The Role of Externalities in Social and Economic Networks

Inauguraldissertation zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaften (Dr. rer. pol.) an der Fakultät für Wirtschaftswissenschaften der Universität Bielefeld

vorgelegt von

Philipp Möhlmeier
Dipl.-Wirt. Math., Dipl.-Kfm.

September 2017

Erster Betreuer: Prof. Dr. Herbert Dawid
Zweite Betreuerin: Agnieszka Rusinowska, Ph. D.

Acknowledgment

First and foremost, I would like to thank Herbert Dawid and Agnieszka Rusinowska for agreeing to supervise this thesis and for the countless discussions about my research during the last years. I am deeply indebted to you for your input and strong support during all periods of time.

Herbert Dawid who already supervised my diploma theses in Mathematical Economics and Business Administration has encouraged me to apply for the international research training group Economic Behavior and Interaction Models (EBIM). He always gave me the freedom to work in the areas in which I was interested. Due to his impressively broad and deep knowledge in countless areas of economics, Herbert Dawid was always able to provide extremely valuable and immediate feedback.

Agnieszka Rusinowska and I met during my research stay at Université Paris I Panthéon-Sorbonne where she organized a seminar on Models of Influence and Network Theory (MINT). From the very first day, Agnieszka Rusinowska was open to discuss my research ideas and agreed to supervise me. Consequently, later on two joint projects together with her and Emily Tanimura arose out of this early connection from Paris. Agnieszka Rusinowska always impressed me by her enthusiasm for research, her drive to make progress and her optimism for publication.

Furthermore, although I decided to found the company Talents' Friends GmbH during my doctoral studies – which implied I had to finish this thesis in part time – both stood behind me and encouraged me to proceed on the way I had chosen.

Since none of this can be taken for granted, I would like to thank the two of you, Herbert and Agnieszka, once more for everything you did for me!

Besides my two supervisors, I would like to explicitly thank many other people who accompanied me on my way at Bielefeld University, Purdue University and Université Paris I Panthéon-Sorbonne. Let me just list their names in alphabetical order and omit all the dots connecting us: Charalambos D. Aliprantis, Patrick Beißner, Volker Böhm, Jean-Marc Bonnisseau, Gabriele Camera, Oliver Claas, Christoph Diehl, Nils Mattis Görnemann, Michel Grabisch, Claus-Jochen Haake, Tim Hellmann, Andreas Karpf, Sonja Recker, Frank Riedel, Alexandre Skoda, Emily Tanimura and Walter Trockel.

Furthermore, I acknowledge the academic environment and financial support by the German Research Foundation (DFG), the International Research Training Group Economic Behavior and Interaction Models (EBIM), the Bielefeld Graduate School of Economics and Management (BiGSEM) and the Center for Mathematical Economics (IMW).

As an entrepreneur, I would like to thank everyone who contributed to the success of Talents' Friends GmbH during the last years. In particular my co-founders Alexander Bongartz and Jan Philipp Platenius, our experienced and supportive shareholders Thomas Bentz, Rudolf Delius, Hartmut Ostrowski, Burkhard Schwenker and Ralf Struthoff, as well as our fabulous team.

Last, but not least, I should mention the people that are the closest to me: My parents Christa and Heinz Möhlmeier, my brother Benjamin together with his wife WeiWei and my fiancé Isabella. Without their support, understanding and constant trust, I would not have been able to live my dreams and to finally finish this dissertation project. They play the most central role in my life. There are no words to explain my gratitude to them.

Bielefeld, Philipp Möhlmeier
September 2017

Contents

Chapter 1

Introduction

This thesis contributes to the area of network theory in economics. It especially deals with the role of externalities, either positive or negative, in social and economic networks. The implementation of the network approach in economics is a relatively recent development that has strongly flourished over the last two decades. The books by Goyal (2007), Jackson (2008) and Bramoullé et al. (2016) provide an excellent overview of the theory of social and economic networks. Network theory is an interdisciplinary field and links various research areas together, like economics, sociology, computer science, physics and mathematics. Many of the basic ideas and concepts go back to the early roots of graph theory in mathematics.

Social and economic networks play a prominent role in many areas of our daily life. Consider for example a network of professional contacts. There is strong empirical evidence that the better a person is positioned in such a network, the more valuable information about open vacancies she will receive when she is looking for a job (see e.g. Granovetter (1974) and Cingano and Rosolia (2012)). Further applications are e.g. opinion formation and influence in social networks, buyer-seller networks or R&D networks in which companies form joint collaborations to reduce costs, benefit by knowledge spillovers and foster innovation. All these applications have also been studied from a theoretical perspective: See e.g. Calvó-Armengol (2004) and Cahuc and Fontaine (2009) for job networks; Grabisch and Rusinowska (2010) and Grabisch et al. (2017) for opinion formation and influence in social networks; Kranton and Minehart (2001) and Wang and Watts (2006) for buyer-seller networks; and Goyal and Moraga-González (2001) and Dawid and Hellmann (2014) for R&D networks.

A *network* is given by a set of nodes and a set of links which connect the nodes. The network explicitly indicates the relationships and distances (by the length of the paths) between the nodes. When new links are formed or existing ones are deleted, the situation may (substantially) change, not only for the nodes directly involved, but also for the other nodes in the network. These effects are called *externalities*. Consequently, some nodes may then favor to form/delete (different) links than before and the network structure could significantly change over time. Typical questions which arise and which are usually discussed in the relevant literature are the following: Are there positive and/or negative externalities and if so, how severe are they for whom under which conditions? Which networks are (pairwise) stable and which are (strongly) efficient? Which structures are equilibria that are likely to be observed in the long run?

For this thesis, the groundbreaking paper by Jackson and Wolinsky (1996) plays a central role. We consequently build up on it and we frequently relate our results to their outcomes. Jackson and Wolinsky (1996) introduce two fundamental models, the connections model and the co-author model and provide conclusions regarding the pairwise stability and strong efficiency of specific architectures in these settings.

In the *connections model*, nodes establish links to each other and receive a certain benefit for it. Establishing/Maintaining a link is associated with a cost and depending on the cost-benefit comparison, adding, maintaining or deleting a specific link might be individually rational. Benefits spill over to a node from all nodes it is directly and indirectly connected to. The spillovers from indirect connections are discounted and dependent on distance. If two nodes form a new link, this may only reduce the distance between other nodes. Hence, the connections model is a framework with purely positive externalities by link formation.

In contrast, the *co-author model* is a framework with purely negative externalities by link formation. In the co-author model, the nodes are interpreted as researchers who put effort in (joint) research projects. On the one hand, to be involved in a further co-authored project provides an additional option to pubish, but on the other hand, being increasingly connected also comes with a downside. The more projects a researcher is involved in, the less time he spends on every single project, since his time budget is fixed. Hence, if a researcher decides to join an additional project, link formation always induces negative externalities on his existing co-authors.

Jackson and Wolinsky (1996) als introduce the notion of pairwise stability. A network is said to be *pairwise stable*, if no single node wants to delete an existing link and no two nodes mutually would like to establish a link between each other. This stability concept has been widely applied in the literature afterwards (see e.g. in Johnson and Gilles (2000) and Goyal and Joshi (2003)). It has been critized, however, due to the restriction that only one link can be altered at the same time. Therefore, different (stronger) stability concepts were suggested, like the notion of *strong stability* where a network is stable against changes in links by any coalition of individuals (see e.g. Dutta and Mutuswami (1997) and Jackson and van den Nouweland (2005)).

Besides that, there exist many further refinements of stability. Buechel and Hellmann (2012) provide a comprehensive overview of the most common notions and elaborate important results regarding the connectedness of networks and the role of externalities in strategic network formation. First, the authors show that if a profile of utility functions satisfies positive externalities, then no pairwise Nash stable network is over-connected with respect to any monotonic welfare function. Second, Buechel and Hellmann (2012) prove that if a profile of utility functions satisfies negative externalities and concavity, then no network which is pairwise stable with transfers, is under-connected with respect to the utilitarian welfare function. Additional contributions on the role of externalities in social and economic networks have been provided, e.g. by Goyal and Joshi (2006), Bloch and Jackson (2007), Morrill (2011) and Hellmann (2013).

While many models in the aforementioned literature consider either positive or negative externalities, in this thesis we will focus in large parts on modifications that contain both types of externalities – positive and negative – at the same time. Then, given a network structure and the perspective of a specific node, the critical question usually is, which effect overweighs. It is important to understand better the implications of having both types of externalities simultaneously because models with purely positive or negative externalities appear to be somehow restrictive. As a concrete example, displaying either externailities, consider a stylized academic job market in which information about job opportunities and candidates is shared within a network of scientists (the nodes). In such a network, some scientists offer vacancies which they cannot fill internally, while others need to place team members, e.g. their job market candidates or untenured faculty. Establishing and maintaining a connection to a colleague (i.e., a link) is costly, but increases the probability of receiving valuable information. Information received from a neighboring node is passed on to all neighbors, with its value depreciating. For a given node, an additional link induces a positive externality (if it reduces the distance to other nodes) and/or a negative externality (if it better connects remote nodes, i.e., if it gives theses scientists a relative advantage). The analysis of such a situation would be strongly simplified by considering only one specific type of externality, but like in many other real-world applications it does not to appear very coherent. Depending on the structure of the network, the perspective of a specific node and the nodes involved in the link formation process, either positive or negative externalities may overweigh.

Throughout this thesis, we are going to introduce various (specific and generalized) models that capture both types of externalities. Our main goals are to describe the structural implications induced hereby and to understand better the role of externalities in social and economics networks. We will derive results about (asymptotic) pairwise stability and strong efficiency, draw conclusions depending on the underlying framework and relate our results to the existing literature on externalities. Following this introductory chapter, this thesis contains four additional chapters: Chapters 2 to 5 are all self coherent research papers using their own notation. Chapters 2 and 3 are joint work with Agnieszka Rusinowska and Emily Tanimura, chapter 4 is work on my own and chapter 5 is joint work with Claus-Jochen Haake and Sonja Recker. Chapter 2 is already published in the Journal of Public Economic Theory (JPET) and chapter 3 is forthcoming in Mathematical Social Sciences (MSS). In the following let me highlight more details regarding the specific contents of the chapters 2 to 5.

In chapter 2 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. Consider a situation in which people are involved in projects. They generate some kind of knowledge by themselves and receive some from others. If an agent is involved in many projects, he will have less time to generate output by himself. However, the more connections he has, the more knowledge he will receive from neighbors, neighbors of neighbors and so on. 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 changes certain results on stability and efficiency in comparison to the original connections model. In particular, we see the emergence of new stable structures, such as a star with links between peripheral nodes. Our analysis focuses mainly on structures with short diameters, but also considers cases with extreme levels of decay. We also identify structures, for example, certain disconnected networks that are efficient in our model, but which cannot 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.

In chapter 3 we deal with network formation frameworks, 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 situation, 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. In the second model we do not assume that there is competition for the access to information, but rather that the use of information is rivalrous. In this case, it is assumed that when people are closer to the sender than an agent, the harmful effect is greater than when others are at the same distance to the sender as that agent. In both models we analyze pairwise stability and examine if the stability of a structure is preserved when the number of agents becomes very large. This leads to a new concept that we call asymptotic pairwise stability. We show that there exists a tension between asymptotic pairwise stability and efficiency. The results allow us to compare and contrast the effects of two kinds of competition for information.

While in chapter 2 the connections model by Jackson and Wolinsky (1996) is modified to a degree-distance-based variation, in chapter 4 we present another modification of the connections model that is closely related and takes account of negative externalities by overall connectivity. The idea for this approach goes back to Jackson and Wolinsky (1996) who mention in their seminal paper that "... one might have a decreasing value for each connection (direct or indirect) as the total amount of connectedness increases." (p. 53.). Taking this as a starting point, we add a weighting factor depending on overall connectivity to the functional form of the original connections model. This weighting factor is independent of own links, but benefits received from direct and indirect connections are reduced by increasing overall connectivity of the other nodes in the network. In this context, we solve for pairwise stable and asymptotically pairwise stable networks and analyze strongly efficient networks. We compare the results and indicate the similarities and differences of the connections model with purely positive link externalities and the adjusted version with negative externalities by overall connectivity. What appears to be striking in the overall connectivity model is the role of the star network. It turns out to be pairwise stable, asymptotically pairwise stable and to be a very well performing structure in terms of strong efficiency. The reason for this is that it combines short distances between all nodes with a minimal number of links. Hence, all nodes receive many spillovers with low decay and relatively low punishment by overall connectivity.

Chapter 5 is more applied and we investigate a duopoly with horizontal product differentiation, in which firms strategically form costly links to customers. Such a link to a customer may be interpreted as the firm granting access to trade its product. Altering the network of links changes the structure of competition. This results in externalities and influences the equilibrium quantities and profits. We investigate in how far the degree of substitutability of the firms' products and the costs of link formation influence equilibrium profits and thus the incentives to form or delete links. We illustrate which networks are locally and Nash stable for which regions of costs/substitutability combinations. For networks with an arbitrary number of customers we analyze local stability regions for selected networks and determine their limits as the number of customers becomes large. We also relate local and Nash stability for selected networks with n customers. For networks with three customers we entirely characterize locally stable networks. In particular, existence is guaranteed for any degree of substitutability and any cost value.

Chapter 2

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

This chapter is based on a joint work with Agnieszka Rusinowska and Emily Tanimura, both from Université Paris I Panthéon-Sorbonne, Centre d'Economie de la Sorbonne. It is published in the Journal of Public Economic Theory (JPET), volume 18, pages 168–192, 2016.

2.1 Introduction

The connections model, introduced in the seminal paper of Jackson and Wolinsky (1996) is a setting in which only direct contacts are costly but discounted benefits spill over from indirect neighbors. A natural interpretation is that benefits result from the access to a resource conveyed by the network, such as information or knowledge provided by indirect contacts.

An appealing feature of networks is that they 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). In the connections model, network externalities are positive. An additional link formed by some pair of individuals (weakly) benefits all other agents by providing access to new indirect contacts or by reducing the distance that information has to travel. Such positive aspects of increased connectivity are certainly important. However, in many situations increased connectivity can also have negative side effects. Studying such cases is what motivates the analysis in this paper: we consider a model in which agents benefit from indirect contacts as in the original connections model but in which the connectivity of an agent may also exert a negative externality on his direct and indirect neighbors.

Contexts where this is the case abound. For example, learning of a job opening may be less useful if the information has been communicated to many others. When there is competition for some resource transmitted by the network, the benefits from indirect contacts are reduced when the latter have many connections. In our model, 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. However this might not fully account for the negative impact of all other individuals in the communication chain who receive the information. Hence, our model should be viewed only as a simplified or approximate description of the negative effects of connectivity when there is competition for information.

Another negative effect of high connectivity that our model perfectly captures arises because the busyness of agents reduces their availability or productivity. The connections model of Jackson and Wolinsky (1996) could also be interpreted as follows: nodes generate output by themselves but also forward output from others. Now interpret this as a situation in which individuals are involved in projects and generate knowledge by themselves but also receive knowledge from others. Then, a person involved in many projects will have less time to generate output. On the other hand, the more well connected he is, the more knowledge he will receive and forward to his neighborhood. Stated in a provocative way, "well connected people are often great talkers, but networking is time consuming and reduces one's productive time so that the main work is done by others". Nevertheless, the role of such well connected agents is very important: not that they contribute a lot by their own knowledge production, but they provide access to the output of many others.

By integrating the negative effect of the busyness of agents, at first sight, our model looks similar to the well-known co-author model (also introduced in Jackson and Wolinsky (1996)) where the time devoted to a single project decreases with the total number of projects a co-author is involved in. In our version, the nature of the negative externality is similar but information spills over from indirect connections. The co-author model only considers direct collaborations and thus conveys negative effects solely through the busyness of direct co-authors. Hence, our model combines aspects of the connections model and the co-author model. Externalities resulting from additional links can be both positive and negative. New links are useful for reaching indirect partners, but the latter will be more busy, less productive and thus less valuable per se although more efficient as intermediaries. Exploring the tradeoff of these effects and their impact on the stable and efficient network architectures that can arise is the object of our analysis.

A host of papers modify the original connections model in different directions, but most are not directly related to the issues we explore. Generalizing Jackson and Wolinsky (1996), Bloch and Jackson (2007) show that the results of the latter still hold when decay takes a more general functional form. Besides the aforementioned co-author model, the study of negative externalities that is most closely related to ours is the model of Morrill (2011) in which the benefits of a link is a decreasing function of the partner's degree. In fact, our setting generalizes Morrill (2011), where there are no benefits from indirect contacts and externalities are purely negative. Goyal and Joshi (2006) consider a framework where connectivity can generate positive or negative externalities, depending on whether the other agent is a direct, indirect or non-neighbor. They investigate two specific models. The first one captures negative externalities due to overall connectivity and thus addresses situations somewhat different from the ones that motivate our work. In their second model, the marginal benefit of forming a link to some agent is affected both by his and by one's own degree. The authors characterize stable structures, both in cases where the marginal benefits of a link increase with the potential partner's degree and in cases where they decrease as a function of it. Billand et al. (2012) prove the existence of a pairwise stable network in a local spillover game, when the marginal benefit of linking to an agent is decreasing with the degree of the latter and increasing in the own degree. The two aforementioned studies focus on the interplay between an agent's own degree and the degree of his neighbors when spillovers are local, whereas our analysis explores the tradeoff between degree-based negative externalities, reflecting busyness, and the possibility of receiving positive spillovers from distant partners.

We consider a network formation game whose payoffs have a functional form similar to that of the generalized connections model by Bloch and Jackson (2007) – the distance-based utility model, so as to facilitate comparison with the latter, but involve a penalty resulting from the degree of direct and indirect contacts. In this 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.

As in the original connections model, there is multiplicity of pairwise stable structures. We do not give a complete characterization but focus on some cases where outcomes can be compared and contrasted with those of the original connections model. In particular, we analyze stable structures with short diameters. The Jackson and Wolinsky model exhibits two such pairwise stable structures, the star and the complete network with a swift transition from one to the other. In our model, these structures can also be stable, but more interestingly we find new pairwise stable structures with short diameters which could not arise in the original connections model.

The nature of these new structures raises the following question: when direct and indirect contacts are evaluated based on two criteria, their capacity to be intermediaries and to what extent they are available (i.e., not too busy), how should these two roles be distributed? Will we see specialization so that some highly connected agents are valuable only as efficient intermediaries whereas other contacts are counted on for availability, or will we see a more equal distribution of roles where all agents are moderately busy and play a moderate role as intermediaries?

Indeed, we identify two types of architectures both of which ensure short communication paths between all agents but which are organized quite differently. One resembles a star but all peripheral agents also have a "local" neighborhood of direct contacts. In this case, the agents in the network occupy different roles. The center is specialized in the role of intermediary but is too busy to be of much value per se. The agents in the local neighborhood are not useful as intermediaries, but, being less busy, they are valuable in their own right. In the second structure, all agents have similar degrees and contribute equally to providing indirect contacts.

We derive stability conditions for the stable architectures in the original connections model (star/complete/empty) and for the aforementioned new structures. For a given payoff function, these conditions tell us whether one of these structures is stable. However, they provide little intuition for what determines the stability.

We pursue our analysis under the additional fairly natural assumption that decay is independent of degree. The concavity of the benefit function with respect to degree is then sufficient to ensure the existence of a range of link costs for which at least one of the new structures with short diameter is stable. More generally, the concavity/convexity of the benefit function with respect to degree is seen to play an important role in determining which structures with short diameters are stable. We also analyze stability and efficiency for extreme levels of decay. For high decay, our stable structures coincide with those in Morrill (2011) which is natural since our model approximates his when decay is large. For small levels of decay, stable structures will be minimally connected with some constraints on degrees.

As shown in Jackson and Wolinsky (1996) strongly efficient networks may not be stable, and conversely networks need not be efficient even when they are uniquely stable. Buechel and Hellmann (2012) show that inefficient outcomes can be related to the nature of 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 no stable network can be over-connected. Negative externalities tend to induce over-connected networks, and under some additional conditions, no stable network can be under-connected. In our model, the new structures with short diameters, while stable, would typically not be efficient when the network is large. The same is true for the complete graph. We show this without actually characterizing *the* efficient network. Finally, we show that under certain conditions the star will be uniquely efficient. The conditions required in our proof are quite restrictive but compatible with the stability of the star structures with peripheral links. Thus, our model can indeed generate over-connectedness as defined by Buechel and Hellmann (2012), which could never occur in the original connections model.

The rest of the paper is structured as follows. In Section 2.2 we recapitulate some preliminaries on networks, the Jackson and Wolinsky connections model and existing extensions. In Section 2.3 we present our model. Pairwise stability and efficiency are studied in Sections 2.4 and 2.5, respectively. We begin by providing some illustrating examples in networks of small size and then turn to the general analysis of stability and efficiency. In Section 2.6 we mention some possible extensions. Some proofs are presented in the Appendix.

2.2 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), and the model with degree-based utility functions by Morrill (2011).

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 $G := \{g \mid g \subseteq g^N\}$.

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.

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 $\{i_1i_2, i_2i_3, \ldots, i_{K-1}i_K\} \subseteq g$.

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

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

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

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 $g' \subseteq g^N$.

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 $u_j(g) > u_j(g + ij)$.

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

$$u_i^{JW}(g) = \sum_{j \neq i} \delta^{l_{ij}(g)} - c\eta_i(g) \tag{2.1}$$

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.

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 $\delta + \frac{(n-2)\delta^2}{2} < c$.

They also examine pairwise stability in the symmetric connection model. By virtue of Jackson and Wolinsky (1996) (Proposition 2), in the symmetric connections model:

(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

$$u_i^{co}(g) = \sum_{j:ij\in g} w_i(n_i, j, n_j) - c(n_i) \tag{2.2}$$

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.

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

$$u_i^{dist}(g) = \sum_{j\neq i} f(l_{ij}(g)) - c\eta_i(g) \tag{2.3}$$

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. The 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

$$u_i^{deg}(g) = \sum_{j:ij\in g} \phi(\eta_j(g)) - c\eta_i(g) \tag{2.4}$$

2.3 A degree-distance-based connections model

In Jackson and Wolinsky (1996), an additional link induces only positive externalities. We suggest a modification that also generates negative externalities due to increasing connectivity. Every agent benefits from his direct and indirect connections, but it is additionally assumed that the higher the degree of a direct or indirect partner, the less valuable is this connection. In order to remain close to the connections model we follow existing generalizations of Jackson and Wolinsky (1996), by considering the utility of agent i given by

$$\widetilde{u}_i(g) = \sum_{j \neq i} b(l_{ij}(g), \eta_j(g)) - c\eta_i(g) \tag{2.5}$$

where $b : \{1, \ldots, n-1\}^2 \to \mathbb{R}^+$ is the net benefit that an agent receives from the direct and indirect connections, and $c > 0$ is the cost for a direct connection. 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 $\eta_j \in \{0, 1, \ldots, n-1\}$. In particular, $\widetilde{u}_i(g^{\emptyset}) = 0$ for every $i \in N$.

In the original connections model the benefit term is expressed using 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 we aim to analyze negative externalities resulting from the connectivity of direct and indirect neighbors, we will assume 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.

Our framework also generalizes the degree-based model by Morrill (2011). We have

$$\phi(\eta_j(g)) = b(1, \eta_j(g)), \text{ for all } \eta_j(g) \in \{1, \ldots, n-1\} \tag{2.6}$$

The generalized model defined by (2.5) also extends the distance-based model considered in Bloch and Jackson (2007) and recapitulated in (2.3). In other words, we consider a two-variable (instead of one-variable) benefit function.

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

$$u_i(g) = \sum_{j \neq i} \frac{1}{1 + \eta_j(g)} \delta^{l_{ij}(g)} - c\eta_i(g) \tag{2.7}$$

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

An idea somewhat similar to the one expressed by our model is presented in Haller (2012) who studies a non-cooperative model of network formation. He considers two examples with negative network externalities in which the values of information are endogenously determined and depend on the network. This is in line with the idea that it is harder to access the information from an agent who has more direct neighbors.

2.4 Pairwise stability in the model

2.4.1 Stability of the star, the complete 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

$$\widetilde{f}(k) := b(1,k) - b(2,k-1) \text{ for } k \in \{2, \ldots, n-1\}, \text{ and } \widetilde{f}(1) := b(1,1) \tag{2.8}$$

Note that in the original connections model $\widetilde{f}(k) = \delta - \delta^2$ for each k. We have the following results on pairwise stability of the three prominent architectures.

Proposition 2.1. *In the degree-distance-based connections model defined by* (2.5)*:*

(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

$$\widetilde{f}(2) \leq c \leq \min\left(\widetilde{f}(1), b(1, n-1) + (n-2)b(2,1)\right) \tag{2.9}$$

This cost range is non-empty whenever $\widetilde{f}(2) \leq b(1, n-1) + (n-2)b(2,1)$.

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

$$c \leq \widetilde{f}(n-1) \tag{2.10}$$

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

$$c < \min_{1 \leq \eta_k \leq n-2} \widetilde{f}(\eta_k + 1) \tag{2.11}$$

(v) g^* and g^N are simultaneously PS if and only if $\widetilde{f}(2) \leq c \leq \widetilde{f}(n-1)$. This cost range is non-empty whenever $\widetilde{f}(2) \leq \widetilde{f}(n-1)$. In particular, if $\widetilde{f}(n-1) < c < \widetilde{f}(2)$, then 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 $\widetilde{f}(2) \le c$.

Hence, we get $\widetilde{f}(2) \le c \le \min\left(\widetilde{f}(1), b(1,n-1) + (n-2)b(2,1)\right)$.

(iii) Let $n \ge 3$. 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 $\widetilde{u}_i(g+ij) - \widetilde{u}_i(g) \ge b(1,\eta_j+1) - b(2,\eta_j) - c = \widetilde{f}(\eta_j+1) - c$ and $\widetilde{u}_j(g+ij) - \widetilde{u}_j(g) \ge \widetilde{f}(\eta_i+1) - c$. If $\eta_i\eta_j = 0$, then $\widetilde{u}_i(g+ij) - \widetilde{u}_i(g) \ge b(1,\eta_j+1) - c$ and $\widetilde{u}_j(g+ij) - \widetilde{u}_j(g) \ge b(1,\eta_i+1) - c$. Hence, if $c < \min_{1 \le \eta_k \le n-2} \widetilde{f}(\eta_k+1)$, then any two agents who are not directly connected benefit from forming a link.

(v) We have $\widetilde{f}(n-1) = b(1,n-1) - b(2,n-2) < b(1,n-1) + (n-2)b(2,1)$. Moreover, from the nonincreasingness of b in degree, $\widetilde{f}(n-1) < b(1,1) = \widetilde{f}(1)$. Hence, from (2.9) and (2.10), g^* and g^N are simultaneously PS if and only if $\widetilde{f}(2) \le c \le \widetilde{f}(n-1)$.

(vi) The general existence of pairwise stable disconnected structures is given in Proposition 2.8. In small networks we can also find other types of architectures. Consider for example g given in Figure 2.1 and the utility function given by (2.7).

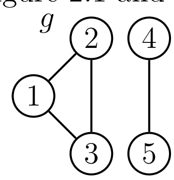


Figure 2.1: A PS network with two components in the degree-distance-based connections model

Network g is PS if $\frac{1}{4}\delta + \frac{2}{3}\delta^2 < c \le \frac{1}{3}\delta - \frac{1}{2}\delta^2$, 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:
$u_2(g) - u_2(g \setminus 23) = \frac{2}{3}\delta - 2c - \left(\frac{1}{3}\delta + \frac{1}{2}\delta^2 - c\right) = \frac{1}{3}\delta - \frac{1}{2}\delta^2 - c \ge 0$
$u_4(g) - u_4(g \setminus 45) = \frac{1}{2}\delta - c \ge 0$
$u_2(g+24) - u_2(g) = \frac{1}{3}\delta + \frac{1}{2}\delta^2 - c$ and $u_4(g+24) - u_4(g) = \frac{1}{4}\delta + \frac{2}{3}\delta^2 - c$

Note that if $\frac{1}{3}\delta - \frac{1}{2}\delta^2 - c \ge 0$, 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 $\frac{1}{3}\delta - \frac{1}{2}\delta^2 \ge c$ and $\frac{1}{4}\delta + \frac{2}{3}\delta^2 < c$. ■

Note that Proposition 2.1 confirms, in particular, the results on pairwise stability of g^N, g^* and $g^{\emptyset}$ in 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 (2.7). 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 for the same δ and n. For $\delta < \frac{1}{2}$, g^* cannot be PS in both frameworks at the same time, but for $\delta \geq \frac{1}{2}$ such an overlap of costs under which the star is PS is non-empty[2].

Proposition 2.1(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 (2.7), $\lim_{n\to\infty} \widetilde{f}(n-1) = 0$ so that the possible cost range for which the complete network and the star are simultaneously PS is very small in large networks.

Figure 2.2 (left) illustrates the pairwise stability regions for the three simple architectures for the model given by (2.7) with $n = 9$, $\delta \in (0, 1)$ and $c \in (0, 0.5]$[3]. 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.

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

Next, we will be interested in PS architectures of the degree-distance-based connections model, other than those analyzed in the previous section. An example of a PS structure that can occur in the white area in Figure 2.2 (left) is the windmill

[2] In the Jackson and Wolinsky model, g^* is PS if $\delta - \delta^2 < c < \delta$. In our model (2.7), g^* is PS if $\frac{\delta}{3} - \frac{\delta^2}{2} \leq c \leq \min\left(\frac{\delta}{2}, \frac{\delta}{n} + \frac{(n-2)\delta^2}{2}\right)$. As $\frac{\delta}{3} - \frac{\delta^2}{2} < \delta - \delta^2$, the costs range under which g^* is PS in both frameworks is $\delta - \delta^2 \leq c \leq \min\left(\frac{\delta}{2}, \frac{\delta}{n} + \frac{(n-2)\delta^2}{2}\right)$. Note that $\delta - \delta^2 \leq \frac{\delta}{2}$ if and only if $\delta \geq \frac{1}{2}$, and therefore g^* cannot be PS if $\delta < \frac{1}{2}$. For $\delta \geq \frac{1}{2}$, the overlap of costs is non-empty, as $\delta - \delta^2 \leq \min\left(\frac{\delta}{2}, \frac{\delta}{n} + \frac{(n-2)\delta^2}{2}\right)$.

[3] The calculations have been done in *Mathematica*. The details can be provided upon request.

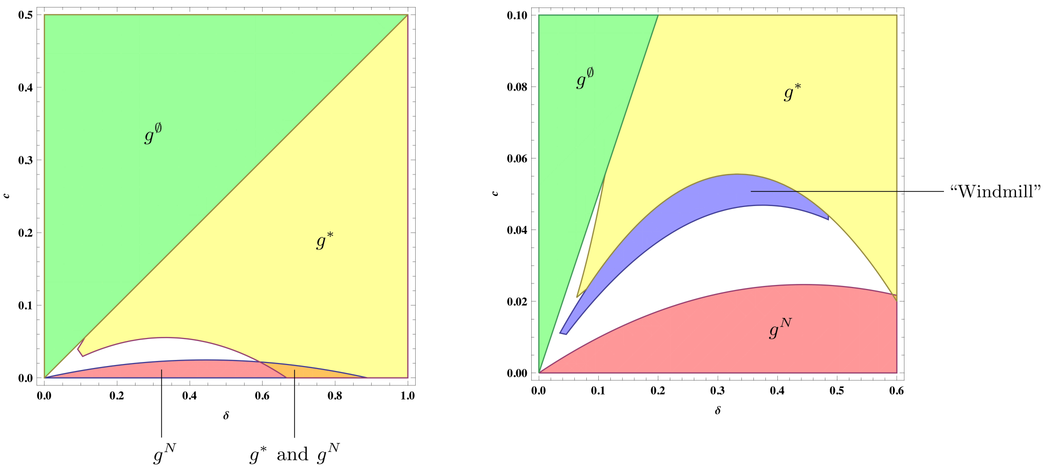


Figure 2.2: PS regions in the degree-distance-based connections model given by (2.7) ($n = 9$): only $g^\emptyset$ (green area), only g^* (yellow area), only g^N (red area), none of these three (white area); *Left* – g^* and g^N simultaneously (orange area); *Right* – "Windmill" (blue region)

structure shown in Figure 2.3. This is a specific example of what we will call a *core-periphery structure*.

Definition 2.1. *In a core periphery structure with periphery degree* η_m*, one node, the center, is linked to all other nodes, and every node other than the center has the same degree* η_m*.*

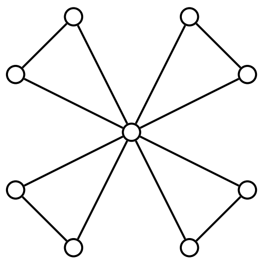


Figure 2.3: Windmill as an example of a PS network in the degree-distance-based connections model

To get a more precise feeling for the range of parameters in which such a windmill structure is stable, consider Figure 2.2 (right). Compared to Figure 2.2 (left), Figure 2.2 (right) is zoomed in and shows an additional blue area in which the windmill structure is PS. The green, yellow, red and white areas have the same meaning as before.

Figure 2.4 presents all PS networks for $n = 3$ with the corresponding parameters for the model (2.7). 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. Figure 2.4 confirms nicely Proposition 2.1(v). In particular, g^* and g^N are simultaneously PS only for $c = \frac{\delta}{3} - \frac{\delta^2}{2}$.

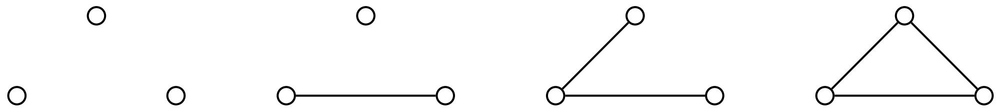


Figure 2.4: PS networks in the degree-distance-based connections model given by (2.7) (from left to right): (i) $\delta \leq 2c$, (ii) $0 < \delta < \frac{1}{3}$ and $\frac{\delta}{3} + \frac{\delta^2}{2} < c \leq \frac{\delta}{2}$, (iii) $(0 < \delta < \frac{1}{3}, \frac{\delta}{3} - \frac{\delta^2}{2} \leq c \leq \frac{\delta}{3} + \frac{\delta^2}{2})$ or $(\frac{1}{3} \leq \delta < 1, \frac{\delta}{3} - \frac{\delta^2}{2} \leq c \leq \frac{\delta}{2})$, (iv) $c \leq \frac{\delta}{3} - \frac{\delta^2}{2}$

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 2.5 and 2.6 show some examples of different PS structures for $n = 5$ and $n = 6$, respectively, for the model given by (2.7) 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 never be PS in the original connections model. Note that four of these networks contain two components. In Figure 2.5 these are the first network (on the left) that has been used in the proof of Proposition 2.1(vi), and the second network with one isolated player and four players, each having the degree equal to 2. In Figure 2.6 these are the first and the second network, both having two non-empty components.

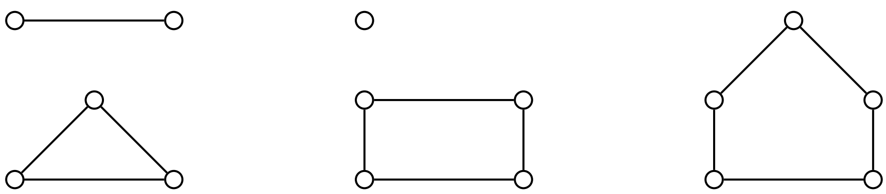


Figure 2.5: Some PS networks in the degree-distance-based connections model given by (2.7) for $n = 5$, $\delta = \frac{1}{15}$ and $c = \frac{107}{5400}$

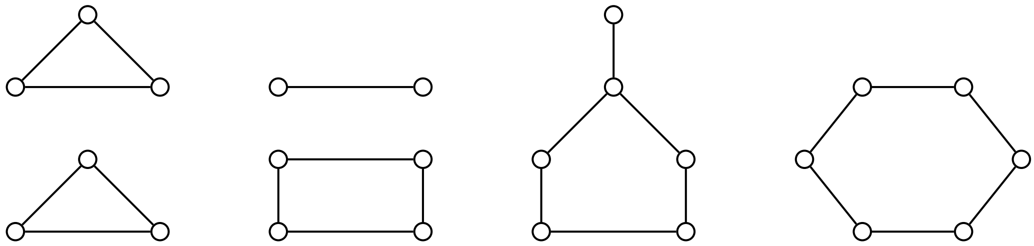


Figure 2.6: Some PS networks in the degree-distance-based connections model given by (2.7) for $n = 6$, $\delta = \frac{1}{15}$ and $c = \frac{107}{5400}$

2.4.3 Stability analysis for new architectures with short diameters

In the previous sections, we provided examples of networks that are PS in our framework for small values of n. These illustrate some of the general results about architectures that can be PS for an arbitrary size n 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 g is PS and $g^* \subset g \subset g^N$. This is never possible in the original Jackson and Wolinsky model. Figure 2.3 depicted an example of such a structure, the windmill 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 about the stability of stars with peripheral links.

Proposition 2.2. *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 $\widetilde{f}(n-1) + (n - 1 - \eta_m - (\eta_m - 1)(\eta_m - 2))\,(b(2, \eta_m) - b(3, \eta_m)) > c$. Then any core-periphery structure with periphery degree η_m is PS.*

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 for the following reason: Without a link to the center, an agent can reach at most $(\eta_m - 1)$ neighbors. Each neighbor has $(\eta_m - 2)$ links to nodes different from the center and the respective agent. Hence, without a link to the center, an agent can reach at most $(\eta_m - 1)(\eta_m - 2) + 1$ nodes at distance 2 (including one indirect link to the center). To identify the nodes who are at distance 3 in absence of a link to the center, one has to subtract from the overall number of nodes n the agent himself, his direct neighbors and the neighbors of degree 2. Consequently, by breaking a link to the center, at least $n - 1 - (\eta_m - 1) - ((\eta_m - 1)(\eta_m - 2) + 1) = n - 1 - \eta_m - (\eta_m - 1)(\eta_m - 2)$ nodes move to distance 3. Thus $\widetilde{u}_j(g) - \widetilde{u}_j(g - ij) \geq b(1, n-1) - b(2, n-2) + (n - 1 - \eta_m - (\eta_m - 1)(\eta_m - 2))\,(b(2, \eta_m) - b(3, \eta_m)) - c = \widetilde{f}(n-1) + (n - 1 - \eta_m - (\eta_m - 1)(\eta_m - 2))\,(b(2, \eta_m) - b(3, \eta_m)) - c$.

We show that for given b and c, we can find n such that

$$\widetilde{f}(n-1) + (n - 1 - \eta_m - (\eta_m - 1)(\eta_m - 2))\,(b(2, \eta_m) - b(3, \eta_m)) - c > 0 \quad (2.12)$$

Note that $b(2, \eta_m) > b(3, \eta_m)$. Moreover, $b(1, k) \geq 0$ for all k, and therefore from (2.8) 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 $n > \frac{c + b(2,1)}{b(2,\eta_m) - b(3,\eta_m)} + 1 + \eta_m + (\eta_m - 1)(\eta_m - 2)$ satisfies (2.12). ■

Note that in the original connections model, where $\widetilde{f}$ is constant, the first assumption of Proposition 2.2 which requires a decreasing $\widetilde{f}$ is never satisfied.

The stable structure with inhomogeneous degrees given in Proposition 2.2 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 provides indirect benefits. We also have examples of stable structures where the degree distribution is essentially homogeneous.

Proposition 2.3. *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 $\widetilde{f}(\eta_j) > c$. In particular, for the utility function given by* (2.7)*, this cost range is non-empty if and only if $k > \frac{\delta + \sqrt{\delta}}{1 - \delta}$ and $\widetilde{f}(\eta_j) > c$.*

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. Moreover, since $l_{ij} \leq 2$, the only path that is shortened is the one to the player one links to, and this path is reduced by one link. No agent wants to drop a link to agent j since $\widetilde{f}(\eta_j) > c$. For the utility function given by (2.7), 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 Proposition 2.3 clearly exist. We can give some (non exhaustive) examples.

Proposition 2.4. *Pairwise stable networks with equal degree: For n, M and l such that $M = \sqrt{n} \in N$, and l is a divisor of M, there exists a pairwise stable network g such that $l_{ij}(g) \leq 2$ for all i, j and all nodes have a degree equal to $k = (M-1) + (M-1)l$. It consists of M completely connected components or islands. Each node is linked to exactly l nodes on all other islands.*

Proof: Divide $n = M^2$ into M disjoint sets of size M, which we index by $m = 1, \ldots, M$. Divide each set of M agents into disjoint sets of size l $(S_i^m)_{i=1}^{M/l}$. Link agents from distinct islands if they have the same i. This ensures that each agent has exactly $M - 1 + (M-1)l$ neighbors. Moreover, $l_{ij}(g) \leq 2$ for all i, j. ■

We note that since $k = (M-1) + (M-1)l \geq 2(\sqrt{n} - 1)$, the structures defined above only exist for fairly large degrees. We also note that whenever $l > 1$, the resulting homogeneous network is such that the removal of a single link does not change the diameter.

Architectures with small diameters: further interpretation and analysis of the stability conditions

We established conditions for the stability of the main PS structures considered in Jackson and Wolinsky (1996) as well as other structures with small diameters: windmill structures and equal degree island models. However, these conditions did not provide much intuition for what really lies behind the stability of the various structures.

The conditions did show that the stability of some of the structures was only compatible with certain behaviors of the function $\widetilde{f}$. For example, Propositions 2.2, 2.3 require $\widetilde{f}$ to be decreasing. In this section we show that, under an additional assumption, namely that decay is independent of degree, the behavior of the function $\widetilde{f}$ (increasing or decreasing) can be related to the convexity/concavity of the benefit with respect to degree. Conditions related to concavity/convexity are easy to interpret and the results in this section should make it apparent that the conditions ensuring the stability of the new structures with short diameters analyzed in the previous section are not difficult to satisfy.

Throughout this section, we make the assumption that decay of the benefit is independent of the degree of the neighbor in the sense that $\frac{b(l+1,k)}{b(l,k)} =: \delta_l$ for every k, that is, the decay of the benefit may vary with the distance l but not with the degree k.

The following proposition links the behavior of the function $\widetilde{f}$ to the convexity/concavity of the benefit with respect to degree.

Proposition 2.5. *We have the following:*

- $\widetilde{f}(k)$ *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 $k \in 1, \ldots, n-1$. 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 2.2 and 2.3 are compatible with a benefit function that is concave in degree and with one that is convex in degree only if decay is large. Other results, such as the existence of simultaneous stability of the star and the complete network (see Proposition 2.1) cannot occur for a b concave in degree, but are compatible with a b convex in degree, for which the function $\widetilde{f}$ exhibits a wider range of behaviors.

The behavior of the function $\widetilde{f}$ is crucial for selecting the stable structures with diameter 2. We will now return to the stability conditions of the structures considered in the previous section. When $\widetilde{f}$ is increasing, we have a precise characterization of the stable networks with diameter 2 which have the property that the removal of a single link does not increase the diameter.

Proposition 2.6. *If $\widetilde{f}$ is increasing, then:*

- *there is a cost range for which the star is stable*
- *there is a cost range for which the complete graph is stable*
- *there is a cost range for which the star and the complete graph are stable*
- *a stable network g such that $l_{ij}(g-kl) \leq 2$ for all i,j and all kl must be a complete network.*

Proof: The function $\widetilde{f}$ is increasing on $[2, n-1]$. However, $\widetilde{f}(1) = b(1,1) > \widetilde{f}(2), \ldots, \widetilde{f}(1) > \widetilde{f}(n-1)$. The complete network is stable if $c < \widetilde{f}(n-1)$. The star is stable if $\widetilde{f}(2) < c < \widetilde{f}(1)$ and if $\widetilde{f}(n-1) + (n-2)b(2,1) > c$. The condition $\widetilde{f}(2) < c < \widetilde{f}(n-1) < \widetilde{f}(1)$ is sufficient for simultaneous stability of the star and the complete network.

For the last property, consider a stable network g such that $l_{ij}(g-kl) \leq 2$ for all i,j and all kl. The existence of a node of degree 1 contradicts the property $l_{ij}(g-kl) \leq 2$ for all i,j and all kl, because the removal of such a node disconnects the network. First note that if there are two nodes in g whose degrees are different from $n-1$, these nodes must be directly linked. Indeed, suppose that their degrees are η and η'. Neighbors of these agents do not lose any indirect benefits by breaking with them. Hence, stability requires that $\widetilde{f}(\eta) > c$ and $\widetilde{f}(\eta') > c$. But then $\widetilde{f}(\eta+1) > c$ and $\widetilde{f}(\eta'+1) > c$. Consequently, if the agents of degree η and η' are not connected, they would like to form a link and g would not be stable. Let $S = \{i \in N(g) | 1 < \eta_i < n-1\}$. All agents in S are linked by the above. The agents in $N \setminus S$ have degree $n-1$, they are linked to all agents. An agent in S is linked to all other agents in S and to agents in $N \setminus S$. But then an agent in S has degree $n-1$ contrary to hypothesis. Thus S is empty. Consequently, the stable networks

g such that $l_{ij}(g - kl) \leq 2$ for all i, j and all kl have only nodes with degree $n - 1$ and are thus complete graphs. ■

Corollary 2.1. *If $\widetilde{f}$ is increasing, then:*

- *no core periphery structure is stable*
- *no homogeneous island structure where each agent has at least two links to the other islands is stable.*

Having established these results, we will now return to the stability of the new structures with short diameters in the case where $\widetilde{f}$ is decreasing. Recall that by Proposition 2.2, the core-periphery structure in Proposition 2.2 with periphery degree η is PS if

- $\widetilde{f}(\eta + 1) < c \leq \widetilde{f}(\eta)$
- $\widetilde{f}(n-1) + (n - 1 - \eta - (\eta - 1)(\eta - 2))[b(2, \eta) - b(3, \eta)] > c$

These conditions can only be satisfied if $\widetilde{f}$ is decreasing. When this is the case, we note that it is fairly easy to find a cost range such that the conditions in Proposition 2.2 are satisfied when the periphery degree η is small compared to n. To satisfy the first condition we must take a cost $c > \widetilde{f}(\eta + 1) \geq \widetilde{f}(n - 1)$. Therefore the link to the center is only maintained if the indirect benefit term, bounded below by $(n - 1 - \eta - (\eta - 1)(\eta - 2))[b(2, \eta) - b(3, \eta)]$ is sufficiently large. When the network size n is large and η is small compared to n ($\eta \lll \sqrt{n}$), this term is large unless $[b(2, \eta) - b(3, \eta)]$ is very small, that is unless there is hardly any decay. In other words, when $\widetilde{f}$ is decreasing, it is easy to find cost ranges with PS core-periphery structures in which each peripheral agent has a “local” neighborhood that is relatively small compared to the whole network.

It is more difficult to obtain a stable core-periphery structure when the periphery degree η is large. Indeed, suppose for example that $n/2 + 1 < \eta < n - 1$. In this case the only benefit that is lost when a peripheral agent breaks with the center is the direct utility of the link to the center. When $n/2 + 1 < \eta < n - 1$, the peripheral agent can reach the whole network at distance 2 without going through the center. Peripheral agent i has at least $n/2 + 1$ contacts other than the center. Suppose that there were some agent k that he could not reach at distance 2 through these contacts. But agent k has at least $n/2 + 1$ links that are not to the center. At least one of these must lead to a neighbor of i. The benefit of conserving the link to the center is then reduced to $b(1, n-1) - b(2, n-2) = \widetilde{f}(n-1) < \widetilde{f}(\eta) < c$ and so the core-periphery structure is not stable. It is quite natural that a core periphery structure where agents have very large peripheral neighborhoods cannot be stable. Indeed, the peripheral agents can then reach a large fraction of the network without going through the center which then becomes superfluous and cannot be maintained.

While there may not be any stable core periphery structure with a large periphery degree η, if η is sufficiently large, then we can always find a cost range in which there is a stable homogeneous island structure (Proposition 2.4) with degree η when $\widetilde{f}$ is decreasing.

Proposition 2.7. *For degrees η such that the structure in Proposition 2.4 exists, this structure is PS if $\widetilde{f}(\eta+1) < c \leq \widetilde{f}(\eta)$. Such a c always exists when $\widetilde{f}$ is decreasing.*

Proof: We revisit the conditions of Proposition 2.4. When $\widetilde{f}$ is decreasing, we can always find a cost satisfying $\widetilde{f}(\eta+1) < c \leq \widetilde{f}(\eta)$. When $\widetilde{f}$ is decreasing, the latter also implies that $f(d) > c$ for any $d < \eta$. ■

The propositions in this section highlight the importance of whether the function $\widetilde{f}$ is decreasing or increasing. This can in turn be related to the concavity/convexity of the benefit function with respect to degree. However, even without concavity/convexity of the latter, it is easy to compute directly the function $\widetilde{f}$ whose behavior allows us to pin down the stable structures with short diameters. Structures with short diameters other than the star and the complete network, such as the core-periphery network and the homogeneous degree network, cannot be stable when $\widetilde{f}$ is increasing. On the other hand, when $\widetilde{f}$ is decreasing, we always have a cost range where the homogeneous island structure is stable. A cost range where the core-periphery structure is stable should also exist for many reasonable benefit functions at least when the number of peripheral links is small compared to the whole network. This shows that our model exhibits new pairwise stable structures in some cost ranges under reasonable and not too demanding assumptions. Indeed, the function $\widetilde{f}$ will be decreasing if the benefit function is concave with respect to the degree, or even when it is convex in this variable, provided that decay is high enough.

We also observe that the comparison of link costs and the marginal benefit of linking to an agent of degree η, i.e., $\widetilde{f}(\eta)$ is important for determining whether a core-periphery or a homogeneous structure emerges. If $\widetilde{f}(\eta)$ exceeds the cost even for large degrees η, then it is difficult to maintain a core-periphery structure. The peripheral agents want to form many links in the periphery, but by doing so, they are able to circumvent the center which becomes superfluous and cannot be maintained. On the other hand, the homogeneous organization where all agents have an equal and fairly high but not maximal degree will be stable. Such a structure is not possible if the cost exceeds $\widetilde{f}(\eta)$ for a small degree. In this case, however, we can have a stable core-periphery structure with only a small number of peripheral links.

2.4.4 Stability analysis for extreme values of the decay parameter

We complete our analysis of pairwise stability by considering the two extreme cases where decay is very large (i.e., δ is very small for the model given by (2.7)), or where decay is very small (i.e., δ is close to one for the model defined by (2.7)). We will show that in both of these cases, the degree-distance-based model can exhibit a very large number of PS architectures, which are only restricted by the fact that (most) nodes must have the same degree. When decay is large, the pairwise stable architectures include a number of disconnected structures with constraints on the degrees. These structures can be seen to coincide with those that are shown to be stable in Morrill (2011) (Proposition 2, p. 372). This is a natural since the benefits of a direct link in our model coincide with the benefits of a link in Morrill (2011) and that the indirect benefits can be neglected when decay is large.

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

Proposition 2.8. *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 (such a network is what Morrill (2011) calls a maximal nearly k-regular network).*

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 $\tilde{u}_j(g+ij) - \tilde{u}_j(g) < b(1, r+1) + \epsilon - c < 0$. 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: $\tilde{u}_j(g - ij) - \tilde{u}_j(g) < c + \epsilon - b(1, \eta_i) \leq c + \epsilon - b(1, r) < 0$. ■

The family of PS networks described in Proposition 2.8 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. There is also an abundance of disconnected structures that satisfy the condition stated in Proposition 2.8.

One example is that of a number of disconnected "islands" of identical size.

Corollary 2.2. *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 2.8.*

The class of PS networks described in Proposition 2.8 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 $(n-2)b(2,1) < \epsilon < b(1,r) \leq b(1,1)$. The possible values of the cost for which these structures exist shrink the larger n and r 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 (2.7). 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 2.9. *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? The indirect benefit of link formation 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 $\widetilde{u}_i(g-ij)-\widetilde{u}_i(g)<c-b(1,\eta_j)<0$. ■

De Jaegher and Kamphorst (2015) also study a model where agents' payoffs are based on their access to information in a setting with small decay. Information is defined as the sum of all decayed paths to indirect contacts, i.e., exactly the benefit term in the connections model. Payoffs differ from those of Jackson and Wolinsky (1996) because agents apply a possibly non-linear function to evaluate this aggregate quantity. Our model differs from Jackson and Wolinsky (1996) because we impose a "penalty" on the value received from each indirect contact *before* aggregating the value of all indirect contacts by addition. Because of these differences, our setting and that of De Jaegher and Kamphorst (2015) are not directly comparable. However, in both cases, the small decay assumption leads to stable structures that are minimally connected, reflecting the low benefit of forming a link to someone who is already in the same connected component. We obtain this result when the decay parameter approaches 1, whereas De Jaegher and Kamphorst (2015) assume small but not vanishing decay.

2.5 Efficiency in the model

Next, we analyze strong efficiency (SE) in the model, also sometimes referred to as efficiency in the paper. In the degree-distance-based model, there is a much wider range of possible SE architectures than in 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 for some parameters. To see this immediately, consider the 3-player example presented in Section 2.4.2. Figure 2.7 shows all strongly efficient networks in this model given by (2.7).

In this simple 3-player example we can already see an additional difference between

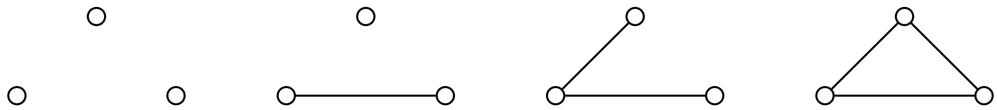


Figure 2.7: Unique SE networks in the degree-distance-based connections model given by (2.7) (from left to right): (i) ($0 < \delta < \frac{1}{3}$ and $2c > \delta$) or ($\frac{1}{3} \leq \delta < 1$ and $2c > \frac{1}{2}\delta^2 + \frac{5}{6}\delta$), (ii) $0 < \delta < \frac{1}{3}$ and $\delta^2 + \frac{2}{3}\delta < 2c < \delta$, (iii) ($0 < \delta < \frac{1}{3}$, $\frac{1}{3}\delta - \delta^2 < 2c < \delta^2 + \frac{2}{3}\delta$) or ($\frac{1}{3} \leq \delta < 1$, $2c < \frac{1}{2}\delta^2 + \frac{5}{6}\delta$), (iv) $2c < \frac{1}{3}\delta - \delta^2$

our model and the original connections model. In the latter, the cost ranges of PS and SE for the complete network coincide. In the degree-distance-based model defined by (2.7), g^N can be PS but not efficient: for $n = 3$ it is the case for $\frac{1}{3}\delta - \delta^2 < 2c < \frac{2}{3}\delta - \delta^2$. 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 neither be 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 (2.7) and $n = 4$ is for instance that the line is the unique SE network if $0 < \delta < \frac{4-\sqrt{13}}{6}$ and $\frac{1}{3}\delta - \frac{1}{3}\delta^2 - \delta^3 < 2c < \frac{1}{3}\delta + \frac{5}{3}\delta^2 + \delta^3$; for the calculations, see Appendix 2.A.1.

After presenting these examples, we turn to the theoretical analysis. Contrary to the original connections model, disconnected networks may be pairwise stable. Let us show that they can also be efficient.

Proposition 2.10. *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 2.8, 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 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. In this parameter range, an efficient network is such that each agent has exactly degree 1. This is achieved uniquely by a network consisting of disjoint 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 2.11 shows that when network size is large, the complete network is not strongly efficient when it is stable.

Proposition 2.11. *Whenever* $(n-1)\,(b(1, n-2) - b(1, n-1)) > b(1, n-2) - b(2, n-2)$*, the complete network is not strongly efficient for any cost* $c > 0$*. For the model defined by* (2.7)*,* 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 $\sum_{i=1}^{n} \left(\tilde{u}_i(g^N) - \tilde{u}_i(g^N - ij)\right)$. 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, $2(n-2)\,(b(1, n-2) - b(1, n-1))$. In total, the change in social utility is

$$\begin{aligned}
&\sum_{i=1}^{n} \left(\tilde{u}_i(g^N - ij) - \tilde{u}_i(g^N)\right) \\
&= 2(n-2)\,(b(1, n-2) - b(1, n-1)) - 2b(1, n-1) + 2b(2, n-2) \\
&= 2(n-1)\,(b(1, n-2) - b(1, n-1)) + 2\,(b(2, n-2) - b(1, n-2))\,.
\end{aligned}$$

This quantity is positive whenever $(n-1)\,(b(1, n-2) - b(1, n-1)) > b(1, n-2) - b(2, n-2)$ meaning that the complete network g^N is not strongly efficient. Except for very large decay (when $b(1, n-2) \gg b(2, n-2)$), the complete network is inefficient even when n is rather small. Solving $(n-1)\,(b(1, n-2) - b(1, n-1)) > b(1, n-2) - b(2, n-2)$ for the model defined by (2.7) leads equivalently to the condition $n > \frac{1}{\delta}$. ■

Proposition 2.12. *Let g be a network in which $l_{ij}(g) \leq 2$ for all i, j and let kl be a link such that $l_{ij}(g - kl) \leq 2$ for all i, j. Then:*

(i) *$g - kl$ has higher overall utility than g for any $c > 0$ when n is such that $2(n-2) > \frac{K}{\min(\alpha,\beta)}$ where $K = b(1, \eta_k) - b(2, \eta_k - 1) + b(1, \eta_l) - b(2, \eta_l - 1)$, $\alpha := \min_{1<l<n}[b(1, l-1) - b(1, l)] > 0$ and $\beta := \min_{1<l<n}[b(2, l-1) - b(2, l)] > 0$*

(ii) *In particular, for the model defined by (2.7), $g - kl$ is strictly more efficient than g for any $c > 0$ if $\delta > 1/n$.*

Corollary 2.3. *Suppose n satisfies the condition in Proposition 2.12. Then, neither the windmill (presented in Proposition 2.2), nor the complete graph, nor the multiple islands model (presented in Proposition 2.4 with $l > 1$) are efficient (even when they are PS).*

Indeed, it is readily verified that these structures contain a link whose removal conserves a maximal network diameter of 2.

Proof of Proposition 2.12: Consider two nodes k, l, such that the maximal distance in $g - kl$ is still 2. We must have $1 < \eta_k \leq n - 1$ and $1 < \eta_l \leq n - 1$. Let us show that overall utility increases when this link is removed.

First consider the change in utility for k and l: $\tilde{u}_k(g) - \tilde{u}_k(g - kl) + \tilde{u}_l(g) - \tilde{u}_l(g - kl) = b(1, \eta_k) - b(2, \eta_k - 1) - c + b(1, \eta_l) - b(2, \eta_l - 1) - c$. This quantity is positive when g is pairwise stable, since k and l have an incentive to maintain the link. It is bounded by K.

The presence of the link kl has a negative impact on all other agents. Agent k has degree η_k. Thus he has $\eta_k - 1$ direct neighbors, excluding l already accounted for above. Besides k himself and his η_k direct neighbors, the remaining $n - 1 - \eta_k$ agents are at distance 2. The utility loss for these agents is $(\eta_k - 1)(b(1, \eta_k - 1) - b(1, \eta_k)) + (n - 1 - \eta_k)(b(2, \eta_k - 1) - b(2, \eta_k))$ and similarly for agent l, replacing η_k by η_l.

We can now compare the overall utility of g and $g - kl$. It is

$$\begin{aligned}
&\sum_{i=1}^{n} (\tilde{u}_i(g) - \tilde{u}_i(g - kl)) \\
= &\; b(1, \eta_k) - b(2, \eta_k - 1) + b(1, \eta_l) - b(2, \eta_l - 1) - 2c \\
&\quad - (\eta_k - 1)(b(1, \eta_k - 1) - b(1, \eta_k)) \\
&\quad - (n - 1 - \eta_k)(b(2, \eta_k - 1) - b(2, \eta_k)) - (\eta_l - 1)(b(1, \eta_l - 1) \\
&\quad - b(1, \eta_l)) - (n - 1 - \eta_l)(b(2, \eta_l - 1) - b(2, \eta_l))
\end{aligned}$$

Define $\alpha := \min_{1<l<n}[b(1, l-1) - b(1, l)] > 0$ and $\beta := \min_{1<l<n}[b(2, l-1) - b(2, l)] > 0$.

$\sum_{i=1}^{n} (\tilde{u}_i(g) - \tilde{u}_i(g - kl)) \leq K - 2(n-2)\min(\alpha, \beta) - 2c < K - 2(n-2)\min(\alpha, \beta)$, which is negative provided $2(n-2) > \frac{K}{\min(\alpha,\beta)}$.

In particular, for the model defined by (2.7), $\sum_{i=1}^{n} (u_i(g) - u_i(g - kl))$ is negative if $F(\eta_k) + F(\eta_l) < 0$ with $F(\eta_k) = \frac{\delta}{\eta_k+1} - \frac{\delta^2}{\eta_k} - (\eta_k - 1)\left[\frac{\delta}{\eta_k} - \frac{\delta}{\eta_k+1}\right] - (n - 1 - \eta_k)\left[\frac{\delta^2}{\eta_k} - \frac{\delta^2}{\eta_k+1}\right]$

$F(\eta_k) < 0 \iff \eta_k - \delta(\eta_k + 1) - (\eta_k - 1) - (n - 1 - \eta_k)\delta < 0 \iff 1 < n\delta$. Similarly, $F(\eta_l) < 0$ if and only if $n > \frac{1}{\delta}$. Thus, whenever $n > \frac{1}{\delta}$ the network g is not efficient. ■

This result is easy to understand. In a network in which all agents are at distance at most 2 when kl is removed, the link kl benefits only agents k and l themselves and exerts a negative externality on a large number of agents at distance 2. Provided n is large, this outweighs the positive effects of the link.

Next, we focus on the conditions for efficiency of the star. Let us define the function

$$h(m) = mb(1, m) + (n - 1 - m)b(2, m) \tag{2.13}$$

This function represents an upper bound of the social utility that an agent with degree m provides to others. This upper bound is attained for example for a star, but generally it is not attained. For the model defined by (2.7), we have $h(m) = m\frac{\delta}{1+m} + (n - 1 - m)\frac{\delta^2}{1+m}$.

We assume the following conditions:

Condition 2.1. *Function h defined by* (2.13) *is decreasing.*

Condition 2.2. *For any $k > 2$, let $a+d = k$ and $a^{'}+d^{'} = k$. Then if $|a-d| \geq |a^{'}-d^{'}|$ then $h(a) + h(d) \geq h(a^{'}) + h(d^{'})$.*

Condition 2.2 means that when keeping fixed the sum of degrees of two agents, the total maximal social utility provided by the two agents, as measured by the function h is greater if the agents have dissimilar degrees. Note that Conditions 2.1 and 2.2 are satisfied for the model defined by (2.7), when $n > \frac{1}{\delta}$; for the proof, see Appendix 2.A.2.

Condition 2.1 is not demanding and holds easily if the network size n is large. Condition 2.2 seems to have a convexity flavor. One agent with high degree and one agent with low degree provide more social utility (or at least the upper bound given by h is greater) than two agents with intermediate degrees. In fact the proposition below shows under some conditions that Condition 2.2 cannot hold if b is concave with respect to degree. However, convexity is not sufficient to guarantee Condition 2.2 because a convex function whose behavior is very close to that of a linear function would not satisfy it.

Proposition 2.13. *Suppose that for any degree η, $b(2, \eta) = \delta b(1, \eta)$ for some $\delta \leq 1$ (decay independent of degree) and that $b(1, \eta)$ is differentiable. Then Condition 2.2 never holds if $b(1, \eta)$ is concave. Moreover Condition 2.2 does not hold for all convex $b(1, \eta)$.*

This proposition is proved in Appendix 2.A.3.

Proposition 2.14. *Let g be a connected structure. Whenever h given by* (2.13) *satisfies Conditions 2.1 and 2.2, we have $v(g) \leq v(g^*)$.*

Proof: 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 2.1 (presented in Appendix 2.A.4), the degrees of the nodes in S and the distances between all pairs of nodes. By Lemma 2.1, all nodes in J_1 have degree one. There are m nodes in S whose degrees are $(2 + \alpha_i)_{i=1}^{m}$ (without loss of generality we let the m nodes in S be $1, 2, \ldots, m$). The remaining $n - j_1 - m$ nodes have degree 2. Therefore by Lemma 2.2 (presented in Appendix 2.A.5), we have
$v(g) \leq \sum_{i=1}^{n} h(\eta_i(g)) = j_1 h(1) + \sum_{i=1}^{m} h(2 + \alpha_i) + (n - j_1 - m)h(2)$.
We apply Condition 2.2 to obtain $h(\alpha_i + \alpha_j + 2) + h(2) \geq h(\alpha_i + 2) + h(\alpha_j + 2)$. Thus

we have $\sum_{i=1}^{m} h(\alpha_i+2)+(n-m-j_1)h(2) = \sum_{i=1}^{m-2} h(\alpha_i+2)+(n-m-j_1)h(2)+h(\alpha_{m-1}+2)+h(\alpha_m+2) \leq \sum_{i=1}^{m-2} h(\alpha_i+2)+(n-m-j_1)h(2)+h(\alpha_{m-1}+\alpha_m+2)+h(2) = \sum_{i=1}^{m-2} h(\alpha_i+2)+h(\alpha_{m-1}+\alpha_m+2)+(n-m-j_1+1)h(2) \leq \ldots. \leq h(\sum_{i=1}^{m}\alpha_i+2)+(n-j_1-1)h(2)$.

Then, applying again repeatedly Condition 2.2, $v(g) \leq j_1h(1)+h(\sum_{i=1}^{m}\alpha_i+2)+(n-j_1-1)h(2) = j_1h(1)+h(j_1)+(n-j_1-1)h(2) = j_1h(1)+(n-j_1-2)h(2)+h(2)+h(j_1) \leq j_1h(1)+(n-j_1-2)h(2)+h(1)+h(j_1+1) = (j_1+1)h(1)+h(j_1+1)+(n-j_1-2)h(2) \leq \ldots \leq (n-1)h(1)+h(n-1) = v(g^*)$.

As we have shown, any minimally connected structure g has a degree sequence $(\eta_i)_{i=1}^n$ such that $v(g) \leq \sum_{i=1}^{n} h(\eta_i) \leq v(g^*)$. 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 2.2 we have $v(g_K) \leq \sum_{i=1}^{n} h(\eta_i(g_K))$. We will show that $\sum_{i=1}^{n} h(\eta_i(g_K)) \leq \sum_{i=1}^{n} h(\eta_i(g))$. Let us consider h being decreasing. For the model given by (2.7), 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:

$\sum_{i=1}^{n} h(\eta_i(g_K)) = \sum_{i=1}^{n} h(\eta_i(g)+\mu_i) \leq \sum_{i=1}^{n} h(\eta_i(g)+(\mu_i-1)) \leq \sum_{i=1}^{n} h(\eta_i(g)+(\mu_i-2)) \leq \ldots. \leq \sum_{i=1}^{n} h(\eta_i(g))$. Since $\sum_{i=1}^{n} h(\eta_i) \leq v(g^*)$, we conclude that $v(g_K) \leq \sum_{i=1}^{n} h(\eta_i) \leq v(g^*)$. ■

Consequently, for the model defined by (2.7), $v(g) \leq v(g^*)$ whenever $n > \frac{1}{\delta}$. We now show that under some fairly weak assumptions on the payoffs, we also have $v(g^*) \geq v(g)$ for any disconnected network g. Under these conditions, the star will then be efficient.

Proposition 2.15. *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), where n_k is the cardinality of g_k^*, $v(g_1^* \cup g_2^* + ij) \geq v(g_1^* \cup g_2^*)$. 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 2.14, 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 2.1 and 2.2, the star is efficient. ■

The condition in the above proposition is sufficient but not necessary and can 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 2.14 and 2.15, and adding condition that $b(1,1) > c$ (which ensures that the empty network is not efficient), we obtain the following.

Proposition 2.16. *Let the benefit function b satisfy Conditions 2.1 and 2.2, $(\tilde{n}-1)[b(1,\tilde{n})+b(2,1)-b(1,\tilde{n}-1)] \geq c$ for all $2 \leq \tilde{n} \leq n$ and let $b(1,1) > c$. Then the star is efficient, and is uniquely efficient whenever a strict inequality holds in Condition 2.2.*

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 2.1(iv) and 2.2). This is most readily verified by checking the respective conditions for the function (2.7). Indeed, the assumptions under which we proved the star to be efficient exclude concavity of the benefit function with respect to degree, but the new structures with short diameters could be stable both when the benefit function is concave and when it is convex in degree. 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 shows that the efficient network can be 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.

2.6 Conclusion

In this paper, we analyzed network formation in the presence of negative externalities in a model that combined the presence of indirect benefits and a penalty resulting from the connectivity of direct and indirect neighbors. Our analysis focused mainly on the case of structures with short diameters but also considered cases with extreme levels of decay. It would be interesting but more challenging to extend it beyond these cases. While remaining in the framework with global positive spillovers, we could also consider somewhat different models that capture other types of negative externalities. 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. We could also gear the model more specifically towards the competition for information, by letting payoffs depend on the number of informed agents in a communication chain. Finally, we note that in the model we considered, as well as in the potential extensions, we are likely to have 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.

2.A Appendix

2.A.1 Proof for efficiency of the line for $n = 4$

The line is the unique SE network for $n = 4$, model (2.7) and some δ, c.

Proof: Let $n = 4$. Let $U^{(k)} = \sum_{i=1}^{4} u_i(g^{(k)})$, where $k \in \mathbb{N}$ and $1 \leq k \leq 8$. Similarly, the sum of the players' utilities for the line, the empty graph and the complete graph is denoted by U^L, $U^{\emptyset}$ and U^N, respectively. We will determine the parameters δ and c under which the line g^L is the unique SE network. We have: $U^L = \delta^3 + \frac{5}{3}\delta^2 + \frac{7}{3}\delta - 6c$
$U^{\emptyset} = 0$, $U^N = 3\delta - 12c$, for the graph with one link: $U^{(1)} = \delta - 2c$
for the graphs with two links: $U^{(2)} = 2\delta - 4c$, $U^{(3)} = \delta^2 + \frac{5}{3}\delta - 4c$
for the graphs with 3 links (different from g^L): $U^{(4)} = 3\delta^2 + \frac{9}{4}\delta - 6c$ (star), $U^{(5)} = 2\delta - 6c$
for the graphs with 4 links: $U^{(6)} = \frac{4}{3}\delta^2 + \frac{8}{3}\delta - 8c$, $U^{(7)} = \frac{5}{3}\delta^2 + \frac{31}{12}\delta - 8c$
for the graph with 5 links: $U^{(8)} = \frac{2}{3}\delta^2 + \frac{17}{6}\delta - 10c$.
$U^L > U^{(4)}$ iff $\delta^2 - \frac{4}{3}\delta + \frac{1}{12} > 0$ iff $0 < \delta < \frac{4-\sqrt{13}}{6}$. We solve: $U^L > U^{(2)}$ and $U^L > U^{(6)}$ and $0 < \delta < \frac{4-\sqrt{13}}{6}$, which gives $0 < \delta < \frac{4-\sqrt{13}}{6}$ and $\frac{1}{3}\delta - \frac{1}{3}\delta^2 - \delta^3 < 2c < \frac{1}{3}\delta + \frac{5}{3}\delta^2 + \delta^3$. For such δ and c, we have, in particular, $\delta > 2c$. Hence, $U^L > U^{(1)}$, $U^L > U^{\emptyset}$, and also $U^L > U^{(3)}$, $U^L > U^{(5)}$, $U^L > U^{(8)}$, $U^L > U^{(7)}$, $U^L > U^N$. ■

2.A.2 Proof regarding Conditions 2.1 and 2.2

Conditions 2.1 and 2.2 are satisfied for the model (2.7), when $n > \frac{1}{\delta}$.

Proof: Consider $h(m) = m\frac{\delta}{1+m} + (n-1-m)\frac{\delta^2}{1+m}$. We have $h'(m) = \frac{\delta(1-\delta n)}{(1+m)^2} < 0$ for $n > \frac{1}{\delta}$. Consider a, d such that $a + d = k$, and therefore $d = k - a$. For the model defined by (2.7), we have $h(a) + h(d) = h(a) + h(k-a) = \frac{\delta a}{1+a} + (n-1-a)\frac{\delta^2}{1+a} + (k-a)\frac{\delta}{1+k-a} + (n-1-(k-a))\frac{\delta^2}{1+k-a} = \frac{(1+a)(\delta-\delta^2)}{1+a} + \frac{(1+k-a)(\delta-\delta^2)}{1+k-a} + \delta\left[\frac{n\delta-1}{1+a} + \frac{n\delta-1}{1+k-a}\right]$. Thus we can write $h(a) + h(d) = c(\delta) + \delta G(a)$, where $G(a) := \frac{n\delta-1}{1+a} + \frac{n\delta-1}{1+k-a}$. The derivative of this function is $G'(a) = \frac{1-n\delta}{(1+a)^2} + \frac{n\delta-1}{(1+k-a)^2}$ which is zero at $a = k/2$. Moreover, we can show that this zero corresponds to a minimum provided that $n\delta > 1$. Indeed, when this is the case, $G(a) \geq 0$ and $G(0) = (n\delta - 1)\left(1 + \frac{1}{1+k}\right) \geq G(k/2) = (n\delta - 1)\left(\frac{2}{1+k/2}\right)$ for every $k \geq 2$. We also note that by symmetry, $G(a) = G(k-a)$. This implies that $G(a)$ is decreasing on $[0, k/2]$ and increasing on $[k/2, k]$. Since

$|a-d| = |a-(k-a)| = |2a-k|$, if $|a-d| > |a^{'}-d^{'}|$, then $|a-k/2| > |a^{'}-k/2|$, which implies $G(a) > G(a^{'})$, which implies $h(a)+h(d) > h(a^{'})+h(d^{'})$. ■

2.A.3 Proof of Proposition 2.13

Proof: Condition 2.2 holds if and only if $h(a)+h(d) = h(a)+h(k-a) =: H(a)$ is decreasing on $[0, k/2]$. We have $b(2,\eta) = \delta b(1,\eta)$ for some $\delta \leq 1$. We compute $H^{'}(a)$ to find $H^{'}(a) = (1-\delta)[b(1,a)-b(1,k-a)] + \delta(n-1)[b^{'}(1,a)-b^{'}(1,k-a)] + (1-\delta)[ab^{'}(1,a)-(k-a)b^{'}(1,k-a)]$. Now $a \leq k-a$. Consequently if $b(1,\eta)$ is concave in degree, $b'(1,k-a) < b'(1,a) < 0$ and $H^{'}(a) > 0$. Thus Condition 2.2 is not compatible with concavity of b with respect to degree. We can also see that in the limit case between concavity and convexity where $b(1,\eta)$ is linear in degree, we would have $H^{'}(a) \geq (1-\delta)[b(1,a)-b(1,k-a)] > 0$. Thus convexity is not sufficient to guarantee Condition 2.2. ■

2.A.4 Proof of Lemma 2.1

Lemma 2.1. *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. Any minimally connected structure of size $n+1$ can be obtained by adding one node $n+1$ and one link between $n+1$ and some $j < n+1$ in a minimally connected network of size n. Suppose that g_n $(n \geq 4)$ verifies the induction hypothesis. If $S(g_n) = \emptyset$, g_n is a line. If we link $n+1$ to a node of degree 1, $j_1(g_{n+1}) = j_1(g_n) = 2$ and $S(g_{n+1})$ is still empty. If we add a link between $n+1$ and a node of degree 2, there will be 3 nodes with degree 1 and 1 node of degree 3. Thus $j_1(g_{n+1}) = 3$, and $\sum_{i\in S(g_{n+1})} \alpha_i(g_{n+1}) = 1 = j_1(g_{n+1})-2$. If g_n verifies the induction hypothesis and $S(g_n) \neq \emptyset$, there are several possibilities. 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 so $j_1(g_{n+1}) = j_1(g_n)$. The number of nodes in S does not change and $\eta_j(g_{n+1}) = 2$. Thus $\sum_{i\in S(g_{n+1})} \alpha_i(g_{n+1}) = \sum_{i\in S(g_n)} \alpha_i(g_n) = j_1(g_n) - 2 = j_1(g_{n+1}) - 2$. If $\eta_j(g_n) = 2$ then $\eta_j(g_{n+1}) = 3$ and $card(S(g_{n+1})) = card(S(g_n)) + 1$ and $j_1(g_{n+1}) = j_1(g_n) + 1$. Thus $\sum_{i\in S(g_{n+1})} \alpha_i(g_{n+1}) = \sum_{i\in S(g_n)} \alpha_i(g_n) + \alpha_j(g_{n+1}) =$

$j_1(g_n) - 2 + 1 = j_1(g_{n+1}) - 2$. Finally, if $n + 1$ links to j such that $\eta_j(g_n) > 2$, then $j_1(g_{n+1}) = j_1(g_n) + 1$, $card(S(g_n)) = card(S(g_{n+1}))$ and $\eta_j(g_{n+1}) = \eta_j(g_n) + 1$. Thus $\sum_{i \in S(g_{n+1})} \alpha_i(g_{n+1}) = \sum_{i \in S(g_n)} \alpha_i(g_{n+1}) = \sum_{i \in S(g_n)} \alpha_i(g_n) + 1 = j_1(g_n) - 2 + 1 = j_1(g_{n+1}) - 2$. This concludes the proof of the induction step. The induction hypothesis is verified when $n = 4$. There are two minimally connected structures: a line and a star. In a line, there are two elements of degree 1, thus $j_1 = 2$ and $S = \emptyset$. In a star with $n = 4$, $j_1 = 3$, S consists of the center with degree 3 and indeed $\sum_{i \in S(g_n)} \alpha_i(g_n) = 3 - 2 = j_1(g_n) - 2$. ■

2.A.5 Proof of Lemma 2.2

Lemma 2.2. *Let g be a network with degree sequence $(\eta_i(g))_{i=1}^n$. Then the value of g is $v(g) \leq \sum_{i=1}^n h(\eta_i(g))$.*

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 $b(2, \eta_i(g))$. ■

Chapter 3

Competition for the access to and use of information in networks

This chapter is based on a joint work with Agnieszka Rusinowska and Emily Tanimura, both from Université Paris I Panthéon-Sorbonne, Centre d'Economie de la Sorbonne. It is forthcoming in Mathematical Social Sciences (MSS).

3.1 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); see also Boorman (1975) for transmission of job information via strong and weak contacts. 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 certain 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 several "standard" architectures, in particular, the star, circle, complete and empty networks, and a structure of separate pairs. We also examine k-regular networks in the second model which is more tractable with respect to this analysis. Depending on model parameters, a variety of structures can be pairwise stable. We introduce a new stability concept that we call "asymptotic pairwise stability", i.e., we are interested in the network structures that remain PS when the number of agents becomes very large. This approach, which consists in studying the asymptotic properties of a sequence of graphs of increasing size is standard in the analysis of random graphs (see, e.g., Erdős and Rényi (1960) and Bollobas (2001)) where it is motivated by the need to use probabilistic limit theorems. Our framework is the same but in the special case where the sequence of graphs is deterministic and the asymptotic properties we seek to study are related to the payoff parameters of our model. From an empirical point of view, focusing on network structures that are stable when the number of agents is large but not necessarily otherwise seems relatively reasonable in our model which aims to describe individuals' access to useful information. To illustrate this point through a numerical example, empirical studies such as e.g., Killworth and Bernhard (1978) estimate the number of contacts where a contact is defined as somebody from whom one can ask for a favor to 200 on average. In an empirical study about finding a job, Granovetter (1974) found that the information that allowed an individual to find a job typically originated from indirect contacts up to three steps away. Even if we assume that there is much overlap, so that each of one's 200 contacts does not provide 200 distinct contacts, conservatively estimating the number of non-overlapping contacts to 100 places the number of agents relevant to the transmission of information at $100^3 = 1\,000\,000$. More generally, the notion of asymptotic pairwise stability could be relevant for studying various large networks such as peer to peer file sharing networks, Facebook, etc. In our setting, asymptotic pairwise stability allows us to obtain sharper predictions. Furthermore, for both models we show a tension between the asymptotic pairwise stability and efficiency, in the sense that some network architectures are not efficient when being asymptotically pairwise stable.

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 overspills 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. Calvó-Armengol (2004) considers job contacts networks and rivalry in the access through network links, but investigates a different framework of strategic network formation and studies the Nash equilibria of a non-cooperative game of network formation. 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) or in the communication networks where the link-strength is endogenously chosen by the agents (e.g., Bloch and Dutta, 2009).

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 are recalled in Section 3.2. Section 3.3 concerns competition for the access to information. In Section 3.4 we study the competition for the use of information. Section 3.5 presents concluding remarks and a comparison of the two kinds of competition for information. Longer proofs of the results are presented in the Appendix.

3.2 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). For simplicity and w.l.g. we assume that n is even. 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 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 $ij \in g$. We denote by $N_i(g)$ the neighborhood of i in g, i.e., $N_i(g) = \{j \in N \mid ij \in g\}$.

The *degree* $d_i(g)$ of agent i counts the number of links i has in g, i.e., $d_i(g) = |N_i(g)| = |\{j \in N \mid ij \in g\}|$.

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.

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 $\{i_1 i_2, i_2 i_3, \dots, i_{K-1} i_K\} \subseteq g$. We denote by $p(i_1 i_K)$ a path from i_1 to i_K and by $P(i_1 i_K)$ the set of all paths from i_1 to i_K. We write $j \in p(i_1 i_K)$ if path $p(i_1 i_K)$ passes through j.

By $l_{ij}(g)$ we denote the geodesic distance between i and j, i.e., the number of links

[1] We do not allow for loops in this setting.

in the shortest path between i and j in g. If there is no path connecting i and j in g, then we set $l_{ij}(g) = \infty$.

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 $u_i : G \to \mathbb{R}$ 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) $\forall\ ij \notin g$, if $u_i(g) < u_i(g + ij)$ then $u_j(g) > u_j(g + ij)$.

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

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

3.3 The model of competition for the access to information

3.3.1 Description of the model

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 a piece of 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

$$u_i^{CA}(g) = \sum_{j \neq i} \sum_{p(ij) \in P(ij)} \delta^{|p(ij)|} \prod_{k \in p(ij), k \neq i} f(d_k(g)) - cd_i(g) \tag{3.1}$$

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 a piece of 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 value all 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

$$\widetilde{u}_i^{CA}(g) = \sum_{j \neq i} \sum_{p(ij) \in P(ij)} \delta^{|p(ij)|} \prod_{k \in p(ij), k \neq i} \frac{1}{d_k} - c d_i(g) \tag{3.2}$$

It is clear that our model differs from the connections model (Jackson and Wolinsky, 1996) and the degree-distance-based connections model (Möhlmeier et al., 2016) in several respects. In the original Jackson-Wolinsky model, the benefit of i from the information sent by j in g is equal to $\delta^{l_{ij}(g)}$. In the degree-distance-based connections model, such a benefit is determined by $\hat{b}(d_{ij}(g), d_j(g))$, where the benefit function $\hat{b}$ is nonincreasing in each of the two variables (distance, degree) and $\hat{b}(\infty, d_j) = 0$ for every $d_j \in \{0, 1, \ldots, n-1\}$. In the present model, 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 3.1 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 δ^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 the Jackson-Wolinsky model. Similarly in the degree-distance-based connections model (and if we set a particular functional form of the benefit function $\hat{b}(d_{ij}(g), d_j(g)) = \frac{1}{1+d_j(g)} \delta^{d_{ij}(g)}$ 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 (3.1) and (3.2) 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 (3.2) 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 3.1 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 (3.2)). Note that in g', for sufficiently large δ this benefit $\frac{\delta^3}{8}$ is greater than the benefit $\frac{\delta^2}{14}$ from the information passed via the shortest path.

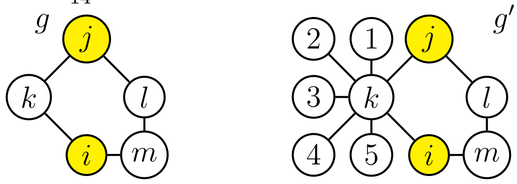


Figure 3.1: The models of competition for information versus the connections and degree-distance-based connections models

In the next sections we analyze pairwise stability in the model given by (3.1) and (3.2).

3.3.2 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 (3.1). 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:

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

  where $p^{-i}(kj)$ is a path from k to j that does not pass through i and $P^{-i}(kj)$

denotes the set of such paths.

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

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

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

Lemma 3.1. *Let $ij \in g$. Agent i prefers network g to network $g - ij$ if and only if*

$$\delta f(d_j(g)) \left(1 + I^{j-i}(g)\right) - I^{i,j}(g - ij) \left(1 - \frac{f(d_j(g))}{f(d_j(g) - 1)}\right) > c \qquad (3.3)$$

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 (3.3). ■

It results from the above formulation that if the communication technology is inefficient for high degrees ($\lim_{d \to \infty} f(d) = 0$) 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 3.2. *Suppose that the degree d_i of every agent i in the network satisfies $\underline{d} \leq d_i \leq \bar{d}$. 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 3.2 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 3.1. *Suppose that the communication technology satisfies $f(d)d < \alpha < 1$ 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 3.2 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 $\lim_{n\to\infty}\sum_{l=1}^{n-1}(kf(k))^l = \frac{kf(k)}{1-kf(k)}$. 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 high (on the order of magnitude of the total network size) and fairly homogeneous (i.e., whose variation is bounded by the constraint $k \leq d_i \leq \frac{k}{\delta}$) degrees. ■

The condition $k \leq d_i \leq \frac{k}{\delta}$ for each i puts a bound on the variation in degree of the nodes. 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 degree of every node i must verify $k \leq d_i \leq \frac{k}{\delta}$).

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.

3.3.3 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 3.2. *In the model defined by* (3.1) *the following holds:*

(i) The empty network $g^{\emptyset}$ is PS if $f(1)\delta \leq c$.

(ii) The star g^* with $n \geq 3$ is PS if

$$f(2)\delta + f(n-1)\left(2f(2) - f(1)\right)\delta^2 + (n-3)f(1)f(2)f(n-1)\delta^3 \leq c$$

and

$$c \leq \min\left(f(1)\delta, f(n-1)\delta + (n-2)f(1)f(n-1)\delta^2\right). \tag{3.4}$$

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

$$\begin{aligned} c \leq\ & f(n-1)\delta \\ & + f(n-1)\delta \sum_{k=1}^{n-2} \delta^k \frac{(n-2)!}{(n-2-k)!} f^{k-1}(n-1)\left((k+1)f(n-1) - kf(n-2)\right). \end{aligned} \tag{3.5}$$

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

$$c \leq \sum_{k=1}^{n-1} f^k(2)\delta^k + f^{n-2}(2)\delta^{n-1}\left(f(2) - f(1)\right) \tag{3.6}$$

and

$$c \geq f(3)\delta + 2f(3)\sum_{k=1}^{n-2} f^k(2)\delta^{k+1} + 2f(3)f^{\frac{n}{2}-1}(2)\delta^{\frac{n}{2}} - 2\sum_{k=\frac{n}{2}}^{n-1} f^k(2)\delta^k. \tag{3.7}$$

(v) The structure of $\frac{n}{2}$ separate pairs is PS if

$$f(2)\delta(1 + f(1)\delta) \leq c \leq f(1)\delta \tag{3.8}$$

The cost range for the stability is non-empty whenever

$$\delta \leq \frac{f(1) - f(2)}{f(1)f(2)}. \tag{3.9}$$

See the proof in Appendix 3.A.1.

The conditions for pairwise stability obviously involve the cost c, the decay δ 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. Conditions for pairwise stability of the network structures of n nodes usually depend on n, with the exception of the empty network and the structure of separate pairs. It is interesting to see if the stability of these structures is preserved when the number of agents becomes very large. For instance, if we consider the particular model defined in (3.2), then one can clearly see that the cost range for stability of the star decreases with the number of agents and becomes empty when $n \to \infty$. The same remark holds for the stability of the complete network. The next subsection is devoted to the analysis of this issue.

3.3.4 Asymptotic pairwise stability

As we have explained in the introduction, it is natural to assume that the number of agents who participate in the information exchange described by our models is large. For this reason, it is natural to focus on networks that are PS when n is large. To this effect, we introduce a notion of asymptotic (with respect to network size) pairwise stability.

Definition 3.1. *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, $i \neq j$*
$$\lim_{n\to+\infty} (u_i(g_n) - u_i(g_n - ij)) \geq 0$$
and if
$$\lim_{n\to+\infty} (u_i(g_n + ij) - u_i(g_n)) > 0,$$
then
$$\lim_{n\to+\infty} (u_j(g_n + ij) - u_j(g_n)) \leq 0.$$

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 3.1. *In the model* (3.1) *the asymptotic stability range is* (c, δ, f)*, i.e., it* *is determined by the cost* $c > 0$, $0 < \delta < 1$ *and a function* f *defined in* (3.1)*, i.e., verifying* $f(d) > 0$ *for all* $d \in \mathbb{N}_+$ *and* $f(1) \leq 1$.

We should note that APS is not a refinement of PS, nor is it a weaker concept. A certain network structure can be PS for some fixed n but not APS. It is also possible that a certain network structure is APS but not PS for small values of n. This will be illustrated by several examples further on. The main interest of the concept of APS is to reduce the parameter space since the parameter n disappears. The conditions for APS tend to be less involved than those for PS which may depend on n.

Proposition 3.3. *In the model* (3.1) *the following holds:*

(i) The empty network $g^{\emptyset}$ *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, when* $\lim_{n\to+\infty} f(n)n = 0$*, for any* $\delta < 1$.
If f *satisfies* $0 < \lim_{n\to+\infty} f(n)n \leq 1$*, then there exists a non-empty positive cost range for which the star* g^* *is APS if and only if*

$$1 \geq \lim_{n\to+\infty} f(n)n > \frac{f(2)}{f(1)(1-f(2))}.$$

In particular, for the model (3.2)*, i.e., when* $f(d) = 1/d$*, such a non-empty 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* $\lim_{n\to+\infty} f(n)n \leq 1$*, for any* $\delta < 1$.

(iv) There exists a non-empty positive cost range for which the circle g^c *is APS if and only if*

$$\delta \leq \frac{f(2) - f(3)}{f(2)f(3)}. \tag{3.10}$$

In particular, for the model (3.2) *such a non-empty cost range exists for any* $0 < \delta < 1$ *and is given by*

$$\frac{\delta}{3} + \frac{2\delta^2}{3(2-\delta)} \leq c \leq \frac{\delta}{2-\delta}. \tag{3.11}$$

(v) The structure of $\frac{n}{2}$ *separate pairs is APS whenever it is PS, i.e., when* (3.9) *is satisfied. In particular, for model* (3.2) *such a non-empty cost range exists*

for any $0 < \delta < 1$.

See the proof in Appendix 3.A.2.

Let us return to the possible and "ruled-out" PS categories of networks listed in Section 3.3.2. 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 3.3(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 3.3(iii) can also be deduced from Section 3.3.2, 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 nodes that 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)).

Asymptotic pairwise stability versus pairwise stability Clearly, if conditions for PS depend on n, PS and APS do not coincide. On the one hand, a structure can be PS but not APS, i.e., the cost range for stability becomes empty when the number of agents grows. An example is the complete network with n nodes which is PS under condition (3.5) but not APS for inefficient communication technology. Similarly, the star with n nodes is PS in the model (3.2) under condition (3.4) but not APS. On the other hand, a structure can be APS but not PS, which is the case for the circle, as shown in the following example.

Example 3.1. *In the model* (3.2)*, for any* $\delta \in (0, 1)$ *there exists a non-empty cost range for which the circle is APS and this cost range is given by* (3.11)*. On the other hand, by virtue of* (3.6) *and* (3.7)*, the cost range for which the circle with 4 nodes is PS in the model* (3.2) *is given by*

$$\frac{\delta}{3} + \frac{\delta^2}{6} - \frac{\delta^3}{12} \leq c \leq \frac{\delta}{2} + \frac{\delta^2}{4}.$$

Note that $\frac{\delta}{2}+\frac{\delta^2}{4}<\frac{\delta}{2-\delta}$ *for any* $\delta\in(0,1)$. *Hence, for the cost such that*

$$\max\left(\frac{\delta}{2}+\frac{\delta^2}{4},\frac{\delta}{3}+\frac{2\delta^2}{3(2-\delta)}\right)<c\leq\frac{\delta}{2-\delta}$$

the circle is APS (for large n*) but not PS for* $n=4$.

Tension between asymptotic pairwise stability and efficiency Under competition for the access to information, it might happen that a structure is APS but not SE in some cost range. To see that, first we calculate the following result.

Lemma 3.3. *The structure of separate pairs is more efficient than the circle for very large number of agents when* $c>\frac{2f(2)\delta}{1-\delta f(2)}-\delta f(1)$.

See the proof in Appendix 3.A.3.

From Lemma 3.3 and Proposition 3.3 we have the following conclusion.

Conclusion 3.1. *In the model* (3.2) *the circle* g^c *is not efficient for the whole cost range where it is APS, i.e., for* c *satisfying* (3.11).

Proof: In the model (3.2) the condition $c>\frac{2f(2)\delta}{1-\delta f(2)}-\delta f(1)$ under which the structure of separate pairs is more efficient than the circle for very large n is equivalent to $c>\frac{\delta^2}{2-\delta}$. On the other hand, note that for every $\delta\in(0,1)$, $\frac{\delta^2}{2-\delta}<\frac{\delta}{3}+\frac{2\delta^2}{3(2-\delta)}<\frac{\delta}{2-\delta}$. Hence, when the circle is APS, that is, for c such that $\frac{\delta}{3}+\frac{2\delta^2}{3(2-\delta)}\leq c\leq\frac{\delta}{2-\delta}$, it is not efficient. ■

3.3.5 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 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 the degree-distance-based connections model, 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 (3.1), 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 CA model, 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 (3.2) 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. If we denote by u_i^{JW} the utility of i in the Jackson-Wolinsky model, 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 (3.2), where $\widetilde{u}_i^{CA}(g^*) - \widetilde{u}_i^{CA}(g^* - ij) = \delta - c$ and $\widetilde{u}_j^{CA}(g^*) - \widetilde{u}_j^{CA}(g^* - ij) = \frac{\delta}{n-1} + \frac{(n-2)\delta^2}{n-1} - c$, for every δ and $n \geq 3$, there exists c such that $\widetilde{u}_i^{CA}(g^*) > \widetilde{u}_i^{CA}(g^* - ij)$ but $\widetilde{u}_j^{CA}(g^*) < \widetilde{u}_j^{CA}(g^* - ij)$.

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 non-empty component, in model (3.1) 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.

3.4 A model of competition for the use of information

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

$$u_i^{CU}(g) = \sum_{\{j \neq i | l_{ij}(g) < \infty\}} b(l_{ij}(g), x_{ij}(g), y_{ij}(g)) - cd_i(g) \tag{3.12}$$

where $b(l_{ij}(g), x_{ij}(g), y_{ij}(g))$ is a three variable function $b : \mathbb{N}_+ \times \mathbb{N}^2 \to \mathbb{R}_+$ for the value of the information that i receives from j, $l_{ij}(g)$ 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 $\sum_{l=1}^{+\infty} b(l, x, y) < +\infty$ for all $x, y \in \mathbb{N}$.

We note that if $l_{ij}(g) = 1$, then necessarily $x_{ij}(d) = 0$. Due to the fact that it is worse to an agent when others are closer to the sender than when they are at the same distance to the sender as the agent, we also assume

$$b(l, x, y) > b(l, x + k, y - k) \text{ for all } k \geq 1 \tag{3.13}$$

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: an agent does not care if others are informed, and the case where b reaches 0 if many others learn the information before the agent. Moreover, if the utility declines rapidly with distance, it seems reasonable that b should decline more rapidly with respect to agents who are closer to the sender than the given agent.

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 (3.12) 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, where this benefit is determined by $\hat{b}(1, d_x)$, with $\hat{b}(l_{ij}(g), d_j(g))$ depending on the distance between the sender j and the receiver i, and on the degree of the sender. In the distance-based model (Bloch and Jackson (2007)), which is an extension of the original connections model with the nonincreasing benefit function $\widetilde{b}(l_{ij}(g))$, the benefit from the information sent by a neighbor is always the same and is equal to $\widetilde{b}(1)$. For example, in Figure 3.1 the benefit of agent i from the information sent by his neighbor k in the degree-distance-based connections model is given by $\hat{b}(1, 2)$ and $\hat{b}(1, 7)$ in g and g', respectively, and in model (3.12) 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 3.1, for instance, the benefits of i from information sent by j are equal to $\widetilde{b}(2)$, $\hat{b}(2, 2)$ and $b(2, 2, 1)$ in network g, and $\widetilde{b}(2)$, $\hat{b}(2, 2)$ and $b(2, 2, 6)$ in network g'.

3.4.2 Pairwise stability of some "standard" architectures

We start our analysis of the model defined in (3.12) by proving conditions for pairwise stability of the prominent structures.

Proposition 3.4. *In the model defined by* (3.12) *the following holds:*

(i) The empty network $g^{\emptyset}$ *is PS if* $b(1, 0, 0) \leq c$.

(ii) The star g^* *with* $n \geq 3$ *is PS if*

$$c \geq b(1, 0, 1) - b(2, 1, n - 3) \tag{3.14}$$

and

$$c \leq \min\left(b(1, 0, 0), b(1, 0, n - 2) + (n - 2)b(2, 1, n - 3)\right). \tag{3.15}$$

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

$$c \leq b(1, 0, n-2) - b(2, n-2, 0). \tag{3.16}$$

(iv) *The circle g^c of $n > 3$ nodes is PS if*[2]

$$c \leq \sum_{k=1}^{\frac{n}{2}-1} b(k, 2k-2, 1) - \sum_{k=\frac{n}{2}+1}^{n-1} b(k, n-2, 0) \tag{3.17}$$

and for $n > 6$

$$c \geq b(1,0,2) - b(\frac{n}{2}, n-2, 0) + 2\sum_{k=2}^{\lfloor \frac{n}{4} \rfloor} b(k, 2k-2, 2) - 2\sum_{k=\lceil \frac{n}{4} \rceil + 1}^{\frac{n}{2}-1} b(k, 2k-2, 1) \tag{3.18}$$

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

(v) *The structure of $\frac{n}{2}$ separate pairs is PS if*

$$b(1, 0, 1) + b(2, 1, 0) \leq c \leq b(1, 0, 0). \tag{3.19}$$

See the proof in Appendix 3.A.4.

All the conditions for pairwise stability stated in Proposition 3.4 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 receiving an information alone. For other structures, a non-empty 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 (3.13), 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 at the same

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

distance to the sender as himself to a case where all other agents are closer to the sender, 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 at the same distance to the sender as himself to a situation with two more agents being informed. There exists a non-empty 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 is at the same distance to the sender.

3.4.3 Asymptotic pairwise stability of some "standard" architectures

Next we analyze the asymptotic pairwise stability (APS) as introduced in Definition 3.1. We check if the PS networks listed in Proposition 3.4 remain PS if the number of agents becomes very large.

Remark 3.2. *In the model* (3.12) *the asymptotic stability range is determined by* (c, b)*, i.e., by the cost* $c > 0$ *and a function* b *defined in* (3.12)*, i.e., verifying* $b(l, x, y) > 0$ *for all* l, x, y*.*

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

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

$$(A2) \qquad \lim_{k \to +\infty} b(l, x, k) = 0 \text{ for all } l \geq 2,\ x \in \mathbb{N}.$$

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

Proposition 3.5. *In the model* (3.12) *the following holds:*

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

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

$$b(1,0,1) - \lim_{n\to+\infty} b(2,1,n-3) \leq c \leq b(1,0,0). \tag{3.20}$$

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

$$\lim_{n\to+\infty} (n-2)b(2,1,n-3) \geq b(1,0,1) \tag{3.21}$$

and then the cost range must satisfy

$$b(1,0,1) \leq c \leq \min\left(b(1,0,0), \lim_{n\to+\infty}(n-2)b(2,1,n-3)\right). \tag{3.22}$$

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

$$c \leq \lim_{n\to+\infty} b(1,0,n-2) - \lim_{n\to+\infty} b(2,n-2,0). \tag{3.23}$$

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

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

$$b(1,0,2) + 2\sum_{k=2}^{+\infty} b(k,2k-2,2) \leq c \leq \sum_{k=1}^{+\infty} b(k,2k-2,1) \tag{3.24}$$

(v) *The structure of separate pairs is APS whenever it is PS, i.e., when* (3.19) *is satisfied.*

See the proof in Appendix 3.A.5.

Consider the case where the number of agents becomes very large and an agent still derives some benefit from the information even if many others are closer to the sender than himself. Then there always exists a non-empty 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 being at the same distance to the sender, the larger is this cost range. For the complete network, there exists a non-empty cost range for its pairwise stability if an agent strictly prefers to get the information directly and to be at the same distance to the sender 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.

3.4.4 Tension between asymptotic pairwise stability and efficiency

We will show that under competition in the use of information, some structures are APS but not SE. First, we will establish conditions under which the star is more efficient than some other structures.

Lemma 3.4. *The star g^* is more efficient than:*

(i) The empty network $g^{\emptyset}$ for sufficiently small costs, i.e., if

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

and under assumption (A1) it is the case for any $c > 0$ when n is sufficiently large.

(ii) The structure of disjoint pairs when

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

and under assumption (A1) it is the case for any $c > 0$ when n is sufficiently large.

(iii) The complete network g^N for sufficiently large costs, i.e., if

$$c > \frac{n-1}{n-2}b(1,0,n-2) - b(2,1,n-3) - \frac{b(1,0,0)}{n-2}. \tag{3.25}$$

Proof: Let $\widetilde{g}$ denote the structure of disjoint pairs. We have:

$$\sum_{i\in N} u_i^{CU}(g^{\emptyset}) = 0,\ \sum_{i\in N} u_i^{CU}(\widetilde{g}) = n\left(b(1,0,0) - c\right),$$
$$\sum_{i\in N} u_i^{CU}(g^*) = (n-1)\left(b(1,0,0) + b(1,0,n-2) + (n-2)b(2,1,n-3) - 2c\right)$$

and

$$\sum_{i \in N} u_i^{CU}(g^N) = n(n-1)\left(b(1,0,n-2) - c\right).$$

All parts result immediately from the comparison of the sums given above. ■

From Proposition 3.5 and Lemma 3.4 we can write the following conclusions.

Conclusion 3.2. *In the model of competition in the use of information:*

(i) Under assumption (A1) for sufficiently large n, the empty network is not efficient when being APS, i.e., for c satisfying $c \geq b(1,0,0)$.

(ii) Under assumption (A1) for sufficiently large n, the structure of disjoint pairs is not efficient when being APS, i.e., for c satisfying (3.19).

(iii) There exists a non-empty cost range for which g^N is APS but not efficient if

$$\lim_{n \to +\infty} b(2,1,n-3) > \lim_{n \to +\infty} b(2,n-2,0). \tag{3.26}$$

See the proof in Appendix 3.A.6.

3.4.5 Connectedness and degree homogeneity in (asymptotically) pairwise stable networks

Given the difficulty of obtaining a full characterization of the PS networks in our model, we will focus on exploring two of their properties, which are likely to be affected by the presence of the negative externality associated with connectivity, namely their connectedness and their degree distribution. Indeed, it is natural to ask whether the desire to access more exclusive information would tend to generate disconnected networks contrary to the original connections model. Secondly, we can ask what is the impact on the degree distribution? We have already seen that agents can still form networks with highly unequal degrees such as the star. Will they also tend to form networks in which everyone has a similar degree, so as to distribute the "nuisance of connectivity" more evenly? In this section, we provide a condition for the connectedness of PS networks and then turn to analyzing under which conditions k-regular networks can be stable.

We already know from Proposition 3.4(v) that the structure of disjoint pairs is PS under condition (3.19). However, the following proposition which gives a sufficient but not necessary condition for ensuring that only connected networks can be PS shows that the conditions under which disconnected networks can be PS are quite restrictive.

Proposition 3.6. *Suppose that for all $d \geq 1$, $b(1,0,d-1) - b(1,0,d) < b(2,d,d^2)$. Then no network containing more than one non-trivial component can be PS.*

Proof: Suppose that l and i are in different components in g and that i and l are the nodes with the highest degrees in their respective components. Since i is in a non-trivial connected component, he has at least one neighbor j (and similarly for l). If l forms a link to i, he gains at least $u_j^{CU}(g) - u_j^{CU}(g-ij) - b(1,0,d_i-1) + b(1,0,d_i) + b(2,d_j,|N_j^2(g)|) \geq u_j^{CU}(g) - u_j^{CU}(g-ij) - b(1,0,d_i-1) + b(1,0,d_i) + b(2,d_i,d_i^2)$, where $N_j^2(g)$ is the set of nodes which are exactly at distance 2 from node j in g. Indeed, the direct benefit of the link to i is lower for l than for j: $b(1,0,d_i)$ instead of $b(1,0,d_i-1)$. Otherwise, l gains at least as much from adding the link il as j gains from adding ij (strictly more if i and j are not disconnected in $g - ij$) and he also gains $b(2,d_j,|N_j^2(g)|)$ because he can reach j at distance 2. Since d_i is the maximal degree in i and j's component, $b(2,d_j,|N_j^2(g)|) \geq b(2,d_i,d_i^2)$. By the assumption, $u_l^{CU}(g+il) - u_l^{CU}(g) \geq u_j^{CU}(g) - u_j^{CU}(g-ij) - b(1,0,d_i-1) + b(1,0,d_i) + b(2,d_i,d_i^2) > u_j^{CU}(g) - u_j^{CU}(g-ij) \geq 0$. Thus agent l would like to form the link il. One can apply an identical argument to show that agent i wants to link to l implying that g is not PS. ■

The situation under competition for the use of information is not that different from the standard Jackson-Wolinsky framework, where an agent on a different island gains more than neighbors on the same island gain from maintaining the link. The only difference is a small disutility due to one additional link to the partner. If this loss is small compared to the gain in indirect benefits (that we underestimate in the proof), no network with several non-trivial components can be stable.

Note that Proposition 3.6 is consistent with condition (3.19) which implies that there exists a non-empty cost range for pairwise stability of the structure of separate pairs if $b(1,0,1) + b(2,1,0) \leq b(1,0,0)$. Taking $d = 1$ in Proposition 3.6 leads to $b(1,0,0) < b(1,0,1) + b(2,1,1)$ which implies $b(1,0,0) < b(1,0,1) + b(2,1,0)$.

We turn now to the analysis of k-regular networks. We start by introducing some concepts and conditions that will appear in our first result.

Definition 3.2. *An agent i has a growth dummy neighbor if there exists j such that $ij \in g$ and $N_j(g) \subset N_i(g) \cup \bigcup_{\{k|ik\in g, k\neq i\}} N_k(g)$.*

A growth dummy neighbor of i is a neighbor who does not contribute to i's neighborhood growth. Growth dummy agents exist in networks with much cohesion and "group structure" in which an agent's neighbors are likely to be connected to each other and have many neighbors in common.

Condition 3.1. *We say that Condition 3.1 is verified for d if for every $k \geq 3$ and $0 \leq x \leq k^2$, $k, x \in \mathbb{N}$, it holds that:*

(1) $b(1,0,k-1) - b(1,0,k) < k[b(2,k,x) - b(d-1,k,x)]$

(2) $b(2,k-1,k^2) \geq b(d,k,0)$

(3) $b(1,0,l-1) - b(1,0,k-1) > b(d,l-1,0) - b(2,k-1,k^2)$ *for every $l < k$.*

Condition 3.2. *For every $k \geq 3$ and $0 \leq x \leq k^2$, $b(1,0,k-1) - b(1,0,k) < kb(2,k,x)$.*

If Condition 3.1 holds for d, then it also holds for any $\tilde{d} > d$. In particular, when $d \to \infty$, parts (2) and (3) are always verified, and part (1) becomes Condition 3.2. When d is large, Condition 3.1 is not very demanding. It requires only that one additional neighbor does not reduce benefits too sharply and that there are sufficient benefits at distance 2.

Proposition 3.7. *If Condition 3.1 is verified for d, then no network g containing an agent i who has a growth dummy neighbor and an agent l such that $d_l(g) \leq d_i(g)$ and $d_{il}(g) \geq d$ can be PS.*

See the proof in Appendix 3.A.7.

Corollary 3.1. *We have the following:*

(i) If Condition 3.1 is verified for d, then no k-regular network g such that its diameter $diam(g) > 2d$ and which contains an agent who has a growth dummy neighbor can be PS.

(ii) If Condition 3.2 is verified and k is bounded independently of n, then no k-regular network g which contains an agent who has a growth dummy neighbor

can be APS.

Proof: (i) Let i be the agent with a growth dummy neighbor. Since $diam(g) > 2d$, there exists some agent l such that $d_{il} \geq d$. Since the network is k-regular, $d_i(g) = d_l(g)$. The conditions of Proposition 3.7 are fulfilled.
(ii) Let i be the agent with a growth dummy neighbor. There is an agent l at distance at least $diam(g_n)/2$ from i, $d_l(g) = d_i(g)$. By replacing d_{il} in Proposition 3.7 by $diam(g_n)/2$, we obtain that g_n cannot be PS if $b(1, 0, k-1) - b(2, k-1, k^2) < b(1, 0, k) - b(diam(g_n)/2, k, 0) + \sum_{a \in N_l(g)} b(2, k, x_a) - b(diam(g_n)/2 - 1, k, x_a)$. If k is bounded independently of n, then we have $\lim_{n\to\infty} diam(g_n) = \infty$. Assuming that the sum of the indirect benefits converges, we have $\lim_{n\to\infty} b(diam(g_n)/2, k, 0) = \lim_{n\to\infty} b(diam(g_n)/2 - 1, k, x_a) = 0$. Moreover, $|N_l(g)| = k$. Since $x_a \leq k^2$ (it is the number of agents at distance 2 from l), by Condition 3.2, g cannot be APS. ■

The types of networks that are ruled out (under fairly weak assumptions on the benefits) are networks with many locally redundant links and large diameters. Networks of this type are for example networks based on some notion of proximity (whether it is geographical or in terms of similar attributes): agents make links to others who are geographically close and do not form long range links. Let us give some examples of networks which contain dummy players. In particular, several types of k-regular networks have this properties and thus cannot be stable.

Example 3.2. *One example is the "geography based" circle network. Let the agents* $1, \ldots, n$ *be located on a circular graph and let agent* i *be connected to the agents numbered* $i-m, i-(m-1), \ldots, i+1, \ldots, i+m$ $(mod(n))$. *This network is* k*-regular with* $k = 2m$ *and has the growth dummy player property when* $m \geq 2$. *For example, agent* $i+1$ *is a growth dummy for agent* i. *(This well known model is in fact a special case of the circulant graph example given below).*

Example 3.3. *Another example is a certain type of* k*-regular bi-partite graph. Suppose that* $n = 2m$. *Divide the nodes into two disjoint sets of size* m. *Label the nodes in the sets* $a_1, \ldots, a_m$ *and* $b_1, \ldots, b_m$*, respectively. Let* $a_j b_l \in g \iff |j-l| \leq c$ $(mod(n))$. *This is a* $k = 2c$ *regular network and agent* b_j *is a growth dummy player for player* a_j*: the neighbors of* b_j *are* $a_{j-c}, \ldots, a_{j+c}$*, but these agents are also neighbors of* $b_{j-c}, \ldots, b_{j+c}$ *who are neighbors of* a_j.

Example 3.4. *Consider the following* k*-regular circulant graph. Let the agents numbered* $1, \ldots, n$ *be located on a circular graph and let* $ij \in g \iff j = i+z, i+$

$z+1,\ldots,i+z+k,i-z,i-(z+1),\ldots,i-(z+k)$ $(mod(n))$ *with* $k \geq 3$. *This is a* k*-regular graph and* $j = i+z+1$ *is a growth dummy player for* i *because his neighbors are also neighbors of either* $i+z$ *or* $i+z+2$.

The previous subsection provided examples of different types of k-regular networks which in virtue of Proposition 3.7 cannot be PS or APS under the weak assumptions on the benefits. Now we will show that there are other constructions of k-regular networks that can be PS/APS for the same degree k and under the same assumptions on the benefits.

Example 3.5. *We construct the following* $(2m-2)$*-regular network* g. *Let* $n = m(m+1)$ *with* $m \in \mathbb{N}$. *Divide the agents into* $m+1$ *islands consisting of* m *agents. Number the agents on island 1 by* $a_{1,1},\ldots,a_{1,m}$ *and the agents on island* l *by* $a_{l,1},\ldots,a_{l,m}$. *Let* $a_{x,y}a_{p,q} \in g$ *if* $x = p$, *i.e., if agents belong to the same island. On island* 1, *let agent* $a_{1,j}$ *be linked to agent* $a_{v,j}$ *on every island* v *except* $v = j+1$ $(mod(m))$. *On island* 2 *let agent* $a_{2,j}$ *be linked to agent* $a_{v,j}$ *on every island* v *except* $v = j+2$ $(mod(m))$. *On island* t *let agent* $a_{t,j}$ *be linked to agent* $a_{v,j}$ *on every island* v *except* $v = j+t$ $(mod(m))$.*This ensures that each agent on the same island has a distinct island he is not linked to. Moreover, if an agent* $a_{x,y}$ *is not linked to the agent with label* y *on island* v, *he is linked to an agent who is linked to this agent. Network* g *is* $(2m-2)$*-regular. We can verify that* $diam(g) = 2$. *Indeed, an agent is linked directly to some agent on every island different from his own except one. On this island, the agents with different labels are reached by the agents on his island. The agent with his label is reached by some agent on another island with the same label.*

Proposition 3.8. *The* $(2m-2)$*-regular network defined in Example 3.5 is PS in a non-empty cost range whenever* $b(d,.,.)$ *is strictly decreasing in* d. *It is not APS if condition* $(A1)$ *does not hold.*

See the proof in Appendix 3.A.8.

For $n = m(m+1)$ we can therefore construct $(2m-2)$-regular networks that are PS in a non-empty cost range as long as the benefit is strictly decreasing with respect to distance, whereas other $(2m-2)$-regular networks of the types in the previous examples (nearest neighbor circle, bipartite graph, circulant graph) are not PS. These observations indicate that k-regularity itself is not sufficient to determine whether a network is PS. This is not so surprising since k-regular networks for the same k comprise very different structures. In particular, the diameter of such networks can be very different.

Proposition 3.8 shows that there exists a k-regular network with k large and small diameter ($diam = 2$) that can be PS in our model under very weak assumptions. Now we will determine the conditions under which k-regular networks with k not too large (i.e., such that the diameter of the network is large) can be APS. We introduce the following condition.

Condition 3.3. *The network does not contain a cycle of length $l \leq 2d + 1$.*

We can show the following results.

Proposition 3.9. *Assume that $b(\tilde{d}, ., .) = 0$ whenever $\tilde{d} > d$. Let g be a k-regular network, where k is bounded independently of n.*

- *If $b(1, 0, k-1) - b(1, 0, k) < b(2, k, k^2)$, then g cannot be APS.*
- *If g verifies Condition 3.3 and $b(1, 0, k-1) - b(1, 0, k) > \sum_{l=2}^{l=d} b(l, k^{l-1}, k^l)$, then g is APS in a cost range of size*

$$b(1, 0, k-1) - b(1, 0, k) - \sum_{l=2}^{l=d} b(l, k^{l-1}, k^l).$$

See the proof in Appendix 3.A.9.

3.4.6 Connectedness and degree homogeneity in asymptotically pairwise stable networks as a function of the level of aversion to others being informed

In this section we give some results on the connectedness and degree homogeneity of APS structures under different assumptions about the level of aversion to others being informed. We start with the case where there is moderate aversion to others being informed, in the sense that the benefit declines if a very large number of agents are closer to the sender than oneself but not if they are at the same distance to the sender as oneself (NON A1). In this case we have a clear characterization in the case where the cost is small compared to the direct benefit. The second case in which we obtain some interesting results is when aversion to others being informed is very strong: the benefit must not only go to zero when others are at the same distance to the sender as oneself (A2) but it must decline at a sufficiently rapid pace. We also have a proposition in the intermediary case, which mainly shows that in this case there are no clear predictions about degree homogeneity/heterogeneity without making additional assumptions.

The following definition will be important for characterizing the behavior in the first case (NON A1).

Definition 3.3. *Note by $B(g_n)$ the set of nodes whose degrees in g_n are bounded independently of n. Let $B^C(g_n)$ be the complement of $B(g_n)$, so that $\{1, \ldots, n\} = B(g_n) \cup B^C(g_n)$. Let SG be the set of connected networks g verifying the following properties: g is connected, $B^C(g_n) \neq \emptyset$. If $|B(g_n)|$ is not bounded, then there exists at least one node l in $B^C(g_n)$ for which $lim_{n\to\infty}|\{j \in B(g_n) \mid lj \in g_n\}| = \infty$, that is, a node l which has a "large" number of links to nodes in $B(g_n)$.*

The networks in SG are basically of two types. Either there is a very large completely connected component of nodes whose degrees are not bounded independently of n and a "small" (size bounded independently of n) group of nodes with low (bounded) degrees. The network is connected, so links exist between the "high" and "low" degree group. If there is a large (size not bounded independently of n) number of low-degree nodes, then there is at least one high degree node who has an unbounded number of links to nodes in the low degree group. In the first case, we have a structure reminiscent of a complete graph, since the nodes who are not part of the large connected component make up only a negligible fraction of the total number of nodes. 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 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 a large number of agents being closer to the sender and under a condition on direct benefits.

Proposition 3.10. *If $\lim_{x\to\infty} b(2, x, 0) = 0$ and $\lim_{n\to\infty} b(1, 0, n-2) > c$, then:*

- *The complete graph is PS.*

- *Any g, other than the complete graph that is APS belongs to the class SG introduced in Definition 3.3.*
- *In particular, if g is PS, then it is connected and its diameter is bounded independently of n.*

See the proof in Appendix 3.A.10.

The intuition for this result is that with the given conditions on the payoff parameters, the diameter has to be small. If the diameter is small and all agents have similar degrees, these degrees have to be high. However, under assumption (NON A1) an agent wants to form a direct link to another agent with a very high degree to avoid being informed after all his neighbors. Thus a network with similar and high degrees will always "collapse" to a complete network. Therefore, if the network is not the complete graph, there must be degree inequality. A network with unequal degrees can be stable because the low degree nodes want to link to the high degree nodes but the converse is not true for all payoff parameters.

If the direct benefits are lower compared to the cost, many different types of structures can be PS and additional assumptions are needed to determine which ones arise. The following proposition is mainly for the sake of comparison with the other cases.

Proposition 3.11. *The assumptions that* $\lim_{x\to\infty} b(2,x,0) = 0$ *and* $0 < \lim_{n\to\infty} b(1,0,n) < c$ *are:*

- *not compatible with APS of the complete graph;*
- *compatible with the APS of the star;*
- *compatible with the APS of networks with small (bounded independently of n) and large (unbounded) diameters.*

We omit the proof which is not complicated. We now consider the case where there is stronger aversion to others being informed, in the sense that utility decreases to zero also when a large number of others are of the same distance from the sender, i.e., when Assumption (A2) holds. However, the conditions under which we obtain PS structures that are clearly different from those with weaker aversion to others being informed are stronger than just assumption (A2). We also make an assumption on the rate of decline.

Proposition 3.12. *Suppose that Assumption (A2) holds and that*

$$\lim_{n\to\infty} nb(2,0,n) < c.$$

Then any g that is APS is such that:

- *The degrees of all nodes in g are bounded independently of n.*
- *The network diameter is not bounded independently of n.*
- *Under certain assumptions on the benefit parameters non connected networks can be stable.*

See the proof in Appendix 3.A.11.

3.5 Discussion and concluding remarks

Since the payoff functions in our two models of a 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 (asymptotically) 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. We have shown that if an agent sees little value in an information that many others have received before him, and if costs are low enough, 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 although 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 a piece of 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. Moreover, we have shown that in both models some structures are not efficient when being APS. 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 information 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 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.

3.A Appendix

3.A.1 Proof of Proposition 3.2

Proof: (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 (3.4).

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

$$\begin{aligned}
&u_i^{CA}(g^N) - u_i^{CA}(g^N - ij) \\
&= f(n-1)\delta + (n-2)f(n-1)\,(2f(n-1) - f(n-2))\,\delta^2 \\
&\quad + f^2(n-1)(n-2)(n-3)\,(3f(n-1) - 2f(n-2))\,\delta^3 + \cdots \\
&\quad + f^{n-2}(n-1)(n-2)!((n-1)f(n-1) - (n-2)f(n-2))\delta^{n-1} - c
\end{aligned}$$

which leads to (3.5).

(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 $2f(2)\delta + 2f^2(2)\delta^2 + \ldots + 2f^{n-1}(2)\delta^{n-1} - 2c \geq f(2)\delta + f^2(2)\delta^2 + \ldots + f^{n-2}(2)\delta^{n-2} + f^{n-2}(2)f(1)\delta^{n-1} - c$ iff condition (3.6) 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. Then $u_i^{CA}(g^c) \geq u_i^{CA}(g^c + ik)$ iff

$$\begin{aligned}
&2f(2)\delta + 2f^2(2)\delta^2 + \ldots + 2f^{n-1}(2)\delta^{n-1} - 2c \geq \\
&2f(2)\delta + f(3)\delta + 2f^2(2)\delta^2 + 2f(2)f(3)\delta^2 + \ldots \\
&\quad + 2f^{\frac{n}{2}-1}(2)\delta^{\frac{n}{2}-1} + 2f^{\frac{n}{2}-2}(2)f(3)\delta^{\frac{n}{2}-1} + 4f^{\frac{n}{2}-1}(2)f(3)\delta^{\frac{n}{2}} + 2f^{\frac{n}{2}}(2)f(3)\delta^{\frac{n}{2}+1} + \ldots \\
&\quad + 2f^{n-3}(2)f(3)\delta^{n-2} + 2f^{n-2}(2)f(3)\delta^{n-1} - 3c
\end{aligned}$$

iff condition (3.7) is satisfied.

(v) 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:

$$\begin{aligned}
u_i^{CA}(g) - u_i^{CA}(g - ij) &= f(1)\delta - c \geq 0 \\
u_i^{CA}(g + ik) - u_i^{CA}(g) &= f(2)\delta + f(1)f(2)\delta^2 - c \leq 0
\end{aligned}$$

Hence, g is PS iff (3.8) holds. The cost range is non-empty whenever $f(2)\delta(1+$

$f(1)\delta) \leq f(1)\delta$ which gives condition (3.9). ■

3.A.2 Proof of Proposition 3.3

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

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

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

Such a positive cost exists when

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

Let $\lim_{n\to+\infty} f(n)n := a \in (0,1]$. We need to consider the inequality

$$af(1)f(2)\delta^2 - af(1)\delta + f(2) \leq 0$$

We have $\Delta = a^2 f^2(1) - 4af(1)f^2(2) = af(1)(af(1) - 4f^2(2))$, and $\Delta \geq 0$ iff $2f(2) \leq \sqrt{af(1)} \leq 1$. Since $\delta \in (0,1)$, we have the condition $\frac{af(1)-\sqrt{\Delta}}{2af(1)f(2)} < 1$ which is equivalent to $a > \frac{f(2)}{f(1)(1-f(2))}$. In particular, the condition is not satisfied for $f(d) = \frac{1}{d}$.

(iii) Stability of the complete network requires

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

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

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

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

Thus

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

We have

$$f(n-2)f^{k-1}(n-1)\frac{(n-2)!}{(n-2-k)!} = (n-2-k+1)\ldots(n-3)(n-2)f(n-2)f^{k-1}(n-1) < ((n-2)f(n-2))^k \leq 1.$$

Thus

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

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

(iv) The circle is APS if

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

or equivalently

$$\begin{aligned} & f(3)\delta + 2f(3)\delta\lim_{n\to+\infty}\left[\frac{f(2)\delta-(f(2)\delta)^{n-1}}{1-f(2)\delta}\right] \leq c \leq \lim_{n\to+\infty}\frac{f(2)\delta-(f(2)\delta)^n}{1-f(2)\delta} \\ \Longleftrightarrow\ & f(3)\delta + 2f(3)\delta\left[\frac{f(2)\delta}{1-f(2)\delta}\right] \leq c \leq \frac{f(2)\delta}{1-f(2)\delta} \qquad (3.27) \\ \Longleftrightarrow\ & f(3)+f(3)f(2)\delta \leq \frac{c(1-f(2)\delta)}{\delta} \leq f(2). \end{aligned}$$

Such a positive cost exists whenever $f(3)+f(3)f(2)\delta \leq f(2) \iff \delta \leq \frac{f(2)-f(3)}{f(2)f(3)}$. Moreover, if we apply $f(d) = \frac{1}{d}$ to (3.27), we get the cost range given by (3.11). ∎

3.A.3 Proof of Lemma 3.3

Proof: We compare the circle with the disjoint pair structure denoted by $\widetilde{g}$. Consider

$$\lim_{n\to\infty}\left(\sum_{i\in N} u_i^{CA}(\widetilde{g}) - \sum_{i\in N} u_i^{CA}(g^c)\right)$$

We have

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

and

$$\lim_{n\to\infty}\left(n\delta f(1) - nc - \frac{2nf(2)\delta\,(1-(\delta f(2))^{n-1})}{1-\delta f(2)} + 2nc\right) =$$
$$\lim_{n\to\infty}\left(n\delta f(1) + nc - \frac{2nf(2)\delta}{1-\delta f(2)} - \right) = \lim_{n\to\infty}\left(f(1)\delta + c - \frac{2f(2)\delta}{1-\delta f(2)}\right) n$$

This quantity is positive if $c > \frac{2f(2)\delta}{1-\delta f(2)} - \delta f(1)$. ■

3.A.4 Proof of Proposition 3.4

Proof: (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 conditions (3.14) and (3.15).

(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 (3.17).

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 $n \geq 8$. The condition $u_i^{CU}(g^c) - u_i^{CU}(g^c + ik) \geq 0$ for $ik \notin g$ is equivalent to (3.18). 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 (3.18) 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.

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

3.A.5 Proof of Proposition 3.5

Proof: (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 $\lim_{k \to +\infty} b(l, x, k) > 0$ for all l, x. The star is APS whenever (3.20) is satisfied, as $b(1,0,0) < \lim_{n \to +\infty} (b(1,0,n-2) + (n-2)b(2,1,n-3))$. Note that the cost range is non-empty, since $b(1,0,1) - \lim_{n \to +\infty} b(2,1,n-3) < b(1,0,0)$.

Suppose now that assumption (A2) is satisfied. The star is APS whenever (3.22) is satisfied. If $\lim_{n \to +\infty} (n-2)b(2,1,n-3) < b(1,0,1)$, then this cost range is empty. On the contrary, if $\lim_{n \to +\infty} (n-2)b(2,1,n-3) \geq b(1,0,1)$, then there exists some cost that satisfies (3.22), since $b(1,0,1) < b(1,0,0)$.

(iii) Stability of the complete network requires (3.23) to be satisfied. If we have

$\lim_{n\to+\infty} b(1,0,n-2) > \lim_{n\to+\infty} b(2,n-2,0)$ then the right hand side of (3.23) is positive and g^N is APS. If $\lim_{n\to+\infty} b(1,0,n-2) = \lim_{n\to+\infty} b(2,n-2,0)$, then the right hand side of (3.23) 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 $n \to +\infty$ in the right hand expressions in (3.17) and (3.18), we obtain condition (3.24). ■

3.A.6 Proof of Conclusion 3.2

Proof: (i) By virtue of Lemma 3.4, g^* is more efficient than $g^\emptyset$ if

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

and when moving to the limit with $n \to +\infty$, under assumption (A1), g^* is always more efficient than $g^\emptyset$, and therefore also for the cost range when $g^\emptyset$ is APS.

(ii) By virtue of Lemma 3.4, g^* is more efficient than the structure of disjoint pairs if

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

which is equivalent to

$$c < \frac{n-1}{n-2} b(1,0,n-2) + (n-1)b(2,1,n-3) - \frac{b(1,0,0)}{n-2}.$$

When moving to the limit with $n \to +\infty$, this condition is always satisfied under assumption (A1), in particular when the structure of disjoint pairs is APS.

(iii) When moving to the limit in (3.25), by virtue of Lemma 3.4, g^* is more efficient than g^N if

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

From monotonicity of function b, and assumptions (3.13) and (3.26) we have

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

and therefore from Proposition 3.5(iii), g^N is APS and the cost range for its pairwise stability satisfies (3.23). Hence, if the cost range is such that

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

and

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

then g^N is APS but not efficient. This cost range is non-empty under assumption (3.26). ■

3.A.7 Proof of Proposition 3.7

Proof: Let i be the agent who has a growth dummy neighbor j. If g is PS, then $u_i^{CU}(g) - u_i^{CU}(g-ij) \geq 0$. Let $d_i = k$. The growth dummy condition implies that the only loss from cutting the link to j is that j himself moves further away: $u_i^{CU}(g) - u_i^{CU}(g-ij) \leq b(1,0,k-1) - b(2,k-1,k^2) - c$. By the assumption, there exists l such that $d_{il} \geq d$ and $d_l \leq d_i$. Suppose first that $d_l = d_i$. PS requires that $u_i^{CU}(g+il) - u_i^{CU}(g) \leq 0$.

We have $u_i^{CU}(g+il) - u_i^{CU}(g) \geq b(1,0,k) - b(d_{il},k,0) + \sum_{a\in N_l(g)} b(2,k,x_a) - b(d_{il}-1,k,x_a) - c$, where $x_a = \sum_{m\in N} \mathbf{1}_{d_{ml}(g)=2}$. Forming a link to l brings the k direct neighbors of l to distance 2 from i, while they were previously at distance $d_{il}-1$. By symmetry, we also have $u_l^{CU}(g+il) - u_l^{CU}(g) \geq b(1,0,k) - b(d_{il},k,0) + \sum_{a\in N_l(g)} b(2,k,x_a) - b(d_{il}-1,k,x_a) - c$. Thus pairwise stability of g will fail to hold if

$$u_i^{CU}(g) - u_i^{CU}(g-ij) \leq$$

$$b(1,0,k) - b(d_{il},k,0) + \sum_{a\in N_l(g)} b(2,k,x_a) - b(d_{il}-1,k,x_a) - c \iff$$

$$b(1,0,k-1) - b(2,k-1,k^2) \leq$$

$$b(1,0,k) - b(d_{il},k,0) + \sum_{a\in N_l(g)} b(2,k,x_a) - b(d_{il}-1,k,x_a)$$

Since $x_a \leq k^2$, $\sum_{a\in N_l(g)} b(2,k,x_a) - b(d_{il}-1,k,x_a) \geq k \min_{1\leq x\leq k^2}[b(2,k,x) - b(3,k,x)]$ and by Conditions 3.1 and 3.2 we conclude that $u_i^{CU}(g) - u_i^{CU}(g-ij) \leq b(1,0,k) - b(d_{il},k,0) + \sum_{a\in N_l(g)} b(2,k,x_a) - b(d_{il},k,x_a) - c \leq \min[u_l^{CU}(g+il) - u_l^{CU}(g), u_i^{CU}(g+il) - u_i^{CU}(g)]$. In other words, either i does not want to maintain the link $ij \in g$

or i and l both wish to add the link $il \notin g$, which contradicts g being PS. Finally, if $d_l < d_i$ and $d_l = k'$, then i gains at least $b(1,0,k') - b(d_{il}, k'-1, 0) - c > b(1,0,k-1) - b(2,k-1,k^2) - c = u_i^{CU}(g) - u_i^{CU}(g-ij)$. Hence, in this case as well by Condition 3.1, i would like to add il whenever he wants to maintain ij. ■

3.A.8 Proof of Proposition 3.8

Proof: Consider agent $a_{x,y}$. He will maintain a link to an agent on his own island $a_{x,z}$ if $u_{a_{x,y}}^{CU}(g) - u_{a_{x,y}}^{CU}(g - a_{x,y}a_{x,z}) \geq 0$. $u_{a_{x,y}}^{CU}(g) - u_{a_{x,y}}^{CU}(g - a_{x,y}a_{x,z}) = b(1,0,2m-1) - b(2,2m-1,n-(2m+1)) + b(2,2m-1,n-(2m+1)) - b(3,n-2,0) - c$. Note that by breaking the link to $a_{x,z}$, there is an agent that moves from distance 2 to 3, namely agent $a_{y+x(mod(m)),z}$. Indeed, agent $a_{x,y}$ is not directly linked to this island. He can reach agent $a_{y+x(mod(m)),y}$ in two steps through some y label agent on another island and then $a_{y+x(mod(m)),z}$ in another step, or reach some agent $a_{y+x(mod(m)),r}$ in two steps, but reaching $a_{y+x(mod(m)),z}$ now requires 3 steps. Suppose that agent $a_{x,y}$ breaks his inter island link to some $a_{v,y}$. We have $u_{a_{x,y}}^{CU}(g) - u_{a_{x,y}}^{CU}(g - a_{x,y}a_{v,y}) \geq 0$ and $u_{a_{x,y}}^{CU}(g) - u_{a_{x,y}}^{CU}(g - a_{x,y}a_{x,z}) = b(1,0,2m-1) - b(2,2m-1,n-(2m+1)) + b(2,2m-1,n-(2m+1)) - b(3,n-2,0) - c$. Indeed, there is an agent that moves from distance 2 to 3, namely agent $a_{v,z}$ with $v = x+z\ (mod(m))$. By construction, agent $a_{x,z}$ is not linked to agent $a_{v,z}$. After breaking the link to island v, it requires 3 steps to reach $a_{v,z}$. Now the utility of adding a link is $u_{a_{x,y}}^{CU}(g + a_{x,y}a_{x+y(mod(m)),l}) - u_{a_{x,y}}^{CU}(g) = b(1,0,2m-1) - b(2,2m-1,N-(2m-1)) - c$. No agent except the one he links to moves closer since everyone is already at distance 2. If $b(1,0,2m-1) - b(2,2m-1,n-(2m+1)) < c < b(1,0,2m+1) - b(2,2m-1,n-(2m+1)) + b(2,2m-1,n-(2m+1)) - b(3,n-2,0)$, then this network is PS. A cost range where this holds will exist if $b(2,2m-1,n-(2m+1)) > b(3,n-2,0)$, and $b(2,2m-1,n-(2m+1)) > b(2,2m-1+n-(2m+1),0) = b(2,n-2,0) \geq b(3,n-2,0)$. As long as the benefit at distance 2 is strictly greater than at distance 3, such a cost range will exist. For finite n, this k-regular network is therefore PS for a non-empty cost range, whereas other k-regular networks cannot be PS when benefits satisfy Condition 3.1. However, the cost range for which the island network is PS goes to zero if condition $(A1)$ does not hold. Therefore it is not always APS. ■

3.A.9 Proof of Proposition 3.9

Proof: Suppose that k is bounded independently of n. Thus $\lim_{n\to\infty} diam(g_n) = \infty$. To simplify the analysis we will assume that there exists a cut-off level d (d large) above which benefits are zero. Suppose that $l_{ij}(g) > d$. If i links to j, then he gains $b(1,0,k)$ in direct benefits. All the indirect neighbors of j were previously at distance greater than d from i. Note by $n_l^k(g) =: |N_l^k(g)|$, the number of agents at distance exactly k from l. The indirect benefits are $\sum_{k=2}^{k=d}\sum_{\{l\in N_j^{k-1}(g)\}} b(k, \sum_{r=1}^{r=k-1} n_l^r(g), n_l^k(g))$. We should note that for all $1 \leq k \leq d$ and $l \in N_j^k(g)$, $n_l^k(g) = n_l^k(g+ij)$. Now, let m be a neighbor of j. If m breaks the link to j, then his loss of direct benefits is $b(1,0,k-1)$. His loss of indirect benefits can potentially be much smaller than the gain of i if many of the indirect neighbors of j are also indirect neighbors of m in g. However an upper bound on the loss is $\sum_{k=2}^{k=d}\sum_{\{l\in N_j^{k-1}(g)\}}[b(k, \sum_{r=1}^{r=k-1} n_l^r(g), n_l^k(g)) - b(2,k,k^2)]$. Indeed, if i links to j, then agent m is at distance 2. Agent m himself does not obtain this benefit from linking to j. Stability requires that m wishes to maintain mj and i does not want to form ij. We have

$$u_i^{CU}(g+ij) - u_i^{CU}(g) = b(1,0,k) + \sum_{k=2}^{k=d}\sum_{\{l\in N_j^{k-1}(g)\}} b(k, \sum_{r=1}^{r=k-1} n_l^r(g), n_l^k(g)) - c$$

$$u_m^{CU}(g) - u_i^{CU}(g-jm) \leq b(1,0,k-1)$$
$$+\sum_{k=2}^{k=d}\sum_{\{l\in N_j^{k-1}(g)\}}[b(k, \sum_{r=1}^{r=k-1} n_l^r(g), n_l^k(g)) - b(2,k,k^2)] - c$$

So g cannot be pairwise stable if $b(1,0,k-1) - b(1,0,k) < b(2,k,k^2)$. Now suppose that g satisfies Condition 3.3. An implication of Condition 3.3 is that if $ab \in g$, then $(\bigcup_{l\leq d} N_a^l(g)) \cap (\bigcup_{l\leq d} N_b^l(g)) = \emptyset$. If Condition 3.3 holds, then

$$u_m^{CU}(g) - u_i^{CU}(g-jm) = b(1,0,k-1)+$$

$$+\sum_{k=2}^{k=d}\sum_{\{l\in N_j^{k-1}(g)\}} b(k, \sum_{r=1}^{r=k-1} n_l^r(g), n_l^k(g)) - \sum_{l=2}^{l=d} b(l, k^{l-1}, k^l) - c$$

Indeed, Condition 3.3 implies that none of m's indirect neighbors at distance not greater than d are indirect neighbors of j at distance not greater than d. Like before, stability requires that m wishes to conserve mj and that i does not want to form

ij. This is also a sufficient condition, because Condition 3.3 implies that neighbor growth is tree like everywhere up to level d, and if agent i does not want to form link ij, no agent wants to add a link that is not in g. This holds if $b(1,0,k-1)-b(1,0,k) > \sum_{l=2}^{l=d} b(l,k^{l-1},k^l)$. ■

3.A.10 Proof of Proposition 3.10

Proof: 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 $\lim_{n\to\infty} \bar{d}(g_n) = \infty$. 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 $\lim_{\bar{d}\to\infty} b(\bar{d},x,y) = 0$ for any x and y, due to the assumption that the sum of utilities converges. Moreover, we know that $\lim_{n\to\infty} b(1,0,n-2) > c$. This contradicts the stability of the network with diameter $\bar{d}$. From this it also follows that $B^C(g_n)$ cannot be empty. If it were, all agents would be in $B(g_n)$ and have degrees bounded independently of n. The diameter of such a network is not bounded independently of n which is, by the previous argument, impossible. To show that all agents with asymptotically unbounded degree must be linked, note that the benefit of forming a link with $i \in B^C(g_n)$ if it does not exist is at least $b(1,0,n-2) - b(2,d_i,0)$. Since $\lim_{n\to\infty} d_i = \infty$, $\lim_{n\to\infty} b(2,d_i,0) = 0$. Moreover, if the complete graph is APS, then $b(1,0,n-2) - b(2,n-2,0) > c$ and $\lim_{n\to\infty} b(2,n-2,0) = 0$.

Suppose that $\lim_{n\to\infty} |B(g_n)| = \infty$. Let $M =: max_{i\in B(g_n)} d_i < \infty$. Set $k(n) =: \frac{log(|B(g_n)|/2)}{log(M)}$. Let g_r be the restriction of g to B. For every $i \in B$, $|N^k(i)| \leq M^k \leq |B^C(g_n)|/2$. It follows that if we define $S_i =: \{j|d_{ij}(g_r) > k\}$, $|S_i| \geq |B(g_n)| - |B(g_n)|/2 = |B(g_n)|/2$. We deduce that for every $j \in S_i$, i and j must have a common neighbor in B^C. Suppose that this is not the case. If there is not a path that goes through B^C and that is shorter than k, agent i gains $b(1,d_j,0) - b(k(n),c_1,c_2)$ from linking to j and the gain for j is similar. Since $\lim_{n\to\infty} b(k(n),c_1,c_2) = 0$, the link is profitable. If $|B^C(g_n)|$ is bounded independently of n, then automatically every node in this set has an unbounded number of links to nodes in $B(g_n)$. Thus suppose $\lim_{n\to\infty} |B^C(g_n)| = \infty$. If there is a path between i and j that goes through B^C that is shorter than k but involves at least two distinct nodes in B^C, then there are at least $|B^C(g_n)|$ nodes who are closer to i than j and vice versa. Then the gain for i and j of forming a direct link is at least $b(1,d_j,0) - b(2,|B^C(g_n)|,c_2)$ and since

$\lim_{n\to\infty} b(2, |B^C(g_n)|, c_2)$, the link is profitable. This contradicts the stability of g. Therefore there must be some node in $l \in B^C$ that is linked to i and to every $j \in S_i$. Since $\lim_{n\to\infty} |S_i| = \infty$, the result follows. ■

3.A.11 Proof of Proposition 3.12

Proof: Assume to the contrary that some node i has a degree that is not bounded independently of n: $\lim_{n\to\infty} d_i(g_n) = \infty$. Let j be a neighbor of i. j's direct benefit of maintaining link ij is $b(1, 0, d_i(g_n) - 1)$ which goes to zero as n grows. The greatest possible indirect benefit of the link ij is achieved if i is the center of a star (in all other configurations indirect benefits are smaller). In this case $u_j^{CU}(g) - u_j^{CU}(g - ij) = b(1, 0, d_i(g_n) - 1) + (n-2)b(2, 1, n-2) - c$ but $\lim_{n\to\infty} c - b(1, 0, d_i(g_n) - 1) - (n-2)b(2, 1, n-2) = -c < 0$. Thus all nodes must have degrees bounded independently of n. The result about the diameter comes from arguments given previously.

Assumption (A2) in itself is not sufficient to ensure this weak degree heterogeneity. For example, the star can still be APS. To see this, note that the conditions are that the center does not want to break with the periphery (i.e., $b(1, 0, 0) \geq c$), the periphery does not want to break with the center (i.e., $\lim_{n\to\infty}((n-2)b(2, 1, n-2) + b(1, 0, n-2)) - c \geq 0$) and two peripheral nodes do not want to form a link to each other (i.e., $\lim_{n\to\infty}(b(1, 0, 1) - b(2, 1, n-2)) - c \leq 0$). This gives $b(1, 0, 1) \leq c \leq b(1, 0, 0)$ and $\lim_{n\to\infty} b(1, 0, n) \geq c$. If the term $b(1, 0, n)$ does not decline too rapidly, the star can be APS but the cost range $b(1, 0, 1) \leq c \leq b(1, 0, 0)$ is very small. There are basically two effects that destabilize the star. The peripheral nodes may want to break with the center if $b(1, 0, n)$ declines rapidly with the number of people being at the same distance to the sender. Moreover, agents' incentive to form direct links in the periphery is greater than in the Jackson-Wolinsky model since the information that two peripheral agents receive from each other in the star is shared with everyone else which gives them incentive to link directly. ■

Chapter 4

A modification of the connections model with negative externalities by overall connectivity

This chapter is based on work of my own and single-authored.

4.1 Introduction

Jackson and Wolinsky (1996) introduce in their seminal paper the so called *connections model*. It is an example for social communication between individuals where benefits and costs for each individual are determined by the direct and indirect connections among them. Each direct connection is costly and provides a certain benefit. Additionally, (discounted) benefits spill over from and to more distant partners to which only an indirect connection exists. Jackson and Wolinsky (1996) focus on identifying pairwise stable and strongly efficient networks. A network is said to be *pairwise stable* if no agent wants to sever a link, and if no two agents both want to add a link. A network is said to be *strongly efficient* if it maximizes the total utility of all agents. Jackson and Wolinsky (1996) point out that 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 Buechel and Hellmann (2012). For directed communication networks also see Dutta and Jackson

(2000) and for directed connections and hybrid connections models Bala and Goyal (2000).

There are numerous extensions of the connections model. Jackson and Rogers (2005) assume geographic costs of forming links. Players are grouped on the so called islands, and the costs of connecting to each other are low within an island and high across islands. Other variations of the connections model with geographic costs are investigated in, e.g., Johnson and Gilles (2000), Carayol and Roux (2005) and Carayol and Roux (2009). Jackson and Watts (2002), Watts (2001) as well as Watts (2002) embed the connections model in a dynamic framework. Although the connections model was modified intensively over time, the issue of *negative externalities* has been hardly considered in this framework.

Morrill (2011) introduces a *degree-based utility* and implements the idea that the more connections a direct neighbor has, the less utility is provided through a linkage. A simple example for this is the *co-author model* by Jackson and Wolinsky (1996): A researcher benefits a lot from a connection to a co-author, but the more projects he is already involved in, the less time is devoted to a single connection. Möhlmeier et al. (2016) build up on that and propose a *degree-distance-based* extension of the model by Morrill (2011) to capture the idea that increasing *busyness of neighbors* (and neighbors of neighbors) causes negative externalities. Their generalized *degree-distance-based* model subsumes the *degree-based* model by Morrill (2011) as well as the *distance-based model* by Bloch and Jackson (2007) as special cases. Hence, Möhlmeier et al. (2016) combine in the basic ideas from the connections model with the ones from the co-auther model and provide a quite general framework with both types of externalities, positive and negative.

Möhlmeier et al. (2017) integrate externalities due to connectivity associated with two types of effects: First, competition for the access to information and second, rivalrous use of information. Competition for the access to information can arise if an agent with many contacts must share his time between his contacts and thus has fewer/shorter opportunities to pass on information to each particular contact. The main idea is that the probability that every neighbor receives the information decreases with the number of contacts the sender has. In the second model there is no competition for the access to information but the use of information is rivalrous. It is assumed that when other agents receive the information before me, the harmful effect is greater than when they receive the information at the same time as myself.

Additional contributions on the role of externalities in social and economic networks are provided, e.g. by Goyal and Joshi (2006),Currarini (2007), Billand et al. (2013) and Hellmann (2013). Currarini (2007) 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 the spillovers. Besides that, Billand et al. (2013) provide existence results for a game with local spillovers, where the payoff function simultaneously satisfies the convexity and the strategic substitutes property. They use the notion of a pairwise stable network (Jackson and Wolinsky (1996)) and its refinement, called *pairwise equilibrium network* (Goyal and Joshi (2006)). A network is said to be a pairwise equilibrium network if there is a Nash equilibrium strategy profile which supports the network, and no two agents both want to add a link. Billand et al. (2013) characterize the architecture of a pairwise stable network and the architecture of a pairwise equilibrium network. Hellmann (2013) studies how externalities between links affect the existence and uniqueness of pairwise stable networks.

Jackson and Wolinsky (1996) mention already that "... one might have a decreasing value for each connection (direct or indirect) as the total amount of connectedness increases." (p. 53.). Taking this as a starting point, we introduce the *overall connectivity* model that incorporates the idea of adding negative externalities from increasing overall connectivity to the connections model. Additional links may generate positive externalities by shorter distances, but also negative externalities, since the total amount of connectedness increases.

Goyal and Joshi (2006) investigate two specific models which are closely related. The first model 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. The second one 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.

The overall connectivity model is not covered either by the playing the field game nor by the local spillovers game, and it identifies situations which cannot be distinguished in the other two frameworks. In the overall connectivity model, an agent's utility aggregates benefits from direct and indirect connections and weighs them by a factor considering the aggregate number of links of the other players in the network.

As an example, consider a stylized academic job market in which information about job opportunities and candidates is distributed within a network of scientists (the nodes). In this network, some scientists offer vacancies which they cannot fill internally, while others need to place team members, e.g. their job market candidates or untenured faculty. Establishing and maintaining a connection to a colleague (i.e., a link) is costly, but increases the probability of receiving valuable information. Information received from a neighboring node is passed on to all neighbors, but its value is depreciated. For a given node, an additional link induces a positive externality (if it reduces the distance to other nodes) and/or a negative externality (if it better connects remote nodes, i.e., if it gives theses scientists a relative advantage).

For the overall connectivity model, we provide results on pairwise stable, asymptotically pairwise stable and strongly efficient networks and compare them with the ones from the original connections model by Jackson and Wolinsky (1996). Our main findings are the following:

As in Jackson and Wolinsky (1996), we provide the conditions for pairwise stability of the empty network $g^{\emptyset}$, the complete network g^N, the star network g^s and the circle network g^c. All these structures are pairwise stable in the overall connectivity model, but usually for smaller costs compared to the ones from the connections model. This is due to the weighting factor from overall connectivity which reduces the benefit terms. We show that pairwise stable networks with homogeneous degree distribution, called regular networks, consist of at most one (non-empty) component in the overall connectivity model. Furthermore, we prove that a regular network is pairwise stable in the connections model for costs c if and only if it is pairwise stable in the overall connectivity model for the fraction $c' = \frac{1}{1+L(g_{-i})}c$. For large n we show that the star network g^s is asymptotically pairwise stable, while the empty network $g^{\emptyset}$, the complete network g^N as well as the circle network g^c are never asymptotically pairwise stable. The set of strongly efficient networks may differ from the architectures identified in Jackson and Wolinsky (1996). However, for general n, we are not able to fully characterize the set of strongly efficient architectures, but the star network g^s appears to be a very good candidate. In terms of aggregate utility, star network g^s strictly dominates the complete network g^N and the circle network g^c. The star network g^s provides (relatively) short distances with a small number of links and hence, (relatively) small negative externalities due to overall connectivity.

The rest of the paper is structured as follows. In Section 4.2 we recapitulate some preliminaries on networks, the connections model and the co-author model by Jackson and Wolinsky (1996) as well as some of the related extensions. Section 4.3 is concerned with our modification based on overall connectivity. We indicate the central differences to the existing modifications and provide results on pairwise stability, asymptotic pairwise stability and strong efficiency. In Section 4.4 we finish with some concluding remarks and present ideas for further research.

4.2 Notation and selected models with externalities

In this section we first present the preliminaries on networks (see, e.g., Jackson and Wolinsky (1996) and Jackson (2008)) and then briefly recapitulate the models which are points of departure for our work: the connections model and the co-author model introduced in Jackson and Wolinsky (1996), the playing the field game and the local spillovers game presented by Goyal and Joshi (2006), the model with degree-based utility functions by Morrill (2011) and the model with degree-distance-based utility functions by Möhlmeier et al. (2016).

4.2.1 Definitions

Let $N = \{1, 2, \ldots, n\}$ denote the set of nodes, often also called agents or players. A *network* g is a set of pairs $\{i, j\}$, denoted for convenience by ij, with $\{i, j\} \in N^2$, $i \neq j$[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.,

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

As a convention, the terms graphs and networks are used as synonyms in this framework.

Two simple structures are 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

[1] We do not allow for loo in this setting.

of all possible networks g on N is

$$G := \{g \mid g \subseteq g^N\}.$$

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

Let $N(g)$ ($n(g)$, respectively) denote the set (the number, respectively) of players with at least one link, i.e.,

$$N(g) = \{i \mid \exists j \text{ s.t. } ij \in g\}, \quad n(g) = |N(g)|.$$

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 $\{i_1i_2, i_2i_3, \ldots, i_{K-1}i_K\} \subseteq g$. A network g is *connected* if there exists 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 $ij \in g'$. Consequently, a network is connected if and only if it consists of a single component.

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 to 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 said to be *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 $u_j(g) > u_j(g + ij)$.

Additionally, we are interested in structures which are stable when n tends to be large. Möhlmeier et al. (2017) suggest a notion of asymptotic (with respect to network size) pairwise stability.

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

(i) 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 of network S, and

(ii) there exist fixed admissible parameters of the utility functions $(u_i)_{i=1}^n$ such that for all i, j, $i \neq j$

(a) $\lim_{n\to+\infty}(u_i(g_n) - u_i(g_n - ij)) \geq 0$ and

(b) if $\lim_{n\to+\infty}(u_i(g_n + ij) - u_i(g_n)) > 0 \Rightarrow \lim_{n\to+\infty}(u_j(g_n + ij) - u_j(g_n)) \leq 0$.

The set of admissible specifications (parameters) of the utility function for which the network is APS is the *asymptotic stability range*. In the overall connectivity model the asymptotic stability range is (c, δ), i.e., it is determined by the cost $c > 0$ and $0 < \delta < 1$.

APS is neither a refinement of PS, nor it is a weaker concept. A certain network structure can be PS for some fixed n but not APS. The main interest of the concept of APS is to reduce the parameter space since the parameter n disappears. The conditions for APS tend to be less involved than those for PS which may depend on n.

4.2.2 The connections model and the co-author model

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

$$u_i^{JW}(g) = \sum_{j\neq i}\delta^{t_{ij}} - \sum_{j:ij\in g} c = \sum_{j\neq i}\delta^{t_{ij}} - c\eta_i(g) \tag{4.1}$$

where $0 < \delta < 1$ denotes the undiscounted valuation of a connection, t_{ij} describes the number of links in the shortest path between i and j (with $t_{ij} = \infty$, if there is no path connecting i and j) and $c > 0$ are 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 is reduced by the costs of maintaining the direct connections.

Jackson and Wolinsky (1996) (Proposition 1) show that the complete network g^N, the empty network $g^\emptyset$ or the star network g^s can be uniquely SE (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) the star network g^s if $\delta - \delta^2 < c < \delta + \frac{(n-2)\delta^2}{2}$ and

(iii) the emtpy network $g^\emptyset$ if $\delta + \frac{(n-2)\delta^2}{2} < c$.

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

(i) A PS graph has at most one (non-empty) component.

(ii) For $c < \delta - \delta^2$, the unique PS network is the complete network g^N.

(iii) For $\delta - \delta^2 < c < \delta$, a star network g^s 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 the players are interpreted as researchers and a link represents a collaboration between two researchers. The amount of time each researcher spends on a collaborations is inversely related to the number of projects in which he is involved in. The utility function of player i in network g is given by

$$u_i^{co}(g) = \sum_{j:ij\in g} w_i(n_i, j, n_j) - c(n_i) \tag{4.2}$$

where $w_i(n_i, j, n_j)$ is the benefit of i derived from a link with j when i and j are involved in n_i and n_j projects, respectively, and $c(n_i)$ are the costs to i of maintaining n_i links.

Jackson and Wolinsky (1996) analyze the following specific utility function:

$$u_i^{co}(g) = \sum_{j:ij\in g}\left[\frac{1}{n_i} + \frac{1}{n_j} + \frac{1}{n_i n_j}\right] = 1 + \left(1 + \frac{1}{n_i}\right)\sum_{j:ij\in g}\frac{1}{n_j} \tag{4.3}$$

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

Although there are no direct costs of a connection, every new link decreases the strength of the existing links. Jackson and Wolinsky (1996) (Proposition 4) prove for the model (4.3) the following:

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

(ii) a PS network can be partitioned into fully intraconnected components, each of which has a different number of members.

4.2.3 Related frameworks to model externalities

As mentioned in the introduction, there are related frameworks which build up on the connections model as well as the co-author model and which model externalities in networks. In the following, we will present a selection of them in more detail.

Goyal and Joshi (2006) introduce two specific models of network formation. In the first one, called a *playing the field game*, an agent's aggregate payoff of an 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 the function

$$\pi_i^{pfg}(g) = \Phi(\eta_i(g), L(g_{-i})) \tag{4.4}$$

and its net payoff by

$$\Pi_i^{pfg}(g) = \Phi(\eta_i(g), L(g_{-i})) - \eta_i(g)c \tag{4.5}$$

where $\eta_i(g)$ is the degree of agent i and

$$L(g_{-i}) = \sum_{j \neq i} \eta_j(g_{-i}). \tag{4.6}$$

Note that g_{-i} is obtained by deleting i and all his links from g. It is assumed that for all $L(g_{-i})$, $\Phi(k, L(g_{-i}))$ is strictly increasing in own links k. Goyal and Joshi (2006) study two externality effects – across links of the same player and across links of different players (which are either positive or negative).

The second game investigated in Goyal and Joshi (2006) is called the *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

$$\pi_i^{lsg}(g) = \Psi_1(\eta_i(g)) + \sum_{j:ij\in g} \Psi_2(\eta_j(g)) + \sum_{j:ij\notin g} \Psi_3(\eta_j(g)). \tag{4.7}$$

Goyal and Joshi (2006) note that the marginal payoff to i from a link with j, $ij \notin g$, is given by

$$\pi_i^{lsg}(g+ij) - \pi_i^{lsg}(g) = \Psi_1(\eta_i(g)+1) - \Psi_1(\eta_i(g)) + \left[\Psi_2(\eta_j(g)+1) - \Psi_3(\eta_j(g))\right].$$

Hence, it depends only on the number of links of i and j and is independent of the number of links of $k \neq i, j$.

Morrill (2011) models situations in which adding links causes negative externalities. The 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

$$u_i^{deg}(g) = \sum_{ij\in g} \phi(\eta_j(g)) - c\eta_i(g). \tag{4.8}$$

Möhlmeier et al. (2016) introduce a *degree-distance-based* variation that additionally accounts for negative externalities by link addition of agents that are indirectly connected to the relevant player. The utility of agent i is given by

$$u_i^{deg-dis}(g) = \sum_{j\neq i} b(l_{ij}(g), \eta_j(g)) - c\eta_i(g) \tag{4.9}$$

where $b : \{1,\dots,n-1\}^2 \to \mathbb{R}^+$ is the benefit that an agent receives from a connection and $c > 0$ are the costs of one direct connection. It is assumed that $b(l_{ij}(g), k)$ is nonincreasing in degree k for all $l_{ij}(g)$ and nonincreasing in distance l for all $\eta_j(g)$. 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 $\eta_j \in \{0, 1, \dots, n-1\}$. In particular, $\widetilde{u}_i(g^{\emptyset}) = 0$ for every $i \in N$. This generalizes the degree-based model by Morrill (2011) since

$$\phi(\eta_j(g)) = b(1, \eta_j(g)), \text{ for all } \eta_j(g) \in \{1, \dots, n-1\}. \tag{4.10}$$

4.3 The overall connectivity model

We propose a modification of the connections model which takes into account the overall connectivity in a network. Similar to Jackson and Wolinsky (1996), an agent benefits from his direct and indirect connections, less distant connections are more valuable than more distant ones and direct connections are costly. Our modification additionally considers the aggregate number of links of the remaining agents in a network. The higher the overall connectivity of the other agents, the smaller is the benefit of an agent has from his own (direct and indirect) connections. The idea of implementing negative externalities from increasing overall connectivity can be motivated, for instance, by the academic job market example presented in the introduction. In our model, which we call *overall connectivity model*, the utility of agent i is given by

$$u_i^{oc}(g) = \sum_{j \neq i} \frac{1}{1 + L(g_{-i})} \delta^{t_{ij}} - \sum_{j:ij \in g} c = \frac{1}{1 + L(g_{-i})} \sum_{j \neq i} \delta^{t_{ij}} - c\eta_i(g) \tag{4.11}$$

where $0 < \delta < 1$, t_{ij} is the number of links in the shortest path between i and j, $c > 0$ are the costs of a direct connection and $L(g_{-i})$ is defined in (4.6).

To see that the overall connectivity model is not a special case of one of the existing models, consider a few simple examples. The following figures show networks which generate identical levels of utility for player 1 in at least one of the existing models, but different (greater) levels of utility for player 1 in the overall connectivity model.

Figure 4.1: The overall connectivity model versus the connections model, the co-author model and the degree-based utility

If my co-author and me are the only ones who work on a specific topic, the benefit from that collaboration should be greater than the benefit as if there was another "couple" working on that topic. Figure 4.1 indicates these two networks. While $u_1^{JW}(g) = u_1^{JW}(g') = \delta - c$, $u_1^{co}(g) = u_1^{co}(g') = w_1(1,2,1) - c(1)$ and $u_1^{deg}(g) = u_1^{deg}(g') = \phi(1) - c$, the overall connectivity results in $u_1^{oc}(g) = \delta - c > \frac{\delta}{3} - c = u_1^{oc}(g')$.

For Figure 4.2 let us compare the overall connectivity model with the playing the field game. We have $\Pi_1^{pfg}(g) = \Pi_1^{pfg}(g') = \Phi(1,4) - c$, but $u_1^{oc}(g) = \frac{1}{5}(\delta + 2\delta^2) - c >$

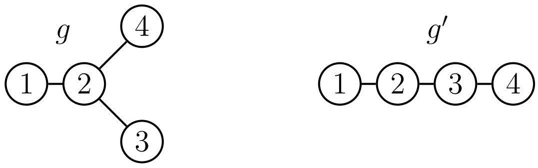


Figure 4.2: Overall connectivity model versus the playing the field game

$\frac{1}{5}(\delta + \delta^2 + \delta^3) - c = u_1^{oc}(g')$ for every $\delta \in (0,1)$. Here, although both models account for overall connectivity, the overall connectivity model additionally distinguishes between different patterns of indirect connections.

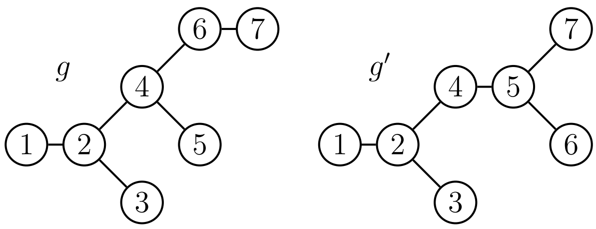


Figure 4.3: Overall connectivity model versus the local spillovers game

A comparison between the overall connectivity model and the local spillovers game is illustrated in Figure 4.3. We see that $\pi_1^{lsg}(g) = \pi_1^{lsg}(g') = \Psi_1(1) + \Psi_2(3) + 3\Psi_3(1) + \Psi_3(3) + \Psi_3(2)$, but $u_1^{oc}(g) = \frac{1}{10}(\delta + 2\delta^2 + 2\delta^3 + \delta^4) - c > \frac{1}{10}(\delta + 2\delta^2 + \delta^3 + 2\delta^4) - c = u_1^{oc}(g')$ for every $\delta \in (0,1)$. Again, while the local spillovers game treats equally the indirect connections of agents with the same degree, the overall connectivity model takes into account the exact length of the paths to all (indirectly connected) of agent 1. Note that, moreover, g and g' of Figure 4.3 are not distinguished also by the remaining models mentioned before (except by the connections model).

4.3.1 Pairwise stability

First, we examine pairwise stability of the typical network structures, namely the empty network $g^{\emptyset}$, the complete network g^N, the star network g^s and the circle network g^c. Furthermore, we are going to show that a PS regular network always consists of (at most) one non-empty component. Additionally, we relate the conditions for pairwise stability of regular networks in the overall connectivity model to the ones in the original connections model.

Proposition 4.1. *Let the utility be defined by (4.11). The empty network $g^\emptyset$ is PS if and only if $\delta \leq c$.*

Proof: Consider any two agents $i, j \in g^\emptyset$. Forming a link would for each of agents i, j result in $u_i^{oc}(g^\emptyset + ij) - u_i^{oc}(g^\emptyset) = u_j^{oc}(g^\emptyset + ij) - u_j^{oc}(g^\emptyset) = \delta - c$. Hence, if $\delta > c$, then $u_i^{oc}(g^\emptyset + ij) - u_i^{oc}(g^\emptyset) > 0$ and $u_j^{oc}(g^\emptyset + ij) - u_j^{oc}(g^\emptyset) > 0$ which implies that both players will profit from establishing a link, and therefore $g^\emptyset$ is not PS. If $\delta \leq c$, then $u_i^{oc}(g^\emptyset + ij) - u_i^{oc}(g^\emptyset) \leq 0$ and $u_j^{oc}(g^\emptyset + ij) - u_j^{oc}(g^\emptyset) \leq 0$ which means that $g^\emptyset$ is PS. ■

For high costs, $c \geq \delta$, no node wants to connect to another node so that the stability range for the empty network $g^\emptyset$ in the overall connectivity model is the same as in the connections model. This comes from the fact that the benefit term in the utility function from (4.11) reduces to the one from (4.1) when checking for pairwise stability of the empty network $g^\emptyset$, which means formally $L(g^\emptyset_{-i}) = 0 \Rightarrow u_i^{oc}(g^\emptyset) = u_i^{JW}(g^\emptyset)$ for every agent i.

Next, we analyze the pairwise stability of the complete network g^N.

Proposition 4.2. *Let the utility be defined by (4.11).*

(i) The complete network g^N with $n = 2$ is PS if and only if $\delta \geq c$.

(ii) The complete network g^N with $n \geq 3$ is PS if and only if the following conditions hold:

$$c \leq \frac{1}{4(n^2 - 3n + 3)} \leq \frac{1}{12} \quad \textit{and} \quad \delta_1 \leq \delta \leq \delta_2, \quad \textit{where} \tag{4.12}$$

$$\delta_1 = \frac{1 - \sqrt{1 - 4c(n^2 - 3n + 3)}}{2} > 0 \tag{4.13}$$

$$\delta_2 = \frac{1 + \sqrt{1 - 4c(n^2 - 3n + 3)}}{2} < 1. \tag{4.14}$$

Proof: (i) Let $n = 2$. Then $u_i^{oc}(g^N) - u_i^{oc}(g^N - ij) = u_j^{oc}(g^N) - u_j^{oc}(g^N - ij) = \delta - c$, and therefore g^N is PS if and only if $\delta \geq c$.

(ii) Let $n \geq 3$. Consider any two agents $i, j \in g^N$. Because of symmetry, it suffices to verify $u_i^{oc}(g^N) - u_i^{oc}(g^N - ij) \geq 0$. We have

$$u_i^{oc}(g^N) - u_i^{oc}(g^N - ij) = \frac{(n-1)\delta}{1+(n-1)(n-2)} - c(n-1) - \frac{(n-2)\delta + \delta^2}{1+(n-1)(n-2)} + c(n-2)$$

and therefore

$$u_i^{oc}(g^N) - u_i^{oc}(g^N - ij) = u_j^{oc}(g^N) - u_j^{oc}(g^N - ij) = \frac{\delta(1-\delta)}{n^2 - 3n + 3} - c \tag{4.15}$$

This term is nonnegative if and only if $\frac{\delta(1-\delta)}{n^2-3n+3} \geq c \Leftrightarrow \delta^2 - \delta + c(n^2 - 3n + 3) \leq 0$.

We therefore consider $W(\delta) = \delta^2 - \delta + c(n^2 - 3n + 3)$ with $n \geq 3$ as a parameter, and calculate for which δ the inequality $W(\delta) \leq 0$ holds.

Let $\Delta = 1 - 4c(n^2 - 3n + 3)$. Since $n^2 - 3n + 3 \geq 3$ for each $n \geq 3$, we have $\Delta \geq 0 \Leftrightarrow c \leq \frac{1}{4(n^2-3n+3)} \leq \frac{1}{12}$ which gives (4.12). Moreover, $W(\delta) \leq 0 \Leftrightarrow \Delta \geq 0$ and $\delta_1 \leq \delta \leq \delta_2$, where δ_1, δ_2 are given by (4.13) and (4.14), respectively. ■

Corollary 4.1. *Let the utility be defined by (4.11) and* $n \geq 3$.

(i) If $\delta(1-\delta) < 3c$*, then the complete network* g^N *is not PS (for arbitrary* $n \geq 3$*).*

(ii) If the costs of a direct connection are too high, namely if $c > \frac{1}{12}$*, then the complete network* g^N *is never PS (independently of* $\delta \in (0,1)$ *and* $n \geq 3$*).*

(iii) For every fixed $n \geq 3$*, there exist* δ_n *and sufficiently small* $c_n > 0$ *such that* g^N *is PS. However, the larger the number of agents in the network is, the smaller must be these required maximal costs* $c_n > 0$ *(which are drastically decreasing with* n*).*

Proof: (i) The complete network g^N is PS iff $\frac{\delta(1-\delta)}{n^2-3n+3} \geq c \Leftrightarrow n^2 - 3n + 3 \leq \frac{\delta(1-\delta)}{c}$. Since $n^2 - 3n + 3 \geq 3$ for each $n \geq 3$, if $\frac{\delta(1-\delta)}{c} < 3$, then it must be $n^2 - 3n + 3 > \frac{\delta(1-\delta)}{c}$. Consequently, the complete network g^N is never PS.

(ii) This comes immediately from condition (4.12).

(iii) Given $n \geq 3$, the "proper" $c > 0$ is determined from (4.12), and $\delta_1 \leq \delta \leq \delta_2$ from (4.13) and (4.14). Note that $\left(\frac{1}{4(n^2-3n+3)}\right)$ is decreasing in n and $\lim_{n\to+\infty} \frac{1}{4(n^2-3n+3)} = 0$. ■

Compared to the original connections model, note two additional points about the stability conditions for the complete network g^N. First, the stability region shrinks down and gets more tiny. In (4.11) g^N is PS iff $c < \frac{\delta-\delta^2}{n^2-3n+3}$ while in (4.1) g^N is PS iff $c < \delta - \delta^2$. Second, the multiplicative constant $\frac{1}{1+L(g_{-i})}$ from the functional form of the overall connectivity model directly appears in the pairwise stability condition: $c < \frac{\delta-\delta^2}{n^2-3n+3} = \frac{\delta-\delta^2}{1+(n-1)(n-2)} = \frac{1}{1+L(g_{-i})}(\delta - \delta^2)$. Hence, for the complete network g^N the stability region in (4.11) is equivalent to the one in (4.1) multiplied by the factor $\frac{1}{1+L(g_{-i})}$. As we will see later on, this equivalence for the pairwise stability conditions will hold true for arbitrary regular networks in the two frameworks.

Next, we determine the sufficient and necessary conditions for the star network g^s to be PS in the overall connectivity model.

Proposition 4.3. *Let the utility be defined by (4.11). The star network g^s with $n \geq 3$ is PS if and only if:*

$$\frac{\delta(1-\delta)}{1+2(n-2)} \leq c \leq \frac{\delta[1+(n-2)\delta]}{1+2(n-2)}. \tag{4.16}$$

Proof: Take the center i of the star network g^s and two arbitrary peripheral agents j, k, $j \neq k$, such that $ij \in g^s$, but $jk \notin g^s$. For pairwise stability the following conditions must hold:
(A) $u_i^{oc}(g^s) - u_i^{oc}(g^s - ij) \geq 0$ and
(B) $u_j^{oc}(g^s) - u_j^{oc}(g^s - ij) \geq 0$ and
(C) $u_j^{oc}(g^s + jk) - u_j^{oc}(g^s) \leq 0$.

Condition (A): $u_i^{oc}(g^s) - u_i^{oc}(g^s - ij) = (n-1)\delta - (n-1)c - (n-2)\delta + (n-2)c = \delta - c$.
Hence, $u_i^{oc}(g^s) - u_i^{oc}(g^s - ij) \geq 0 \;\Leftrightarrow\; \delta \geq c$.
Condition (B): $u_j^{oc}(g^s) - u_j^{oc}(g^s - ij) = \frac{\delta+(n-2)\delta^2}{1+2(n-2)} - c$. Hence, $u_j^{oc}(g^s) - u_j^{oc}(g^s - ij) \geq 0 \;\Leftrightarrow\; \frac{\delta[1+(n-2)\delta]}{1+2(n-2)} \geq c$. Note that if condition (B) is satisfied, then also condition (A) is satisfied, since $\delta < 1$ and therefore $c \leq \frac{\delta[1+(n-2)\delta]}{1+2(n-2)} < \delta$.
Condition (C): $u_j^{oc}(g^s + jk) - u_j^{oc}(g^s) = \frac{2\delta+(n-3)\delta^2}{1+2(n-2)} - 2c - \frac{\delta+(n-2)\delta^2}{1+2(n-2)} + c = \frac{\delta-\delta^2}{1+2(n-2)} - c$.
Hence, $u_j^{oc}(g^s + jk) - u_j^{oc}(g^s) \leq 0 \;\Leftrightarrow\; \frac{\delta(1-\delta)}{1+2(n-2)} \leq c$. ■

Overall, in the star network g^s costs must be high enough so that peripheral nodes do not want to connect with each other, but at the same time small enough that peripheral nodes want to stay linked to the center.

Corollary 4.2. *Let the utility be defined by (4.11) and $n \geq 3$.*

(i) *If* $\delta < c$*, then the star network* g^s *is never PS.*

(ii) *If* $\delta \geq \frac{\delta^2}{2} \geq c$*, then the star network* g^s *is PS for sufficiently large* n *(i.e., for* $n \geq \frac{3}{2} + \frac{\delta(1-\delta)}{2c}$*). In particular, if* $\delta \geq \frac{\delta^2}{2} \geq c$ *and* $\delta(1-\delta) \leq 3c$*, then the star network* g^s *is PS for every* $n \geq 3$*.*

(iii) *If* $\delta \geq c \geq \frac{\delta^2}{2}$ *and* $\delta(1+\delta) < 3c$*, then the star network* g^s *is never PS.*

(iv) *If* $\delta \geq c \geq \frac{\delta^2}{2}$ *and* $\delta(1+\delta) \geq 3c$*, then the star network* g^s *is PS for* $3 \leq n \leq 2 + \frac{\delta-c}{2c-\delta^2}$*.*

Proof: (i) If $\delta < c$, then also $\frac{\delta[1+(n-2)\delta]}{1+2(n-2)} < c$, and therefore condition (4.16) is not satisfied.

(ii)-(iv) Conditions (4.16) are written equivalently as follows:

$$\frac{\delta(1-\delta)}{1+2(n-2)} \leq c \;\Leftrightarrow\; n \geq \frac{3}{2} + \frac{\delta(1-\delta)}{2c} \tag{4.17}$$

$$c \leq \frac{\delta[1+(n-2)\delta]}{1+2(n-2)} \;\Leftrightarrow\; (n-2)(\delta^2-2c) \geq c-\delta \tag{4.18}$$

Let $\delta \geq \frac{\delta^2}{2} \geq c$. Then (4.18) holds for every $n \geq 3$, since $c-\delta \leq 0$ and $(\delta^2-2c) \geq 0$. In particular, if also $\delta(1-\delta) \leq 3c$, then $\frac{3}{2} + \frac{\delta(1-\delta)}{2c} \leq 3$, and therefore (4.17) holds for every $n \geq 3$.

Let $\delta \geq c \geq \frac{\delta^2}{2}$. Then (4.18) is equivalent to $n \leq 2 + \frac{\delta-c}{2c-\delta^2}$. If $\delta(1+\delta) < 3c$, then $\frac{\delta-c}{2c-\delta^2} < 1$, and hence (4.18) holds only for $n < 3$. If $\delta(1+\delta) \geq 3c$, then $\frac{\delta-c}{2c-\delta^2} \geq 1$, so (4.18) holds for $3 \leq n \leq 2 + \frac{\delta-c}{2c-\delta^2}$. Moreover, if $\delta(1+\delta) \geq 3c$, then $\frac{\delta(1-\delta)}{2c} \geq \frac{3}{2} - \frac{\delta^2}{c}$, so $\frac{3}{2} + \frac{\delta(1-\delta)}{2c} \geq 3 - \frac{\delta^2}{c}$, so (4.17) holds for all $n \geq 3$. ■

In the connections model, the star network g^s is PS under the condition $\delta - \delta^2 < c < \delta$. In comparison to that, the star network g^s is PS in the overall connectivity model whenever $\frac{\delta(1-\delta)}{1+2(n-2)} \leq c \leq \frac{\delta[1+(n-2)\delta]}{1+2(n-2)}$. We directly observe that this condition depends on the network size n and that the upper as well as the lower bound lie below the ones from the connections model. Intuitively, this makes sense again, since the functional form differs by $\frac{1}{1+L(g_{-i})}$ that reduces the benefit terms by overall connectivity. Hence, for given benefits δ, costs c can possibly be higher in the connections model to guarantee pairwise stability of the star network g^s. In contrast, for given benefits δ, there exist costs c (small enough) such that the star network g^s

is PS in the overall connectivity model, but not in the connections model.

Next, we will look at the pairwise stability of structures with more than one component. To start with an example, consider an arbitrary network g, existing of (multiple) disconnected completely connected components. As we will see, this type of structure cannot be PS in the overall connectivity model.

Example 4.1. *Let the utility be defined by (4.11). Take an arbitrary agent i_a in component a and an arbitrary agent i_b in component b. Without loss of generality assume that the network g consists only of two completely connected components and component a is of size n_a and component b is of size n_b. For pairwise stability, we have to check for the no link deletion and no link addition conditions:*

(i) No link deletion condition:

(a) If $n_a = 2$ and $n_b \geq 2$, it must hold that:
$u^{oc}_{i_a}(g) - u^{oc}_{i_a}(g - i_a j_a) \geq 0 \Leftrightarrow \frac{\delta}{1+n_b(n_b-1)} - c - 0 \geq 0 \Leftrightarrow \frac{\delta}{1+n_b(n_b-1)} \geq c$

(b) If $n_a \geq 3$ and $n_b \geq 2$, it must hold that:
$u^{oc}_{i_a}(g) - u^{oc}_{i_a}(g - i_a j_a) \geq 0$
$\Leftrightarrow \frac{(n_a-1)\delta}{1+(n_a-1)(n_a-2)+n_b(n_b-1)} - (n_a-1)c - \frac{(n_a-2)\delta+\delta^2}{1+(n_a-1)(n_a-2)+n_b(n_b-1)} + (n_a-2)c \geq 0$
$\Leftrightarrow \frac{\delta-\delta^2}{1+(n_a-1)(n_a-2)+n_b(n_b-1)} \geq c$

(ii) No link addition condition (no case differentiation necessary, following form is valid for all $n_a \geq 2$ and $n_b \geq 2$):
$u^{oc}_{i_a}(g + i_a i_b) - u^{oc}_{i_a}(g) \leq 0$
$\Leftrightarrow \frac{n_a\delta+(n_b-1)\delta^2}{1+(n_a-1)(n_a-2)+n_b(n_b-1)} - n_a c - \frac{(n_a-1)\delta}{1+(n_a-1)(n_a-2)+n_b(n_b-1)} + (n_a-1)c \leq 0$
$\Leftrightarrow \frac{\delta+(n_b-1)\delta^2}{1+(n_a-1)(n_a-2)+n_b(n_b-1)} \leq c$

Obviously, the no link deletion and no link addition conditions can never hold jointly. Therefore, any network g, consisting of disconnected completely connected components, cannot be PS.

This result is in line with the connections model. There, also any network g, consisting of disconnected completely connected components, cannot be PS.

Next, we provide additional results for networks with homogeneous degree distribution, called regular networks. A regular (non-empty) network g^{reg}_d is a network with equal degree d ($1 \leq d \leq n-1$) for every node $i \in N$. We show that a PS regular network has at most one (non-empty) component in the overall connectivity model.

Proposition 4.4. *Let the utility be defined by (4.11). A PS regular network has at most one (non-empty) component.*

Proof: The proof follows the same line of thoughts as in Jackson and Wolinsky (1996), p. 51. Suppose that g_d^{reg} is PS and has two or more (non-empty) components. Let Δ_i^{ij} denote the marginal utility which node i receives from a link to node j, keeping the rest of network g_d^{reg} fixed. So, if $ij \notin g_d^{reg}$ it follows that $\Delta_i^{ij} = u_i^{oc}(g_d^{reg} + ij) - u_i^{oc}(g_d^{reg})$ and if $ij \in g_d^{reg}$ we have $\Delta_i^{ij} = u_i^{oc}(g_d^{reg}) - u_i^{oc}(g_d^{reg} - ij)$. Consider $ij \in g_d^{reg}$, then $\Delta_i^{ij} \geq 0$. Consider a link kl which belongs to another component. Since node i is already in a component with node j, but k is not, it follows that $\Delta_k^{kj} > \Delta_i^{ij} \geq 0$. This holds true, since node k will also receive $\frac{\delta^2}{1+L(g_d^{reg}{}_{-k})}$ in value from the indirect connection to node i, which is not included in Δ_i^{ij}. For similar reasons, it follows that $\Delta_j^{jk} > \Delta_l^{lk} \geq 0$. This contradicts pairwise stability, since the link $jk \notin g_d^{reg}$. ■

Hence, in the connections model as well as the overall connectivity model, a PS regular network has at most one (non-empty) component.

After having discussed the conditions for pairwise stability of the empty network $g^{\emptyset}$ (regular network of degree 0) and the complete network g^N (regular network of degree $n-1$), we will additionally provide the conditions for the pairwise stability of the circle g^c (regular network of degree 2).

Proposition 4.5. *Let the utility be defined by (4.11). The circle g^c is PS if and only if the following conditions hold:*

(i) For n even:

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

and

$$c \geq \frac{1}{2n-3}\left(\delta + 2\sum_{k=2}^{\lfloor\frac{n}{4}\rfloor}\delta^k - 2\sum_{k=\lceil\frac{n}{4}\rceil+1}^{\frac{n}{2}-1}\delta^k - \delta^{\frac{n}{2}}\right)$$

(ii) For n uneven:

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

and

$$c \geq \frac{1}{2n-3}\left(\sum_{k=1}^{\lfloor\frac{n-1}{4}\rfloor}\delta^k + \sum_{k=2}^{\lceil\frac{n-1}{4}\rceil}\delta^k - \sum_{k=\lceil\frac{n-1}{4}\rceil+1}^{\frac{n-1}{2}}\delta^k - \sum_{k=\lfloor\frac{n-1}{4}\rfloor+2}^{\frac{n-1}{2}}\delta^k\right)$$

Proof: (i) For n even:

- No link deletion condition:
  $u_i^{oc}(g^c) - u_i^{oc}(g^c - ij) \geq 0$
  $\Leftrightarrow \frac{1}{2n-3}\left(2\sum_{k=1}^{\frac{n}{2}-1}\delta^k + \delta^{\frac{n}{2}} - \sum_{k=1}^{n-1}\delta^k\right) - c \geq 0$
  $\Leftrightarrow \frac{1}{2n-3}\left(\sum_{k=1}^{\frac{n}{2}-1}\delta^k - \sum_{k=\frac{n}{2}+1}^{n-1}\delta^k\right) \geq c$
- No link addition condition (n divisible by 4):
  $u_i^{oc}(g^c) - u_i^{oc}(g^c + il) \geq 0$
  $\Leftrightarrow \frac{1}{2n-3}\left(2\sum_{k=1}^{\frac{n}{2}-1}\delta^k + \delta^{\frac{n}{2}} - \left(3\delta + 4\sum_{k=2}^{\frac{n}{4}}\delta^k\right)\right) + c \geq 0$
  $\Leftrightarrow c \geq \frac{1}{2n-3}\left(\delta + 2\sum_{k=2}^{\frac{n}{4}}\delta^k - 2\sum_{k=\frac{n}{4}+1}^{\frac{n}{2}-1}\delta^k - \delta^{\frac{n}{2}}\right)$
  $\Leftrightarrow c \geq \frac{1}{2n-3}\left(\delta + 2\sum_{k=2}^{\lfloor\frac{n}{4}\rfloor}\delta^k - 2\sum_{k=\lceil\frac{n}{4}\rceil+1}^{\frac{n}{2}-1}\delta^k - \delta^{\frac{n}{2}}\right)$
  No link addition condition (n not divisible by 4):
  $u_i^{oc}(g^c) - u_i^{oc}(g^c + il) \geq 0$
  $\Leftrightarrow \frac{1}{2n-3}\left(2\sum_{k=1}^{\frac{n}{2}-1}\delta^k + \delta^{\frac{n}{2}} - \left(3\delta + 4\sum_{k=2}^{\lfloor\frac{n}{4}\rfloor}\delta^k + 2\delta^{\lceil\frac{n}{4}\rceil}\right)\right) + c \geq 0$
  $\Leftrightarrow c \geq \frac{1}{2n-3}\left(\delta + 2\sum_{k=2}^{\lfloor\frac{n}{4}\rfloor}\delta^k - 2\sum_{k=\lceil\frac{n}{4}\rceil+1}^{\frac{n}{2}-1}\delta^k - \delta^{\frac{n}{2}}\right)$

(ii) For n uneven:

- No link deletion condition:
  $u_i^{oc}(g^c) - u_i^{oc}(g^c - ij) \geq 0$
  $\Leftrightarrow \frac{1}{2n-3}\left(2\sum_{k=1}^{\frac{n-1}{2}}\delta^k - \sum_{k=1}^{n-1}\delta^k\right) - c \geq 0$
  $\Leftrightarrow \frac{1}{2n-3}\left(\sum_{k=1}^{\frac{n-1}{2}}\delta^k - \sum_{k=\frac{n-1}{2}+1}^{n-1}\delta^k\right) \geq c$
- No link addition condition ($n-1$ divisible by 4):
  $u_i^{oc}(g^c) - u_i^{oc}(g^c + il) \geq 0$
  $\Leftrightarrow \frac{1}{2n-3}\left(2\sum_{k=1}^{\frac{n-1}{2}}\delta^k - \left(3\delta + 4\sum_{k=2}^{\frac{n-1}{4}}\delta^k + \delta^{\frac{n-1}{4}+1}\right)\right) + c \geq 0$
  $\Leftrightarrow c \geq \frac{1}{2n-3}\left(\delta + 2\sum_{k=2}^{\frac{n-1}{4}}\delta^k - \delta^{\frac{n-1}{4}+1} - 2\sum_{k=\frac{n-1}{4}+2}^{\frac{n-1}{2}}\delta^k\right)$
  $\Leftrightarrow c \geq \frac{1}{2n-3}\left(\sum_{k=1}^{\lfloor\frac{n-1}{4}\rfloor}\delta^k + \sum_{k=2}^{\lceil\frac{n-1}{4}\rceil}\delta^k - \sum_{k=\lceil\frac{n-1}{4}\rceil+1}^{\frac{n-1}{2}}\delta^k - \sum_{k=\lfloor\frac{n-1}{4}\rfloor+2}^{\frac{n-1}{2}}\delta^k\right)$
  No link addition condition ($n-1$ not divisible by 4):

$$
\begin{aligned}
& u_i^{oc}(g^c) - u_i^{oc}(g^c + il) \geq 0 \\
\Leftrightarrow\ & \tfrac{1}{2n-3}\left(2\sum_{k=1}^{\frac{n-1}{2}}\delta^k - \left(3\delta + 4\sum_{k=2}^{\lfloor\frac{n-1}{4}\rfloor}\delta^k + 3\delta^{\lceil\frac{n-1}{4}\rceil}\right)\right) + c \geq 0 \\
\Leftrightarrow\ & c \geq \tfrac{1}{2n-3}\left(\delta + 2\sum_{k=2}^{\lfloor\frac{n-1}{4}\rfloor}\delta^k + \delta^{\lceil\frac{n-1}{4}\rceil} - 2\sum_{k=\lceil\frac{n-1}{4}\rceil+1}^{\frac{n-1}{2}}\delta^k\right) \\
\Leftrightarrow\ & c \geq \tfrac{1}{2n-3}\left(\sum_{k=1}^{\lfloor\frac{n-1}{4}\rfloor}\delta^k + \sum_{k=2}^{\lceil\frac{n-1}{4}\rceil}\delta^k - \sum_{k=\lceil\frac{n-1}{4}\rceil+1}^{\frac{n-1}{2}}\delta^k - \sum_{k=\lfloor\frac{n-1}{4}\rfloor+2}^{\frac{n-1}{2}}\delta^k\right) \quad \blacksquare
\end{aligned}
$$

While Jackson and Wolinsky (1996) do not compute pairwise stability conditions for the circle g^c, we will not provide a direct comparison here. However, as mentioned before while analyzing the pairwise stability conditions for the complete network g^N, we conjecture that the stability regions for regular networks in (4.11) are equivalent to the ones in (4.1) multiplied by the factor $\frac{1}{1+L(g_{-i})}$. The next proposition proves this result.

Proposition 4.6. *A regular (non-empty) network is PS in the connections model for costs $c = (1 + L(g_{-i}))c'$ if and only if it is PS in the overall connectivity model for costs c'.*

Proof: Suppose that g_d^{reg} that is PS in Jackson and Wolinsky (1996). Due to symmetry and pairwise stability, for any two agents i and j who are directly connected it must hold that $u_i^{JW}(g_d^{reg}) - u_i^{JW}(g_d^{reg} - ij) = u_j^{JW}(g_d^{reg}) - u_j^{JW}(g_d^{reg} - ij) \geq 0$.

Let $\Delta := u_j^{JW}(g_d^{reg}) - u_j^{JW}(g_d^{reg} - ij) + c$. With this, it follows that $u_j^{JW}(g_d^{reg}) - u_j^{JW}(g_d^{reg} - ij) \geq 0 \Leftrightarrow u_j^{JW}(g_d^{reg}) - u_j^{JW}(g_d^{reg} - ij) + c \geq c \Leftrightarrow \Delta \geq c \Leftrightarrow \frac{1}{1+L(g_{-i})}\Delta \geq \frac{1}{1+L(g_{-i})}c \Leftrightarrow \frac{1}{1+L(g_{-i})}(u_j^{JW}(g_d^{reg}) - u_j^{JW}(g_d^{reg} - ij) + c) \geq \frac{1}{1+L(g_{-i})}c \Leftrightarrow u_j^{oc}(g_d^{reg}) - u_j^{oc}(g_d^{reg} - ij) + c' \geq c' \Leftrightarrow u_j^{oc}(g_d^{reg}) - u_j^{oc}(g_d^{reg} - ij) \geq 0$.

Therefore, the no link deletion condition is satisfied for costs $c' = \frac{1}{1+L(g_{-i})}c$ in the overall connectivity model. The no link addition condition follows straightforward.

Since the above calculations are all equivalences, it holds true that a regular (non-empty) network is PS in the connections model for costs $c = (1 + L(g_{-i}))c'$ if and only if it is PS in the overall connectivity model for costs c'. $\blacksquare$

After having analyzed pairwise stability of some standard architectures and comparing the results with the ones from the connections model, we will consider asymptotic pairwise stability and strong efficiency of networks in the overall connectivity model in the next section.

4.3.2 Asymptotic pairwise stability

Asymptotic pairwise stability was introduced by Möhlmeier et al. (2017) and deals with the pairwise stability of networks when the number of nodes tends to be very large. Möhlmeier et al. (2017) show that structures which are PS, do not need to be APS.

The following proposition checks whether, respectively under which conditions, the standard architectures are APS in the overall connectivity model. As it turns out, the empty network $g^{\emptyset}$ is APS whenever it is PS, the star network g^s is APS under some specific condition and the complete network g^N as well as the circle network g^c are never APS.

Proposition 4.7. *Let the utility be defined by (4.11).*

(i) The empty network $g^{\emptyset}$ is APS whenever it is PS.

(ii) The star network g^s is APS whenever $c \leq \frac{\delta^2}{2}$.

(iii) The complete network g^N is never APS.

(iv) The circle network g^c is never APS.

Proof: (i) This is obvious since the condition for pairwise stability does not depend on n.

(ii) g^s is PS $\Leftrightarrow \frac{\delta(1-\delta)}{1+2(n-2)} \leq c \leq \frac{\delta(1+(n-2)\delta)}{1+2(n-2)}$. Hence, g^s is APS whenerver $c \leq \frac{\delta^2}{2}$.

(iii) g^N is PS $\Leftrightarrow \frac{\delta(1-\delta)}{n^2-3n+3} \geq c$. From this and due to the fact that $c > 0$ by assumption it directly follows that the complete network g^N is never APS.

(iv) For n even, the circle g^c is PS if the following two conditions hold:

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

and

$$c > \frac{1}{2n-3}\left(\delta + 2\sum_{k=2}^{\lfloor\frac{n}{4}\rfloor}\delta^k - 2\sum_{k=\lceil\frac{n}{4}\rceil+1}^{\frac{n}{2}-1}\delta^k - \delta^{\frac{n}{2}}\right).$$

Since $\lim\limits_{n\to\infty}\frac{1}{2n-3} = 0$ and for $n \to \infty$ the sum terms converge (because $\delta < 1$), the circle g^c is never APS. For n uneven, the result follows analogously. ■

In the following, let us provide some further intuition for the results.

The empty network $g^{\emptyset}$ will always be APS whenever it is PS since the condition for pairwise stability is independent of the network size n. To guarantee PS and APS as well, one has to make sure that no node wants to establish a connection to another node, which simply means that the costs must be at least as high as the potential benefit ($c \geq \delta$). Since there are no possible links to be deleted in the empty network $g^{\emptyset}$, this is the only condition which has to be satisfied.

For the star network g^s the situation is different. For pairwise stability one has to make sure that neither center nor the peripherals want to delete their links and the peripherals do not want to add a link between them. By looking at the asymptotic stability range, we see that the lower bound for the costs c goes to 0 and the upper bound converges to $\frac{\delta^2}{2}$. This upper bound guarantees that no peripheral node wants to delete its connection to the center. On the one hand, for increasing n, the direct benefit from the center starts to vanish, but on the other hand there are still benefits coming in from the peripheral nodes. As long as the sum of the benefits is high enough ($c \leq \frac{\delta^2}{2}$), the peripheral nodes maintain their links to the center while n is getting large. The positive impact of an additional peripheral node is δ^2 divided by the overall connectivity. By an additional peripheral node, the overall connectivity increases by 2 what provides even more intuition for the condition $c \leq \frac{\delta^2}{2}$. In Figure 4.4 the (asymptotic) pairwise stability regions are plotted. We directly see that there are parameter regions, in which the star network g^s is PS, but not APS, APS but not PS as well as PS and APS simultaneously.

Regarding the complete network g^N we only have an upper bound on the costs c to consider. As we see in Figure 4.5, the stability region quickly decreases and vanishes when n is getting large. The intuition for this result is that due to the high connectivity of the complete network g^N, a node does not want to maintain a direct connection to another node because it is able to reach the node at distance 2 for sure. With increasing n, the marginal benefit from a direct connection strictly decreases, converging to 0. Hence, the complete network g^N is never APS.

For the circle network g^c the situation is similar to the complete network g^N. The stability region steadily shrinks and finally vanishes when n is increasing. However, in Figure 4.6 we see that the speed of convergence is (relatively) slow. Additionally, we observe that stability for increasing network size n is only possible for high values of δ. Intuitively, this makes sense since then a node receives a high spillover and

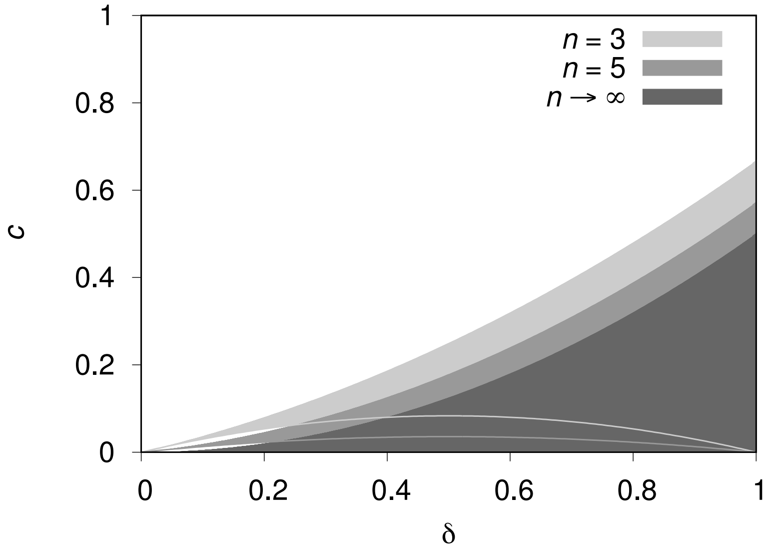


Figure 4.4: (Asymptotic) pairwise stability of the star network

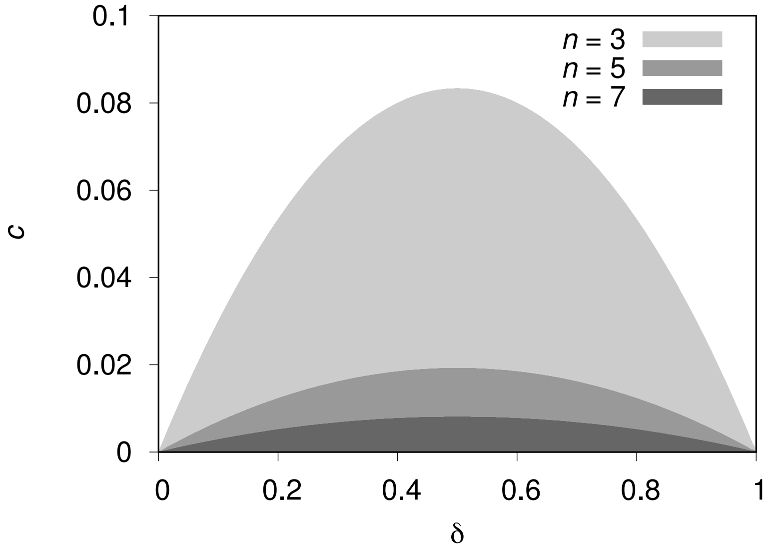


Figure 4.5: (Asymptotic) pairwise stability of the complete network

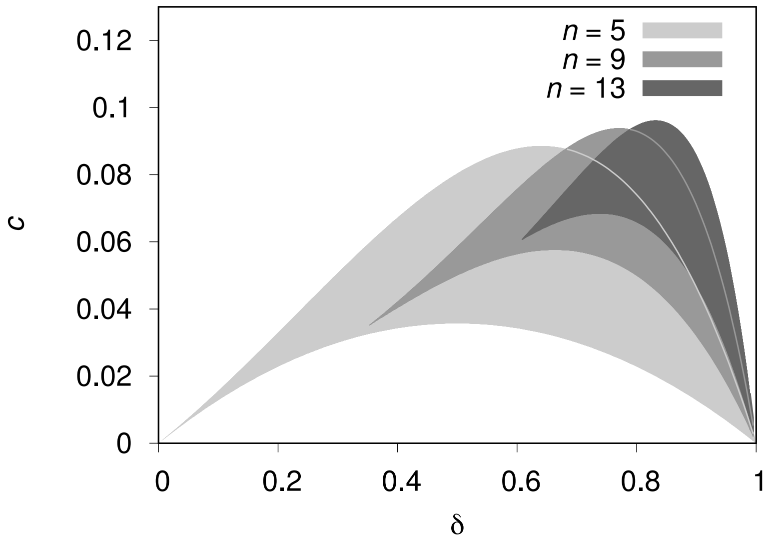


Figure 4.6: (Asymptotic) pairwise stability of the circle network

even does not want to add a link to the node(s) most far away in the circle network g^c. Furthermore, a node does not want to delete one of its existing links (as long as the costs c are low enough for increasing δ) because by maintaining one of its links it brings (roughly) half of the nodes closer to it.

4.3.3 Strong efficiency

Next, we focus on the analysis of strong efficiency in the overall connectivity model. To start, we look at strong efficiency for $n = 3$.

Example 4.2. *Let the utility be defined by (4.11) and $n = 3$. The values for the different possible architectures are the following (g^I denotes the network consisting of only one link):*

$u_1^{oc}(g^\emptyset) = u_2^{oc}(g^\emptyset) = u_3^{oc}(g^\emptyset) = 0$ *and hence* $\sum_{i=1}^{3} u_i^{oc}(g^\emptyset) = 0$.

$u_1^{oc}(g^I) = u_2^{oc}(g^I) = \delta - c$, $u_3^{oc}(g^I) = 0$ *and hence* $\sum_{i=1}^{3} u_i^{oc}(g^I) = 2\delta - 2c$.

$u_1^{oc}(g^s) = u_2^{oc}(g^s) = \frac{\delta+\delta^2}{3} - c$, $u_3^{oc}(g^s) = 2\delta - 2c$ *and hence* $\sum_{i=1}^{3} u_i^{oc}(g^s) = \frac{8}{3}\delta + \frac{2}{3}\delta^2 - 4c$.

$u_1^{oc}(g^N) = u_2(g^N) = u_3^{oc}(g^N) = \frac{2}{3}\delta - 2c$ *and hence* $\sum_{i=1}^{3} u_i^{oc}(g^N) = 2\delta - 6c$.

(i) The complete network g^N is never SE.

This is obvious, since the value of the complete network g^N is strictly dominated by the value of the star network g^s and the network g^I.

(ii) The empty network $g^{\emptyset}$ is SE whenever $c \geq \delta$.

$\sum_{i=1}^{3} u_i^{oc}(g^{\emptyset}) \geq \sum_{i=1}^{3} u_i^{oc}(g^I) \Leftrightarrow 0 \geq 2\delta - 2c \Leftrightarrow c \geq \delta$. *For $c \geq \delta$ it follows that* $\sum_{i=1}^{3} u_i^{oc}(g^s) = \frac{8}{3}\delta + \frac{2}{3}\delta^2 - 4c < 0 = \sum_{i=1}^{3} u_i^{oc}(g^{\emptyset})$ *and hence, the empty network $g^{\emptyset}$ is SE.*

(iii) The network g^I is SE whenever $\delta > c > \frac{1}{3}\delta + \frac{1}{3}\delta^2$.

For $c < \delta$ it follows that $\sum_{i=1}^{3} u_i^{oc}(g^I) > \sum_{i=1}^{3} u_i^{oc}(g^s) \Leftrightarrow 2\delta - 2c > \frac{8}{3}\delta + \frac{2}{3}\delta^2 - 4c \Leftrightarrow 2c > \frac{2}{3}\delta + \frac{2}{3}\delta^2 \Leftrightarrow c > \frac{1}{3}\delta + \frac{1}{3}\delta^2$. *Therefore, the network g^I is SE whenever $\delta > c > \frac{1}{3}\delta + \frac{1}{3}\delta^2$.*

(iv) The star network g^s is SE whenever $c \leq \frac{1}{3}\delta + \frac{1}{3}\delta^2$.

For $c < \delta$ it follows that $\sum_{i=1}^{3} u_i^{oc}(g^I) \leq \sum_{i=1}^{3} u_i^{oc}(g^s) \Leftrightarrow 2\delta - 2c \leq \frac{8}{3}\delta + \frac{2}{3}\delta^2 - 4c \Leftrightarrow 2c \leq \frac{2}{3}\delta + \frac{2}{3}\delta^2 \Leftrightarrow c \leq \frac{1}{3}\delta + \frac{1}{3}\delta^2$. *Therefore, the star network g^s is SE whenever $c \leq \frac{1}{3}\delta + \frac{1}{3}\delta^2$.*

Corollary 4.3. *Let the utility be defined by (4.11) and $n = 3$.*

(i) The empty network $g^{\emptyset}$ is SE if and only if it is PS.

(ii) If the star network g^s is PS, it is SE.

(iii) If the complete network g^N is PS, it cannot be SE.

In Jackson and Wolinsky (1996) only three network structures, the empty network $g^{\emptyset}$, the complete network g^N and the star network g^s turned out to be SE. In the overall connectivity model, we already observe for $n = 3$ differences with respect to this result. As shown before, we observe that the empty network $g^{\emptyset}$, the star network g^s and the network g^I consisting of only one link, can be SE. In contrast, the complete network g^N turns out to be never SE.

For $n = 3$, the complete network g^N is strictly dominated by the star network g^s. While it turns out to be relatively hard to obtain results on strong efficiency for general n in the overall connectivity model due to the negative externalities, we are at least able to compare the values of the star network g^s, the complete network g^N and the circle network g^c with each other. In the following, we show that the value of the star network g^s strictly dominates the value of the complete network g^N as well as the circle network g^c for general n.

Proposition 4.8. *Let the utility be defined by (4.11) and $n \geq 3$.*

(i) The value of the star network g^s is always higher than the one of the complete network g^N. Hence, the complete network g^N is never SE in the overall connectivity model.

(ii) The value of the star network g^s is always higher than the one of the circle network g^c. Hence, the circle network g^c is never SE in the overall connectivity model.

Proof: (i) The values of the star network g^s and the complete network g^N are given by

$\sum_{i=1}^{n} u_i^{oc}(g^s) = (n-1)(\delta - c) + (n-1)\left(\frac{\delta+(n-2)\delta^2}{1+2(n-2)} - c\right)$ and

$\sum_{i=1}^{n} u_i^{oc}(g^N) = n(n-1)\left(\frac{\delta}{1+(n-1)(n-2)} - c\right)$.

Comparing these values yields:

$\sum_{i=1}^{n} u_i^{oc}(g^s) > \sum_{i=1}^{n} u_i^{oc}(g^N)$

$(n-1)(\delta - c) + (n-1)\left(\frac{\delta+(n-2)\delta^2}{1+2(n-2)} - c\right) > n(n-1)\left(\frac{\delta}{1+(n-1)(n-2)} - c\right)$

$\Leftrightarrow c(n-1)(n-2) > \frac{n(n-1)\delta}{1+(n-1)(n-2)} - (n-1)\delta - \left(\frac{(n-1)(\delta+(n-2)\delta^2)}{1+2(n-2)}\right)$

$\Leftrightarrow c(n-2) > \frac{n\delta}{1+(n-1)(n-2)} - \delta - \left(\frac{\delta+(n-2)\delta^2}{1+2(n-2)}\right)$

$\Leftrightarrow \frac{c(n-2)}{\delta} > \frac{n}{n^2-3n+3} - 1 - \left(\frac{1+(n-2)\delta}{2n-3}\right)$

$\Leftrightarrow \frac{c(n-2)(2n-3)(n^2-3n+3)}{\delta} > n(2n-3) - (2n-3)(n^2-3n+3) - (1+(n-2)\delta)(n^2-3n+3)$

For $n \geq 3$ and $\delta > 0$ the expression on the left hand side is positive. Next, we consider the expression on the right hand side and check when it is 0 at most:

$n(2n-3) - (2n-3)(n^2-3n+3) - (1+(n-2)\delta)(n^2-3n+3) \leq 0$

$\Leftrightarrow n(2n-3) - (n^2-3n+3)(2n-3+1+(n-2)\delta) \leq 0$

This inequality is fulfilled if:

$n^2 - 3n + 3 \geq n$

$\Leftrightarrow n^2 - 4n + 3 \geq 0$

$\Leftrightarrow (n-3)(n-1) \geq 0$

$\Leftrightarrow n \geq 3$

Hence, $\frac{c(n-2)(2n-3)(n^2-3n+3)}{\delta} \geq n(2n-3) - (2n-3)(n^2-3n+3) - (1+(n-2)\delta)(n^2-3n+3)$ holds true for every $n \geq 3$ and $\delta > 0$. Therefore, the value of the star network g^s strictly dominates the one of the complete network g^N.

(ii) The values of the star network g^s and the circle network g^c are given by
$\sum_{i=1}^{n} u_i^{oc}(g^s) = (n-1)(\delta - c) + (n-1)\left(\frac{\delta+(n-2)\delta^2}{1+2(n-2)} - c\right)$ and
$\sum_{i=1}^{n} u_i^{oc}(g^c) = \frac{2n}{2n-3}\sum_{k=1}^{\frac{n-1}{2}} \delta^k - 2nc$ (for n uneven).
Comparing these values yields:
$\sum_{i=1}^{n} u_i^{oc}(g^s) > \sum_{i=1}^{n} u_i^{oc}(g^c)$
$\Leftrightarrow (n-1)(\delta - c) + (n-1)\left(\frac{\delta+(n-2)\delta^2}{1+2(n-2)} - c\right) > \frac{2n}{2n-3}\sum_{k=1}^{\frac{n-1}{2}} \delta^k - 2nc$
$\Leftrightarrow (n-1)\delta + \frac{(n-1)(\delta+(n-2)\delta^2)}{2n-3} + 2c > \frac{2n}{2n-3}\sum_{k=1}^{\frac{n-1}{2}} \delta^k$
$\Leftrightarrow (2n-3)(n-1)\delta + (n-1)(\delta + (n-2)\delta^2) + 2c(2n-3) > 2n\sum_{k=1}^{\frac{n-1}{2}} \delta^k$
Since $2c(2n-3) > 0$ for $n \geq 3$ and $\delta > 0$, let us ignore it and compare the benefit terms on both sides with each other:
$(2n-3)(n-1)\delta + (n-1)(\delta + (n-2)\delta^2) \geq 2n\sum_{k=1}^{\frac{n-1}{2}} \delta^k$
$\Leftrightarrow (n-1)(\delta + (n-2)\delta^2 + (2n-3)\delta) \geq 2n\sum_{k=1}^{\frac{n-1}{2}} \delta^k$
Next, we analyze the number and size of the benefit terms on both sides. On the left hand side, we have $(n-1)(3n-4)$ elements that are discounted by at most distance 2. On the right hand side, we have $(n-1)n$ elements which are (partly) additionally discounted due to higher distances. Obviously, the term on the left hand side is always larger than the one on the right hand side.
We omit the calculations for the circle network g^c with n even since they are similar and lead to the same conclusions. Consequently, the value of the star network g^s strictly dominates the one of the circle network g^c. ■

For $n \geq 3$, the value of the star network g^s is always higher than the one of the complete network g^N as well as the circle network g^c. Regarding the circle network g^c, this result is in line with Jackson and Wolinsky (1996) where this structure cannot be SE, either (for $n \geq 4$, since for $n = 3$, the circle network g^c is structurally identical to the complete network g^N). However, the result for the complete network g^N is in sharp contrast to Jackson and Wolinsky (1996) where it is (uniquely) SE whenever $c < \delta - \delta^2$. In the overall connectivity model, this cannot be the case, since the complete network g^N is overconnected and always strictly dominated by the star network g^s.

Why does the star network g^s perform so well in the overall connectivity model? Intuitively, there are two reasons for this. First, it offers relatively short distances. The center reaches every peripheral node by distance 1 and the peripheral nodes every other node by distance 2 at most. Hence, the discount factor for the benefit terms in the aggreagte utility is at most 2. Second, the star network g^s is minimally connected. From graph theory we know that a graph is minimally connected if and only if it is a tree. Any connected graph with n nodes and $n-1$ links is a tree. So, the star network g^s is a tree and hence, minimally connected. Due to this, the negative externalities from overall connectivity stay in total (relatively) small. Nevertheless, the impact of the negative externality by overall connectivity is extremely different for the two different types of nodes, namely the center and the peripheral nodes. For the center, the factor $\frac{1}{1+L(g_{-i})}$ is simply 1 and hence, the benefit terms in the utility are not reduced at all by overall connectivity. For the peripheral nodes, the factor $\frac{1}{1+L(g_{-i})}$ in the utility function is equal to $\frac{1}{1+2(n-2)}$ and hence, their benefit terms are quite heavily reduced in comparison to the center. So, overall, the main part of the value of the star network g^s is contributed by the center. In contrast, in the other structures we looked at, such as the circle network g^c and the complete network g^N, the benefits of every node are (negatively) affected by overall connectivity. Due to the regularity of both structures, the induced weight factor $\frac{1}{1+L(g_{-i})}$ for the benefit terms is the same for all nodes in the relevant network. Apparently, some structural properties like an uneven degree distribution, connectedness, a small number of nodes and short distances play a central role for the total value of a network. The star network g^s combines all of these structural properties simultaneously and therefore performs so well.

However, we shall point out again that we did not prove that for general n the star network g^s is SE and if so, under which conditions. We only showed that it dominates some other (connected) structures, like the circle network g^c and the complete network g^N. Furthermore, we explained why we believe that the star network g^s is such a good candidate for being SE in the overall connectivity model. Unfortunately, compared to Jackson and Wolinsky (1996) a general proof appears to be much more complicated due to the negative externalities by link formation on other nodes. Additionally, there are usually also positive externalities by link formation on other nodes and then the question is always, which effect overweighs in terms of total value of the network. In Jackson and Wolinsky (1996), the proof on SE networks is quite straightforward because the externalities by link formation on other nodes are purely positive and thus the calculations are (much) easier. We tried to apply a similar argument as in the proof of Proposition 1(ii) in Jackson and Wolinsky (1996), but general statements about the value are very hard to achieve due the individual weights $\frac{1}{1+L(g_{-i})}$ in the utility functions that are strongly dependent on the underlying network structure.

4.4 Concluding remarks and further research

Based on the connections model by Jackson and Wolinsky (1996) we introduce a modification that accounts for negative externalities by overall connectivity. The functional form of the utilities in the overall connectivity model is very closely related to the ones in the connections model. We implement the externalities by additionally weighting the benefit terms of the connections model by a factor depending on the overall connectivity. Then, we discuss pairwise stable, asymptotically pairwise stable and strongly efficient networks in this framework. By comparing the two models, we observe some important similarities and differences between the connections model (with purely positive link externalities) and the overall connectivity model with negative externalities by overall connectivity.

As we have seen in the section on pairwise stability, the empty network $g^{\emptyset}$, the complete network g^N and the star network g^s can be pairwise stable. However, compared to the connections model, the stability regions become more tiny, since the benefit terms are (partially) reduced by overall connectivity. Similarly to the connections model, a pairwise stable regular network always consists of at most one (non-empty) component in the overall connectivity model. Furthermore, we identify the conditions for pairwise stability of the circle network g^c and show that all regular networks are PS in the connections model if and only if they are pairwise stable in the overall connectivity model for a specific fraction of the costs.

In the section on asymptotic pairwise stability we provide conclusions regarding the pairwise stability of networks when the number of nodes becomes large. We find out that the empty network $g^{\emptyset}$ and star network g^s can be asymptotically pairwise stable, while the complete network g^N and the circle network g^c cannot be asymptotically pairwise stable in the overall connectivity model.

Analyzing strong efficiency indicates already for $n = 3$ some central differences to the connections model. First, a disconnected structure can be strongly efficient and second, the complete network g^N cannot be efficient in the overall connectivity model. For $n \geq 3$ we show that the value of the star network g^s strictly dominates the one of the complete network g^N as well as the the circle network g^c. Unfortunately, we are not able to provide the precise conditions to guarantee strong efficiency of the star network g^s. Due to tractability, we leave this and a full characterization of the class of strongly efficient networks for further research.

Nevertheless, what is striking in the overall connectivity model is the fact that the star network g^s performs quite well in all areas we analyzed. The usual tension between pairwise stability and strong efficiency seems to be reduced here. However, to become more specific on this conjecture, we need to clarify additional points. Thus, for further research we suggest to identify conditions under which the star network g^s is (uniquely) strongly efficient. As described before, it appears to be a very good candidate for a strongly efficient network in the overall connectivity model. The star network g^s is strongly efficient (under some conditions) in the connections model with purely positive externalitites. In a situation with negative externalities due to overall connectivity it appears to be an even more favorable architecture, since it combines short distances with a small number of links. As we have seen for $n = 3$, even a disconnected structure can be strongly efficient in the overall connectivity model. Hence, we need to find out whether this may also happen for larger n and under which conditions a strongly efficient network is connected. An additional hypothesis we have is that (connected) regular networks are always dominated by the star network g^s. We looked at the two extreme cases, the circle network g^c with degree 2 and the complete network g^N with degree $n - 1$ of every node, and eventually all regular networks with in-between degree, are dominated by the star network g^s as well.

Overall, a general discussion of strong efficiency in models with both, positive and negative externalities, appears to be a fruitful area for further research and would provide a valuable contribution to the existing literature.

Chapter 5

Stability of customer relationship networks

This chapter is based on a joint work with Claus-Jochen Haake and Sonja Recker, both from Paderborn University, Faculty of Business Administration and Economics.

5.1 Introduction

In many real-world economic situations, a company has to invest in customer relationships to sell their products, e.g. in form of organizing access to their distribution channels. As examples one may think of a free product support, maintenance of user accounts or royalties that have to be paid for each customer. Regarding competitors and their products in the market, such a link formation problem is not a bare optimization exercise, as customers' demands play an important role for the incentives to form links. An interesting question is, which link structure evolves under specific market conditions. To answer this question, we investigate a duopoly, which is on the one hand characterized by Cournot competition and horizontal product differentiation and on the other hand by the network structure that is built up by the firms to serve customers' demands. Beyond the scope of the current model, the own optimal positioning strategy might also be depending on additional factors, such as the mode of competition, the connectedness of other market participants, the market size or the degree of substitutability of the own product compared to those of other competitors.

More precisely, in this article we focus on two firms' decisions to form links to their customers, explicitly taking link costs and the degree of substitutability into account. That means, the firms make strategic choices on which customer relationships they are willing to establish at given link formation costs and to what degree customers view their products as substitutes. One would expect that if customers consider the two firms' products as complements, then they shall be linked to both firms. However, if the firms' products are considered to be substitutes, customers will rather split and only trade with one firm. Our particular interest is in the network structure that evolves under given link costs and degree of product differentiation. For this, we analyze stable networks, which we define in two ways: *Local stability* requires that neither firm benefits from rendering the link status of a single customer, whereas *Nash stability* assures this for multiple customers. In the analysis we first concentrate on the general case with n customers. As the number of potential changes of the network structure significantly increases with the number of customers, we put emphasis on specific prominent network architectures and determine as well as depict the regions of local stability for arbitrary n. In the limit case (Proposition 5.2) it turns out that already a setting with few customers is a good proxy. This implies that much of the market structure and the forces that drive the results can already be understood in a model with few customers. Additionally, for networks with an arbitrary number of customers we establish the existence of locally stable networks for complementary or independent products (Proposition 5.3). For these product types either the empty, the complete or both networks are locally stable. Moreover, we link local and Nash stability for selected networks (Proposition 5.4) and demonstrate for which degrees of substitutability the model is sound enough to produce valid results (Propositions 5.1 and 5.5). Motivated by the fact that stability regions quickly converge, we devote a whole section on the scenario with three customers. The main result is a complete characterization of the set of stable networks, identifying for each possible network the combinations of link costs and degrees of substitutability under which it is stable. We display and compare such regions of stability for the different networks and identify adjacency relations. It turns out that in the case of complementary products, no asymmetric networks prevail and that the firms are generally willing to accept higher costs of customer acquisition as in the case of substitutes (Proposition 5.6) and that existence of locally stable networks is guaranteed for any degree of substitutability (Proposition 5.7).

Our work is linked to research in industrial organization, network and matching theory. The model of a differentiated duopoly we use is most closely related to Häckner (2000), which is based on the seminal work by Singh and Vives (1984). In our article we basically adapt this framework, adding a customer relationship network and putting emphasis on the impact of link costs and product differentiation on its stability. A different strand of related literature deals with the strategic formation of networks for research and development collaborations. Departing from the model of Goyal and Joshi (2003), it has been studied, e.g., by Goyal and Moraga-González (2001), by Dawid and Hellmann (2014) or by Dawid and Hellmann (2016). In these articles a link describes a research cooperation and the network is used to model the externalities that are induced by transferring knowledge through such a partnership. Forming a link reduces the production costs and, hence, endogenously introduces asymmetries. A different type of network formation is studied in Billand et al. (2016). They combine oligopolistic multimarket competition with a network structure among firms, in which links represent spying activities between firms. The article analyzes equilibria of the network formation game, in which payoffs are influenced by the impact of the network structure on product qualities. In contrast to these network models, we interpret links as customer relationships that purely enable trade and which directly impact equilibrium prices and equilibrium quantities on the market. Likewise, Kranton and Minehart (2000a, 2001, 2000b) and Corominas-Bosch (2004) interpret a link in this way. To model trade in buyer-seller networks Kranton and Minehart (2000a) focus on the analysis of competitive prices whereas Kranton and Minehart (2001) use an ascending bid auction and Corominas-Bosch (2004) assumes non-cooperative bilateral bargaining. For vertically integrated firms Kranton and Minehart (2000b) analyze equilibrium industrial structures including a model of link formation. Also Wang and Watts (2006) analyze the formation of links between buyers and sellers for quality differentiated products. Apart from similarities in the research question, our work especially differs with respect to the competition framework. We assume oligopolistic quantity competition between the sellers and thus, drop the assumption of unit-supply to investigate perfectly divisible goods. A more closely related contribution on oligopolistic competition and an underlying network structure is Bimpikis et al. (2016). They directly investigate quantity competition together with a network structure between m markets and n firms. The main assumption connecting separate markets is a convex, i.e., quadratic, cost function while at the same time products are assumed to be perfectly substitutable and the demand structure is linear. In contrast, we directly link different markets by the degree of product differentiation between two firms using the demand function and assuming that a firm charges the same price to all its customers. Moreover, our cost structure is supposed to be linear. In recent years, supply-chain (networks) have been investigated in matching theory with contracts, e.g., in Ostrovsky (2008), Westkamp (2010) and Hatfield and Kominers (2012). A common feature in all these matching models is that preferences over the set of available contracts (trade relations) are assumed to be exogenous and do not result from strategic interaction. Although various single aspects were studied in the literature, our work newly brings together the features of competition under product differentiation of divisible products and strategic formation and stability of customer relationship networks. In the latter, decisions depend on the competition and the degree of product substitutability in the market.

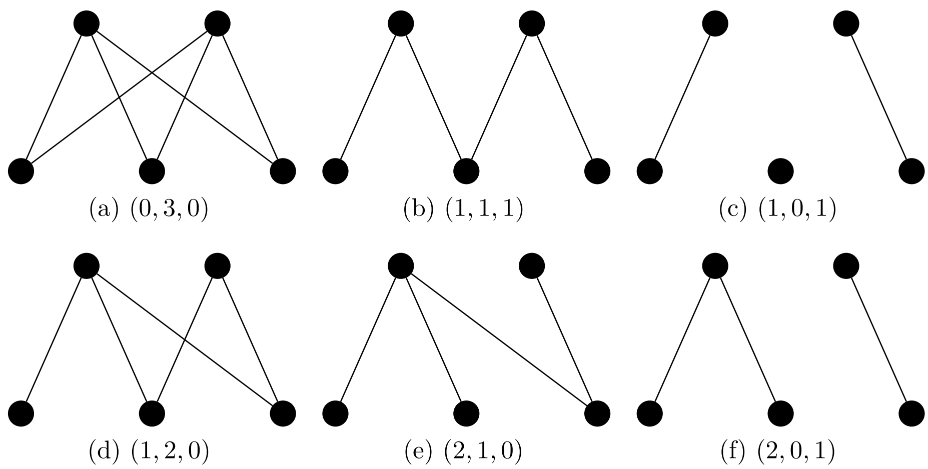
Our article is structured as follows: Section 5.2 presents the model with product differentiation for two firms and discusses (Cournot-Nash) equilibria for an arbitrary number and network of customers as well as the used notions of stability (local and Nash stability). The analysis for an arbitrary number of customers is in Section 5.3. Here we discuss the shape of the stability regions for specific architectures when n is getting large. In Section 5.4 we focus on the case of three customers and completely characterize stable network architectures. Section 5.5 concludes. All proofs are relegated to the Appendix.

5.2 The model

This section discusses the essential ingredients of our model such as the network architecture, competition and resulting equilibria as well as the notions of network stability. There are two firms and $n > 1$ customers in the market. The two firms offer differentiated products to the customers. We assume that prior to sales a customer has to be (costly) linked to the firm he wants to purchase from. Establishing links are in the firms' responsibilities and therefore costs are taken by them.

5.2.1 Networks of customer relationships

When two firms form links to customers there are three different types of customers that have at least one link to a firm. Customers that are only linked to firm 1 or only linked to firm 2 are termed *exclusive customers* (of firm 1 or 2, respectively), whereas customers who are linked to both firms are called *joint customers*. With n customers, we describe a *customer relationship network architecture* (or simply *network*) by the numbers n_1^e of exclusive customers of firm 1, n^j of joint customers and n_2^e of exclusive customers of firm 2, i.e., by a triplet $(n_1^e, n^j, n_2^e) \in \mathbb{N}_+^3$ with $n_1^e + n^j + n_2^e \leq n$. We denote the set of all networks with n customers by $\mathcal{N}_n = \left\{(n_1^e, n^j, n_2^e) \in \mathbb{N}_+^3 | n_1^e + n^j + n_2^e \leq n\right\}$. Examples for networks of customer relationships with $n = 3$ customers are given in Figure 5.1 (the two firms are represented by the two nodes on the top, customers are represented by nodes on the bottom). Note that to describe a network we always write the number of joint customers in the middle and the exclusive customers of the two firms left and right. This notation is intuitively related to the illustrations in Figure 5.1.



(a) $(0,3,0)$ (b) $(1,1,1)$ (c) $(1,0,1)$

(d) $(1,2,0)$ (e) $(2,1,0)$ (f) $(2,0,1)$

Figure 5.1: Examples of customer relationship networks

From a market perspective, we are interested in the structure of the network rather than which particular customer is linked to which firm, as customers as well as firms are assumed to be identical. Therefore, one natural distinction of networks is one between *symmetric* networks, in which $n_1^e = n_2^e$ holds, and *asymmetric* networks with $n_1^e \neq n_2^e$. The networks in Figures 5.1a, 5.1b and 5.1c are symmetric, in contrast to asymmetric networks in Figures 5.1d, 5.1e and 5.1f. Figure 5.1a shows a network of customer relationships where all customers are linked to both firms. This implies that each firm has exactly three links and all customers are joint customers. In contrast, in Figure 5.1f all customers are linked to either firm 1 or to firm 2 and in Figure 5.1c there is a customer without any link. Figure 5.1b illustrates a symmetric network where both firms have exactly one exclusive customer, and, in addition, there is one joint customer linked to both firms. The networks in Figure 5.1d and Figure 5.1e are asymmetric and only one firm has some exclusive customers. However, a network with no joint customers may also be asymmetric, as in Figure 5.1f. These networks of customer relationships are just some examples for $n = 3$ which are analyzed in full detail in Section 5.4 after the general model has been introduced and investigated.

5.2.2 Customers' demands

To describe the customers' preferences we use a standard utility function on product differentiation as in Singh and Vives (1984) and Häckner (2000). For this, we need to specify a customer's utility function, depending on whether he is an exclusive or a joint customer. Recall that exclusive customers can only derive utility from consuming the corresponding firm's product, whereas both products enter a joint customer's utility function. Phrased differently, we understand a link as granting access to a customer to purchase at that firm. In fact, creating (or deleting) a link to a customer will alter his utility function and hence his demand for products.

Type $e1$ customers are exclusive customers from firm 1 and consume an amount of $\mathbf{q_1^e} = (q_1^e)$, type j customers are joint customers and consume $\mathbf{q^j} = (q_1^j, q_2^j)$, whereas type $e2$ customers are exclusively consuming $\mathbf{q_2^e} = (q_2^e)$ from firm 2. So, q_i^t denotes the amount of firm i's product that a customer of type t consumes ($t \in \{e, j\}$). The vector $\mathbf{q} = (\mathbf{q_1^e}, \mathbf{q^j}, \mathbf{q_2^e}) = (q_1^e, q_1^j, q_2^j, q_2^e)$ collects all such quantities. Customers' utilities are as follows:[1]

$$u_i^e(\mathbf{q_i^e}, I, \gamma) = aq_i^e - \frac{1}{2}\left(q_i^e\right)^2 + I, \qquad \text{(type } ei\text{)}$$

$$u^j(\mathbf{q^j}, I, \gamma) = aq_1^j + aq_2^j - \frac{1}{2}\left(\left(q_1^j\right)^2 + \left(q_2^j\right)^2 + 2\gamma q_1^j q_2^j\right) + I, \qquad \text{(type } j\text{)}$$

[1]Singh and Vives (1984) postulate the utility function for joint customers, as all customers have access to trade with both firms in their model. Besides that, we newly consider exclusive customers.

for $i \in \{1,2\}$, where $a > 0$ can be interpreted as the products' common quality level and I as the consumption (expenditures) for other goods.[2] The parameter $\gamma \in [-1,1]$ reflects the *degree of horizontal product differentiation* (or the degree of substitutability/complementarity) and only affects the utility of joint customers. For $\gamma = 1$, the utility function u^j effectively depends on the sum $q_1^j + q_2^j$, showing that a joint customer views the products as perfect substitutes. On the contrary, at $\gamma = -1$ he considers two products as complements, as there the squared difference of the quantities given by $(q_1^j - q_2^j)^2$ enters negatively in the utility function u^j. Finally, at $\gamma = 0$ the joint customers' utility function is additively separable and thus, the two firms can be considered as monopolists for their products. Note that the utility function u^j of a joint customer reduces to the utility function of an exclusive customer, u_i^e, when the according quantity q_{3-i}^j is zero. Solving customers' utility maximization problems[3] at given firms' prices p_1 and p_2 in a network of customer relationships $(n_1^e, n^j, n_2^e) \in \mathcal{N}_n$, the first order conditions of firm i, $i = 1,2$ read

$$a - q_i^e - p_i = 0 \quad \text{for } n_i^e \text{ customers of type } ei, \tag{5.1}$$

$$a - q_i^j - \gamma q_{3-i}^j - p_i = 0 \quad \text{for } n^j \text{ customers of type } j. \tag{5.2}$$

For $n_1^e, n^j, n_2^e > 0$ the above system of equations is equivalent to

$$p_i = a - \frac{n_i^e q_i^e + n^j q_i^j + n^j \gamma q_{3-i}^j}{n_i^e + n^j} \quad \text{and} \quad q_i^e = q_i^j + \gamma q_{3-i}^j \quad \text{for } i = 1,2. \tag{5.3}$$

The inverse demand functions $p_i(\mathbf{q})$ in the first part of (5.3) will next serve as primitive for the Cournot competition among firms. Due to differentiated products they may be charged at different prices, but each price depends on the demands for both firms' products. The second part of (5.3) reflects the relation between the demands of exclusive and joint customers from the firms' perspective, given the customers' optimizing behavior.

[2] We assume that customers' income is high enough, so that due to quasilinearity of the utility function, expenditure I will not alter product demands.

[3] i.e., $\max_{\mathbf{q_i^e}} u_i^e(\mathbf{q_i^e}, I, \gamma) - \mathbf{p}\mathbf{q_i^e}$ and $\max_{\mathbf{q^j}} u^j(\mathbf{q^j}, I, \gamma) - \mathbf{p}\mathbf{q^j}$ for appropriate price vector $\mathbf{p}$.

5.2.3 Duopoly with differentiated products

We now investigate the equilibrium prices and equilibrium quantities of two firms, taking a network of customer relationships $(n_1^e, n^j, n_2^e) \in \mathcal{N}_n$ as given. On the market the two firms compete in quantities. Although there are exclusive and joint customers, a firm charges the same price to all his customers. The firms maximize profits given by the difference of price and costs times total quantity sold to either exclusive or joint customers. The price follows the inverse demand as given in (5.3). Firms have identical and constant marginal costs denoted by k.

Computing customers' demands is necessary to determine firms' (Cournot-Nash) equilibrium profits for a given network and, hence, to analyze incentives for creating new or deleting existing links. For this, we impose a technical assumption on γ to ensure non-negativity of linked customers' demands. The problem can be seen as follows. Suppose both firms sell their product to their exclusive and joint customers such that these demand according to equations (5.1) and (5.2) and there exists at least one customer of each type, i.e., $n_1^e, n^j, n_2^e > 0$. If both firms sell positive quantities to joint customers, then according to the second part of equation (5.3) the quantities sold to exclusive customers are positive for $\gamma \in [0, 1]$. However, they may become negative if $\gamma \in [-1, 0)$. Therefore, to obtain non-negative equilibrium quantities we require that for $\gamma \in [-1, 0)$ and $n_{3-i}^e > n_i^e > 0$

$$(1+\gamma)\left(n^j + n_{3-i}^e\right)\left[(2-\gamma)\left(n_i^e + n^j\right) - \gamma n_i^e\right] + \gamma^2 n^j \left(n_i^e - n_{3-i}^e\right) \geq 0 \quad \text{for } i = 1, 2. \tag{5.4}$$

These inequalities ensure that the quantities traded in equilibrium will be non-negative.[4] As we will see below, they are vacuously satisfied for $n \leq 3$ and satisfied for a wide range of γ when $n \geq 4$.

The profits of firm i are

$$\pi_i^{\left(n_1^e, n^j, n_2^e\right)}(\mathbf{q}) = \left(a - \frac{n_i^e q_i^e + n^j q_i^j + n^j \gamma q_{3-i}^j}{n_i^e + n^j} - k\right)\left(n_i^e q_i^e + n^j q_i^j\right),$$

[4]See also equation (5.7) below.

and, hence, the first order conditions are

$$\frac{\partial \pi_i^{\left(n_1^e,n^j,n_2^e\right)}(\mathbf{q})}{\partial q_i^e} = n_i^e a - \frac{2\left(n_i^e\right)^2 q_i^e + 2n_i^e n^j q_i^j + n_i^e n^j \gamma q_{3-i}^j}{n_i^e + n^j} - n_i^e k = 0, \tag{5.5}$$

$$\frac{\partial \pi_i^{\left(n_1^e,n^j,n_2^e\right)}(\mathbf{q})}{\partial q_i^j} = n^j a - \frac{2n_i^e n^j q_i^e + 2\left(n^j\right)^2 q_i^j + \left(n^j\right)^2 \gamma q_{3-i}^j}{n_i^e + n^j} - n^j k = 0 \quad (i \in \{1,2\}). \tag{5.6}$$

Note that equations (5.5) and (5.6) are linearly dependent. Using the first order conditions resulting from (5.5) or (5.6) and the second part of equation (5.3) we have for the equilibrium quantities

$$q_i^{e*}(k) = (a-k)\frac{(1+\gamma)\left(n^j + n_{3-i}^e\right)\left[(2-\gamma)\left(n_i^e + n^j\right) - \gamma n_i^e\right] + \gamma^2 n^j \left(n_i^e - n_{3-i}^e\right)}{4\left(n_i^e + n^j\right)\left(n^j + n_{3-i}^e\right) - \gamma^2\left(2n_i^e + n^j\right)\left(n^j + 2n_{3-i}^e\right)}, \tag{5.7}$$

$$q_i^{j*}(k) = (a-k)\frac{\left(n^j + n_{3-i}^e\right)\left[(2-\gamma)\left(n_i^e + n^j\right) - \gamma n_i^e\right]}{4\left(n_i^e + n^j\right)\left(n^j + n_{3-i}^e\right) - \gamma^2\left(2n_i^e + n^j\right)\left(n^j + 2n_{3-i}^e\right)} \quad \text{for } i = 1,2. \tag{5.8}$$

For $a \geq k$, these equilibrium quantities are non-negative for $\gamma \in [0,1]$ by the non-negativity of $q_i^{j*}(k)$ and the second part of equation (5.3). Moreover, with the condition (5.4) on γ and (n_1^e, n^j, n_2^e) this also holds for $q_i^{e*}(k)$ for $a \geq k$. The price of firm i, $i = 1, 2$, is

$$p_i(\mathbf{q}^*) = a - (a-k)\frac{(1+\gamma)(n^j + n_{3-i}^e)[(2-\gamma)(n_i^e + n^j) - \gamma n_i^e] + \gamma^2 n^j (n_i^e - n_{3-i}^e)}{4(n_i^e + n^j)(n^j + n_{3-i}^e) - \gamma^2(2n_i^e + n^j)(n^j + 2n_{3-i}^e)}. \tag{5.9}$$

It is easy to verify that for $a \geq k$ these prices are indeed positive.

If $n^j > 0$ and $n_1^e = 0$ or $n_2^e = 0$, to find the equilibrium quantities and prices we drop the first order conditions on the according quantities and on the relationship of the quantities. The according equilibrium quantities are as in equations (5.7) and (5.8) and the equilibrium prices are as in equations (5.9) inserting the according values for n_1^e and n_2^e.

If $n^j = 0$, there is no joint customer and thus, there is no dependence of the products of the firms. Using the first order condition (5.5) of the exclusive customer the equilibrium quantities reduce to $q_i^{e*}(k) = \frac{a-k}{2}$, and equilibrium prices are given by

$p_i(\mathbf{q}^*) = \frac{a+k}{2}$ for $i = 1, 2$.

We return to condition (5.4) that ensures non-negative equilibrium quantities. In fact it can be viewed as a restriction on the degree of substitutability γ that may be effective for $\gamma \in [-1, 0)$. However, for a large range they are satisfied for all network structures.

Proposition 5.1. *For all networks of customer relationships with $n \geq 3$ customers there exists $\gamma_{\min}^n < 0$ such that for all $\gamma \in [\gamma_{\min}^n, 0)$ condition (5.4) always holds. More precisely, $\gamma_{\min}^3 = -1$ and $\gamma_{\min}^n < -\frac{3}{4}$ for all $n \geq 4$.*

The bound $-\frac{3}{4}$ is valid for all $n \geq 4$, though it is not tight. As indicated in the proof, in specific networks we may have positive equilibrium quantities even for a greater range of complementarities.

5.2.4 Network formation and stability concepts

In the previous subsections we in particular calculated the firms' profits in (Cournot-Nash) equilibrium for a given network $(n_1^e, n^j, n_2^e) \in \mathcal{N}_n$. Comparing its profits for different networks reveals a firm's incentives to create or delete links to customers. Thereby, establishing or deleting a link are solely the firms' strategic decisions. Moreover, firms pay the entire linking costs. Phrased differently, we analyze firms' strategic decisions in customer acquisition. In our model, a customer cannot directly influence whether a link is established or not. One way to motivate this is that customers cannot influence receiving marketing benefits and there is no reason for customers to reject such links.. In that sense one-sided link formation should be interpreted as advertising a product that customers cannot discover themselves. A well-known model of one-sided link formation in this spirit is Bala and Goyal (2000), for instance.

Arguably, the central notion in network formation is stability. We are interested in networks in which no firm has an incentive to alter the network and to enjoy a higher profit. To define a stability concept, we need to specify which alterations of networks are feasible for an firm. In our setup, deviations from a given network of customer relationships are achieved by adding further links to acquire new exclusive or joint customers or by deleting existing links to own exclusive or joint customers.

Figure 5.2 illustrates deviations with the addition or deletion of single links. Along the solid lines firm 1 alters exactly one link, whereas dashed lines correspond to changes from firm 2. Deviations that involve exactly one customer and one firm are called *local*.

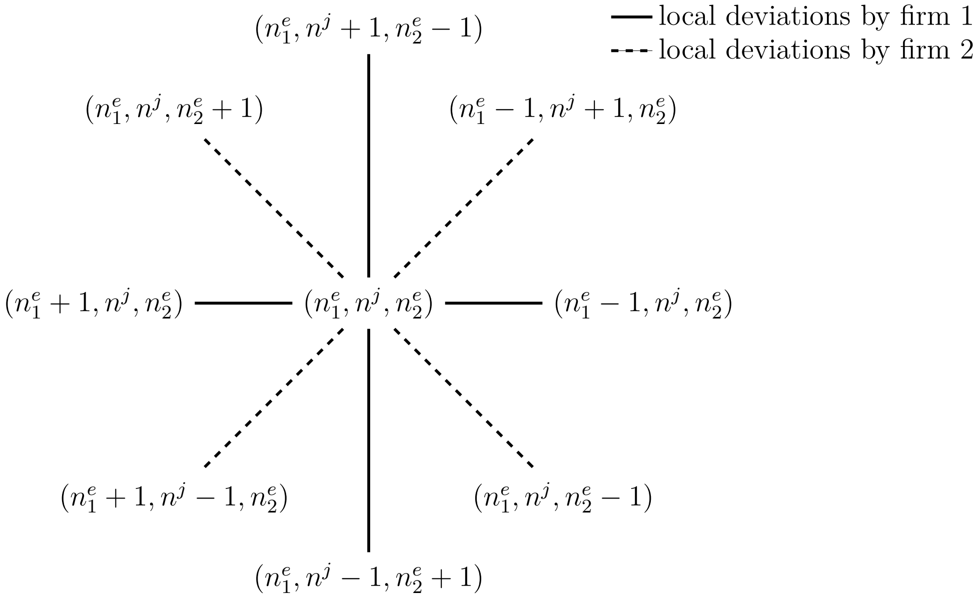


Figure 5.2: Local deviations from a network $(n_1^e, n^j, n_2^e) \in \mathcal{N}_n$

Apparently, not all of these local deviations are feasible for every network. The reason is that given a particular network there has to be an according customer to form or to delete a link. For example, as we consider the total number of customers to be fixed, it is not possible to add a further link if all customers are already linked to both firms. This is captured by defining which deviations are considered to be *locally feasible*.

Definition 5.1. *Given a total number of customers $n \in \mathbb{N}_{++}$ and a network of customer relationships $(n_1^e, n^j, n_2^e) \in \mathcal{N}_n$.*
*A **feasible local deviation of firm 1** from the network (n_1^e, n^j, n_2^e) to a network $(\tilde{n}_1^e, \tilde{n}^j, \tilde{n}_2^e)$ requires*

$$\left(\tilde{n}_1^e, \tilde{n}^j, \tilde{n}_2^e\right) \in \mathcal{N}_n \cap \left\{\left(n_1^e \pm 1, n^j, n_2^e\right), \left(n_1^e, n^j \pm 1, n_2^e \mp 1\right)\right\}$$

*Analogously, a **feasible local deviation of firm 2** from the network (n_1^e, n^j, n_2^e)*

to a network $(\tilde{n}_1^e, \tilde{n}^j, \tilde{n}_2^e)$ *requires*

$$\left(\tilde{n}_1^e, \tilde{n}^j, \tilde{n}_2^e\right) \in \mathcal{N}_n \cap \left\{\left(n_1^e, n^j, n_2^e \pm 1\right), \left(n_1^e \mp 1, n^j \pm 1, n_2^e\right)\right\}.$$

We next turn to the notion of stability. For this, consider firm i's equilibrium profit $\pi_i^{(n_1^e, n^j, n_2^e)}$ at network $(n_1^e, n^j, n_2^e) \in \mathcal{N}_n$. Recall that it is derived using the equilibrium prices and quantities on the market from Subsection 5.2.3. By $\tilde{\pi}_i^{(n_1^e, n^j, n_2^e)}$ denote the equilibrium profits including costs of link formation. Assuming that each link incurs a constant cost $c \geq 0$, this implies

$$\tilde{\pi}_i^{(n_1^e, n^j, n_2^e)} = \pi_i^{(n_1^e, n^j, n_2^e)} - \left(n_i^e + n^j\right) c \quad \text{for } i = 1, 2.$$

Roughly speaking, a network is stable if no firm has an incentive to feasibly modify the network by adding or deleting one or several links. Changing links has an effect on customers' demands, hence on the equilibrium prices, quantities, and profits. We discuss two versons of stability, distinguished by the number of allowed modifications.

The first concept, termed *local stability*, addresses modifications of single links, i.e., feasible local deviations as defined above. *Nash stability* considers the case in which an firm may manipulate several links at the same time, or, put in other words, iterates feasible local deviations.

Definition 5.2. *A network of customer relationships* $(n_1^e, n^j, n_2^e) \in \mathcal{N}_n$ *is* ***locally stable*** *if for both* $i = 1, 2$, $\tilde{\pi}_i^{(n_1^e, n^j, n_2^e)} \geq \tilde{\pi}_i^{(\tilde{n}_1^e, \tilde{n}^j, \tilde{n}_2^e)}$ *for all networks* $(\tilde{n}_1^e, \tilde{n}^j, \tilde{n}_2^e)$ *that result from a feasible local deviation of firm i .*

The basic question we raise is at what combinations of link cost c and degree of substitutability γ a given network is locally stable. For the analysis we start at the network and determine the minimal and maximal c (as a function of γ) such that adding or deleting a link does not result in a higher profit.

For given $(n_1^e, n^j, n_2^e) \in \mathcal{N}_n$ denote by $(n_1^e, n^j, n_2^e)_{i+}$ a network that is obtained if firm i feasibly adds one link and by $(n_1^e, n^j, n_2^e)_{i-}$ a network if firm i feasibly deletes one link. Then, the conditions imposed on the costs c of link formation are

$$\pi_i^{(n_1^e,n^j,n_2^e)} \geq \pi_i^{(n_1^e,n^j,n_2^e)_{i+}} - c, \tag{5.10}$$

$$\pi_i^{(n_1^e,n^j,n_2^e)} - c \geq \pi_i^{(n_1^e,n^j,n_2^e)_{i-}} \quad \text{for } i = 1, 2. \tag{5.11}$$

Inequality (5.10) defines a lower bound for the costs of link formation for the network $(n_1^e, n^j, n_2^e)_{i+}$, and inequality (5.11) provides an upper bound for the network $(n_1^e, n^j, n_2^e)_{i-}$. As there may be several possibilities to feasibly add or to delete links we have to take the maximal lower and minimal upper bound. Therefore, define

$$\underline{c}^{(n_1^e,n^j,n_2^e)} := \max\left\{\pi_i^{(n_1^e,n^j,n_2^e)_{i+}} - \pi_i^{(n_1^e,n^j,n_2^e)} \,|\, (n_1^e, n^j, n_2^e)_{i+},\ i = 1, 2\right\},$$

$$\overline{c}^{(n_1^e,n^j,n_2^e)} := \min\left\{\pi_i^{(n_1^e,n^j,n_2^e)} - \pi_i^{(n_1^e,n^j,n_2^e)_{i-}} \,|\, (n_1^e, n^j, n_2^e)_{i-},\ i = 1, 2\right\}.$$

Hence, the network (n_1^e, n^j, n_2^e) is locally stable if the costs of link formation are $c \in \left[\underline{c}^{(n_1^e,n^j,n_2^e)}, \overline{c}^{(n_1^e,n^j,n_2^e)}\right]$. Intuitively, c must not be too low, so that adding one link is not attractive, and it must not be too high, so that deletion would become favorable. Besides the case that networks should be stable against manipulation of single links, we also address the case of multiple changes and define *Nash stability.*

Definition 5.3. *A network of customer relationships $(n_1^e, n^j, n_2^e) \in \mathcal{N}_n$ is* ***Nash stable*** *if $\tilde{\pi}_i^{(n_1^e,n^j,n_2^e)} \geq \tilde{\pi}_i^{(\tilde{n}_1^e,\tilde{n}^j,\tilde{n}_2^e)}$ for all networks $(\tilde{n}_1^e, \tilde{n}^j, \tilde{n}_2^e)$ that result from iterated feasible local deviations of firm i for all $i = 1, 2$.*

Note that Bala and Goyal (2000) refer to these networks as “Nash networks”. Local deviations are depicted in Figure 5.2. The possibility for firm i to add or to delete several links at the same time actually partitions the set of all networks. Two networks are in the same set of this partition if there is a sequence of feasible local deviations of firm i that leads from the one network to the other one. More precisely, consider the network $(0, 0, \bar{n})$, in which firm 2 has exactly $\bar{n}$ links, all of them to exclusive customers. Then firm 1 may iteratively add a further link to one of these exclusive customers of firm 2, and/or to a new own exclusive customer who is not yet linked. As firm 1 is only able to manipulate own links, it follows that exactly all networks in which firm 2 has precisely $\bar{n}$ links are reachable by iterated feasible local deviations of firm 1. Hence, the partition Π_i^n of networks with n customers induced by local deviations of firm i contains $n + 1$ sets, indexed by the number of links of the other firm. Formally, we may describe the partition induced by firm i

by $\Pi_i^n := \{\Pi_i^n(0), \ldots, \Pi_i^n(n)\}$ with $\Pi_i^n(\bar{n}) := \{(n_1^e, n^j, n_2^e) \in \mathcal{N}_n | \, n^j + n_2^e = \bar{n}\}$ for $0 \leq \bar{n} \leq n$ for $i = 1, 2$. Then, a network $(n_1^e, n^j, n_2^e) \in \mathcal{N}_n$ is Nash stable if it provides the highest profit (including link costs) to firm 1 among all networks in $\Pi_1^n(n^j + n_2^e)$ and the highest profit to firm 2 among the networks in $\Pi_2^n(n_1^e + n^j)$. As above, for fixed unit link costs the resulting inequalities can be translated into bounds for c, such that the given network is Nash stable. Note that we may use this partition of the networks to figure out how many networks exist for a given number of customers n. Consider the partition for firm 1. Then, the set $\Pi_1^n(\bar{n})$ contains those networks in which firm 2 has exactly $n^j + n_2^e = \bar{n}$ links. The cardinality of $\Pi_1^n(\bar{n})$ is given by $|\Pi_1^n(\bar{n})| = (n + 1 - \bar{n})\,(\bar{n} + 1)$. There are $(\bar{n} + 1)$ possibilities how the links of firm 2 are split into joint and exclusive customers. This is, for firm 2 there may be $\bar{n}$ exclusive customers and no joint customers, $\bar{n} - 1$ exclusive customers and 1 joint customer and so on until there are no exclusive customers and $\bar{n}$ joint customers of firm 2. As for each given $\bar{n} = \bar{n}_2 + \bar{n}_3$ the links to joint customers need also to be established by the other firm, firm 1 has $n + 1 - \bar{n}_2 - \bar{n}_3 = n + 1 - \bar{n}$ possibilities to further establish links to obtain own exclusive customers. Hence, we may compute the number of existing networks for n customers by summing over all elements of the partition for firm 1, given by

$$|\mathcal{N}_n| = \sum_{\bar{n}=0}^{n} |\Pi_1^n(\bar{n})| = \sum_{\bar{n}=0}^{n} (n + 1 - \bar{n})\,(\bar{n} + 1) = \frac{(n+3)\,(n+2)\,(n+1)}{6}$$

where the last equation can be show by induction. Thus, the series containing the number of networks is given by

$$(|\mathcal{N}_n|)_{n \geq 1} = (4, 10, 20, 35, 56, 84, 120, 165, 220, 286, 364, 455, 560, 680, 816, \ldots)\,.$$

Using these insights about how to partition $\mathcal{N}_n$ we can rephrase the conditions for Nash stability considering the network $(n_1^e, n^j, n_2^e) \in \mathcal{N}_n$ for firm 1 as

$$\begin{gathered}\pi_1^{\left(n_1^e, n^j, n_2^e\right)} - \pi_1^{\left(\tilde{n}_1^e, \tilde{n}^j, \tilde{n}_2^e\right)} \geq \left(n_1^e + n^j - \tilde{n}_1^e - \tilde{n}_1^j\right) c \\ \text{for all } \left(\tilde{n}_1^e, \tilde{n}^j, \tilde{n}_2^e\right) \in \Pi_1^n \left(n^j + n_2^e\right) \setminus \left\{\left(n_1^e, n^j, n_2^e\right)\right\}.\end{gathered}$$

If the analogous conditions also hold for firm 2, then the network is Nash stable.

5.3 Stable networks

In this section, we examine local and Nash stability of networks with an arbitrary number of customers. As the number of possible networks and, hence, the number of possible deviations is large, we concentrate on specific networks and calculate and display the regions of local stability. Surprisingly, the qualitative picture even in the limit for $n \to \infty$ is very similar to what we see for few customers. By Proposition 5.1, we can ensure that the equilibrium quantities computed in equations (5.7) and (5.8) are non-negative only for $\gamma \in \left[-\frac{3}{4}, 1\right]$. Then, condition (5.4) is satisfied for all possible networks of customer relationships.

Our next Proposition collects local stability results of specific networks with n customers, namely the complete network with only joint customers $(0, n, 0)$ (*"perfect" oligopoly with two firms*), the empty network $(0, 0, 0)$, the asymmetric network with exclusive customers for one firm given by $(n, 0, 0)$ and $(0, 0, n)$ (*natural monopoly*), and, finally, the network with an equal number of exclusive customers for both firms $\left(\frac{n}{2}, 0, \frac{n}{2}\right)$ (*two coexisting monopolies*). Moreover, the limiting cases for growing n is included and some locally unstable networks are identified.

Proposition 5.2. *Consider* $(n_1^e, n^j, n_2^e) \in \mathcal{N}_n$ *and* $\gamma \in \left[-\frac{3}{4}, 1\right]$.

(i) The complete network with $(n_1^e, n^j, n_2^e) = (0, n, 0)$ *is locally stable if and only if the link costs c satisfy*

$$c \leq \frac{(a-k)^2 \, n \left(16n - 16n\gamma^2 + (3n+1)\,\gamma^4\right)}{(2+\gamma)^2 \left(4n - (n+1)\,\gamma^2\right)^2}.$$

This upper bound converges to

$$\lim_{n\to\infty} \frac{(a-k)^2 \, n \left(16n - 16n\gamma^2 + (3n+1)\,\gamma^4\right)}{(2+\gamma)^2 \left(4n - (n+1)\,\gamma^2\right)^2} = \frac{(a-k)^2 \left(4 - 3\gamma^2\right)}{(2-\gamma)\,(2+\gamma)^3}.$$

(ii) The empty network with $(n_1^e, n^j, n_2^e) = (0, 0, 0)$ *is locally stable if and only if the link costs c satisfy* $c \geq \frac{(a-k)^2}{4}$.

(iii) An asymmetric network with $(n_1^e, n^j, n_2^e) = (0, 0, n)$ *or* $(n_1^e, n^j, n_2^e) = (n, 0, 0)$

is locally stable if and only if the link costs c satisfy

$$c \in \left[\frac{(a-k)^2 n^2 (2-\gamma)^2}{(4n-(2n-1)\gamma^2)^2}, \frac{(a-k)^2}{4} \right].$$

This lower bound converges to

$$\lim_{n\to\infty} \frac{(a-k)^2 n^2 (2-\gamma)^2}{(4n-(2n-1)\gamma^2)^2} = \frac{(a-k)^2 (2-\gamma)^2}{4(2-\gamma^2)^2}.$$

(iv) For n even, the symmetric network with $(n_1^e, n^j, n_2^e) = \left(\frac{n}{2}, 0, \frac{n}{2}\right)$ *is locally stable if and only if the link costs c satisfy*

$$c \in \left[\frac{(a-k)^2 n^2 (n+2)(2+n-\gamma-(n-1)\gamma^2)^2}{8(n(n+2)-(n-1)(n+1)\gamma^2)^2} - \frac{(a-k)^2 n}{8}, \frac{(a-k)^2}{4} \right].$$

This lower bound converges to

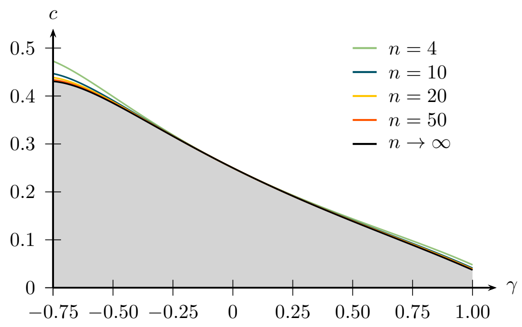
$$\lim_{n\to\infty} \frac{(a-k)^2 n^2 (n+2)(2+n-\gamma-(n-1)\gamma^2)^2}{8(n(n+2)-(n-1)(n+1)\gamma^2)^2} - \frac{(a-k)^2 n}{8} = \frac{(a-k)^2}{4(1+\gamma)}.$$

(v) For $n \geq 3$ *the asymmetric networks with* $(n_1^e, n^j, n_2^e) = (0, 1, n-1)$ *and* $(n_1^e, n^j, n_2^e) = (n-1, 1, 0)$ *are locally unstable for all* $\gamma \neq 0$ *and for* $\gamma = 0$ *for all* $c \neq \frac{(a-k)^2}{4}$.

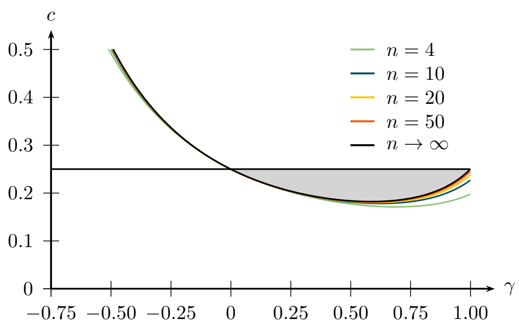
(vi) Networks $(n_1^e, n^j, n_2^e) \in \mathcal{N}_n$ *with* $n^j = 0$ *and* $0 < n_1^e + n_2^e < n$ *are locally unstable for* $c \neq \frac{(a-k)^2}{4}$.

(vii) For $\gamma = 0$ *and* $c = \frac{(a-k)^2}{4}$ *all networks* $(n_1^e, n^j, n_2^e) \in \mathcal{N}_n$ *are locally stable.*

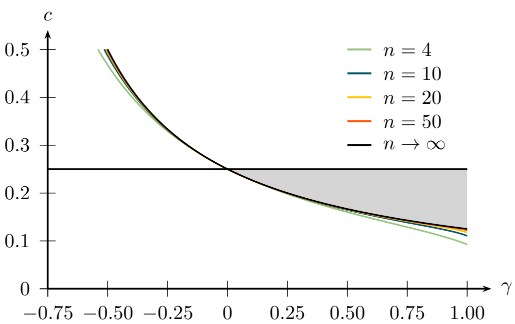
Figure 5.3 illustrates some of the findings of Proposition 5.2. The stability regions for these networks for relatively small n already appear to be a good approximation to the stability regions when the number of customers n becomes large. Graphically, already for $n = 20$ there is hardly any visible difference compared to the limit region. We observe that the perfect oligopoly with two firms may be locally stable for any degree of product differentiation. In comparison the natural monopoly or two coexisting monopolies of identical size are only locally stable for substitutable products. The costs of link formation need to be relatively high for the monopolies to be locally stable compared to the corresponding oligopoly (for a given degree of product differentiation for substitutable products).



(a) $(0, n, 0)$



(b) $(n, 0, 0)$, $(0, 0, n)$



(c) $(\frac{n}{2}, 0, \frac{n}{2})$

Figure 5.3: Stable regions of networks with n customers for $\gamma \in \left[-\frac{3}{4}, 1\right]$ $(a - k = 1)$

In general for the existence of locally stable networks we know the following.

Proposition 5.3. *For all $n \in \mathbb{N}_+$, all combinations of degrees of substitutability for complementary or independent products with $\gamma \in \left[-\frac{3}{4}, 0\right]$, and all costs of link formation $c \geq 0$ there exists at least one locally stable network. More precisely, either the network $(0, n, 0)$, the network $(0, 0, 0)$ or both are locally stable. When products are substitutable, then for all $\gamma \in (0, 1]$ there exist costs of link formation $c \geq 0$, for which neither the complete nor the empty network is locally stable.*

Figure 5.4 illustrates these findings graphically. Proposition 5.3 shows that for complementary or independent products we are sure that a locally stable network always exists. However, we also establish that for substitutable products this issue is more complex. Among others, we further investigate this question in case of three customers in Section 5.4.

The next proposition relates local and Nash stability for some particular networks

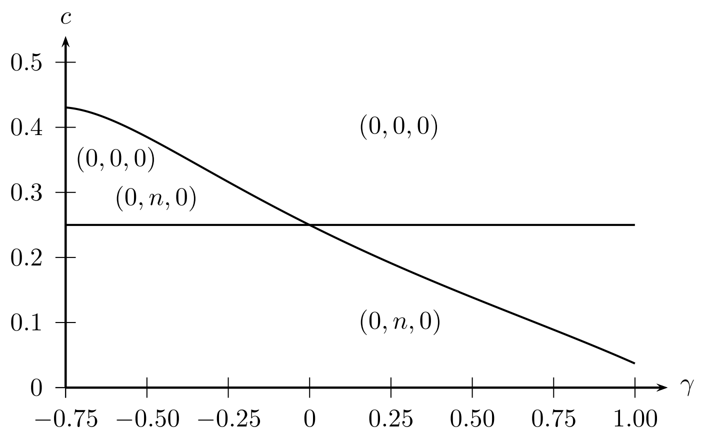


Figure 5.4: Limit regions for local stability of the networks $(0, n, 0)$ and $(0, 0, 0)$ $(a - k = 1)$

from Proposition 5.2.

Proposition 5.4.

(i) Consider networks $(n_1^e, n^j, n_2^e) \in \mathcal{N}_n$ *and* $\gamma \in [\gamma_{\min}^n, 1]$*. The networks* $(0, n, 0)$*,* $(0, 0, n) / (n, 0, 0)$ *and* $(0, 0, 0)$ *are Nash stable if and only if they are locally stable.*

(ii) For $\gamma = 0$ *and* $c = \frac{(a-k)^2}{4}$ *all networks* $(n_1^e, n^j, n_2^e) \in \mathcal{N}_n$ *are Nash stable.*

(iii) Local and Nash stability do not coincide.

Proposition 5.4 (i) exemplarily shows that there is no difference between the concepts of local and Nash stability for the prominent networks with only exclusive, only joint, or no customers, i.e., the regions of stability for local and Nash stability coincide. The proof in particular shows that if there is a beneficial deviation for a firm that involves several links, then there must be a feasible local deviation that increases the firm's profit. The main intuition behind Proposition 5.4 (i) is that the incentives for feasible deviations are provided by networks that are directly neighboring. For networks within the set $\Pi_i^n(0)$ a local deviation by firm i $(i \in \{1, 2\})$ triggers the same change in profits. Hence, they are increasing or decreasing in the number of exclusive customers and therefore the conditions imposed by local and Nash stability for deviations of firm i coincide. Proposition 5.4 (iii) reveals that the coincidence of local and Nash stability that was established in Proposition 5.4 (i) for selected networks with n customers does not generalize to arbitrary networks of customer relationships. We close this section by returning to the issue for which degrees of substitutability γ our analysis in Proposition 5.2 is valid.

Proposition 5.5. *The results on local stability for the networks* $(0,n,0)$*,* $(0,0,0)$*,* $(n,0,0)$ *and* $(0,0,n)$ *as in Proposition 5.2 (i) to (iii) are valid for all* $\gamma \in [-1,1]$ *and those for the network* $\left(\frac{n}{2},0,\frac{n}{2}\right)$ *as in Proposition 5.2 (iv) are valid for all* $\gamma \in \left[-\frac{\sqrt{3}\sqrt{163}-3}{20},1\right] \approx [-0.96,1]$.

Proposition 5.5 states that for the specific networks we consider in Proposition 5.2 (i) to (iv), equilibrium quantities are well defined for a broader range of γ than given in Proposition 5.1. Readers that feel uncomfortable with not having the full range $[-1,1]$ may be reconciled with the fact that there is not much difference between a utility function of a joint customer for $\gamma=-1$ and $\gamma=-\frac{3}{4}$. Unlike the use of Leontief type utility functions that precisely capture the notion of perfect complements, the utility functions used here (as well as in the literature on product differentiation) still allow for some substitutability at $\gamma=-1$. Although it is only a good proxy for perfect complements, the main advantage over taking a Leontief type utility function is that equilibrium quantities are well-defined and, hence, we get a sound basis for our analysis of stable customer relationship networks.

5.4 Networks with three customers

In this section we closer investigate a special case that allows us a complete characterization of locally stable networks. More precisely, for networks with three customers we completely identify conditions on link costs and the degree of substitutability that renders specific networks locally stable. As shown in Proposition 5.1 we may consider $\gamma \in [-1,1]$ for $n=3$. With three customers there are in total 20 different networks, presented in Figure 5.5. The complete network $(0,3,0)$ and empty network $(0,0,0)$ are at the top and bottom. Symmetric networks ($n_1^e = n_2^e$) are found in the center “column”. The relations between those networks through feasible local deviations from firm 1 are depicted by solid lines, those of firm 2 by dashed lines. Observe that the partition Π_1^3 resulting from the iteration of feasible local deviations precisely contains the connected components when only considering connections along solid lines (firm 1). Here, there are four sets in this partition.

Our first stability result investigates local stability and distinguishes between the

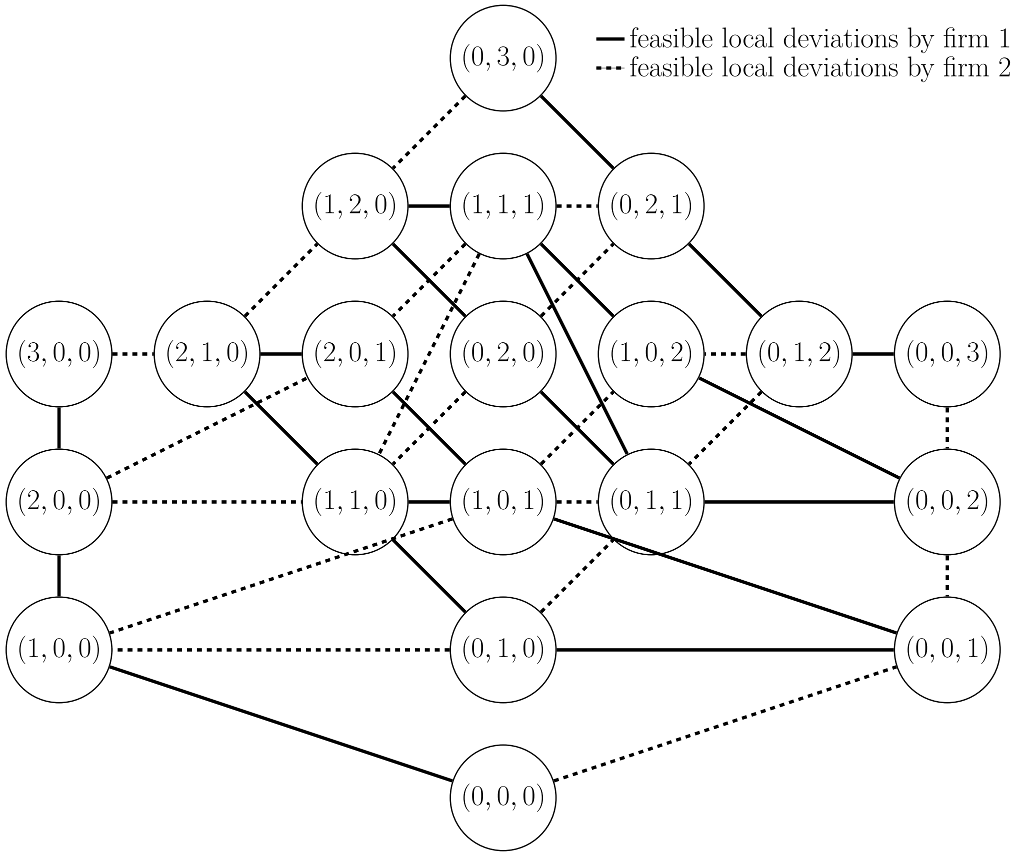


Figure 5.5: Networks with $(n_1^e, n^j, n_2^e) \in \mathcal{N}_3$

case with substitutes ($\gamma > 0$) and complements ($\gamma < 0$).

Proposition 5.6. *Consider networks $(n_1^e, n^j, n_2^e) \in \mathcal{N}_3$.*

(i) For substitutable products with $\gamma \in (0, 1]$ there exist costs of link formation such that there are symmetric ($n_1^e = n_2^e$) locally stable networks different from the empty and complete network and there exist costs of link formation such that there are asymmetric ($n_1^e \neq n_2^e$) locally stable networks.

(ii) For complementary products with $\gamma \in [-1, 0)$ there are no asymmetric locally stable networks for any costs $c \neq \frac{(a-k)^2}{4}$ of link formation. Moreover, there exist costs of link formation such that there are symmetric locally stable networks with $0 < n_1^e + n^j + n_2^e < 3$.

Apart from the results stated in Proposition 5.6, we have analytically determined the regions (in γ/c space) of stability for all networks with three customers. Figure 5.6

shows the regions of stability for stable networks with three customers (for $a - k = 1$). The displayed functions in γ are upper and lower bounds on c imposed by local stability. The gray-shaded area indicates the region of (γ, c) combinations for which the according network is locally stable. Note that the local instability of the networks $(2, 1, 0)$ and $(0, 1, 2)$ has already been established in Proposition 5.2(v) for a general number of customers. The networks not shown in Figure 5.6 are locally stable only for very specific values of γ and c. More precisely, the networks $(1, 1, 0)$ and $(0, 1, 1)$ are locally stable only for $\gamma = 0$ and $c = \frac{(a-k)^2}{4}$ and the remaining networks $(2, 0, 0)$, $(0, 0, 2)$, $(1, 0, 0)$ and $(0, 0, 1)$ for all $\gamma \in [-1, 1]$ and $c = \frac{(a-k)^2}{4}$. In view of Proposition 5.6 and Figure 5.6, we observe that with complementary products we rather have symmetric networks among the locally stable ones, whereas for substitutable products, asymmetric networks appear in the set of stable networks. This confirms the intuition that complementary products trigger joint customers, whereas substitutable products rather lead to exclusiveness.

Looking at Figure 5.6 there is a striking similarity between the shape of stability regions depicted in Figure 5.3 for n customers in comparison to the corresponding regions displayed for three customers. Qualitatively there is no difference when comparing the regions for the complete network, the network with only one firm being linked to all customers, or a proper dispersion of exclusive customers. For the latter, compare the networks $(\frac{n}{2}, 0, \frac{n}{2})$ and $(2, 0, 1)/(1, 0, 2)$. We may therefore conclude that the basic insights in stability of customer relationship networks seem to be already observable in the $n = 3$ case. Comparing the different graphs in Figure 5.6, we observe that the regions of stability may overlap or be adjacent for different networks. As a consequence, for fixed costs of link formation and a fixed degree of substitutability, there might be more than one locally stable network. However, even if in some cases the regions of local stability intersect, we also observe that there exist cases in which they are directly adjacent. The next proposition in particular establishes that at least one locally stable network always exists.

Proposition 5.7. *Consider networks* $(n_1^e, n^j, n_2^e) \in \mathcal{N}_3$. *For all combinations of degrees of substitutability* $\gamma \in [-1, 1]$ *and all costs of link formation* $c \geq 0$ *there exists at least one locally stable network. More precisely, if we consider substitutable products with* $\gamma \in (0, 1]$ *and order the networks by increasing costs of link formation, then the regions of local stability of the networks* $(0, 3, 0)$, $(1, 2, 0) / (0, 2, 1)$, $(1, 1, 1)$, $(1, 0, 2) / (2, 0, 1)$ *and* $(0, 0, 0)$ *are directly adjacent. For complementary products*

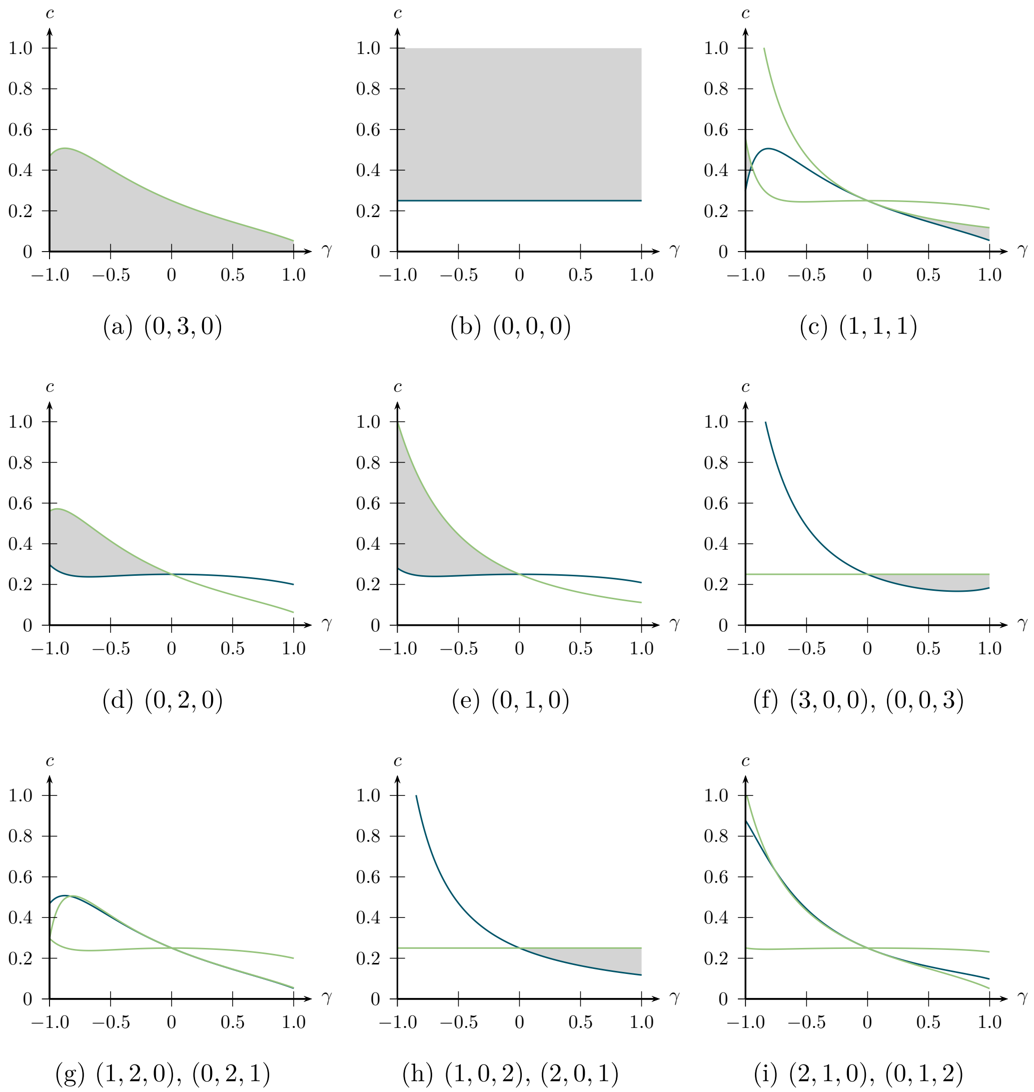


Figure 5.6: Regions of stability for networks with three customers ($a - k = 1$)

with $\gamma \in [-1, 0)$ *or independent products with* $\gamma = 0$ *either the network* $(0, 3, 0)$*, the network* $(0, 0, 0)$ *or both are locally stable.*

Figure 5.7 illustrates the stability regions of the networks from Proposition 5.7.

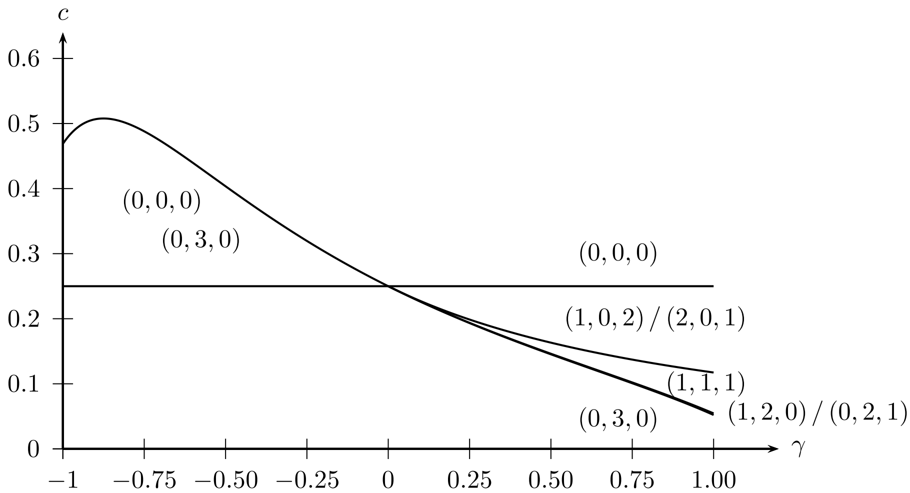


Figure 5.7: Existence of locally stable networks for $n = 3$ $(a - k = 1)$

Note that the stability region of the networks $(1, 2, 0)$ and $(0, 2, 1)$ in Figure 5.7 also Figure 5.6g is hardly visible. However, in the proof of Proposition 5.7, we establish that it is indeed located between the stability region of the networks $(0, 3, 0)$ and $(1, 1, 1)$. To see the adjacency result, first consider substitutable products with $\gamma \in (0, 1]$ in Figure 5.7. At very low costs of link formation c, the network $(0, 3, 0)$ is locally stable. Then, when the costs of link formation increase, first all links to joint customers are successively deleted going from $(0, 3, 0)$ to $(1, 2, 0)$ / $(0, 2, 1)$, then to $(1, 1, 1)$, and finally to $(1, 0, 2)$ / $(2, 0, 1)$. One may think of firms alternatingly deleting links to joint customers. Finally, for relatively high costs of link formation, also the links to exclusive customers are deleted, ending up with the empty network $(0, 0, 0)$. In contrast, for complementary or independent products with $\gamma \in [-1, 0]$, as indicated in Figure 5.7, the empty and the complete network already suffice to find a locally stable network for any (γ, c). At intermediate costs of link formation even both of them are locally stable. We add a final remark on the shape of the local stability regions in Figures 5.6 and 5.7 to close our discussion on local stability. When looking at lower and upper bounds, the profits in two networks (the stable one and a deviation network) are considered. However, one should not fail not recall that with a change of the structure of exclusive and joint customers, the environment for competition among firms may completely change. Intuitively, such changes take more effect, the more complementary the products are. We use this as an explanation for the fact that the functions describing lower and upper bounds are not monotonic, especially for γ sufficiently close to -1. Lastly, we briefly recall the results from Proposition 5.4 on the second notion of stability. From Figure 5.5 we readily identify the partitions from both firms induced by iterated feasible local deviations. The proof of Proposition 5.4 (iii) indeed gives a counter example with $n = 3$ showing that for a range of complementary products that the network $(1, 1, 1)$ is locally but not Nash stable as there is an incentive to deviate to the network $(0, 2, 0)$. This can also be observed in Figure 5.6. However, by Proposition 5.4 (i) we also know that the networks $(0, 3, 0)$, $(0, 0, 3) \,/\, (3, 0, 0)$ and $(0, 0, 0)$ are Nash stable if and only if they are locally stable.

5.5 Conclusion

Interpreting links to customers as relationships that enable trade to sell a product, we analyzed the stability of networks where two firms strategically form costly links to customers. Given a network of customer relationships $(n_1^e, n^j, n_2^e) \in \mathbb{N}_+^3$ with n customers and $n_1^e + n^j + n_2^e \leq n$, we determined the equilibrium prices and quantities for quantity competition between two firms. We identified a lower bound on the products' substitutability that is needed to ensure interior solutions in equilibrium and holds for all networks with $n \geq 3$ customers (Proposition 5.1 and also Proposition 5.5). Furthermore, we observed that the substitutability of the firms' products, γ, and the costs of link formation, c, influence the firms' equilibrium profits and, thus, have an impact on the incentives to strategically form relationships to customers. To analyze these incentives we introduced a notion of local and of Nash stability to identify for a given network of customer relationships regions of (γ, c), in which the given network is stable. In the general case with an arbitrary number of customers, we determined the stability regions for selected networks and presented the limit regions of stability when n goes to infinity (Proposition 5.2). It turned out that the shape of the stability regions does not significantly change compared to a setting with relatively small n. This implies that already for small n we obtain a good picture of the general scenario. For the existence of locally stable networks with n customers we showed that for complementary or independent products either the empty, the complete or both networks are locally stable, which is not true for substitutable products (Proposition 5.3). In addition, for selected networks we established that the stability regions for Nash and local stability indeed coincide (Proposition 5.4). Thus, allowing for the addition or deletion of several links does not influence the stability. For networks of customer relationships with three customers, we observed that there is a tendency to have symmetric networks among the locally stable ones for complementary products, whereas rather asymmetric networks are stable for substitutable products (Proposition 5.6). Moreover, a locally stable network always exists (Proposition 5.7). In sum, the novelty of our model compared to the existing literature is that instead of an indivisible good being exchanged along a link, it is possible to sell perfectly divisible goods and the division of surplus between buyer and seller depends on equilibrium prices and quantities.

We close with two implications that should be mentioned. From a managerial perspective, the successful sales manager of a company should carefully observe these parameters when deciding on which customers should be acquired next. One might either go for exclusive or joint (i.e., shared) customers or even build up a specific mixture of both customer types or stop acquiring new customers if the acquisition costs are becoming too high. Given a network of customer relationships and a specific degree of substitutability, we will be able to provide conclusions regarding the optimal behavior of the firms and the stability of the network. From a market design perspective, influencing link costs (e.g., through legal restrictions) ultimately has an effect on the network structure in the market. If one, for instance, thinks of the market as organized on a centralized (Internet) platform, link costs and, hence, stable trade relations can be influenced by the platform owner. As we found out, this will in particular be true, when products are well substitutable.

5.A Appendix A: Proofs

5.A.1 Proof of Proposition 5.1

Proof: For $n = 3$ there does not exist a network of customer relationships with $n^j > 0$ and $n_2^e > n_1^e > 0$ or $n_1^e > n_2^e > 0$ and, hence, condition (5.4) is vacuously satisfied. Therefore, we have to consider networks with $n \geq 4$ customers to see when condition (5.4) may be violated.

Consider a network (n_1^e, n^j, n_2^e) with $n \geq 4$, $n^j > 0$ and $n_2^e > n_1^e > 0$, i.e., $i = 1$ in condition (5.4). A necessary condition to find a network such that condition (5.4) does not hold is a negative coefficient in front of n_2^e, which is

$$(1+\gamma)\left[(2-\gamma)\left(n_1^e+n^j\right)-\gamma n_1^e\right]-\gamma^2 n^j = 2\left(1-\gamma^2\right)\left(n_1^e+n^j\right)+\gamma n^j < 0.$$

Then, if there exists an n_2^e large enough in comparison to n_1^e and n^j, condition (5.4) no longer holds. However, the coefficient in front of n_2^e in condition (5.4) is always positive for $\gamma > -\frac{\sqrt{16\left(n_1^e+n^j\right)^2+(n^j)^2}-n^j}{4\left(n_1^e+n^j\right)}$ which is bounded above for any n_1^e and n^j by

$$-\frac{\sqrt{16\left(n_1^e+n^j\right)^2+\left(n^j\right)^2}-n^j}{4\left(n_1^e+n^j\right)} - \frac{\sqrt{16\left(n_1^e+n^j\right)^2}-\left(n^j+n_1^e\right)}{4\left(n_1^e+n^j\right)} = -\frac{3}{4}.$$

This means for $\gamma \geq -\frac{3}{4}$ that there does not exist a network with $n \geq 4$ that violates condition (5.4). Note that, however, this bound is not tight. We obtain for networks with $n \geq 4$ and $n_1^e > n_2^e > 0$ for condition (5.4) an analogous bound depending on n_2^e and n^j.

Suppose now n is fixed. Consider again networks with $n^j > 0$ and $n_2^e > n_1^e > 0$ or $n_1^e > n_2^e > 0$, i.e., $i = 1$ in condition (5.4). Let $\gamma^{\left(n_1^e, n^j, n_2^e\right)}$ denote the value for γ when condition (5.4) is satisfied with equality (for $i = 1$). To find the network over all networks with at most n customers, which imposes the maximal lower bound on γ, we choose $n_1^e = 1$ and n^j as large as possible as

$$\frac{\partial\left(-\frac{\sqrt{16\left(n_1^e+n^j\right)^2+(n^j)^2}-n^j}{4\left(n_1^e+n^j\right)}\right)}{\partial n_1^e} = -\frac{n^j\left(\sqrt{16\left(n_1^e+n^j\right)^2+\left(n^j\right)^2}-n^j\right)}{4\left(n_1^e+n^j\right)^2\sqrt{16\left(n_1^e+n^j\right)^2+\left(n^j\right)^2}} \leq 0,$$

$$\frac{\partial\left(-\frac{\sqrt{16\left(n_1^e+n^j\right)^2+(n^j)^2}-n^j}{4\left(n_1^e+n^j\right)}\right)}{\partial n^j} = \frac{n_1^e\left(\sqrt{16\left(n_1^e+n^j\right)^2+\left(n^j\right)^2}-n^j\right)}{4\left(n_1^e+n^j\right)^2\sqrt{16\left(n_1^e+n^j\right)^2+\left(n^j\right)^2}} \geq 0.$$

As we consider now networks with at most n customers, this observation implies that there is a tradeoff between choosing n^j and n_2^e large enough for finding the network that yields the lower bound for γ. This implies that the network that puts the lower bound on γ is of the form $(n_1^e, n^j, n_2^e) = (1, n-n_2^e-1, n_2^e)$ with $2 \leq n_2^e \leq n-2$. For symmetry reasons this holds analogously for $i = 2$ and $n_1^e > n_2^e > 0$ with networks of the form $(n_1^e, n^j, n_2^e) = (n_1^e, n - n_1^e - 1, 1)$ with $2 \leq n_1^e \leq n-2$. Therefore,

$$\gamma_{\min}^n = \max\left\{\max_{2\leq n_2^e\leq n-2} \gamma^{\left(1,n-n_2^e-1,n_2^e\right)}, \max_{2\leq n_1^e\leq n-2} \gamma^{\left(n_1^e,n-n_1^e-1,1\right)}\right\}$$

$$= \max_{2\leq n_1^e\leq n-2} -\frac{\sqrt{16\,(n-1)^2 + (n-n_1^e-1)^2} - (n-n_1^e-1)}{4\,(n-1)}. \qquad \blacksquare$$

5.A.2 Proof of Proposition 5.2

Suppose $\gamma \in \left[-\frac{3}{4}, 1\right]$. The profits needed for the proof can be found in Table 5.1.

Table 5.1: Equilibrium profits for networks with n customers

(n_1^e, n^j, n_2^e)	$\pi_1^{\left(n_1^e,n^j,n_2^e\right)}$	$\pi_2^{\left(n_1^e,n^j,n_2^e\right)}$
$(0,n,0)$	$\frac{(a-k)^2 n}{(2+\gamma)^2}$	$\frac{(a-k)^2 n}{(2+\gamma)^2}$
$(0,n-1,1)$	$\frac{(a-k)^2 n^2 (2-\gamma)^2 (n-1)}{(4n-(n+1)\gamma^2)^2}$	$\frac{(a-k)^2 n\left(2n-(n-1)\gamma-\gamma^2\right)^2}{(4n-(n+1)\gamma^2)^2}$
$(0,1,n-1)$	$\frac{(a-k)^2 n^2 (2-\gamma)^2}{(4n-(2n-1)\gamma^2)^2}$	$\frac{(a-k)^2 n\left(2n-\gamma-(n-1)\gamma^2\right)^2}{(4n-(2n-1)\gamma^2)^2}$
$(0,2,n-2)$	$\frac{(a-k)^2 n^2 (2-\gamma)^2}{2(2n-(n-1)\gamma^2)^2}$	$\frac{(a-k)^2 n\left(2n-2\gamma-(n-2)\gamma^2\right)^2}{4(2n-(n-1)\gamma^2)^2}$
$(0,1,n-2)$	$\frac{(a-k)^2 (2-\gamma)^2 (n-1)^2}{(4(n-1)-(2n-3)\gamma^2)^2}$	$\frac{(a-k)^2 (2-\gamma)^2 (n-1)\left(2(n-1)-\gamma-(n-2)\gamma^2\right)}{(4(n-1)-(2n-3)\gamma^2)^2}$
$\left(\frac{n}{2},0,\frac{n}{2}\right)$	$\frac{(a-k)^2 n}{8}$	$\frac{(a-k)^2 n}{8}$
$\left(\frac{n}{2},1,\frac{n}{2}-1\right)$	$\frac{(a-k)^2 n^2 (n+2)\left(2+n-\gamma-(n-1)\gamma^2\right)^2}{8(n(n+2)-(n-1)(n+1)\gamma^2)^2}$	$\frac{(a-k)^2 n\left(n(n+2)-(n+2)\gamma-(n-2)(n+1)\gamma^2\right)^2}{8(n(n+2)-(n-1)(n+1)\gamma^2)^2}$
$\left(\frac{n}{2}-1,0,\frac{n}{2}\right)$	$\frac{(a-k)^2 (n-2)}{8}$	$\frac{(a-k)^2 n}{8}$
$(0,0,n)$	0	$\frac{(a-k)^2 n}{4}$
$(0,0,n-1)$	0	$\frac{(a-k)^2 (n-1)}{4}$
$(0,0,0)$	0	0

Proof: (i) Because of symmetry we just consider within this proof deviations of firm 1. From a local perspective, possible deviations from the complete network $(0,n,0)$ are if firm 1 deletes a link to a joint customer yielding to the network $(0,n-1,1)$. For local stability we require $\pi_1^{(0,n,0)} - c \geq \pi_1^{(0,n-1,1)}$. Using the profits from Table 5.1 we obtain

$$c \leq \pi_1^{(0,n,0)} - \pi_1^{(0,n-1,1)} = \frac{(a-k)^2\, n\,(16n\,(1-\gamma^2) + (3n+1)\,\gamma^4)}{(2+\gamma)^2\,(4n-(n+1)\,\gamma^2)^2}. \tag{5.12}$$

For the convergence we use L'Hôpital's rule. First note that for $\gamma \in [-1,1]$ we have

$$\lim_{n\to\infty} (a-k)^2 n \left(16n\left(1-\gamma^2\right)+(3n+1)\gamma^4\right) = \infty,$$
$$\lim_{n\to\infty} (a-k)^2 \left(32n\left(1-\gamma^2\right)+6n\gamma^4+\gamma^4\right) = \infty,$$
$$\lim_{n\to\infty} (2+\gamma)^2 \left(4n-(n+1)\gamma^2\right)^2 = \infty,$$
$$\lim_{n\to\infty} 2(2+\gamma)^2 \left(4n-(n+1)\gamma^2\right)\left(4-\gamma^2\right) = \infty.$$

Therefore, applying L'Hôpital's rule twice yields

$$\lim_{n\to\infty} \frac{(a-k)^2 n\left(16n\left(1-\gamma^2\right)+(3n+1)\gamma^4\right)}{(2+\gamma)^2\left(4n-(n+1)\gamma^2\right)^2} = \frac{(a-k)^2\left(4-3\gamma^2\right)}{(2-\gamma)(2+\gamma)^3}.$$

(ii) Consider the empty network with $(0,0,0)$. Possible local deviations from the empty network are if firm 1 forms a link yielding $(1,0,0)$ (see Figure 5.5). In this case, profits raise from zero to $\frac{(a-k)^2}{4}$ (as computed in Table 5.1). Therefore, if and only if the costs of link formation are $c \geq \frac{(a-k)^2}{4}$, then the empty network with $(0,0,0)$ is locally stable.

(iii) Consider now the network $(0,0,n)$. The possible deviations are for firm 1 to form a link to an exclusive customer of firm 2 yielding the network $(0,1,n-1)$. The only possible deviation for firm 2 is to delete a link to an own exclusive customer yielding the network $(0,0,n-1)$. Therefore, we require

$$\pi_1^{(0,0,n)} \geq \pi_1^{(0,1,n-1)} - c, \quad \pi_2^{(0,0,n)} - c \geq \pi_2^{(0,0,n-1)}. \tag{5.13}$$

Using the profits from Table 5.1 this is equivalent to requiring from the first inequality in (5.13) $c \geq \frac{(a-k)^2 n^2 (2-\gamma)^2}{(4n-(2n-1)\gamma^2)^2}$ and from the second inequality in (5.13) $c \leq \frac{(a-k)^2}{4}$. Thus, we have found an upper and a lower bound for the costs of link formation such that the networks $(0,0,n)$ an $(n,0,0)$ are locally stable. For the convergence we use again L'Hôpital's rule. First note that for $\gamma \in [-1,1]$ we have

$$\lim_{n\to\infty} (a-k)^2 n^2 (2-\gamma)^2 = \infty, \quad \lim_{n\to\infty} 2(a-k)^2 n (2-\gamma)^2 = \infty,$$
$$\lim_{n\to\infty} \left(4n-(2n-1)\gamma^2\right)^2 = \infty, \quad \lim_{n\to\infty} 4\left(4n-(2n-1)\gamma^2\right)\left(2-\gamma^2\right) = \infty.$$

Therefore, applying L'Hôpital's rule twice yields

$$\lim_{n\to\infty} \frac{(a-k)^2 n^2 (2-\gamma)^2}{(4n-(2n-1)\gamma^2)^2} = \frac{(a-k)^2 (2-\gamma)^2}{4(2-\gamma^2)^2}.$$

(iv) Because of symmetry we just consider within this proof deviations of firm 1. From a local perspective, possible deviations from the $(\frac{n}{2}, 0, \frac{n}{2})$ are if firm 1 forms a link to an exclusive customer of firm 2 yielding $(\frac{n}{2}, 1, \frac{n}{2}-1)$ or deletes a link to an own exclusive customer yielding to the network $(\frac{n}{2}-1, 0, \frac{n}{2})$. For local stability we require

$$\pi_1^{\left(\frac{n}{2},0,\frac{n}{2}\right)} \geq \pi_1^{\left(\frac{n}{2},1,\frac{n}{2}-1\right)} - c, \quad \pi_1^{\left(\frac{n}{2},0,\frac{n}{2}\right)} - c \geq \pi_1^{\left(\frac{n}{2}-1,0,\frac{n}{2}\right)} \tag{5.14}$$

Using the profits from Table 5.1 this is equivalent to requiring from the first inequality in (5.14)

$$c \geq \frac{(a-k)^2 n^2 (n+2)(2+n-\gamma-(n-1)\gamma^2)^2}{8(n(n+2)-(n-1)(n+1)\gamma^2)^2} - \frac{(a-k)^2 n}{8}$$

and from the second inequality in (5.14) $c \leq \frac{(a-k)^2}{4}$. Thus, we have found an upper and a lower bound for the costs of link formation such that the network $(\frac{n}{2}, 0, \frac{n}{2})$ is locally stable. For the convergence we use again L'Hôpital's rule. First note that for $\gamma \in [-1, 1]$ the numerator of the lower bound is a polynomial in n with non-negative coefficients given by

$$(a-k)^2 n\Big(2(1+\gamma)(1-\gamma)^2 n^3 + (1-\gamma)\left(8-\gamma^2+\gamma^3\right) n^2 \\ +2\left(4-4\gamma+3\gamma^2-2\gamma^3+\gamma^4\right) n - \gamma^4\Big)$$

and, analogously, the denominator is also a polynomial in n with non-negative coefficients given by

$$8\left(\left(1-\gamma^2\right)^2 n^4 + 4\left(1-\gamma^2\right) n^3 + 2\left(2-\gamma^2\right)\left(1+\gamma^2\right) n^2 + 4\gamma^2 n + \gamma^4\right).$$

Hence, for $n \to \infty$ the numerator as well as its first three derivatives, and also the denominator as well as its first three derivatives, tend to ∞. Applying

L'Hôpital's rule four times yields

$$\lim_{n\to\infty} \frac{(a-k)^2 n^2 (n+2) \left(2+n-\gamma-(n-1)\gamma^2\right)^2}{8\left(n(n+2)-(n-1)(n+1)\gamma^2\right)^2} - \frac{(a-k)^2 n}{8}$$
$$= \lim_{n\to\infty} \frac{(a-k)^2}{4(1+\gamma)}.$$

(v) For the network $(0,1,n-1)$ to be locally stable we need to have that firm 1 is not willing to form a link to an exclusive customer of firm 2. Moreover, firm 2 should not have an incentive to delete a link to neither a joint nor to an exclusive customer. This requires

$$\pi_1^{(0,1,n-1)} \geq \pi_1^{(0,2,n-2)} - c, \tag{5.15}$$
$$\pi_2^{(0,1,n-1)} - c \geq \pi_2^{(1,0,n-1)}, \quad \pi_2^{(0,1,n-1)} - c \geq \pi_2^{(0,1,n-2)}. \tag{5.16}$$

Inequality (5.15) gives us a lower bound for the costs of link formation and the inequalities in (5.16) define an upper bound. We use this conditions to establish a contradiction between the requirements needed for local stability. More precisely, we show that for all $\gamma \neq 0$ one of the upper bounds always exceeds the lower bound for the costs of link formation. We obtain

$$\pi_1^{(0,2,n-2)} - \pi_1^{(0,1,n-1)} = \frac{(a-k)^2 n^2 (2-\gamma)^2 \left(2\left(2-\gamma^2\right)^2 n^2 - \gamma^4\right)}{2\left(2n-(n-1)\gamma^2\right)^2 \left(4n-(2n-1)\gamma^2\right)^2}, \tag{5.17}$$

$$\pi_2^{(0,1,n-1)} - \pi_2^{(1,0,n-1)} = \frac{(a-k)^2 n \left(2n-\gamma-(n-1)\gamma^2\right)^2}{\left(4n-(2n-1)\gamma^2\right)^2} - \frac{(a-k)^2 (n-1)}{4}, \tag{5.18}$$

$$\pi_2^{(0,1,n-1)} - \pi_2^{(0,1,n-2)} \tag{5.19}$$
$$= \frac{(a-k)^2 n \left(2n-\gamma-(n-1)\gamma^2\right)^2}{\left(4n-(2n-1)\gamma^2\right)^2} \tag{5.20}$$
$$- \frac{(a-k)^2 (2-\gamma)^2 (n-1) \left(2(n-1)-\gamma-(n-2)\gamma^2\right)}{\left(4(n-1)-(2n-3)\gamma^2\right)^2}.$$

First note that the lower bound on the costs of link formation in equation (5.17) is increasing in n whereas the upper bounds in equations (5.18) and (5.21) are decreasing in n as

$$\frac{\partial\left[\pi_1^{(0,2,n-2)}-\pi_1^{(0,1,n-1)}\right]}{\partial n} = \frac{(a-k)^2\, n\,(2-\gamma)^2\left(6\left(2-\gamma^2\right)^3 n^3+6\gamma^2\left(2-\gamma^2\right)^3 n^2-\gamma^6\right)}{2\left(2n-(n-1)\,\gamma^2\right)^3\left(4n-(2n-1)\,\gamma^2\right)^3} \geq 0$$

and

$$\frac{\partial\pi_2^{(0,1,n-1)}}{\partial n} = \frac{(a-k)^2\left(\gamma+(n-1)\,\gamma^2-2n\right)\left(2\left(2-\gamma^2\right)^2 n^2-\gamma\,(2+\gamma)\left(2-\gamma^2\right)n\right)}{\left(4n-(2n-1)\,\gamma^2\right)^3} + \frac{(a-k)^2\left(\gamma+(n-1)\,\gamma^2-2n\right)\gamma^3\,(1-\gamma)}{\left(4n-(2n-1)\,\gamma^2\right)^3} \leq 0.$$

Thus, it suffices to consider small n to establish instability. For $n=3$ we have

$$\pi_1^{(0,2,1)}-\pi_1^{(0,1,2)} = \frac{9\,(a-k)^2\,(2-\gamma)^2\left(72-72\gamma^2+17\gamma^4\right)}{8\left(3-\gamma^2\right)^2\left(12-5\gamma^2\right)^2}, \tag{5.21}$$

$$\pi_2^{(0,1,2)}-\pi_2^{(1,0,2)} = \frac{(a-k)^2\left(72-72\gamma-18\gamma^2+24\gamma^3-\gamma^4\right)}{2\left(12-5\gamma^2\right)^2},$$

$$\pi_2^{(0,1,2)}-\pi_2^{(0,1,1)} = \frac{(a-k)^2\left(2304-3744\gamma^2+2284\gamma^4-20\gamma^5-607\gamma^6+8\gamma^7+58\gamma^8\right)}{\left(12-5\gamma^2\right)^2\left(8-3\gamma^2\right)^2},$$

$$\left(\pi_2^{(0,1,2)}-\pi_2^{(1,0,2)}\right)-\left(\pi_1^{(0,2,1)}-\pi_1^{(0,1,2)}\right) = -\frac{\gamma^2\,(a-k)^2\left(432-720\gamma^2+252\gamma^3+201\gamma^4-96\gamma^5+4\gamma^6\right)}{8\left(3-\gamma^2\right)^2\left(12-5\gamma^2\right)^2},$$

$$\left(\pi_2^{(0,1,2)}-\pi_2^{(0,1,1)}\right)-\left(\pi_1^{(0,2,1)}-\pi_1^{(0,1,2)}\right) = \frac{\gamma\,(a-k)^2\left(165888-131328\gamma-290304\gamma^2+248256\gamma^3\right)}{8\left(3-\gamma^2\right)^2\left(8-3\gamma^2\right)^2} + \frac{\gamma\,(a-k)^2\left(185472\gamma^4-177312\gamma^5-51168\gamma^6+59252\gamma^7\right)}{8\left(3-\gamma^2\right)^2\left(8-3\gamma^2\right)^2} + \frac{\gamma\,(a-k)^2\left(4964\gamma^8-9017\gamma^9+64\gamma^{10}+464\gamma^{11}\right)}{8\left(3-\gamma^2\right)^2\left(8-3\gamma^2\right)^2}.$$

Note that the polynomial in the numerator of the first equation does not possess a zero for $\gamma \in (0, 1]$ and is strictly positive for $\gamma = 1$.

Hence, the first equation is strictly negative for $\gamma \in (0, 1]$. The second equation is strictly negative for $\gamma \in [-1, 0)$. Thus, for complementary as well as for substitutable products there is always one lower bound imposed on the costs of link formation that exceeds an upper bound. This means that the network $(0, 1, 2)$ is locally unstable for all $\gamma \neq 0$. Also considering $\gamma = 0$ and inspecting the conditions we observe local instability for $c \neq \frac{(a-k)^2}{4}$.

(vi) The networks $(n_1^e, n^j, n_2^e) \in \mathcal{N}_n$ with $n^j = 0$ and $0 < n_1^e + n_2^e < n$ are locally unstable for $c \neq \frac{(a-k)^2}{4}$, as there is always a profitable deviation of one firm to either add or delete a link to an exclusive customer. Therefore, they may only be locally stable for very specific costs, i.e., $c = \frac{(a-k)^2}{4}$ (for all $\gamma \in [-1, 1]$).

(vii) For $\gamma = 0$ the profits of the intermediaries reduce to

$$\pi_1^{\left(n_1^e, n^j, n_2^e\right)} = \frac{\left(n_1^e + n^j\right)(a-k)^2}{4}, \quad \pi_2^{\left(n_1^e, n^j, n_2^e\right)} = \frac{\left(n^j + n_2^e\right)(a-k)^2}{4}.$$

Therefore, for $c = \frac{(a-k)^2}{4}$ both intermediaries are indifferent between deleting and adding links. ■

5.A.3 Proof of Proposition 5.3

Proof: We show that there is an overlap of the local stability regions of the complete network $(0, n, 0)$ and the empty network $(0, 0, 0)$. First note that the stability region for the network $(0, n, 0)$ shrinks as n grows. This is taking $\overline{c}^{(0,n,0)}$ from equation (5.12) and

$$\begin{aligned}
\frac{\partial \overline{c}^{(0,n,0)}}{\partial n} &= \frac{\partial \left[\frac{(a-k)^2 n\left(16n\left(1-\gamma^2\right)+(3n+1)\gamma^4\right)}{(2+\gamma)^2(4n-(n+1)\gamma^2)^2}\right]}{\partial n} \\
&= \frac{(a-k)^2 \gamma^2 \left(-32n + 28\gamma^2 n - \gamma^4 (5n+1)\right)}{(2+\gamma)^2 \left(4n - (n+1)\gamma^3\right)^3} \leq 0
\end{aligned}$$

(for $n \geq 1$). We investigate the difference of the limit of $\overline{c}^{(0,n,0)}$ and $\underline{c}^{(0,0,0)}$

$$\lim_{n\to\infty} \overline{c}^{(0,n,0)} - \underline{c}^{(0,0,0)} = \frac{\gamma\,(a-k)^2\,(-16-12\gamma+4\gamma^2+\gamma^3)}{4\,(2-\gamma)\,(2+\gamma)^3} \geq 0 \quad \text{for } \gamma \in [-1,0].$$

This implies that there is an overlap between the stability regions of the networks $(0,n,0)$ and $(0,0,0)$ for complementary or independent products. Note that for substitutable products the above inequality shows that the regions are not even directly adjacent. ■

5.A.4 Proof of Proposition 5.4

Suppose $\gamma \in [\gamma^n_{\min}, 1]$. The profits needed for the proof can be found in Table 5.2.

Table 5.2: Equilibrium profits for networks with n customers

(n^e_1, n^j, n^e_2)	$\pi_1^{(n^e_1,n^j,n^e_2)}$	$\pi_2^{(n^e_1,n^j,n^e_2)}$
$(0, n-n^e_2, n^e_2)$	$\frac{(a-k)^2 n^2 (2-\gamma)^2 \left(n-n^e_2\right)}{\left(4n-\gamma^2\left(n+n^e_2\right)\right)^2}$	$\frac{(a-k)^2 n \left(n(2-\gamma)+n^e_2\gamma(1-\gamma)\right)^2}{\left(4n-\gamma^2\left(n+n^e_2\right)\right)^2}$
$(0, n^j, n-n^j)$	$\frac{(a-k)^2 n^2 (2-\gamma)^2 n^j}{(4n-\gamma^2(2n-n^j))^2}$	$\frac{(a-k)^2 n \left(n(2-\gamma)+\left(n-n^j\right)\gamma(1-\gamma)\right)^2}{(4n-\gamma^2(2n-n^j))^2}$

Proof: (i) Consider the complete network with $(0,n,0)$. For the Nash stability, we compare the according profits with those of a network $(0, n-n^e_2, n^e_2)$ in which n^e_2 links have been deleted by firm 1. These are all the networks in $\Pi^{n^e_1}(n)$ The network $(0,n,0)$ is Nash stable if none of the firms is willing to delete a link. This requires for firm 1 $\pi_1^{(0,n,0)} - \pi_1^{\left(0,n-n^e_2,n^e_2\right)} \geq n^e_2 c$ for all $0 < n^e_2 \leq n$.

Using the profits from Table 5.2 we obtain for

$$\frac{\pi_1^{(0,n,0)} - \pi_1^{\left(0,n-n^e_2,n^e_2\right)}}{n^e_2} = \frac{(a-k)^2\, n\,((4-\gamma^2)\,(4-3\gamma^2)\,n + \gamma^4 n^e_2)}{(2+\gamma)^2\,(4n-\gamma^2\,(n+n^e_2))^2}$$

and observe that this difference is increasing in n_2^e by noting that

$$\frac{\partial\left[(a-k)^2 n\left(\left(4-\gamma^2\right)\left(4-3\gamma^2\right)n+\gamma^4 n_2^e\right)\right]}{\partial n_2^e}=(a-k)^2 n\gamma^4 \geq 0$$

$$\frac{\partial\left[n_2^e(2+\gamma)^2\left(4n-\gamma^2\left(n+n_2^e\right)\right)^2\right]}{\partial n_2^e}=-2\gamma^2(2+\gamma)^2\left(4n-\gamma^2\left(n+n_2^e\right)\right)\leq 0$$

and applying the quotient rule. Thus, for the Nash stability region we require for firm 1

$$\min_{0<n_2^e\leq n}\frac{\pi_1^{(0,n,0)}-\pi_1^{\left(0,n-n_2^e,n_2^e\right)}}{n_2^e}\geq c.$$

Using the observation that this difference is increasing in n_2^e this minimum is attained at $n_2^e=1$. This means the network that defines the Nash stability region for the network $(0,n,0)$ considering deviations of firm 1 is the the directly neighboring network $(0,n-1,1)$. Thus, by symmetry of the firms for the network $(0,n,0)$ the stability regions for local and Nash stability coincide.

Consider the empty network with $(0,0,0)$. There is no incentive of firm 1 to add further links if $\pi_1^{(0,0,0)}-\pi_1^{\left(n_1^e,0,0\right)}\geq -n_1^e c$. We obtain for

$$\frac{\pi_1^{(0,0,0)}-\pi_1^{\left(n_1^e,0,0\right)}}{n_1^e}=\frac{0-(a-k)^2 n_1^e}{-4n_1^e}=\frac{(a-k)^2}{4}.$$

which implies for the Nash stability region considering deviations of firm 1 $c\geq\frac{(a-k)^2}{4}$. This means that the change of marginal profits is constant between all networks in $\Pi^{n_1^e}(0)$. Thus, by symmetry of the firms for the network $(0,0,0)$ the stability regions for local and Nash stability coincide.

Consider the network with $(0,0,n)$. This network is asymmetric. Therefore, we have consider the conditions imposed on Nash stability separately for both firms. Firm 1 may add links to customers of firm 2. This is we have to establish for the Nash stability of the network $(0,0,n)$ considering deviations of firm 1 $\pi_1^{\left(0,n^j,n-n^j\right)}-\pi_1^{(0,0,n)}\leq n^j c$ for all $0<n^j\leq n$. We look at

$$\frac{\pi_1^{\left(0,n^j,n-n^j\right)}-\pi_1^{(0,0,n)}}{n^j}=\frac{(a-k)^2 n^2(2-\gamma)^2 n^j}{n^j\left(4n-\gamma^2\left(2n-n^j\right)\right)^2}=\frac{(a-k)^2 n^2(2-\gamma)^2}{\left(4n-\gamma^2\left(2n-n^j\right)\right)^2}$$

and observe that this difference is decreasing in n^j by noting that

$$\frac{\partial\left[\frac{(a-k)^2 n^2 (2-\gamma)^2}{(4n-\gamma^2(2n-n^j))^2}\right]}{\partial n^j} = -\frac{2\gamma^2 (a-k)^2 n^2 (2-\gamma)^2}{\left(4n-\gamma^2 (2n-n^j)\right)^3} \leq 0.$$

For the Nash stability region we require for firm 1

$$\max_{0<n^j\leq n} \frac{\pi_1^{\left(0,n^j,n-n^j\right)} - \pi_1^{(0,0,n)}}{n^j} \leq c$$

and this maximum is attained at $n^j = 1$, which is the network $(0, 1, n-1)$. Firm 2 already has n links in the network $(0, 0, n)$. Therefore, firm 2 may only delete links. As already observed the change of marginal profits is constant between all networks in $\Pi^{n^j}(0)$, we require for the Nash stability for the network $(0, 0, n)$ for firm 2 $c \leq \frac{(a-k)^2}{4}$.

Summing up the Nash stability region for the network $(0, 0, n)$ is determined by deviations to the directly adjacent networks $(0, 1, n-1)$ for firm 1 and $(0, 0, n-1)$ for firm 2. Thus, for the networks $(0, 0, n)$ and $(n, 0, 0)$ local and Nash stability coincide.

(ii) For $\gamma = 0$ and $c = \frac{(a-k)^2}{4}$ we have already established in Proposition 5.2 (vii) that the firms are indifferent between adding and deleting links and it is straightforward to see that all networks are Nash stable.

(iii) Consider the network $(n_1^e, n^j, n_2^e) = (1, 1, 1)$ and $\gamma \in (-1, 0)$. For local stability we require

$$c \in \left[\frac{(a-k)^2 \left(576 + 288\gamma - 660\gamma^2 - 264\gamma^3 + 184\gamma^4 + 52\gamma^5 - 5\gamma^6\right)}{16\left(3-\gamma^2\right)^2 (4+3\gamma)^2}, \frac{2(a-k)^2 \left(2-\gamma^2\right)\left(64 + 96\gamma + 12\gamma^2 - 36\gamma^3 - 9\gamma^4\right)}{(4+3\gamma)^2 \left(8-3\gamma^2\right)^2}\right].$$

The difference between the upper and lower bound is equal to $(a-k)^2 \gamma$ multiplied by

$$\frac{36864+33792\gamma-54528\gamma^2-45504\gamma^3+30176\gamma^4}{16\left(3-\gamma^2\right)^2\left(4+3\gamma\right)^2\left(8-3\gamma^2\right)^2} + \frac{22164\gamma^5-7416\gamma^6-4584\gamma^7+684\gamma^8+333\gamma^9}{16\left(3-\gamma^2\right)^2\left(4+3\gamma\right)^2\left(8-3\gamma^2\right)^2}$$

and has just one zero for $\gamma \in (-1,0)$ at $\gamma \approx -0.95$. Noting that for $\gamma = -1$ this expression is positive, we know that the network $(1,1,1)$ is locally stable for $\gamma \in [-1, 0.95]$. This can also be seen directly looking at Figure 5.6c. However, the network $(1,1,1)$ is not Nash stable for $\gamma \in (-1, 0.95]$ as

$$\begin{aligned}\pi_1^{(1,1,1)} - \pi_1^{(0,2,0)} &= \frac{2\left(a-k\right)^2\left(2+\gamma\right)^2}{\left(4+3\gamma\right)^2} - \frac{2\left(a-k\right)^2}{\left(2+\gamma\right)^2} \\ &= \frac{2\gamma\left(a-k\right)^2\left(1+\gamma\right)\left(8+7\gamma+\gamma^2\right)}{\left(4+3\gamma\right)^2\left(2+\gamma\right)^2} < 0\end{aligned}$$

for $\gamma \in (-1,0)$ and firm 1 has an incentive to first delete a link to an exclusive customer and then to establish a link to a customer of firm 2 yielding the network $(0,2,0)$. ■

5.A.5 Proof of Proposition 5.5

Proof: Among the directly neighboring networks of $(0,n,0)$, $(0,0,0)$, $(n,0,0)$ and $(0,0,n)$, there is no network in which condition (5.4) is not satisfied. Therefore, we may immediately consider $\gamma \in [-1,1]$. However, for the network $\left(\frac{n}{2},0,\frac{n}{2}\right)$, the two neighboring networks $\left(\frac{n}{2},1,\frac{n}{2}-1\right)$ and $\left(\frac{n}{2}-1,1,\frac{n}{2}\right)$ may violate condition (5.4). For the two networks $\left(\frac{n}{2},1,\frac{n}{2}-1\right)$ and $\left(\frac{n}{2}-1,1,\frac{n}{2}\right)$ condition (5.4) reduces to

$$\frac{1}{2}\left[n\left(n+2+\gamma\right)-\gamma^2 n\left(n+1\right)+2\gamma\right] \tag{5.22}$$

with a zero for $\gamma \in [-1,1]$ of $-\frac{\sqrt{4n^4+12n^3+9n^2+4n+4}-n-2}{2n(n+1)}$. We have that (5.22) is increasing in n as

$$\frac{(n+1)\left[(n+3+\gamma)-\gamma^2(n+2)\right]+2\gamma-\left[n(n+2+\gamma)-\gamma^2 n(n+1)+2\gamma\right]}{2}$$
$$=\frac{(1+\gamma)(2n(1-\gamma)+3-2\gamma)}{2}\geq 0 \quad \text{for } n\geq 4 \text{ and } \gamma\in[-1,1].$$

Moreover, (5.22) as a function in γ is a parabola that opens downward with a maximum at $\gamma=\frac{n+2}{2n(n+1)}$ as

$$\frac{\partial\left[\frac{1}{2}\left[n(n+2+\gamma)-\gamma^2 n(n+1)+2\gamma\right]\right]}{\partial\gamma}=\frac{n-2\gamma n(n+1)+2}{2},$$
$$\frac{\partial^2\left[\frac{1}{2}\left[n(n+2+\gamma)-\gamma^2 n(n+1)+2\gamma\right]\right]}{\partial\gamma^2}=-n(n+1)<0.$$

These two observations imply that the lower bound for γ over all networks with $n\geq 4$ is at $n=4$ for the networks $(1,1,2)$ and $(2,1,1)$ given by $-\frac{\sqrt{3}\sqrt{163}-3}{20}\approx -0.96$. ■

5.A.6 Proof of Proposition 5.6

For this proof we use the lower and upper bounds that local stability imposes on the costs of link formation.

Proof: (i) A symmetric network that is locally stable for appropriately chosen costs of link formation is the network $\left(\mathbf{n_1^e},\mathbf{n^j},\mathbf{n_2^e}\right)=(\mathbf{1},\mathbf{1},\mathbf{1})$. For local stability we require

$$\pi_1^{(1,1,1)}\geq \pi_1^{(1,2,0)}-c,\quad \pi_1^{(1,1,1)}-c\geq \pi_1^{(1,0,2)},\quad \pi_1^{(1,1,1)}-c\geq \pi_1^{(0,1,1)}.$$

Note that for symmetry reasons we just investigate deviations of firm 1. Thus, for $\gamma\in[0,1]$ local stability imposes the costs of link formation

$$c\in\left[\frac{(a-k)^2\left(576+288\gamma-660\gamma^2-264\gamma^3+184\gamma^4+52\gamma^5-5\gamma^6\right)}{16\left(3-\gamma^2\right)^2(4+3\gamma)^2},\right.$$
$$\left.\frac{(a-k)^2\left(16+8\gamma-\gamma^2\right)}{4(4+3\gamma)^2}\right].$$

The difference between the upper and the lower bound imposed for local stability on the costs of link formation for $\gamma \in [0,1]$ is

$$\begin{aligned}
&\frac{(a-k)^2 \left(16+8\gamma-\gamma^2\right)}{4\,(4+3\gamma)^2} \\
&\qquad - \frac{(a-k)^2 \left(576+288\gamma-660\gamma^2-264\gamma^3+184\gamma^4+52\gamma^5-5\gamma^6\right)}{16\left(3-\gamma^2\right)^2 (4+3\gamma)^2} \\
=&\frac{(a-k)^2 \gamma^2 \left(240+72\gamma-96\gamma^2-20\gamma^3+\gamma^4\right)}{16\left(3-\gamma^2\right)^2 (4+3\gamma)^2} \geq 0.
\end{aligned}$$

For an asymmetric network consider $\left(\mathbf{n_1^e}, \mathbf{n^j}, \mathbf{n_2^e}\right) = (\mathbf{3}, \mathbf{0}, \mathbf{0})$. For local stability we require

$$\pi_1^{(0,0,3)} \geq \pi_1^{(0,1,2)} - c, \quad \pi_2^{(0,0,3)} - c \geq \pi_2^{(0,0,2)}.$$

Thus, local stability imposes the costs of link formation

$$c \in \left[\frac{9\,(a-k)^2 (2-\gamma)^2}{\left(12-5\gamma^2\right)^2}, \frac{(a-k)^2}{4}\right].$$

The difference between the upper and the lower bound imposed for local stability on the costs of link formation is

$$\frac{(a-k)^2}{4} - \frac{9\,(a-k)^2 (2-\gamma)^2}{\left(12-5\gamma^2\right)^2} = \frac{(a-k)^2 \gamma\,(6-5\gamma)\left(24-6\gamma-5\gamma^2\right)}{4\left(12-5\gamma^2\right)^2} \geq 0$$

for $\gamma \in [0,1]$.

(ii) Suppose $n_1^e < n_2^e$. Then, asymmetric networks are

$$\left(n_1^e, n^j, n_2^e\right) \in \{(0,0,1), (0,0,2), (0,1,1), (0,0,3), (0,1,2), (0,2,1), (1,0,2)\}.$$

The networks $(0,0,1)\,/\,(0,0,2)$ are always locally unstable for $c \neq \frac{(a-k)^2}{4}$. There is always a profitable deviation of firm 2 to either add or delete a link to an exclusive customer. The local instability of the network $(0,1,2)$ for $\gamma \neq 0$ and for $\gamma = 0$ and $c \neq \frac{(a-k)^2}{4}$ has already been established in Proposition 5.2(v) for n customers. From (i) we immediately observe that for the network $(0,0,3)$ the interval for the conditions of local stability is empty for

$\gamma \in [-1, 0)$. For the network $(0, 1, 1)$ to be locally stable we require for the costs of link formation

$$c \geq \max\left\{\pi_1^{(1,1,1)} - \pi_1^{(0,1,1)}, \pi_1^{(0,2,0)} - \pi_1^{(0,1,1)}, \pi_2^{(0,1,2)} - \pi_2^{(0,1,1)}\right\},$$
$$c \leq \min\left\{\pi_1^{(0,1,1)} - \pi_1^{(0,0,2)}, \pi_2^{(0,1,1)} - \pi_2^{(0,1,0)}, \pi_2^{(0,1,1)} - \pi_2^{(1,0,1)}\right\}.$$

However, these conditions imposed on the costs of link formation cannot hold simultaneously for complementary products as

$$\left(\pi_1^{(0,2,0)} - \pi_1^{(0,1,1)}\right) - \left(\pi_2^{(0,1,1)} - \pi_2^{(0,1,0)}\right)$$
$$= -\frac{(a-k)^2 \gamma \left(64 + 24\gamma - 56\gamma^2 - 13\gamma^3 + 12\gamma^4 + 2\gamma^5\right)}{(2+\gamma)^2 \left(8 - 3\gamma^2\right)^2} > 0 \quad \text{for } \gamma \in [-1, 0).$$

For the network $(0, 2, 1)$ to be locally stable we require for the costs of link formation

$$c \geq \pi_1^{(0,3,0)} - \pi_1^{(0,2,1)},$$
$$c \leq \min\left\{\pi_1^{(0,2,1)} - \pi_1^{(0,1,2)}, \pi_2^{(0,2,1)} - \pi_2^{(1,1,1)}, \pi_2^{(0,2,1)} - \pi_2^{(0,2,0)}\right\}.$$

However, these conditions imposed on the costs of link formation cannot hold simultaneously for complementary products as

$$\left(\pi_1^{(0,3,0)} - \pi_1^{(0,2,1)}\right) - \left(\pi_2^{(0,2,1)} - \pi_2^{(0,2,0)}\right)$$
$$= -\frac{(a-k)^2 \gamma \left(144 + 60\gamma - 120\gamma^2 - 26\gamma^3 + 24\gamma^4 + 3\gamma^5\right)}{16\,(2+\gamma)^2 \left(3 - \gamma^2\right)^2} > 0 \quad \text{for } \gamma \in [-1, 0).$$

For the network $(1, 0, 2)$ to be locally stable we require for the costs of link formation

$$c \geq \max\left\{\pi_1^{(1,1,1)} - \pi_1^{(1,0,2)}, \pi_2^{(0,1,2)} - \pi_2^{(1,0,2)}\right\},$$
$$c \leq \min\left\{\pi_1^{(1,0,2)} - \pi_1^{(0,0,2)}, \pi_2^{(1,0,2)} - \pi_2^{(1,0,1)}\right\}.$$

However, these conditions imposed on the costs of link formation cannot hold simultaneously for complementary products as

$$\left(\pi_1^{(1,1,1)} - \pi_1^{(1,0,2)}\right) - \left(\pi_1^{(1,0,2)} - \pi_1^{(0,0,2)}\right) = -\frac{(a-k)^2 \gamma (8 + 5\gamma)}{2\,(4 + 3\gamma)^2} > 0$$

for $\gamma \in [-1, 0)$. Examples for the second claim are the networks $(0, 2, 0)$ and $(0, 1, 0)$. ■

5.A.7 Proof of Proposition 5.7

Proof: Table 5.3 summarizes the networks imposing conditions on the costs of link formation. For complementary products this has already be shown in Proposition 5.3. For substitutable products with $\gamma \in (0, 1]$, the statement of Proposition 5.7 can be directly seen from Table 5.3. The according networks are marked in bold. The non-emptiness of the adjacency regions can be seen in Figures 5.6a, 5.6c, 5.6g, 5.6h and 5.6b. As the stability region in Figure 5.6g is hardly visible, Figure 5.8 redraws the area for $\gamma \in [0.9, 1]$. ■

Table 5.3: Networks imposing conditions on the costs of link formation

	networks defining	
(n_1^e, n^j, n_2^e)	$\underline{c}^{(n_1^e, n^j, n_2^e)}$	$\overline{c}^{(n_1^e, n^j, n_2^e)}$
$(\mathbf{0}, \mathbf{3}, \mathbf{0})$	—	$(1, 2, 0)$, $(0, 2, 1)$
$(\mathbf{1}, \mathbf{1}, \mathbf{1})$	$(1, 2, 0)$, $(0, 2, 1)$	$(1, 1, 0)$, $(0, 1, 1)$ (for $\gamma \in [-1, 0]$), $(2, 0, 1)$, $(1, 0, 2)$ (for $\gamma \in [0, 1]$)
$(0, 2, 0)$	$(1, 2, 0)$, $(0, 2, 1)$	$(1, 1, 0)$, $(0, 1, 1)$
$(0, 1, 0)$	$(1, 1, 0)$, $(0, 1, 1)$	$(1, 0, 0)$, $(0, 0, 1)$
$(\mathbf{1}, \mathbf{2}, \mathbf{0})$	$(0, 3, 0)$	$(0, 2, 0)$ (for $\gamma \in [-1, 0]$), $(1, 1, 1)$ (for $\gamma \in [0, 1]$)
$(\mathbf{0}, \mathbf{2}, \mathbf{1})$	$(0, 3, 0)$	$(0, 2, 0)$ (for $\gamma \in [-1, 0]$), $(1, 1, 1)$ (for $\gamma \in [0, 1]$)
$(\mathbf{2}, \mathbf{0}, \mathbf{1})$	$(1, 1, 1)$	$(2, 0, 0)$, $(1, 0, 1)$
$(\mathbf{1}, \mathbf{0}, \mathbf{2})$	$(1, 1, 1)$	$(0, 0, 2)$, $(1, 0, 1)$
$(3, 0, 0)$	$(2, 1, 0)$	$(2, 0, 0)$
$(0, 0, 3)$	$(0, 1, 2)$	$(0, 0, 2)$
$(\mathbf{0}, \mathbf{0}, \mathbf{0})$	$(1, 0, 0)$, $(0, 0, 1)$	—

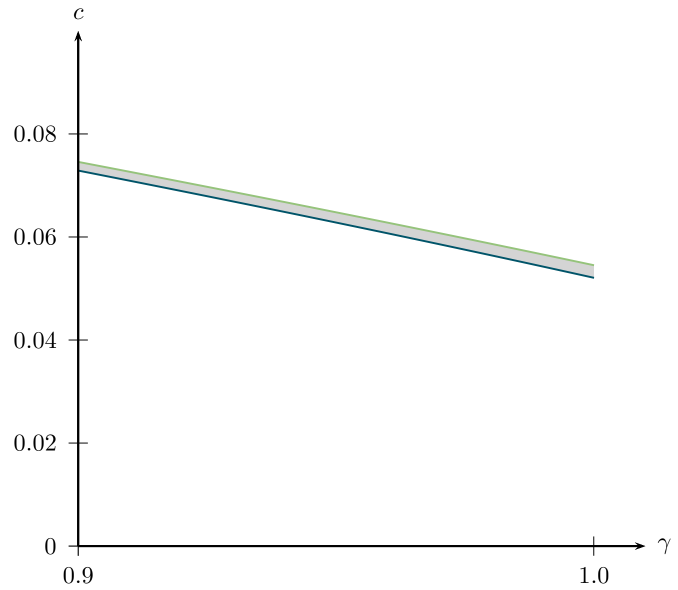


Figure 5.8: Local stability of $(0,2,1)$, $(1,2,0)$ for $\gamma \in [0.9,1]$ $(a-k=1)$

5.B Appendix B: Conditions on the costs of link formation for local stability for networks with three customers

5.B.1 Equilibrium profits

The equilibrium profits are summarized in Table 5.4. Indeed, from an individual perspective and with no costs of link formation the complete network with three joint customers yields the highest payoff for firm 1 for complementary products, whereas this is the network with three own exclusive customers for substitutable products.

5.B.2 Symmetric networks

Because of symmetry we just consider within this subsection deviations of firm 1.

- $\left(\mathbf{n_1^e}, \mathbf{n^j}, \mathbf{n_2^e}\right) = (\mathbf{0}, \mathbf{3}, \mathbf{0})$
  For local stability we require $\pi_1^{(0,3,0)} - \pi_1^{(0,2,1)} \geq c$. Using inequality (5.12) from

Table 5.4: Equilibrium profits of firm 1 for networks with three customers

(n_1^e, n^j, n_2^e)	$\pi_1^{(n_1^e, n^j, n_2^e)}$	(n_1^e, n^j, n_2^e)	$\pi_1^{(n_1^e, n^j, n_2^e)}$
$(0,3,0)$	$\frac{3(a-k)^2}{(2+\gamma)^2}$	$(1,0,1)$	$\frac{(a-k)^2}{4}$
$(1,1,1)$	$\frac{2(a-k)^2(2+\gamma)^2}{(4+3\gamma)^2}$	$(2,0,1)$	$\frac{2(a-k)^2}{4}$
$(0,2,0)$	$\frac{2(a-k)^2}{(2+\gamma)^2}$	$(1,0,2)$	$\frac{(a-k)^2}{4}$
$(0,1,0)$	$\frac{(a-k)^2}{(2+\gamma)^2}$	$(3,0,0)$	$\frac{3(a-k)^2}{4}$
$(1,2,0)$	$\frac{3(a-k)^2\left(6-2\gamma-\gamma^2\right)^2}{16(3-\gamma^2)^2}$	$(2,0,0)$	$\frac{2(a-k)^2}{4}$
$(2,1,0)$	$\frac{3(a-k)^2(2+\gamma)^2(3-2\gamma)^2}{(12-5\gamma^2)^2}$	$(1,0,0)$	$\frac{(a-k)^2}{4}$
$(1,1,0)$	$\frac{2(a-k)^2\left(4-\gamma-\gamma^2\right)^2}{(8-3\gamma^2)^2}$	$(0,0,3)$	0
$(0,2,1)$	$\frac{18(a-k)^2(2-\gamma)^2}{16(3-\gamma^2)^2}$	$(0,0,2)$	0
$(0,1,2)$	$\frac{9(a-k)^2(2-\gamma)^2}{(12-5\gamma^2)^2}$	$(0,0,1)$	0
$(0,1,1)$	$\frac{4(a-k)^2(2-\gamma)^2}{(8-3\gamma^2)^2}$	$(0,0,0)$	0

the poof of Proposition 5.2 we obtain

$$\begin{aligned}\pi_1^{(0,3,0)} - \pi_1^{(0,2,1)} &= \frac{3\,(a-k)^2}{(2+\gamma)^2} - \frac{9\,(a-k)^2\,(2-\gamma^2)^2}{8\,(3-\gamma^2)^2} \\ &= \frac{3\,(a-k)^2\,(24\,(1-\gamma^2)+5\gamma^4)}{8\,(3-\gamma^2)^2\,(2+\gamma)^2}.\end{aligned}$$

- $\left(\mathbf{n_1^e}, \mathbf{n^j}, \mathbf{n_2^e}\right) = (\mathbf{0}, \mathbf{0}, \mathbf{0})$

  For local stability we require $\pi_1^{(0,0,0)} - \pi_1^{(1,0,0)} \geq c$ which is $c \geq \frac{(a-k)^2}{4}$.

- $\left(\mathbf{n_1^e}, \mathbf{n^j}, \mathbf{n_2^e}\right) = (\mathbf{1}, \mathbf{1}, \mathbf{1})$

  Firm 1 may add a link to an exclusive customer of firm 2 delete a link to a joint customer or to an own exclusive customer. For local stability we require

$$\pi_1^{(1,1,1)} \geq \pi_1^{(1,2,0)} - c, \tag{5.23}$$

$$\pi_1^{(1,1,1)} - c \geq \pi_1^{(1,0,2)}, \quad \pi_1^{(1,1,1)} - c \geq \pi_1^{(0,1,1)}. \tag{5.24}$$

Inequality (5.23) defines a lower bound for the costs of link formation

$$c \geq \frac{(a-k)^2\left(576+288\gamma-660\gamma^2-264\gamma^3+184\gamma^4+52\gamma^5-5\gamma^6\right)}{16\left(3-\gamma^2\right)^2(4+3\gamma)^2},$$

and the inequalities in (5.24) give an upper bound. First note

$$\max\left\{\pi_1^{(1,0,2)},\pi_1^{(0,1,1)}\right\}=\begin{cases}\frac{4(a-k)^2(2-\gamma)^2}{(8-3\gamma^2)^2} & \text{for } \gamma\in[-1,0],\\ \frac{(a-k)^2}{4} & \text{for } \gamma\in[0,1].\end{cases}$$

This implies

$$\begin{aligned} c &\leq \pi_1^{(1,1,1)}-\max\left\{\pi_1^{(1,0,2)},\pi_1^{(0,1,1)}\right\}\\ &=\begin{cases}\frac{2(a-k)^2\left(2-\gamma^2\right)\left(64+96\gamma+12\gamma^2-36\gamma^3-9\gamma^4\right)}{(4+3\gamma)^2(8-3\gamma^2)^2} & \text{for } \gamma\in[-1,0],\\ \frac{(a-k)^2\left(16+8\gamma-\gamma^2\right)}{4(4+3\gamma)^2} & \text{for } \gamma\in[0,1].\end{cases}\end{aligned}$$

- $\left(\mathbf{n_1^e},\mathbf{n^j},\mathbf{n_2^e}\right)=(\mathbf{0},\mathbf{2},\mathbf{0})$
  Firm 1 may add a link to obtain an exclusive customer or has the possibility to delete a link to a joint customer. Thus, for local stability we require

$$\pi_1^{(0,2,0)} \geq \pi_1^{(1,2,0)}-c, \tag{5.25}$$
$$\pi_1^{(0,2,0)}-c \geq \pi_1^{(0,1,1)}. \tag{5.26}$$

  Inequality (5.25) defines a lower bound for the costs of link formation

$$c \geq \frac{(a-k)^2\left(144+144\gamma-84\gamma^2-120\gamma^3+4\gamma^4+24\gamma^5+3\gamma^6\right)}{16\left(3-\gamma^2\right)^2(2+\gamma)^2},$$

  and inequality (5.26) gives an upper bound

$$c \leq \frac{2(a-k)^2\left(32-32\gamma^2+7\gamma^4\right)}{(2+\gamma)^2\left(8-3\gamma^2\right)^2}.$$

- $\left(\mathbf{n_1^e},\mathbf{n^j},\mathbf{n_2^e}\right)=(\mathbf{0},\mathbf{1},\mathbf{0})$
  Firm 1 may add a link to obtain an exclusive customer or firm 1 has the possibility to delete a link to a joint customer. Thus, for local stability we

require

$$\pi_1^{(0,1,0)} \geq \pi_1^{(1,1,0)} - c, \tag{5.27}$$

$$\pi_1^{(0,1,0)} - c \geq \pi_1^{(0,0,1)}. \tag{5.28}$$

Inequality (5.27) defines a lower bound for the costs of link formation

$$c \geq \frac{(a-k)^2 \left(64 + 64\gamma - 40\gamma^2 - 56\gamma^3 + \gamma^4 + 12\gamma^5 + 2\gamma^6\right)}{\left(8 - 3\gamma^2\right)^2 (2+\gamma)^2},$$

and inequality (5.28) gives an upper bound $c \leq \frac{(a-k)^2}{(2+\gamma)^2}$.

- $\left(\mathbf{n_1^e}, \mathbf{n^j}, \mathbf{n_2^e}\right) = (\mathbf{1}, \mathbf{0}, \mathbf{1})$
  For $n^j = 0$ the profits from Figure 5.4 are linearly increasing if a link is added by firm 1 and are linearly decreasing if a link is deleted. Therefore, if the costs of link formation are sufficiently small, $c < \frac{(a-k)^2}{4}$, then a further link is added, and for $c > \frac{(a-k)^2}{4}$ a link is deleted.

5.B.3 Asymmetric networks

- $\left(\mathbf{n_1^e}, \mathbf{n^j}, \mathbf{n_2^e}\right) \in \{(\mathbf{0}, \mathbf{0}, \mathbf{3}), (\mathbf{3}, \mathbf{0}, \mathbf{0})\}$
  For local stability we require

$$\pi_1^{(0,0,3)} \geq \pi_1^{(0,1,2)} - c, \tag{5.29}$$

$$\pi_2^{(0,0,3)} - c \geq \pi_2^{(0,0,2)}. \tag{5.30}$$

  From the poof of Proposition 5.2 we obtain $c \geq \frac{9(a-k)^2(2-\gamma)^2}{(12-5\gamma^2)^2}$ and from inequality (5.30) $c \leq \frac{(a-k)^2}{4}$.

- $\left(\mathbf{n_1^e}, \mathbf{n^j}, \mathbf{n_2^e}\right) \in \{(\mathbf{0}, \mathbf{2}, \mathbf{1}), (\mathbf{1}, \mathbf{2}, \mathbf{0})\}$
  Consider the network $(0, 2, 1)$. Possible deviations are for firm 1 to add or to delete link to a joint customer and for firm 2 to delete a link either to a joint or to an exclusive customer. For local stability of the network $(0, 2, 1)$

we therefore require for the costs of link formation

$$\pi_1^{(0,2,1)} \geq \pi_1^{(0,3,0)} - c, \tag{5.31}$$

$$\pi_1^{(0,2,1)} - c \geq \pi_1^{(0,1,2)}, \quad \pi_2^{(0,2,1)} - c \geq \pi_2^{(1,1,1)}, \quad \pi_2^{(0,2,1)} - c \geq \pi_2^{(0,2,0)}. \tag{5.32}$$

Inequality (5.31) poses an lower bound on the costs of link formation given by

$$c \geq \pi_1^{(0,3,0)} - \pi_1^{(0,2,1)} = \frac{3\,(a-k)^2\,(24 - 24\gamma^2 + 5\gamma^4)}{8\,(3-\gamma^2)^2\,(2+\gamma)^2}.$$

The inequalities in (5.32) give us a upper bound. Thus, for local stability of $(0,2,1)$ we need to have

$$\begin{aligned} c &\leq \min\left\{\pi_1^{(0,2,1)} - \pi_1^{(0,1,2)}, \pi_2^{(0,2,1)} - \pi_2^{(1,1,1)}, \pi_2^{(0,2,1)} - \pi_2^{(0,2,0)}\right\} \\ &= \begin{cases} \frac{(a-k)^2\left(144+144\gamma-84\gamma^2-120\gamma^3+4\gamma^4+24\gamma^5+3\gamma^6\right)}{16(3-\gamma^2)^2(2+\gamma)^2} & \text{for } \gamma \in [-1,0], \\ \frac{(a-k)^2\left(576+288\gamma-660\gamma^2-264\gamma^3+184\gamma^4+52\gamma^5-5\gamma^6\right)}{16(3-\gamma^2)^2(4+3\gamma)^2} & \text{for } \gamma \in [0,1]. \end{cases} \end{aligned}$$

- $\left(\mathbf{n_1^e}, \mathbf{n^j}, \mathbf{n_2^e}\right) \in \{(\mathbf{1},\mathbf{0},\mathbf{2}), (\mathbf{2},\mathbf{0},\mathbf{1})\}$

Consider the network $(1,0,2)$. Possible deviations are for firm 1 to add a link to an exclusive customer of firm 2 or to delete a link to an own exclusive customer. firm 2 has the analogous deviation possibilities. For local stability of the network $(1,0,2)$ we therefore require for the costs of link formation

$$\pi_1^{(1,0,2)} \geq \pi_1^{(1,1,1)} - c, \quad \pi_2^{(1,0,2)} \geq \pi_2^{(0,1,2)} - c, \tag{5.33}$$

$$\pi_1^{(1,0,2)} - c \geq \pi_1^{(0,0,2)}, \quad \pi_2^{(1,0,2)} - c \geq \pi_2^{(1,0,1)}. \tag{5.34}$$

The inequalities in (5.33) pose a lower bound on the costs of link formation given by

$$c \geq \max\left\{\pi_1^{(1,1,1)} - \pi_1^{(1,0,2)}, \pi_2^{(0,1,2)} - \pi_2^{(1,0,2)}\right\} = \frac{(a-k)^2\,(16+8\gamma-\gamma^2)}{4\,(4+3\gamma)^2}.$$

The inequalities in (5.34) pose an upper bound on the costs of link formation given by

$$c \leq \min\left\{\pi_1^{(1,0,2)} - \pi_1^{(0,0,2)}, \pi_2^{(1,0,2)} - \pi_2^{(1,0,1)}\right\} = \frac{(a-k)^2}{4}.$$

- $\left(\mathbf{n_1^e}, \mathbf{n^j}, \mathbf{n_2^e}\right) \in \{(\mathbf{0}, \mathbf{1}, \mathbf{1}), (\mathbf{1}, \mathbf{1}, \mathbf{0})\}$
  Consider the network $(0, 1, 1)$. Possible deviations are for firm 1 to add a link to an own exclusive customer, or to an exclusive customer of firm 2, or to delete a link to a joint customer. firm 2 may add a link to an own exclusive customer or delete a link to either an own exclusive or a joint customer. For local stability of the network $(0, 1, 1)$ we therefore require for the costs of link formation

$$\pi_1^{(0,1,1)} \geq \pi_1^{(1,1,1)} - c, \quad \pi_1^{(0,1,1)} \geq \pi_1^{(0,2,0)} - c, \quad \pi_2^{(0,1,1)} \geq \pi_2^{(0,1,2)} - c, \tag{5.35}$$
$$\pi_1^{(0,1,1)} - c \geq \pi_1^{(0,0,2)}, \quad \pi_2^{(0,1,1)} - c \geq \pi_2^{(0,1,0)}, \quad \pi_2^{(0,1,1)} - c \geq \pi_2^{(1,0,1)}. \tag{5.36}$$

  The inequalities in (5.35) pose a lower bound on the costs of link formation given by

$$c \geq \max\left\{\pi_1^{(1,1,1)} - \pi_1^{(0,1,1)}, \pi_1^{(0,2,0)} - \pi_1^{(0,1,1)}, \pi_2^{(0,1,2)} - \pi_2^{(0,1,1)}\right\}.$$

  The inequalities in (5.36) pose an upper bound on the costs of link formation given by

$$c \leq \min\left\{\pi_1^{(0,1,1)} - \pi_1^{(0,0,2)}, \pi_2^{(0,1,1)} - \pi_2^{(0,1,0)}, \pi_2^{(0,1,1)} - \pi_2^{(1,0,1)}\right\}.$$

  In the proof of Proposition 5.6 we have already shown that the network $(0, 1, 1)$ is locally unstable for $\gamma \in (0, 1]$. We now show that it is also locally unstable for $\gamma \in (0, 1]$. Figure 5.9 graphically shows these conditions imposed by local stability on the costs of link formation.

  To see this, consider

$$\begin{aligned}&\left(\pi_1^{(1,1,1)} - \pi_1^{(0,1,1)}\right) - \left(\pi_1^{(0,1,1)} - \pi_1^{(0,0,2)}\right)\\ &\quad = \frac{2\,(a-k)^2\,\gamma\,(4-3\gamma^2)\,(32+12\gamma-12\gamma^2-3\gamma^3)}{(4+3\gamma)^2\,(8-3\gamma^2)^2} > 0\end{aligned}$$

  for $\gamma \in (0, 1]$. This means that the conditions imposed on the costs of link formation cannot hold simultaneously for substitutable products with $\gamma \in$

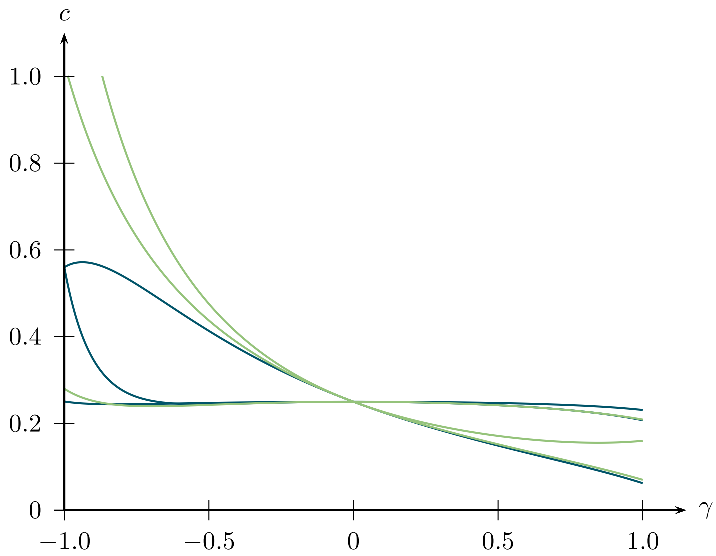


Figure 5.9: Conditions for local stability for $(n_1^e, n^j, n_2^e) = (1,1,0)\,,\,(0,1,1)$ $(a-k = 1)$

$(0,1]$. For $\gamma = 0$ we have

$$\max\left\{\pi_1^{(1,1,1)} - \pi_1^{(0,1,1)}, \pi_1^{(0,2,0)} - \pi_1^{(0,1,1)}, \pi_2^{(0,1,2)} - \pi_2^{(0,1,1)}\right\} = \frac{(a-k)^2}{4},$$
$$\min\left\{\pi_1^{(0,1,1)} - \pi_1^{(0,0,2)}, \pi_2^{(0,1,1)} - \pi_2^{(0,1,0)}, \pi_2^{(0,1,1)} - \pi_2^{(1,0,1)}\right\} = \frac{(a-k)^2}{4}.$$

Therefore, the network $(0,1,1)$ is locally stable for $\gamma = 0$ and $c = \frac{(a-k)^2}{4}$.

- $\left(\mathbf{n_1^e}, \mathbf{n^j}, \mathbf{n_2^e}\right) \in \{(\mathbf{0},\mathbf{1},\mathbf{2})\,,(\mathbf{2},\mathbf{1},\mathbf{0})\}$
  The local instability of the network $(0,1,2)$ for $\gamma \neq 0$ and for $\gamma = 0$ and $c \neq \frac{(a-k)^2}{4}$ has already been established in Proposition 5.2(v) for n customers.
- $\left(\mathbf{n_1^e}, \mathbf{n^j}, \mathbf{n_2^e}\right) \in \{(\mathbf{2},\mathbf{0},\mathbf{0})\,,(\mathbf{0},\mathbf{0},\mathbf{2})\,,(\mathbf{1},\mathbf{0},\mathbf{0})\,,(\mathbf{0},\mathbf{0},\mathbf{1})\}$
  These networks are locally unstable for $c \neq \frac{(a-k)^2}{4}$, as there is always a profitable deviation of one firm to either add or delete a link to an exclusive customer.

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(6):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.

P. Billand, C. Bravard, S. Chakrabarti, and S. Sarangi. Business intelligence and multimarket competition. *Journal of Public Economic Theory*, 18(2):248–267, 2016.

K. Bimpikis, S. Ehsani, and R. Ilkılıç. Cournot competition in networked markets. Stanford University, Working Paper, 2016.

F. Bloch and B. Dutta. Communication networks with endogenous link strength. *Games and Economic Behavior*, 66(1):39–56, 2009.

F. Bloch and M. O. Jackson. The formation of networks with transfers among players. *Journal of Economic Theory*, 133(1):83–110, 2007.

B. Bollobas. *Random Graphs*. New York: Academic Press, 2001.

S. A. Boorman. A combinatorial optimization model for transmission of job information through contact networks. *The Bell Journal of Economics*, 6(1):216–249, 1975.

Y. Bramoullé, A. Galeotti, and B. Rogers. *The Oxford Handbook of the Economics of Networks*. Oxford University Press, 2016.

B. Buechel and T. Hellmann. Under-connected and over-connected networks: the role of externalities in strategic network formation. *Review of Economic Design*, 16(1):71–87, 2012.

P. Cahuc and F. Fontaine. On the efficiency of job search with social networks. *Journal of Economic Behavior and Organization*, 11(3):411–439, 2009.

A. Calvó-Armengol. Job contact networks. *Journal of Economic Theory*, 115(1): 191–206, 2004.

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 Public Economic Theory*, 71(2): 414–427, 2009.

B. Charoensook. Network formation with productivity as decay. Munich Personal RePEc Archive, Working Paper, MPRA Paper No. 37099, 2012.

F. Cingano and A. Rosolia. People i know: Job search and social networks. *Journal of Labor Economics*, 30(2):291–332, 2012.

M. Corominas-Bosch. Bargaining in a network of buyers and sellers. *Journal of Economic Theory*, 115(1):35–77, 2004.

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

H. Dawid and T. Hellmann. The evolution of R&D networks. *Journal of Economic Behavior and Organization*, 105:158–172, 2014.

H. Dawid and T. Hellmann. R&D investments under endogenous cluster formation. Center for Mathematical Economics, Working Paper, No. 555, 2016.

K. De Jaegher and J. J. A. Kamphorst. Two-way flow networks with small decay. *Journal of Economic Behavior and Organization*, 109:217–239, 2015.

B. Dutta and M. O. Jackson. The stability and efficiency of directed communication networks. *Review of Economic Design*, 5(3):251–272, 2000.

B. Dutta and S. Mutuswami. Stable networks. *Journal of Economic Theory*, 76(2): 322–344, 1997.

P. Erdős and A Rényi. On the evolution of random graphs. In *Publication of the mathematical institute of the hungarian academy of sciences*, pages 17–61, 1960.

S. Goyal. *Connections: An Introduction to the Economics of Networks*. Princeton University Press, 2007.

S. Goyal and S. Joshi. Networks of collaboration in oligopoly. *Games and Economic Behavior*, 43(1):57–85, 2003.

S. Goyal and S. Joshi. Unequal connections. *International Journal of Game Theory*, 34(3):319–349, 2006.

S. Goyal and J. L. Moraga-González. R&D networks. *Rand Journal of Economic*, 31(4):686–707, 2001.

M. Grabisch and A. Rusinowska. A model of influence in a social network. *Theory and Decision*, 69(1):69–96, 2010.

M. Grabisch, A. Mandel, A. Rusinowska, and E. Tanimura. Strategic influence in social networks. *Mathematics of Operations Research*, 2017. Forthcoming.

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, 1974.

J. Häckner. A note on price and quantity competition in differentiated oligopolies. *Journal of Economic Theory*, 93(2):233–239, 2000.

H. Haller. Network extension. *Mathematical Social Sciences*, 64(2):166–172, 2012.

J. W. Hatfield and S. D. Kominers. Matching in networks with bilateral contracts. *American Economic Journal: Microeconomics*, 4(1):176–208, 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 B. W. Rogers. The economics of small worlds. *Journal of the European Economic Association*, 3(2-3):617–627, 2005.

M. O. Jackson and A. van den Nouweland. Strongly stable networks. *Games and Economic Behavior*, 51(2):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(1):44–74, 1996.

C. Johnson and R. Gilles. Spatial social networks. *Review of Economic Design*, 5 (3):273–299, 2000.

P. D. Killworth and H. R. Bernhard. The reversal small-world wxperiment. *Social Networks*, 1(1978/79):159–192, 1978.

E. R. Kranton and D. F. Minehart. Networks versus vertical integration. *The RAND Journal of Economics*, 31(3):570–601, 2000a.

E. R. Kranton and D. F. Minehart. Competition for goods in buyer-seller networks. *Review of Economic Design*, 5(3):301–331, 2000b.

E. R. Kranton and D. F. Minehart. A theory of buyer-seller networks. *The American Economic Review*, 91(3):485–508, 2001.

Y. Lim, A. Ozdaglar, and A. Teytelboym. A simple model of cascades in networks. Laboratory for Information and Decision Systems, MIT, 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*, 18(2):168–192, 2016.

P. Möhlmeier, A. Rusinowska, and E. Tanimura. Competition for the access to and use of information in networks. *Mathematical Social Sciences*, 2017. Forthcoming.

T. Morrill. Network formation under negative degree-based externalities. *International Journal of Game Theory*, 40(2):367–385, 2011.

M. Ostrovsky. Stability in supply chain networks. *The American Economic Review*, 98(3):897–923, 2008.

N. Singh and X. Vives. Price and quantity competition in a differentiated duopoly. *The RAND Journal of Economics*, 15(4):546–554, 1984.

P. Wang and A. Watts. Formation of buyer-seller trade networks in a quality-differentiated product market. *Canadian Journal of Economics/Revue canadienne d'économique*, 39(3):971–1004, 2006.

A. Watts. A dynamic model of network formation. *Games and Economic Behavior*, 34(2):331–341, 2001.

A. Watts. Non-myopic formation of circle networks. *Economic Letters*, 74(2):277–282, 2002.

A. Westkamp. Market structure and matching with contracts. *Journal of Economic Theory*, 145(5):1724–1738, 2010.

Short Curriculum Vitae of Philipp Möhlmeier

Education

2010-2017 **Member of the International Research Training Group Economic Behavior and Interaction Models (EBIM)** at Bielefeld University, Germany, in collaboration with Université Paris 1 – Panthéon-Sorbonne, France.
Member of the Bielefeld Graduate School of Economics and Management (BiGSEM) at Bielefeld University, Germany.
Member of the Center for Mathematical Economics (IMW) at Bielefeld University, Germany.

2005-2011 **Diploma in Business Administration** at Bielefeld University, Germany.

2003-2010 **Diploma in Mathematical Economics** at Bielefeld University, Germany.

2003 **Abitur** at Erich-Gutenberg-Berufskolleg Bünde, Germany.

Publications

2017 **Competition for the access to and use of information in networks**, Mathematical Social Sciences, forthcoming, 2017 (joint with A. Rusinowska and E. Tanimura).

2016 **A degree-distance-based connections model with negative and positive externalities**, Journal of Public Economic Theory, 18(2), pp. 168-192, 2016 (joint with A. Rusinowska and E. Tanimura).

This thesis was printed on non-ageing paper in compliance with DIN-ISO 9706.