HAL
archives-ouvertes.fr

A degree-distance-based connections model with negative and positive externalities

Philipp Möhlmeier, Agnieszka Rusinowska, Emily Tanimura

▶ **To cite this version:**

Philipp Möhlmeier, Agnieszka Rusinowska, Emily Tanimura. A degree-distance-based connections model with negative and positive externalities. Documents de travail du Centre d'Economie de la Sorbonne 2013.40 - ISSN : 1955-611X. 2013. <halshs-00825266>

HAL Id: halshs-00825266

https://halshs.archives-ouvertes.fr/halshs-00825266

Submitted on 23 May 2013

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.

Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers

A degree-distance-based connections model with negative and positive externalities

Philipp MÖHLMEIER, Agnieszka RUSINOWSKA, Emily TANIMURA

2013.40

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
http://centredeconomiesorbonne.univ-paris1.fr/bandeau-haut/documents-de-travail/
ISSN : 1955-611X

A degree-distance-based connections model with negative and positive externalities*

Philipp Möhlmeier[1], Agnieszka Rusinowska[2], and Emily Tanimura[3]**

[1] Bielefeld University, BiGSEM, Center for Mathematical Economics, Bielefeld, Germany
philipp.moehlmeier@uni-bielefeld.de
[2] Paris School of Economics - CNRS, Centre d'Economie de la Sorbonne
agnieszka.rusinowska@univ-paris1.fr
[3] Université Paris I Panthéon-Sorbonne, Centre d'Economie de la Sorbonne
106-112 Bd de l'Hôpital, 75647 Paris, France
Emily.Tanimura@univ-paris1.fr

Abstract. We develop a modification of the connections model by Jackson and Wolinsky (1996) that takes into account negative externalities arising from the connectivity of direct and indirect neighbors, thus combining aspects of the connections model and the co-author model. We consider a general functional form for agents' utility that incorporates both the effects of distance and of neighbors' degree. Consequently, we introduce a framework that can be seen as a degree-distance-based connections model with both negative and positive externalities. Our analysis shows how the introduction of negative externalities modifies certain results about stability and efficiency compared to the original connections model. In particular, we see the emergence of new stable structures, such as a star with links between peripheral nodes. We also identify structures, for example, certain disconnected networks, that are efficient in our model but which could not be efficient in the original connections model. While our results are proved for the general utility function, some of them are illustrated by using a specific functional form of the degree-distance-based utility.

JEL Classification: D85, C70

Keywords: connections model, degree, distance, negative externalities, positive externalities, pairwise stability, efficiency

Running title: Degree-distance-based connections model

1 Introduction

1.1 Motivation of the paper

In the seminal paper of Jackson and Wolinsky (1996) the so called connections model was introduced. It is an example of social communication between individuals where benefits and costs for each individual are determined by the direct and indirect connections among them. In the model, only direct connections are costly, but (discounted) benefits spill over from more distant partners. Jackson and Wolinsky (1996) focus on the identification of pairwise stable networks (those where no agent wants to sever a link and no two agents

* We would like to thank participants of the conference 'The Role of Externalities in Networks' (Baton Rouge, February 23-24, 2013), in particular, Hans Haller, Ahmed Saber Mahmud and Sudipta Sarangi, and also Herbert Dawid and Tim Hellmann, for helpful comments. Agnieszka Rusinowska and Emily Tanimura acknowledge the support by the National Agency for Research (Agence Nationale de la Recherche), Reference: ANR-09-BLAN-0321-01.

** Corresponding author.

both want to add a link) and (strongly) efficient networks (those which maximize the total utility of all agents).

One of the particularly appealing features of networks is the fact that they allow us to capture the *externalities* that 'occur when the utility of or payoff to an individual is affected by the actions of others, although those actions do not directly involve the individual in question' (Jackson (2008), p. 162). Naturally, the nature of the externalities that operate in a network are important for determining which types of networks can form and can be maintained.

In the Jackson and Wolinsky model, network externalities are positive. A link formed by some pair of individuals only has (weakly) positive effects for everybody else. A natural interpretation of the Jackson and Wolinsky model is that benefits result from the access to a valuable resource that flows throughout the network, such as information or knowledge provided by indirect contacts. Increased connectivity is then beneficial to all agents in the network because it gives access to new indirect contacts from whom information is obtained or because information arrives more rapidly when paths are shorter. These positive aspects of increased connectivity are certainly important. However, in many situations, it is likely that increased connectivity also has negative side effects. Studying such cases is what motivates our extensions. In this paper, we introduce and analyze a model in which agents benefit from indirect contacts as in the original connections model but which also takes into account that there may be negative externalities arising from the connectivity of direct and indirect neighbors.

Indeed, there are several aspects of increased connectivity that could affect agents negatively. On the one hand, information received from indirect contacts is valuable but on the other hand it may be less valuable when it is too widely shared. For example, learning of a job opening through an indirect contact may be less useful if the same information has been communicated to many others. In many situations of interest, it is thus natural to assume that the benefits of having indirect contacts are reduced when the latter have a large number of connections. In our model, the form of the negative externality is such that the utility an agent derives from an indirect contact, viewed as the initial sender of an information, is reduced when the latter has a high degree and thus sends the information to many others. While we do not take into account the full effect of all other indirect contacts of the sender, our model could thus be viewed as a simplified or approximate description of the negative effects of connectivity when there is competition for information.

Another type of negative effect from increased connectivity that is well represented by our model is related to the fact that the busyness of agents reduces their availability or productivity for their other direct and indirect partners. The model of Jackson and Wolinsky (1996) could also be interpreted as follows: nodes can generate some output by themselves and forward output from other nodes. Output that is generated by other nodes arrives at distant nodes with some decay depending on the length of the path it has to wander through. Now consider this as a situation in which people are involved in projects and generate some kind of knowledge by themselves and receive some from others. Then, a natural perception would be to say that if a person is involved in many projects, she will have less time to generate output by herself. On the other hand, the more well connected she is, the more knowledge she will receive from others which she can then forward to her neighborhood. A provocative way to state this fact is to say that "well

connected people are often great talkers, but networking is time consuming and reduces the own productive time such that the main work is done by others". Nevertheless, the role of such well connected agents is very important: not in the sense that they contribute a lot by their own results, but because they provide access to the output of many other people.

The original connections model captures information overspill from indirect neighbors but does not take into account that the busyness of the agents may have a negative effect on the own productivity and hence to all other agents that receive this (reduced amount of) agent specific output. By integrating this aspect, on a first sight, our model looks similar to the well-known co-author model of Jackson and Wolinsky (1996), where depending on the number of projects a co-author is already involved in, the amount of time that he can devote to a single connection is decreasing. Indeed, the nature of the negative externality is similar but a major difference is that in our adjusted version, information spills over from indirect connections. The co-author model only takes into account the benefit of direct scientific collaborations and consequently conveys negative effects solely through the busyness of direct co-authors. Increased connectivity of a neighbor is then only a bad thing since there are no spillovers from indirect neighbors.

Hence, our model combines different aspects of the connections model and the co-author model, and can be seen as a degree and distance-based model in which the externalities resulting from additional links in the network are both positive and negative: they allow other agents to reach new indirect partners or to reach existing ones in fewer steps, however they also reduce the productivity of the agents that are reached by making them more busy and thus less valuable per se although more efficient as intermediaries. Exploring the tradeoff of these effects is the object of this analysis.

1.2 Related literature

There are numerous extensions of the connections model in the literature on networks. A certain variation of the connections model is presented in Jackson and Rogers (2005), where geographic costs of forming links are assumed. Players are grouped on the so called islands, and the connection costs are low within an island and high across islands. Other variations of the connections model with geographic costs are investigated, e.g., in Johnson and Gilles (2000), Carayol and Roux (2005, 2009). Bloch and Jackson (2007) generalize the symmetric connections model by introducing the distance-based utility model, in which agents obtain utility from their direct and indirect connections, and the utility decreases with the distance between the agents. Extensions of the Jackson-Wolinsky connections model to a dynamic framework are presented, e.g., in Jackson and Watts (2002), Watts (2001, 2002).

While several related works in the literature similarly deal with network formation in the presence of negative externalities, the form of the externality considered in those does not allow us to account for situations such as the one mentioned previously in Section 1.1. For example, Morrill (2011) introduces a degree-based utility and implements the idea of a negative externality that comes from the connectivity of direct neighbors. The more connections a direct neighbor has, the less utility is provided through the link. A simple example for this is the co-author model by Jackson and Wolinsky (1996).

Another related work in the area by Goyal and Joshi (2006) distinguishes in a quite general framework negative externalities resulting from the connectivity of direct neigh-

bors, indirect neighbors and agents to whom one is not linked. The authors investigate two specific models and assume that the externalities are either positive or negative. One of them is a playing the field game in which the payoff of an agent depends on the number of his links and the aggregate number of links of the remaining agents. Their second model is a local spillovers game in which the payoff of an agent depends on the distribution of links of all agents and the identity of neighbors. Billand et al. (2012a) provide existence results for a game with local spillovers, where the payoff function satisfies both the convexity and the strategic substitutes property. They characterize the architecture of a pairwise stable network and of a pairwise equilibrium network. While Goyal and Joshi (2006) deal with undirected networks, Billand et al. (2012b) focus on directed networks. They also examine the model with global spillovers and the model with local spillovers, but they use the notion of Nash equilibrium to characterize the equilibrium networks.

Currarini (2007) also considers the impact of both positive and negative externalities on network formation. He investigates a game theoretic model of cooperation, in which critical structural features of an organization (which is represented by a connected network) depend on the sign of spillovers.

As already shown in Jackson and Wolinsky (1996) strongly efficient networks may not be stable. This potential conflict between stability and efficiency of networks is further analyzed in Dutta and Mutuswami (1997), and for directed communication networks in Dutta and Jackson (2000); see also Bala and Goyal (2000a) for the directed connections and hybrid connections models, and Jackson (2005) for a survey. Buechel and Hellmann (2012) analyze the role of positive and negative externalities of link formation, and relate inefficient outcomes to the externalities. They introduce the notion of over-connected (under-connected) networks – those which can be socially improved by the deletion (addition) of links. The authors prove that for positive externalities there is no stable network that is over-connected. Situations with negative externalities tend to induce over-connected networks, but for negative externalities no stable network can be under-connected only when some other conditions hold. Hellmann (2012) shows how externalities between links affect the existence and uniqueness of pairwise stable networks. He focuses on the properties of ordinal convexity/concavity and ordinal strategic complements/substitutes of utility functions on networks.

Bala and Goyal (2000a) formulate the network formation process as a non-cooperative game and examine Nash and strict Nash networks (no agent has an incentive to unilaterally add, remove or replace a link). Hojman and Szeidl (2008) analyze a decay model for Nash networks with a general benefit function, where the benefits depends on the network distance and agents who are farther away than a certain communication threshold yield no benefit. They show that Nash networks tend to have a core-periphery structure. Billand et al. (2012c) study the role of heterogeneity in the Jackson-Wolinsky framework by introducing two different benefit formulations: a model with the homogeneous decay parameter and heterogeneous players, and a model with heterogeneous links. They examine the existence and characterization of Nash networks and strict Nash networks in these models.

Haller and Sarangi (2005) develop a generalization of a non-cooperative model of network formation by Bala and Goyal (2000b) by introducing heterogeneity in the model. While Bala and Goyal (2000b) permit links to fail with a certain common probability, in Haller and Sarangi (2005) this probability of link failure can differ across links. The

authors focus on Nash networks, also for some alternative model specifications. In particular, they consider endogeneous link probabilities and show examples in which some novel Nash networks may arise due to negative link externalities. Haller (2012) considers a model of strategic network formation in which the endogeneously formed network is built around a pre-existing network. He uses a non-cooperative approach by Bala and Goyal (2000a), where a player's decision of adding or deleting a link is an unilateral decision. After introducing a benchmark model à la Galeotti et al. (2006), where the worth of information received by an agent from another agent depends only on the two agents involved, Haller (2012) studies network formation under internal constraints and Nash networks. He also considers two examples with negative network externalities, in which the values of information are endogeneously determined and depend on the network. This is in line with the idea that it is harder to access the information from an agent which has more direct neighbors.

1.3 The present paper

We investigate a model in which the benefits but also the negative externality resulting from an indirect neighbor's connectivity depend on exact network distance, and which thus differs from the settings studied both by Morrill (2011) and by Goyal and Joshi (2006). While Morrill (2011) considers "busyness" of only direct neighbors, Goyal and Joshi (2006) take into account degrees of all other agents, but they neither distinguish between indirect connections of different distances nor between indirect neighbors and agents that are never reached. In this sense, we remain close to the original connections model. In the latter, benefits depend on network distance and it seems natural that the same should be true for the negative externalities resulting from the connectivity of indirect neighbors. We thus consider a network formation game whose payoffs have a functional form close to the generalization of the connections model by Bloch and Jackson (2007) – the distance-based utility model, so as to facilitate comparison with the latter, but involves a penalty resulting from the degree of direct and indirect contacts. In our degree-distance-based connections model, we assume a two-variable (instead of one-variable) benefit function which depends on distances to and degrees of (direct and indirect) neighbors. We analyze pairwise stability and efficiency in order to understand how the negative externality modifies the results of the original connections model. There are natural assumptions with a straightforward economic interpretation, namely that decay is independent of degree and some conditions related to concavity/convexity of the benefit function with respect to degree, that would ensure that many of our results hold. However, since these conditions are not necessary we give weaker sufficient conditions on the benefit function.

Our way of expressing negative externalities is similar to the one advanced in examples given by Haller (2012). While his examples are presented in a non-cooperative approach with the analysis of Nash networks, we investigate an extension of the original connections model by Jackson and Wolinsky (1996) and study efficiency and pairwise stability, in which adding a link in a network is a bilateral decision by the two agents involved.

As it is typically the case in many network formation games including the original connections model, there is multiplicity of the pairwise stable structures. We do not give a complete characterization of the pairwise stable networks but focus on cases which allow us to compare outcomes with the original connections model. The complete graph,

the star and the empty network which were the pairwise stable networks identified in Jackson and Wolinsky (1996) are also stable for some parameter ranges in our model, but we also find new pairwise stable structures which could not arise in the original connections model.

For example, we find a variety of stable structures in which the distance between any pair of agents is at most two. In particular, there is a parameter range where neither the star nor the complete graph is stable but for which we find stable graphs in between the two (in the sense of link inclusion), which is not possible in the original connections model. These graphs, while stable, are never efficient. Structurally, these networks contain two types of nodes, highly connected ones which must provide much indirect benefits to all their direct neighbors due to the high degree penalty and nodes which may provide little indirect benefit but which are acceptable because the degree penalty is low. We can also exhibit examples of stable networks with short distances and even degree distribution. Indeed, it can be noted that to some extent, our model favors degree homogeneity in stable networks. Since high degrees are penalizing, agents who provide similar indirect benefits to their neighbors must have similar degrees. However, this observation needs to be tempered, since it remains possible for a node which is an efficient intermediary to have a considerably higher degree and remain attractive, as is the case in the star network.

Furthermore, in our setting, contrary to the Jackson and Wolinsky model, a pairwise stable network may have more than one component. This occurs only when decay is large so that indirect benefits have little value. Disconnected components may then be stable if all agents already have a degree high enough to make direct benefits from an additional link low. For all network sizes, such stable disconnected networks exist for some costs but the cost range becomes very small as the network size grows. Our results on stability and efficiency of disconnected structures are coherent with those of the co-author model, with which our model shares some features when decay is large.

In terms of tension between efficiency and stability, the Jackson and Wolinsky model always had agreement between the two in the low cost range where the complete network is uniquely pairwise stable, and potentially in an intermediate cost range where the star is efficient and pairwise stable, although not uniquely so. In our model, stability and efficiency are opposed precisely in the parameter range which is a stability basin for the star and the complete network. Indeed, in our setting, in contrast with the results in Jackson and Wolinsky, the complete graph is not necessarily efficient when it is stable. When the network is large this problem is generic. We also have the appearance of stable graphs between the star and the complete network which are never efficient when the number of agents is large. This is coherent with the general results of Buechel and Hellmann (2012) who show that in the presence of negative externalities a pairwise stable network can be overconnected in the sense that it strictly contains the efficient network. We note, however, that our model does not exhibit purely negative externalities, as defined by Buechel and Hellmann (2012) but combines both negative and positive externalities. These inefficiency results can be established even without characterizing the efficient network.

In terms of positive results about efficiency, we show, under an additional condition on the nature of the degree penalties, that the star is efficient when decay and connection

cost are not too great. The hypothesis about the degree penalty is related to the fact that marginal utility loss is greater for small degrees.

Besides the star, the complete, and the empty network, we also establish the efficiency of some structures that could never be efficient in the original connections model. Indeed, some of the disconnected structures shown to be pairwise stable in some parameter ranges can also be efficient in subsets of these.

The rest of the paper is structured as follows. In Section 2 we recapitulate some basic preliminaries on networks, the Jackson-Wolinsky connections model and some of its established extensions. In Section 3 we present our degree-distance-based modification of the connections model. Pairwise stability and efficiency in the model are studied in Sections 4 and 5, respectively. In our analysis of stability and efficiency, we begin by providing some illustrating examples in networks of small size, we then turn to the analysis of stability and efficiency in networks of any size and for different levels of decay and link cost. In Section 6 we mention some possible extensions.

2 The connections model and its modifications

In this section we first present the preliminaries on networks (see, e.g., Jackson and Wolinsky (1996); Jackson (2008)) and then briefly recapitulate some models related to our work: the connections model and the co-author model of Jackson and Wolinsky (1996), the distance-based model by Bloch and Jackson (2007), the model with degree-based utility functions by Morrill (2011), playing the field games and the local spillovers games by Goyal and Joshi (2006).

2.1 Definitions

Let $N = \{1, 2, \ldots, n\}$ denote the set of players (agents). A *network* g is a set of pairs $\{i, j\}$ denoted for convenience by ij, with $i, j \in N$, $i \neq j$,[1] where ij indicates the presence of a pairwise relationship and is referred to as a *link* between players i and j. Nodes i and j are directly connected if and only if $ij \in g$.

A *degree* $\eta_i(g)$ of agent i counts the number of links i has in g, i.e.,

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

We can identify two particular network relationships among players in N: the *empty network* $g^\emptyset$ without any link between players, and the *complete network* g^N which is the set of all subsets of N of size 2. The set of all possible networks g on N is $G := \{g | g \subseteq g^N\}$.

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.

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* in g connecting i_1 and i_K is a set of distinct nodes $\{i_1, i_2 \ldots, i_K\} \subseteq N(g)$ such that $\{i_1 i_2, i_2 i_3, \ldots, i_{K-1} i_K\} \subseteq g$.

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

[1] Loop ii is not a possibility in this setting.

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 $ij \in g'$.

A *star* g^* is a connected network in which there exists some node i (referred to as the *center* of the star) such that every link in the network involves node i.

The *value* of a graph is represented by $v : G \to \mathbb{R}$. By V we denote the set of all such functions. In what follows we will assume that the value of a graph is an aggregate of individual utilities, i.e., $v(g) = \sum_{i \in N} u_i(g)$, where $u_i : G \to \mathbb{R}$.

A network $g \subseteq g^N$ is *strongly efficient (SE)* if $v(g) \geq v(g')$ for all $g' \subseteq g^N$.

An *allocation rule* $Y : G \times V \to \mathbb{R}^N$ describes how the value of a network is distributed among the players. We will examine the allocation rule $Y_i(g) = u_i(g)$, which might correspond to models without side payment.

A network $g \in G$ is *pairwise stable (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) $\forall\ ij \notin g$, if $u_i(g) < u_i(g + ij)$ then $u_j(g) > u_j(g + ij)$.

2.2 The connections and co-author models

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

$$u_i^{JW}(g) = \sum_{j \neq i} \delta^{l_{ij}(g)} - c\eta_i(g) \tag{1}$$

where $0 < \delta < 1$ denotes the undiscounted valuation of a connection, $l_{ij}(g)$ denotes the distance between i and j in terms of the number of links in the shortest path between them in g (with $l_{ij}(g) = \infty$, if there is no path connecting i and j in g) and $c > 0$ determines the costs for a direct connection. Hence, the first sum determines the benefits agent i receives via direct and indirect connections, while the overall utility has to be reduced by the costs for the direct connections.

Jackson and Wolinsky (1996) (Proposition 1) show that the complete, the empty or the star graph can be uniquely strongly efficient (depending on c and δ). More precisely, they prove that the unique SE network in the symmetric connections model is:

(i) the complete network g^N if $c < \delta - \delta^2$
(ii) a star g^* if $\delta - \delta^2 < c < \delta + \frac{(n-2)\delta^2}{2}$
(iii) no links if $\delta + \frac{(n-2)\delta^2}{2} < c$.

Furthermore they examine pairwise stability in the symmetric connection model without side payments. By virtue of Jackson and Wolinsky (1996) (Proposition 2), in the symmetric connections model with $Y_i(g) = u_i^{JW}(g)$:

(i) A pairwise stable graph has at most one (non-empty) component.
(ii) For $c < \delta - \delta^2$, the unique PS network is the complete graph g^N.
(iii) For $\delta - \delta^2 < c < \delta$, a star g^* encompassing all players is PS, but not necessarily the unique PS graph.

(iv) For $\delta < c$, any PS network which is non-empty is such that every player has at least two links (and thus is inefficient).

Jackson and Wolinsky (1996) also present the *co-author model*, in which nodes are interpreted as researchers and a link represents a collaboration between two researchers. The utility function of each player i in network g is given by

$$u_i^{co}(g) = \sum_{j:ij\in g} w_i(n_i, j, n_j) - c(n_i) \tag{2}$$

where $w_i(n_i, j, n_j)$ is the utility of i derived from a link with j when i and j are involved in n_i and n_j projects, respectively, and $c(n_i)$ is the cost to i of maintaining n_i links. They analyze a specific version of this model where utility of i is given by

$$u_i^{co}(g) = \sum_{j:ij\in g} \left[\frac{1}{n_i} + \frac{1}{n_j} + \frac{1}{n_i n_j}\right] \tag{3}$$

for $n_i > 0$, and $u_i^{co}(g) = 0$ for $n_i = 0$.

Jackson and Wolinsky (1996) (Proposition 4) show that in the co-author model:

(i) If n is even, then the strongly efficient network is a graph consisting of $\frac{n}{2}$ separate pairs.

(ii) A pairwise stable network can be partitioned into fully intraconnected components, each of which has a different number of members. If m is the number of members of one such component and n is the next largest in size, then $m > n^2$.

2.3 Some established modifications of the connections model

Bloch and Jackson (2007) introduce an extension of the original connections model – the distance-based model, where the utility of agent i is given by

$$u_i^{dist}(g) = \sum_{j\neq i} f(l_{ij}(g)) - c\eta_i(g) \tag{4}$$

with f nonincreasing in $l_{ij}(g)$; see also Jackson (2008) for the presentation of this distance-based model.

Morrill (2011) models situations in which any new relationship causes negative externalities on the rest of the network. In his model, a payoff of each player from a link is a decreasing function of the number of links maintained by his partner. A utility function is *degree-based* if there exists a decreasing function ϕ such that

$$u_i^{deg}(g) = \sum_{j:ij\in g} \phi(\eta_j(g)) - c\eta_i(g) \tag{5}$$

Goyal and Joshi (2006) introduce two specific models of network formation. In the first one, called a *playing the field game*, the aggregate payoff of an agent depends only on the number of his links and the aggregate number of links of the remaining agents. More precisely, the gross payoff of each player i is given by

$$\pi_i^{pfg}(g) = \Phi(\eta_i(g), L(g_{-i})) \tag{6}$$

and its net payoff is given by

$$u_i^{pfg}(g) = \Phi(\eta_i(g), L(g_{-i})) - c\eta_i(g) \tag{7}$$

where $L(g_{-i}) = \sum_{j \neq i} \eta_j(g_{-i})$. It is assumed that for all $L(g_{-i})$, $\Phi(k, L(g_{-i}))$ is strictly increasing in own links k.

The second game investigated in Goyal and Joshi (2006) is called a *local spillovers game*. In that game, the aggregate payoff of an agent depends on the distribution of links of all players and the identity of neighbors. More precisely, the aggregate gross payoff of each player i is given by

$$\pi_i^{lsg}(g) = \Psi_1(\eta_i(g)) + \sum_{j:ij \in g} \Psi_2(\eta_j(g)) + \sum_{j:ij \notin g} \Psi_3(\eta_j(g)). \tag{8}$$

3 A degree-distance-based connections model

In the original connections model, forming an additional connection may only induce a positive externality on other players as it eventually shortens the distance to other agents. In contrast to this approach, we suggest a modification that generates also negative externalities by increasing connectivity. In this model, similar as in Jackson and Wolinsky (1996) every agent benefits from his direct and indirect connections, but it is additionally assumed that the higher degree of a direct or indirect partner, the less valuable this connection is to the agent. While any information received from indirect contacts is valuable, it may be less valuable when it is shared with many other individuals.

In order to conserve the close relation to the original connections model, we follow existing generalizations of Jackson and Wolinsky (1996), by considering the utility of agent i given by

$$\widetilde{u}_i(g) = \sum_{j \neq i} b(l_{ij}(g), \eta_j(g)) - c\eta_i(g) \tag{9}$$

where $b : \{1, ..., n-1\}^2 \to \mathbb{R}^+$ is the net benefit that an agent receives from the direct and indirect connections. It is assumed that for all $l_{ij}(g)$, $b(l_{ij}(g), k)$ is nonincreasing in degree k, and for all $\eta_j(g)$, $b(l, \eta_j(g))$ is nonincreasing in distance l. Moreover, if there is no path connecting i and j in g, i.e., if $l_{ij}(g) = \infty$, then we set $b(\infty, \eta_j) = 0$ for every $\eta_j \in \{0, 1, ..., n-1\}$. In particular, $\widetilde{u}_i(g^{\emptyset}) = 0$ for every $i \in N$.

In the original connections model the benefit term is covered by a single parameter δ. If we expressed the benefit function in our model using parameters that regulate decay with distance and utility loss due to degree, we could write

$$b(l+1, \eta_j(g)) = \delta_{l,\eta_j(g)} b(l, \eta_j(g)), \quad b(l, \eta_j(g)+1) = c_{l,\eta_j(g)} b(l, \eta_j(g))$$

where $\delta_{l,\eta_j(g)} \in (0,1)$ expresses the decay between distance l and $(l+1)$ for a fixed degree $\eta_j(g)$, and $c_{l,\eta_j(g)} \in (0,1)$ expresses the utility loss due to an additional link increasing the degree from $\eta_j(g)$ to $(\eta_j(g)+1)$ for a fixed distance l. This gives much versatility, in particular, decay does not need to be constant with distance.

Since in the paper we aim to analyze negative externalities resulting from the connectivity of direct and indirect neighbors, we will be assuming that the benefit function b

is decreasing in degree (and in distance), except when mentioning explicitly the original connections model as a particular case of the degree-distance-based model.

In our model, the dispersion of information is negatively influenced by the number of direct links a (distant) neighbor has, that is, the busyness of a potential neighbor influences the valuation to form a link with him. Consider a co-author relationship. On the one hand, the more co-authors an author has, the more knowledge may spill over from the projects he is involved in. On the other hand, the more connections the author maintains, the busier he is so that he may not be able to dedicate enough time to every co-author to transfer all information. Hence, the model incorporates positive and negative externalities of link formation. Morrill (2011) introduced a model which applies this perception to direct neighbors. Our framework generalizes the degree-based model by Morrill (2011), as it covers in addition externalities from indirect connections. We can write

$$\phi(\eta_j(g)) = b(1, \eta_j(g)), \text{ for all } \eta_j(g) \in \{1, ..., n-1\} \tag{10}$$

The generalized model defined by (9) also extends the distance-based model considered in Bloch and Jackson (2007) and recapitulated in (4). In other words, we consider a two-variable (instead of one-variable) benefit function.

Concerning the relation of our model with the model (8) of Goyal and Joshi (2006), we observe that neither model is generalized by the other. However, in the particular case when $b(l, \eta_j(g)) = 0$ for all $l > 1$, we have

$$b(1, \eta_j(g)) =: \Psi_2(\eta_j(g)) \text{ for every } \eta_j(g)$$

Furthermore, without decay, i.e., if $b(l, \eta_j(g)) = b(l', \eta_j(g))$ for all $l, l' < \infty$, we have

$$\widetilde{u}_i(g) = \sum_{j:ij \in g} b(l, \eta_j(g)) + \sum_{j:ij \notin g, l<\infty} b(l, \eta_j(g))$$

which is similar to the model (8), except that we do not take into account own degree, but we do distinguish between agents who are in our own connected component and those who are not.

In some of our examples, we will use a specific form of the degree-distance-based utility, which is very close to the definition of the original connections model, except of incorporating an additional information on the degree of direct and indirect neighbors. More precisely, in order to illustrate some of our results, the following utility of agent i will be used:

$$u_i(g) = \sum_{j \neq i} \frac{1}{1 + \eta_j(g)} \delta^{l_{ij}(g)} - c\eta_i(g) \tag{11}$$

that is, we will set $b(l_{ij}(g), \eta_j(g)) = \frac{1}{1+\eta_j(g)} \delta^{l_{ij}(g)}$.

4 Pairwise stability in the model

4.1 Stability of the star, the complete graph and the empty graph

Next, we examine pairwise stability (PS) in the model. In order to compare results in our model with those of Jackson and Wolinsky (1996), we start by analyzing the stability

of the architectures which were prominent there: the empty network, the star and the complete graph. Furthermore, we check if a non-empty PS network must be connected. In the connections model of Jackson and Wolinsky (1996), any pairwise stable graph has at most one (non-empty) component. We will show that it is not necessarily the case in our model. After establishing the stability conditions for $g^{\emptyset}$, g^* and g^N, we will look for other PS structures. We will begin by considering networks in which all agents are at distance at most 2 from each other. In such a network, the benefit of adding a link to an agent of degree $k-1$ is

$$\widetilde{f}(k) := b(1,k) - b(2,k-1) \text{ for } k \in \{2,...,n-1\}, \text{ and } \widetilde{f}(1) := b(1,1) \tag{12}$$

Note that in the original connections model $\widetilde{f}(k) = \delta - \delta^2$ for each k.

Let us show that whether the function $\widetilde{f}$ is increasing or decreasing can be related to some fairly weak and natural assumptions about decay and the nature of the degree penalty.

Proposition 1 *Assume that the decay is defined by* $\frac{b(l+1,k)}{b(l,k)} := \delta_l$ *for every* k*, that is, decay may depend on the distance* l *but not on the degree* k*. Then:*

- $\widetilde{f}(k)$ *is decreasing whenever* $b(l,k)$ *is concave in the degree* k*.*
- *If* $b(l,k)$ *is convex in the degree* k*, then there exists a level of decay* $0 < \delta_m < 1$ *such that* $\widetilde{f}(k)$ *is decreasing whenever* $\delta_1 \leq \delta_m$*, and there exists a* $0 < \delta_M < 1$ *such that* $\widetilde{f}(k)$ *is increasing whenever* $\delta_1 \geq \delta_M$*.*

Proof: Consider $\widetilde{f}(k) - \widetilde{f}(k+1) = b(1,k) - b(2,k-1) - (b(1,k+1) - b(2,k)) = b(1,k) - b(1,k+1) - \delta_1 (b(1,k-1) - b(1,k))$. This quantity is positive if and only if $\frac{b(1,k)-b(1,k+1)}{b(1,k-1)-b(1,k)} \geq \delta_1$. If b is concave in degree, then $\frac{b(1,k)-b(1,k+1)}{b(1,k-1)-b(1,k)} \geq 1 \geq \delta_1$, so that $\widetilde{f}(k)$ will always be decreasing. Now, if b is convex, then $0 < \frac{b(1,k)-b(1,k+1)}{b(1,k-1)-b(1,k)} < 1$ for all $k \in 1,...,n-1$. Define $\delta_M =: \max_k \frac{b(1,k)-b(1,k+1)}{b(1,k-1)-b(1,k)}$ and $\delta_m =: \min_k \frac{b(1,k)-b(1,k+1)}{b(1,k-1)-b(1,k)}$, then $\widetilde{f}(k)$ is decreasing for $\delta_1 \leq \delta_m$ and increasing when $\delta_1 \geq \delta_M$. ■

Consequently, the conditions in Propositions 3, 4 and 5 are verified in particular under the assumptions stated in Proposition 1, although these conditions are not necessary. We note that some results, such as the existence of simultaneous stability of the star and the complete network (see Proposition 2) are compatible with a b convex in degree, for which the function $\widetilde{f}$ exhibits a wider range of behaviors.

We have the following results on pairwise stability of the three prominent architectures.

Proposition 2 *In the degree-distance-based connections model defined by (9):*

(i) The empty network $g^{\emptyset}$ *is PS if and only if* $\widetilde{f}(1) \leq c$*.*
(ii) The star g^* *with* $n \geq 3$ *encompassing all players is PS if and only if*

$$\widetilde{f}(2) \leq c \leq \min\left(\widetilde{f}(1), b(1,n-1) + (n-2)b(2,1)\right) \tag{13}$$

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

$$c \leq \widetilde{f}(n-1) \tag{14}$$

(iv) The unique PS network is the complete network g^N if

$$c < \min_{1 \leq \eta_k \leq n-2} \widetilde{f}(\eta_k + 1) \tag{15}$$

(v) g^* and g^N are simultaneously PS if and only if $\widetilde{f}(2) \leq c \leq \widetilde{f}(n-1)$. Whenever $\widetilde{f}(n-1) < c < \widetilde{f}(2)$, neither the complete graph nor the star is PS.

(vi) A PS network may have more than one (non-empty) component.

Proof: (i) Consider any two agents $i, j \in g^\emptyset$. $\widetilde{u}_i(g^\emptyset + ij) - \widetilde{u}_i(g^\emptyset) = \widetilde{u}_j(g^\emptyset + ij) - \widetilde{u}_j(g^\emptyset) = b(1,1) - c = \widetilde{f}(1) - c$. Hence, if $\widetilde{f}(1) > c$, then both players profit from establishing the link, and therefore $g^\emptyset$ is not PS. If $\widetilde{f}(1) \leq c$, then $\widetilde{u}_i(g^\emptyset + ij) - \widetilde{u}_i(g^\emptyset) \leq 0$ and $\widetilde{u}_j(g^\emptyset + ij) - \widetilde{u}_j(g^\emptyset) \leq 0$ which means that $g^\emptyset$ is PS.

(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) $\widetilde{u}_i(g^*) - \widetilde{u}_i(g^* \setminus ij) \geq 0$ and (B) $\widetilde{u}_j(g^*) - \widetilde{u}_j(g^* \setminus ij) \geq 0$ and (C) $\widetilde{u}_j(g^* + jk) - \widetilde{u}_j(g^*) \leq 0$.
(A): $\widetilde{u}_i(g^*) - \widetilde{u}_i(g^* \setminus ij) = b(1,1) - c = \widetilde{f}(1) - c$. Hence, (A) holds iff $\widetilde{f}(1) \geq c$.
(B): $\widetilde{u}_j(g^*) - \widetilde{u}_j(g^* \setminus ij) = b(1, n-1) + (n-2)b(2,1) - c$. Hence, (B) holds iff $b(1, n-1) + (n-2)b(2,1) \geq c$.
(C): $\widetilde{u}_j(g^* + jk) - \widetilde{u}_j(g^*) = b(1,2) - b(2,1) - c = \widetilde{f}(2) - c$. Hence, (C) holds iff $\widetilde{f}(2) \leq c$.

(iii) Let $n \geq 3$. Consider any two agents $i, j \in g^N$. We have $\widetilde{u}_i(g^N) - \widetilde{u}_i(g^N - ij) = \widetilde{u}_j(g^N) - \widetilde{u}_j(g^N - ij) = b(1, n-1) - b(2, n-2) - c = \widetilde{f}(n-1) - c$.

(iv) Consider two arbitrary agents i and j, $j \neq i$ such that $ij \notin g$, $\eta_i > 0$ and $\eta_j > 0$. Then we have $\widetilde{u}_i(g + ij) - \widetilde{u}_i(g) \geq b(1, \eta_j + 1) - b(2, \eta_j) - c = \widetilde{f}(\eta_j + 1) - c$ and $\widetilde{u}_j(g + ij) - \widetilde{u}_j(g) \geq \widetilde{f}(\eta_i + 1) - c$. If $\eta_i \eta_j = 0$, then $\widetilde{u}_i(g + ij) - \widetilde{u}_i(g) \geq b(1, \eta_j + 1) - c$ and $\widetilde{u}_j(g + ij) - \widetilde{u}_j(g) \geq b(1, \eta_i + 1) - c$. Hence, if $c < \min_{1 \leq \eta_k \leq n-2} \widetilde{f}(\eta_k + 1)$, then any two agents who are not directly connected benefit from forming a link.

(v) This comes immediately from results (ii) and (iii).

(vi) The general existence of pairwise stable disconnected structures is given in Proposition 6. In small networks we can also find other types of architectures. Consider for example g given in Figure 1.

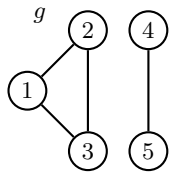

Fig. 1. A PS network with two components in the degree-distance-based connections model

Consider the utility function given by (11). Network g is PS if $\frac{1}{4}\delta + \frac{2}{3}\delta^2 < c \leq \frac{1}{3}\delta - \frac{1}{2}\delta^2$, e.g., for $\delta = \frac{1}{15}$ and $c = \frac{107}{5400}$. Since we have two groups of symmetric agents $(1, 2, 3$ and $4, 5)$, we only need to calculate the following:
$u_2(g) - u_2(g \setminus 23) = \frac{2}{3}\delta - 2c - \left(\frac{1}{3}\delta + \frac{1}{2}\delta^2 - c\right) = \frac{1}{3}\delta - \frac{1}{2}\delta^2 - c \geq 0$
$u_4(g) - u_4(g \setminus 45) = \frac{1}{2}\delta - c \geq 0$
$u_2(g + 24) - u_2(g) = \frac{1}{3}\delta + \frac{1}{2}\delta^2 - c$
$u_4(g + 24) - u_4(g) = \frac{1}{4}\delta + \frac{2}{3}\delta^2 - c$
Note that if $\frac{1}{3}\delta - \frac{1}{2}\delta^2 - c \geq 0$, then $\frac{1}{2}\delta - c > 0$ and $\frac{1}{3}\delta + \frac{1}{2}\delta^2 - c > 0$. Hence, g will be PS if $\frac{1}{3}\delta - \frac{1}{2}\delta^2 \geq c$ and $\frac{1}{4}\delta + \frac{2}{3}\delta^2 < c$.

■

Note that Proposition 2 confirms, in particular, the results on pairwise stability of g^N, g^* and $g^\emptyset$ for the Jackson and Wolinsky model. Assume now that the benefit function b is strictly decreasing in degree – for simplicity, take the degree-distance-based model given by (11). Naturally, since $\delta - \delta^2 > \frac{\delta}{n} - \frac{\delta^2}{n-1}$ for any $n \geq 2$, if g^N is PS in our model, then it is also PS in the Jackson and Wolinsky framework under the same parameters. Moreover, if $g^\emptyset$ is PS in the original connections model, then it is also PS in our framework. Roughly speaking, the costs under which the star g^* is PS in our model are rather lower than the costs under which g^* is PS in the Jackson and Wolinsky model under the same δ and n. For $\delta < \frac{1}{2}$, g^* cannot be PS in both frameworks in the same time. For $\delta \geq \frac{1}{2}$ and $n > 3$ such an overlap of costs under which the star is PS is non-empty.

Proposition 2(v) shows the existence of a region where the star and the complete graph are simultaneously PS. This could never occur in the original connections model where the regions of stability for these two structures were disjoint. However, we should note that, for instance, for the degree-distance-based utility given by (11), $\lim_{n\to\infty} \widetilde{f}(n-1) = 0$ so that the possible cost range for which the complete network and the star are simultaneously PS is very small in large networks.

Figure 2 illustrates the pairwise stability regions for the three simple architectures for the model given by (11) with $n = 9$, $\delta \in (0, 1)$ and $c \in (0, 0.5]$. In this figure, the green area indicates the stability region of $g^\emptyset$, the red area the one for g^N and the yellow area the one for g^*. The overlapping (quite small) orange area indicates the parameter region in which g^N and g^* are simultaneously PS, and the white area in which none of the three simple structures are PS.

4.2 Other pairwise stable structures in the degree-distance-based model: examples and illustration

Next, we will be interested in PS architectures, other than those analyzed in the previous section that can exist in the degree-distance-based connections model. An example of a PS structure that can occur in the white area in Figure 2 is the "windmill or flower" structure shown in Figure 3.

To get a more precise feeling for the range of parameters in which such "windmill or flower" structure is stable, consider Figure 4. Compared to Figure 2, Figure 4 is zoomed in and shows an additional blue area in which the "windmill or flower" structure is PS. The green, yellow, red and white areas have the same meaning as before.

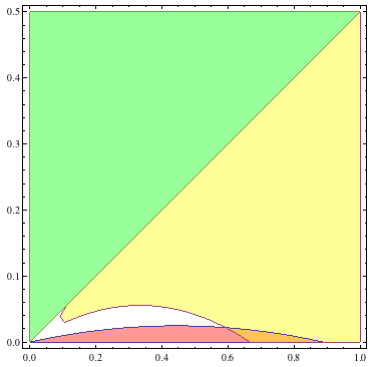

Fig. 2. PS regions in the degree-distance-based connections model given by (11) ($n = 9$): only $g^{\emptyset}$ (green area), only g^* (yellow area), only g^N (red area), g^* and g^N simultaneously (orange area), none of these three (white area)

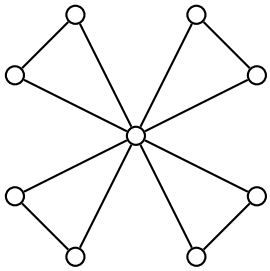

Fig. 3. "Windmill or flower" as an example of a PS network in the degree-distance-based connections model

Consider a very simple example of network formation. Imagine a call for scientific proposals of 2-person projects opened by a research agency. Suppose the call does not require submitting only one project by researcher, but two researchers can submit together only one project. Usually preparing an application for research grants is time and energy consuming, and working on two applications in the same time can considerably decrease the effort that a person can put on one project. The amount of funding proposed by the research agency depends on the quality of the application. Consider a small team of 3 researchers who decide how many applications to prepare. The example is obviously very simple and does not stress that much the issue of negative externalities coming from indirect connections, but it does show the intuition and flavor of situations where our model can be useful. Figure 5 presents all PS networks with the corresponding parameters for such a collaboration formation modeled by (11). Note that a network with one link can be PS in our framework (with b being strictly decreasing in degree), contrary to the Jackson and Wolinsky model. It is possible that under relatively high costs and low valuation of a connection, two researchers prefer to work on a project, and while each of them would be also interested in preparing a second project with another person, the latter will prefer to save his time and energy and not to work on the project at all. The involvement in a project with a very busy colleague who already works on another application could mean that the costs finally exceed the benefits, despite the fact that the researcher could also benefit a bit from the 'indirect project'. Figure 5 confirms nicely Proposition 2(v). In particular, g^* and g^N are simultaneously PS only for $c = \frac{\delta}{3} - \frac{\delta^2}{2}$.

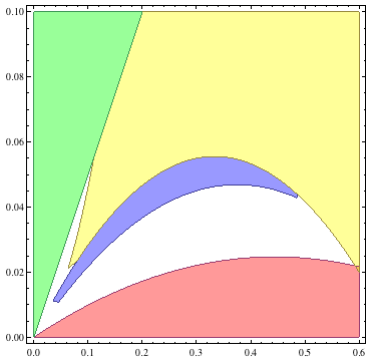

Fig. 4. PS regions in the degree-distance-based connections model given by (11) ($n = 9$): "windmill or flower" (blue region), other regions indicate the same PS networks as in Figure 2

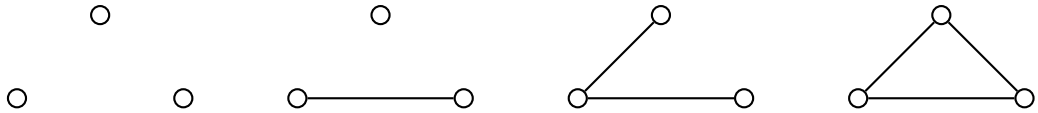

Fig. 5. PS networks in the degree-distance-based connections model given by (11) (from left to right): (i) $\delta \leq 2c$, (ii) $0 < \delta < \frac{1}{3}$ and $\frac{\delta}{3} + \frac{\delta^2}{2} < c \leq \frac{\delta}{2}$, (iii) $(0 < \delta < \frac{1}{3},\ \frac{\delta}{3} - \frac{\delta^2}{2} \leq c \leq \frac{\delta}{3} + \frac{\delta^2}{2})$ or $(\frac{1}{3} \leq \delta < 1,\ \frac{\delta}{3} - \frac{\delta^2}{2} \leq c \leq \frac{\delta}{2})$, (iv) $0 < c \leq \frac{\delta}{3} - \frac{\delta^2}{2}$

Let us now illustrate, for some small network sizes, examples of other PS structures that can appear. In some cases, these architectures are not stable in the original connections model which make them interesting per se. Many of the architectures that we see in these examples can also be shown to exist in a network of arbitrary size n in some parameter range, as will be shown in a later section. Figures 6 and 7 show some examples of different PS structures for $n = 5$ and $n = 6$, respectively, for the model given by (11) with $\delta = \frac{1}{15}$ and $c = \frac{107}{5400}$. From among these examples, only the two regular networks (with the degree $\eta_i = 2$ for every $i \in N$) can be PS under some parameters in the original connections model. The remaining networks which are PS in our framework could be never PS in the original connections model. Note that four of these networks contain two components. In Figure 6 these are the first network (on the left) that has been used in the proof of Proposition 2(vi), and the second network with one isolated player and four players, each having the degree equal to 2. In Figure 7 these are the first and the second network, both having two non-empty components.

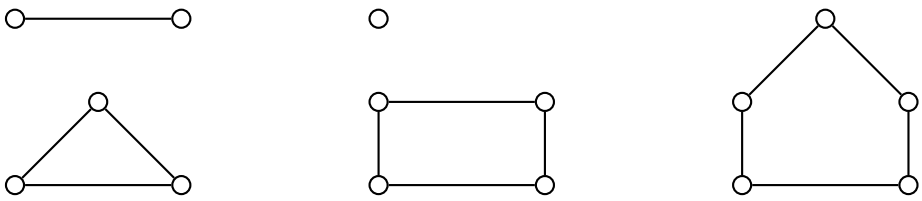

Fig. 6. Some PS networks in the degree-distance-based connections model given by (11) for $n = 5$, $\delta = \frac{1}{15}$ and $c = \frac{107}{5400}$

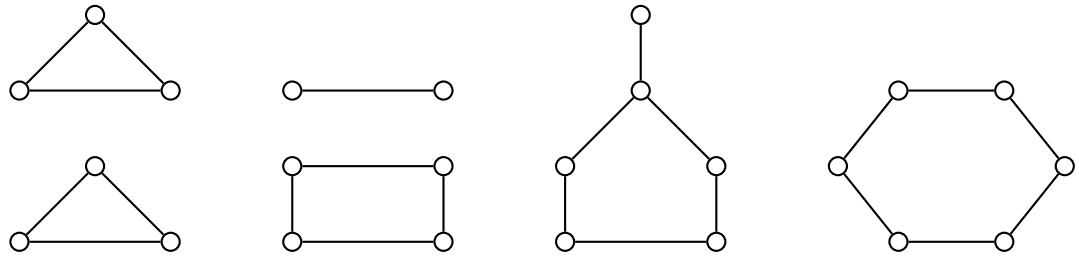

Fig. 7. Some PS networks in the degree-distance-based connections model given by (11) for $n = 6$, $\delta = \frac{1}{15}$ and $c = \frac{107}{5400}$

4.3 Stability analysis for intermediate levels of decay

These examples of networks of small size that can be PS in our framework illustrate some of the general results about architectures that can be PS for arbitrarily large networks and for some values of the decay and cost parameters. In contrast, most of them can never be PS in the original connections model, We have already seen that there exist PS structures other than the star and the complete graph, in particular, ones that are intermediary between these two in the sense of link inclusion. That is, there exists g PS such that $g^* \subset g \subset g^N$. This also indicates a difference with Jackson and Wolinsky (1996) where no intermediary network between the star and the complete graph could possibly be stable. In Figure 3 we have already shown "Windmill or flower" which is an example of a PS structure that contains a star but which also comprises links between the nodes in the periphery. We prove the following result.

Proposition 3 *Assume that the benefit b and the cost c are such that there exists an $\eta_m \in N$ such that $\widetilde{f}(\eta_m + 1) < c \leq \widetilde{f}(\eta_m)$. Assume that moreover, n is such that $\widetilde{f}(n-1) + (n - 1 - \eta_m - (\eta_m - 1)^2)\,(b(2, \eta_m) - b(3, \eta_m)) > c$. Then any graph satisfying the following property is PS. One node i, the center, is linked to all other nodes. Every node other than the center, $j \neq i$, has the same degree η_m.*

Proof: Let us show that under the conditions stated the described structure is pairwise stable. Every peripheral node has degree η_m. Breaking the link to another peripheral node does not modify the benefits from nodes that can be reached at distance 2. Only direct benefits are lost. None of those peripheral nodes wants to break a local link with another peripheral agent of degree η_m because $\widetilde{u}_j(g) - \widetilde{u}_j(g - lj) = b(1, \eta_m) - b(2, \eta_m - 1) - c = \widetilde{f}(\eta_m) - c \geq 0$ by assumption.
No agent wants to add a link to a peripheral agent whose degree in g is η_m because doing this would decrease the utility, as $\widetilde{u}_j(g + jm) - \widetilde{u}_j(g) = b(1, \eta_m + 1) - b(2, \eta_m) - c = \widetilde{f}(\eta_m + 1) - c < 0$ by assumption.
Every peripheral node has an incentive to maintain a link to the center due to the following fact: without a link to the center, each agent can reach at most $(\eta_m - 1)^2$ nodes at distance 2. By breaking with the center, $n-1-\eta_m-(\eta_m-1)^2$ nodes would move to distance 3. Thus $\widetilde{u}_j(g) - \widetilde{u}_j(g - ij) \geq b(1, n-1) - b(2, n-2) + (n - 1 - \eta_m - (\eta_m - 1)^2)\,(b(2, \eta_m) - b(3, \eta_m)) - c = \widetilde{f}(n-1) + (n - 1 - \eta_m - (\eta_m - 1)^2)\,(b(2, \eta_m) - b(3, \eta_m)) - c$.
We show that for given b and c, we can find n such that

$$\widetilde{f}(n-1) + \left(n - 1 - \eta_m - (\eta_m - 1)^2\right)(b(2, \eta_m) - b(3, \eta_m)) - c > 0 \tag{16}$$

Note that $b(2,\eta_m) > b(3,\eta_m)$. Moreover, $b(1,k) \geq 0$ for all k, and therefore from (12) we have $\widetilde{f}(n-1) \geq -b(2,n-2) \geq -b(2,1)$. Hence, $c - \widetilde{f}(n-1) \leq c + b(2,1)$ and then every n such that $n > \frac{c+b(2,1)}{b(2,\eta_m)-b(3,\eta_m)} + 1 + \eta_m + (\eta_m - 1)^2$ satisfies (16). ■

Note that for the original connections model, where $\widetilde{f}$ is constant, the first assumption of Proposition 3 is never satisfied.

The stable structure with inhomogeneous degrees given in Proposition 3 illustrates the fact that in the degree-distance based model, a node can be attractive either because it is not too "busy" with other neighbors or because it is highly connected and it provides indirect benefits.

We also have examples of stable structures where the degree distribution is essentially homogeneous.

Proposition 4 *Let g be a network such that $l_{ij}(g) \leq 2$ for all i,j, and there are at least $(n-1)$ nodes with identical degree k and at most one node j such that $\eta_j < k$. Then g is PS in the cost range $\widetilde{f}(k+1) < c < \widetilde{f}(k)$ and $\widetilde{f}(\eta_j - 1) > c$. In particular, for the utility function given by (11), this cost range is non empty if and only if $k > \frac{\delta+\sqrt{\delta}}{1-\delta}$ and $\widetilde{f}(\eta_j - 1) > c$.*

Proof: No agent wants to add a link to an agent with degree k since $\widetilde{f}(k+1) < c$. Moreover, no agent wants to delete a link to an agent with degree k since the loss of utility is at least $\widetilde{f}(k)$ which exceeds the saving of c. Note that there may be a single agent j such that $\eta_j < k$. This agent does not want to form a link to any other agent since all other agents have degree k. No agent wants to drop a link to agent j since $\widetilde{f}(\eta_j - 1) > c$. For the utility function given by (11), there exists a cost c such that $\widetilde{f}(k+1) < c < \widetilde{f}(k)$ if k belongs to the interval where $\widetilde{f}(k)$ is decreasing, i.e., on $[\frac{\delta+\sqrt{\delta}}{1-\delta}, n-1]$. ■

Structures satisfying the properties stated in the above proposition clearly exist. We can give some (non exhaustive) examples.

Proposition 5 *Pairwise stable networks with equal degree: let $n = M^2$. Then there exists a pairwise stable network g such that $l_{ij}(g) \leq 2$ for all i,j and all nodes have identical degree equal to $k = 2M - 2$ in the cost range $\widetilde{f}(k) < c < \widetilde{f}(k-1)$.*

Proof: It is sufficient to show that there exists a network g such that $l_{ij}(g) \leq 2$ for all i,j, and $l_{ij}(g - lm) = l_{ij}(g)$ for all i,j,l,m. Divide $n = M^2$ into M disjoint sets of size M. Let each agent be linked to all agents in own group and to exactly one agent in every other group in such a way that each agent has exactly $(M-1)$ links outside his group. ■

The above results concerned stability for intermediate levels of decay (moderately high values of δ in the model given by (11)), for which indirect contacts are of interest for agents.

4.4 Pairwise stability for extreme values of the decay parameter

We will now consider pairwise stability in the two extreme cases where decay is very large (i.e., δ is very small for the model given by (11)), or where decay is very small (i.e., δ is close to one for the model defined by (11)). We will show that in both these cases, the degree-distance-based model can exhibit a very large number of different PS architectures, which are only restricted by the fact that (most) nodes must have the same degree. In the case where decay is large, the pairwise stable architectures include a number of disconnected structures.

The case of the benefit function where decay is very large can be expressed by the condition $b(1,k) >>> b(2,k)$ for every k.

Proposition 6 *Let n be a fixed network size. Let $\epsilon > 0$. Then there exists $\underline{b} > 0$ such that for any function b with $b(2,1) < \underline{b}$ and any cost such that $b(1,r+1) + \epsilon < c < b(1,r) - \epsilon$, a network g satisfying the following properties is pairwise stable: in g, $n-k$ nodes, where $k \leq r$, have an identical degree r. The remaining k nodes are all linked to each other.*

Proof: Fix $\epsilon > 0$. Consider the maximal indirect benefit an agent can gain from a link. This benefit is bounded by $(n-2)b(2,1)$. Let $\underline{b} = \frac{\epsilon}{n-2}$. For any $0 < b(2,1) < \underline{b}$, the benefit of indirect links is inferior to ϵ. Basically, we can now neglect utility from indirect contacts. Let us establish that no pair of agents i, j in g can establish a mutually beneficial link. Let i be a node of degree $\eta_i = r$. Then $\widetilde{u}_j(g+ij) - \widetilde{u}_j(g) < b(1,m+1) + \epsilon - c < 0$. Therefore no agent wishes to form a link to an agent who already has degree r. Let $ij \in g$ with $\eta_i \leq r$ and $\eta_j \leq r$. Neither i nor j wishes to break this link: $\widetilde{u}_j(g-ij) - \widetilde{u}_j(g) < c + \epsilon - b(1,\eta_i) \leq c + \epsilon - b(1,r) < 0$. ■

We note that the family of PS networks described in Proposition 6 is very large and includes in particular all structures where agents have identical degrees, such as the circle or a generalized circle with agents linked to their m nearest neighbors. We note that there is also an abundance of disconnected structures that satisfy the condition stated in Proposition 6.

Corollary 1 *Under the conditions stated in Proposition 6, there exist disconnected PS networks.*

There is a large number of such disconnected structures. One example is that of a number of disconnected "islands" of identical size.

Corollary 2 *Consider a network of size n. Let m be a divisor of n. The network consisting of n/m completely connected components of size m is PS under the conditions stated in Proposition 6.*

The class of PS networks described in Proposition 6 exists for some decay values and for some cost range for every network size. However, in order for these structures to appear, decay must be large, and all the more so when n is large, as we have $(n-2)b(2,1) < \epsilon < b(1,m) \leq b(1,1)$. The possible values of the cost for which these structures exist shrink the larger n and m are. In practice, the most likely decay is very large so that indirect benefits are almost negligible. The larger the size of the components, the smaller is the

possible cost range. Among the possible structures, the one that we are most likely to see emerge in practice is thus that in which all agents have degree one or two, and the network size is not too large.

Let us now consider the other extreme case where decay is very small so that δ is close to one in the model defined by (11). In other words, we consider the benefit functions such that $b(l,k) \approx b(l',k)$ for all $k, l, l' \neq \infty$.

Proposition 7 *Let n be a fixed network size. Let $c > 0$. Then there exist $\epsilon > 0$ and b such that $|b(l,k) - b(l',k)| \leq \epsilon$ for all $k, l, l' \neq \infty$, and any network g satisfying the following properties is pairwise stable: g is minimally connected and satisfies $\min_{j \in N} b(1, \eta_j(g)) \geq c > (n-1)\epsilon$.*

Proof: Fix $c > 0$. What is the maximal benefit an agent can derive from an additional link in an already connected structure? This benefit is bounded above by $(n-2)\max_{1 \leq x \leq n-1}(b(1,x) - b(n-1,x)) \leq (n-2)\epsilon$. Since g is connected, any additional link provides a utility $\widetilde{u}_i(g+ij) - \widetilde{u}_i(g) \leq b(1, \eta_j + 1) - b(n-1, \eta_j) + (n-2)\epsilon - c \leq (n-2)\epsilon + b(1, \eta_j + 1) - b(n-1, \eta_j + 1) + b(n-1, \eta_j + 1) - b(n-1, \eta_j) - c \leq (n-1)\epsilon - c < 0$. Let us establish that no agent wants to remove a link. If $ij \in g$, then neither i nor j wishes to break this link. Indeed, because g is minimally connected, i and j are not in the same connected component in $g - ij$. Therefore $\widetilde{u}_i(g - ij) - \widetilde{u}_i(g) < c - b(1, \eta_j) < 0$. ■

When the decay is very small, minimally connected networks, i. e. those with a tree-structure are thus pairwise stable.

5 Efficiency in the model

Next, we analyze strong efficiency (SE) in the model, also sometimes referred to as efficiency in the paper. One of the crucial differences between our model and the original connections model is the enlargement of the possible types of SE architectures in the degree-distance-based model with respect to the Jackson and Wolinsky model. While in the latter, only $g^{\emptyset}$, g^* and g^N can be SE, in our model additional structures can be SE under specific parameters. To see this immediately, let us come back to the 3-researcher example concerning formation of collaboration to prepare 2-person applications for a grant. Figure 8 presents all strongly efficient networks in this collaboration formation modeled by (11).

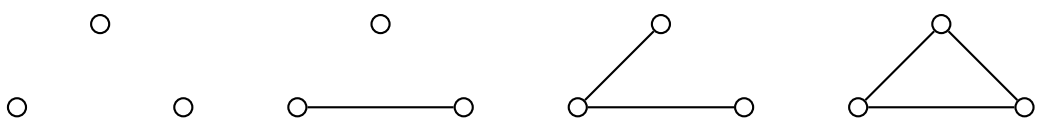

Fig. 8. Unique SE networks in the degree-distance-based connections model given by (11) (from left to right): (i) ($0 < \delta < \frac{1}{3}$ and $2c > \delta$) or ($\frac{1}{3} \leq \delta < 1$ and $2c > \frac{1}{2}\delta^2 + \frac{5}{6}\delta$), (ii) $0 < \delta < \frac{1}{3}$ and $\delta^2 + \frac{2}{3}\delta < 2c < \delta$, (iii) ($0 < \delta < \frac{1}{3}$, $\frac{1}{3}\delta - \delta^2 < 2c < \delta^2 + \frac{2}{3}\delta$) or ($\frac{1}{3} \leq \delta < 1$, $0 < 2c < \frac{1}{2}\delta^2 + \frac{5}{6}\delta$), (iv) $0 < \delta < \frac{1}{3}$ and $0 < 2c < \frac{1}{3}\delta - \delta^2$

Documents de Travail du Centre d'Economie de la Sorbonne - 2013.40

As we can see from Figure 8, it is possible that preparing only one 2-person project and leaving the third researcher "free" is the best collective option (under some restrictions on the parameters), i.e., it gives the highest aggregate of individual utilities. Indeed, it confirms the intuition that if preparing an application requires a lot of energy and time, and the benefit from a project that a researcher is not personally involved in so far is not attractive to that researcher, it would be better for the overall welfare of the team to submit only one (but one of a higher quality) research project.

Note that already in this simple 3-player example we can see an additional difference between our model and the original connections model. While in the latter, the cost ranges of PS and SE for the complete network coincide, in the degree-distance-based model defined by (11), g^N can be PS but not efficient: for $n = 3$ it is the case for $\frac{1}{3}\delta - \delta^2 < 2c < \frac{2}{3}\delta - \delta^2$ and $\delta < \frac{2}{3}$. Similarly to Jackson and Wolinsky, we exhibit the contradiction between stability and efficiency in the higher cost ranges: the empty network is PS but not SE for $\delta < 2c < \frac{1}{2}\delta^2 + \frac{5}{6}\delta$ and $\frac{1}{3} < \delta < 1$. The structure which could be neither PS nor SE in the original connections model – the network containing one link and one isolated player – is PS and SE in our model in the same cost range.

Another interesting observation in the model given by (11) and $n = 4$ is for instance that the line is the unique SE network if $0 < \delta < \frac{4-\sqrt{13}}{6}$ and $\frac{1}{3}\delta - \frac{1}{3}\delta^2 - \delta^3 < 2c < \frac{1}{3}\delta + \frac{5}{3}\delta^2 + \delta^3$.

After presenting the examples, we turn to the theoretical analysis.

We have seen that contrary to the original connections model, disconnected networks may be pairwise stable. Let us show that they can also be efficient.

Proposition 8 *Let n be even and fixed, and $\epsilon > 0$. There exists $\underline{b} > 0$ such that for any function b with $b(2,1) < \underline{b}$, the network described in Proposition 6, consisting of $n/2$ disjoint completely connected components with $m = 2$ is uniquely efficient in the cost range $\frac{b(1,1)+b(1,2)}{2} + \epsilon < c < b(1,1) - \epsilon$.*

Proof: Fix $\epsilon > 0$. Consider the maximal indirect benefit an agent can gain from a link. This benefit is bounded by $(n-2)b(2,1)$. Thus the total social utility from indirect links is bounded by $n(n-2)b(2,1)$. Let $\underline{b} = \frac{\epsilon}{n(n-2)}$. For any $0 < b(2,1) < \underline{b}$, the total social utility of indirect links is inferior to ϵ. Basically, we can now neglect utility from indirect contacts. Note then that if there are at least two nodes with no link, forming a link between them increases total utility since $c < b(1,1)$. Assume now that a network contains some node i such that $\eta_i = k \geq 2$. Let ij belong to the network. Let us show that removing ij is efficiency improving. Let $\eta_j \geq 1$. Since indirect benefits are negligible, only i and j lose by removing the link ij. This loss is $b(1,\eta_i) + b(1,\eta_j) \leq b(1,2) + b(1,1) < 2c$. Therefore removing the link of an agent with degree greater than one is efficiency improving. This shows that in this parameter range, an efficient network is such that each agent has exactly degree 1. This is achieved uniquely by the network consisting of disjoint completely connected components of size two. ■

When decay is very large, benefits from indirect contacts are negligible and the negative impact of an increased degree dominates. Thus it is not socially desirable to connect two components.

Proposition 9 shows that when network size is large, the complete network is not strongly efficient when it is stable.

Proposition 9 *Whenever* $(n-1)\left(b(1,n-2)-b(1,n-1)\right) > b(1,n-2)-b(2,n-2)$, *the complete network is not strongly efficient for any cost* $c>0$. *For the model defined by (11),* g^N *is not SE whenever* $n > \frac{1}{\delta}$. *In particular, the complete network is not strongly efficient when it is uniquely PS.*

Proof: We note that the total link cost is always greater in g^N than in $g^N - ij$ when $c > 0$. Assume $c = 0$. Let us consider the difference in total utility between g^N and $g^N - ij$, that is $\sum_{i=1}^{n}\left(u_i(g^N) - u_i(g^N - ij)\right)$. For agents i, j the utility loss is $b(1,n-1) - b(2,n-2)$. The remaining $n-2$ agents gain from the connectivity decrease of their direct neighbors i and j, $2(n-2)\left(b(1,n-2)-b(1,n-1)\right)$. In total, the change in social utility is $\sum_{i=1}^{n}\left(u_i(g^N - ij) - u_i(g^N)\right) = 2(n-2)\left(b(1,n-2)-b(1,n-1)\right) - 2b(1,n-1) + 2b(2,n-2) = 2(n-1)\left(b(1,n-2)-b(1,n-1)\right) + 2\left(b(2,n-2)-b(1,n-2)\right)$. This quantity is positive whenever $(n-1)\left(b(1,n-2)-b(1,n-1)\right) > b(1,n-2)-b(2,n-2)$ meaning that the complete network g^N is not strongly efficient. Except for very large decay (when $b(1,n-2) >>> b(2,n-2)$), the complete network is inefficient even when n is rather small. Solving $(n-1)\left(b(1,n-2)-b(1,n-1)\right) > b(1,n-2)-b(2,n-2)$ for the model defined by (11) leads equivalently to the condition $n > \frac{1}{\delta}$. ■

Proposition 10 says that a pairwise stable network comprised strictly between the star and the complete network cannot be SE. We show this without characterizing the efficient network.

Proposition 10 *Let* g *be a network in which* $l_{ij}(g) \leq 2$ *for all* i, j. *Let* $g^* \subset g \subset g^N$. *Then* g *is never efficient for sufficiently large* n. *In particular, for the model defined by (11),* g *is never efficient when* $n > \frac{2}{\delta}$, *even in the cost range where it is pairwise stable.*

Proof: A stable network g containing the star is such that all nodes except the center and possibly one more node have the same degree. Suppose that g is a network containing the star. Consider two nodes i, j, both different from the center which are linked in g. They are such that $1 < \eta_i = \eta_j \leq n-1$. Let us show that social utility increases when this link is removed. First consider the change in utility for i and j:
$u_i(g) - u_i(g-ij) + u_j(g) - u_j(g-ij) = 2\left(b(1,\eta_i) - b(2,\eta_i - 1) - c\right)$.
Now, both i and j have degree η_i. Thus there are $(2\eta_i - 2)$ links to other agents in the network. It does not matter here whether the neighbors of i and j are the same and gain both from the decrease in degree of i and j or two different neighbors one of whom gains from the decrease in degree of i and the other one from the decrease in degree of j. The total benefit to the neighbors of i and j, in $g-ij$ compared to g is $(2\eta_i - 2)\left(b(1,\eta_i - 1) - b(1,\eta_i)\right)$. All agents who are no immediate neighbors of i or j are at distance 2 from i and j. This concerns at least $M \geq n-2-2\eta_i$ agents. The total gain of these agents in $g-ij$ compared to g is $M\left(b(2,\eta_i - 1) - b(2,\eta_i)\right)$. We can now compare the social utility of g and $g-ij$. It is $\sum_{l=1}^{n}\left(u_l(g) - u_l(g-ij)\right) = 2\left(b(1,\eta_i) - b(2,\eta_i - 1) - c\right) - (2\eta_i - 2)\left(b(1,\eta_i - 1) - b(1,\eta_i)\right) - M\left(b(2,\eta_i - 1) - b(2,\eta_i)\right) \leq 2\left(b(1,\eta_i) - b(2,\eta_i - 1) - c\right) - (2\eta_i - 2)\left(b(1,\eta_i - 1) - b(1,\eta_i)\right) - (n-2-2\eta_i)\left(b(2,\eta_i - 1) - b(2,\eta_i)\right)$. This quantity is negative

for sufficiently large n. In particular, for the model defined by (11), this quantity is negative if $2(\eta_i-\delta(\eta_i+1))-(2\eta_i-2)-M\delta < -2\delta(\eta_i+1)+2-(n-2-2\eta_i)\delta < 0 \iff n\delta > 2$. Thus, whenever $n > \frac{2}{\delta}$ the network g is not efficient. ■

Next, we focus on the conditions for efficiency of the star.

Lemma 1 *Let g be a minimally connected network of size $n > 3$. Let J_1 be the set of nodes of degree 1 and j_1 the number of elements in this set. Then, whenever $j_1 > 2$, there exists a set S containing $1 \leq m \leq j_1$ nodes such that for all $i \in S$, $\eta_i \geq 3$ and $\sum_{i\in S} \alpha_i = j_1 - 2$, with $\alpha_i = \eta_i - 2$.*

Proof: We prove this by induction on the network size. Note that any minimally connected structure can be obtained from a minimally connected structure of size $n-1$ by adding one node and one link. The induction hypothesis is verified for $n = 3$. The only minimally connected structure is a line. There are two elements of degree 1, thus $j_1 = 2$. The set $S = \emptyset$ and all remaining nodes in $N - J_1$ have degree 2. Assume that the induction hypothesis is verified for all minimally connected structures of size n. Let g_n be a minimally connected network of size n, thus verifying the induction hypothesis. Any minimally connected structure of size $n+1$ can be obtained by adding one node $n+1$ and one link between $n+1$ and $i < n+1$ to some minimally connected network of size n. There are several possible cases. Either the link from $n+1$ goes to a node j such that $\eta_j(g_n) = 1$. Then the number of nodes with degree one does not change. The number of nodes in S does not change and $\eta_j(g_{n+1}) = 2$. The induction hypothesis is verified. If $n+1$ links to j such that $\eta_j(g_n) \geq 2$ then $j_1(g_{n+1}) = j_1(g_n) + 1$ and $j \in S(g_{n+1})$. Thus $\sum_{i\in S} \eta_i(g_{n+1}) = \sum_{i\in S} \eta_i(g_n) + 1 = j_1(g_n) + 1 = j_1(g_{n+1})$. ■

Let us define the function

$$h(m) = mb(1, m) + (n - 1 - m)b(2, m) \tag{17}$$

which is the maximal social utility that an agent with degree m provides to others. For the model defined by (11), we have $h(m) = m\frac{\delta}{1+m} + (n-1-m)\frac{\delta^2}{1+m}$.

Lemma 2 *Let g be a network with degree sequence $(\eta_i(g))_{i=1}^{i=n}$. Then the value of g is $v(g) \leq \sum_{i=1}^{i=n} h(\eta_i(g))$.*

Proof: The $\eta_i(g)$ immediate neighbors of i derive the utility $b(1, \eta_i(g))$ from the link to i. The remaining $n - 1 - \eta_i(g)$ nodes are at distance at least 2 from i and therefore the utility obtained from i is bounded by $b(2, \eta_i(g))$. ■

Let us assume the following conditions:

Condition 1 *Function h defined by (17) is decreasing.*

Condition 2 *For any $k > 2$, let $a + b = k$ and $a' + b' = k$. Then if $|a - b| \geq |a' - b'|$ then $h(a) + h(b) \geq h(a') + h(b')$.*

Condition 2 means that when keeping fixed the sum of degrees of two agents, the total maximal social utility provided by the two agents is greater the larger is the difference between the degrees of the agents. Note that Condition 2 is satisfied for the model defined by (11), when $n \geq \frac{1}{\delta}$.

Proposition 11 *Let g be a connected structure. Whenever h given by (17) satisfies Conditions 1 and 2, we have $v(g) \leq v(g^*)$.*

Proof: We will prove this proposition in two steps.

Step 1: *First we show that $v(g) \leq v(g^*)$ for any minimally connected structure g.*
Let g be a minimally connected structure. It is thus characterized by m, J_1 and S in Lemma 1, the degrees of the nodes in S and the distances between all pairs of nodes. By Lemma 1, all nodes in J_1 have degree one. There are m nodes in S whose degrees are $(2+\alpha_i)_{i=1}^{i=m}$ (without loss of generality we let the m nodes in S be $1, 2, ...m$). The remaining $n - j_1 - m$ nodes have degree 2. Therefore by Lemma 2, we have
$v(g) \leq \sum_{i=1}^{i=n} h(\eta_i(g)) = j_1 h(1) + \sum_{i=1}^{i=m} h(2+\alpha_i) + (n - j_1 - m)h(2)$.
We apply Lemma 2 to obtain $h(\alpha_i + \alpha_j + 2) + h(2) \geq h(\alpha_i + 2) + h(\alpha_j + 2)$. Thus we have $\sum_{i=1}^{i=m} h(\alpha_i+2) + (n-m-j_1)h(2) = \sum_{i=1}^{i=m-2} h(\alpha_i+2) + (n-m-j_1)h(2) + h(\alpha_{m-1}+2) + h(\alpha_m+2) \leq \sum_{i=1}^{i=m-2} h(\alpha_i+2) + (n-m-j_1)h(2) + h(\alpha_{m-1}+\alpha_m) + h(2) = \sum_{i=1}^{i=m-2} h(\alpha_i + 2) + h(\alpha_{m-1}+\alpha_m) + (n-m-j_1+1)h(2) \leq .... \leq h(\sum_{i=1}^{i=n} \alpha_i + 2) + (n - j_1 - 1)h(2)$.
Then, applying again repeatedly Lemma 2, $v(g) \leq j_1 h(1) + h(\sum_{i=1}^{i=n} \alpha_i + 2) + (n - j_1 - 1)h(2) = j_1 h(1) + h(j_1) + (n - j_1 - 1)h(2) = j_1 h(1) + (n - j_1 - 2)h(2) + h(2) + h(j_1) \leq j_1 g(1) + (n - j_1 - 2)h(2) + h(1) + h(j_1 + 1) = (j_1 + 1)h(1) + h(j_1 + 1) + (n - j_1 - 2)h(2) \leq ... \leq (n-1)h(1) + h(n-1) = v(g^*)$.

Step 2: *Next, we prove that $v(g) \leq v(g^*)$ for any connected structure g.*
As we have shown, any minimally connected structure g has a degree sequence $(\eta_i)_{i=1}^{i=n}$ such that $v(g) \leq \sum_{i=1}^{i=n} h(\eta_i) \leq v(g^*)$. Any connected structure g_K is a superset of a minimally connected network g. Let $g \subset g_K$ and let $\mu_i = \eta_i(g_K) - \eta_i(g) \geq 0$. By Lemma 2 we have $v(g_K) \leq \sum_{i=1}^{i=n} h(\eta_i(g_K))$. We will show that $\sum_{i=1}^{i=n} h(\eta_i(g_K)) \leq \sum_{i=1}^{i=n} h(\eta_i(g))$. Let us consider h being decreasing. For the model given by (11), we verify that for all $m > 1$, $h(m+1) - h(m) \leq 0 \iff \delta \geq \frac{1}{n}$. We use this now to show successively that: $\sum_{i=1}^{i=n} h(\eta_i(g_K)) = \sum_{i=1}^{i=n} h(\eta_i(g) + \mu_i) \leq \sum_{i=1}^{i=n} h(\eta_i(g) + (\mu_i - 1)) \leq \sum_{i=1}^{i=n} h(\eta_i(g) + (\mu_i - 2)) \leq .... \leq \sum_{i=1}^{i=n} h(\eta_i(g))$. Since $\sum_{i=1}^{i=n} h(\eta_i) \leq v(g^*)$, we conclude that $v(g_K) \leq \sum_{i=1}^{i=n} h(\eta_i) \leq v(g^*)$. ■

It can be proved that for the model defined by (11), $v(g) \leq v(g^*)$ whenever $n \geq \frac{1}{\delta}$.

We now show that under some fairly weak assumptions on the payoffs, we will also have $v(g^*) \geq v(g)$ for any disconnected network g. Under these conditions, the star will then be efficient.

Proposition 12 *Let g_1^* and g_2^* be two disjoint stars with centers i and j. Whenever $(n_k - 1)[b(1, n_k) + b(2, 1) - b(1, n_k - 1)] \geq c$ for $k = 1, 2$ (sufficient but not necessary condition), $v(g_1^* \cup g_2^* + ij) \geq v(g_1^* \cup g_2^*)$. In particular, this cost range exists when $b(1, n_k) + b(2, 1) > b(1, n_k - 1)$.*

Proof: Under the assumptions in Proposition 11, the value of a star is not smaller than the value of any connected structure. Let g be a disconnected network. Then $v(g)$ is maximized when g is the union of star components. Let us show that under some weak conditions, social utility is increased by connecting two stars.

Indeed, let i and j be the centers of two stars of size n_1 and n_2, with $2 \leq n_1 \leq n$ and $2 \leq n_2 \leq n$. If we add a link between the centers, then the change in utility is $(n_1-1)[b(2,n_2)+(n_2-1)b(3,1)+b(1,n_1)-b(1,n_1-1)]+b(1,n_2)+(n_2-1)b(2,1)+(n_2-1)[b(2,n_1)+(n_1-1)b(3,1)+b(1,n_2)-b(1,n_2-1)]+b(1,n_1)+(n_1-1)b(2,1)-2c \geq (n_1-1)[b(2,n_2)+b(1,n_1)+b(2,1)-b(1,n_1-1)]+(n_2-1)[b(2,n_1)+b(1,n_2)+b(2,1)-b(1,n_2-1)]-2c$

This quantity is positive under a fairly weak condition: it is sufficient that decay with distance is not too great and utility decrease with respect to the neighbor's degree is not too great: $(n_k-1)[b(1,n_k)+b(2,1)-b(1,n_k-1)] \geq c$ for $k=1,2$.

If connecting two disjoint stars is efficiency improving, the efficient network cannot be disconnected and so under Conditions 1 and 2, the star is efficient. ■

The condition in the above proposition is sufficient but not necessary and could easily be improved. However, the star is not uniquely efficient. The complete graph, the empty one, and also the line or a disconnected graph with connected components of size two can all be efficient for some choices of b. The results we have are sufficient to show that the star is uniquely stable under conditions which are compatible with the (in some cases unique) pairwise stability of other structures than the star.

From Propositions 11 and 12, and adding condition that $b(1,1) > c$ (which ensures that the empty network is not efficient), we obtain the following.

Proposition 13 *Let the benefit function b satisfy Conditions 1 and 2, $(\widetilde{n}-1)[b(1,\widetilde{n})+b(2,1)-b(1,\widetilde{n}-1)] \geq c$ for all $2 \leq \widetilde{n} \leq n$ and let $b(1,1) > c$, then the star is uniquely efficient.*

Having established the efficiency of the star under some conditions, we can now compare with the pairwise stable structures characterized in the previous section. The conditions for efficiency of the star can be compatible with the stability conditions of the complete network or with that of the windmill network (Propositions 2(iv) and 3). This is most readily verified by checking the respective conditions for the function (10). This implies the existence of benefit functions and cost ranges verifying our general assumptions for which the star is efficient but not pairwise stable and the pairwise stable or even uniquely pairwise stable network is not efficient. We have already seen that for large n the windmill or complete network are not efficient in their stability region. The result about the efficiency of the star also show the existence of benefit functions for which the efficient network is strictly contained in a (or the, in the case of uniqueness) pairwise stable network. This implies that our model can give rise to overconnectivity in the strong sense defined by Buechel and Hellmann (2012), which could never occur in the original connections model.

6 Conclusion

Negative externalities in networks are an important and interesting issue which has many aspects that remain to be explored. In this paper, we analyzed a model of network formation that combined on one hand the presence of indirect benefits and on the other hand, a penalty resulting from the connectivity of direct and indirect neighbors. In order to show the consequences of modifying the Jackson and Wolinsky model by implementing negative externalities from the degree of indirect contacts, we defined the degree-distance-based utility close to the distance-based utility model by Bloch and Jackson (2007), but assumed a two-variable (instead of one-variable) benefit function which depends on distances to and degrees of direct or indirect neighbors. Indeed our obtained results show that new strategic considerations induced by the presence of a trade off between gains from increased spillovers and penalties from increased busyness of indirect contacts leads to different outcomes than in the original connections model. This applies to both parts of the analysis, stability and efficiency. Thus, this initial study confirms that it is indeed relevant to take into account the effects of negative externalities which do alter agents' strategic behavior and the social efficiency of various outcomes.

There are several issues in our agenda for the follow-up research on the connections model. Other possible extensions of the present work concern the types of negative externalities that we want our model to capture. As we discussed in the introduction the model we proposed here is a good fit for a situation that we could see as a "generalized" co-author model where knowledge spills over from more distant parts of the network. We can also think of situations where benefits could spill over from distant neighbors but be reduced by overall connectivity. One might see a link as a consumer good whose value is based on its rarity and which thus decreases the more common or widespread it is. Moreover, we note that as in many network formation games, our model exhibits a high multiplicity of equilibria, suggesting that the use of stronger stability concepts, e.g., strongly stable networks (those which are stable against changes in links by any coalition of individuals; see Jackson and van den Nouweland (2005)) could be helpful for equilibrium selection. We could also gear this general model more specifically towards some of the interesting social and economic problems on which the proposed framework can shed some light, notably the competition for information.

Bibliography

V. Bala and S. Goyal. A noncooperative model of network formation. *Econometrica*, 68(5):1181–1229, 2000a.

V. Bala and S. Goyal. A strategic analysis of network reliability. *Review of Economic Design*, 5:205–228, 2000b.

P. Billand, C. Bravard, and S. Sarangi. A note on local spillovers, convexity and the strategic substitutes property in networks. *Theory and Decision*, 2012a. Forthcoming.

P. Billand, C. Bravard, and S. Sarangi. Directed networks with spillovers. *Journal of Public Economic Theory*, 14:849–878, 2012b.

P. Billand, C. Bravard, and S. Sarangi. On the interaction between heterogeneity and decay in two-way flow networks. *Theory and Decision*, 73:525–538, 2012c.

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.

N. Carayol and P. Roux. 'Collective innovation' in a model of network formation with preferential meeting. In T. Lux, E. Samanidou, and S. Reitz, editors, *Nonlinear Dynamics and Heterogeneous Interacting Agents*, pages 139–153. Lecture Notes in Economics and Mathematical Systems, Vol. 550, Springer, 2005.

N. Carayol and P. Roux. Knowledge flows and the geography of networks: A strategic model of small world formation. *Journal of Economic Behavior and Organization*, 71: 414–427, 2009.

S. Currarini. Network design in games with spillovers. *Review of Economic Design*, 10(4): 305–326, 2007.

B. Dutta and M. O. Jackson. The stability and efficiency of directed communication networks. *Review of Economic Design*, 5:251–272, 2000.

B. Dutta and S. Mutuswami. Stable networks. *Journal of Economic Theory*, 76:322–344, 1997.

A. Galeotti, S. Goyal, and J. Kamphorst. Network formation with heterogeneous players. *Games and Economic Behavior*, 54:353–372, 2006.

S. Goyal and S. Joshi. Unequal connections. *International Journal of Game Theory*, 34: 319–349, 2006.

H. Haller. Network extension. *Mathematical Social Sciences*, 64:166–172, 2012.

H. Haller and S. Sarangi. Nash networks with heterogeneous links. *Mathematical Social Sciences*, 50:181–201, 2005.

T. Hellmann. On the existence and uniqueness of pairwise stable networks. *International Journal of Game Theory*, 2012. Forthcoming.

D. Hojman and A. Szeidl. Core and periphery in networks. *Journal of Economic Theory*, 139:295–309, 2008.

M. O. Jackson. A survey of network formation models: stability and efficiency. In G. Demange and M. Wooders, editors, *Group Formation in Economics: Networks, Clubs, and Coalitions*, pages 11–49. Cambridge University Press, Cambridge, UK, 2005.

M. O. Jackson. *Social and Economic Networks*. Princeton University Press, 2008.

M. O. Jackson and B. W. Rogers. The economics of small worlds. *Journal of the European Economic Association*, 3:617–627, 2005.

M. O. Jackson and A. van den Nouweland. Strongly stable networks. *Games and Economic Behavior*, 51:420–444, 2005.

M. O. Jackson and A. Watts. The evolution of social and economic networks. *Journal of Economic Theory*, 106(2):265–295, 2002.

M. O. Jackson and A. Wolinsky. A strategic model of social and economic networks. *Journal of Economic Theory*, 71:44–74, 1996.

C. Johnson and R. Gilles. Spatial social networks. *Review of Economic Design*, 5:273–299, 2000.

T. Morrill. Network formation under negative degree-based externalities. *International Journal of Game Theory*, 40:367–385, 2011.

A. Watts. A dynamic model of network formation. *Games and Economic Behavior*, 34: 331–341, 2001.

A. Watts. Non-myopic formation of circle networks. *Economic Letters*, 74:277–282, 2002.