Peer Effects in Endogenous Networks

Timo Hiller*

September 18, 2013

Abstract

This paper presents a simple model of strategic network formation with local complementarities in effort levels and positive local externalities for a general class of payoff functions. Results are obtained for one-sided and two-sided link creation. In both cases (pairwise) Nash equilibrium networks are *nested split graphs*, which are a strict subset of *core-periphery networks*. The relevance of the convexity of the *value function* (gross payoffs as a function of neighbors' effort levels when best responding) in obtaining nested split graphs is highlighted. Under additional assumptions on payoffs, we show that the only efficient networks are the complete and the empty network. Furthermore, there exists a range of linking cost such that any (pairwise) Nash equilibrium is inefficient and for a strict subset of this range any (pairwise) Nash equilibrium network structure is different from the efficient network. These findings are relevant for a wide range of social and economic phenomena, such as educational attainment, criminal activity, labor market participation, and R&D expenditures of firms.

Key Words: Strategic network formation, peer effects, strategic complements, positive externalities. **JEL Codes:** D62, D85.

1 Introduction

Peer effects and social structure play an important role in determining individual behavior and aggregate outcomes in many social and economic settings. This has been documented by a large body of empirical work, which finds peer effects and network position crucial for decisions concerning educational attainment, criminal activity, labor market participation

*I am grateful to my supervisors Fernando Vega-Redondo and Massimo Morelli for their support and guidance. I also thank Sebastian Bervoets, Sergio Currarini, Andrea Galeotti, Sanjeev Goyal, In-Uck Park, Paolo Pin, Brian Rogers, Tim Schmidt-Eisenlohr and participants of the Networks Working Group at the European University Institute, the UECE Lisbon Meetings 2010 and the CTN 2013 for helpful comments. All remaining errors are mine. Contact: timo.hiller@bristol.ac.uk, Address: Department of Economics, University of Bristol, 8 Woodland Road, BS81TN, UK.

and R&D expenditures of firms. In these settings an agent's optimal action and payoff is thought to depend directly on the action or payoff of others (peer effects), while the relevant reference group is determined by the network of bilateral relationships between agents (social structure).

This paper endogenizes the network in a setting which accounts for peer effects. In accordance with empirical studies, peer effects are assumed to induce positive local externalities and strategic complementarities.[1] The setup is simple. Agents simultaneously choose a non-negative, continuous effort level and create links at a cost. The significance of a link is that direct neighbors in the network benefit from each other's effort levels (positive local externalities) and an agent's incentive to exert effort is strictly increasing in the sum of his neighbor's effort levels (strict strategic complementarities). Furthermore, we assume payoff functions for which the corresponding *value function* is convex. That is, when best responding, own payoffs are convex in the sum of effort levels of direct neighbors. Results for two specifications of the model are presented. First, two-sided link formation, where linking cost are shared equally. For this case we solve for pairwise Nash equilibrium, which reflects the bilateral nature of creating a link (and sharing its cost). Second, one-sided link formation, where linking cost are borne unilaterally. This specification allows us to employ Nash equilibrium.[2]

For both, one-sided and two-sided link formation, we show that equilibrium networks are *nested split graphs*, a special case of core-periphery networks.[3] Core-periphery structures are frequently observed in empirical work on networks.[4] More recently, nested split graphs have gained increased attention and empirical support has been established in a variety of contexts.[5] Typically, some agents in the periphery are connected to agents in the core and we

[1]See Hoxby (2000), Sacerdote (2001) for a treatment of peer effects in education, Glaeser, Sacerdote and Scheinkman (1996), Case and Katz (1991) and Ludwig et al (2001) for criminal and delinquent behavior, Topa (2001) and Conley and Topa (2001) for labor markets and Cohen and Levinthal (1989, 1990) and Levin and Reiss (1988) for R&D expenditure of firms.

[2]Pairwise Nash equilibrium was first discussed in Jackson and Wolinsky (1996). For applications see, for example, Goyal and Joshi (2003) and Belleflamme and Bloch (2004). The one-sided specification follows Bala and Goyal (2000).

[3]There are varying definitions of core-periphery networks in the social sciences literature. Typically, a core-periphery structure indicates the existence of one group that is internally densely connected and another group that is internally sparsely connected (Borgatti and Everett (1999)). We follow the definition given in Bramoullé (2007): A *core-periphery network* is such that the set of agents can be partitioned into two sets, called the core and the periphery, where all agents in the core are connected to each other, while no pair of agents in the periphery is connected. A *nested split graph* is a graph such that, if the link between i and l exists and the degree of k is at least as high as the degree of l, then the link between i and k also exists. All nested split graphs are core-periphery networks, while the converse is not true. For a detailed account of nested split graphs see Mahadev and Peled (1995).

[4]Empirical examples of core-periphery networks relevant to this paper are Adamic and Adam (2003) for friendship, Canter (2004) for crime and Powell et al (1996), Gulati and Gargiulo (1999) and Baker et al (2004) for R&D networks.

[5]See Akerman and Larsson (2010) for an example in global arms trade and May et al (2008) for networks of banks. König, Tessone and Zenou (2012) provide a discussion of nested split graphs with further references and empirical evidence, briefly discussed below.

therefore investigate the possibility of such equilibria to arise. For the case of linear-quadratic payoffs we provide necessary and sufficient conditions for the existence of a star network for both link formation specifications. Turning again to the more general class of payoff functions, we show that effort levels and gross payoffs are higher for more central agents. This is in accordance with the empirical literature on networks.[6] Perhaps surprisingly, however, if link formation is two-sided, then the center of a star may obtain strictly lower net payoffs (i.e., net of linking cost). In these equilibria higher linking cost outweigh higher gross payoffs. Finally, we provide results for social welfare and show that, if own payoffs are weakly convex in effort levels of neighbors, then the only efficient network structures are the empty and the complete network. Moreover, there exists a range of linking cost for which all (pairwise) Nash equilibria are inefficient. For a strict subset of this range not only is any (pairwise) Nash equilibrium inefficient, but also any (pairwise) Nash equilibrium network structure is different from the efficient network. Note that the range of linking cost for which any (pairwise) Nash equilibrium is inefficient may be thought of as large, in the sense that a (pairwise) Nash equilibrium may only be efficient if the network is empty.

Two empirical papers that are closely related are Calvó-Armengol, Pattacchini and Zenou (2005 and 2009). The authors use a detailed data set on friendship networks in U.S. high schools (AddHealth) to test a structural model on a fixed network. This allows for the measurement of peer effects in education and delinquent behavior, respectively. Calvó-Armengol et al find a positive relationship between grades and delinquency rates on the one hand and centrality on the other hand. Network position turns out to be a key determinant for an individual's effort level. This emphasizes the importance of social structure for peer influences, as opposed to average in-group effects. In both papers positive local spillovers and strategic complementarities in effort levels are observed. Note that the payoff function in Calvó-Armengol et al (2005 and 2009) is the same as the one in the linear-quadratic case presented here and we therefore endogenize the network for these papers.

The one-sided specification presented in our paper can be seen as the conceptual counterpart to Galeotti and Goyal (2010). They also solve a simultaneous move game, where agents choose a non-negative, continuous effort level and link formation is one-sided. Externalities are positive and local, but different from our paper, Galeotti and Goyal (2010) assume strategic substitutes. The only strict equilibria in their paper are core-periphery networks, where agents in the periphery extend links to agents in the core. Note that the star network, for which we provide necessary and sufficient conditions, is a special case of the core-periphery structures obtained in Galeotti and Goyal (2010). We thereby show that core-periphery networks, in which agents in the periphery are connected to agents in the core, are not a feature

[6] Baldwin et al (1997) finds grades to be positively correlated with centrality for students of an MBA course. Lin (2001) analyzes job market data and links centrality to obtaining better job offers. Powell et al (1996) and Ahuja (2000) show that more central firms tend to issue more patents, have higher earnings and a lower probability of going bankrupt.

of strategic substitutes alone, but may also be obtained under strategic complementarities.[7] The model presented by Baetz (2012) is also related. Baetz' setup is as in the one-sided link formation specification of our paper, but instead of payoff functions with a convex value function, Baetz (2012) assumes payoff functions such that the corresponding value function is concave. A full characterization is not available yet. It can be shown, however, that a wide range of equilibria may be sustained in equilibrium, such as regular networks, biregular bipartite graphs, multipartite graphs, and (under fairly restrictive assumptions on payoffs) also core-periphery networks. That is, if the value function is convex, then *all* equilibria are nested split graphs, while if the value function is concave, then many more structures may arise (under one-sided link formation). To emphasize the relevance of the convexity of the value function, we indicate in the final part of the paper how our equilibrium characterization extends to the case of strategic substitutes.[8] More specifically, if link formation is one-sided and the value function is convex, then all Nash equilibria are nested split graphs, irrespective of the nature of strategic interaction.

A recent paper by König, Tessone, and Zenou (2012) addresses link formation for the linear-quadratic specification. Note that the authors interpret their model as one of financial and trade networks and, of course, their interpretation carries over to our model as well.[9] The link formation process in König et al (2012) is very different from the one presented here. The setup is dynamic and in each time period players play a two-stage game. In the first stage, agents choose their effort levels on a fixed network, while in the second stage a randomly selected player may create a new link in the current network, at zero cost. Links decay over time, with more valuable links decaying at a slower rate. König et al (2012) then introduce noise into the model and derive the stochastically stable networks. These are shown to be nested split graphs. The authors test their model using four different data sets (two banking networks, a global trade network and an arms trade network) and find evidence for nestedness and an overall good fit.

Finally, the paper by Ballester, Calvó-Armengol and Zenou (2006) is also related. Again the presence of a link in the network (after a payoff decomposition) allows agents to benefit from each other's effort levels and payoffs are linear-quadratic. Different from our paper, however, Ballester et al (2006) assume not only local strategic complementarities, but also allow for global strategic substitutes. The authors then link equilibrium actions to Bonacich

[7] Note that the payoff function in Galeotti and Goyal (2010) is given by $\pi(x_i, y_i) = f(x_i + \sum_{j \in N_i(\bar{\mathbf{g}})} x_j) - c(x_i)$, where f is assumed to be concave and c is linear. By making appropriate assumptions on f and c, we can generate a model of positive externalities and strategic complements that fits our setup. Arguably the simplest such specification is $\pi(x_i, y_i) = (x_i + \sum_{j \in N_i(\bar{\mathbf{g}})} x_j)^2 - x_i^3$.

[8] Note that in Galeotti and Goyal (2010) the value function is linear as long as own effort level is positive and strictly concave thereafter.

[9] See Appendix C in König, Tessone, and Zenou (2012) for a derivation of the linear-quadratic payoff specification in the context of networks of banks operating in loan markets and networks of trade relationships among countries.

centrality on a fixed network.

The rest of the paper is organized as follows: Section 2 describes the two-sided model and then presents the analysis. Section 3 concludes. Some of the proofs for the two-sided specification are relegated to Appendix A, while the one-sided model is presented in its entirety in Appendix B.

2 The Two-Sided Model

2.1 Model Description

Let $N = \{1, 2, ..., n\}$ be the set of players with $n \geq 3$. Each agent i chooses an effort level $x_i \in X$ and announces a set of agents to whom he wishes to be linked to, which is represented by a row vector $\mathbf{g}_i = (g_{i,1}, ..., g_{i,i-1}, g_{i,i+1}, ..., g_{i,n-1})$, with $g_{i,j} \in \{0, 1\}$ for each $j \in N \backslash \{i\}$. Assume $X = [0, +\infty)$ and $\mathbf{g}_i \in G_i = \{0, 1\}^{n-1}$. The set of agent i's strategies is denoted by $S_i = X \times G_i$ and the set of strategies of all players by $S = S_1 \times S_2 \times ... \times S_n$. A strategy profile $\mathbf{s} = (\mathbf{x}, \mathbf{g}) \in S$ then specifies the individual effort level for each player, $\mathbf{x} = (x_1, x_2, ..., x_n)$, and the set of intended links, $\mathbf{g} = (\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_n)$. A link between i and j, denoted with $\bar{g}_{i,j} = 1$, is created if and only if *both* agents intend to create a link. That is, $\bar{g}_{i,j} = 1$ if $g_{i,j} = g_{j,i} = 1$, and $\bar{g}_{i,j} = 0$ otherwise. From $\mathbf{g}$ we thereby obtain the undirected graph $\bar{\mathbf{g}}$ with $\bar{g}_{i,j} = \bar{g}_{j,i}$. The presence of a link $\bar{g}_{i,j} = 1$ allows players to directly benefit from the effort level exerted by the respective other agent involved in the link. Let $N_i(\mathbf{g}) = \{j \in N : g_{i,j} = 1\}$ be the set of agents to which agent i extends a link and denote the corresponding cardinality with $\eta_i(\mathbf{g}) = |N_i(\mathbf{g})|$. Define the set of i's neighbors in $\bar{\mathbf{g}}$ (i.e., the set of agents to which i reciprocates the announcement of a link) with $N_i(\bar{\mathbf{g}}) = \{j \in N : \bar{g}_{i,j} = 1\}$ and define $\eta_i(\bar{\mathbf{g}}) = |N_i(\bar{\mathbf{g}})|$. The aggregate effort level of agent i's neighbors in $\bar{\mathbf{g}}$ is written as $y_i = \sum_{j \in N_i(\bar{\mathbf{g}})} x_j$. We sometimes drop the subscript of y_i when it is clear from the context. Given a network $\bar{\mathbf{g}}$, $\bar{\mathbf{g}} + \bar{g}_{i,j}$ and $\bar{\mathbf{g}} - \bar{g}_{i,j}$ have the following interpretation. When $\bar{g}_{i,j} = 0$ in $\bar{\mathbf{g}}$, $\bar{\mathbf{g}} + \bar{g}_{i,j}$ adds the link $\bar{g}_{i,j} = 1$, while if $\bar{g}_{i,j} = 1$ in $\bar{\mathbf{g}}$, then $\bar{\mathbf{g}} + \bar{g}_{i,j} = \bar{\mathbf{g}}$. Similarly, if $\bar{g}_{i,j} = 1$ in $\bar{\mathbf{g}}$, $\bar{\mathbf{g}} - \bar{g}_{i,j}$ deletes the link $\bar{g}_{i,j}$, while if $\bar{g}_{i,j} = 0$ in $\bar{\mathbf{g}}$, then $\bar{\mathbf{g}} - \bar{g}_{i,j} = \bar{\mathbf{g}}$. The network is called empty and denoted with $\bar{\mathbf{g}}^e$, if $\bar{g}_{i,j} = 0 \; \forall i, j \in N$ and complete and denoted with $\bar{\mathbf{g}}^c$ if $\bar{g}_{i,j} = 1 \; \forall i, j \in N$.

Payoffs of player i under strategy profile $\mathbf{s} = (\mathbf{x},\mathbf{g})$ are given by

$$\Pi_i(\mathbf{s}) = \pi(x_i, y_i) - \eta_i(\mathbf{g})k,$$

where k denotes the cost of extending a link, with $k > 0$. Gross payoffs $\pi(x_i, y_i)$ are a function of own effort, x_i, and the sum of effort levels of direct neighbors, $y_i = \sum_{j \in N_i(\bar{\mathbf{g}})} x_j$. We assume strict positive local externalities and strict strategic complementarities in effort levels, so that $\frac{\partial \pi(x,y)}{\partial y} > 0$ and $\frac{\partial^2 \pi(x,y)}{\partial x \partial y} > 0$. Further assume that $\frac{\partial^2 \pi(x,y)}{\partial^2 x} < 0$. The latter

assumption, together with the convexity of X, guarantees a unique maximizer, which is denoted by $\bar{x}(y)$. We also assume $\bar{x}(y) > 0$.[10] From $\frac{\partial^2 \pi(x,y)}{\partial x \partial y} > 0$ (strict positive externalities) we know that best response functions are strictly increasing in the sum of effort level accessed, so that $\frac{\partial \bar{x}(y)}{\partial y} > 0$. Best response functions are assumed to be either linear or concave, i.e., $\frac{\partial^2 \bar{x}(y)}{\partial^2 y} = 0$ or $\frac{\partial^2 \bar{x}(y)}{\partial^2 y} < 0$. Denote the value function with $v(y) = \pi(\bar{x}(y), y)$. The value function yields an agent's gross payoffs when best responding to a given sum of effort level accessed. Throughout we assume $v(y)$ to be convex, $\frac{\partial^2 v(y)}{\partial^2 y} > 0$. In order to guarantee existence, we further assume that there exists a value of y such that $\frac{\partial \bar{x}(y)}{\partial y} < \frac{1}{n-1}$.

In the following we define pairwise Nash equilibrium (PNE). A strategy profile $\mathbf{s} = (\mathbf{x}, \mathbf{g})$ is a pairwise Nash equilibrium *iff*

- $\mathbf{s}$ is a Nash Equilibrium, and
- for all $\bar{g}_{i,j} = 0$, if $\Pi_i(x'_i, x'_j, \mathbf{x}_{-i,-j}, \bar{\mathbf{g}} + \bar{g}_{i,j}) > \Pi_i(\mathbf{s})$, then $\Pi_j(x'_i, x'_j, \mathbf{x}_{-i,-j}, \bar{\mathbf{g}} + \bar{g}_{ij}) < \Pi_j(\mathbf{s})$, $\forall x'_i, x'_j \in X$.

A pairwise Nash equilibrium is therefore both, a Nash equilibrium and pairwise stable.

Social welfare is defined as the sum of individual payoffs. For any strategy profile $\mathbf{s}$, social welfare is given by

$$W(\mathbf{s}) = \sum_{i \in N} \Pi_i(\mathbf{s}).$$

A profile $\tilde{\mathbf{s}}$ is socially efficient if $W(\tilde{\mathbf{s}}) \geq W(\mathbf{s})$, $\forall \mathbf{s} \in S$.

2.2 Analysis

We start the analysis by providing a proof for the existence and uniqueness of a Nash equilibrium on a fixed network. For the case of linear best response functions we employ the existence result provided by Ballester et al (2006). For the case of concave best response functions we make use of a fixed point theorem provided by Kennan (2001). As in Kennan's paper, we define a vector $\mathbf{b}$ to be larger than a vector $\mathbf{a}$ if and only if $b_i > a_i$ $\forall i \in N$.

Proposition 1: *For any fixed network, $\bar{\mathbf{g}}$, there exists a unique NE in effort levels.*

Proof. *See the Appendix.*

[10] This assumption guarantees that there does not always, i.e., for any linking cost k, exist a (pairwise) Nash equilibrium such that the network is empty.

Lemma 1 shows that agents in a complete component exert equal effort levels.

Lemma 1: *NE effort levels are equal for all players in a complete component.*

Proof. Assume to the contrary that there exists a pair of players k and l in a complete component, such that $x_k^* \neq x_l^*$ and, without loss of generality, that $x_k^* > x_l^*$. Note that in a complete component, k and l are connected to the same sets of third agents, such that $N_k(\bar{\mathbf{g}}) \setminus \{l\} = N_l(\bar{\mathbf{g}}) \setminus \{k\}$. But then, for $x_k^* > x_l^*$, the sum of effort levels of l's neighbors, $\sum_{j \in N_l(\bar{\mathbf{g}})} x_j^*$, is larger than the sum of effort levels of k's neighbors, $\sum_{j \in N_k(\bar{\mathbf{g}})} x_j^*$. We have reached a contradiction, since due to strict strategic complementarities $\sum_{j \in N_l(\bar{\mathbf{g}})} x_j^* > \sum_{j \in N_k(\bar{\mathbf{g}})} x_j^*$ implies $x_l^* > x_k^*$. *Q.E.D.*

In Lemma 2 we show that effort levels are maximal in the complete network.

Lemma 2: *NE effort levels are maximal in the complete network.*

Proof. Denote the Nash equilibrium effort level in the complete network, $\bar{\mathbf{g}}^c$, with $\mathbf{x}^{c*}$. From Lemma 1 we know that $x^{c*} = x_i^*(\bar{\mathbf{g}}^c)\ \forall i \in N$. Start by deleting a link $\bar{g}_{i,j}$ from $\bar{\mathbf{g}}^c$ and consider each player's best response to $\mathbf{x}^{c*}$ in $\bar{\mathbf{g}}^c - \bar{g}_{i,j}$. Agent i's initial best response will be lower in $\bar{\mathbf{g}}^c - \bar{g}_{i,j}$ than in $\bar{\mathbf{g}}^c$, as $\sum_{j \in N_i(\bar{\mathbf{g}}^c - \bar{g}_{i,j})} x_j^* < \sum_{j \in N_i(\bar{\mathbf{g}}^c)} x_j^*$. Iterating on best responses, any agent l with $\bar{g}_{i,l}^* = 1$ decreases his effort level, and those sustaining links with l decrease their effort levels in turn, and so forth. The effort level of each agent is a decreasing sequence of real numbers, which is bounded below by $\bar{x}(0)$. Effort levels therefore convergence to a new equilibrium in $\bar{\mathbf{g}}^c - \bar{g}_{i,j}$, with $x_l^*(\bar{\mathbf{g}}^c - \bar{g}_{i,j}) < x^*(\bar{\mathbf{g}}^c)\ \forall l \in N$. Note that any network $\bar{\mathbf{g}} \neq \bar{\mathbf{g}}^c$ can be obtained from $\bar{\mathbf{g}}^c$ by deleting a sequence of links. Effort levels are weakly decreasing at each step and strictly for the deletion of the first link in a complete network. *Q.E.D.*

Next, we define two cost threshold cost, k^1 and k^2. The first threshold, k^1, is given by the gross marginal payoffs (net of linking cost) when a pair of agents creates a link in the empty network. Note that under pairwise Nash equilibrium we allow both agents creating a new link to adjust their effort levels. The second threshold, k^2, is defined as the average gross marginal payoffs of linking to $(n-1)$ agents in the complete network. Both thresholds are expressed in terms of the value function.

Definition 1: $k^1 = v_i(x_j^*(\bar{\mathbf{g}}^e + \bar{g}_{i,j})) - v_i(0)$ and $k^2 = \frac{v((n-1)x^*(\bar{\mathbf{g}}^c)) - v(0)}{n-1}$.

Proposition 2 shows that, for linking cost smaller than k^1, the unique pairwise Nash equilibrium is the complete network, while for linking cost larger than or equal to k^1, there exists a pairwise Nash equilibrium such that the network is empty. To see this, note that with strategic complementarities and a convex value function, the marginal value of adding a link (and adjusting effort levels) is lowest for a pair of agents with no links. That is, if

linking cost is smaller than k^1, then in any network that is not complete, there exists a pair of agents who finds it profitable to create a link (and to adjust their effort levels). Therefore, the unique pairwise Nash equilibrium is the complete network. Conversely, if linking cost is larger than or equal to k^1, then no pair of agents in the empty network finds it profitable to create a link (and to adjust their effort levels) and there exists a pairwise Nash equilibrium such that the network is empty.

Proposition 2: *If $k < k^1$, then the unique PNE is the complete network. If $k \geq k^1$, then there exists a PNE such that the network is empty.*

Proof. *See the Appendix.*

Proposition 3 shows that for linking cost larger than k^2, the unique pairwise Nash equilibrium is the empty network, while for linking cost smaller than or equal to k^2, there exists a pairwise Nash equilibrium such that the network is complete. Note that the (non-empty) pairwise Nash equilibrium network that can be sustained at the highest linking cost is the complete network. This follows from strategic complementarities and the convexity of the value function, together with our results from Lemma 1 and Lemma 2, that effort levels are equal and maximal in the complete network. The threshold value of k^2 is such that, if linking cost is larger than k^2, then an agent in the complete network finds it profitable to delete his links (and adjust his effort level). The unique pairwise Nash equilibrium is therefore the empty network. For linking cost smaller than or equal to k^2, no agent in the complete network finds it profitable to sever any subset of links (and to adjust his effort level). Therefore, there exists a pairwise Nash equilibrium such that the network is complete.

Proposition 3: *If $k > k^2$, then the unique PNE is the empty network. If $k \leq k^2$, then there exists a PNE such that the network is complete.*

Proof. *See the Appendix.*

Lemma 3 shows that $k^1 < k^2$.

Lemma 3: $k^1 < k^2$.

Proof. *See the Appendix.*

To summarize, for linking cost smaller than k^1, the unique PNE is the complete network, while for linking cost larger than k^2 the unique PNE network is the empty network. For linking cost $k \in [k^1, k^2]$ the complete and the empty network are pairwise Nash equilibria.

Next, we formally define *core − periphery networks* (which are called *split graphs* in the mathematical graph theory literature) and *nested split graphs*. In a core-periphery network the set of agents can be partitioned into two sets, such that all pairs of agents within the first set (the core) are connected and no pair of agents within the second set (the periphery) is connected. Note that the definition does not state anything about links between pairs of agents where one agent is in the core and the other is in the periphery. The *star network* is a special case of a core-periphery network. In a star network there exists a partition of agents such that the core consists of a single agent and all agents in the periphery are connected to the the core. A *nested split graph* is a network such that, if a link between agents i and l exists, and agent k has a higher or equal number of links (degree) than l, then the link between i and k also exists. That is, neighborhoods are nested. Nested split graphs display a core-periphery structure.[11] Note that the complete and empty network are nested split graphs and therefore core-periphery networks (take the periphery or core to be the empty set, respectively).

Definition 2: A network $\bar{\mathbf{g}}$ is a *core − periphery network* (*split graph*) if the set of agents N can be partitioned into two sets, $C(\bar{\mathbf{g}})$ (the core) and $P(\bar{\mathbf{g}})$ (the periphery), such that $\bar{g}_{i,j} = 1\ \forall i, j \in C(\bar{\mathbf{g}})$ and $\bar{g}_{i,j} = 0\ \forall i, j \in P(\bar{\mathbf{g}})$. A *star* is a *core − periphery network* in which the set of agents can be partitioned into two sets, $C(\bar{\mathbf{g}})$ and $P(\bar{\mathbf{g}})$, such that $\mid C(\bar{\mathbf{g}}) \mid = 1$ and $\bar{g}_{i,j} = 1\ \forall i \in P(\bar{\mathbf{g}})$ and $\forall j \in C(\bar{\mathbf{g}})$.

Definition 3: A network $\bar{\mathbf{g}}$ is a *nested split graph* if and only if

$$[\bar{g}_{i,l} = 1 \wedge \eta_k(\bar{\mathbf{g}}) \geq \eta_l(\bar{\mathbf{g}})] \Rightarrow \bar{g}_{i,k} = 1.$$

To illustrate the difference between a core-periphery network and a nested split graph we provide a simple example in Figure 1 below. The numbers in the graph indicate the degree of an agent.

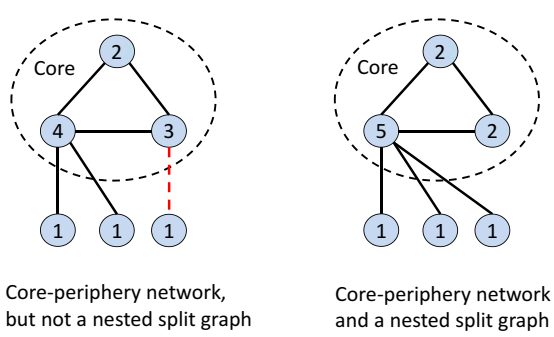

Figure 1

[11] See König et al (2011) and Mahadev and Peled (1995) for properties of nested split graphs.

In the following we provide three Lemmas which are useful when establishing Proposition 4 and Corollary 1. In Proposition 4 we prove that in any PNE the network is a nested split graph. Corollary 1 shows that, if all agents display equal effort levels, then the network is either empty or complete, while if there exists a pair of agents with different effort levels, then the network displays a core-periphery structure (which is neither the empty nor the complete network). In Lemma 4 we prove that in any PNE, if an agent i is linked to agent l, then agent i must also be linked to any agent k with higher or equal effort level than agent l. This is a direct consequence of the convexity of the value function. Lemma 5 shows that in any PNE agents with same effort levels must be connected to the same set of agents, while in Lemma 6 we prove that the neighborhoods of agents with lower effort levels are contained in the neighborhoods of agents with higher effort levels.

Lemma 4: *If $\bar{g}^*_{i,l} = 1$, then $\bar{g}^*_{i,k} = 1$ for all agents k with $x^*_k \geq x^*_l$.*

Proof. Assume to the contrary of the above that $\bar{g}^*_{i,l} = 1$ and $\bar{g}^*_{i,k} = 0$ for some agent k with $x^*_k \geq x^*_l$. Note first that for $\bar{g}^*_{i,l} = 1$ to be part of a PNE, it must be that agent i and agent j can not profitably deviate by deleting the link. For agent i this condition reads $v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x^*_j) - v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x^*_j - x^*_l) \geq k$. If linking to agent l is profitable for agent i, then linking to any agent k with $x^*_k \geq x^*_l$ is likewise profitable. To see this, note that $v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x^*_j + x^*_k) - v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x^*_j) > v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x^*_j) - v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x^*_j - x^*_l) \geq k$, which follows from the convexity of the value function and $x^*_k \geq x^*_l$. Therefore, for a link between agents i and k to be absent in PNE, we need that agent k does not find it profitable to link to agent i. In the following we show that this can not be the case. For $\bar{g}^*_{i,l} = 1$ to be in place in a PNE, it must be that agent l can not profitably deviate by deleting his link with agent i. That is, we must have $v(\sum_{j \in N_l(\bar{\mathbf{g}}^*)} x^*_j) - v(\sum_{j \in N_l(\bar{\mathbf{g}}^*)} x^*_j - x^*_i) \geq k$. Note next that, due to $x^*_k \geq x^*_l$ and strict strategic complementarities, $\sum_{j \in N_k(\bar{\mathbf{g}}^*)} x^*_j \geq \sum_{j \in N_l(\bar{\mathbf{g}}^*)} x^*_j$ holds. We can now write $v(\sum_{j \in N_k(\bar{\mathbf{g}}^*)} x^*_j + x^*_i) - v(\sum_{j \in N_k(\bar{\mathbf{g}}^*)} x^*_j) > v(\sum_{j \in N_l(\bar{\mathbf{g}}^*)} x^*_j) - v(\sum_{j \in N_l(\bar{\mathbf{g}}^*)} x^*_j - x_i) \geq k$. The inequalities again follow from the convexity of the value function and $\sum_{j \in N_k(\bar{\mathbf{g}}^*)} x^*_j \geq \sum_{j \in N_l(\bar{\mathbf{g}}^*)} x^*_j$. We have thereby shown that it is profitable for agent k to link to agent i and have reached a contradiction. That is, if $\bar{g}^*_{i,l} = 1$, then agent i finds it profitable to link to any agent k with $x^*_k \geq x^*_l$, while any agent k finds it profitable to link to agent i and therefore $\bar{g}^*_{i,k} = 1$ for all agents k with $x^*_k \geq x^*_l$. *Q.E.D.*

Lemma 5: *In any PNE, $x^*_i = x^*_k \Leftrightarrow N_i(\bar{\mathbf{g}}^*) \setminus \{k\} = N_k(\bar{\mathbf{g}}^*) \setminus \{i\}$.*

Proof. First, $N_i(\bar{\mathbf{g}}^*) \setminus \{k\} = N_k(\bar{\mathbf{g}}^*) \setminus \{i\} \Rightarrow x^*_i = x^*_k$. If $\bar{g}^*_{i,k} = 0$, then the sum of i and k's neighbors' effort levels is the same, i.e. $y^*_i = \sum_{j \in N_i(\bar{\mathbf{g}}^*)} x^*_j = y^*_k = \sum_{j \in N_k(\bar{\mathbf{g}}^*)} x^*_j$ and therefore $x^*_i = x^*_k$. Assume next that $\bar{g}^*_{i,k} = 1$ and, without loss of generality, that $x^*_i > x^*_k$. But then k accesses a higher effort level than i, $y^*_i = \sum_{j \in N_i(\bar{\mathbf{g}}^*)} x^*_j < y^*_k = \sum_{j \in N_k(\bar{\mathbf{g}}^*)} x^*_j$, and we have reached a contradiction. Second, $x^*_i = x^*_k \Rightarrow N_i(\bar{\mathbf{g}}^*) \setminus \{k\} = N_k(\bar{\mathbf{g}}^*) \setminus \{i\}$.

Assume to the contrary that $x_i^* = x_k^*$ and $N_i(\bar{\mathbf{g}}^*) \setminus \{k\} \neq N_k(\bar{\mathbf{g}}^*) \setminus \{i\}$. Note that for $x_i^* = x_k^*$, effort levels accessed must be equal by strict strategic complementarities, so that $y_i^* = y_k^*$. There must therefore exist an agent l, such that $l \in N_k(\bar{\mathbf{g}}^*)$ and $l \notin N_i(\bar{\mathbf{g}}^*)$. For the link $\bar{g}_{k,l}^* = 1$ to be in place in $\bar{\mathbf{g}}^*$, we must have that $v(\sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^*) - v(\sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^* - x_l^*) \geq k$. From $y_i^* = y_k^*$ and the convexity of the value function we then reach a contradiction since $v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^* + x_l^*) - v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^*) > v(\sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^*) - v(\sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^* - x_l^*) \geq k$. *Q.E.D.*

Lemma 6: *In any PNE,* $x_i^* \leq x_k^* \Leftrightarrow N_i(\bar{\mathbf{g}}^*) \setminus \{k\} \subseteq N_k(\bar{\mathbf{g}}^*) \setminus \{i\}$.

Proof. First, $N_i(\bar{\mathbf{g}}^*) \setminus \{k\} \subseteq N_k(\bar{\mathbf{g}}^*) \setminus \{i\} \Rightarrow x_i^* \leq x_k^*$. If $\bar{g}_{i,k}^* = 0$, then k accesses a weakly higher effort level, i.e. $y_i = \sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^* \leq y_k = \sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^*$ and therefore $x_i^* \leq x_k^*$. Assume next that $\bar{g}_{i,k}^* = 1$ and, without loss of generality, that $x_i^* > x_k^*$. But then k accesses a strictly higher effort level than i, $y_i^* = \sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^* < y_k^* = \sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^*$, and we have reached a contradiction. Second, $x_i^* \leq x_k^* \Rightarrow N_i(\mathbf{g}^*) \setminus \{k\} \subseteq N_k(\mathbf{g}^*) \setminus \{i\}$. Assume to the contrary that $x_i^* \leq x_k^*$ and there exists an agent l such that $l \in N_i(\bar{\mathbf{g}}^*)$ and $l \notin N_k(\bar{\mathbf{g}}^*)$. For the link $\bar{g}_{i,l}^* = 1$ to be in place in $\bar{\mathbf{g}}^*$, we must have that $v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^*) - v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^* - x_l^*) \geq k$. But from $y_i^* \leq y_k^*$ and the convexity of the value function it follows that $v(\sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^* + x_l) - v(\sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^*) > v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^*) - v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^* - x_l^*) \geq k$ and we have reached a contradiction. *Q.E.D.*

We are now in the position to prove Proposition 4, which states that in any *PNE* the network is a nested split graph.

Proposition 4: *In any PNE the network is a nested split graph.*

Proof. Note first that the complete and the empty network are nested split graphs. We start by showing that in any *PNE*, if $d_k(\bar{\mathbf{g}}^*) \geq d_l(\bar{\mathbf{g}}^*)$, then $x_k^* \geq x_l^*$. Assume to the contrary that $x_l^* > x_k^*$. We distinguish two cases. First, $d_k(\bar{\mathbf{g}}^*) > d_l(\bar{\mathbf{g}}^*)$. There then exists an agent $m \in N_k(\bar{\mathbf{g}}^*)$ and $m \notin N_l(\bar{\mathbf{g}}^*)$. But, since $\bar{g}_{m,k}^* = 1$ and $x_l^* > x_k^*$, by Lemma 4 $\bar{g}_{m,l}^* = 1$ and we have reached a contradiction. Assume next that $d_k(\bar{\mathbf{g}}^*) = d_l(\bar{\mathbf{g}}^*)$. If $N_k(\bar{\mathbf{g}}^*) \setminus \{l\} = N_l(\bar{\mathbf{g}}^*) \setminus \{k\}$, then $x_k^* = x_l^*$ by Lemma 5 and we have reached a contradiction. If $N_k(\bar{\mathbf{g}}^*) \setminus \{l\} \neq N_l(\bar{\mathbf{g}}^*) \setminus \{k\}$, then there exists an agent $m \in N_k(\bar{\mathbf{g}}^*) \setminus \{l\}$ and $m \notin N_l(\bar{\mathbf{g}}^*) \setminus \{k\}$. Again by Lemma 4, since $\bar{g}_{m,k}^* = 1$ and $x_l^* > x_k^*$, it follows that $\bar{g}_{l,m}^* = 1$ and we have reached a contradiction. We have established that in any *PNE*, if $d_k(\bar{\mathbf{g}}^*) \geq d_l(\bar{\mathbf{g}}^*)$ then $x_k^* \geq x_l^*$. Therefore, in any *PNE*, if $\bar{g}_{i,l}^* = 1$ and $d_k(\bar{\mathbf{g}}^*) \geq d_l(\bar{\mathbf{g}}^*)$, then $x_k^* \geq x_l^*$ and by Lemma 4 $\bar{g}_{i,k}^* = 1$ and $\bar{\mathbf{g}}^*$ is a nested split graph. *Q.E.D.*

We know from the literature in graph theory that nested split graphs are core-periphery networks (split graphs).[12] Below, we provide an alternative proof, which highlights agent's incentives of linking and exerting effort. Furthermore, a relationship between a property of the PNE effort levels and PNE network structure is established.

Corollary 1: *In any PNE*

- *if $x_i^* = x_j^*$ $\forall i, j \in N$, then the network is either empty or complete,*
- *if $x_i^* \neq x_j^*$ for some $i, j \in N$, then the network is a core-periphery network (which is neither complete nor empty).*

Proof. Lemma 4 directly implies that if $x_i^* = x_j^*$ $\forall i, j \in N$, then in any PNE the network is either empty or complete. Next, we focus on the case such that there exists a pair of agents i and j, with $x_i^* \neq x_j^*$. We show that in any PNE the network displays a core-periphery structure (other than the complete and empty network). Rank agents by their effort levels in increasing order, such that $x_1^* \leq x_2^* \leq ... \leq x_{n-1}^* \leq x_n^*$. We know from Lemma 1 that the network is not complete, since there exists a pair of agents i and j such that $x_i^* \neq x_j^*$. The network is not empty, as agents have identical payoff functions and singleton agents display equal effort levels, $\bar{x}(0)$. First, we show that the two lowest ranked agents, agent 1 and agent 2, are not connected. Assume to the contrary that $\bar{g}_{1,2}^* = 1$. From Lemma 4 we know that in this case agent i must be connected to all agents, since $x_j^* \geq x_2^*$ $\forall j \geq 2$. Lemma 6 then implies the network is complete, since $N_i(\bar{\mathbf{g}}^*) \setminus \{j\} \subseteq N_j(\bar{\mathbf{g}}^*) \setminus \{i\}$ holds for agents j with $x_j^* \geq x_1^*$ $\forall j \geq 1$. But then $x_i^* = x_j^*$ $\forall i, j \in N$ by Lemma 1 and we have reached a contradiction. Since the network is neither empty nor complete, at least one link exists. Pick the agent i with the lowest subscript who has a link. If i has more than one link, pick the link to the agent with the lowest subscript j. We discern two cases. First, agent i and j are adjacent in the ranking. As i is the agent with the lowest subscript to sustain a link, all agents with lower subscripts have no links. All agents with a subscript higher or equal to i are connected to each other: By Lemma 4, agent i is connected to all agents with subscripts higher or equal than j and by Lemma 6, $\bar{g}_{l,m}^* = 1$ $\forall l, m \geq i$. The periphery, $P(\bar{\mathbf{g}}^*)$, consists of agents with subscripts $k < i$, while the core, $C(\bar{\mathbf{g}}^*)$, consists of agents with subscripts $k \geq i$. Second, agent i and j are not adjacent. Note that since $\bar{g}_{i,j}^* = 1$ and $x_i^* \leq x_{j-1}^*$, we know by Lemma 6 that the link between $j-1$ and j, $\bar{g}_{j-1,j}^* = 1$, also exists. Next, check for the link $\bar{g}_{j-2,j-1}^*$. If $\bar{g}_{j-2,j-1}^* = 0$, then by Lemma 6 no agent with a subscript lower than $j-2$ is connected to $j-1$. Furthermore, no pair of agents with subscripts of lower or equal than $j-2$ is connected. Assume to the contrary that there exists a pair of nodes l, m with $l \leq m < j-2$ and $\bar{g}_{l,m}^* = 1$. By Lemma 4 we must then have that $\bar{g}_{l,j-1}^* = 1$. This, however, contradicts

[12] See, for example, Mahadev and Peled (1995).

Lemma 6, since $\bar{g}^*_{j-2,j-1} = 0$. The periphery, $P(\bar{\mathbf{g}}^*)$, consists of agents with subscripts $k < j$, while the core, $C(\bar{\mathbf{g}}^*)$, consists of agents with subscripts $k \geq j$. If $\bar{g}^*_{j-2,j-1} = 1$, check for the link $\bar{g}^*_{j-3,j-2}$. If $\bar{g}^*_{j-3,j-2} = 0$, then by above argument the periphery, $P(\bar{\mathbf{g}}^*)$, consists of agents with subscripts $k < j-1$, while the core, $C(\bar{\mathbf{g}}^*)$, consists of agents with subscripts $k \geq j-1$. If $\bar{g}^*_{j-3,j-2} = 1$, proceed in descending order until a pair of adjacent agents is found that is not connected and define the core and periphery accordingly. Note that such a pair exists, since i and j were assumed to be not adjacent and therefore $\bar{g}^*_{i,i+1} = 0$. This completes the proof. *Q.E.D.*

Proposition 5 provides an existence result for core-periphery networks that are neither complete nor empty and in which there are no links between agents in the core and the periphery. More specifically, assume $lim_{\frac{\partial\pi(x,y)}{\partial x\partial y}\to 0}\frac{\partial^2\pi(x,y)}{\partial^2 y} > 0$ and $lim_{\frac{\partial\pi(x,y)}{\partial x\partial y}\to 0}\bar{x}(0) > 0$. Then, for $\frac{\partial\pi(x,y)}{\partial x\partial y}$ sufficiently small, a core-periphery network exists with at least three agents in the core (for appropriately chosen k). Denote the cardinality of the core with $c(\bar{\mathbf{g}}^*) =\mid C(\bar{\mathbf{g}}^*) \mid$ and the cardinality of the periphery with $p(\bar{\mathbf{g}}^*) =\mid P(\bar{\mathbf{g}}^*) \mid$. For ease of notation we write c and p, respectively.

Proposition 5: *If $lim_{\frac{\partial\pi(x,y)}{\partial x\partial y}\to 0}\frac{\partial^2\pi(x,y)}{\partial^2 y} > 0$ and $lim_{\frac{\partial\pi(x,y)}{\partial x\partial y}\to 0}\bar{x}(0) > 0$, then for $\frac{\partial\pi(x,y)}{\partial x\partial y}$ sufficiently small, there exists a linking cost k, such that a PNE displays a core$-$periphery network with $\mid C(\bar{\mathbf{g}}^*) \mid\geq 3$.*

Proof. Partition the set of agents into the core, $C(\bar{\mathbf{g}}^*)$, with $\bar{g}^*_{i,j} = 1\ \forall i,j \in C(\bar{\mathbf{g}}^*)$, and the periphery, $P(\bar{\mathbf{g}}^*)$, with $\bar{g}^*_{i,j} = 0\ \forall i,j \in P(\bar{\mathbf{g}}^*)$. Further assume that $\bar{g}^*_{i,j} = 0\ \forall i \in C(\bar{\mathbf{g}}^*)$ and $\forall j \in P(\bar{\mathbf{g}}^*)$. That is, we assume a complete component, consisting of $C(\bar{\mathbf{g}}^*)$, and a set of singletons, $P(\bar{\mathbf{g}}^*)$. Denote the PNE effort level of an agent in the core of size c with x^*_c. The condition for an agent in the core to not delete all his links (recall the argument from Proposition 3) is given by $\frac{v((c-1)x^*_c)-v(0)}{c-1} \geq k$. Denote with x'_c the effort level in a deviation where a pair of agents $p \in P(\bar{\mathbf{g}}^*)$ and $c \in C(\bar{\mathbf{g}}^*)$ creates a link. The condition for an agent in the periphery to not find it profitable to link to an agent in the core is given by $k \geq v(x'_c) - v(0)$. (These are the only two conditions one needs to consider, since payoffs for an agent in the periphery linking to another agent in periphery are lower). For $\frac{\partial\pi(x,y)}{\partial x\partial y}$ sufficiently small, effort levels of all agents are arbitrarily close to $\bar{x}(0) > 0$ and $lim_{\frac{\partial\pi(x,y)}{\partial x\partial y}\to 0}\frac{\partial^2 v(y)}{\partial^2 y} = lim_{\frac{\partial\pi(x,y)}{\partial x\partial y}\to 0}\frac{\partial^2\pi(x,y)}{\partial^2 y} > 0$. That is, $lim_{\frac{\partial\pi(x,y)}{\partial x\partial y}\to 0}\frac{\partial^2 v(y)}{\partial^2 y}$ is bounded away from *zero*. Note also that the value function v and best response function $\bar{x}(0)$ are continuous. For $\frac{\partial\pi(x,y)}{\partial x\partial y}$ sufficiently small we have $\frac{v((c-1)x^*_c)-v(0)}{c-1} > v(x'_c) - v(0)$, where the inequality follows from $\frac{\partial^2 v(y)}{\partial^2 y} > 0$, $c \geq 3$ and all effort levels being arbitrarily close to $\bar{x}(0) > 0$. We can therefore find a value of k such that $\frac{v((c-1)x^*_c)-v(0)}{c-1} > k > v(x'_c) - v(0)$. *Q.E.D.*

In the linear-quadratic specification of Calvó-Armengol et al (2005 and 2009) $\frac{\partial^2 \pi(x,y)}{\partial^2 y} = 0$ and Proposition 5 does not apply. Necessary and sufficient conditions for the type of core-periphery networks described in Proposition 5 can be derived easily, but we omit them here. Instead, Proposition 6 provides necessary and sufficient conditions for a PNE star network. The appeal of this proposition is to show the existence of a PNE core-periphery network such that all agents in the periphery are connected to the core. Recall that the payoff function in Calvó-Armengol et al (2005 and 2009) is given by $\pi(x_i, y_i) = x_i - \frac{\beta}{2}x_i^2 + \lambda x_i \sum_{j \in N_i(\bar{\mathbf{g}})} x_j$, where $\frac{\lambda}{\beta} < \frac{1}{n-1}$.[13]

Note first that a necessary condition for a PNE to exist is that effort levels are a NE on the corresponding fixed network. Ballester et al (2006) show that a Nash equilibrium on a fixed network exists if and only if $\beta > \lambda\mu_1(\bar{\mathbf{g}})$, where $\mu_1(\bar{\mathbf{g}})$ is the largest eigenvector of the corresponding adjacency matrix. For a star network the relevant condition reads $\beta > \lambda\sqrt{n-1}$, which is stated below as $n < 1 + \frac{\lambda^2}{\beta^2}$. The first condition of Proposition 6, $\beta > \lambda(2+\sqrt{2})$, imposes an upper bound on λ relative to β, which is independent of k and n. That is, if the parameter governing strategic complementarities, λ, is sufficiently large relative to the convexity of the cost function, β, then a PNE star network does not exist (for any combination of k and n). Note next that for a PNE star network to exist, linking cost k must be in an intermediate range. The lower bound on linking cost k is given by the marginal payoffs of two agents in the periphery linking to each other. The upper bound, in turn, is given by the average marginal payoffs of the center of the star linking to the remaining $n-1$ agents. Both bounds are increasing in n. For $\beta > \lambda\sqrt{n-1}$ the marginal payoff of a peripheral agent linking to the center is always higher than the center's average marginal payoff from his $n-1$ links to agents in the periphery. Therefore, the requirement that an agent in the periphery does not find it profitable to delete his link to the center is satisfied if k is below the aforementioned upper bound. The lower bound on n (together with $\beta > \lambda(2+\sqrt{2})$) then ensures that the upper bound on linking cost is, in fact, (weakly) larger than the lower bound. The last condition presents the knife-edge case where the lower bound on linking cost k coincides with the upper bound.

Proposition 6: *If best response functions are linear, then there exists a PNE with a star network if and only if*

$\beta > \lambda(2+\sqrt{2})$ and

- $\frac{\beta(2\beta-\lambda)\lambda(\beta+\lambda)^2}{2(\beta-\lambda)^2(\beta^2-(n-1)\lambda^2)^2} \leq k \leq \frac{\lambda(2\beta^3+(n+1)\beta^2\lambda-(n-1)\lambda^3)}{2\beta(\beta^2-(n-1)\lambda^2)^2}$ *and*

[13] We briefly show that $\pi(x_i, y_i) = x_i - \frac{\beta}{2}x_i^2 + \lambda x_i \sum_{j \in N_i(\bar{\mathbf{g}})} x_j$ fulfills the conditions on our payoff functions for $\frac{\lambda}{\beta} < \frac{1}{n-1}$. Note that $\frac{\partial \pi(x,y)}{\partial y} = \lambda x > 0$ and $\frac{\partial^2 \pi(x,y)}{\partial x \partial y} = \lambda > 0$. The best response function is given by $\bar{x}(y) = \frac{1+\lambda y}{\beta}$ with $\frac{\partial \bar{x}(y)}{\partial y} = \frac{\lambda}{\beta} > 0$. The value function is given by $v(y) = \frac{(1+\lambda y)^2}{2\beta}$ and $\frac{\partial^2 v(y)}{\partial^2 y} = \frac{\lambda^2}{\beta} > 0$. Furthermore, $\frac{\partial \pi(x,y)}{\partial y} = \lambda x > 0$ (note that $\bar{x}(0) = \frac{1}{\beta}$) and $\frac{\partial^2 \pi(x,y)}{\partial x \partial y} = \lambda > 0$. For existence we require that $\frac{\lambda}{\beta} < \frac{1}{n-1}$, so that $\frac{\partial \bar{x}(y)}{\partial y} = \frac{\lambda}{\beta} < \frac{1}{n-1}$.

$1+\frac{\beta^2(5\beta-3\lambda)}{(\beta-\lambda)^3} < n < 1+\frac{\beta^2}{\lambda^2}$, *or*

- $k=\frac{\beta(2\beta-\lambda)\lambda(\beta+\lambda)^2}{2(\beta-\lambda)^2(\beta^2-(n-1)\lambda^2)^2}$ *and* $n=1+\frac{\beta^2(5\beta-3\lambda)}{(\beta-\lambda)^3}$ *holds.*

Proof. *See the Appendix.*

In Proposition 7 we show that in any PNE such that two agents have different effort levels, there exists a core-periphery partition such that agents in the core exert strictly higher effort levels and obtain strictly higher gross payoffs. The reason for expressing the statement in terms of the existence of such a partition can be illustrated with a star network. In a star network there are two ways to define a core-periphery network. First, the center of the star as the core and the remaining spokes as the periphery. Second, the center of the star and one spoke as the core and the remaining spokes as the periphery. Note that the center of a star exerts strictly higher effort levels than the remaining spokes (this follows from Proposition 7). In the first case agents in core exert strictly higher effort levels, while in the second case there exists an agent in the core, such that his effort level is equal to those of the periphery.

Proposition 7: *In any PNE with a pair of agents i and j, such that $x_i^* \neq x_j^*$, there exists a core-periphery partition of agents, such that $x_i^* > x_j^*$ and $\pi_i^* > \pi_j^*$ $\forall i \in C(\bar{\mathbf{g}}^*)$ and $\forall j \in P(\bar{\mathbf{g}}^*)$.*

Proof. We start by proving the first part of the statement. Rank again agents by their effort levels in increasing order, such that $x_1^* \leq x_2^* \leq ... \leq x_{n-1}^* \leq x_n^*$. From Proposition 4 we know that there exists an agent x_k^* such that $\bar{g}_{i-1,i}^* = 0$ $\forall i \leq k$ and $\bar{g}_{i,i+1}^* = 1$ $\forall i \geq k$. We discern two cases. First, $x_k^* = x_{k+1}^*$. Define $C(\bar{\mathbf{g}}^*) = \{i : x_i \geq x_k\}$ and $P(\bar{\mathbf{g}}^*) = \{i : x_i < x_k\}$ and assume, contrary to the above, that $x_{k-1}^* \geq x_k^* = x_{k+1}^*$. From Lemma 6 we know that $N_{k+1}(\bar{\mathbf{g}}^*)\backslash\{k-1\} \subseteq N_{k-1}(\bar{\mathbf{g}}^*)\backslash\{k\}$ must hold. That is, if $x_{k-1}^* \geq x_{k+1}^*$, then all of agent $k+1$'s neighbors are also $k-1$'s neighbors. But then, since $k+1$ is connected with k, $k-1$ is also connected with k and $\bar{g}_{k-1,k}^* = 1$. We have reached a contradiction. Next, $x_k^* < x_{k+1}^*$. In this case simply define $C(\bar{\mathbf{g}}^*) = \{i : x_i \geq x_{k+1}\}$ and $P(\bar{\mathbf{g}}^*) = \{i : x_i < x_{k+1}\}$. For the second part of the statement note that from $x_i^* > x_j^*$ we know that $y_i^* = \sum_{k \in N_i(\bar{\mathbf{g}}^*)} x_k^* > \sum_{k \in N_j(\bar{\mathbf{g}}^*)} x_k^* = y_j^*$ holds. We can then write $\pi_i^*(x_i^*, y_i^*) > \pi_i(x_j^*, y_i^*) > \pi_j^*(x_j^*, y_j^*)$, where the inequalities follow from $y_i^* > y_j^*$, positive externalities and a unique maximizer. *Q.E.D.*

Proposition 7 shows that more central agents obtain higher gross payoffs. In the following we provide an example which illustrates that the central agent may, however, obtain strictly lower net payoffs. In this case the cost of sustaining a higher number of links outweighs higher gross payoffs.

Example 1: Assume $n = 10$, $\beta = \frac{2}{3}$ and $\lambda = \frac{1}{10}$.

From plugging in the parameter values into the conditions in Proposition 6 we know that the star network is a PNE for $\frac{725}{11907} \leq k \leq \frac{16}{243}$. Payoffs are $\frac{\beta(\beta+\lambda(n-1))^2}{2(\beta^2-\lambda^2(n-1))^2} - (n-1)k$ for the center and $\frac{\beta(\beta+\lambda)^2}{2(\beta^2-\lambda^2(n-1))^2} - k$ for an agent in the periphery. It can easily be checked that payoffs for an agent in the center are strictly higher than for an agent in the core for $\frac{725}{11907} \leq k < \frac{125}{1944}$, while they are strictly higher for an agent in the periphery if $\frac{125}{1944} < k \leq \frac{16}{243}$.

Next we turn to social welfare. Denote the vector of efficient effort levels for a given network $\bar{\mathbf{g}}$ with $\mathbf{x}_*(\bar{\mathbf{g}})$ and agent i's entry in vector $\mathbf{x}_*(\bar{\mathbf{g}})$ with $x_{i*}(\bar{\mathbf{g}})$. That is, $\mathbf{x}_*(\bar{\mathbf{g}})$ yields the weakly highest sum of payoffs for a given network $\bar{\mathbf{g}}$, so that $W(\mathbf{x}_*(\bar{\mathbf{g}}), \bar{\mathbf{g}}) \geq W(\mathbf{x}(\bar{\mathbf{g}}), \bar{\mathbf{g}})$ $\forall \mathbf{x} \in X_1 \times X_2 \times ... \times X_n$. We show in Proposition 8 that, if payoffs are bounded above in the complete network and $\frac{\partial^2 \pi(x,y)}{\partial^2 y} \geq 0$, then there exists a linking cost k_*, such that for linking cost above k_*, the efficient configuration is given by the empty network and each agent exerting effort level $\bar{x}(0)$. For linking cost below k_*, the efficient configuration is given by the complete network and each agent exerting the same effort level $x_{i*}(\bar{\mathbf{g}}^c) = x_*(\bar{\mathbf{g}}^c) > \bar{x}(0)$ $\forall i \in N$.

Note that if payoffs are bounded in the complete network, then they are bounded in any other network. A sufficient condition for bounded payoffs is therefore that $\pi(x, (n-1)x)$ is concave in x and that there exists a stationary point. The linear-quadratic specification satisfies $\frac{\partial^2 \pi(x,y)}{\partial^2 y} \geq 0$ and one can easily show that in this case $k_* = \frac{2}{n}(1 - (\frac{\beta}{2} - \lambda)(n-1))$, while the condition that efficient payoffs are bounded in the complete network is given by $\frac{\lambda}{\beta} < \frac{1}{2(n-1)}$.

Since the link formation process is not relevant for efficiency, Proposition 8 and Corollary 2 also hold in the one-sided case. Note, however, that the cost of a link in the one-sided case is k, while in the two-sided specification it is $2k$ and the respective bounds have to be adjusted accordingly.

Proposition 8: *If $\frac{\partial^2 \pi(x,y)}{\partial^2 y} \geq 0$ and payoffs are bounded above in the complete network, then there exists a k_*, such that for linking cost $k \leq k_*$ the efficient configuration is given by the complete network and all agents exerting same effort level $x_*(\bar{\mathbf{g}}^c) > \bar{x}(0)$. For $k \geq k_*$ the efficient configuration is given by the empty network and all agents exerting effort level $\bar{x}(0)$.*

Proof. *See the Appendix.*

Corollary 2 states that for $k \in (0, k_*)$ all PNE are inefficient. Note that if $k \in (0, k^2]$, then there exists a PNE network that is complete and therefore coincides with the efficient network. However, due to positive externalities, PNE effort levels are lower than in the efficient configuration and any PNE is therefore inefficient. For $k \in (k^2, k_*)$ the unique

PNE is the empty network, while the unique efficient network is complete. For $k \geq k_*$ the unique PNE is efficient and consists of the empty network and each agent exerting effort level $\bar{x}(0)$. It is also clear that $k^2 < k_*$ and we omit the proof. Note that the range of linking cost for which any PNE is inefficient may be thought of as large, since for a PNE to be efficient, the network must be empty.

Corollary 2: *If $k \in (0, k_*)$, then all PNE configurations are inefficient. If $k \in (k^2, k_*)$, then all PNE networks are different from the efficient network, while if $k \geq k_*$, then the unique PNE configuration is efficient.*

Proof. *Omitted.*

Remark. We briefly comment on implications of the convexity of the value function. More specifically, under one-sided link formation, the argument for our equilibrium characterization also goes through with strategic substitutes. To see this, note that when links are one-sided, then only the agent extending a link in a deviation may adjust his effort level. Therefore, Proposition 11 and Corollary 3 also hold for payoff functions such that effort levels are strategic substitutes, as long as the value function is convex (and link formation is one-sided). Under two-sided link formation both agents may adjust their effort level when creating a new link and whether Proposition 4 and Corollary 1 carry over is not that clear. However, it appears plausible that, if the payoff function is such that the value function is sufficiently convex and agents decrease their effort sufficiently little when creating a new link, then the arguments of Proposition 4 and Corollary 1 go through for this case as well.

3 Conclusion

This paper provides a model of endogenous network formation with peer effects, where peer effects are assumed to induce positive local externalities and strategic complementarities in effort levels. These features are descriptive of a wide range of social and economic and phenomena, such as educational attainment, crime, labor market participation and R&D expenditures of firms. We solve the model for a two-sided specification, where both agents need to agree to form a link, and a one-sided specification, where links are created unilaterally. In both cases the only (pairwise) Nash equilibrium network structures are nested split graphs, which are a special case of core-periphery networks, often observed in empirical work. Gross payoffs and effort levels are shown to be higher for central agents. Agents in the periphery may, however, obtain strictly higher payoffs net of linking cost. This may be informative for empirical studies, which frequently report higher payoffs and effort levels for more central agents, while typically disregarding linking cost. We show, under additional assumptions on payoffs, that the only efficient networks are the complete and the empty network. Moreover,

there exists a range of linking cost for which any (pairwise) Nash equilibrium is inefficient. For a strict subset of this range not only is any (pairwise) Nash equilibrium inefficient, but also any (pairwise) Nash equilibrium network structure is different from the efficient network. Finally, we emphasize the relevance of the convexity of the value function in obtaining core-periphery networks and, more specifically, nested split graphs. While for payoff functions such that the value function is convex, *all* (pairwise) Nash equilibria are nested split graphs, many more equilibrium structures may arise in the case of concave value functions (Baetz, 2012). Furthermore, we indicate that, as long as the value function is convex and link formation is one-sided, all Nash equilibrium networks are nested split graphs, irrespective of whether we consider strategic substitutes or strategic complements.

4 APPENDIX A - The Two-Sided Model

Proposition 1: *For any fixed network, $\bar{\mathbf{g}}$, there exists a unique NE in effort levels.*

Proof. We discern two cases. First, linear best response functions, which allow us to use the existence result provided by Ballester, Calvó-Armengol and Zenou (2006). Write the best response function as $\bar{x}_i(\sum_{j\in N_i(\bar{\mathbf{g}})} x_j) = \frac{\lambda}{\beta}\sum_{j\in N_i(\bar{\mathbf{g}})} x_j + \frac{1}{\beta}$ (where β governs own concavity and λ the strength of local interaction). A NE exists and is unique if and only if $\beta > \lambda\mu_1(\bar{\mathbf{g}})$, where $\mu_1(\bar{\mathbf{g}})$ is the largest eigenvalue of the adjacency matrix of $\bar{\mathbf{g}}$. One can show that the largest eigenvector for a graph lies between the following bounds $max\{d_{avg}(\bar{\mathbf{g}}), \sqrt{d_{max}(\bar{\mathbf{g}})}\} \leq \mu_1(\bar{\mathbf{g}}) \leq d_{max}(\bar{\mathbf{g}})$,[14] where $d_{max}(\bar{\mathbf{g}})$ is the maximum degree and $d_{avg}(\bar{\mathbf{g}})$ the average degree in network $\bar{\mathbf{g}}$. Note that then the largest eigenvector for a graph with n agents is at most $n-1$ (and maximal and equal to $n-1$ in the complete network, $\bar{\mathbf{g}}^c$). For the existence of a unique NE we therefore only need $\frac{\lambda}{\beta} < \frac{1}{n-1}$ to hold, where $\frac{\lambda}{\beta}$ is the slope of the best response function. Second, strictly concave best response functions. Define the function $f_{\bar{\mathbf{g}}} : X^n \to X^n$ as

$$f_{\bar{\mathbf{g}}}(\mathbf{x}) = \begin{pmatrix} \bar{x}(\sum\limits_{j\in N_1(\bar{\mathbf{g}})} x_j) \\ \vdots \\ \bar{x}(\sum\limits_{j\in N_n(\bar{\mathbf{g}})} x_j) \end{pmatrix}.$$

From strategic complementarities we know that $\bar{x}(y)$ is strictly increasing and, together with strict concavity of $\bar{x}(y)$, f is increasing and strictly concave. We can now apply Kennan's result (Theorem 3.3 in Kennan (2001)), which is restated here. Suppose f is an increasing and strictly concave function from $\mathbb{R}^n$ to $\mathbb{R}^n$, such that $f(\mathbf{0}) \geq \mathbf{0}$, $f(\mathbf{a}) > \mathbf{a}$ for some positive vector $\mathbf{a}$, and $f(\mathbf{b}) < \mathbf{b}$ for some vector $\mathbf{b} > \mathbf{a}$. Then f has a unique positive fixed point. Recall that we assumed $\bar{x}(0) > 0$ and therefore $f(\mathbf{0}) > \mathbf{0}$. To see that there exists a vector $\mathbf{a}$ such that $f(\mathbf{a}) > \mathbf{a}$, choose $\mathbf{a} = (\varepsilon_1, ..., \varepsilon_n)$ such that $\varepsilon_i = \varepsilon < \frac{\bar{x}(0)}{n-1}$ $\forall i \in N$ and $\varepsilon > 0$. The sum of neighbor's effort levels of an agent with $\eta_i(\bar{\mathbf{g}})$ neighbors is given by $\eta_i(\bar{\mathbf{g}})\varepsilon$. Note that we then have $\bar{x}(\eta_i(\bar{\mathbf{g}})\varepsilon) > \bar{x}(0) > \eta_i(\bar{\mathbf{g}})\varepsilon$. The first inequality holds because $\bar{x}$ is strictly increasing, while the second inequality follows from $\bar{x}(0) > \varepsilon(n-1)$ and $n-1 \geq \eta_i(\bar{\mathbf{g}})$. Next we show that there exists a vector $\mathbf{b}$ with $f(\mathbf{b}) < \mathbf{b}$. Choose again a vector such that all entries are equal, i.e., $\mathbf{b} = (b_1, ..., b_n)$ with $b = b_i$ $\forall i \in N$. The condition $f(\mathbf{b}) < \mathbf{b}$ can be written as $\bar{x}(\eta_i(\bar{\mathbf{g}})b) < b$ $\forall i \in N$. For b sufficiently large, $\bar{x}(\eta_i(\bar{\mathbf{g}})b) < b$ $\forall i \in N$ holds due to the assumption that $\frac{\partial \bar{x}(y)}{\partial y} < \frac{1}{n-1}$ for some value of y, the strict concavity of $\bar{x}(y)$ and $n \geq 3$. To show that $\mathbf{b} > \mathbf{a}$, note that we can choose ε (and therefore $\mathbf{a}$) arbitrarily close to *zero* for $\bar{x}(\eta_i(\bar{\mathbf{g}})\varepsilon) > \bar{x}(0) > \eta_i(\bar{\mathbf{g}})\varepsilon$ to hold. *Q.E.D.*

[14] See, for example, L. Lovasz, Geometric Representations of Graphs (2009).

Proposition 2: *If $k < k^1$, then the unique PNE is the complete network. If $k \geq k^1$, then there exists a PNE such that the network is empty.*

Proof. From Proposition 1 we know that there exists a unique equilibrium on a fixed network, where the only link is the one between agents i and j. Since i and j form a complete component, $x^* = x_i^* = x_j^*$ (from Lemma 1) and the corresponding gross payoffs are given by $v(x_j^*(\bar{\mathbf{g}}^e + \bar{g}_{i,j})) = \pi_i^*(x_i^*, x_j^*)$. If $k < v_i(x_j^*(\bar{\mathbf{g}}^e + \bar{g}_{i,j})) - v_i(0) = k^1$, then a pair of agents i and j finds it profitable to create the link $\bar{g}_{i,j}^*$ and create effort level x^*. Note that this is the least profitable link in any network, due to the convexity of the value function v and strict strategic complementarities in effort levels. Therefore, every pair of agents must be connected for any $k < k^1$ and the unique PNE is the complete network. If, on the other hand, $k \geq v_i(x_j^*(\bar{\mathbf{g}}^e + \bar{g}_{i,j})) - v_i(0) = k^1$, then no pair of agents can profitably deviate in the empty network. Therefore, for $k \geq k^1$ a PNE exists such that the network is empty. $Q.E.D.$

Proposition 3: *If $k > k^2$, then the unique PNE is the empty network. If $k \leq k^2$, then there exists a PNE such that the network is complete.*

Proof. Note first that the relevant deviation to consider in a complete network is an agent deleting all his links. To see this, note that due to the convexity of v, $v(hx^{c*}) - v((h-1)x^{c*}) < v((n-1)x^{c*}) - v((n-2)x^{c*})$ for all $0 < h < n-1$. That is, marginal payoffs are strictly increasing in the number of links and an agent wants to delete all his links, if any. For linking cost $k \leq k^2 = \frac{v((n-1)x^{c*}) - v(0)}{n-1}$ there therefore exists a complete PNE network, while for $k > k^2$ there does not, as an agent finds it profitable to delete all his links (and to decrease his effort level). Next we show that for $k = k^2$, the only PNE networks are the complete and the empty network. Assume to the contrary that there exists a PNE with a network that is neither empty nor complete. Consider the most profitable deviation of an agent i in network $\bar{\mathbf{g}}^* \neq \bar{\mathbf{g}}^c$, consisting of deleting h of his $\eta_i(\bar{\mathbf{g}}^*)$ links. Note that for any agent i, $n-1 \geq \eta_i(\bar{\mathbf{g}}^*) \geq h$ holds. Denote the network after proposed deviation with $\bar{\mathbf{g}}'$ and the set of agents whose links are deleted in the deviation of agent i with $H = \{j : \bar{g}_{i,j}^* = 1 \wedge \bar{g}'_{i,j} = 0\}$. We can then compare average payoffs per link in the complete network $\bar{\mathbf{g}}^{c*}$ with payoffs in $\bar{\mathbf{g}}^* \neq \bar{\mathbf{g}}^c$ and write

$$\frac{v((n-1)x^{c*}) - v(0)}{n-1} \geq \frac{v(\eta_i(\bar{\mathbf{g}}^*)x^{c*}) - v(\eta_i(\bar{\mathbf{g}}^*)x^{c*} - hx^{c*})}{\eta_i(\bar{\mathbf{g}}^*) - h} > \frac{v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^*) - v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^* - \sum_{j \in H} x_j^*)}{\eta_i(\bar{\mathbf{g}}^*) - h}.$$

The first inequality follows from the convexity of v, $n-1 \geq \eta_i(\bar{\mathbf{g}}^*)$ and $h \geq 0$. The second inequality follows from the convexity of v and $\eta_i(\bar{\mathbf{g}}^*)x^{c*} - hx^{c*} > \sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^* - \sum_{j \in H} x_j^*$ and $hx^{c*} > \sum_{j \in H} x_j^*$ (for the last two inequalities, recall that effort levels are maximal in the complete network). The threshold k^2 was defined as $k^2 = \frac{v((n-1)x^{c*}) - v(0)}{n-1}$ and the only PNE networks are therefore either the complete network and (from Proposition 2) the empty network. For linking cost $k > k^2$, agents find it profitable to delete their links and the unique PNE is the empty network. $Q.E.D.$

Lemma 3: $k^1 < k^2$.

Proof. $k^2 - k^1 = v((n-1)x^*(\bar{\mathbf{g}}^c)) - v_i(x_j^*(\bar{\mathbf{g}}^e + \bar{g}_{i,j}))$. From Lemma 2 we know that $x^*(\bar{\mathbf{g}}^c) > x_j^*(\mathbf{g}^e + \bar{g}_{i,j})$ and since v is strictly increasing, $k^2 - k^1 > 0$. *Q.E.D.*

Proposition 6: *If best response functions are linear, then there exists a PNE with a star network if and only if* $\beta > \lambda(2+\sqrt{2})$ *and either*

- $\frac{\beta(2\beta-\lambda)\lambda(\beta+\lambda)^2}{2(\beta-\lambda)^2(\beta^2-(n-1)\lambda^2)^2} \leq k \leq \frac{\lambda(2\beta^3+(n+1)\beta^2\lambda-(n-1)\lambda^3)}{2\beta(\beta^2-(n-1)\lambda^2)^2}$ *and* $1 + \frac{\beta^2(5\beta-3\lambda)}{(\beta-\lambda)^3} < n < 1 + \frac{\beta^2}{\lambda^2}$, *or*
- $k = \frac{\beta(2\beta-\lambda)\lambda(\beta+\lambda)^2}{2(\beta-\lambda)^2(\beta^2-(n-1)\lambda^2)^2}$ *and* $n = 1 + \frac{\beta^2(5\beta-3\lambda)}{(\beta-\lambda)^3}$ *holds.*

Proof. A necessary condition for a PNE star network is that effort levels are a NE on a fixed star network. From Ballester (2006) we know that a NE on a star network with n agents exists if and only if $\beta > \lambda\sqrt{n-1}$, which we write as $n < 1 + \frac{\beta^2}{\lambda^2}$ below. Note next that in a star network, all agents in the periphery access the same effort level, x_c^*, and therefore all agents in the periphery display the same effort level, x_p^*. The agent in the center, c, therefore maximizes $x_c \in argmax_{x_c \in X} x_c - \frac{\beta}{2}x_c^2 + \lambda x_c(n-1)x_p$, while for an agent in the periphery we have $x_p \in argmax_{x_p \in X} x_p - \frac{\beta}{2}x_p^2 + \lambda x_p x_c$. The reaction functions are given by $x_c(x_p) = \frac{1+\lambda x_p(n-1)}{\beta}$ and $x_p(x_c) = \frac{1+\lambda x_c}{\beta}$. Equilibrium effort levels are given by $x_c^* = \frac{\beta+\lambda(n-1)}{\beta^2-\lambda^2(n-1)}$ and $x_p^* = \frac{\beta+\lambda}{\beta^2-\lambda^2(n-1)}$. Plugging equilibrium effort levels back into the payoff function yields equilibrium gross payoffs of $\pi_c^* = \frac{\beta(\beta+\lambda(n-1))^2}{2(\beta^2-\lambda^2(n-1))^2}$ and $\pi_p^* = \frac{\beta(\beta+\lambda)^2}{2(\beta^2-\lambda^2(n-1))^2}$. Next, we calculate deviation payoffs of when two agents in the periphery create a link. Denote with $x_p^{1'}$ and $x_p^{2'}$ the effort levels of the two agents involved in the new link. A deviating agent maximizes $x_p^{1'} \in argmax_{x_p^{1'} \in X} x_p^{1'} - \frac{\beta}{2}(x_p^{1'}) + \lambda x_p^{1'}(x_c + x_p^{2'})$, which yields the following reaction function $x_p^{1'}(x_c^*, x_p^{2'}) = \frac{1+\lambda(x_c^*+x_p^{2'})}{\beta}$. Due to symmetry, deviation effort levels are given by $x_p^{1'}(x_c^*, x_p^{2'}) = x_p^{2'}(x_c^*, x_p^{1'}) = x_p' = \frac{\beta^2+\lambda\beta}{(\beta-\lambda)(\beta^2-\lambda^2(n-1))}$ and corresponding deviation gross payoffs by $\pi_p' = \frac{\beta^3(\beta+\lambda)^2}{2(\beta-\lambda)^2(\beta^2-\lambda^2(n-1))^2}$. For the existence of a star network we now need three conditions to hold simultaneously. First, linking cost k must be such that an agent in the periphery finds it profitable to link to the center of the star. Second, given the link with the center, an agent in the periphery does not find it profitable to link to another agent in the periphery. We can write this as $\frac{\beta(\beta+\lambda)^2}{2(\beta^2-\lambda^2(n-1))^2} - \frac{1}{2\beta} \geq k$ and $k \geq \frac{\beta^3(\beta+\lambda)^2}{2(\beta-\lambda)^2(\beta^2-\lambda^2(n-1))^2} - \frac{\beta(\beta+\lambda)^2}{2(\beta^2-\lambda^2(n-1))^2} = \frac{\beta(2\beta-\lambda)\lambda(\beta+\lambda)^2}{2(\beta-\lambda)^2(\beta^2-(n-1)\lambda^2)^2}$, respectively. Note that $\frac{1}{2\beta}$ are the payoffs of a singleton. Third, linking cost k must be such that the center of the star finds it profitable to link to the periphery. We can write this as $(\frac{\beta(\beta+\lambda(n-1))^2}{2(\beta^2-\lambda^2(n-1))^2} - \frac{1}{2\beta})/(n-1) = \frac{\lambda(2\beta^3+(n+1)\beta^2\lambda-(n-1)\lambda^3)}{2\beta(\beta^2-(n-1)\lambda^2)^2} \geq k$. First we show that for $\beta > \sqrt{n-1}\lambda$ the marginal payoffs of an agent in the periphery linking to the center always larger than average marginal payoffs of the center when linking to the $n-1$ agents in the periphery. Subtracting the latter from the former one obtains $\frac{(n-1)\lambda^2}{2(\beta^3-(n-1)\beta\lambda^2)}$.

The nominator is positive and we can focus on the denominator $2\beta(\beta^2-(n-1)\lambda^2)$, which is positive for $\beta > \sqrt{n-1}\lambda$. Combining the two remaining inequalities one can show after some algebraic manipulation that they simultaneously hold if and only if $\beta > (2+\sqrt{2})\lambda$ and either $\frac{\beta(2\beta-\lambda)\lambda(\beta+\lambda)^2}{2(\beta-\lambda)^2(\beta^2-(n-1)\lambda^2)^2} \leq k \leq \frac{\lambda(2\beta^3+(n+1)\beta^2\lambda-(n-1)\lambda^3)}{2\beta(\beta^2-(n-1)\lambda^2)^2}$ and $1+\frac{\beta^2(5\beta-3\lambda)}{(\beta-\lambda)^3} < n < 1+\frac{\beta^2}{\lambda^2}$ or $k = \frac{\beta(2\beta-\lambda)\lambda(\beta+\lambda)^2}{2(\beta-\lambda)^2(\beta^2-(n-1)\lambda^2)^2}$ and $n = 1+\frac{\beta^2(5\beta-3\lambda)}{(\beta-\lambda)^3}$ holds.[15] *Q.E.D.*

Proposition 8: *If $\frac{\partial^2\pi(x,y)}{\partial^2 y} \geq 0$ and payoffs are bounded above in the complete network, then there exists a k_*, such that for linking cost $k \leq k_*$ the efficient configuration is given by the complete network and all agents exerting same effort level $x_*(\bar{\mathbf{g}}^c) > \bar{x}(0)$. For $k \geq k_*$ the efficient configuration is given by the empty network and all agents exerting effort level $\bar{x}(0)$.*

Proof. We show the above result in three steps.

Step 1: A network that is not empty or complete is not efficient for any linking cost k.

Assume to the contrary that the efficient network $\bar{\mathbf{g}}_* \notin \{\bar{\mathbf{g}}^e, \bar{\mathbf{g}}^c\}$. We discern two cases.

Case 1: The efficient vector of effort levels, $\mathbf{x}_*(\bar{\mathbf{g}}_*)$, is such that all agents exert the same effort level, $x_* = x_{i*}(\bar{\mathbf{g}}_*)\ \forall i \in N$.

Note that for $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$ to be maximal it must be that $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*) \geq W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_* - \bar{g}_{i,j})$ for any existing link in $\bar{\mathbf{g}}_*$. Otherwise, deleting a link increases total payoffs and $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_* - \bar{g}_{i,j}) > W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$, contradicting our assumption that $\bar{\mathbf{g}}_*$ is efficient. That is, the marginal gross benefit of a link in $\bar{\mathbf{g}}_*$ is weakly larger than the cost of linking, even when effort levels (suboptimally) remain the same after deleting the link. But then for any pair of agents l and m such that $\bar{g}_{*l,m} = 0$, we have that $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_* + \bar{g}_{l,m}) \geq W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$, which follows directly from $\frac{\partial^2\pi(x,y)}{\partial^2 y} \geq 0$ and all agents exerting equal effort. Due to $\frac{\partial^2\pi(x,y)}{\partial x \partial y} > 0$, $\mathbf{x}_*(\bar{\mathbf{g}}_*)$ does not satisfy the first order conditions for $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_* + \bar{g}_{l,m})$ and there exists a vector of effort levels $\mathbf{x}'$ such that $W(\mathbf{x}', \bar{\mathbf{g}}_* + \bar{g}_{l,m}) > W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$.

Case 2: The efficient vector of effort levels, $\mathbf{x}_*(\bar{\mathbf{g}}_*)$, is such that there exists a pair of agents i and j with $x_{i*}(\bar{\mathbf{g}}_*) \neq x_{j*}(\bar{\mathbf{g}}_*)$.

Denote with i an agent with the weakly highest effort level and with j an agent such that $x_{j*}(\bar{\mathbf{g}}_*) < x_{i*}(\bar{\mathbf{g}}_*)$. We start by showing that $N_j(\bar{\mathbf{g}}_*) \subset N_i(\bar{\mathbf{g}}_*)$. First note that for $\bar{\mathbf{g}}_*$ to be efficient, there can be no agent l such that $l \in N_j(\bar{\mathbf{g}}_*)$ and $l \notin N_i(\bar{\mathbf{g}}_*)$. Otherwise, due to $\frac{\partial^2\pi(x,y)}{\partial^2 y} \geq 0$, switching the link of l with j to a link from l to i strictly increases payoffs. A configuration in which $N_j(\bar{\mathbf{g}}_*) = N_i(\bar{\mathbf{g}}_*)$ and $x_{j*} < x_{i*}$ is also not efficient, which follows directly from W's first order conditions for x_j and x_i. Therefore, $N_j(\bar{\mathbf{g}}_*) \subset N_i(\bar{\mathbf{g}}_*)$ holds for $x_{j*}(\bar{\mathbf{g}}_*) < x_{i*}(\bar{\mathbf{g}}_*)$. In the following we discern two subcases.

Case 2, Subcase 1: Agent i and j are not connected, $\bar{g}_{*i,j} = 0$.

We start by showing that for any configuration with $\bar{\mathbf{g}}_* \notin \{\bar{\mathbf{g}}^e, \bar{\mathbf{g}}^c\}$ and $x_{j*}(\bar{\mathbf{g}}_*) < x_{i*}(\bar{\mathbf{g}}_*)$, there exists an alternative configuration, denoted with $\mathbf{x}'$ and $\bar{\mathbf{g}}'$, which yields weakly larger

[15] The calculations were executed with Mathematica and the codes are available upon request.

payoffs than $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$. Define $\mathbf{x}'$ as $\mathbf{x}' = (x_{1*}, ..., x_{j-1*}, x_{i*}, x_{j+1*}, ..., x_{n*})$ and $\bar{\mathbf{g}}'$ as $\bar{\mathbf{g}}' = \bar{\mathbf{g}}_* + \sum_{l \in N_i(\bar{\mathbf{g}}_*)} \bar{g}_{jl}$. That is, $\mathbf{x}'$ and $\mathbf{g}'$ is a configuration where, relative to $\mathbf{x}_*$ and $\bar{\mathbf{g}}_*$, the effort level of j is increased to i's effort level in $\mathbf{x}_*$ and j is linked in $\bar{\mathbf{g}}'$ to all agents to whom i is linked to in $\bar{\mathbf{g}}_*$. Note first that for $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$ to be efficient, it must be that $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*) \geq W(\mathbf{x}'', \bar{\mathbf{g}}'')$, where $\mathbf{x}'' = (x_{1*}, ..., x_{i-1*}, x_{j*}, x_{i+1*}, ..., x_{n*})$ and $\bar{\mathbf{g}}'' = \bar{\mathbf{g}}_* - \sum_{l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)} \bar{g}_{il}$. That is, payoffs $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$ must be at least as large as in a configuration where, relative to $\mathbf{x}_*$ and $\bar{\mathbf{g}}_*$, agent i's effort level in $\mathbf{x}''$ is decreased to $j's$ effort level in $\bar{\mathbf{g}}_*$ and i is linked to agents in $\bar{\mathbf{g}}''$, to which j sustains links in $\bar{\mathbf{g}}_*$. The difference in total payoffs of configuration $\mathbf{x}_*(\bar{\mathbf{g}}_*)$, $\bar{\mathbf{g}}_*$ and $\mathbf{x}''$, $\bar{\mathbf{g}}''$ is given by

$$W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*) - W(\mathbf{x}'', \bar{\mathbf{g}}'') =$$

$$\pi_i(x_{i*}, \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)} x_{l*} + \sum_{l \in N_j(\bar{\mathbf{g}}_*)} x_{l*}) - \pi_i(x_{j*}, \sum_{l \in N_j(\bar{\mathbf{g}}_*)} x_{l*})$$

$$+ \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)} \pi_l(x_{l*}, \sum\limits_{k \in N_l(\bar{\mathbf{g}}_*)} x_{k*}) - \sum_{l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)} \pi_l(x_{l*}, \sum_{k \in N_l(\bar{\mathbf{g}}_*)} x_{k*} - x_{i*})$$

$$- \mid N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*) \mid k \geq 0.$$

The difference of the first and second term yields the gross marginal effect on i's payoffs from increasing the effort level from x_{j*} to x_{i*} and accessing effort levels of agents $l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)$. The difference of the third and fourth term yields the effect on payoffs of agents $l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)$ accessing $i's$ effort level x_{i*} in $\bar{\mathbf{g}}_*$, while the last term are the cost of additional links in $\bar{\mathbf{g}}_*$ as compared to $\bar{\mathbf{g}}''$. The inequality just follows from the assumption that the configuration $\mathbf{x}_*(\bar{\mathbf{g}}_*)$, $\bar{\mathbf{g}}_*$ is efficient. We can now write an analogous expression for $W(\mathbf{x}', \bar{\mathbf{g}}') - W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$, which is given by

$$W(\mathbf{x}', \bar{\mathbf{g}}') - W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*) =$$

$$= \pi_j(x_{i*}, \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)} x_{l*} + \sum_{l \in N_j(\bar{\mathbf{g}}_*)} x_{l*}) - \pi_j(x_{j*}, \sum_{l \in N_j(\bar{\mathbf{g}}_*)} x_{l*})$$

$$+ \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)} \pi_l(x_{l*}, \sum_{k \in N_l(\bar{\mathbf{g}}_*)} x_{k*} + x_{i*}) - \sum_{l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)} \pi_l(x_{l*}, \sum_{k \in N_l(\bar{\mathbf{g}}_*)} x_{k*})$$

$$- \mid N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*) \mid .$$

Note that the first difference (between the first and second term) in $W(\mathbf{x}', \bar{\mathbf{g}}') - W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$ is the same as in $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*) - W(\mathbf{x}'', \bar{\mathbf{g}}'')$. Likewise, cost of linking are the same, so that we can focus on the differences of the third and fourth terms, which stem from the effect on agents $l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)$. From $\frac{\partial^2 \pi(x,y)}{\partial^2 y} \geq 0$ we know that the difference of the third and fourth term is weakly larger in $W(\mathbf{x}', \bar{\mathbf{g}}') - W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$ than in $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*) - W(\mathbf{x}'', \bar{\mathbf{g}}'')$ and therefore $W(\mathbf{x}', \bar{\mathbf{g}}') - W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*) \geq W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*) - W(\mathbf{x}'', \bar{\mathbf{g}}'')$. Note that $\mathbf{x}'$, $\bar{\mathbf{g}}'$ is suboptimal since $\frac{\partial^2 \pi(x,y)}{\partial x \partial y} > 0$ and the first order conditions for agents $l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)$

are not satisfied in $\mathbf{x}'$, $\bar{\mathbf{g}}'$. That is, there exists a vector of effort levels $\mathbf{x}_*(\bar{\mathbf{g}}')$ such that $W(\mathbf{x}_*(\bar{\mathbf{g}}'), \bar{\mathbf{g}}') > W(\mathbf{x}', \bar{\mathbf{g}}') \geq W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$. We have reached a contradiction.

Case 2, Subcase 2: Agent i and j are connected, $\bar{g}_{*i,j} = 1$.

By a similar argument as above, we show that a profile of effort levels $\mathbf{x}'$ and a network $\bar{\mathbf{g}}'$ yields strictly larger payoffs than $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$, so that $W(\mathbf{x}', \bar{\mathbf{g}}') > W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$, where $\mathbf{x}' = (x_{1*}, ..., x_{j-1*}, x_{i*}, x_{j+1*}, ..., x_{n*})$ and $\bar{\mathbf{g}}' = \bar{\mathbf{g}}_* + \sum_{l \in N_i(\bar{\mathbf{g}}_*)} \bar{g}_{jl}$. That is, a configuration $\mathbf{x}'$ and $\bar{\mathbf{g}}'$, where (relative to $\mathbf{x}_*$ and $\bar{\mathbf{g}}_*$) the effort level of j is increased to i's effort level in $\mathbf{x}_*$ and j is linked in $\bar{\mathbf{g}}'$ to all agents to whom i is linked to in $\bar{\mathbf{g}}_*$. Note that for $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$ to be efficient, it must be that $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*) \geq W(\mathbf{x}'', \bar{\mathbf{g}}'')$, where $\mathbf{x}'' = (x_{1*}, ..., x_{i-1*}, x_{j*}, x_{i+1*}, ..., x_{n*})$ and $\bar{\mathbf{g}}'' = \bar{\mathbf{g}}_* - \sum_{l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)} \bar{g}_{il}$. That is, payoffs $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$ must be at least as large as in a configuration where, relative to $\mathbf{x}_*$ and $\bar{\mathbf{g}}_*$, agent i's effort level in $\mathbf{x}''$ is decreased to j's effort level in $\bar{\mathbf{g}}_*$ and i links to agents in $\bar{\mathbf{g}}''$, to which j sustains links in $\bar{\mathbf{g}}_*$. Since $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$ is assumed to be efficient, it must be that $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*) - W(\mathbf{x}'', \bar{\mathbf{g}}'') \geq 0$. We can write for $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*) - W(\mathbf{x}'', \bar{\mathbf{g}}'')$

$$W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*) - W(\mathbf{x}'', \bar{\mathbf{g}}'') =$$

$$= \pi_i(x_{i*}, \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)} x_{l*} + \sum_{l \in N_j(\bar{\mathbf{g}}_*) \setminus \{i\}} x_{l*} + x_{j*}) - \pi_i(x_{j*}, \sum_{l \in N_j(\bar{\mathbf{g}}_*) \setminus \{i\}} x_{l*} + x_{j*})$$

$$+ \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)} \pi_l(x_{l*}, \sum_{k \in N_l(\bar{\mathbf{g}}_*)} x_{k*}) - \sum_{l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)} \pi_l(x_{l*}, \sum_{k \in N_l(\bar{\mathbf{g}}_*)} x_{k*} - x_{i*})$$

$$- \mid N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*) \mid k \geq 0.$$

While for $W(\mathbf{x}', \bar{\mathbf{g}}') - W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$ we have

$$W(\mathbf{x}', \bar{\mathbf{g}}') - W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*) =$$

$$= \pi_j(x_{i*}, \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)} x_{l*} + \sum_{l \in N_j(\bar{\mathbf{g}}_*) \setminus \{i\}} x_{l*} + x_{i*}) - \pi_j(x_{j*}, \sum_{l \in N_j(\bar{\mathbf{g}}_*) \setminus \{i\}} x_{l*} + x_{i*})$$

$$+ \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)} \pi_l(x_{l*}, \sum_{k \in N_l(\bar{\mathbf{g}}_*)} x_{k*} + x_{i*}) - \sum_{l \in N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*)} \pi_l(x_{l*}, \sum_{k \in N_l(\bar{\mathbf{g}}_*)} x_{k*})$$

$$- \mid N_i(\bar{\mathbf{g}}_*) \setminus N_j(\bar{\mathbf{g}}_*) \mid .$$

Again the linking cost is the same in both expressions and the difference between the third and fourth term is, due to $\frac{\partial^2 \pi(x,y)}{\partial^2 y} \geq 0$, weakly larger in $W(\mathbf{x}', \bar{\mathbf{g}}') - W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$ than in $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*) - W(\mathbf{x}'', \bar{\mathbf{g}}'')$. However, the first difference is strictly larger in $W(\mathbf{x}', \bar{\mathbf{g}}') - W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$ than in $W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*) - W(\mathbf{x}'', \bar{\mathbf{g}}'')$. This follows from $x_{i*}(\bar{\mathbf{g}}_*) > x_{j*}(\bar{\mathbf{g}}_*)$ and strict strategic complementarities (strictly increasing differences of π). Therefore, $W(\mathbf{x}', \bar{\mathbf{g}}') > W(\mathbf{x}_*(\bar{\mathbf{g}}_*), \bar{\mathbf{g}}_*)$ and the efficient network is either complete or empty.

Step 2: Efficient effort levels in the complete network are equal.

Assume to the contrary that there exist a pair of agents such that $x_{i*}(\bar{\mathbf{g}}^c) \neq x_{j*}(\bar{\mathbf{g}}^c)$ and assume without loss of generality that $x_{i*}(\bar{\mathbf{g}}^c) > x_{j*}(\bar{\mathbf{g}}^c)$. For $W(x_{i*}, x_{j*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c)$ to be

efficient we need that $W(x_{i*}, x_{j*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c) - W(x_{j*}, x_{j*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c) \geq 0$. In the following we show that

$$W(x_{i*}, x_{i*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c) - W(x_{i*}, x_{j*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c) >$$

$$W(x_{i*}, x_{j*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c) - W(x_{j*}, x_{j*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c) \geq 0$$

and therefore $W(x_{i*}, x_{i*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c) > W(x_{i*}, x_{j*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c)$. Again we write out the respective expressions. First,

$$W(x_{i*}, x_{i*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c) - W(x_{i*}, x_{j*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c) =$$

$$= \pi_i(x_{i*}, \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*)\setminus\{j\}} x_{l*} + x_{i*}) - \pi_i(x_{i*}, \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*)\setminus\{j\}} x_{l*} + x_{j*})$$

$$+\pi_j(x_{i*}, \textstyle\sum_{l \in N_j(\bar{\mathbf{g}}_*)\setminus\{i\}} x_{l*} + x_{i*}) - \pi_j(x_{j*}, \textstyle\sum_{l \in N_j(\bar{\mathbf{g}}_*)\setminus\{i\}} x_{l*} + x_{i*})$$

$$+ \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*)\setminus\{i,j\}} \pi_l(x_{l*}, x_{j*}, \textstyle\sum_{k \in N_l(\bar{\mathbf{g}}_*)\setminus\{i,j\}} x_{k*} + x_{i*} + x_{i*})$$

$$- \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*)\setminus\{i,j\}} \pi_l(x_{l*}, x_{j*}, \textstyle\sum_{k \in N_l(\bar{\mathbf{g}}_*)\setminus\{i,j\}} x_{k*} + x_{i*} + x_{j*})$$

Second,

$$W(x_{i*}, x_{j*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c) - W(x_{j*}, x_{j*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c) =$$

$$= \pi_j(x_{j*}, \textstyle\sum_{l \in N_j(\bar{\mathbf{g}}_*)\setminus\{i\}} x_{l*} + x_{i*}) - \pi_j(x_{j*}, \textstyle\sum_{l \in N_j(\bar{\mathbf{g}}_*)\setminus\{i\}} x_{l*} + x_{j*})$$

$$+\pi_i(x_{i*}, \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*)\setminus\{j\}} x_{l*} + x_{j*}) - \pi_i(x_{j*}, \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*)\setminus\{j\}} x_{l*} + x_{j*})$$

$$+ \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*)\setminus\{i,j\}} \pi_l(x_{l*}, x_{j*}, \textstyle\sum_{k \in N_l(\bar{\mathbf{g}}_*)\setminus\{i,j\}} x_{k*} + x_{i*} + x_{j*})$$

$$- \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*)\setminus\{i,j\}} \pi_l(x_{l*}, x_{j*}, \textstyle\sum_{k \in N_l(\bar{\mathbf{g}}_*)\setminus\{i,j\}} x_{k*} + x_{j*} + x_{j*})$$

To evaluate

$$W(x_{i*}, x_{i*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c) - W(x_{i*}, x_{j*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c) >$$

$$W(x_{i*}, x_{j*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c) - W(x_{j*}, x_{j*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c)$$

we again compare across pairs of terms of the expressions. Comparing the respective first and second terms we have

$$\pi_i(x_{i*}, \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*)\setminus\{j\}} x_{l*} + x_{i*}) - \pi_i(x_{i*}, \textstyle\sum_{l \in N_i(\bar{\mathbf{g}}_*)\setminus\{j\}} x_{l*} + x_{j*}) >$$

$$\pi_j(x_{j*}, \textstyle\sum_{l \in N_j(\bar{\mathbf{g}}_*)\setminus\{i\}} x_{l*} + x_{i*}) - \pi_j(x_{j*}, \textstyle\sum_{l \in N_j(\bar{\mathbf{g}}_*)\setminus\{i\}} x_{l*} + x_{j*}).$$

The inequality follows from $\frac{\partial^2 \pi(x,y)}{\partial x \partial y} > 0$. For the third and fourth term we can write

$\pi_j(x_{i*}, \sum_{l\in N_j(\bar{\mathbf{g}}_*)\setminus\{i\}} x_{l*} + x_{i*}) - \pi_j(x_{j*}, \sum_{l\in N_j(\bar{\mathbf{g}}_*)\setminus\{i\}} x_{l*} + x_{i*}) >$

$\pi_i(x_{i*}, \sum_{l\in N_i(\bar{\mathbf{g}}_*)\setminus\{j\}} x_{l*} + x_{j*}) - \pi_i(x_{j*}, \sum_{l\in N_i(\bar{\mathbf{g}}_*)\setminus\{j\}} x_{l*} + x_{j*}),$

again this follows from $\frac{\partial^2\pi(x,y)}{\partial x\partial y} > 0$. For the difference of the fifth and sixth term we have

$\sum_{l\in N_i(\bar{\mathbf{g}}_*)\setminus\{i,j\}} \pi_l(x_{l*}, x_{j*}, \sum_{k\in N_l(\bar{\mathbf{g}}_*)\setminus\{i,j\}} x_{k*} + x_{i*} + x_{i*})$

$-\sum_{l\in N_i(\bar{\mathbf{g}}_*)\setminus\{i,j\}} \pi_l(x_{l*}, x_{j*}, \sum_{k\in N_l(\bar{\mathbf{g}}_*)\setminus\{i,j\}} x_{k*} + x_{i*} + x_{j*}) \geq$

$\sum_{l\in N_i(\bar{\mathbf{g}}_*)\setminus\{i,j\}} \pi_l(x_{l*}, x_{j*}, \sum_{k\in N_l(\bar{\mathbf{g}}_*)\setminus\{i,j\}} x_{k*} + x_{i*} + x_{j*})$

$-\sum_{l\in N_i(\bar{\mathbf{g}}_*)\setminus\{i,j\}} \pi_l(x_{l*}, x_{j*}, \sum_{k\in N_l(\bar{\mathbf{g}}_*)\setminus\{i,j\}} x_{k*} + x_{j*} + x_{j*}),$

which follows form $\frac{\partial^2\pi(x,y)}{\partial^2 y} \geq 0$. Therefore, $W(x_{i*}, x_{i*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c) > W(x_{i*}, x_{j*}, \mathbf{x}_{-i,-j*}, \bar{\mathbf{g}}^c)$ and we have reached a contradiction. That is, in the complete network effort levels must be such that $x_*(\bar{\mathbf{g}}^c) = x_{i*}\ \forall i \in N$.

Step 3: There exists a linking cost k_*, such that for $k \leq k_*$ the efficient network is complete, while for $k \geq k_*$ the efficient network is empty.

Note first that for any vector of effort levels $\mathbf{x}$, total gross payoffs are larger in the complete network than in any other network, while the difference in linking cost is bounded by $kn(n-1)$. That is, if payoffs are bounded in the complete network, then they are bounded for any other network. From the first order conditions and $\frac{\partial^2\pi(x,y)}{\partial x\partial y} > 0$ we know that $x_*(\bar{\mathbf{g}}^c) > \bar{x}(0)$. Since payoffs are bounded and total linking cost is linearly increasing in k for any network that is not empty, there exists a linking cost k_*, such that for linking cost k smaller or equal than k_*, the efficient network is complete, while for linking cost k larger or equal than k_*, the efficient network is empty (with agents exerting effort level $\bar{x}(0)$). *Q.E.D.*

5 APPENDIX B - The One-Sided Model

5.1 Model Description

The one-sided specification differs from the two-sided model in that only one agent needs to extend a link (and bear the cost), in order for a pair of agents to benefit from each other's effort level. This allows us to use Nash equilibrium as equilibrium concept. Note that under pairwise Nash equilibrium, pairs of agents can create only *one* link at a time and *both* agents may adjust their effort levels. Under Nash equilibrium, we consider deviations where an agent may extend *multiple* links (and simultaneously delete any subset of existing ones), but only the (single) deviating agent may adjust effort levels.

Let again $N = \{1, 2, ..., n\}$ be the set of players, with $n \geq 3$. As before, each player i chooses a personal effort level $x_i \in X$ and a set of links, which are represented as a row vector

$\mathbf{g}_i = (g_{i,1},..., g_{ii-1}, g_{ii+1},..., g_{in})$, where $g_{ij} \in \{0,1\}$ for each $j \in N\backslash\{i\}$. Assume $X = [0, +\infty)$ and $\mathbf{g}_i \in G_i = \{0,1\}^{n-1}$. The set of strategies of i is denoted by $S_i = X \times G_i$ and the set of strategies of all players by $S = S_1 \times S_2 \times ... \times S_n$. A strategy profile $\mathbf{s} = (\mathbf{x}, \mathbf{g}) \in S$ again specifies the individual effort level of each player, $\mathbf{x} = (x_1, x_2, ..., x_n)$, and a set of links $\mathbf{g} = (\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_n)$. Agent i is said to sustain or extend a link to j, if $g_{i,j} = 1$ and to receive a link from j, if $g_{j,i} = 1$. The network of relations $\mathbf{g}$ is a directed graph, i.e., it is possible that $g_{i,j} \neq g_{j,i}$. Let $N_i(\mathbf{g}) = \{j \in N : g_{i,j} = 1\}$ be the set of agents i has extended a link to and define $\eta_i(\mathbf{g}) = |N_i(\mathbf{g})|$. Call the closure of $\mathbf{g}$ an undirected network, denoted by $\bar{\mathbf{g}} = cl(\mathbf{g})$, where $\bar{g}_{i,j} = \max\{g_{i,j}, g_{j,i}\}$ for each i and j in N. Denote with $N_i(\bar{\mathbf{g}}) = \{j \in N : \bar{g}_{i,j} = 1\}$ the set of players that are directly connected to i. The effort level of i's direct neighbors can then be written as $y_i = \sum_{j \in N_i(\bar{\mathbf{g}})} x_j$. We will drop the subscript of y_i when it is clear from the context. Given a network $\mathbf{g}$, $\mathbf{g} + g_{i,j}$ and $\mathbf{g} - g_{i,j}$ have the following interpretation. When $g_{i,j} = 0$ in $\mathbf{g}$, $\mathbf{g} + g_{i,j}$ adds the link $g_{i,j} = 1$, while if $g_{i,j} = 1$ in $\mathbf{g}$, then $\mathbf{g} + g_{i,j} = \mathbf{g}$. Similarly, if $g_{i,j} = 1$ in $\mathbf{g}$, $\mathbf{g} - g_{i,j}$ deletes the link $g_{i,j}$, while if $g_{i,j} = 0$ in $\mathbf{g}$, then $\mathbf{g} - g_{i,j} = \mathbf{g}$. The network is said to be empty and denoted by $\bar{\mathbf{g}}^e$ if $\bar{g}_{i,j} = 0 \ \forall i, j \in N$ and complete and denoted by $\bar{\mathbf{g}}^c$ if $\bar{g}_{i,j} = 1 \ \forall i, j \in N$.

Payoffs of player i under strategy profile $\mathbf{s} = (\mathbf{x}, \mathbf{g})$ are given by

$$\Pi_i(\mathbf{s}) = \pi(x_i, y_i) - \eta_i(\mathbf{g})k,$$

where k denotes the cost of extending a link. The assumptions on the payoff function are as in the one-sided specification. A Nash equilibrium is a strategy profile $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$ such that

$$\Pi_i(\mathbf{s}_i^*, \mathbf{s}_{-i}^*) \geq \Pi_i(\mathbf{s}_i, \mathbf{s}_{-i}^*), \ \forall \mathbf{s}_i \in S_i, \forall i \in N.$$

Denote the directed equilibrium network by $\mathbf{g}^*$ and the undirected equilibrium network by $\bar{\mathbf{g}}^*$. Social welfare is again defined as the sum of individual payoffs. For any strategy profile $\mathbf{s}$, social welfare is given by

$$W(\mathbf{s}) = \sum_{i \in N} \Pi_i(\mathbf{s}).$$

A profile $\widetilde{\mathbf{s}}$ is socially efficient if $W(\widetilde{\mathbf{s}}) \geq W(\mathbf{s})$, $\forall \mathbf{s} \in S$.

5.2 Analysis

Note that in the main part of the paper we assume the network to be fixed for Proposition 1, Lemma 1, Lemma 2 and Lemma 3 and therefore these results do not depend on the specification of link formation. Similarly, for the efficiency result in Proposition 8, the link formation process is not relevant and the result from the two-sided case carries over. Note,

however, that the cost for a link in the one-sided case is k, while in the two-sided specification it is $2k$ and the respective bounds have to be adjusted accordingly. We start by showing that in any NE there can be at most one link between any pair of players.

Lemma 7: *In any NE $\mathbf{s}^*=(\mathbf{x}^*, \mathbf{g}^*)$, there is at most one directed link between any pair of agents $i, j \in N$.*

Proof. Assume that $\mathbf{s}^*=(\mathbf{x}^*, \mathbf{g}^*)$ is a Nash equilibrium and that $g_{i,j} = g_{j,i} = 1$. But then i can profitably deviate by cutting the link to j, such that $g_{i,j} = 0$. Gross payoffs remain unchanged, while i's linking total cost decrease by k. *Q.E.D.*

In Lemma 8 we show, again using the convexity of the value function, that in any Nash equilibrium, if i extends a link to l, then i must also be connected to agent k, for any k such that $x_k^* \geq x_l^*$. Note that we do not require that i extends a link to k, but only that i and k are connected. That is, it may be agent k extending the link to agent i.

Lemma 8: *In any NE $\mathbf{s}^*=(\mathbf{x}^*, \mathbf{g}^*)$, if $g_{i,l}^* = 1$ then $\bar{g}_{i,k}^* = 1$ $\forall k : x_k^* \geq x_l^*$.*

Proof. For $g_{i,j}^* = 1$ to be part of a NE, it must be that $v(\sum_{j \in N_i(\bar{\mathbf{g}})} x_j^*) - v(\sum_{j \in N_i(\bar{\mathbf{g}})} x_j^* - x_l^*) \geq k$. Assume, contrary to the above statement, that $\bar{g}_{i,k}^* = 0$ for some k with $x_k^* \geq x_l^*$. This, however, can not be a NE, since i then finds it profitable to extend a link to agent k. To see this, note that $v(\sum_{j \in N_i(\mathbf{g})} x_j^* + x_k) - v(\sum_{j \in N_i(\mathbf{g})} x_j^*) > v(\sum_{j \in N_i(\mathbf{g})} x_j^*) - v(\sum_{j \in N_i(\mathbf{g})} x_j^* - x_l^*) \geq k$, where the inequalities follow from the convexity of the value function. We have reached a contradiction and therefore $\bar{g}_{i,k}^* = 1$ for all agents k with $x_k^* \geq x_l^*$. *Q.E.D.*

The following Lemma shows that if i extends a link to l, then any agent k with a higher or equal effort level than i must also be connected to l. Again this follows from the convexity of the value function.

Lemma 9: *In any NE $\mathbf{s}^*=(\mathbf{x}^*, \mathbf{g}^*)$, if $g_{i,l}^* = 1$ then $\bar{g}_{k,l}^* = 1$ $\forall k : x_k^* \geq x_i^*$.*

Proof. For $g_{i,j}^* = 1$ to be part of a NE, it must be that $v(\sum_{j \in N_i(\bar{\mathbf{g}})} x_j^*) - v(\sum_{j \in N_i(\bar{\mathbf{g}})} x_j^* - x_l^*) \geq k$. Assume, contrary to the above statement, that $\bar{g}_{k,l}^* = 0$ for some k with $x_k^* \geq x_i^*$. Note next that, for $x_k^* \geq x_i^*$ to hold we must have $\sum_{j \in N_k(\bar{\mathbf{g}})} x_j^* \geq \sum_{j \in N_i(\bar{\mathbf{g}})} x_j^*$, which follows directly from strict strategic complementarities. Therefore, $v(\sum_{j \in N_k(\mathbf{g})} x_j^* + x_l^*) - v(\sum_{j \in N_k(\mathbf{g})} x_j^*) > v(\sum_{j \in N_i(\mathbf{g})} x_j^*) - v(\sum_{j \in N_i(\mathbf{g})} x_j^* - x_l) \geq k$, where the inequalities follow from the convexity of the value function and we have reached a contradiction. *Q.E.D.*

Similar to the two-sided specification, we again define two bounds, k^1 and k^2. In Proposition 9 we show that for k smaller than k^1, the unique Nash equilibrium is such that the network is complete, while for k larger or equal than k^1, there exists a Nash equilibrium

such that the network is empty. Proposition 10 shows that for linking cost larger than k^2, the unique Nash equilibrium is such that the network is empty, while for k smaller or equal to k^2, there exists a Nash equilibrium such that the network is complete. Note that the thresholds k^1 and k^2 are defined differently from the two-sided specification. Due to the convexity of the value function, the most profitable deviation in the empty network is to extend a link to all remaining agents (where no agent other than the deviating agent adjust his effort level). The second threshold, k^2, is the maximal linking cost that can be sustained in the complete network. Due to the convexity of the value function the agent with the fewest incoming links has the greatest incentives to deviate. That is, the network that may sustain the maximum linking cost is the one where incoming and outgoing links are evenly distributed. With n agents there are $\frac{n(n-1)}{2}$ pairs of agents. For n odd this implies that when incoming and outgoing links are evenly distributed, each agent has $\frac{n-1}{2}$ incoming and $\frac{n-1}{2}$ outgoing links. For n even, $\frac{n}{2}$ agents have $\frac{n}{2}$ incoming and $\frac{n-2}{2}$ outgoing links and $\frac{n}{2}$ agents have $\frac{n-2}{2}$ incoming and $\frac{n}{2}$ outgoing links. For simplicity we assume in the following that n is odd. Analogous results are easily derived for n even.

Definition 2: $k^1 = \frac{v((n-1)\bar{x}(0))-v(0)}{n-1}$ and $k^2 = \frac{2}{n-1}(v((n-1)x^*(\mathbf{g}^c)) - v(\frac{n-1}{2}x^*(\mathbf{g}^c))$.

Proposition 9: *If $k < k^1$, then the unique NE network is the complete network. If $k \geq k^1$, then there exists a NE such that the network is empty.*

Proof. If $k < k^1$ then an agent finds it profitable to create a link to all remaining $n-1$ agents in an empty network, since average payoffs per link are given by $\frac{v((n-1)\bar{x}(0))-v(0)}{n-1}$ with $\frac{v((n-1)\bar{x}(0))-v(0)}{n-1} > k$. This is the most profitable deviation in an empty network, due to the convexity of the value function. Assume there exists a $\mathbf{g}^* \notin \{\mathbf{g}^e, \mathbf{g}^c\}$ with $k < k^1$. Consider the deviation of an agent i, with $\eta_i(\bar{\mathbf{g}}^*) < n-1$, who links to all agents he is not connected to in $\mathbf{g}^*$, i.e., $k \notin N_i(\bar{\mathbf{g}}^*)$. To simplify notation, we write η_i for $\eta_i(\bar{\mathbf{g}}^*)$ in the following. Average marginal payoffs per link of proposed deviation are given by

$$\frac{v(\sum_{j\in N_i\setminus\{i\}} x_j^*)-v(\sum_{j\in N_i(\bar{\mathbf{g}}^*)} x_j^*)}{n-1-\eta_i}.$$

We can now write

$$\frac{v(\sum_{j\in N_i\setminus\{i\}} x_j^*)-v(\sum_{j\in N_i(\bar{\mathbf{g}}^*)} x_j^*)}{n-1-\eta_i} > \frac{v((n-1)\bar{x}(0))-v(\eta_i\bar{x}(0))}{n-1-\eta_i}.$$

To see that the inequality holds, note first that $\sum_{j\in N_i(\bar{\mathbf{g}}^*)} x_j^* > \eta_i\bar{x}(0)$, since $\bar{x}(0)$ is the effort level of a singleton and connected agents display strictly larger effort levels in any NE. Second, that $\sum_{j\in N_i\setminus\{i\}} x_j^* - \sum_{j\in N_i(\bar{\mathbf{g}}^*)} x_j^* \geq (n-1-\eta_i)\bar{x}(0)$. That is, when linking to the remaining $n-1-\eta_i$ agents in proposed deviation, the minimum additional effort level accessed is given by $(n-1-\eta_i)\bar{x}(0)$. The condition above then follows from the convexity of the value function. Note next that

$$\frac{v((n-1)\bar{x}(0))-v((\eta_i\bar{x}(0))}{n-1-\eta_i} > \frac{v((n-1)\bar{x}(0))-v(0)}{n-1}$$

also holds, again from the convexity of the value function, and we therefore have

$$\frac{v(\sum_{j\in N_i\setminus\{i\}} x_j^*)-v(\sum_{j\in N_i(\bar{\mathbf{g}}^*)} x_j^*)}{n-1-\eta_i} > \frac{v((n-1)\bar{x}(0))-v(0)}{n-1} > k.$$

Therefore, if $k < k^1$, a profitable deviation exists in any $\mathbf{g}^* \neq \mathbf{g}^c$. It is easy to see that there then exists a NE with $\mathbf{g}^* = \mathbf{g}^c$ for $k < k^1$. If, on the other hand $k \geq k^1$, then no agent can profitably deviate in the empty network, and a NE exists such that $\mathbf{g}^* = \mathbf{g}^e$. *Q.E.D.*

Before proceeding to Proposition 10, we show that in any Nash equilibrium such that the network is neither empty nor complete, there exists an agent that extends at least one link and has less than $\frac{n-1}{2}$ incoming links. This result is useful when proving that the network that can be sustained at the highest linking cost is the complete network with evenly distributed incoming links.

Lemma 10: *In any NE network that is neither empty nor complete, there exists an agent with $\eta_i(\mathbf{g}) \geq 1$ and $\eta_i(\bar{\mathbf{g}}) - \eta_i(\mathbf{g}) < \frac{n-1}{2}$.*

Proof. We discern two cases. First, every agent extends at least one link and receives at least $\frac{n-1}{2}$ links. That is, $\eta_i(\mathbf{g}) \geq 1$ and $\eta_i(\bar{\mathbf{g}}) - \eta_i(\mathbf{g}) \geq \frac{n-1}{2}$ $\forall i \in N$. But then there are at least $\frac{n(n-1)}{2}$ links in the network and the network is complete. Second, not everyone extends a link. Assume there are $k < n$ agents who extend a link. Since there are no incoming links from the remaining $n-k$ agents, the maximum number of incoming links among the k agents extending a link is given by $\frac{k(k-1)}{2}$. That is, on average an agent has $\frac{k-1}{2}$ incoming links. The maximum of the minimum number of incoming links is therefore given by $\frac{k-1}{2}$ (if the network among the k agents is complete and incoming links are distributed evenly). Since $k < n$, there must be one agent with at most $\frac{k-1}{2} < \frac{n-1}{2}$ incoming links who extends a link. *Q.E.D.*

Proposition 10: *If $k > k^2$, then the unique NE is the empty network. If $k \leq k^2$, then there exists a NE such that the network is complete.*

Proof. We will first show that the highest cost that can be sustained under the complete network is given by k^2. Denote the NE effort level in a complete network with x^{c*}. In the complete network the agent extending the highest number of links (and therefore receiving the fewest number of links) is the one with the highest incentives to delete his links. To see this, write

$$\frac{v((n-1)x^{c*})-v((n-1-h)x^{c*})}{n-1-h} > \frac{v((n-1)x^{c*})-v((n-1-h')x^{c*})}{n-1-h'},$$

where $h' > h > 0$. The inequality holds by the convexity of the value function. The network that minimizes the maximum number of links extended by agents in a network is such that each agent extends $\frac{n-1}{2}$ links (and receives $\frac{n-1}{2}$ links). Therefore, the highest linking cost that can be sustained in a complete network are given by $k = k^2$. It is easy to see that for $k < k^2$ there exists a NE such that $\bar{\mathbf{g}}^* = \bar{\mathbf{g}}^c$. Assume next, and contrary to the above statement, that for $k > k^2$ there exists a NE such that $\bar{\mathbf{g}}^* \neq \bar{\mathbf{g}}^e$. k^2 was derived as the maximal payoffs sustainable in a complete network, and therefore for $k > k^2$, there does not exist a NE such that $\bar{\mathbf{g}}^* = \bar{\mathbf{g}}^c$. Next, we show that for $k > k^2$ there also does not exist a NE with $\bar{\mathbf{g}}^* \notin \{\bar{\mathbf{g}}^e, \bar{\mathbf{g}}^c\}$. Assume the contrary. Pick an agent with less than $\frac{n-1}{2}$ incoming links, $\eta_i(\bar{\mathbf{g}}^*) - \eta_i(\mathbf{g}^*) < \frac{n-1}{2}$, and at least one outgoing link, $\eta_i(\mathbf{g}^*) \geq 1$. We know from Lemma 8 that such an agent exists in $\bar{\mathbf{g}}^* \notin \{\bar{\mathbf{g}}^e, \bar{\mathbf{g}}^c\}$. We consider a deviation where this agent deletes all his links. To see that the deviation considered is profitable, note that in the complete network the average marginal payoff from extending links to all remaining agents is larger for an agent with $\frac{n-1}{2}$ incoming links, than for an agent with fewer incoming links, i.e., for an agent with$\eta_i(\bar{\mathbf{g}}^*) - \eta_i(\mathbf{g}^*) < \frac{n-1}{2}$. From $\eta_i(\bar{\mathbf{g}}^*) - \eta_i(\mathbf{g}^*) < \frac{n-1}{2}$ we have $n - 1 - (\eta_i(\bar{\mathbf{g}}^*) - \eta_i(\mathbf{g}^*)) > \frac{n-1}{2}$. The following inequality then holds, again by the convexity of v,

$$\frac{v((n-1)x^{c*}) - v(\frac{n-1}{2}x^{c*})}{\frac{n-1}{2}} > \frac{v((n-1)x^{c*}) - v((\eta_i(\bar{\mathbf{g}}^*) - \eta_i(\mathbf{g}^*))x^{c*})}{n-1-(\eta_i(\bar{\mathbf{g}}^*) - \eta_i(\mathbf{g}^*))}.$$

Note that, given $\eta_i(\bar{\mathbf{g}}^*) - \eta_i(\mathbf{g}^*) < \frac{n-1}{2}$ incoming links, average marginal payoffs are highest when linking to all remaining agents with effort level x^{c*}. As $\eta_i(\bar{\mathbf{g}}^*)$ is at most $n-1$, we can write

$$\frac{v((n-1)x^{c*}) - v((\eta_i(\bar{\mathbf{g}}^*) - \eta_i(\mathbf{g}^*))x^{c*})}{n-1-(\eta_i(\bar{\mathbf{g}}^*) - \eta_i(\mathbf{g}^*))} \geq \frac{v(\eta_i(\bar{\mathbf{g}}^*)x^{c*}) - v((\eta_i(\bar{\mathbf{g}}^*) - \eta_i(\mathbf{g}^*))x^{c*})}{\eta_i(\mathbf{g}^*)}.$$

Finally, note that by Lemma 2 effort levels are maximal in the complete component and by the convexity of v we can therefore write

$$\frac{v(\eta_i(\bar{\mathbf{g}}^*)x^{c*}) - v((\eta_i(\bar{\mathbf{g}}^*) - \eta_i(\mathbf{g}^*))x^{c*})}{\eta_i(\mathbf{g}^*)} > \frac{v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^*) - v(\sum_{j \in N_i(\mathbf{g}^*)} x_j^*)}{\eta_i(\mathbf{g}^*)}.$$

Average marginal payoffs are highest in the complete network where each agent extends $\frac{n-1}{2}$ links and therefore for $k > k^2$ the empty network is the unique NE network. *Q.E.D.*

Lemma 11 shows that $k^1 < k^2$. We have therefore shown that, for linking cost smaller than k^1, the unique NE is the complete network, while, for linking cost larger than k^2, the unique NE network is the empty network. For linking cost $k \in [k^1, k^2]$ the complete and the empty network are Nash equilibria.

Lemma 11: $0 < k^1 < k^2$.

Proof. Recall the definitions of $k^1 = \frac{v((n-1)\bar{x}(0))-v(0)}{n-1}$ and $k^2 = \frac{2(v((n-1)x^{c*})-v((\frac{n-1}{2})x^{c*}))}{n-1}$. The inequalities then follow from $\bar{x}(0) > 0$, $x^{c*} > \bar{x}(0)$ and the convexity of the value function. *Q.E.D.*

The following Lemma shows that in any Nash equilibrium, if a pair of agents exert same effort levels, then they must share the same neighborhoods. The proof is a direct consequence of the convexity of the value function.

Lemma 12: *In any NE* $\mathbf{s}^*=(\mathbf{x}^*, \mathbf{g}^*)$, $x_i^* = x_k^* \Leftrightarrow N_i(\bar{\mathbf{g}}^*) \setminus \{k\} = N_k(\bar{\mathbf{g}}^*) \setminus \{i\}$.

Proof. First, $N_i(\bar{\mathbf{g}}^*) \setminus \{k\} = N_k(\bar{\mathbf{g}}^*) \setminus \{i\} \Rightarrow x_i^* = x_k^*$. If $\bar{g}_{i,k}^* = 0$, then i and k access the same effort level, i.e., $y_i = \sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^* = y_k = \sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^*$ and therefore $x_i^* = x_k^*$. Assume next that $\bar{g}_{i,k}^* = 1$ and, without loss of generality that $x_i^* > x_k^*$. But then k accesses a higher effort level than i, $y_i = \sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^* < y_k = \sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^*$, and we have reached a contradiction. Second, $x_i^* = x_k^* \Rightarrow N_i(\bar{\mathbf{g}}^*) \setminus \{k\} = N_k(\bar{\mathbf{g}}^*) \setminus \{i\}$. Assume to the contrary that $x_i^* = x_k^*$ and $N_i(\bar{\mathbf{g}}^*) \setminus \{k\} \neq N_k(\bar{\mathbf{g}}^*) \setminus \{i\}$. Note that for $x_i^* = x_k^*$, effort levels accessed must be equal by strict strategic complementarities, so that $y_i = y_k$. For $N_i(\bar{\mathbf{g}}^*) \setminus \{k\} \neq N_k(\bar{\mathbf{g}}^*) \setminus \{i\}$ to hold, there must exist an agent l, such that $l \in N_k(\bar{\mathbf{g}}^*)$ and $l \notin N_i(\bar{\mathbf{g}}^*)$. For the link $\bar{g}_{k,l}^* = 1$ to be in place in $\bar{\mathbf{g}}^*$ we must have that $v(\sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^*) - v(\sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^* - x_l^*) \geq k$. But from $y_i = y_k$ and the convexity of the value function $v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^* + x_l) - v(\sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^*) > v(\sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^*) - v(\sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^* - x_l^*) \geq k$ holds and we reach a contradiction. *Q.E.D.*

Lemma 13 shows that in any Nash equilibrium, if an agent i exerts a weakly lower effort level than another agent k, then agent i's neighborhood is contained in k's neighborhood.

Lemma 13: *In any NE* $\mathbf{s}^*=(\mathbf{x}^*, \mathbf{g}^*)$, $x_i^* \leq x_k^* \Leftrightarrow N_i(\bar{\mathbf{g}}^*) \setminus \{k\} \subseteq N_k(\bar{\mathbf{g}}^*) \setminus \{i\}$.

Proof. First, $N_i(\bar{\mathbf{g}}^*) \setminus \{k\} \subseteq N_k(\bar{\mathbf{g}}^*) \setminus \{i\} \Rightarrow x_i^* \leq x_k^*$. If $\bar{g}_{i,k}^* = 0$, then k accesses a weakly higher effort level, i.e., $y_i = \sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^* \leq y_k = \sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^*$ and therefore $x_i^* \leq x_k^*$. Assume next that $\bar{g}_{i,k}^* = 1$ and, without loss of generality, that $x_i^* > x_k^*$. But then k accesses a strictly higher effort level than i, $y_i = \sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^* < y_k = \sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^*$, and we have reached a contradiction. Second, $x_i^* \leq x_k^* \Rightarrow N_i(\bar{\mathbf{g}}^*) \setminus \{k\} \subseteq N_k(\bar{\mathbf{g}}^*) \setminus \{i\}$. Assume to the contrary that $x_i^* \leq x_k^*$ and there exists an agent l such that $l \in N_i(\bar{\mathbf{g}}^*)$ and $l \notin N_k(\bar{\mathbf{g}}^*)$. For the link $\bar{g}_{i,l}^* = 1$ to be in place in $\bar{\mathbf{g}}^*$, either $g_{i,l}^* = 1$ or $g_{l,i}^* = 1$. If $g_{i,l}^* = 1$, then $v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^*) - v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^* - x_l^*) \geq k$ must hold. But from $y_i \leq y_k$ and the convexity of the value function can write $v(\sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^* + x_l) - v(\sum_{j \in N_k(\bar{\mathbf{g}}^*)} x_j^*) > v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^*) - v(\sum_{j \in N_i(\bar{\mathbf{g}}^*)} x_j^* - x_l^*) \geq k$ and we have reached a contradiction. We can apply an analogous argument for $g_{l,i}^* = 1$. *Q.E.D.*

In Proposition 11 we show that all Nash equilibria are nested split graphs.

Proposition 11: *In any NE the network is a nested split graph.*

Proof. Note first that the complete and the empty network are nested split graphs. We start by showing that in any NE, if $d_k(\bar{\mathbf{g}}^*) \geq d_l(\bar{\mathbf{g}}^*)$, then $x_k^* \geq x_l^*$. Assume to the contrary that $x_l^* > x_k^*$. We distinguish two cases. First, $d_k(\bar{\mathbf{g}}^*) > d_l(\bar{\mathbf{g}}^*)$. Then there exists an agent $m \in N_k(\bar{\mathbf{g}}^*)$ and $m \notin N_l(\bar{\mathbf{g}}^*)$. Distinguish two subcases. First, agent m extends a link to k and $g_{m,k}^* = 1$. But then, by $x_l^* > x_k^*$ and Lemma 8, $\bar{g}_{m,l}^* = 1$ and we have reached a contradiction. Second, agent k extends a link to m and $g_{k,m}^* = 1$. But then, by Lemma 9, $\bar{g}_{l,m}^* = 1$ must hold and we have reached a contradiction. Assume next that $d_k(\bar{\mathbf{g}}^*) = d_l(\bar{\mathbf{g}}^*)$. If $N_k(\bar{\mathbf{g}}^*) \setminus \{l\} = N_l(\bar{\mathbf{g}}^*) \setminus \{k\}$, then $x_k^* = x_l^*$ by Lemma 12 and we have reached a contradiction. If $N_k(\bar{\mathbf{g}}^*) \setminus \{l\} \neq N_l(\bar{\mathbf{g}}^*) \setminus \{k\}$, then there exists an agent $m \in N_k(\bar{\mathbf{g}}^*) \setminus \{l\}$ and $m \notin N_l(\bar{\mathbf{g}}^*) \setminus \{k\}$. The argument that, in fact, $m \in N_l(\bar{\mathbf{g}}^*) \setminus \{k\}$ is analogous to the previous case and we again reach a contradiction. We have established that in any NE, if $d_k(\bar{\mathbf{g}}^*) \geq d_l(\bar{\mathbf{g}}^*)$, then $x_k^* \geq x_l^*$. Next we show that in any NE, if $\bar{g}_{i,l}^* = 1$ and $d_k(\bar{\mathbf{g}}^*) \geq d_l(\bar{\mathbf{g}}^*)$ (and therefore $x_k^* \geq x_l^*$), then $\bar{g}_{i,k}^* = 1$. We distinguish two cases. First, $g_{i,l}^* = 1$. Then, by Lemma 8 $\bar{g}_{i,k}^* = 1$. Second, $g_{l,i}^* = 1$. Then, by Lemma 9 $\bar{g}_{i,k}^* = 1$. That is, $\bar{\mathbf{g}}^*$ is a nested split graph. *Q.E.D.*

We know from the literature in graph theory that nested split graphs are core-periphery networks (split graphs).[16] Below, we provide a separate proof, which highlights agent's incentives of linking and exerting effort. Furthermore, a relationship between a property of the NE effort levels and NE network structure is established.

Corollary 3: *In any NE*

- *if $x_i^* = x_j^*$ $\forall i, j \in N$, then the network is either empty or complete,*
- *if $x_i^* \neq x_j^*$ for some $i, j \in N$, then the network is a core-periphery network (that is neither complete nor empty).*

Proof. Lemma 9 directly implies that the only networks such that $x_i^* = x_j^*$ $\forall i, j \in N$ are the empty and the complete network. Next, we show that in any NE such that there exists a pair of agents i and j, such that $x_i^* \neq x_j^*$, the network displays a core-periphery structure (other than the complete or empty network). Rank agents by their effort levels in increasing order, such that $x_1^* \leq x_2^* \leq ... \leq x_{n-1}^* \leq x_n^*$. We know from Lemma 1 that the network is not complete, since there exists a pair of agents i and j such that $x_i^* \neq x_j^*$. The network is not empty, as agents have identical payoff functions and singleton agents display same effort

[16] See, for example, Mahadev and Peled (1995)

levels, $\bar{x}(0)$. We start by showing that the two lowest ranked agents, agent 1 and agent 2, are not connected. Two cases are to be discerned. First, $g^*_{1,2} = 1$. From Lemma 8 we know that agent 1 must be connected to all agents remaining agents, since $x^*_j \geq x^*_2 \; \forall j \geq 2$. Lemma 13 implies that the network is complete since $N_i(\mathbf{g}^*) \setminus \{j\} \subseteq N_j(\mathbf{g}^*) \setminus \{i\}$ holds for agents j with $x^*_j \geq x^*_1 \; \forall j \geq 1$. But then $x^*_i = x^*_j \; \forall i, j \in N$ by Lemma 1 and we have reached a contradiction. Second, $g^*_{2,1} = 1$. From Lemma 9 we know that agent 1 is connected to all agents, since $x^*_j \geq x^*_2 \; \forall j \geq 2$ and the above argument applies. Since the network is neither empty, nor complete, at least one link exists. Pick the agent i with the lowest subscript that is involved in a link and, if i is involved in more than one link, consider the link to the agent with the lowest subscript j. We discern two cases, $g^*_{i,j} = 1$ and $g^*_{j,i} = 1$. First, $g^*_{i,j} = 1$. We can discern two subcases. First, agent i and j are adjacent. As i is the agent with the lowest subscript involved in a link, all agents with lower subscripts have no links. All agents with a subscript higher or equal to i are connected to each other. To see this, note that by Lemma 8, agent i is connected to all agents with a subscript higher or equal than j. But then by Lemma 13, $\bar{g}^*_{l,m} = 1 \; \forall l, m \geq i$. The periphery, $P(\mathbf{g}^*)$, consists of agents with subscripts $k < i$, while the core, $C(\mathbf{g}^*)$, consists of agents with subscripts $k \geq i$. The argument for the case where $g^*_{j,i} = 1$ is analogous. Assume next that i and j are not adjacent. Note that since $g^*_{i,j} = 1$ and from $x^*_i \leq x^*_{j-1}$, we know by Lemma 9 that the link between $j-1$ and j, $\bar{g}^*_{j-1,j} = 1$, also exists. Next, check for the link $\bar{g}^*_{j-2,j-1}$. If $\bar{g}^*_{j-2,j-1} = 0$, then by Lemma 9 no agent with a subscript lower than $j-2$ is connected to $j-1$. Furthermore, no pair of agents with subscripts of lower or equal than $j-2$ is connected. Assume to the contrary that there exists a pair of nodes l, m with $l \leq m < j-2$ and $\bar{g}^*_{l,m} = 1$. By Lemma 8 we must then have that $\bar{g}^*_{l,j-1} = 1$. This, however, contradicts Lemma 3, since $\bar{g}^*_{j-2,j-1} = 0$. The periphery, $P(\mathbf{g}^*)$, consists of agents with subscripts $k < j$, while the core, $C(\mathbf{g}^*)$, consists of agents with subscripts $k \geq j$. If $\bar{g}^*_{j-2,j-1} = 1$, check for the link $\bar{g}^*_{j-3,j-2}$. If $\bar{g}^*_{j-3,j-2} = 0$, then by above argument the periphery, $P(\mathbf{g}^*)$, consists of agents with subscripts $k < j-1$, while the core, $C(\mathbf{g}^*)$, consists of agents with subscripts $k \geq j-1$. If $\bar{g}^*_{j-3,j-2} = 1$, proceed in descending order until a pair of adjacent agents is found that is not connected and define the core and periphery accordingly. Note that such a pair of agents exists, since i and j were assumed to not be adjacent and therefore $\bar{g}^*_{i,i+1} = 0$. This completes the proof. *Q.E.D.*

In Proposition 12 we show that in any NE such that two agents have different effort levels, there exists a core-periphery partition such that agents in the core exert strictly higher effort levels and obtain strictly higher gross payoffs.

Proposition 12: *In any NE with a pair of agents i and j, such that $x_i^* \neq x_j^*$, there exists a core-periphery partition of agents, such that $x_i^* > x_j^*$ and $\pi_i^* > \pi_j^*$ $\forall i \in C(\bar{\mathbf{g}}^*)$ and $\forall j \in P(\bar{\mathbf{g}}^*)$.*

Proof. Rank again agents by their effort levels in increasing order, such that $x_1^* \leq x_2^* \leq ... \leq x_{n-1}^* \leq x_n^*$. From Proposition 11 we know that there exists an agent x_k^* such that $\bar{g}_{i-1,i}^* = 0$ $\forall i \leq k$ and $\bar{g}_{i,i+1}^* = 1$ $\forall i \geq k$. We discern two cases. First, $x_k^* = x_{k+1}^*$. Define $C(\bar{\mathbf{g}}^*) = \{i : x_i \geq x_k\}$ and $P(\bar{\mathbf{g}}^*) = \{i : x_i < x_k\}$ and assume, contrary to the above, that $x_{k-1}^* \geq x_k^* = x_{k+1}^*$. From Lemma 13 we know that $N_{k+1}(\bar{\mathbf{g}}^*) \setminus \{k-1\} \subseteq N_{k-1}(\bar{\mathbf{g}}^*) \setminus \{k\}$ must hold. That is, if $x_{k-1}^* \geq x_{k+1}^*$, then all of agent $k+1$'s neighbors are also $k-1$'s neighbors. But then, since $k+1$ is connected with k, $k-1$ is also connected with k and $\bar{g}_{k-1,k}^* = 1$. We have reached a contradiction. Next, $x_k^* < x_{k+1}^*$. In this case simply define $C(\bar{\mathbf{g}}^*) = \{i : x_i \geq x_{k+1}\}$ and $P(\bar{\mathbf{g}}^*) = \{i : x_i < x_{k+1}\}$. For the second part of the statement note that from $x_i^* > x_j^*$ we know that $y_i^* = \sum_{k \in N_i(\bar{\mathbf{g}}^*)} x_k^* > \sum_{k \in N_j(\bar{\mathbf{g}}^*)} x_k^* = y_j^*$ holds. We can then write $\pi(x_j^*, y_i^*) > \pi_j^*(x_j^*, y_j^*)$, where the inequality follows from positive externalities. *Q.E.D.*

Next, we define a *periphery-sponsored core-periphery network* as a core-periphery network where all agents in the periphery extend links to all agents in the core. A *core-sponsored core-periphery network* is a core-periphery network where all agents in the core extend links to all agents in the periphery.

Definition 3: A network $\mathbf{g}$ is a *periphery − sponsored core − periphery network* if the set of agents N can be partitioned into two sets, $C(\mathbf{g})$ (the core) and $P(\mathbf{g})$ (the periphery), such that $\bar{g}_{i,j} = 1$ $\forall i, j \in C(\mathbf{g})$, $\bar{g}_{i,j} = 0$ $\forall i, j \in P(\mathbf{g})$ and $g_{i,j} = 1 \forall i \in C(\mathbf{g})$ and $\forall i \in P(\mathbf{g})$. A network $\mathbf{g}$ is a *core − sponsored core − periphery network* if the set of agents N can be partitioned into two sets, $C(\mathbf{g})$ (the core) and $P(\mathbf{g})$ (the periphery), such that $\bar{g}_{i,j} = 1$ $\forall i, j \in C(\mathbf{g})$, $\bar{g}_{i,j} = 0$ $\forall i, j \in P(\mathbf{g})$ and $g_{i,j} = 1 \forall i \in C(\mathbf{g})$ and $\forall i \in P(\mathbf{g})$.

Lemma 14: *There does not exist a NE such that the network is a center-sponsored star.*

Proof. Assume to the contrary that the center extends link to all agents in the periphery. Note that the center extends $n-1$ links to agents with effort level x_p^*. That is, for this to be a NE, $\frac{v((n-1)x_p^*)-v(0)}{n-1} \geq k$ needs to hold. Furthermore, we need that an agent in the periphery does not find it profitable to deviate by extending links to the remaining $n-2$ agents in the periphery. That is, we need $k \geq \frac{v((n-2)x_p^*)-v(x_c^*)}{n-2}$ to hold. However, from Proposition 12 we know that $x_c^* > x_p^*$ (and therefore $(n-2)x_p^* + x_c^* > (n-1)x_p^*$) and then, from the convexity of the value function, $\frac{v((n-2)x_p^*)-v(x_c^*)}{n-2} > \frac{v((n-1)x_p^*)-v(0)}{n-1}$ holds. We have reached a contradiction. *Q.E.D.*

Corollary 3: *For any NE star network in which the center obtains lower payoffs than an agent in the periphery, there exists another NE star network with a core-periphery partition such that the center obtains strictly higher payoffs than agents in the periphery .*

Proof. From Lemma 14 we know that there does not exist a NE star network such that the center extends links to all agents in the periphery. This leaves two configurations. First, periphery-sponsored star networks and second, networks where some agents in the periphery extend links to the center and the center, in turn, extends links to some agents in the periphery. In a periphery-sponsored star network, payoffs of the center are strictly larger than payoffs of an agent in the periphery. To see this, note that from Proposition 12 and strategic complementarities we know that $(n-1)x_p^* > x_c^*$. That is, the center's neighbors not only exert higher effort level and gross payoffs are therefore higher, but the center does also not incur any linking cost. For the second case, note that some of the agents in the periphery extend a link to the central agent. But then there exists another NE which is a periphery-sponsored star network, for which we have just shown that payoffs are strictly larger for an agent in the center. $Q.E.D.$

Proposition 13 presents necessary and sufficient conditions for the existence of a NE star network for the linear-quadratic payoff function proposed in Calvó-Armengol et al (2005 and 2009), which is given by $\pi(x_i, y_i) = x_i - \frac{\beta}{2}x_i^2 + \lambda x_i \sum_{j \in N_i(\bar{\mathbf{g}})} x_j$. Note that for the existence of a NE star network, we only need to check for periphery-sponsored stars. This follows from Lemma 14, which shows that a center-sponsored star does not exist, and therefore in any NE star network at least one agent in the periphery extends a link to the center. That is, if a NE star network exists such that the central agent extends a link, then there also exists a NE with a periphery-sponsored star network, while the converse need not hold. As in Proposition 6, the condition for a NE on a fixed network is given by $n < 1 + \frac{\lambda^2}{\beta^2}$. However, in the one-sided link formation specification this condition does not bind, as the upper bound on n (presented below) is always smaller than $1 + \frac{\lambda^2}{\beta^2}$ for λ and β positive. The first condition of Proposition 13, $\beta \geq 5\lambda$, imposes an upper bound on λ relative to β, which is independent of k and n. That is, if the parameter governing strategic complementarities, λ, is sufficiently large relative to the convexity of the cost function, β, then a NE star network does not exist (for any combination of k and n). Note next that for a NE star network to exist, linking cost k must be in an intermediate range. The lower bound on linking cost k is given by the marginal payoffs of an agent in the periphery linking to all $n-2$ agents in the periphery, while keeping his link to the center. That is, for k larger than the lower bound, an agents in the periphery will not find it profitable to link to the remaining agents in the periphery. The upper bound, in turn, is given by the marginal payoffs of periphery linking to the center. Therefore, for linking cost k smaller than the upper bound, an agent in the periphery does not find it profitable to delete his links with the central agent. Note that bounds on linking

cost k are increasing in n. The bounds on n (together with $\beta \geq 5\lambda$) then ensure that the upper bound on linking cost is, in fact, (weakly) larger than the lower bound. The last two conditions present the knife-edge cases where the lower bound on linking cost k coincides with the upper bound.

Proposition 13: *If best response functions are linear, then a NE with a star network exists if and only if*

- $\beta > 5\lambda$ *and either*
  - $\frac{\lambda(\beta+\lambda)^2(2\beta+(n-2)\lambda)}{2\beta(\beta^2-(n-1)\lambda^2)^2} \leq k \leq \frac{\beta(\beta+\lambda)^2}{2(\beta^2-\lambda^2(n-1))^2} - \frac{1}{2\beta}$ *and* $\frac{(\beta-\lambda)^2}{2\lambda^2} - \sqrt{\frac{(\beta-5\lambda)(\beta-\lambda)(\beta+\lambda)^2}{4\lambda^4}} \leq n \leq \frac{(\beta-\lambda)^2}{2\lambda^2} + \sqrt{\frac{(\beta-5\lambda)(\beta-\lambda)(\beta+\lambda)^2}{4\lambda^4}}$, *or*
  - $\frac{\lambda(\beta+\lambda)^2(2\beta+(n-2)\lambda)}{2\beta(\beta^2-(n-1)\lambda^2)^2} = k$ *and* $\frac{(\beta-\lambda)^2}{2\lambda^2} + \sqrt{\frac{(\beta-5\lambda)(\beta-\lambda)(\beta+\lambda)^2}{4\lambda^2}} = n$, *or*
- $\beta \geq 5\lambda$, $\frac{\beta(\beta+\lambda)^2}{2(\beta^2-\lambda^2(n-1))^2} - \frac{1}{2\beta} = k$ *and* $\frac{(\beta-\lambda)^2}{2\lambda^2} - \sqrt{\frac{(\beta-5\lambda)(\beta-\lambda)(\beta+\lambda)^2}{4\lambda^2}} = n$.

Proof. From Lemma 14 we know that in a NE star network there must be at least one peripheral player extending a link to the center and we can therefore focus on periphery sponsored core-periphery networks. Note next that that in a star network, all agents in the periphery access the same effort level, x_c^*, and therefore all agents in the periphery display the same effort level, x_p^*. The agents in the core maximize $x_c \in argmax_{x_c \in X} x_c - \frac{\beta}{2}x_c^2 + \lambda x_c(n-1)x_p$, where x_p is the effort level of agents in the periphery. For an agent in the periphery we have $x_p \in argmax_{x_p \in X} x_p - \frac{\beta}{2}x_p^2 + \lambda x_p x_c$. The reaction functions are given by $x_c(x_p, \hat{x}_c) = \frac{1+\lambda x_p(n-1)}{\beta}$ and $x_p(x_c) = \frac{1+\lambda x_c}{\beta}$, respectively. Equilibrium effort levels are given by $x_c^* = \frac{\beta+\lambda(n-1)}{\beta^2-\lambda^2(n-1)}$ and $x_p^* = \frac{\beta+\lambda}{\beta^2-\lambda^2(n-1)}$. Plugging equilibrium effort levels back into the payoff function, yields equilibrium gross payoffs of $\pi_c^* = \frac{\beta(\beta+\lambda(n-1))^2}{2(\beta^2-\lambda^2(n-1)}$ and $\pi_p^* = \frac{\beta(\beta+\lambda)^2}{2(\beta^2-\lambda^2(n-1))^2}$. Next, we calculate deviation payoffs of an agent in the periphery linking to all $n-2$ remaining agents in the periphery. A deviating agent maximizes $x_p^d \in argmax_{x_p^d \in X} x_p^d - \frac{\beta}{2}(x_p^d)^2 + \lambda x_p^d(x_c^* + (n-2)x_p^*)$, which yields the following reaction function $x_p^d(x_c^*, x_p^*) = \frac{1+\lambda x_c^* + \lambda x_p^*(n-2)}{\beta}$. The deviation effort level is given by $x_p^d(x_c^*, x_p^*) = \frac{(\beta+\lambda)(\beta+\lambda(n-2))}{\beta(\beta^2-\lambda^2(n-1))}$ and corresponding deviation gross payoffs by $\pi_p^d = \frac{(\beta+\lambda)^2(\beta+\lambda(n-2))^2}{2\beta(\beta^2-\lambda^2(n-1))^2}$. For the existence of a periphery-sponsored core-periphery network we need two conditions to hold. First, linking cost k must be such that an agent in the periphery does not find it profitable to delete his link to the central agent and, second, agents in the periphery do not find it profitable to link to the remaining $(n-1)$ agents in the periphery, given the link to the center. The first condition then reads $\frac{\beta(\beta+\lambda)^2}{2(\beta^2-\lambda^2(n-1))^2} - \frac{1}{2\beta} \geq k$, where $\frac{1}{2\beta}$ are the payoffs of an agent without any links. The second condition is given by $k \geq (\frac{(\beta+\lambda)^2(\beta+\lambda(n-2))^2}{2\beta(\beta^2-\lambda^2(n-1))^2} - \frac{\beta(\beta+\lambda)^2}{2(\beta^2-\lambda^2(n-1))^2})/(n-2) = \frac{\lambda(\beta+\lambda)^2(2\beta+(n-2)\lambda)}{2\beta(\beta^2-(n-1)\lambda^2)^2}$. Combining the two inequalities one can show after some algebraic manipulation that these two conditions hold

if and only if $\beta > 5\lambda$ and either $\frac{\lambda(\beta+\lambda)^2(2\beta+(n-2)\lambda)}{2\beta(\beta^2-(n-1)\lambda^2)^2} \leq k \leq \frac{\beta(\beta+\lambda)^2}{2(\beta^2-\lambda^2(n-1))^2} - \frac{1}{2\beta}$ and $\frac{(\beta-\lambda)^2}{2\lambda^2} - \sqrt{\frac{(\beta-5\lambda)(\beta-\lambda)(\beta+\lambda)^2}{4\lambda^4}} \leq n \leq \frac{(\beta-\lambda)^2}{2\lambda^2} + \sqrt{\frac{(\beta-5\lambda)(\beta-\lambda)(\beta+\lambda)^2}{4\lambda^4}}$ or, one of two knife-edge cases, where the upper and lower bound on k coincide, i.e. $\beta > 5\lambda$ and $\frac{\lambda(\beta+\lambda)^2(2\beta+(n-2)\lambda)}{2\beta(\beta^2-(n-1)\lambda^2)^2} = k$ and $\frac{(\beta-\lambda)^2}{2\lambda^2} + \sqrt{\frac{(\beta-5\lambda)(\beta-\lambda)(\beta+\lambda)^2}{4\lambda^2}} = n$ or $\beta \geq 5\lambda$, $\frac{\beta(\beta+\lambda)^2}{2(\beta^2-\lambda^2(n-1))^2} - \frac{1}{2\beta} = k$ and $\frac{(\beta-\lambda)^2}{2\lambda^2} - \sqrt{\frac{(\beta-5\lambda)(\beta-\lambda)(\beta+\lambda)^2}{4\lambda^2}} = n$.[17] *Q.E.D.*

[17] The calculations were executed with Mathematica and the codes are available upon request.

6 References

1. Adamic, L.A. and E. Adam, (2003), "Friends and neighbors on the web", Social Networks 25, pp. 211.230.

2. Ahuja, G., (2000), "Collaboration networks, structural holes and innovation: A longitudinal study", Administrative Science Quarterly, Vol. 45, pp. 425455.

3. Baetz, O., (2012), “Social Activity and Network Formation”, Working Paper.

4. Baker, G., R. Gibbons, and K.J. Murphy (2004): "Strategic alliances: bridges between islands of conscious power", mimeo, MIT.

5. Bala, V. and S. Goyal, (2000), "A non-cooperative model of network formation", Econometrica, Vol. 68, No. 5, pp. 1181-1229.

6. Baldwin, T.T., M.D. Bedell, and J.L. Johnson, (1997), "The social fabric of a team-based M.B.A. program: network effects on student satisfaction and performance", Academy of Management Journal, Vol. 40, pp. 1369.1397.

7. Ballester, C., A. Calvó-Armengol, and Y. Zenou, (2006), “Who is Who in Networks. Wanted: The Key Player”, Econometrica 74, pp.1403-1417.

8. Belleflamme, P., and F. Bloch, (2004), “Market Sharing Agreements and Stable Collusive Networks”, International Economic Review 45 Vol. 2, pp. 387-411.

9. Bramoullé, Y., (2007), "Anti-Coordination and Social Interactions", Games and Economic Behavior, Vol. 58, pp. 30-49.

10. Calvó-Armengol, A., E. Pattacchini and Y. Zenou, (2009), "Peer Effects and Social Networks in Education", Review of Economic Studies, Vol. 76, pp. 1239-1267.

11. Calvó-Armengol, A., Patacchini, E. and Y. Zenou (2005), “Peer effects and social networks in education and crime”, CEPR Discussion Paper No. 5244.

12. Canter, D. (2004), "A Partial Order Scalogram Analysis of Criminal Network Structures", Behaviormetrika, Vol. 31, No.2, pp. 131-152.

13. Case, A. C. and L. F. Katz, (1991), “The Company you Keep: The Effects of Family and Neighborhood on Disadvantaged Youths”, NBER Working Paper No. 3705.

14. Cohen, W. and D. Levinthal, (1989), "Innovation and learning: The two faces of R&D", The Economic Journal, Volume 99, pp. 569-596.

15. Cohen, W. and D. Levinthal (1990), "Absorptive capacity: A new perspective on learning and innovation", Administration Science Quarterly, Vol. 35, pp.128-152

16. Cohen-Cole, E., E. Patacchini, Y. Zenou (2012), “Systemic Risk and Network Formation in the Interbank Market”, working paper.

17. Conley, T. G. and G. Topa, (2002), "Socio-Economic Distance and Spatial Patterns in Unemployment", Journal of Applied Econometrics, Volume 17, Issue 4, pp. 303-327.

18. De Benedictis, L., L. Tajoli, (2011), “The world trade network”, The World Economy 34 (8), pp. 1317-1454.

19. Galeotti, A. and S. Goyal, (2010), "The Law of the Few", American Economic Review, Vol. 100, No. 4, pp. 1468-92.

20. Glaeser, E., B. Sacerdote and J. Scheinkman (1996), "Crime and Social Interactions", The Quarterly Journal of Economics, Vol. 111, No. 2, pp. 507-48.

21. Goyal, S., and S. Joshi, (2003), “Networks of Collaboration in Oligopoly”, Games and Economic Behavior 43, pp. 57-85.

22. Gulati, R., and M. Gargiulo, (1999), "Where do interorganizational networks come from?", American Journal of Sociology, Vol. 104, pp. 177-231.

23. Hoxby, C., (2000), "Peer Effects in the Classroom: Learning from Gender and Race Variation", NBER Working Paper 7867.

24. Jackson, M.O., and J. Wolinsky (1996), “A Strategic Model of Social and Economic Networks”, Journal of Economic Theory, Vol. 71 (1), pp. 44-74.

25. Kennan, J., (2001), “Uniqueness of Positive Fixed Points for Increasing, Concave Functions on $\mathbb{R}^n$: An Elementary Result.”, Review of Economic Dynamics, 4, pp. 893-899.

26. König, M., C. J. Tessone, and Y. Zenou, (2012), “Nestedness in Networks: A Theoretical Model and Some Applications”, CEPR Discussion Paper No. 8807.

27. Levin, R. C. and Reiss, P. C. (1988), "Cost-reducing and demand-creating R& D with spillovers", Rand Journal of Economics, Vol. 19, pp. 538-56.

28. Lin, M., (2001), "Social Capital: A Theory of Social Structure and Action", Cambridge University Press, Cambridge.

29. Lovasz, L., (2009), “Geometric Representations of Graphs”, http://www.cs.elte.hu/~lovasz/geomrep.pdf

30. Mahadev, N. and U. Peled, (1995), Threshold Graphs and Related Topics, North Holland.

31. Powell, W.W., K.W. Koput, L. Smith-Doerr and J. Owen-Smith, (1999), "Network position and firm performance: organizational returns to collaboration in the biotechnology industry", in: S. Andrews, D. Knoke (Eds.), Networks in and Around Organizations, JAI Press, Greenwich, pp. 129.160.

32. Sacerdote, B., (2001), "Peer Effects with Random Assignment: Results for Dartmouth Roommates", Quarterly Journal of Economics, Vol. 116, No. 2, pp. 681-704.

33. Soramaki, K., M. Bech, J. Arnold, R. Glass, W. Beyeler, (2007), “The topology of interbank payment flows. Physica A: Statistical Mechanics and its Applications 379 (1), pp. 317-333.

34. Topa, G., (2001), "Social Interactions, Local Spillovers and Unemployment", Review of Economic Studies, Vol. 68, No. 2, pp. 261-95.