertion that the equivalence of strong pairwise stability and pairwise stability holds if and only if  $\pi$  satisfies  $\alpha$ -convexity. (Calvó-Armengol and Ilkiliç 2004, Theorem 1)

**Proposition 2.4** Let  $\varphi$  be some network payoff structure on  $\mathbb{G}^N$ . Then  $\mathcal{L}_s(\varphi) = \mathcal{L}(\varphi)$  if and only if for every player  $i \in N$  the network payoff function  $\varphi_i$  is network convex on every link deletion proof network  $g \in \mathcal{L}(\varphi)$ .

For a proof of this assertion we refer to Section 6.

**Example 2.5** We conclude our discussion with an example which delineates the different link-wise stability concepts and shows a situation in which link deletion proofness and strong link deletion proofness lead to different results.

Consider the network payoffs given in the following table:

Network	$\varphi_1(g)$	$\varphi_2(g)$	$\varphi_3(g)$	Stability
$g_0 = \emptyset$	0	0	0	$L_s$
$g_1 = \{12\}$	-1	-1	-1	
$g_2 = \{13\}$	-1	-1	-1	
$g_3 = \{23\}$	5	3	3	$L_s$
$g_4 = \{12, 13\}$	1	1	1	$L_s$
$g_5 = \{12, 23\}$	0	4	0	
$g_6 = \{13, 23\}$	0	0	4	
$g_7 = g_N$	1	5	5	$L$

In the table  $L$  stands for link deletion proofness and  $L_s$  for strong link deletion proofness. The main feature here is that network  $g_7$  is link deletion proof, but not strong link deletion proof. To make the differences between the various possibilities more clear we provide an overview of the marginal benefits:

Network	$D(g, 12)$	$D(g, 13)$	$D(g, 23)$
$g_0 = \emptyset$	—	—	—
$g_1 = \{12\}$	$-1, -1, -1$	—	—
$g_2 = \{13\}$	—	$-1, -1, -1$	—
$g_3 = \{23\}$	—	—	$5, 3, 3$
$g_4 = \{12, 13\}$	$2, 2, 2$	$2, 2, 2$	—
$g_5 = \{12, 23\}$	$-5, 1, -3$	—	$0, 4, 0$
$g_6 = \{13, 23\}$	—	$-5, -3, 1$	$0, 0, 4$
$g_7 = g_N$	$1, 5, 1$	$1, 1, 5$	$0, 3, 3$

Note that  $D(g_7, 12) + D(g_7, 13) = (2, 6, 6)$  and that  $D(g_7, \{12, 13\}) = (-4, 2, 2)$ . Hence, the case of the removal of the links 12 and 13 from network  $g_7$  shows that  $\varphi$  is not network convex.

In  $g_7$  player 1 is stuck with bad company if she could delete only a single link at the time; she would like to break links with *both* players 2 and 3 and improve her payoff from 1 unit to 5 units. However, deleting either of these two links separately would make her only worse off. In this regard network convexity requires that no player is in such a bad company situation.  $\blacklozenge$

### 3 Two-sided link formation costs

In this section we present the first of two game-theoretic models of costly network formation. Let  $N = \{1, \dots, n\}$  be a given set of players and  $\varphi: \mathbb{G}^N \rightarrow \mathbb{R}^N$  be a fixed, but arbitrary network payoff function representing the gross benefits that accrue to the players in a network. For every player  $i \in N$  we introduce individualized link formation costs represented by  $c_i = (c_{ij})_{j \neq i} \in \mathbb{R}_+^{N \setminus \{i\}}$ . (Recall that for some links  $ij \in g_N$  it might hold that  $c_{ij} \neq c_{ji}$ .) Thus, the pair  $\langle \varphi, c \rangle$  represents the basic payoffs and costs of network formation to the individuals in  $N$ .

A simple, fundamental model of network formation has been introduced by Myerson (1991, page 448) and is based on the idea that pairs of players approach each other on equal footing and both have to consent to form a link. Myerson (1991) based the benefits from network formation on an underlying cooperative game.<sup>7</sup> Here we extend this framework further to incorporate costs of link formation for arbitrary network payoff functions. We model link formation costs in two ways: Costs can be *two-sided*, i.e., both players incur costs while approaching each other to form a link, or costs can be *one-sided*. In the latter case costs are only incurred by the initiating player, not the responding player.

<sup>7</sup>This cooperative benefits model has been extended by Slikker and van den Nouweland (2000) and Garratt and Qin (2003) to incorporate link formation costs. Their formulation only allowed them to develop a complete and exhaustive description of the resulting networks for situations with up to four individuals.

We first address the formalization of the standard model with two-sided link formation costs. For every player  $i \in N$  we introduce an action set

$$A_i^a = \{(\ell_{ij})_{j \neq i} \mid \ell_{ij} \in \{0, 1\}\} \quad (3)$$

Player  $i$  seeks contact with player  $j$  if  $\ell_{ij} = 1$ . A link is formed if both players seek contact, i.e.,  $\ell_{ij} = \ell_{ji} = 1$ .

Let  $A^a = \prod_{i \in N} A_i^a$  where  $\ell \in A^a$ . Then the resulting network is given by

$$g^a(\ell) = \{ij \in g_N \mid \ell_{ij} = \ell_{ji} = 1\}. \quad (4)$$

Link formation is costly. Approaching player  $j$  to form a link costs player  $i$  an amount  $c_{ij} \geq 0$ . This results in the following net payoff function for player  $i$ :

$$\pi_i^a(\ell) = \varphi_i(g^a(\ell)) - \sum_{j \neq i} \ell_{ij} \cdot c_{ij} \quad (5)$$

where  $c$  is the link formation cost introduced at the beginning of this section.

The pair  $\langle \varphi, c \rangle$  thus generates the non-cooperative game  $(A^a, \pi^a)$  as described above. We call this non-cooperative game the *standard model of network formation with two-sided link formation costs*.

Now a network  $g \in \mathbb{G}^N$  is called *individually stable under two-sided link formation costs* if there exists a Nash equilibrium action tuple  $\hat{\ell} \in A^a$  in the standard model with two-sided link formation costs  $(A^a, \pi^a)$  such that  $g^a(\hat{\ell}) = g$ . Hence, individually stable networks are those networks supported through Nash equilibrium strategies.

We are able to provide a complete characterization of individual stability under two-sided link formation costs.

**Proposition 3.1** *Let  $\varphi$  and  $c \geq 0$  be given as above. A network  $g \subset g_N$  is individually stable under two-sided link formation costs if and only if  $g$  is strong link deletion proof for the net payoff function  $\varphi^a$  given by*

$$\varphi_i^a(g) = \varphi_i(g) - \sum_{j \in N_i(g)} c_{ij}.$$

For a proof of this result we refer to Section 6.

Proposition 3.1 gives a complete characterization of the individually stable networks in the standard model with two-sided costs of link formation. Note that regardless of the cost structure, the empty network is always individually stable. The next corollary strengthens this insight by showing that the empty network is actually “strictly” individually stable for positive costs.

**Corollary 3.2** *If  $c \gg 0$ , then the empty network is supported by a strict Nash equilibrium of the standard model with two-sided link formation costs.*

**Proof.** First, for every  $i \in N$  and  $\ell \in A^a$  we define  $h_i(\ell) = \{ij \in g_N \mid \ell_{ij} = 1 \text{ and } \ell_{ji} = 0\}$ . We now show that  $\ell^\emptyset$  is a strict Nash equilibrium in the game  $(A^a, \pi^a)$ , where  $\ell_{ij}^\emptyset = 0$  for all players  $i, j \in N$  with  $i \neq j$ . Now, for every player  $i \in N$  and  $l_i \neq \ell_i^\emptyset$ :

$$\pi_i^a(l_i, \ell_{-i}^\emptyset) = \varphi_i(\emptyset) - \sum_{ij \in h_i(l_i, \ell_{-i}^\emptyset)} c_{ij} < \varphi_i(\emptyset) = \pi_i^a(\ell^\emptyset)$$

since  $h_i(l_i, \ell_{-i}^\emptyset) \neq \emptyset$ . Hence, we may conclude that indeed  $\ell^\emptyset$  is a strict Nash equilibrium in the link formation game  $(A^a, \pi^a)$ . ■

From Corollary 3.2 it should be clear that if players start from the empty network and link formation costs are positive, then there is no reason to form any links.

Dutta, van den Nouweland, and Tijs (1998) showed that in the cooperative benefits model under costless link formation, every network is individually stable if the network payoff function is “link monotonic”. Proposition 3.1 generalizes this insight for situations with arbitrary network payoff functions. This is stated in the next corollary which proof is immediate from Proposition 3.1.

**Corollary 3.3** *Assume that  $\varphi$  is link monotonic in the sense that  $\varphi_i(g) < \varphi_i(g + ij)$  for all networks  $g$  and players  $i \in N$  with  $ij \notin g$  where  $j \neq i$ . If  $c = 0$ , then every network is individually stable.*

## 4 One-sided link formation costs

Next we address the formalization of the standard model with one-sided link formation costs. Here links are formed by mutual agreement, but one player initiates the formation process and the other player responds to it. The initiator incurs the formation costs of the link, while the respondent incurs no costs.<sup>8</sup> Hence, a different strategy space is called for. Formally, for every player  $i \in N$  we introduce an action set

$$A_i^b = \{(\ell_{ij}, r_{ij})_{j \neq i} \mid \ell_{ij}, r_{ij} \in \{0, 1\}\}. \quad (6)$$

Player  $i$  acts as the initiator in forming a link with player  $j$  if  $\ell_{ij} = 1$ . Player  $j$  responds positively to this initiative if  $r_{ji} = 1$ . A link is established if formation is initiated and accepted, i.e., if  $\ell_{ij} = r_{ji} = 1$ . This is formalized as follows.

---

<sup>8</sup>We remark that a similar link formation structure has been already discussed by Slikker (2000) and Slikker, Gilles, Norde, and Tijs (2004) in the context of the formation of *directed* networks. See also Dutta and Jackson (2000).

Let  $A^b = \prod_{i \in N} A_i^b$ . Given the link formation procedure described, for any  $(\ell, r) \in A^b$ , the resulting network is now given by

$$g^b(\ell, r) = \{ij \in g_N \mid \ell_{ij} = r_{ji} = 1\}. \quad (7)$$

When player  $i$  initiates the formation of a link with player  $j$  she incurs a cost of  $c_{ij} \geq 0$ . Responding to the initiative by another player however, is costless. This results in the following net payoff function for player  $i$ :

$$\pi_i^b(\ell, r) = \varphi_i(g^b(\ell, r)) - \sum_{j \neq i} \ell_{ij} \cdot c_{ij} \quad (8)$$

where  $c$  denotes the link formation costs.

Analogous to the previous model with two-sided link formation costs, the pair  $\langle \varphi, c \rangle$  generates the non-cooperative game  $(A^b, \pi^b)$  introduced above. This game represents the *standard model with one-sided link formation costs*.

Like before, a network  $g \in \mathbb{G}^N$  is called *individually stable under one-sided link formation costs* if there exists a Nash equilibrium action tuple  $(\hat{\ell}, \hat{r}) \in A^b$  in the standard model with one-sided link formation costs  $(A^b, \pi^b)$  such that  $g^b(\hat{\ell}, \hat{r}) = g$ .

The next example discusses the simplest possible case of a single link between two players that illustrates the multitude of individually stable networks under one-sided link formation costs.

**Example 4.1** Let  $N = \{1, 2\}$ . Hence, we have two possible networks,  $g_0 = \emptyset$  and  $g_N = \{12\}$ . Consider the network payoff function  $\varphi$  given by  $\varphi_1(g_0) = \varphi_2(g_0) = 0$ ,  $\varphi_1(g_N) = 10$ , and  $\varphi_2(g_N) = 5$ . We now consider the following (equal) cost structures:

- (a)  $c_{12} = c_{21} \in [0, 5]$

In these cases the class of individually stable networks is the same for one-sided and two-sided link formation costs, namely both feasible networks  $g_0$  and  $g_N$ .

- (b)  $c_{12} = c_{21} \in (5, 10]$

In these cases, under two-sided link formation costs, only the empty network  $g_0$  is individually stable. However, under one-sided link formation costs, again both  $g_0$  and  $g_N$  are individually stable. Indeed, the complete network  $g_N$  is supported by strategy profiles with  $\ell_{12} = 1$ ,  $r_{21} = 1$ , and  $\ell_{21} = 0$ .

These two cases illustrate the main differences between two-sided and one-sided link formation costs. In particular it shows that under one-sided costs there are more individually stable networks.  $\blacklozenge$

The next result generalizes the insight of Example 4.1. For a proof of this proposition we refer to Section 6 of the paper.

**Proposition 4.2** *Let  $\varphi$  and  $c \geq 0$  be given. Any individually stable network under two-sided link formation costs is individually stable under one-sided link formation costs.*

Example 4.1 for link formation costs in the range  $(5, 10]$  shows that the assertion stated in Proposition 4.2 cannot be reversed.

In Proposition 3.1 we characterized the class of individually stable networks under two-sided link formation costs. However, such a complete characterization is not possible with one-sided link formation costs. Instead we provide two inclusions that show the largeness of the set of individually stable networks under one-sided link formation costs.

**Proposition 4.3** *Let  $\varphi$  be arbitrary and let  $c \geq 0$  be such that  $c_{ij} \neq c_{ji}$  for all potential links  $ij \in g_N$ .*

- (a) *If a network  $g \subset g_N$  is strong link deletion proof for the net payoff function  $\varphi^b$  given by*

$$\varphi_i^b(g) = \varphi_i(g) - \sum_{j \in N_i(g): c_{ij} < c_{ji}} c_{ij},$$

*then  $g$  is individually stable under one-sided link formation costs.*

- (b) *If  $g \subset g_N$  is individually stable under one-sided link formation costs, then for all links  $ij \in g$  it holds that*

$$D_i(g, ij) \geq c_{ij} \quad \text{or} \quad D_j(g, ij) \geq c_{ji}.$$

The assertion of Proposition 4.3(a) is proved in Section 6. Proposition 4.3(b) is a rather straightforward application of the definition of the marginal payoffs and individual stability under one-sided link formation costs. A formal proof of this assertion is therefore omitted and left to the reader.

The next example demonstrates that Proposition 4.3(a) cannot be reversed.

**Example 4.4** Again consider  $N = \{1, 2\}$ . As before we let  $g_0 = \emptyset$  and  $g_N = \{12\}$  with  $\varphi_1(g_0) = \varphi_2(g_0) = 0$ ,  $\varphi_1(g_N) = 1$ , and  $\varphi_2(g_N) = 10$ . We consider two different cost structures:

- (a) Consider  $c_{12} = 2 < c_{21} = 5$ .

Now both  $g_0$  and  $g_N$  are individually stable under one-sided link formation costs,

but  $g_N$  is not link deletion proof for  $\varphi^b$ .

Indeed a Nash equilibrium for the standard model with one-sided link formation costs supporting  $g_N$  is given by  $\ell_{12} = 0$ ,  $r_{12} = 1$ ,  $\ell_{21} = 1$ , and  $r_{21} = 0$ . Now,  $g^b(\ell, r) = g_N$ ,  $\pi_1^b(\ell, r) = 1 > 0 = \varphi_1(g_0)$ , and  $\pi_2^b(\ell, r) = 5 > 0 = \varphi_2(g_0)$ . However,  $\varphi_1^b(g_N) = -1 < \varphi_1^b(g_0) = 0$ , which implies that  $g_N$  is *not* link deletion proof with respect to  $\varphi^b$  for player 1.

- (b) Consider  $c_{12} = 11 > c_{21} = 5$ .

In this case again both  $g_0$  and  $g_N$  are individually stable under one-sided link formation costs. However, in this case the inclusion stated in Proposition 4.3(a) is tight.

Indeed, it can be checked that

$$\varphi_1^b(g_0) = \varphi_2^b(g_0) = 0$$

$$\varphi_1^b(g_N) = 1 - 0 = 1$$

$$\varphi_2^b(g_N) = 10 - 5 = 5$$

Hence, both  $g_0$  and  $g_N$  are strong link deletion proof with respect to  $\varphi^b$ . This confirms that in this case the inclusion stated in Proposition 4.3(a) is indeed tight.

Case (a) demonstrates a form of inefficiency in link formation, since in equilibrium higher than necessary costs are incurred. This implies that outside regulation of link formation processes — in the sense that an outside regulator determines who initiates which link — will restore efficiency. In this example, player 1 should be forced to initiate the link with player 2. ♦

With regard to the possibility of the tightness of the inclusion stated in Proposition 4.3(b) we refer to Example 4.1. There it has been shown that the collection of individually stable networks under one-sided link formation costs is exactly equal to the class of networks indicated in 4.3(b) for any cost structure. We refer to Section 5.2 for the discussion of another class of network payoff structures for which this inclusion is tight.

## 5 Equilibria and potential maximizers

Thus far we only considered network formation using arbitrary (network) payoff functions that do not rely on specific payoff structures or even on explicit formulations. In this section we develop the case of link-based network payoffs. We use this straightforward model to illustrate some interesting properties and arrive at some startling conclusions.

We first develop a simple formulation of link-based payoff generation. For that purpose we introduce  $\theta_i: L_i \rightarrow \mathbb{R}_+$  as a *link benefit function* for player  $i \in N$  that assigns to every



potential link  $ij \in L_i$  of player  $i$  a benefit  $\theta_i(ij) \geq 0$ . Next we define the network payoff function  $\Theta_i: \mathbb{G}^N \rightarrow \mathbb{R}_+$  with  $\Theta_i(g) = \sum_{j \in N_i(g)} \theta_i(ij)$ , where  $\theta_i$  is the link benefit function for player  $i$ . The resulting network payoff function  $\Theta$  is called a *link-based network payoff function*.

In this simple model, benefits are only generated from the direct links of a certain individual with other individuals. There are no benefits from being connected to players beyond one's direct neighbors in the network. Thus, there are *no* spillovers in the network. The most immediate example of such link-based payoffs are profits generated by trade relationships between buyers and sellers in a market.

We investigate the properties of this link-based network payoff structure to illustrate the relationships between the different concepts. The link-based payoff structure in this application reflects in particular the benefits obtained from having links with direct neighbors. Interestingly this simple payoff structure is shown to have some insightful properties.

## 5.1 Two-sided link formation costs

First we discuss link-based benefits in the setting of the standard model with two-sided link formation costs. It turns out that this particular case has some interesting and illustrative properties.

**Claim 5.1** *Consider the link-based network payoff function  $\Theta$  based on the link benefit functions  $\theta_i: L_i \rightarrow \mathbb{R}_+$ . Let  $c \geq 0$  be the link formation cost parameter.*

*For network payoff function  $\Theta$  the individually stable networks with two-sided link formation costs are given by  $g \subset \{ij \in g_N \mid \min\{\theta_i(ij), \theta_j(ij)\} \geq \max\{c_{ij}, c_{ji}\}\}$ .*

In other words, individually stable networks consist of links for which the formation costs are covered by their direct benefits. This is exactly as one would expect within this setting.

The properties of the link-based network payoff functions also include a relationship with potential games. This is the subject of our next proposition. We remark that we call the link-based benefit structure *mutual* if there exists a link-based benefit function  $\theta: g_N \rightarrow \mathbb{R}_+$  such that  $\theta_i(ij) = \theta_j(ij) = \theta(ij)$  for all players  $i, j \in N$  with  $i \neq j$ . We are able to show that mutual link-based benefits generate a potential game.

**Proposition 5.2** *If for every player  $i \in N$  the link-based network payoff function  $\Theta_i(g) = \sum_{ij \in L_i(g)} \theta(ij)$  is founded on a mutual link benefit function  $\theta: g_N \rightarrow \mathbb{R}_+$ , then the standard model with two-sided link formation costs is a potential game.*

*Furthermore, in this game the potential maximizing individually stable networks are given by  $g = \widehat{g}_\theta \cup h$ , where  $\widehat{g}_\theta = \{ij \in g_N \mid \theta(ij) > c_{ij} + c_{ji}\}$  and  $h \subset \{ij \in g_N \mid \theta(ij) = c_{ij} + c_{ji}\}$ .*

**Proof.** We proceed by constructing an appropriate interaction potential for the standard model with two-sided link formation costs. By application of Lemma 2.1 it then is established that this model has a potential.

Let  $\ell \in A^a$ . We now introduce an interaction potential for every coalition  $S \subset N$  by

$$\Phi_S(\ell_S) = \begin{cases} -\sum_{j \neq i} \ell_{ij} \cdot c_{ij} & \text{if } S = \{i\} \\ \ell_{ij} \cdot \ell_{ji} \cdot \theta(ij) & \text{if } S = \{i, j\} \\ 0 & \text{otherwise} \end{cases}$$

Observe that this is indeed an interaction potential. The function  $\Phi_{\{i\}}(\ell_i)$  depends only on the variables  $\ell_i$ . The other parts of the definition above are easily checked as well. Also, it holds that

$$\begin{aligned} \pi_i^a(\ell) &= \sum_{j \in N^d(i, g^a(\ell))} (\theta(ij) - c_{ij}) - \sum_{j \notin N^d(i, g^a(\ell))} \ell_{ij} \cdot c_{ij} = \\ &= \sum_{j \neq i} \ell_{ij} \cdot \ell_{ji} \cdot \theta(ij) - \sum_{j \neq i} \ell_{ij} \cdot c_{ij} = \\ &= \sum_{j \neq i} \Phi_{ij}(\ell_{\{i, j\}}) + \Phi_i(\ell_i) = \sum_{S \subset N, i \in S} \Phi_S(\ell_S). \end{aligned}$$

Now from Lemma 2.1 a potential of the game  $(A^a, \pi^a)$  is given by

$$Q(\ell) = \sum_{S \subset N} \Phi_S(\ell_S) = \sum_{ij \in g^a(\ell)} [\theta(ij) - c_{ij} - c_{ji}] - \sum_{ij \notin g^a(\ell)} [\ell_{ij} \cdot c_{ij} + \ell_{ji} \cdot c_{ji}].$$

Hence,  $Q$  is maximal if  $g^a(\ell) = \widehat{g}_\theta \cup h$  with  $h \subset \{ij \in g_N \mid \theta(ij) = c_{ij} + c_{ji}\}$ . ■

From Proposition 5.2 and the previous discussion of Proposition 3.1 and Corollary 3.2, we can draw some important conclusions.

First, in game theory the set of potential maximizers is usually considered to be an important and useful refinement of the Nash equilibrium concept. (More specifically, we refer to Slikker, Dutta, van den Nouweland, and Tijs (2000) for the relationship between network formation and potential maximizers.) Proposition 5.2, however, shows that for mutual link-based benefits and two-sided link formation costs, the set of potential maximizing networks may not be the most interesting class of networks. Indeed, for mutual link-based network payoffs, the largest individually stable network is given by  $g_\theta^* = \{ij \in g_N \mid \theta(ij) > \max\{c_{ij}, c_{ji}\}\}$ . The class of networks identified in Proposition 5.2 does not contain this network. Contrary, this class of networks, in fact, does not have any significantly distinguishing features. It is clear that we have to resort to other modifications of the Nash equilibrium concept in our study of the formation of non-trivial stable networks.

Second, Monderer and Shapley (1996) introduced the notion of an “improvement path” to describe an individually myopic improvement process that results in a Nash equilibrium for a potential game. In the context of the model addressed in Proposition 5.2 such processes are less useful. In particular, starting from the empty network — as the most natural starting point — these improvement paths terminate immediately, thus, rendering the discussion rather pointless. It is apparent that other behavioral rules besides individually myopic behavior have to be introduced in the analysis to support the formation of non-trivial stable networks. Nevertheless, we remark that individual stability of a network remains a basic requirement for the outcome of any game theoretic network formation process.

## 5.2 One-sided link formation costs

In this section we consider the case of one-sided link formation costs for any link-based network payoff function  $\Theta$  introduced above.

**Claim 5.3** *Consider the link-based network payoff function  $\Theta$  based on the link benefit functions  $\theta_i: L_i \rightarrow \mathbb{R}_+$ . Let  $c \geq 0$  be the link formation cost parameter.*

*For network payoff function  $\Theta$  the individually stable networks under one-sided link formation costs are given by  $g \subset \{ij \in g_N \mid \theta_i(ij) \geq c_{ij} \text{ or } \theta_j(ij) \geq c_{ji}\}$ .*

From this claim and the previous analysis it follows immediately that with link-based network payoffs, the class of individually stable networks under two-sided link formation costs is a strict subset of the class of individually stable networks under one-sided link formation costs. The claim also shows that the inclusion stated in Proposition 4.3(b) is tight in this case of link-based network benefits.

In Proposition 5.2 we discussed the class of potential maximizing networks for mutual link-based benefits and two-sided link formation costs. Here we present an analogue of that case for one-sided link formation costs.

**Proposition 5.4** *If for every  $i \in N$  the mutual link-based network payoff function  $\Theta_i(g) = \sum_{j \in N_i(g)} \theta(ij)$  is founded on the mutual link benefit function  $\theta: g_N \rightarrow \mathbb{R}_+$ , then the standard model with one-sided link formation costs is a potential game.*

*Moreover, in this case the potential maximizing individually stable networks are given by  $g = \widetilde{g}_\theta \cup h$ , where  $\widetilde{g}_\theta = \{ij \in g_N \mid \theta(ij) > \min\{c_{ij}, c_{ji}\}\}$  and  $h \subset \{ij \in g_N \mid \theta(ij) = \min\{c_{ij}, c_{ji}\}\}$ .*

**Proof.** Again we proceed by constructing an appropriate interaction potential. By application of Lemma 2.1 it is then established that this model has a potential.

Let  $(\ell, r) \in A^b$ . We now introduce an interaction potential for every coalition  $S \subset N$  as

follows

$$\Phi_S(\ell_S, r_S) = \begin{cases} -\sum_{j \in N_i(g)} \ell_{ij} \cdot c_{ij} & \text{if } S = \{i\} \\ m_{ij}(\ell, r) \cdot \theta(ij) & \text{if } S = \{i, j\} \\ 0 & \text{otherwise,} \end{cases}$$

where  $m_{ij}(\ell, r) = \max\{\ell_{ij} \cdot r_{ji}, r_{ij} \cdot \ell_{ji}\}$ . It is obvious that this defines an interaction potential. Indeed, we have

$$\begin{aligned} \pi_i^b(\ell, r) &= \sum_{j \in N_i(g)} (\theta(ij) - \ell_{ij} \cdot c_{ij}) = \\ &= \sum_{j \neq i} m_{ij}(\ell, r) \cdot \theta(ij) - \sum_{j \in N_i(g)} \ell_{ij} \cdot c_{ij} = \\ &= \sum_{j \neq i} \Phi_{ij}(\ell_{\{i,j\}}, r_{\{i,j\}}) + \Phi_i(\ell_i, r_i) = \sum_{S \subset N, i \in S} \Phi_S(\ell_S). \end{aligned}$$

Using Lemma 2.1, a potential of the standard model with one-sided link formation costs is now given by

$$Q(\ell, r) = \sum_{S \subset N} \Phi_S(\ell_S) = \sum_{ij \in g^b(\ell, r)} \theta(ij) - \sum_{ij \in g_N} [\ell_{ij} \cdot c_{ij} + \ell_{ji} \cdot c_{ji}].$$

From this it is clear that  $Q$  is maximal if  $g^a(\ell) = \widetilde{g}_\theta \cup h$  with  $h \subset \{ij \in g_N \mid \theta(ij) = \min\{c_{ij}, c_{ji}\}\}$ . ■

Compared to the conclusion in Proposition 5.2 the assertion of Proposition 5.4 is much more interesting. It identifies exactly the class of networks that result from the formation of each profitable link, i.e., when link formation is profitable for the individual with the lowest link costs, the link is always formed. Hence, we conclude that the potential maximizer as a refinement of Nash equilibrium, is a more useful tool in explaining the formation of non-trivial networks in the context of one-sided link formation costs.

### 5.3 Some lessons from our analysis

We find an interesting contrast between the two-sided and one-sided link formation costs for the case of link-based network benefits. On the one hand, in general, all equilibria of the two-sided model are also equilibria in the one-sided model. (Proposition 4.2) On the other hand, the example of mutual link-based network benefits (Propositions 5.2 and 5.4) in which all benefits are derived only from direct links, provides interesting additional insights. Under two-sided costs we find that the potential maximizer is not a useful solution concept since it requires that individual payoffs (stemming from individual actions) must cover the link costs of both agents. Yet for one-sided costs the potential maximizer is able

to select the right Nash equilibria since in this case the potential function takes correctly into account the actions and costs of individual players only.

Our analysis also points to some questions that are worth investigating further. In this section we investigated a particular example of link-based network payoffs. Can we extend this class of payoff functions to a more general family with similar properties? Especially it would be interesting to identify larger classes of network payoff functions that generate potential functions within the setting of link formation games. We refer to Durieu, Haller, and Solal (2004) for the discussion of a more general setting in which benefits are principally link-based. In that case the network formation game also has a potential.

## 6 Proofs of the main results

### 6.1 Proof of Proposition 2.4

Obviously from the definitions it follows that in general  $\mathcal{L}_s(\varphi) \subset \mathcal{L}(\varphi)$ .

**Only if:** Suppose that  $g \in \mathcal{L}(\varphi)$  and that  $\varphi_i$  is not network convex on  $g$  for some  $i \in N$  and some link set  $h \subset L_i(g)$ . We show that  $g \notin \mathcal{L}_s(\varphi)$ .

Indeed, from the hypothesis that  $g$  is link deletion proof, we know that  $D_i(g, ij) \geq 0$  for every  $ij \in L_i(g)$ . Then for  $h$  it has to be true that since  $\sum_h D_i(g, ij) \geq 0$ ,  $D_i(g, h) < 0$ . But then this implies that player  $i$  would prefer to sever all links in  $h$ . Hence,  $g$  cannot be strong link deletion proof, i.e.,  $g \notin \mathcal{L}_s(\varphi)$ .

**If:** Let  $g \in \mathcal{L}(\varphi)$  and assume that  $\varphi$  is network convex on  $g$ . Then for every player  $i \in N$  and link  $ij \in L_i(g)$  it has to hold that  $D_i(g, ij) \geq 0$  due to link deletion proofness of  $g$ . In particular, for any link set  $h \subset L_i(g)$ :  $\sum_h D_i(g, \cdot) \geq 0$ . Now by network convexity this implies that  $D_i(g, h) \geq 0$  for every link set  $h \subset L_i(g)$ . In other words,  $g$  is strong link deletion proof, i.e.,  $g \in \mathcal{L}_s(\varphi)$ .

This completes the proof of Proposition 2.4.

### 6.2 Proof of Proposition 3.1

*If.* Suppose that  $g \subset g_N$  is strong deletion proof with respect to the given payoff function  $\varphi^a$ . Define  $\ell^g \in A^a$  by  $\ell_{ij}^g = 1$  if and only if  $ij \in g$ . Now  $g^a(\ell^g) = g$ . We now show that  $\ell^g$  is a Nash equilibrium in  $(A^a, \pi^a)$ . Indeed, from equation (5),

$$\pi_i^a(\ell^g) = \varphi_i(g^a(\ell^g)) - \sum_{j \neq i} \ell_{ij}^g \cdot c_{ij} = \varphi_i(g) - \sum_{j \neq i, ij \in g} c_{ij} = \varphi_i^a(g) \quad (9)$$

Let  $l_i \neq \ell_i^g$  and define  $h_i = \{ij \in g_N \mid \ell_{ij}^g = 1 \text{ and } l_{ij} = 0\}$ . Then it follows that  $h_i = \{ij \in g \mid l_{ij} = 0\}$  and  $g^a(l_i, \ell_{-i}^g) = g \setminus h_i$ . From this, equation (9), and strong link deletion proofness of  $g$  it now follows that

$$\pi_i^a(l_i, \ell_{-i}^g) = \varphi_i^a(g \setminus h_i) \leq \varphi_i^a(g) = \pi_i^a(\ell).$$

*Only if.* Suppose that  $g$  is individually stable. Then, with the definitions above,  $\ell^g$  is a Nash equilibrium in  $(A^a, \pi^a)$ . Let  $M \subset N_i(g)$  and let  $h_M = \{ij \in g \mid j \in M\}$  be the set of all links connecting  $i$  to the players in the set  $M$ . Define  $L_i \in A_i^a$  by

$$L_{ij} = \begin{cases} 1 & \text{if } ij \in g \setminus h_M; \\ 0 & \text{otherwise.} \end{cases}$$

Then with the above it can be concluded that

$$\begin{aligned} \pi_i^a(\ell_{-i}^g, L_i) &= \varphi_i(g \setminus h_M) - \sum_{j \neq i, ij \in g \setminus h_M} c_{ij} = \\ &= \varphi_i^a(g \setminus h_M) \leq \pi_i^a(\ell^g) = \varphi_i^a(g). \end{aligned}$$

From this it can be concluded that  $g$  is indeed strong link deletion proof.

This completes the proof of Proposition 3.1.

### 6.3 Proof of Proposition 4.2

Let  $\widehat{\ell} \in A^a$  be a Nash equilibrium strategy tuple in the standard model with two-sided link formation costs. We construct with  $\widehat{\ell}$  a strategy tuple in the standard model with one-sided link formation generating exactly the same network  $g^a(\widehat{\ell})$  and show that this is a Nash equilibrium in the model with one-sided link formation costs.<sup>9</sup>

First we remark that by the Nash equilibrium requirements on  $\widehat{\ell}$  without loss of generality we may assume that for any  $ij \in g_N$  either  $\widehat{\ell}_{ij} = \widehat{\ell}_{ji} = 1$ , or  $\widehat{\ell}_{ij} = \widehat{\ell}_{ji} = 0$ . In the first case we have that  $ij \in g^a(\widehat{\ell})$  and in the second case we have that  $ij \notin g^a(\widehat{\ell})$ .

For  $\widehat{\ell}$  we define  $(\ell, r) \in A^b$  such that

(A)  $\ell_{ij} = 1$  and  $r_{ij} = 0$  if and only if  $\widehat{\ell}_{ij} = \widehat{\ell}_{ji} = 1$  and

- $c_{ij} < c_{ji}$ , or
- $c_{ij} = c_{ji}$  and  $i < j$ .

(B)  $\ell_{ij} = 0$  and  $r_{ij} = 1$  if and only if  $\widehat{\ell}_{ij} = \widehat{\ell}_{ji} = 1$  and

---

<sup>9</sup>The cases excluded here are for  $c_{ij} = 0$  and/or  $c_{ji} = 0$ . These cases are trivial and no explicit analysis is required.

- $c_{ij} > c_{ji}$ , or
- $c_{ij} = c_{ji}$  and  $i > j$ .

(C)  $\ell_{ij} = r_{ij} = 0$  if and only if  $\widehat{\ell}_{ij} = \widehat{\ell}_{ji} = 0$ .

So,  $(\ell, r) \in A^b$  describes that the lowest link formation cost is paid for the formation of every link  $ij \in g^a(\widehat{\ell}) = g^b(\ell, r)$ .

We now show that  $(\ell, r)$  is indeed a Nash equilibrium of the standard model with one-sided link formation costs.

Let  $(L_i, R_i) \in A_i^b$  be such that  $(L_i, R_i) \neq (\ell_i, r_i)$ . Now we define  $\hat{L}_{ij} = 1$  if and only if  $L_{ij} = 1$  or  $R_{ij} = r_{ij} = 1$ . Otherwise  $\hat{L}_{ij} = 0$ .

Now it holds that  $ij \in g^a(\widehat{\ell}_{-i}, \hat{L}_i)$  if and only if  $\widehat{\ell}_{ij} = \hat{L}_{ij} = 1$  if and only if

1.  $\ell_{ij} = L_{ij} = 1$ ,
2.  $r_{ji} = L_{ij} = 1$ , or
3.  $r_{ij} = R_{ij} = \ell_{ji} = 1$ .

Case 1 implies that  $ij \notin g^b(\ell_{-i}, r_{-i}; L_i, R_i)$ , while cases 2 and 3 imply that  $ij \in g^b(\ell_{-i}, r_{-i}; L_i, R_i)$ . This in turn implies — together with the construction that  $r_{ij} = 0$  implies that  $\ell_{ji} = 0$  — that

$$g^b(\ell_{-i}, r_{-i}; L_i, R_i) \subset g^a(\widehat{\ell}_{-i}, \hat{L}_i) \subset g^a(\widehat{\ell}). \quad (10)$$

Hence, we may conclude from this that

$$\begin{aligned} \pi^b(\ell_{-i}, r_{-i}; L_i, R_i) &= \varphi_i(g^b(\ell_{-i}, r_{-i}; L_i, R_i)) - \sum_{j \neq i} L_{ij} \cdot c_{ij} \\ &= \varphi_i(g^b(\ell_{-i}, r_{-i}; L_i, R_i)) - \sum_{ij \in g^b(\ell_{-i}, r_{-i}; L_i, R_i)} c_{ij} + \sum_{j \neq i} R_{ij} \cdot r_{ij} \cdot c_{ij} \\ &\leq \varphi_i(g^a(\widehat{\ell})) - \sum_{ij \in g^a(\widehat{\ell})} c_{ij} + \sum_{j \neq i} R_{ij} \cdot r_{ij} \cdot c_{ij} \\ &= \varphi_i(g^a(\widehat{\ell})) - \sum_{j \neq i} \ell_{ij} \cdot c_{ij} - \sum_{j \neq i} r_{ij} \cdot c_{ij} + \sum_{j \neq i} R_{ij} \cdot r_{ij} \cdot c_{ij} \\ &\leq \varphi_i(g^b(\ell, r)) - \sum_{j \neq i} \ell_{ij} \cdot c_{ij} = \pi_i^b(\ell, r), \end{aligned}$$

where the first inequality follows from Proposition 3.1 and (10). The second inequality follows from the fact that  $\sum_{j \neq i} r_{ij} \cdot c_{ij} \geq \sum_{j \neq i} R_{ij} \cdot r_{ij} \cdot c_{ij}$ .

The above shows that  $(\ell, r)$  indeed is a Nash equilibrium with regard to the payoff function  $\pi^b$ . Thus,  $g^a(\widehat{\ell})$  is supported as a individually stable network in the standard model with one-sided link formation costs.

This completes the proof of Proposition 4.2.

## 6.4 Proof of Proposition 4.3(a)

Let  $g$  be a strong link deletion proof network under the net payoff function  $\varphi^b$ .

With  $g$  we define the strategy tuple  $(\ell^g, r^g) \in A^b$  as follows:  $\ell_{ij}^g = r_{ji}^g = 1$  if  $ij \in g$  and  $c_{ij} < c_{ji}$ , and  $\ell_{ij}^g = r_{ji}^g = 0$  otherwise.

It is clear that  $(\ell^g, r^g)$  describes the cost minimizing link formation scheme that supports  $g$ , i.e.,  $g^b(\ell^g, r^g) = g$ . We proceed by showing that  $(\ell^g, r^g) \in \text{NE}(A^b, \pi^b)$ . First, remark that

$$\begin{aligned} \pi_i^b(\ell^g, r^g) &= \varphi_i(g^b(\ell^g, r^g)) - \sum_{j \neq i} \ell_{ij}^g \cdot c_{ij} \\ &= \varphi_i(g) - \sum_{j \in N_i(g): c_{ij} < c_{ji}} c_{ij} = \varphi_i^b(g). \end{aligned}$$

Let  $(L_i, R_i) \in A_i^b$  such that  $(L_i, R_i) \neq (\ell_i^g, r_i^g)$ . We now define

$$M = \{j \in N_i(g) \mid L_{ij} = r_{ij}^g = 0\} \cup \{j \in N_i(g) \mid R_{ij} = \ell_{ij}^g = 0\} \neq \emptyset.$$

Then for  $h_M = \{ij \in g \mid j \in M\}$  it is clear that  $g^b(\ell_{-i}^g, r_{-i}^g; L_i, R_i) = g \setminus h_M$ .

From the properties of  $(\ell^g, r^g)$  and the above it follows that  $j \in N^d(i, g \setminus h_M)$  if and only if  $[L_{ij} = \ell_{ij}^g = 1 \text{ and } r_{ij}^g = 0]$  or  $[R_{ij} = r_{ij}^g = 1 \text{ and } \ell_{ij}^g = 0]$ . In the first case  $c_{ij} < c_{ji}$  and in the latter  $c_{ij} > c_{ji}$ .

From this it follows that

$$\sum_{j \in N^d(i, g \setminus h_M)} L_{ij} \cdot c_{ij} \geq \sum_{j \in N^d(i, g \setminus h_M): c_{ij} < c_{ji}} c_{ij} \quad (11)$$

Hence,

$$\begin{aligned} \pi_i^b(\ell_{-i}^g, r_{-i}^g; L_i, R_i) &= \varphi_i(g^b(\ell_{-i}^g, r_{-i}^g; L_i, R_i)) - \sum_{j \neq i} L_{ij} \cdot c_{ij} \leq \\ &\leq \varphi_i(g \setminus h_M) - \sum_{j \in N^d(i, g \setminus h_M)} L_{ij} \cdot c_{ij} \leq \\ &\leq \varphi_i(g \setminus h_M) - \sum_{j \in N^d(i, g \setminus h_M): c_{ij} < c_{ji}} c_{ij} \leq \\ &\leq \varphi_i^b(g) = \pi_i^b(\ell^g, r^g), \end{aligned}$$

where the second inequality follows from (11) and the third inequality from the hypothesis that  $g$  is strong link deletion proof with respect to  $\varphi^b$ .

Since this holds for all  $i \in N$  we conclude that  $(\ell^g, r^g)$  is indeed a Nash equilibrium in  $(A^b, \pi^b)$ .

This completes the proof of Proposition 4.3(a).



## References

- BALA, V., AND S. GOYAL (2000): “A Non-Cooperative Model of Network Formation,” *Econometrica*, 68, 1181–1230.
- BELLEFLAMME, P., AND F. BLOCH (2004): “Market Sharing Agreements and Collusive Networks,” *International Economic Review*, in press.
- BLOCH, F., AND M. O. JACKSON (2004): “The Formation of Networks with Transfers among Players,” Working Paper, GREQAM, Université d’Aix-Marseille, Marseille, France.
- BRUECKNER, J. K. (2003): “Friendship Networks,” Working paper, University of Illinois, Urbana-Champaign, IL.
- CALVÓ-ARMENGOL, A., AND R. ILKILIÇ (2004): “Pairwise Stability and Nash Equilibria in Network Formation,” Working Paper, Universitat Autònoma, Barcelona, Spain.
- DURIEU, JACQUES, M., H. HALLER, AND P. SOLAL (2004): “Nonspecific Networking,” Working paper, Virginia Tech, Blacksburg, VA.
- DUTTA, B., AND M. O. JACKSON (2000): “The Stability and Efficiency of Directed Communication Networks,” *Review of Economic Design*, 5, 251–272.
- DUTTA, B., AND M. O. JACKSON (2003): “On the Formation of Networks and Groups,” in *Models of Strategic Formation of Networks and Groups*, ed. by B. Dutta, and M. O. Jackson, chap. 1. Springer Verlag, Heidelberg, Germany.
- DUTTA, B., A. VAN DEN NOUWELAND, AND S. TIJS (1998): “Link Formation in Cooperative Situations,” *International Journal of Game Theory*, 27, 245–256.
- FRIEDKIN, N. (1980): “A Test of Structural Features of Granovetter’s Strength of Weak Ties Theory,” *Social Networks*, 2, 411–422.
- GARRATT, R., AND C.-Z. QIN (2003): “On Cooperation Structures Resulting From Simultaneous Proposals,” *Economics Bulletin*, 3(5), 1–9.
- GILLES, R. P., S. CHAKRABARTI, S. SARANGI, AND N. BADASYAN (2004): “The Role of Middlemen in Efficient and Strongly Pairwise Stable Networks,” CentER Discussion Paper 2004-64, Center for Economic Research, Tilburg University, Tilburg, The Netherlands.
- GRANOVETTER, M. (1973): “The Strength of Weak Ties,” *American Journal of Sociology*, 78, 1360–1380.
- HART, S., AND A. MAS-COLELL (1989): “Potential, Value, and Consistency,” *Econometrica*, 57, 589–614.
- JACKSON, M. O. (2003): “A Survey of Models of Network Formation: Stability and Efficiency,” Working paper, CalTech, Pasadena, CA.
- JACKSON, M. O., AND A. WATTS (2002): “The Existence of Pairwise Stable Networks,” *Seoul Journal of Economics*, 14(3), 299–321.

- JACKSON, M. O., AND A. WOLINSKY (1996): "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, 71, 44–74.
- KRANTON, R., AND D. MINEHART (2001): "A Theory of Buyer-Seller Networks," *American Economic Review*, 91, 301–332.
- MONDERER, D., AND L. S. SHAPLEY (1996): "Potential Games," *Games and Economic Behavior*, 14, 124–143.
- MYERSON, R. B. (1977): "Graphs and Cooperation in Games," *Mathematics of Operations Research*, 2, 225–229.
- (1991): *Game Theory: Analysis of Conflict*. Harvard University Press, Cambridge, MA.
- NEWMAN, M. (2003): "The structure and function of complex networks," *SIAM Review*, 45, 167–256.
- (2004): "Detecting community structure in networks," *European Physical Journal, B*, in press.
- NEWMAN, M., AND M. GIRVAN (2004): "Finding and evaluating community structure in networks," *Physical Review E*, in press.
- SLIKKER, M. (2000): "Decision Making and Cooperation Structures," Ph.D. thesis, Tilburg University, Tilburg, The Netherlands.
- SLIKKER, M., B. DUTTA, A. VAN DEN NOUWELAND, AND S. TIJS (2000): "Potential Maximizers and Network Formation," *Mathematical Social Sciences*, 39, 55–70.
- SLIKKER, M., R. P. GILLES, H. NORDE, AND S. TIJS (2004): "Directed Networks, Payoff Properties, and Hierarchy Formation," *Mathematical Social Sciences*, in press.
- SLIKKER, M., AND A. VAN DEN NOUWELAND (2000): "Network Formation Models with Costs for Establishing Links," *Review of Economic Design*, 5, 333–362.
- TYLER, J. R., D. M. WILKINSON, AND B. A. HUBERMAN (2003): "Email as Spectroscopy: Automated Discovery of Community Structure within Organizations," arXiv:cond-mat/0303264.
- UI, T. (2000): "A Shapley Value Representation of Potential Games," *Games and Economic Behavior*, 31, 121–135.
- VAN DEN BRINK, R., AND R. P. GILLES (2003): "Explicit and Latent Authority in Hierarchical Organizations," Working paper, Department of Economics, Virginia Tech, Blacksburg, Virginia.
- VAN DEN NOUWELAND, A. (1993): "Games and Graphs in Economic Situations," Ph.D. thesis, Tilburg University, Tilburg, The Netherlands.

WATTS, D. J. (2003): *Six Degrees: The Science of a Connected Age*. W.W. Norton & Co., New York, NY.

WELLMAN, B., P. CARRINGTON, AND A. HALL (1988): "Networks as Personal Communities," in *Social Structures: A Network Approach*, ed. by B. Wellman, and S. Berkowitz. Cambridge University Press, Cambridge, MA.