



**HAL**  
open science

## Competition for the access to and use of information in networks

Philipp Möhlmeier, Agnieszka Rusinowska, Emily Tanimura

► **To cite this version:**

Philipp Möhlmeier, Agnieszka Rusinowska, Emily Tanimura. Competition for the access to and use of information in networks. 2016. halshs-01316936

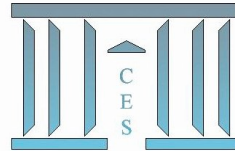
**HAL Id: halshs-01316936**

**<https://shs.hal.science/halshs-01316936v1>**

Submitted on 17 May 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**Competition for the access to and use  
of information in networks**

Philipp MÖHLMEIER, Agnieszka RUSINOWSKA, Emily TANIMURA

**2016.33**



# Competition for the access to and use of information in networks

Philipp Möhlmeier<sup>1</sup>, Agnieszka Rusinowska<sup>2</sup>, and Emily Tanimura<sup>3</sup>

<sup>1</sup> Bielefeld University, BiGSEM, Center for Mathematical Economics, Bielefeld, Germany

[philipp.moehlmeier@uni-bielefeld.de](mailto:philipp.moehlmeier@uni-bielefeld.de)

<sup>2</sup> Paris School of Economics - CNRS, Université Paris I Panthéon-Sorbonne

Centre d'Economie de la Sorbonne

106-112 Bd de l'Hôpital, 75647 Paris, France

Phone: 0033 144 078 212, Fax: 0033 144 078 301

[agnieszka.rusinowska@univ-paris1.fr](mailto:agnieszka.rusinowska@univ-paris1.fr)

<sup>3</sup> Université Paris I Panthéon-Sorbonne, Centre d'Economie de la Sorbonne

[Emily.Tanimura@univ-paris1.fr](mailto:Emily.Tanimura@univ-paris1.fr)

**Abstract.** In a network formation framework, where payoffs reflect an agent's ability to access information from direct and indirect contacts, we integrate negative externalities due to connectivity associated with two types of effects: competition for the access to information, and rivalrous use of information. We consider two separate models to capture the first and the second situations, respectively. In the first model we assume that information is a non-rivalrous good but that there is competition for the access to information, for example because an agent with many contacts must share his time between them and thus has fewer opportunities to pass on information to each particular contact. The main idea is that the probability that each neighbor receives the information decreases with the number of contacts the sender has. In the second model we assume that there is not competition for the access to information but that the use of information is rivalrous. In this case, it is assumed that when people receive the information before me, the harmful effect is greater than when others receive the information at the same time as myself. Our results concern pairwise stability and efficiency in both models and allow us to compare and contrast the effects of two kinds of competition for information.

**JEL Classification:** D85, C70

**Keywords:** network formation, connections model, information, negative externalities, pairwise stability, efficiency

## 1 Introduction

In addition to public information, diffused by sources such as the media, and available to everyone, most of us also receive valuable information that circulates only in a restricted manner, between friends and acquaintances. Access to such decentralized information certainly matters. For example, numerous studies have shown that direct and indirect personal contacts are the most frequent providers of information that leads to finding a job (Granovetter (1973, 1974)). This highlights the importance of an individual's network of friends and acquaintances and justifies the widely recognized idea that the latter is a form of social capital.

Jackson and Wolinsky (1996) were the first to propose and analyze a setting, later generalized by Bloch and Jackson (2007), where individuals seek to maximize the benefits in terms of information flow, that they receive in a network by strategically forming their links. An underlying assumption in their framework is that a denser social network will

benefit everyone by providing more information, at least if we disregard the cost required to maintain the relationships.

However, this relies on the underlying assumption that information is what is known as a non-rivalrous good, that is, a good which, used by several individuals simultaneously, provides each of them with the same utility as if he were the sole user. At a closer look, the validity of this assumption seems to depend on the nature of the information. Going back to the case of employment opportunities, someone who learns of an attractive job opening would certainly prefer that few other people were informed. Similarly, learning of an early sale of coveted concert tickets or where there is available parking in a crowded part of town is more valuable when the information is not widely shared. In other cases, information is truly a non-rivalrous good: I am glad to be informed that rain is predicted in the afternoon because I can bring my umbrella and my well being is in no way reduced by the fact that others find out too and bring theirs. For this reason, it is natural to integrate a negative externality that captures how much an agent's utility declines when he has to share his information with others.

A second point to consider is that even in cases where the use of information is non rival, there may be competition for the access to it due to congestion effects. Agents with many contacts are less likely to spend as much time with each one of them as someone who has a small number of contacts. This may translate into a lower probability of transmitting useful information to each one of the contacts. In this case an agent is not per se unhappy that other people receive an information. He is unhappy if others receive it instead of him.

In this paper we discuss how to model the negative externalities associated with connectivity resulting from the two aforementioned effects in a network formation game framework. We consider two separate models, to capture competitive use of information and competitive access to information, respectively. In practice, the presence of both together is perfectly possible. However, separating the cases facilitates the analysis and allows us to better understand and compare the effects of each assumption.

Our main results concern pairwise stability and efficiency in these two models. In particular, our analysis sheds some light on the essential differences between the effects of competition for the access to information and the competitive use of information. We begin our analysis by determining necessary conditions for pairwise stability in the first model. This allows us to rule out some candidates for pairwise stability and to characterize the possible ones in terms of the quantity of incoming information received by the agents and their importance as intermediaries. In both models, we analyze pairwise stability of the "standard" architectures, like the star, circle, complete and empty networks, and a structure of separate pairs. Depending on model parameters, a variety of structures can be pairwise stable. By focusing on "asymptotic pairwise stability", i.e., the network structures that remain PS when the number of agents becomes very large, we get sharper predictions. Finally, we compare the social efficiency of some different network architectures in these settings.

While our paper is related to a number of other works in the literature on strategic network formation, to the best of our knowledge we are the first ones to propose a framework for studying network formation with information overflows in the case where the use of information is competitive. A model of network formation with information flow and congestion effects is studied by Charoensook (2012), although in a different

framework from the connections model (Jackson and Wolinsky (1996)) and the model we propose, since he builds on Bala and Goyal (2000) and analyzes Nash networks. The problems of competitive use or access to information are not explicitly considered in the modifications and extensions of the original connections model (Jackson and Wolinsky (1996)) that capture the issue of negative and positive externalities in networks (e.g., Billand et al. (2012), Billand et al. (2013), Buechel and Hellmann (2012), Currarini (2007), Goyal and Joshi (2006), Haller (2012), Hellmann (2013), Morrill (2011), Möhlmeier et al. (2016) and the references therein).

In the model of competition for the access to information, for reasons that will be explained when the model is presented, all paths between agents are taken into account, not only the shortest paths as is usually the case in the literature on strategic network formation. An exception is Charoensook (2012) where in case of multiple paths between two agents, the value of information sent between them is given by the optimal paths, i.e., the paths that maximize the value of information obtained via the different paths.

Another work that has a similarity with the present paper in the sense of considering all possible paths is Lim et al. (2015) who investigate a threshold model of cascades in networks. They define a cascade centrality of an agent as the expected number of switches given the agent is the seed, where the expected probability that an agent switches is equal to the sum of the degree sequence products along all the paths from each seed.

In the remainder of the paper we proceed as follows. First, preliminaries on networks and some related models are recalled in Section 2. Section 3 concerns competition for the access to information. In Section 4 we study the competition for the use of information. Section 5 presents concluding remarks and a comparison of the two kinds of competition for information. Longer proofs of the results are presented in the Appendix.

## 2 Preliminaries

First we recall some notations and definitions related to networks that will be used in our analysis; see e.g., Jackson and Wolinsky (1996); Jackson (2008). Let  $N = \{1, 2, \dots, n\}$  denote the set of players (actors, agents). A *network*  $g$  is a set of pairs  $\{i, j\}$  denoted for convenience by  $ij$ , with  $i, j \in N$ ,  $i \neq j$ ,<sup>1</sup> where  $ij$  denotes a *link* between players  $i$  and  $j$ . Nodes  $i$  and  $j$  are directly connected (in other words,  $i$  and  $j$  are neighbors) if and only if  $ij \in g$ .

The *degree*  $d_i(g)$  of agent  $i$  counts the number of links  $i$  has in  $g$ , i.e.,

$$d_i(g) = |\{j \in N \mid ij \in g\}|$$

A network  $g$  is regular if for some  $d \in \{0, 1, \dots, n-1\}$ ,  $d_i(g) = d$  for each  $i \in N$ .

We denote by  $g^\emptyset$ ,  $g^*$ ,  $g^c$  and  $g^N$  the empty network (regular network with  $d = 0$ ), the star (network in which  $d_i = n - 1$  for one node  $i$  (the center) and  $d_j = 1$  for all other (peripheral) nodes  $j \neq i$ ), the circle (regular network with  $d = 2$ ) and the complete network (regular network with  $d = n - 1$ ), respectively. The set of all possible networks  $g$  on  $N$  is denoted by  $G := \{g \mid g \subseteq g^N\}$ .

By  $g + ij$  ( $g - ij$ , respectively) we denote the network obtained by adding link  $ij$  to  $g$  (deleting link  $ij$  from  $g$ , respectively). Furthermore, by  $g_{-i}$  we denote the network obtained by deleting player  $i$  and all his links from the network  $g$ .

<sup>1</sup> We do not allow for loops in this setting.

Let  $N(g)$  ( $n(g)$ , respectively) denote the set (the number, respectively) of players in  $N$  with at least one link, i.e.,  $N(g) = \{i \mid \exists j \text{ s.t. } ij \in g\}$  and  $n(g) = |N(g)|$ .

A *path* connecting  $i_1$  and  $i_K$  is a set of distinct nodes  $\{i_1, i_2, \dots, i_K\} \subseteq N(g)$  such that  $\{i_1i_2, i_2i_3, \dots, i_{K-1}i_K\} \subseteq g$ . We denote by  $p(i_1i_K)$  a path from  $i_1$  to  $i_K$  and by  $P(i_1i_K)$  the set of all paths from  $i_1$  to  $i_K$ . We write  $j \in p(i_1i_K)$  if path  $p(i_1i_K)$  passes through  $j$ .

By  $d_{ij}(g)$  we denote the geodesic distance between  $i$  and  $j$ , i.e., the number of links in the shortest path between  $i$  and  $j$  in  $g$ . If there is no path connecting  $i$  and  $j$  in  $g$ , then we set  $d_{ij}(g) = \infty$ .

A network  $g$  is *connected* if there is a path between any two nodes in  $g$ .

The network  $g' \subseteq g$  is a *component* of  $g$  if for all  $i \in N(g')$  and  $j \in N(g')$ ,  $i \neq j$ , there exists a path in  $g'$  connecting  $i$  and  $j$ , and for any  $i \in N(g')$  and  $j \in N(g)$ ,  $ij \in g$  implies that  $ij \in g'$ .

Let  $u_i : G \rightarrow \mathbb{R}$  denote the utility for player  $i \in N$ . A network  $g \in G$  is *pairwise stable* (denoted by *PS*) if:

- (i)  $\forall ij \in g, u_i(g) \geq u_i(g - ij)$  and  $u_j(g) \geq u_j(g - ij)$  and
- (ii)  $\forall ij \notin g, \text{ if } u_i(g) < u_i(g + ij) \text{ then } u_j(g) > u_j(g + ij)$ .

A network  $g \subseteq g^N$  is *strongly efficient* (denoted by *SE*) if

$$\sum_{i \in N} u_i(g) \geq \sum_{i \in N} u_i(g') \text{ for all } g' \subseteq g^N$$

In the symmetric *connections model* by Jackson and Wolinsky (1996) the utility of each player  $i$  from network  $g$  is defined as

$$u_i^{JW}(g) = \sum_{j \neq i} \delta^{d_{ij}(g)} - cd_i(g) \quad (1)$$

where  $0 < \delta < 1$  denotes undiscounted valuation of a connection and  $c > 0$  is the cost for a direct connection.

Bloch and Jackson (2007) extend the original connections model to the distance-based model (see also Jackson (2008)), where the utility of  $i$  is given by

$$u_i^{BJ}(g) = \sum_{j \neq i} \tilde{b}(d_{ij}(g)) - cd_i(g) \quad (2)$$

with  $\tilde{b}$  nonincreasing in  $d_{ij}(g)$ .

In Möhlmeier et al. (2016) we consider the degree-distance based connections model, where the utility of agent  $i$  is defined as

$$u_i^{MRT}(g) = \sum_{j \neq i} \hat{b}(d_{ij}(g), d_j(g)) - cd_i(g) \quad (3)$$

with  $\hat{b} : \{1, \dots, n-1\}^2 \rightarrow \mathbb{R}_+$  being the net benefit that an agent receives from the direct and indirect connections. It is assumed that for all  $d_{ij}(g)$ ,  $\hat{b}(d_{ij}(g), k)$  is nonincreasing in degree  $k$ , and for all  $d_j(g)$ ,  $\hat{b}(l, d_j(g))$  is nonincreasing in distance  $l$ . Moreover, if there is no path connecting  $i$  and  $j$  in  $g$ , i.e., if  $d_{ij}(g) = \infty$ , then we set  $\hat{b}(\infty, d_j) = 0$  for every

$d_j \in \{0, 1, \dots, n-1\}$ . In particular,  $u_i^{MRT}(g^\emptyset) = 0$  for every  $i \in N$ . In this model, the value of an information that originates from an overly busy (i.e. high degree) sender is reduced. However it is not affected by high degrees of the intermediaries in the communication chain.

To illustrate some of the results, in Möhlmeier et al. (2016) we consider a particular functional form by setting  $\hat{b}(d_{ij}(g), d_j(g)) = \frac{1}{1+d_j(g)}\delta^{d_{ij}(g)}$ , i.e., the model in which the utility of agent  $i$  is given by

$$\tilde{u}_i^{MRT}(g) = \sum_{j \neq i} \frac{1}{1+d_j(g)} \delta^{d_{ij}(g)} - cd_i(g) \quad (4)$$

### 3 The model of competition for the access to information

#### 3.1 Description of the model

Consider a set of agents  $N = \{1, \dots, n\}$ . Each agent possesses a private piece of information which provides other agents with a benefit if they receive it. The information transmission between two agents takes place if a costly link between them is established. We assume two-way communication which can be modeled by an undirected network, where a link between agents  $i$  and  $j$  enables agent  $i$  to access  $j$ 's information and vice-versa. We consider two-sided link formation, i.e., mutual consent is required for forming a link.

In our first model we assume that there is competition for the access to information, for example because an agent with many contacts spends less time with each of his contacts, and thus has fewer opportunities to pass on information to the latter. On the other hand, we assume that the use of the information is non rivalrous. We want to capture the idea that the likelihood that each neighbor receives an information decreases with the number of contacts of the sender. When an agent is not sure that information will actually reach him through the shortest path to the sender, it is natural to value redundancy, that is, the agent gets utility from all paths between him and the sender not just the shortest one. Consequently, the utility of agent  $i$  in the model of competition for the access to information is defined by

$$u_i^{CA}(g) = \sum_{j \neq i} \sum_{p(ij) \in P(ij)} \delta^{|p(ij)|} \prod_{k \in p(ij), k \neq i} f(d_k(g)) - cd_i(g) \quad (5)$$

where  $p(ij)$  is a path from  $i$  to  $j$  of length  $|p(ij)|$ ,  $0 < \delta < 1$ ,  $f$  is a decreasing function of the degree,  $f(d) > 0$  for every  $d \in \mathbb{N}_+$ ,  $f(1) \leq 1$ , and  $c > 0$  is the cost for a direct connection. We have  $u_i^{CA}(g) = 0$  if  $P(ij) = \emptyset$  for each  $j \in N$ .

The modeling of the function  $f$  is important. We can interpret  $f(d)$  as the probability that a neighbor of an agent with degree  $d$  receives the information from the latter. We compute the probability of each possible path along which information can travel to an agent, and multiply it by the utility of receiving the information through that path (i.e. taking into account the path length). Arguably, the value one assigns to receiving an information from some path with a certain probability, might be lower if one also has a probability of receiving the information along other paths. This would be especially true if one was almost sure to receive the information along some of the paths. However, we

are interested in the case where communication efficiency decreases with degree, so that a large number of paths necessarily means that each one has a low probability. We will assume that agents fully value additional possibilities of receiving an information and so we do not apply a concave transformation to  $\sum_{j \neq i} \sum_{p(ij) \in P(ij)} \delta^{|p(ij)|} \prod_{k \in p(ij), k \neq i} f(d_k(g))$ .

The functional form of  $f$  determines how the level of inefficiency of the communication varies with degree. If  $f$  is convex, the decline is more rapid for small degrees, if  $f$  is concave it is the contrary. We can see the quantity  $f(d)d$  as measuring the *efficiency of transmission* to  $d$  contacts. We note that  $f(d)d \leq d$ . Cases where  $f(d)d$  is large (a particular case being  $f(d) = 1$ ) would be for example when an agent uses mailing lists to communicate, he can then successfully send information to an arbitrarily large number of contacts. If he needs to meet contacts in person, we should have  $f(d) \ll \frac{1}{d}$  for large degrees, so that  $f(d)d < 1$ , which can be interpreted as a situation where conveying the information requires a lengthy explanation so that an agent with too many contacts may not convey it successfully to any of them.

The specific choice  $f(d) = 1/d$  can be viewed as a case where the agent splits his time equally with his contacts and the probability of being informed is proportional to this time. Since we will sometimes do computations with this specific functional form, we define the utility

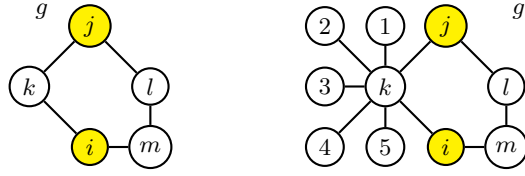
$$\tilde{u}_i^{CA}(g) = \sum_{j \neq i} \sum_{p(ij) \in P(ij)} \delta^{|p(ij)|} \prod_{k \in p(ij), k \neq i} \frac{1}{d_k} - cd_i(g) \quad (6)$$

It is clear that our model differs from the original connections model (Jackson and Wolinsky (1996)) and the degree-distance-based connections model (Möhlmeier et al. (2016)) in several respects. First of all, the function  $f$  reduces the flow of information from high degree senders and passing via high degree intermediaries. Moreover, we take into account multiple paths originating from the same sender, not only the shortest paths. To illustrate it on a simple example, consider networks  $g$  and  $g'$  in Figure 1 and the benefit of agent  $i$  from the information sent by agent  $j$ . In the original connections model, the flow of information from  $j$  to  $i$  is equally beneficial to  $i$  in  $g$  and  $g'$ , and is equal to  $\delta^2$ . The fact that agent  $k$  has fairly more connections in  $g'$  than in  $g$  and hence the information from  $j$  to  $i$  could be conveyed much easier in  $g$  than in  $g'$  is ignored in model (1). Similarly in model (3) (and in (4) which gives  $i$  the benefit  $\frac{\delta^2}{3}$  from the information sent by  $j$ ): while it does take into account the number of connections of the sender of information, it ignores the connections of agents along the paths of the flow of information, i.e., the connections of agent  $k$  in this example. The model of competition for the access to information defined in (5) and (6) assumes that the neighbors of both agents  $j$  and  $k$  have an impact on the efficiency of the information transmission between  $j$  and  $i$ . Agent  $i$  benefits more when the intermediary agent  $k$  has less contacts. In model (6) the benefits that  $i$  gets from  $j$ 's information sent via  $k$  are equal to  $\frac{\delta^2}{4}$  in  $g$  and  $\frac{\delta^2}{14}$  in  $g'$ .

Our model of competition for the access to information also differs from the original connections model in how the information flow is evaluated. It is assumed that each agent passes the piece of information he possesses to all of his contacts, i.e.,  $j$ 's information is sent to  $i$  both via  $k$  and via  $l$ . Since  $i$  cannot always be sure via which path  $j$ 's information will reach him first, not only the shortest paths that the information passes through (or the path that would give the maximum benefit), but all paths of the information flow between  $j$  and  $i$  "contribute" to the benefit. In the example given in Figure 1 this means



that the benefit for  $i$  from the information sent by  $j$  is equal to the sum of the benefits from the shortest path going through  $k$  and from the path passing through  $l$  and  $m$  (the latter being equal to  $\frac{\delta^3}{8}$  in both  $g$  and  $g'$  under model (6)). Note that in  $g'$ , for sufficiently large  $\delta$  this benefit  $\frac{\delta^3}{8}$  is greater than the benefit  $\frac{\delta^2}{14}$  from the information passed via the shortest path.



**Fig. 1.** The models of competition for information versus the connections and degree-distance-based connections models

In the next subsections we analyze pairwise stability in the model given by (5) and (6).

### 3.2 Possible and ruled-out categories of pairwise stable structures

In what follows we analyze necessary conditions for maintaining a link in a PS network in model (5). We show that the value of a link to an agent is determined by two features of this agent's structural position: on one hand the quantity of incoming information he receives and on the other hand his importance as an intermediary. The condition allows us to obtain some characterizations of possible PS structures in terms of these two features - incoming information and importance as an intermediary.

- We define the information obtained by  $j$  without using  $i$  as an intermediary in network  $g$  as:

$$I^{j-i}(g) = \sum_{k \neq j, i} \sum_{p(kj-i) \in P(kj-i)} \delta^{|p(kj-i)|} \prod_{l \in p(kj-i)} f(d_l(g)) \quad (7)$$

where  $p(kj-i)$  is a path from  $k$  to  $j$  that does not pass  $i$ .

- We define the information of  $i$  with  $j$  as an intermediary (or a sender) in network  $g - ij$  as:

$$I^{i,j}(g - ij) = \sum_{k \neq i} \sum_{p(ki,j) \in P(ki,j)} \delta^{|p(ki,j)|} \prod_{l \in p(ki,j)} f(d_l(g - ij)) \quad (8)$$

where  $p(ki,j)$  is a path from  $k$  to  $i$  that passes  $j$  and  $P(ki,j)$  denotes the set of such paths. Moreover,  $p(ji,j)$  and  $P(ji,j)$  denote simply  $p(ji)$  and  $P(ji)$ , respectively.

**Proposition 1** *Let  $ij \in g$ . Agent  $i$  prefers network  $g$  to network  $g - ij$  if and only if*

$$\delta f(d_j(g)) (1 + I^{j-i}(g)) - I^{i,j}(g - ij) \left( 1 - \frac{f(d_j(g))}{f(d_j(g) - 1)} \right) > c \quad (9)$$

**Proof:**  $u_i^{CA}(g) > u_i^{CA}(g - ij)$  is equivalent to  $\delta f(d_j(g)) + I^{j-i}(g)\delta f(d_j(g)) + I^{i,j}(g - ij)\frac{f(d_j(g))}{f(d_j(g)-1)} - c > I^{i,j}(g - ij)$  which leads to (9). ■

It results from the above formulation that if the communication technology is inefficient for high degrees ( $\lim_{d \rightarrow \infty} f(d) = 0$ ) then:

- In a PS network (unless the cost is very low), agents with a high degree must receive a high quantity of information.
- In a PS network (unless the cost is very high), if low degree agents receive a high quantity of information, they must also be important as intermediaries in the network.

One could be inclined to think that agents with a high degree always receive a lot of information, making the first remark above trivial. This is in fact not the case. Even if a node has a high degree, it does not necessarily need to get a lot of incoming information.

**Lemma 1** *Suppose that the degree  $d_i$  of every agent  $i$  in the network verifies  $\underline{d} \leq d_i \leq \bar{d}$ . Then an upper bound on value of the information received by any agent is  $\sum_{l=1}^{n-1} (\bar{d}f(\underline{d})\delta)^l$ . In particular, if  $\bar{d}f(\underline{d})\delta < 1$ , then the upper bound is independent of the network size  $n$ .*

**Proof:** The number of paths (without repetition) of length  $l$  leading to  $i$  is bounded by  $\bar{d}^l$ . Indeed,  $i$  has at most  $\bar{d}$  neighbors and each of them has at most  $\bar{d}$  neighbors and so on. Since the minimal degree of any node is  $\underline{d}$ , the value of an information sent on a path of length  $l$  is bounded by  $(f(\underline{d})\delta)^l$ . Since the maximal length of a path with no repetition is  $n - 1$ , the total value of incoming information is thus bounded by  $\sum_{l=1}^{n-1} (\bar{d}f(\underline{d})\delta)^l$ . ■

It follows from the bound above that the quantity of incoming information of a high degree node can vary greatly depending on the network structure. For example, we can apply Lemma 1 with  $\bar{d} = \underline{d} = n - 1$  to the complete network. If the communication technology is inefficient, the condition  $\bar{d}f(\underline{d})\delta < 1$  holds and so the value of the incoming information for each node in the complete network is bounded independently of  $n$ . When  $n$  is large, it is thus order of magnitude smaller than for the center of a star which receives  $(n - 1)f(1)\delta$ , and this despite identical degrees.

The following proposition can be seen to result from a bound of the incoming information an agent can receive.

**Proposition 2** *Suppose that the communication technology verifies  $f(d)d < \alpha < 1$  when  $d$  is sufficiently large. Consider a network in which each node  $i$  has a degree  $d_i$  such that  $k \leq d_i \leq \frac{k}{\delta}$ . If  $k = cn$  for  $c > 0$ , so that the degrees are of the same order of magnitude as the network size, then such a network cannot be PS when  $n$  is large.*

**Proof:** We apply Lemma 1 with  $\underline{d} = k$  and  $\bar{d} = \frac{k}{\delta}$ . The total value of incoming information is thus bounded by  $\sum_{l=1}^{n-1} (kf(k))^l$ . By assumption  $kf(k) < \alpha < 1$  for large  $k$ , and we have  $\lim_{n \rightarrow \infty} \sum_{l=1}^{n-1} (kf(k))^l = \frac{kf(k)}{1 - kf(k)}$ . Consequently, the requirement  $\delta f(d)[1 + I^{j-i}(g)] > c$  cannot hold for any  $d = cn$ , when  $n$  is large. This rules out networks with (fairly) homogeneous high degrees, i.e., degrees that are on the order of

magnitude of the total network size. ■

The condition  $k \leq d_i \leq \frac{k}{\delta}$  imposes some homogeneity in degree. The result basically says that large networks where agents have high but homogeneous degrees are not PS when communication technology is inefficient for large degrees.

To conclude, the previous results rule out some candidates for PS structures (except for very low costs):

- (1) No PS network can contain high degree nodes receiving a low level of incoming information.
- (2) No PS network can consist only of nodes with a high and fairly homogeneous degree.

The PS networks that are not ruled out by the previous results can thus belong to the following very general categories:

- (a) networks combining high degree nodes with a lot of information and low degree nodes receiving little information;
- (b) networks with only low degree nodes and a lot of information who are all important as intermediaries in the network (this structural constraint is actually rather restrictive since low degree nodes are not naturally important intermediaries);
- (c) networks combining high degree nodes and low degree nodes, where both types receive a lot of information;
- (d) networks with only low degree nodes with little information.

In the next subsection, we will study the pairwise stability of some particular networks which provide examples of structures belonging to the different categories listed above.

### 3.3 Pairwise stability of some “standard” architectures

We analyze pairwise stability of the prominent network structures that were shown to be stable in the Jackson-Wolinsky model and the degree-distance-based connections model, such as the empty network, the star and the complete network, as well as pairwise stability of the circle and some disconnected structures.

**Proposition 3** *In the model defined by (5) the following holds:*

- (i) *The empty network  $g^0$  is PS if  $f(1)\delta \leq c$*
- (ii) *The star  $g^*$  with  $n \geq 3$  is PS if*

$$f(2)\delta + f(n-1)(2f(2) - f(1))\delta^2 + (n-3)f(1)f(2)f(n-1)\delta^3 \leq c \quad \text{and}$$

$$c \leq \min(f(1)\delta, f(n-1)\delta + (n-2)f(1)f(n-1)\delta^2) \quad (10)$$

- (iii) *The complete network  $g^N$  with  $n \geq 3$  is PS if*

$$c \leq f(n-1)\delta \left[ 1 + \sum_{k=1}^{n-2} \delta^k \frac{(n-2)!}{(n-2-k)!} f^{k-1}(n-1) ((k+1)f(n-1) - kf(n-2)) \right] \quad (11)$$

(iv) The circle  $g^c$  of  $n > 3$  nodes is PS if

$$c \leq \sum_{k=1}^{n-1} f^k(2)\delta^k + f^{n-2}(2)\delta^{n-1} (f(2) - f(1)) \quad (12)$$

and for  $n$ -even

$$c \geq f(3)\delta + 2f(3) \sum_{k=1}^{n-2} f^k(2)\delta^{k+1} + 2f(3)f^{\frac{n}{2}-1}(2)\delta^{\frac{n}{2}} - 2 \sum_{k=\frac{n}{2}}^{n-1} f^k(2)\delta^k \quad (13)$$

and for  $n$ -odd

$$c \geq f(3)\delta + 2f(3) \sum_{k=1}^{n-2} f^k(2)\delta^{k+1} + f(3)f^{\frac{n-3}{2}}(2)\delta^{\frac{n-1}{2}} - f^{\frac{n-1}{2}}(2)\delta^{\frac{n-1}{2}} - 2 \sum_{k=\frac{n+1}{2}}^{n-1} f^k(2)\delta^k \quad (14)$$

(v) The structures of  $\frac{n}{2}$  separate pairs for  $n$  even and of  $\frac{n-1}{2}$  separate pairs with an isolated agent for  $n$  odd are PS if

$$f(2)\delta(1 + f(1)\delta) \leq c \leq f(1)\delta \quad (15)$$

The cost range for the stability is nonempty whenever

$$\delta \leq \frac{f(1) - f(2)}{f(1)f(2)} \quad (16)$$

See the proof in the Appendix.

The conditions for pairwise stability obviously involve the cost  $c$ , the decay  $\delta$  and the function  $f$ . Depending on the structure, pairwise stability depends on the efficiency of the communication technology for different degrees. More precisely, we only take into account  $f(1)$  for the empty network,  $f(1)$  and  $f(2)$  for the structure of separate pairs,  $f(1)$ ,  $f(2)$  and  $f(3)$  for the circle,  $f(1)$ ,  $f(2)$  and  $f(n-1)$  for the star with  $n$  nodes, and finally  $f(n-1)$  and  $f(n-2)$  for the complete graph with  $n$  nodes. If we consider the particular model defined in (6), then by inserting  $f(d) = \frac{1}{d}$  in the conditions obtained in Proposition 3 the results on pairwise stability are simplified as follows:

**Conclusion 1** *In the model defined by (6) the following holds:*

- (i) The empty network  $g^0$  is PS if  $\delta \leq c$ .
- (ii) The star  $g^*$  with  $n \geq 3$  is PS if

$$\frac{\delta}{2} + \frac{(n-3)\delta^3}{2(n-1)} \leq c \leq \frac{\delta}{n-1} + \frac{(n-2)\delta^2}{n-1} \quad (17)$$

For  $n = 3$  this cost range is never empty. For  $n > 3$ , this cost range is nonempty whenever  $\delta \geq \frac{n-2-\sqrt{2n-5}}{n-3}$ .

(iii) The complete network  $g^N$  with  $n \geq 3$  is PS if

$$c \leq \frac{\delta}{n-1} \sum_{k=0}^{n-3} \left( \frac{\delta}{n-1} \right)^k \frac{(n-3)!}{(n-3-k)!} \quad (18)$$

(iv) The circle  $g^c$  of  $n > 3$  nodes is PS if

$$c \leq \sum_{k=1}^{n-2} \left( \frac{\delta}{2} \right)^k \quad (19)$$

and for  $n$ -even

$$c \geq \frac{\delta}{3} + \frac{2}{3}\delta \sum_{k=1}^{n-2} \left( \frac{\delta}{2} \right)^k - \frac{2}{3} \left( \frac{\delta}{2} \right)^{\frac{n}{2}} - 2 \sum_{k=\frac{n}{2}+1}^{n-1} \left( \frac{\delta}{2} \right)^k \quad (20)$$

and for  $n$ -odd

$$c \geq \frac{\delta}{3} + \frac{2}{3}\delta \sum_{k=1}^{n-2} \left( \frac{\delta}{2} \right)^k - \frac{1}{3} \left( \frac{\delta}{2} \right)^{\frac{n-1}{2}} - 2 \sum_{k=\frac{n+1}{2}}^{n-1} \left( \frac{\delta}{2} \right)^k \quad (21)$$

(v) The structures of  $\frac{n}{2}$  separate pairs for  $n$  even and of  $\frac{n-1}{2}$  separate pairs with an isolated agent for  $n$  odd are PS if

$$\frac{\delta}{2}(1 + \delta) \leq c \leq \delta \quad (22)$$

This cost range is always nonempty as condition (16) becomes  $\delta \leq 1$ .

For stability of the star, from (10) we have

$$\frac{\delta}{2} + \frac{(n-3)\delta^3}{2(n-1)} \leq c \leq \min \left( \delta, \frac{\delta}{n-1} + \frac{(n-2)\delta^2}{n-1} \right)$$

but since  $\delta > \frac{\delta}{n-1} + \frac{(n-2)\delta^2}{n-1}$  for all  $\delta \in (0, 1)$ , we get (17).

Proposition 3 and Conclusion 1 present conditions for pairwise stability of the network structures of  $n$  nodes, and the conditions usually depend on  $n$ , with the exception of the empty star and the structure of separate pairs. It is therefore interesting to see if the stability of these structures is preserved when the number of agents becomes very large. For instance, from Conclusion 1(ii) the cost range for stability of the star decreases with the number of agents and becomes empty when  $n \rightarrow \infty$ . The same remark holds for the stability of the complete network. The next subsection is devoted to the analysis of this issue.

### 3.4 Asymptotic pairwise stability

As a tool for further selection, we will be interested in the networks that are PS when  $n$  is large, that is, structures that remain plausible when the number of agents is large. For this purpose we introduce a notion of asymptotic (with respect to network size) pairwise stability.

**Definition 1** Let  $S$  be some network structure (e.g., star, complete graph, circle, ...). We say that the structure  $S$  is asymptotically pairwise stable (APS) with respect to the utility function  $u$  if

- it is asymptotically well defined, i.e., we can define a sequence of networks  $(g_{n_k})_{k \geq 1}$  of strictly increasing size  $n_k$  such that every network  $g_{n_k}$  has the structure  $S$ , and
- there exist fixed admissible parameters of the utility functions  $(u_i)_{i=1}^n$  such that for all  $i, j, i \neq j$

$$\lim_{n \rightarrow +\infty} (u_i(g_n) - u_i(g_n - ij)) \geq 0 \quad (23)$$

and

$$\text{if } \lim_{n \rightarrow +\infty} (u_i(g_n + ij) - u_i(g_n)) > 0 \text{ then} \quad (24)$$

$$\lim_{n \rightarrow +\infty} (u_j(g_n + ij) - u_j(g_n)) \leq 0 \quad (25)$$

The set of admissible specifications (parameters) of the utility function for which the network is APS is the asymptotic stability range of the networks.

**Remark 1** In the model (5) the asymptotic stability range is  $(c, \delta, f)$ , i.e., it is determined by the cost  $c > 0$ ,  $0 < \delta < 1$  and a function  $f$  defined in (5), i.e., verifying  $f(d) > 0$  for all  $d \in \mathbb{N}_+$  and  $f(1) \leq 1$ .

**Proposition 4** In the model (5) the following holds:

- (i) The empty network  $g^0$  is APS whenever it is PS, i.e., if  $f(1)\delta \leq c$ .
- (ii) The star  $g^*$  with  $n \geq 3$  is not APS for inefficient communication technology, more precisely, for a function  $f$  such that  $\lim_{n \rightarrow +\infty} f(n)n = 0$ , for any  $\delta < 1$ .  
If a function  $f$  satisfies  $0 < \lim_{n \rightarrow +\infty} f(n)n \leq 1$ , then there exists a nonempty positive cost range for which the star  $g^*$  is APS if and only if there exists  $\delta \in (0, 1)$  such that

$$1 \geq \lim_{n \rightarrow +\infty} f(n)n \geq \frac{f(2)}{\delta f(1)(1 - \delta f(2))} \quad (26)$$

In particular, for model (6), i.e., when  $f(d) = 1/d$ , such a nonempty cost range does not exist.

- (iii) The complete network  $g^N$  is not APS for inefficient communication technology, more precisely, for a function  $f$  such that  $\lim_{n \rightarrow +\infty} f(n)n \leq 1$ , for any  $\delta < 1$ .
- (iv) There exists a nonempty positive cost range for which the circle  $g^c$  is APS if and only if

$$\delta \leq \frac{f(2) - f(3)}{f(2)f(3)} \quad (27)$$

In particular, for model (6) such a nonempty cost range exists for any  $0 < \delta < 1$ .

- (v) The structures of  $\frac{n}{2}$  separate pairs for  $n$  even and of  $\frac{n-1}{2}$  separate pairs with an isolated agent for  $n$  odd are APS whenever they are PS, i.e., when (16) is satisfied. In particular, for model (6) such a nonempty cost range exists for any  $0 < \delta < 1$ .

See the proof in the Appendix.

Let us return to the possible and “ruled-out” PS categories of networks listed in Section 3.2. The star is an example of a PS structure in the category of networks consisting of high degree - high information nodes and low degree - low information nodes (structure (a)). As shown in Proposition 4(ii) it can be APS if communication efficiency declines moderately but not too rapidly in degree. For an inefficient communication technology the cost range for which the complete network is PS becomes vanishingly small as the network size grows. This result given in Proposition 4(iii) can also be deduced from Section 3.2, since  $g^N$  consists only of nodes with a high and homogeneous degree (structure (2)). The circle is an example of a PS structure with only low degree - high information who are all important as intermediaries (structure (b)). The incentive to add links is countered by the fact that each node is an important intermediary but would become less important if additional links are added. The stability of the circle is reinforced when communication efficiency decreases with degree. The separate pairs structure is an example of a PS network with only low degree nodes with low information (structure (d)).

### 3.5 Comparison with other related models

In Jackson and Wolinsky (1996), as one could expect, the empty network is PS for high cost and the complete graph if costs are low. The main finding is that the star is PS for intermediary levels of cost and decay. The degree-distance-based connections model, defined in (4), differs from the original connections model in that it reduces the value of an information originating from an overly busy sender. However, contrary to the CA model, only the busyness of the original sender and not of the intermediaries in the communication chain matters. In this model, the star remained APS, while the complete network is never APS. Similarly, under competition for information, the complete network is never APS if the communication technology is inefficient. This is not unexpected, in light of the result for model (4) since the negative effects of high degrees are stronger in the CA model where they also concern the intermediaries in the communication chain. As for the star, it can be APS in model (5), but only if the communication efficiency does not decline too rapidly with the degree.

We note however, that when a star is formed in the presence of competition for information, the distribution of benefits between center and periphery is different from that in the Jackson-Wolinsky model in the sense that it is now the center who extracts the most benefits. While in the original connections model, if the center of the star does not want to cut a link, then a periphery node will not want to cut the link either, in the model given by (6) this is not necessarily true. In other words, it is possible that the center of the star does not want to cut a link while a peripheral node prefers to do so. To see that, let  $i$  be the center of a star and  $j$  a peripheral node. Then:

$u_i^{JW}(g^*) - u_i^{JW}(g^* - ij) = \delta - c$  and  $u_j^{JW}(g^*) - u_j^{JW}(g^* - ij) = \delta + (n - 2)\delta^2 - c$ , and therefore if  $u_i^{JW}(g^*) > u_i^{JW}(g^* - ij)$  then also  $u_j^{JW}(g^*) > u_j^{JW}(g^* - ij)$ .

On the other hand, for the model (6), where  $\tilde{u}_i^{CA}(g^*) - \tilde{u}_i^{CA}(g^* - ij) = \delta - c$  and  $\tilde{u}_j^{CA}(g^*) - \tilde{u}_j^{CA}(g^* - ij) = \frac{\delta}{n-1} + \frac{(n-2)\delta^2}{n-1} - c$ , for every  $\delta$  and  $n \geq 3$ , there exists  $c$  such that  $\tilde{u}_i^{CA}(g^*) > \tilde{u}_i^{CA}(g^* - ij)$  but  $\tilde{u}_j^{CA}(g^*) < \tilde{u}_j^{CA}(g^* - ij)$ .

Under competition for information, there are in fact structures different from those that were PS in Jackson and Wolinsky (1996) that enjoy greater asymptotic stability

than the star for inefficient communication technologies. If the decline in communication efficiency going from degree two to degree three is large enough, the circle will be APS for a wide range of decay levels. The same is true for the structure with disjoint pairs if the decline in communication efficiency going from degree one to two is large enough. Thus, contrary to the model by Jackson and Wolinsky (1996), where any PS network has at most one nonempty component, in model (5) disconnected structures can now be PS and remain stable when the number of agents becomes large. We note that both for the circle and the disconnected pair structure, the (asymptotic) stability only depends on the behavior of the communication technology for small degrees.

### 3.6 Efficiency in the model

Next, we analyze (strong) efficiency in the model. To this effect, we will compare the efficiency of the circle  $g^c$ , the star  $g^*$ , the complete network  $g^N$ , and the isolated pair structure  $\tilde{g}$ . Generally, the result of this comparison will depend on the parameters. However, when the communication technology is inefficient, the circle turns out to be (asymptotically with respect to network size) more efficient than the other aforementioned structures for a wide range of costs and levels of decay under a rather plausible assumption about the behavior of the communication function  $f$ . The crucial assumption that allows the circle to outperform the other structures is that the decline in communication efficiency is very small when the number of contacts goes from one to two. We note that this only depends on the behavior of  $f$  for very small degrees.

**Proposition 5** *Let  $f(1) = 1$ . Suppose that the communication technology is inefficient for large degrees in the sense that there exists  $n_0$  such that  $(n - 1)f(n - 1) < 1$  when  $n \geq n_0$ . Then*

– for every fixed  $c > 0$  and  $0 < \delta < 1$ , we have

$$\lim_{\substack{n \rightarrow \infty, \\ f(2) \rightarrow f(1)}} \sum_{i \in N} u_i^{CA}(g^c) - \sum_{i \in N} u_i^{CA}(g^N) > 0 \quad (28)$$

$$\lim_{\substack{n \rightarrow \infty, \\ f(2) \rightarrow f(1)}} \sum_{i \in N} u_i^{CA}(g^c) - \sum_{i \in N} u_i^{CA}(g^*) > 0 \quad (29)$$

–  $\lim_{\substack{n \rightarrow \infty, \\ f(2) \rightarrow f(1)}} \sum_{i \in N} u_i^{CA}(g^c) - \sum_{i \in N} u_i^{CA}(\tilde{g}) > 0$  if and only if  $c < \frac{\delta(1+\delta)}{1-\delta}$

The proof is given in the appendix.

One could expect the circle to outperform the complete network which has a much greater number of links if the link cost is high. In fact, the circle will be (asymptotically) more efficient than the complete network and the star for *any* link cost and any level of decay, provided that the communication technology is inefficient for large degrees and that the decline in communication efficiency going from one to two contacts is small enough. More generally, it results from the inequalities that give us Proposition 5 that the closer  $f(2)$  is to  $f(1)$ , the larger the parameter range  $(c, \delta)$  for which the circle outperforms the star and the complete network. In order for the circle to be more efficient than a network of isolated pairs, it is required, in addition, that the cost is not too high and the decay



not too great. This is natural since the benefits of indirect contacts decline if decay is too high and the total cost of links is greater in the circle.

If we do not assume that  $f(2)$  is close to 1, but only that  $2f(2) > 1$ , the circle can still outperform the star and the complete network if decay is high enough.

**Proposition 6** *Let  $n > 3$ . Suppose that the communication technology is inefficient for large degrees in the sense that there exists  $n_0$  such that  $(n - 1)f(n - 1) < 1$  when  $n \geq n_0$  and  $2f(2) > 1$ . Then*

- For any positive cost  $c > 0$ ,  $g^c$  is more efficient than  $g^N$  if  $\delta < \frac{2f(2)-1}{f(2)}$
- there exists a nonempty cost range for which  $g^c$  is more efficient than  $g^*$  if

$$\delta < \frac{2nf(2) - (n - 1)(f(1) + f(n - 1))}{(n - 1)(n - 2)f(1)f(n - 1)} =: F(n)$$

Moreover, we can show that the parameter range for which the circle outperforms the star, given in the second part of Proposition 6 increases as the communication technology becomes more inefficient. This is shown in Proposition 7 below:

**Proposition 7** *Consider inefficient communication technology for large degrees such that  $\lim_{n \rightarrow +\infty} nf(n) \leq 1$ . Let  $2f(2) > 1$ . Then the more inefficient the communication technology, the greater the cost range for which  $g^c$  is more efficient than  $g^*$ , and with a very inefficient communication technology,  $g^c$  will always be more efficient than  $g^*$ .*

Indeed, consider the function  $F(n)$  defined in Proposition 6. We have:

- if  $\lim_{n \rightarrow +\infty} nf(n) = 1$ , then  $\lim_{n \rightarrow +\infty} F(n) = \frac{2f(2)-f(1)}{f(1)} \in (0, 1)$
- if  $\lim_{n \rightarrow +\infty} nf(n) = a \in (0, 1)$ , then  $\lim_{n \rightarrow +\infty} F(n) = \frac{2f(2)-f(1)}{af(1)} > \frac{2f(2)-f(1)}{f(1)} > 0$
- if  $\lim_{n \rightarrow +\infty} nf(n) = 0$ , then  $\lim_{n \rightarrow +\infty} F(n) = +\infty$

It results from Proposition 6 that the efficiency of the circle compared to complete network and the star is enhanced when the level of decay is large. At a first sight this may seem surprising since the agents are typically further away from each other in the circle than in the other two structures. However, we need to bear in mind that we count all paths not just the shortest paths. In the circle there are only two paths between any pair of agents. In the complete network the number of (typically rather long) paths between any pair of agents is large and the value of these numerous paths decreases when decay is large. In the comparison with the star, all paths going through the center of the star have a relatively low value. When decay is large, what matters is direct paths and these have a greater value in the circle under the condition  $2f(2) > 1$ . Whether or not the circle will also be more efficient than the network of disconnected pairs when decay is high will depend on the cost.

## 4 A model of competition for the use of information

### 4.1 Description of the model

We will now assume that there is no competition for the access to information but that the use of information is rivalrous. Moreover, we assume that the disutility inflicted by

those who receive the information before me is greater than the disutility inflicted by those who receive the information at the same time as myself. For the sake of simplicity, we will assume that it does not matter how many steps before me the earlier informed people got the information. In this case, the utility an agent  $i$  derives from network  $g$  is defined by

$$u_i^{CU}(g) = \sum_{\{j \neq i | d_{ij}(g) < \infty\}} b(d_{ij}(g), x_{ij}(g), y_{ij}(g)) - cd_i(g) \quad (30)$$

where  $b(d_{ij}(g), x_{ij}(g), y_{ij}(g))$  is a three variable function  $b : \mathbb{N}_+ \times \mathbb{N}^2 \rightarrow \mathbb{R}_+$  for the value of the information that  $i$  receives from  $j$ ,  $d_{ij}(g)$  is the geodesic distance from  $i$  to  $j$  in  $g$ ,  $x_{ij}(g)$  is the number of agents who are closer to  $j$  than  $i$ , and  $y_{ij}(g)$  is the number of agents who are at the same distance to  $j$  as  $i$ . It is also assumed that  $\sum_{d=1}^{+\infty} b(d, x, y) < +\infty$  for all  $x, y \in \mathbb{N}$ .

We note that if  $d_{ij}(g) = 1$ , then necessarily  $x_{ij}(d) = 0$ . Due to the fact that it is worse that agents get information before me than at the same time, we also assume

$$b(d, x, y) > b(d, x + k, y - k) \text{ for all } k \geq 1 \quad (31)$$

The function  $b$  is decreasing in each of the three variables. The level of decrease with respect to  $x$  and  $y$  captures the level of rivalry in the use, the extreme cases being the constant case: I do not care if others are informed, and the case where  $b$  reaches 0 if many others learn the information before me. Moreover, if the utility declines rapidly with distance, it seems reasonable that  $b$  should decline more rapidly with respect to agents who are closer than myself to the sender.

Before we turn to the analysis of pairwise stability in the model of competition for the use of information, let us compare the benefits from information as modeled in the different frameworks. Note that in model (30) the benefit from the information received by a neighbor  $x$ , i.e.,  $b(1, 0, d_x - 1)$ , depends only on the degree  $d_x$  of the neighbor, as in the degree-distance-based model recalled in (3), where this benefit is determined by  $\hat{b}(1, d_x)$ . In model (2) the benefit from the information sent by a neighbor is always the same and is equal to  $\tilde{b}(1)$ . For example, in Figure 1 the benefit of agent  $i$  from the information sent by his neighbor  $k$  in model (3) is given by  $\hat{b}(1, 2)$  and  $\hat{b}(1, 7)$  in  $g$  and  $g'$ , respectively, and in model (30) by  $b(1, 0, 1)$  and  $b(1, 0, 6)$  in  $g$  and  $g'$ , respectively. However, if we consider the information obtained from indirect contacts, then what is determined in the degree-distance-based connections model is “included” in the model of competition for the use of information. More precisely, the degree of a sender of information obtained by  $x$  which is not the sender’s neighbor is included in the number of agents that receive the information before  $x$  does. In Figure 1, for instance, the benefits of  $i$  from information sent by  $j$  are equal to  $\tilde{b}(2)$ ,  $\hat{b}(2, 2)$  and  $b(2, 2, 1)$  in network  $g$ , and  $\tilde{b}(2)$ ,  $\hat{b}(2, 2)$  and  $b(2, 2, 6)$  in network  $g'$ .

## 4.2 Results on pairwise stability

We start our analysis of the model defined in (30) by proving conditions for pairwise stability of the prominent structures.

**Proposition 8** *In the model defined by (30) the following holds:*

- (i) The empty network  $g^0$  is PS if  $b(1, 0, 0) \leq c$   
(ii) The star  $g^*$  with  $n \geq 3$  is PS if

$$b(1, 0, 1) - b(2, 1, n - 3) \leq c \leq \min(b(1, 0, 0), b(1, 0, n - 2) + (n - 2)b(2, 1, n - 3)) \quad (32)$$

- (iii) The complete network  $g^N$  with  $n \geq 3$  is PS if

$$c \leq b(1, 0, n - 2) - b(2, n - 2, 0) \quad (33)$$

- (iv) The circle  $g^c$  of  $n > 3$  nodes is PS if  
for  $n$ -even

$$c \leq \sum_{k=1}^{\frac{n}{2}-1} b(k, 2k - 2, 1) - \sum_{k=\frac{n}{2}+1}^{n-1} b(k, n - 2, 0) \quad (34)$$

and for  $n$ -even<sup>2</sup>,  $n > 6$

$$c \geq b(1, 0, 2) - b\left(\frac{n}{2}, n - 2, 0\right) + 2 \sum_{k=2}^{\lfloor \frac{n}{4} \rfloor} b(k, 2k - 2, 2) - 2 \sum_{k=\lceil \frac{n}{4} \rceil + 1}^{\frac{n}{2}-1} b(k, 2k - 2, 1) \quad (35)$$

$c \geq b(1, 0, 2) - b\left(\frac{n}{2}, n - 2, 0\right)$  for  $n \in \{4, 6\}$   
and for  $n$ -odd

$$c \leq \sum_{k=1}^{\frac{n-1}{2}} b(k, 2k - 2, 1) - \sum_{k=\frac{n+1}{2}}^{n-1} b(k, n - 2, 0) \quad (36)$$

and for  $n$ -odd,  $n > 7$

$$\begin{aligned} c \geq & b(1, 0, 2) + \sum_{k=2}^{\lfloor \frac{n-1}{4} \rfloor} b(k, 2k - 2, 2) + \sum_{k=2}^{\lfloor \frac{n}{4} \rfloor} b(k, 2k - 2, 2) - \\ & - \sum_{k=\lceil \frac{n-1}{4} \rceil + 1}^{\frac{n-1}{2}} b(k, 2k - 2, 1) - \sum_{k=\lceil \frac{n}{4} \rceil + 1}^{\frac{n-1}{2}} b(k, 2k - 2, 1) \end{aligned} \quad (37)$$

$c \geq b(1, 0, 2) + b(2, 2, 2) - 2b(3, 4, 1)$  for  $n = 7$ ,  $c \geq b(1, 0, 2) - b(2, 2, 1)$  for  $n = 5$ .

- (v) The structures of  $\frac{n}{2}$  separate pairs for  $n$  even and of  $\frac{n-1}{2}$  separate pairs with an isolated agent for  $n$  odd are PS if

$$b(1, 0, 1) + b(2, 1, 0) \leq c \leq b(1, 0, 0) \quad (38)$$

See the proof in the Appendix.

All the conditions for pairwise stability stated in Proposition 8 depend on the benefit function  $b$  and, with the exception of the empty graph and the separate pair structure, also on the number of agents  $n$ . For any benefit function satisfying our general assumptions, the empty graph will be stable if the cost is high enough, i.e. if it exceeds the benefit of

<sup>2</sup> We use the notation  $\lfloor x \rfloor := \max\{y \in \mathbb{N} \mid y \leq x\}$  and  $\lceil x \rceil := \min\{y \in \mathbb{N} \mid y \geq x\}$ .

receiving an information alone. For other structures, a nonempty cost range in which there is pairwise stability does not exist for all benefit functions. To ensure such a cost range for the star, roughly speaking, an agent's utility must not decline too much when he receives an information at the same time as all other agents in the network compared to when he receives it at the same time as only one other agent. From monotonicity of function  $b$  and assumption (31), for a given  $n$  there always exists a cost range for which the complete network is pairwise stable, but this cost range can be small and shrink drastically when  $n$  grows. Nevertheless, if decay is fairly large and/or the benefit of an agent decreases drastically when he moves from a situation where all other agents are informed at the same time as himself to a case where all other agents are informed before him, then the complete network is pairwise stable. The conditions for the pairwise stability of the circle are rather un-plausible. They require that there is a very large loss in benefit for an agent when moving from a case with only one more agent being informed at the same time as himself to a situation with two more agents being informed. There exists a nonempty cost range for pairwise stability of the separate pairs structure when decay is very large and/or when the benefit of an agent decreases drastically when, instead of being the only one informed, another agent receives information at the same time.

### 4.3 Asymptotic pairwise stability

Next we analyze the asymptotic pairwise stability (APS) as introduced in Definition 1. We check if the PS networks listed in Proposition 8 remain PS if the number of agents becomes very large.

**Remark 2** *In the model (30) the asymptotic stability range is determined by  $(c, b)$ , i.e., by the cost  $c > 0$  and a function  $b$  defined in (30), i.e., verifying  $b(d, x, y) > 0$  for all  $d, x, y$ .*

The asymptotic pairwise stability of some structures depends on which of the following two assumptions (A1) or (A2) is made:

$$(A1) \quad \lim_{k \rightarrow +\infty} b(d, k, y) > 0 \text{ for every } d \geq 2, y \in \mathbb{N}$$

$$(A2) \quad \lim_{k \rightarrow +\infty} b(d, x, k) = 0 \text{ for all } d \geq 2, x \in \mathbb{N}$$

Assumption (A1) states that an agent's benefit from getting the information remains positive even if the number of agents who receive this information earlier than himself becomes very large. On the contrary, assumption (A2) means that the agent does not benefit anymore from the information if that is received at the same time by a very large number of agents. Note that assumption (A1) implies that  $\lim_{k \rightarrow +\infty} b(d, x, k) > 0$  for all  $x \in \mathbb{N}$ . Similarly, (A2) implies that  $\lim_{k \rightarrow +\infty} b(d, k, y) = 0$  for all  $y \in \mathbb{N}$ . We get the following results:

**Proposition 9** *In the model (30) the following holds:*

(i) *The empty network  $g^0$  is APS whenever it is PS, i.e., if  $b(1, 0, 0) \leq c$*

(ii) Under assumption (A1), there always exists a nonempty cost range for which the star  $g^*$  with  $n \geq 3$  is APS. The cost  $c$  must satisfy

$$b(1, 0, 1) - \lim_{n \rightarrow +\infty} b(2, 1, n - 3) \leq c \leq b(1, 0, 0) \quad (39)$$

Under assumption (A2), there exists a nonempty cost range for which the star is APS if and only if

$$\lim_{n \rightarrow +\infty} (n - 2)b(2, 1, n - 3) \geq b(1, 0, 1) \quad (40)$$

and then the cost range must satisfy

$$b(1, 0, 1) \leq c \leq \min \left( b(1, 0, 0), \lim_{n \rightarrow +\infty} (n - 2)b(2, 1, n - 3) \right) \quad (41)$$

(iii) There exists a nonempty cost range for which the complete network  $g^N$  is APS iff  $\lim_{n \rightarrow +\infty} b(1, 0, n - 2) > \lim_{n \rightarrow +\infty} b(2, n - 2, 0)$ . The cost must satisfy

$$c \leq \lim_{n \rightarrow +\infty} b(1, 0, n - 2) - \lim_{n \rightarrow +\infty} b(2, n - 2, 0) \quad (42)$$

In particular, under assumption (A2), the complete network is not APS.

(iv) The circle  $g^c$  of  $n > 3$  nodes is APS if

$$b(1, 0, 2) + 2 \sum_{k=2}^{+\infty} b(k, 2k - 2, 2) \leq c \leq \sum_{k=1}^{+\infty} b(k, 2k - 2, 1) \quad (43)$$

(v) The structures of  $\frac{n}{2}$  separate pairs for  $n$  even and of  $\frac{n-1}{2}$  separate pairs with an isolated agent for  $n$  odd are APS whenever they are PS, i.e., when (38) is satisfied.

See the proof in the Appendix.

Consider the case where the number of agents is getting very large and an agent still derives some benefit from the information even if many others received it earlier than himself. Then there always exists a nonempty cost range for pairwise stability of the star and the more the agent benefits from the information which has been also reached by many other agents in the same time, the larger is this cost range. For the complete network, there exists a nonempty cost range for its pairwise stability if an agent strictly prefers to get the information directly and in the same time as all other agents than to get the information after all other agents have got it. The sharper is this benefit difference, the larger is the cost range for pairwise stability of the complete graph.

Next, we present the result which, under some assumption, puts constraints on the possible APS structures formed under competition in the use of information. We begin by introducing the following definition.

**Definition 2** Let  $SG$  be the set of networks verifying the following property:

All nodes whose degrees are not bounded independently of network size  $n$  belong to a complete subgraph  $g_{cs}$  with the set of agents  $N^{g_{cs}} \subseteq N$  such that  $N^{g_{cs}} \neq \emptyset$ . Moreover, either the size of the complete subgraph  $|N^{g_{cs}}|$  verifies  $\lim_{n \rightarrow \infty} \frac{|N^{g_{cs}}|}{n} > 0$ , or there exists at least one node in  $g_{cs}$  with a high number of connections to agents outside of  $g_{cs}$ , in the sense that there exists  $i \in N^{g_{cs}}$  such that  $|\{j \in g \mid j \notin N^{g_{cs}}\}|$  goes to infinity when  $n$  grows.

We note that in the first case, we have a structure reminiscent of a complete graph. There is a large completely connected component which must make up a sizeable fraction of the whole network. Everyone outside of this component has a much lower degree than the agents in the component. The second case is closer to a star structure: The agents with high degree are all connected to each other but they are not necessarily that many of them. At least one (and possibly all) of the high degree agents must be linked to a large number of low degree agents.

The following proposition characterizes the APS networks in the case where there is strong aversion to the fact that a large number of agents are informed before the receiver.

**Proposition 10** *If  $\lim_{x \rightarrow \infty} b(2, x, 0) = 0$  and  $\lim_{n \rightarrow \infty} b(1, 0, n - 2) \geq c$ , then any network that is APS has a diameter that is bounded independently of  $n$  and belongs to the class of networks SG.*

**Proof:** First, let us show that no APS network can have a diameter that is not bounded independently of  $n$ . Suppose that the diameter  $\bar{d}$  verifies  $\lim_{n \rightarrow \infty} \bar{d}(g_n) = \infty$ . Consider two agents who are at the maximal distance  $\bar{d}$  from each other. Each one of them would gain at least  $b(1, 0, n - 2) - b(\bar{d}, x, y) - c$  from forming a link. But  $\lim_{\bar{d} \rightarrow \infty} b(\bar{d}, x, y) = 0$  for any  $x$  and  $y$ , due to the assumption that the sum of utilities converges. Moreover, we know that  $\lim_{n \rightarrow \infty} b(1, 0, n - 2) \geq c$ . This contradicts the stability of the network with diameter  $\bar{d}$ . From this it also follows that  $g_{cs}$  cannot be empty. If  $g_{cs}$  is the empty graph, then all agents have degrees bounded independently of  $n$ . The diameter of such a network is not bounded independently of  $n$  which is impossible. To show that all agents with asymptotically unbounded degree must be linked, note that the benefit of forming a link with  $i$  if it does not exist is at least  $b(1, 0, n - 2) - b(2, d_i, 0)$ . Since  $\lim_{n \rightarrow \infty} d_i = \infty$ ,  $\lim_{n \rightarrow \infty} b(2, d_i, 0) = 0$ . Moreover, if the complete graph is APS, then  $b(1, 0, n - 2) - b(2, n - 2, 0) > c$  and  $\lim_{n \rightarrow \infty} b(2, n - 2, 0) = 0$ , so we must have  $\lim_{n \rightarrow \infty} b(1, 0, n - 2) \geq c$ .

Suppose that  $\lim_{n \rightarrow \infty} \frac{|N^{g_{cs}}|}{n} = 0$ , and that there exists a  $k$  independent of  $n$  such that for all  $i \in N^{g_{cs}}$ ,  $|\{j \in g \mid j \notin N^{g_{cs}}\}|$  is bounded by  $k$ . The network diameter goes to infinity as the network size grows. Indeed, start from an  $i \in N^{g_{cs}}$ . The number of agents he can reach in  $l$  steps is bounded by  $|N^{g_{cs}}| k d^{l-2}$ , where  $d$  is the maximal degree of agents outside of  $g_{cs}$ . By definition of the diameter  $\bar{d}$ ,  $|N^{g_{cs}}| k d^{\bar{d}-2} = n$  implying  $\bar{d} = \frac{\log(\frac{n}{|N^{g_{cs}}| k})}{\log(d)} + 2$  which goes to infinity as  $n$  grows, which is impossible by the previous analysis. ■

#### 4.4 Efficiency in the model

We will now compare the efficiency of the standard network architectures in our second model. The main insight we obtain from this comparison is that the star network performs better than the other structures in terms of efficiency if we assume that there is always a positive benefit to being informed even when many others receive the information at the same time as myself. We do not necessarily compare the efficiency of all the structures. Once it has been established that a structure cannot be efficient because it is dominated in efficiency for example by the star, it is of limited interest to establish whether or not it is more efficient than some other structure.

**Proposition 11** *Let  $\tilde{g}$  denote the structure of  $\frac{n}{2}$  separate pairs for  $n$  even and of  $\frac{n-1}{2}$  separate pairs with an isolated agent for  $n$  odd. We have:*

- (i) *If  $c > b(1, 0, 0)$  then  $g^0$  is more efficient than  $\tilde{g}$ , and if  $c < b(1, 0, 0)$  then  $\tilde{g}$  is more efficient than  $g^0$ .*
- (ii)  *$g^*$  is more efficient than  $\tilde{g}$  when*

$$(n - 1)[b(1, 0, n - 2) + (n - 2)b(2, 1, n - 3)] > c(n - 2) + b(1, 0, 0) \quad (44)$$

*Under assumption (A1), for any  $c > 0$ ,  $g^*$  is more efficient than  $\tilde{g}$  when  $n$  is sufficiently large.*

- (iii)  *$g^*$  is more efficient than  $g^N$  for sufficiently large costs, i.e., if*

$$c > \frac{n - 1}{n - 2}b(1, 0, n - 2) - b(2, 1, n - 3) - \frac{b(1, 0, 0)}{n - 2} \quad (45)$$

*If  $\lim_{n \rightarrow +\infty} b(1, 0, n - 2) = \lim_{n \rightarrow +\infty} b(2, 1, n - 3)$ , in particular, under assumption (A2), the complete graph cannot be efficient for large  $n$ , since  $g^*$  will then always be more efficient than  $g^N$ .*

- (iv)  *$g^*$  is more efficient than  $g^0$  for sufficiently small costs, i.e., if*

$$2c < b(1, 0, 0) + b(1, 0, n - 2) + (n - 2)b(2, 1, n - 3) \quad (46)$$

*Under the assumption (A1), for any  $c > 0$ ,  $g^*$  will be more efficient than  $g^0$  when  $n$  is sufficiently large.*

From Propositions 9(iii) and 11(iii) we can write the following conclusion:

**Conclusion 2** *There exists a nonempty cost range for which  $g^N$  is APS but not efficient if*

$$\lim_{n \rightarrow +\infty} b(2, 1, n - 3) > \lim_{n \rightarrow +\infty} b(2, n - 2, 0) \quad (47)$$

See the proof in the Appendix.

## 5 Discussion and concluding remarks

Since the payoff functions in our two models of competition for information do not involve the same parameters, we cannot always obtain meaningful comparison of the stable structures in the two models. Indeed, in both models many structures can be PS for some choices of payoff parameters and the parameter ranges cannot be compared. If we turn to asymptotic pairwise stability, predictions become sharper, revealing differences between the effects of competition for the access to, and competition in the use of information, in terms of which structures agents are likely to form. We can also contrast the results in these two models with those of the original connections model.

In the original Jackson-Wolinsky model, the star emerged as the uniquely pairwise stable structure in a large parameter range. This is not the case for either one of the models considered here. In the CA model, the star is never APS for inefficient communication technologies. It fares somewhat better in the CU model if the aversion to others being

informed first is not too strong, since each peripheral agent receives a large quantity of information although it is shared with others. It should be noted that the star, when it is formed does not benefit the same agents in the CA or CU model as in the original connections model. In the latter, most of the benefits are extracted by the peripheral agents who always derive a strictly greater utility from being linked to the center than the center does from being linked to them. In short, the center sponsors costly links that mainly benefit the peripheral agents. Therefore, it is the center's decision that is critical for stability: if the center wants to maintain a link so does the peripheral node. When we introduce competition effects, the situation changes. Agents in the periphery suffer the effects of congestion (CA model), or from receiving information that is always shared with many others (CU model). Now, the star typically fails to be stable because the peripheral agents do not want to maintain a link to the center, or because they link directly to each other to receive more "exclusive" information.

Generally speaking, competition in the use of information seems to favor the stability of structures with small diameters and also of densely connected structures. If an agent sees little value in an information that many others have received before him, and if costs are low enough, we have shown (e.g Proposition 10) that the only APS structures are generalized stars, that is structures where a few high degree nodes, linked to each other are linked to a large number of low degree nodes, or structures similar to the complete network consisting of a large completely connected component whose members are linked to lower degree nodes. Competition in the use of information leads to small network diameters since agents get little benefit from information that has passed through a large number of intermediaries. Again, we are not able to characterize the efficient structure but we have found that in a wide parameter range the star outperforms the other usual structures and in particular the complete network. It results from this that the CU model can exhibit over connectedness since the complete network can be dominated by the star in terms of efficiency in its stability range. The reason for this is quite clear: individual agents do not like to receive an information after everybody else and so they will tend to form more links to gain early access. However, by doing so, they reduce the value of the information for those who previously received it first. Ultimately, nobody will receive any information that is not widely shared.

With competition for the access to information, on the other hand, neither the star nor the complete network is APS when the communication technology is inefficient. Instead, we find that the circle, which avoids congestion effects since each agent has only two contacts, can be stable under some conditions on the parameters, mainly that decay is high enough. While we are not able to identify the efficient network in the CA model, a comparison of several "standard" network architectures show that the circle outperforms the others in terms of efficiency for a wide range of costs and levels of decay under a condition that is only related to the communication technology, namely that the communication efficiency is close to being maximal for agents with degree two. This condition which enhances the efficiency of the circle may however be at odds with the conditions that ensure its stability, and which require some loss in communication efficiency even for small degrees. In the CU model, the circle is PS only under very un-plausible assumptions on the payoff function. A network with long communication chains does not satisfy agents who compete for the use of information since most of the informations they receive will have passed through a large number of intermediaries.



Depending on whether they face competition for the access to information or competition in the use of information, agents' network formation strategies will be rather different. In the CU model, the desire to avoid being informed after others will incite agents to create new links to bridge large distances. If costs are low, the stable networks have short diameters and may be densely connected, sometimes too much so from the point of view of efficiency. In the CA model, on the other hand, densely connected structures are not plausible unless costs are very very low. In this context, the source of disutility is congestion, a problem which is aggravated in densely connected networks. Longer communication chains can now be more stable and efficient than structures with short diameters because there is less congestion when information is transmitted.

## 6 Acknowledgements

Agnieszka Rusinowska and Emily Tanimura acknowledge the support by the National Agency for Research (Agence Nationale de la Recherche), Project DynaMITE (ANR-13-BSH1-0010-01). Philipp Möhlmeier carried out this research within the International Research Training Group Economic Behavior and Interaction Models (EBIM) financed by the German Research Foundation (DFG) under contract GRK 1134/2 and within the Bielefeld Graduate School of Economics and Management (BiGSEM). Financial support by EBIM, BiGSEM and the Center for Mathematical Economics is gratefully acknowledged.

# Bibliography

- V. Bala and S. Goyal. A noncooperative model of network formation. *Econometrica*, 68(5):1181–1229, 2000.
- P. Billand, C. Bravard, and S. Sarangi. Directed networks with spillovers. *Journal of Public Economic Theory*, 14:849–878, 2012.
- P. Billand, C. Bravard, and S. Sarangi. A note on local spillovers, convexity and the strategic substitutes property in networks. *Theory and Decision*, 75(2):293–304, 2013.
- F. Bloch and M. O. Jackson. The formation of networks with transfers among players. *Journal of Economic Theory*, 133(1):83–110, 2007.
- B. Buechel and T. Hellmann. Under-connected and over-connected networks: the role of externalities in strategic network formation. *Review of Economic Design*, 16:71–87, 2012.
- B. Charoensook. Network formation with productivity as decay. MPRA Paper No. 37099, 2012.
- S. Currarini. Network design in games with spillovers. *Review of Economic Design*, 10(4):305–326, 2007.
- S. Goyal and S. Joshi. Unequal connections. *International Journal of Game Theory*, 34:319–349, 2006.
- M. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380, 1973.
- M. Granovetter. *Getting a Job: A Study of Contacts and Careers*. The University of Chicago Press, Chicago, 1974.
- H. Haller. Network extension. *Mathematical Social Sciences*, 64:166–172, 2012.
- T. Hellmann. On the existence and uniqueness of pairwise stable networks. *International Journal of Game Theory*, 42(1):211–237, 2013.
- M. O. Jackson. *Social and Economic Networks*. Princeton University Press, 2008.
- M. O. Jackson and A. Wolinsky. A strategic model of social and economic networks. *Journal of Economic Theory*, 71:44–74, 1996.
- Y. Lim, A. Ozdaglar, and A. Teytelboym. A simple model of cascades in networks. Working Paper, 2015.
- P. Möhlmeier, A. Rusinowska, and E. Tanimura. A degree-distance-based connections model with negative and positive externalities. *Journal of Public Economic Theory*, 2016. Forthcoming.
- T. Morrill. Network formation under negative degree-based externalities. *International Journal of Game Theory*, 40:367–385, 2011.

## A Proofs of the results

### A.1 Proof of Proposition 3

(i) Consider any two agents  $i, j \in g^\emptyset$ . We have  $u_i^{CA}(g^\emptyset + ij) - u_i^{CA}(g^\emptyset) = u_j^{CA}(g^\emptyset + ij) - u_j^{CA}(g^\emptyset) = f(1)\delta - c \leq 0$  iff  $f(1)\delta \leq c$ .

(ii) Consider the star  $g^*$  with  $n \geq 3$  agents. Take the center of the star  $i$  and two arbitrary agents  $j, k$ , where  $j \neq i$ ,  $k \neq i$ , and  $j \neq k$ . This means that  $ij \in g^*$  but  $jk \notin g^*$ . For stability the following conditions must hold:

- (A)  $u_i^{CA}(g^*) - u_i^{CA}(g^* \setminus ij) \geq 0$  and
- (B)  $u_j^{CA}(g^*) - u_j^{CA}(g^* \setminus ij) \geq 0$  and
- (C)  $u_j^{CA}(g^* + jk) - u_j^{CA}(g^*) \leq 0$ .

(A):  $u_i^{CA}(g^*) - u_i^{CA}(g^* \setminus ij) = (n-1)f(1)\delta - (n-1)c - (n-2)f(1)\delta + (n-2)c = f(1)\delta - c$ . Hence, (A) holds iff  $f(1)\delta \geq c$ .

(B):  $u_j^{CA}(g^*) - u_j^{CA}(g^* \setminus ij) = f(n-1)\delta + (n-2)f(1)f(n-1)\delta^2 - c$ . Hence, (B) holds iff  $f(n-1)\delta + (n-2)f(1)f(n-1)\delta^2 \geq c$ .

(C):  $u_j^{CA}(g^* + jk) - u_j^{CA}(g^*) = f(2)\delta + f(n-1)\delta^2(2f(2) - f(1)) + (n-3)f(1)f(2)f(n-1)\delta^3 - c$ . Hence, condition (C) holds iff  $f(2)\delta + f(n-1)\delta^2(2f(2) - f(1)) + (n-3)f(1)f(2)f(n-1)\delta^3 \leq c$ .

Hence, (A) and (B) and (C) lead to condition (10).

(iii) Let  $n \geq 3$ . Consider any two agents  $i, j \in g^N$ . We have

$$u_i^{CA}(g^N) - u_i^{CA}(g^N - ij) = f(n-1)\delta + (n-2)f(n-1)(2f(n-1) - f(n-2))\delta^2 + f^2(n-1)(n-2)(n-3)(3f(n-1) - 2f(n-2))\delta^3 + \dots + f^{n-2}(n-1)(n-2)!((n-1)f(n-1) - (n-2)f(n-2))\delta^{n-1} - c$$

which leads to (11).

(iv) Let  $g^c$  be the circle of  $n$  agents. Let  $i, j \in g$ . The ‘no-deletion’ condition  $u_i^{CA}(g^c) \geq u_i^{CA}(g^c - ij)$  holds iff  $2f(2)\delta + 2f^2(2)\delta^2 + \dots + 2f^{n-1}(2)\delta^{n-1} - 2c \geq f(2)\delta + f^2(2)\delta^2 + \dots + f^{n-2}(2)\delta^{n-2} + f^{n-2}(2)f(1)\delta^{n-1} - c$  iff condition (12) is satisfied.

For the ‘no-addition’ condition, it is enough to show that agent  $i$  does not want to add a link to the node that is the most far away from himself. Denote such a node by  $k$ . Let  $n$  be even. Then  $u_i^{CA}(g^c) \geq u_i^{CA}(g^c + ik)$  iff  $2f(2)\delta + 2f^2(2)\delta^2 + \dots + 2f^{n-1}(2)\delta^{n-1} - 2c \geq 2f(2)\delta + f(3)\delta + 2f^2(2)\delta^2 + 2f^2(2)f(3)\delta^2 + \dots + 2f^{\frac{n}{2}-1}(2)\delta^{\frac{n}{2}-1} + 2f^{\frac{n}{2}-2}(2)f(3)\delta^{\frac{n}{2}-1} + 4f^{\frac{n}{2}-1}(2)f(3)\delta^{\frac{n}{2}} + 2f^{\frac{n}{2}}(2)f(3)\delta^{\frac{n}{2}+1} + \dots + 2f^{n-3}(2)f(3)\delta^{n-2} + 2f^{n-2}(2)f(3)\delta^{n-1} - 3c$  iff condition (13) is satisfied.

If  $n$  is odd, then  $u_i^{CA}(g^c) \geq u_i^{CA}(g^c + ik)$  iff  $2f(2)\delta + 2f^2(2)\delta^2 + \dots + 2f^{n-1}(2)\delta^{n-1} - 2c \geq 2f(2)\delta + f(3)\delta + 2f^2(2)\delta^2 + 2f^2(2)f(3)\delta^2 + \dots + 2f^{\frac{n-1}{2}-1}(2)\delta^{\frac{n-1}{2}-1} + 2f^{\frac{n-1}{2}-2}(2)f(3)\delta^{\frac{n-1}{2}-1} + f^{\frac{n-1}{2}}(2)\delta^{\frac{n-1}{2}} + 3f^{\frac{n-1}{2}-1}(2)f(3)\delta^{\frac{n-1}{2}} + 2f^{\frac{n-1}{2}}(2)f(3)\delta^{\frac{n-1}{2}+1} + \dots + 2f^{n-3}(2)f(3)\delta^{n-2} + 2f^{n-2}(2)f(3)\delta^{n-1} - 3c$  iff condition (14) holds.

(v) Let  $n$  be even. Consider network  $g$  consisting of  $\frac{n}{2}$  separate pairs. Take arbitrary  $i, j, k \in N$  such that  $ij \in g$ ,  $ik \notin g$ . We have the following conditions:

$$u_i^{CA}(g) - u_i^{CA}(g - ij) = f(1)\delta - c \geq 0$$

$$u_i^{CA}(g + ik) - u_i^{CA}(g) = f(2)\delta + f(1)f(2)\delta^2 - c \leq 0$$

Hence,  $g$  is PS iff (15) holds. The cost range is nonempty whenever  $f(2)\delta(1 + f(1)\delta) \leq f(1)\delta$  which gives condition (16).

Now, let  $n$  be odd and consider network  $g$  consisting of  $\frac{n-1}{2}$  separate pairs and one isolated agent, say agent  $m$ . Then we have to consider additionally  $u_m^{CA}(g + im) - u_m^{CA}(g) = f(2)\delta + f(1)f(2)\delta^2 - c$ , but from the previous condition we have  $u_m^{CA}(g + im) - u_m^{CA}(g) \leq 0$ , so we get also (15).

## A.2 Proof of Proposition 4

(i) and (v) are obvious, since the conditions for PS do not depend on  $n$ .

(ii) Let function  $f$  be such that  $\lim_{n \rightarrow \infty} f(n)n = 0$ . Then the star is APS if  $f(2)\delta \leq c \leq 0$ . Suppose now that function  $f$  is such that  $0 < \lim_{n \rightarrow +\infty} f(n)n \leq 1$ . Then the star is APS if

$$f(2)\delta + f(1)f(2)\delta^3 \lim_{n \rightarrow +\infty} f(n)n \leq c \leq f(1)\delta^2 \lim_{n \rightarrow +\infty} f(n)n$$

Such a positive cost exists when

$$\lim_{n \rightarrow +\infty} f(n)n \geq \frac{f(2)}{\delta f(1)(1 - \delta f(2))}$$

In particular, if  $f(d) = \frac{1}{d}$ , then the condition becomes  $\frac{1}{2}(\delta - 1)^2 \leq 0$  which is satisfied by no  $\delta \in (0, 1)$ .

(iii) Stability of the complete network requires

$$c \leq f(n-1)\delta \left[ 1 + \sum_{k=1}^{n-2} f^{k-1}(n-1) \frac{(n-2)!}{(n-2-k)!} \delta^k ((k+1)f(n-1) - kf(n-2)) \right] =: S$$

Let us show that the right hand expression of this inequality goes to 0 as  $n \rightarrow +\infty$ . We note that

$$(k+1)f(n-1) - kf(n-2) < (k+1)f(n-2) - kf(n-2) = f(n-2)$$

Since  $\lim_{n \rightarrow +\infty} f(n)n \leq 1$ , there exists an  $M$  such that for all  $n \geq M$ ,  $f(n)n \leq 1$ . Thus

$$S \leq f(n-1)\delta \left[ 1 + \sum_{k=1}^M f^{k-1}(n-1) \frac{(n-2)!}{(n-2-k)!} \delta^k + \sum_{k=M+1}^{n-2} f(n-2) f^{k-1}(n-1) \delta^k \frac{(n-2)!}{(n-2-k)!} \right]$$

We have

$$\begin{aligned} f(n-2) f^{k-1}(n-1) \frac{(n-2)!}{(n-2-k)!} &= (n-2-k+1) \dots (n-3)(n-2) f(n-2) f^{k-1}(n-1) < \\ &< ((n-2)f(n-2))^k \leq 1 \end{aligned}$$

Thus

$$S \leq f(n-1)\delta \left[ 1 + \sum_{k=1}^M f^{k-1}(n-1) \frac{(n-2)!}{(n-2-k)!} \delta^k + \sum_{k=M+1}^{n-2} \delta^k \right]$$

The first sum in the bracket is finite, the second one converges since  $\delta < 1$  and  $\lim_{n \rightarrow +\infty} f(n-1) = 0$ . Consequently,  $S$  tends to zero and no positive cost exists.

(iv) Consider the circle with  $n$  agents, where  $n$  is even. The proof for  $n$  odd is analogous. The (even) circle is APS if

$$\lim_{n \rightarrow +\infty} \left[ f(3)\delta + 2f(3) \sum_{k=1}^{n-2} f^k(2) \delta^{k+1} \right] \leq c \leq \lim_{n \rightarrow +\infty} \sum_{k=1}^{n-1} f^k(2) \delta^k$$

or equivalently

$$f(3)\delta + 2f(3)\delta \lim_{n \rightarrow +\infty} \left[ \frac{f(2)\delta - (f(2)\delta)^{n-1}}{1 - f(2)\delta} \right] \leq c \leq \lim_{n \rightarrow +\infty} \frac{f(2)\delta - (f(2)\delta)^n}{1 - f(2)\delta} \iff (48)$$

$$f(3)\delta + 2f(3)\delta \left[ \frac{f(2)\delta}{1 - f(2)\delta} \right] \leq c \leq \frac{f(2)\delta}{1 - f(2)\delta} \iff (49)$$

$$f(3) + f(3)f(2)\delta \leq \frac{c(1 - f(2)\delta)}{\delta} \leq f(2) \quad (50)$$

Such a positive cost exists whenever  $f(3) + f(3)f(2)\delta \leq f(2) \iff \delta \leq \frac{f(2)-f(3)}{f(2)f(3)}$ .

### A.3 Proof of Proposition 5

Assume that  $f(1) = 1$  and  $(n-1)f(n-1) < 1$ . We begin by proving the first inequality in the first part concerning the circle vs the complete graph. We will show that  $\sum_{i \in N} u_i^{CA}(g^c) > \sum_{i \in N} u_i^{CA}(g^N)$ . We have

$$\begin{aligned} & \sum_{i \in N} u_i^{CA}(g^N) + (n-1)nc = \\ &= n\delta [(n-1)f(n-1) + \delta(n-1)(n-2)f^2(n-1) + \dots + \delta^{n-2}(n-1)!f^{n-1}(n-1)] < \\ &< n\delta [(n-1)f(n-1) + \delta(n-1)^2f^2(n-1) + \dots + \delta^{n-2}(n-1)^{n-1}f^{n-1}(n-1)] < \\ &< n\delta (1 + \delta + \delta^2 + \dots + \delta^{n-2}) = \frac{n\delta(1 - \delta^{n-1})}{1 - \delta} \end{aligned}$$

Furthermore,

$$\begin{aligned} & \sum_{i \in N} u_i^{CA}(g^c) + 2nc = \\ &= 2nf(2)\delta [1 + \delta f(2) + \dots + \delta^{n-2}f^{n-2}(2)] = \frac{2nf(2)\delta(1 - (\delta f(2))^{n-1})}{1 - \delta f(2)} \end{aligned}$$

Hence,

$$\begin{aligned} \sum_{i \in N} u_i^{CA}(g^N) + 2nc &< \sum_{i \in N} u_i^{CA}(g^N) + (n-1)nc < \frac{n\delta(1 - \delta^{n-1})}{1 - \delta} < \\ &< \frac{n\delta(1 - (\delta f(2))^{n-1})}{1 - \delta} < \frac{2nf(2)\delta(1 - (\delta f(2))^{n-1})}{1 - \delta f(2)} \end{aligned}$$

where the last inequality holds iff  $\delta < \frac{2f(2)-1}{f(2)}$  and  $2f(2) > 1$ . Since  $f(2) \rightarrow f(1) = 1$ , this inequality holds for every  $\delta < 1$ .

We then compare the circle and the star. We have

$$\sum_{i \in N} u_i^{CA}(g^c) = \frac{2nf(2)\delta(1 - (\delta f(2))^{n-1})}{1 - \delta f(2)} - 2nc > 2nf(2)\delta - 2nc$$

Furthermore,

$$\sum_{i \in N} u_i^{CA}(g^*) = (n-1)\delta(f(1) + f(n-1) + (n-2)f(1)f(n-1)\delta) - 2(n-1)c$$

Hence, the condition

$$2nf(2)\delta - 2nc > (n-1)\delta(f(1) + f(n-1) + (n-2)f(1)f(n-1)\delta) - 2(n-1)c$$

is equivalent to

$$2c < 2nf(2)\delta - (n-1)\delta(f(1) + f(n-1) + (n-2)f(1)f(n-1)\delta)$$

Let us show that for any  $0 < \delta < 1$ , the right hand expression above tends to infinity when  $n \rightarrow \infty$  and  $f(2) \rightarrow f(1)$ . Indeed we have:

$$\lim_{\substack{n \rightarrow \infty, \\ f(2) \rightarrow f(1)}} \frac{1}{n} [2nf(2)\delta - (n-1)\delta(f(1) + f(n-1) + (n-2)f(1)f(n-1)\delta)] \geq \quad (51)$$

$$\delta [2f(1) - f(1) - f(1)\delta] > 0 \quad (52)$$

Finally, we consider the last part by comparing the circle with the disjoint pair structure. Consider

$$\lim_{\substack{n \rightarrow \infty, \\ f(2) \rightarrow f(1)}} \left[ \sum_{i \in N} u_i^{CA}(g^c) - \sum_{i \in N} u_i^{CA}(\tilde{g}) \right]$$

We have

$$\sum_{i \in N} u_i^{CA}(\tilde{g}) = n\delta f(1) - nc, \quad \sum_{i \in N} u_i^{CA}(g^c) = \frac{2nf(2)\delta(1 - (\delta f(2))^{n-1})}{1 - \delta f(2)} - 2nc$$

$$\lim_{\substack{n \rightarrow \infty, \\ f(2) \rightarrow f(1)}} \left[ \frac{2nf(2)\delta(1 - (\delta f(2))^{n-1})}{1 - \delta f(2)} - 2nc - n\delta f(1) + nc \right] = \quad (53)$$

$$\lim_{\substack{n \rightarrow \infty, \\ f(2) \rightarrow f(1)}} \left[ \frac{2nf(2)\delta}{1 - \delta f(2)} - n\delta f(1) - nc \right] = \lim_{\substack{n \rightarrow \infty, \\ f(2) \rightarrow f(1)}} \left( \frac{2f(1)\delta}{1 - f(1)\delta} - f(1)\delta - c \right) n \quad (54)$$

Since  $f(1) = 1$ , this quantity is positive if

$$c < \frac{\delta(1 + \delta)}{1 - \delta}$$

#### A.4 Proof of Proposition 6

The condition for  $\sum_{i \in N} u_i^{CA}(g^c) - \sum_{i \in N} u_i^{CA}(g^N) > 0$  comes from the proof of Proposition 5. In the proof of Proposition 5 we also saw that  $\sum_{i \in N} u_i^{CA}(g^c) - \sum_{i \in N} u_i^{CA}(g^*) > 0$  if

$$2c < 2nf(2)\delta - (n-1)\delta(f(1) + f(n-1) + (n-2)f(1)f(n-1)\delta)$$

Such a cost range exists whenever

$$2nf(2)\delta - (n-1)\delta(f(1) + f(n-1) + (n-2)f(1)f(n-1)\delta) > 0$$

which due to the fact that

$$2nf(2) - (n-1)(f(1) + f(n-1)) > 0$$

leads to

$$\delta < \frac{2nf(2) - (n-1)(f(1) + f(n-1))}{(n-1)(n-2)f(1)f(n-1)}$$

## A.5 Proof of Proposition 8

(i) Consider any two agents  $i, j \in g^\emptyset$ . We have  $u_i^{CU}(g^\emptyset + ij) - u_i^{CU}(g^\emptyset) = u_j^{CU}(g^\emptyset + ij) - u_j^{CU}(g^\emptyset) = b(1, 0, 0) - c \leq 0$  iff  $b(1, 0, 0) \leq c$ .

(ii) Consider the star  $g^*$  with  $n \geq 3$  agents. Let  $i$  be the center of the star and  $j, k$  two arbitrary agents, where  $j \neq i, k \neq i$ , and  $j \neq k$ . The stability conditions are the following:

(A)  $u_i^{CU}(g^*) - u_i^{CU}(g^* \setminus ij) \geq 0$  and

(B)  $u_j^{CU}(g^*) - u_j^{CU}(g^* \setminus ij) \geq 0$  and

(C)  $u_j^{CU}(g^* + jk) - u_j^{CU}(g^*) \leq 0$ .

(A):  $u_i^{CU}(g^*) - u_i^{CU}(g^* \setminus ij) = (n-1)b(1, 0, 0) - (n-1)c - (n-2)b(1, 0, 0) + (n-2)c = b(1, 0, 0) - c$ . Hence, (A) holds iff  $b(1, 0, 0) \geq c$ .

(B):  $u_j^{CU}(g^*) - u_j^{CU}(g^* \setminus ij) = b(1, 0, n-2) + (n-2)b(2, 1, n-3) - c$ . Hence, (B) holds iff  $b(1, 0, n-2) + (n-2)b(2, 1, n-3) \geq c$ .

(C):  $u_j^{CU}(g^* + jk) - u_j^{CU}(g^*) = b(1, 0, 1) - b(2, 1, n-3) - c$ . Hence, condition (C) holds iff  $c \geq b(1, 0, 1) - b(2, 1, n-3)$ .

Hence, (A) and (B) and (C) give condition (32).

(iii) Let  $n \geq 3$ . Consider any two agents  $i, j \in g^N$ . We have

$u_i^{CU}(g^N) - u_i^{CU}(g^N - ij) = b(1, 0, n-2) - b(2, n-2, 0) - c \geq 0$  iff  $c \leq b(1, 0, n-2) - b(2, n-2, 0)$ .

(iv) Consider the circle  $g^c$  with  $n > 3$  agents. Let  $i, j$  be arbitrary two agents such that  $ij \in g^c$ . The no-link-deletion condition  $u_i^{CU}(g^c) - u_i^{CU}(g^c - ij) \geq 0$  is equivalent to (34) if  $n$  is even and to (36) if  $n$  is odd.

Consider now the no-link-addition condition. It is sufficient to guarantee that a node does not want to form a link with another node which is most far away from that node, as connecting to any node in the circle which is not at a maximal distance would be less profitable. Let  $n$  be even and  $n \geq 8$ . The condition  $u_i^{CU}(g^c) - u_i^{CU}(g^c + ik) \geq 0$  for  $ik \notin g$  is equivalent to (35). The first difference  $(b(1, 0, 2) - b(\frac{n}{2}, n-2, 0))$  on the right hand side of this inequality corresponds to node  $i$ 's gain of being directly connected to node  $k$  which was before at distance  $\frac{n}{2}$  from  $i$ . The second difference on the right hand side of condition (35) corresponds to  $i$ 's gain from all other nodes that can be reached by  $i$  by a shorter distance via node  $k$ . For  $n = 4$  and  $n = 6$  node  $i$ 's total gain consists of the first difference only, that is, the gain of being directly connected to  $k$ . The argumentation follows analogously for the case when  $n$  is odd and leads to (37).

(v) It results immediately from the definition of PS.

## A.6 Proof of Proposition 9

(i) and (v) are obvious, since the conditions for PS do not depend on  $n$ .

(ii) Consider the star  $g^*$  with  $n \geq 3$  agents. Let assumption (A1) be satisfied. Then we have  $\lim_{k \rightarrow +\infty} b(d, x, k) > 0$  for all  $d, x$ . The star is APS whenever (39) is satisfied, because  $b(1, 0, 0) < \lim_{n \rightarrow +\infty} (b(1, 0, n-2) + (n-2)b(2, 1, n-3))$ . Note that the cost range is always nonempty, since  $b(1, 0, 1) - \lim_{n \rightarrow +\infty} b(2, 1, n-3) < b(1, 0, 0)$ .

Suppose now that assumption (A2) is satisfied. The star is APS whenever (41) is satisfied. If  $\lim_{n \rightarrow +\infty} (n-2)b(2, 1, n-3) < b(1, 0, 1)$ , then this cost range is empty. On the contrary,

if  $\lim_{n \rightarrow +\infty} (n-2)b(2, 1, n-3) \geq b(1, 0, 1)$ , then there exists some cost that satisfies (41), since  $b(1, 0, 1) < b(1, 0, 0)$ .

(iii) Stability of the complete network requires (42) to be satisfied. If  $\lim_{n \rightarrow +\infty} b(1, 0, n-2) > \lim_{n \rightarrow +\infty} b(2, n-2, 0)$ , then the right hand side of (42) is positive and  $g^N$  is APS. If  $\lim_{n \rightarrow +\infty} b(1, 0, n-2) = \lim_{n \rightarrow +\infty} b(2, n-2, 0)$ , then the right hand side of (42) is equal to 0, and consequently  $g^N$  is not APS.

(iv) Consider the circle  $g^c$  of  $n > 3$  nodes. When going to the limit under  $n \rightarrow +\infty$  in the right hand expressions in (34), (35), (36) and (37), we obtain condition (43).

## A.7 Proof of Proposition 11

We have:

$$\begin{aligned} \sum_{i \in N} u_i^{CU}(g^\emptyset) &= 0, & \sum_{i \in N} u_i^{CU}(\tilde{g}) &= n(b(1, 0, 0) - c) \\ \sum_{i \in N} u_i^{CU}(g^*) &= (n-1)(b(1, 0, 0) + b(1, 0, n-2) + (n-2)b(2, 1, n-3) - 2c) \\ \sum_{i \in N} u_i^{CU}(g^N) &= n(n-1)(b(1, 0, n-2) - c) \\ \sum_{i \in N} u_i^{CU}(g^c) &= n \left( 2 \sum_{k=1}^{\frac{n-1}{2}} b(k, 2k-2, 1) + b\left(\frac{n}{2}, n-2, 0\right) - 2c \right) \quad \text{for } n \text{ even} \\ \sum_{i \in N} u_i^{CU}(g^c) &= 2n \left( \sum_{k=1}^{\frac{n-1}{2}} b(k, 2k-2, 1) - c \right) \quad \text{for } n \text{ odd} \end{aligned}$$

All parts (i), (ii), (iii) and (iv) result immediately from the comparison of the sums given above.

## A.8 Proof of Conclusion 2

When moving to the limit in (45), by virtue of Proposition 11(iii),  $g^*$  is more efficient than  $g^N$  if

$$c \geq \lim_{n \rightarrow +\infty} b(1, 0, n-2) - \lim_{n \rightarrow +\infty} b(2, 1, n-3)$$

From monotonicity of function  $b$ , and assumptions (31) and (47) we have

$$\lim_{n \rightarrow +\infty} b(1, 0, n-2) \geq \lim_{n \rightarrow +\infty} b(2, 1, n-3) > \lim_{n \rightarrow +\infty} b(2, n-2, 0)$$

and therefore from Proposition 9(iii),  $g^N$  is APS and the cost range for its pairwise stability satisfies (42). Hence, if the cost range is such that

$$\lim_{n \rightarrow +\infty} b(1, 0, n-2) - \lim_{n \rightarrow +\infty} b(2, 1, n-3) \leq c \leq \lim_{n \rightarrow +\infty} b(1, 0, n-2) - \lim_{n \rightarrow +\infty} b(2, n-2, 0)$$

then  $g^N$  is APS but not efficient. This cost range is nonempty under assumption (47).