Social Activity and Network Formation

Oliver Baetz*

Draft, 15th January 2013

Abstract

This paper studies games of social activity in a network environment, in which own activity and cumulative neighbour activity are strategic complements and agents have concave value functions. Typical stylised examples for such a situation include activity and linking in online social networks. We show that there is a unique positive activity equilibrium on exogenous networks under mild conditions. When network formation is endogenised, then: (i) Equilibria display a single positive level of activity iff the network is regular. (ii) There are equilibria with many distinct levels of activity; players with higher activity have more neighbours but sponsor fewer links and the corresponding networks are (with few exceptions) not minimally connected. (iii) In strict equilibria, the network is either a complete multipartite graph or a particular variation of these graphs.

Finally, we show how individual preferences shape the social network in large societies—mirroring some empirical findings e.g. for Facebook and YouTube.

JEL Classification: C72, D00, D85

---

*Faculty of Economics, University of Cambridge. Email: ob251@cam.ac.uk
I thank my supervisor Sanjeev Goyal for his guidance and advice. I am grateful to Ansgar Walther, James Rutt, Oliver Latham, Daniel Quigley, Marcin Dziubinski, Friederike Mengel, Timo Hiller, and participants of the Theory Workshop at Cambridge for their helpful comments. I acknowledge the funding of my postgraduate research by the Economics and Social Research Council (ESRC) and the Cambridge European Trust. Daniel Quigley and 7 others like this paper.

1 Introduction

Online social networks have become increasingly important for everyday life over the last decade and their economic impact steadily rises.[1] On these sites, users typically create a personal profile or channel and then interact through a virtual network represented by a list of friends or followers. Casual observation indicates two regularities:

First, activity of friends or followers (that is time spent on a particular social network site) often displays complementarity: if, for instance, a friend uploads more pictures or sends more messages on Facebook, then there are stronger incentives to spend time tagging people in these photos as well as commenting on them or reading and replying to messages.

Second, users typically participate in a variety of social network sites offering different services instead of one integrated platform (for instance Facebook, Google+, LinkedIn, Twitter, Myspace or even YouTube). This indicates that there are diminishing returns to using particular services.

In this paper, we take these two observations as given and are interested in their implications for activity patterns as well as for the network structure.[2] For this purpose, we propose a "social activity and network formation" model with two corresponding key properties:

First, there are strategic complements between an agent's own activity and the *cumulative* activity of his friends. In particular, an agent desires to be active if one of his friends is active and then increases his activity at a diminishing rate as his friends become more involved. Ugander *et al.* (2011) provide some empirical evidence that user activity on Facebook essentially meets these assumptions.[3]

[1] For instance, the initial public offering of Facebook in May 2012 was (to that date) the third largest public offering in US history giving the company a $104 billion valuation (Dembosky and Demos, 2012). On a broader perspective, the McKinsey Global Institute estimated that firms can raise the productivity of high-skill knowledge workers by 20 to 25 percent by fully implementing (online) social technologies (Chui *et al.*, 2012).

[2] Due to the amount of material, we leave welfare considerations in this paper aside and open for future research.

[3] The authors measure user logins over a 28-day window. Their results show that a user's degree as well as the average number of logins of his neighbours are increasing and (largely) convex functions of his own number of logins (see Figure 7, ibid.). Consequently,

Second, there are limitations to the benefits from social activity: optimising agents—who adjust their activity to their friends' involvement—have diminishing marginal utility from (cumulative) neighbour activity.

We consider two versions of the game, a *pure activity game* on exogenous networks and an activity game with *endogenous network formation*. Both versions of the game are one-shot and simultaneous-move.

Our contribution to the literature is twofold: We extend (i) the theory of games with strategic complementarity on exogenous networks as well as under endogenous network formation and (ii) introduce a simple model that replicates some empirical findings for online social networks.

For the *pure activity game*, we provide a necessary and a sufficient condition on the best response function for the existence of a unique (strictly) positive activity equilibrium.[4]

In the game with *endogenous network formation*, an equilibrium displays a *single positive level of activity* if and only if the network is regular (that is all players have a common number of neighbours/degree). The corresponding equilibrium activity is determined by, and increases in, the degree of the regular network.[5]

In principle, positive activity equilibria can display a large number of distinct activity levels.[6] We find several regularities for these *(generic) multi-level equilibria*: First, players' activity and degree are positively correlated—in accordance with empirical findings for Facebook (see Wilson *et al.* (2009), Figure 8). Second, although high activity players have many neighbours, they tend to sponsor few links themselves. Third, equilibrium networks are (with few exceptions) not minimally connected.

---

a user's number of logins is an increasing and concave function of the *total* logins of his neighbours.

[4]Some properties of the positive activity equilibrium as well as comparative statics are presented in the supplementary material in Appendix B.2, see also Footnote 22 for details.

[5]We discuss the relation of the interval of equilibrium supporting linking costs and the degree of a regular equilibrium network in the supplementary material in Appendix B.6.

[6]As examples, we study two prominent types of networks more in detail: any biregular bipartite graph—such as the star and the complete bipartite graph—is an equilibrium network with two levels of activity for some interval of linking costs. Conversely, core-periphery networks exist in equilibrium only under very particular specifications of the model.

We also study a stronger solution concept: In *strict equilibria*, the network is either a complete multipartite graph or a variation of these graphs in which one partite set is replaced by a group of fully connected players (i.e. a clique).

Finally, we consider large societies. We first show that *some* positive activity equilibrium exists for any linking costs smaller than the supremum value of a single link. We then discuss how individual (user) preferences shape the social network: If users are relatively satiated even from little neighbours' activity then the star is a robust equilibrium in large societies and the complete network ceases to exist. If users get less easily satiated then the converse holds true. Indeed, Wilson *et al.* (2009) show empirically that pure social network sites such as Facebook exhibit significantly higher connectivity than content distribution sites with (less elaborate) social network components such as YouTube.

Our model extends a new and rapidly evolving literature. Originally, network formation, such as Bala and Goyal (2000), and (one shot) games on exogenous networks, as in Bramoullé and Kranton (2007), were two separate strands of the literature. Galeotti and Goyal (2010) combine these two strands of literature—and in fact these two papers—to analyse a local public goods game with simultaneous and endogenous network formation. Their model displays strategic substitutes and the authors show that only a small fraction of players invests into the public good in large societies.[7]

Our model is most closely related to Hiller (2010) who also studies a game with strategic complementarity and endogenous network formation but assumes in contrast to us a *convex* value function. Hiller's assumption leads to some kind of "bang-bang" solution: for sufficiently low linking costs, the complete network is the unique equilibrium and for sufficiently high costs the empty network. The only (additional) structure that might arise for intermediate linking costs are core-periphery networks. In our model, much richer social structures and activity patterns—which are concealed in Hiller (2010) due to convexity—exist in equilibrium.

[7] Zhang *et al.* (2011) extend this research to the case of imperfect substitutes. They model (pure) content production in online social networks where each player has a taste for content from different sources.

There are several other related papers: Ballester *et al.* (2006) assume linear best responses and analyse games with local strategic complements and global substitutes on *exogenous networks*. They link the equilibrium action of any player to his Bonacich centrality in the network.[8] Lagerås and Seim (2012) allow for non-linear best responses but assume, in contrast to us, that the complementarity effect from any two neighbours is mutually reinforcing.[9] They show that a unique interior equilibrium exists on any *exogenous network*. Furthermore, the network structure converges to so-called nested split graphs in a particular *dynamic network formation* game with myopic agents.[10] In contrast, equilibrium networks in our model are typically non-nested graphs as high degree agents sponsor few links (and thus rarely share links between themselves).[11]

The paper proceeds as follows: in the next section, we introduce the model and our solution concept. Section 3 considers the pure activity game on exogenous interaction networks. In Section 4, we characterise equilibria in the activity game with endogenous network formation as comprehensively as possible for a fixed size of society. Subsequently, we consider large societies in Section 5 and discuss how individual (user) preferences shape the social network. Finally, the last section concludes. Proofs from the main body are presented in Appendix A and supplementary material is collected in Appendix B.

[8]Recently, Bramoullé *et al.* (2012) have extended the analysis to arbitrary (bounded) *linear* best responses on exogenous networks. In contrast, we assume concave best responses and (mainly) study endogenous network formation.

[9]They essentially assume best response correspondences with non-negative cross partial derivatives; in our model, the cross partial derivatives are strictly negative. It can be easily checked that Assumption 4 from Lagerås and Seim (2012) is for instance not met for the family of social activity models in Example 1.

[10]In these graphs, the neighbourhood of any lower degree player is contained in the neighbourhood of all higher degree players.

[11]For instance, biregular bipartite networks in which lower degree players have at least two neighbours are non-nested equilibrium networks, see Section 4.2.2.

2 Model

Let $N = \{1, 2, \ldots, n\}$ be the set of agents with $n \geq 2$. All agents $i \in N$ choose a level of activity $x_i \in X = [0, \infty)$ simultaneously; the vector $\mathbf{x}^n = (x_1, x_2, \ldots x_n)$ collects the activity of all n players.

We distinguish two versions of the game: either the social interaction structure is exogenously given or it is formed endogenously and simultaneously with the activity choices. In the latter case, we assume one-sided link formation.[12]

Social Interaction Structure

(a) Link Sponsorship If the social interaction structure is formed endogenously, link sponsorship is represented by a directed network or graph $g \in G^n$ where G^n is the set of Boolean $n \times n$ matrices with zeros on the main diagonal.[13] We say that a player i sponsors, forms or supports a link to some other player j iff $g_{ij} = 1$, where g_{ij} denotes the (i, j)th entry of g. Player i's linking choices are collected in the vector g_i, the ith row vector of g. The set of possible linking choices of agent i—i.e. the set of Boolean row vectors of length n in which the ith entry is zero—is called $\mathfrak{g}_i$.

We need to introduce some further notation: first, the number of links player i forms in g is counted by $\eta_{i,g} = |\{j \in N : g_{ij} = 1\}|$. Second, $g_{-i} \odot g'_i = (g_1^t, \ldots, g_{i-1}^t, g_i'^t, g_{i+1}^t, \ldots g_n^t)^t$ is an operation which alters i's linking decision, that is, the ith row of g is replaced by some vector $g'_i \in \mathfrak{g}_i$ whilst the remaining network is kept fixed. Last, $g \oplus ij$ is an operation with which a single link is added, i.e. the (i, j)th entry of g is set to one keeping the remaining network fixed.

[12] Although in some real world social networks (e.g. on Facebook) "friendship invitations" have to be confirmed, there are relatively small costs to accepting and maintaining links. Instead, finding an offline friend within the social network might take some time—e.g. if he is registered under a pseudonym or if there are several users who share the same name. Thus one-sided link formation is a reasonable simplification.

[13] For simplicity, we identify the network or graph with its adjacency matrix. As the number of agents n is often clear from the context, we suppress the superscript frequently.

(b) Social Interactions Social interactions are two-sided, that is independent of the direction of link sponsorship. We denote the undirected interaction network by $\bar{g} \in \bar{G}^n$ where $\bar{G}^n$ is the set of *symmetric* Boolean $n \times n$ matrices with zeros on the main diagonal. Applying the same notation as above, two distinct agents i and j interact iff $\bar{g}_{ij} = 1$.

In the pure activity game, the interaction network $\bar{g}$ is given exogenously. If the social interaction structure is formed endogenously instead, then $\bar{g}$ is induced by the links formed in g such that entries $\bar{g}_{ij} = \max\{g_{ij,}g_{ji}\}$.

The set of agents with whom agent i shares a link $N_{i,\bar{g}} = \{j \in N : \bar{g}_{ij} = 1\}$ is called the *neighbourhood* of i in $\bar{g}$; the number of i's neighbours $n_{i,\bar{g}} = |N_{i,\bar{g}}|$ is also known as his degree. As the network is often clear from the context, we frequently suppress the subscript $\bar{g}$.

We say that two players i and j are *connected* in $\bar{g}$ if either $\bar{g}_{ij} = 1$ or there is a path between them, i.e. there are players $i_1, \ldots, i_l$ with $\bar{g}_{ii_1} = \bar{g}_{i_1i_2} = \cdots = g_{i_lj} = 1$. A player i is *isolated* or a *singleton* in $\bar{g}$ iff he is not connected to any other player j. A *maximal* non-empty subset of mutually connected players is called a *component* of $\bar{g}$. The *network* $\bar{g}$ is called *connected* if every pair of players is connected, that is $\bar{g}$ consists of a single component. Finally, an *independent set* is a non-empty subset of players who do not share any direct links between themselves.

There are two particularly important undirected graphs: In the *empty network* $\bar{g}_{emp}$, there are no links, so $\bar{g}_{emp,ij} = 0$ for all $i, j \in N$. In the *complete network* $\bar{g}_{com}$, all pairs of nodes share a link, so $\bar{g}_{com,ij} = 1$ for all distinct $i, j \in N$.

Payoffs and Modeling Assumptions

Let $\pi_i(\mathbf{x}, \bar{g})$ denote player i's gross payoff under activity $\mathbf{x}$ on interaction network $\bar{g}$ (excluding any costs for link sponsorship). We follow standard formulations in the literature in two aspects: First, gross payoffs of every player are entirely determined by *own activity* and the *total activity of friends* or neighbours in the network, that is $\pi_i(\mathbf{x}, \bar{g}) = \tilde{\pi}(x_i, y_i)$ with $y_i = \sum_{j \in N_{i,\bar{g}}} x_j$.[14] Second, if the network is formed endogenously, players bear linking costs which increase linearly in the number of their sponsored links so that net payoffs write as

$$\Pi_i(\mathbf{x}, g) = \pi_i(\mathbf{x}, \bar{g}) - \eta_{i,g} k, \ \forall i \in N, \tag{1}$$

where $k \in (0, \infty)$.[15]

However, our model differs from existing models by implementing and combining two crucial premises: First, we assume that own activity and the activity of friends are *strategic complements*. In particular, the *best response function* resembles typical behaviour in online social networks as discussed in the introduction: a player desires to be active if at least one of his friends is active then increases his activity at a diminishing rate when his friends become more active:[16]

Assumption 1 (Strategic Complementarity)**.** *Player i's unique best response to activity* $\mathbf{x}_{-i}$ *on interaction network* $\bar{g}$ *is*

$$x_i^*(\mathbf{x}_{-i}, \bar{g}) = f\Big(\sum_{j \in N_{i,\bar{g}}} x_j\Big), \ \forall i \in N, \tag{2}$$

where $f(0) = 0$*,* $f' > 0$*, and* $f'' < 0$*.*

Second, we assume that using the (particular) social network is *beneficial* with *diminishing marginal utility* for optimising agents (e.g. due to time constraints and unmodeled outside options):[17]

[14] See for instance Bramoullé and Kranton (2007); Bramoullé *et al.* (2012); Galeotti *et al.* (2010).

[15] See for instance Galeotti and Goyal (2010); Zhang *et al.* (2011); Hiller (2010); Bramoullé *et al.* (2004).

[16] We later strengthen our assumption about f' to guarantee existence of a *unique* and *stable* strictly positive activity equilibrium on *any* exogenous connected network (see Section 3). In particular, *sufficient concavity* bounds total activity in the society under strategic complements (i.e. prevents an infinite solution). In contrast, Bramoullé *et al.* (2012) consider as a special case (increasing) linear best response correspondences (on exogenous networks) and introduce for the same purpose an upper bound on activity. Although we can approximate any such best response correspondence with concave functions, some of our results do not extend to the limit case as it introduces new corner solutions.

[17] The opposite case of increasing marginal utility, i.e. when h increases convexly, is essentially covered by the analysis of Hiller (2010).

Assumption 2 (Concave Value Function)**.** *Player i's maximised gross payoff under activity* $\mathbf{x}_{-i}$ *and interaction network* $\bar{g}$ *is*

$$\pi_i^*(x_i^*, \mathbf{x}_{-i}, \bar{g}) = h(\sum_{j \in N_{i,\bar{g}}} x_j), \ \forall i \in N,$$

where $h(0) = 0$*,* $h' > 0$ *and* $h'' < 0$*.*

In Appendix B.1, we introduce a family of social activity models and show that they meet the assumptions above. These models feature gross payoff functions of the following form:

Example 1 (A Family of Social Activity Models)**.**

$$\pi_i(\mathbf{x}, \bar{g}) = \lambda \big[v(\sum_{j \in N_{i,\bar{g}}} x_j) x_i \big]^{\frac{\tau}{\lambda}} - (c x_i)^{\tau},$$

where $\lambda \geq 2$, $\tau > 0$, $c \in (0, \infty)$, $v(0) = 0$, $v' > 0$, $v'' < 0$ and v^{τ} strictly concave.[18]

As an example, we frequently refer to the "baseline model" with a Cobb-Douglas utility function and linear activity costs—i.e. when $\lambda = 2$, $\tau = 1$, and v is a concave power function:

Example 2 (The Baseline Model)**.**

$$\pi_i(\mathbf{x}, g) = 2\sqrt{(\sum_{j \in N_{i,\bar{g}}} x_j)^q x_i} - c x_i, \tag{3}$$

where $q \in (0, 1)$.

Equilibrium Concept

We are interested in characterising activity patterns and network structures that can be sustained in Nash equilibria. As already mentioned above, we

[18] $\lambda \geq 2$ and $v'' < 0$ guarantee concavity of the best response correspondence. Assuming $\lambda = 2$ comes in fact at no loss of generality so that $v'(0) > c^2$ and $\lim_{x \to \infty} v'(x) < \frac{c^2}{n-1}$ ensure existence of a positive activity equilibrium on any (exogenously given) connected network (see Appendix B.1 for details).

consider two versions of the game:

(a) Exogenous Networks In Section 3, we assume that a (connected) interaction network $\bar{g}$ is exogenously given and known to all players. All players choose their activity simultaneously and hence a player's strategy is merely his level of activity x_i.

An activity vector $\mathbf{x}^*$ constitutes a (strict) Nash equilibrium iff it solves the best response functions of all players—as given in Equation (2)—simultaneously:

$$x_i^* = f(\sum_{j \in N_{i,\bar{g}}} x_j^*), \; \forall i \in N. \tag{4}$$

We call a Nash equilibrium with $\mathbf{x}^* = \mathbf{0}$ a zero activity equilibrium and an equilibrium with $\mathbf{x}^* > \mathbf{0}$ a positive activity equilibrium.

(b) Endogenous Network Formation In Sections 4 and 5, the network is formed endogenously: players simultaneously choose a level of activity x_i and a vector of supported links $g_i \in \mathfrak{g}_i$ so that they jointly compose the social interaction network $g = (g_1, \ldots, g_n)^t$. Thus each player's strategy is described by a pair (x_i, g_i) and players' utility is determined by net payoffs in Equation (1).

A pair $(\mathbf{x}^*, g^*)$ constitutes a Nash equilibrium of the game with endogenous network formation iff no player gains from choosing some alternative strategy (x_i', g_i') that is

$$\Pi_i(x_i^*, \mathbf{x}_{-i}^*, g^*) \;\; \geq \;\; \Pi_i(x_i', \mathbf{x}_{-i}^*, g_{-i}^* \odot g_i'), \; \forall i \in N, \, x_i' \geq 0, \, g_i' \in \mathfrak{g}_i.$$

As a simplification, we can disentangle the equilibrium conditions for activity and link formation by the following reasoning: since there is a unique best response activity $x_i^*(\mathbf{x}_{-i}, \bar{g})$ for any given network $\bar{g}$, we know that

$$\pi_i(x_i^*(\mathbf{x}_{-i}^*, \overline{g_{-i}^* \odot g_i'}), \mathbf{x}_{-i}^*, \overline{g_{-i}^* \odot g_i'}) \geq \pi_i(x_i', \mathbf{x}_{-i}^*, \overline{g_{-i}^* \odot g_i'})$$

for all $i \in N, x_i' \in [0, \infty), g_i' \in \mathfrak{g}_i$. That is, if an agent alters his link sponsorship, it is optimal to simultaneously adjust his activity to the new cumulative neighbour activity he accesses.

Thus a pair $(\mathbf{x}^*, g^*)$ is an equilibrium of the game with endogenous network formation if and only if no agent wants to change his level of activity x_i assuming network $\bar{g}^*$ to be fixed and nobody wants to change his sponsored links g_i given simultaneous adjustment of his activity for the new network. Mathematically, these equilibrium conditions read as

$$x_i^* = f(\sum_{j \in N_{i,\bar{g}^*}} x_j^*), \forall i \in N \tag{5}$$

and for all $i \in N$ and $g_i' \in \mathfrak{g}_i$

$$h(\sum_{j \in N_{i,\bar{g}^*}} x_j^*) - \eta_{i,g^*} k \geq h(\sum_{j \in N_{i,\overline{g^*_{-i} \odot g_i'}}} x_j^*) - \eta_{i,g^*_{-i} \odot g_i'} k. \tag{6}$$

We collect all pairs $(\mathbf{x}^*, g^*)$ that constitute an equilibrium of the game with endogenous network formation for some linking costs $k > 0$ in the set $\mathcal{E}$. Throughout the paper, we apply two equilibrium refinements: Our main focus lies on *generic* equilibria—collected in $\mathcal{E}_+ \subseteq \mathcal{E}$—which are robust to small changes in the linking costs.[19] We denote the set of undirected interaction graphs which arise in generic equilibria by $\bar{G}_{\mathcal{E}_+} = \{\bar{g} \in \bar{G} : \exists (\mathbf{x}^*, g^*) \in \mathcal{E}_+ \text{ with } \bar{g}^* = \bar{g}\}$. We also consider *strict* equilibria—collected in $\mathcal{E}_{++} \subseteq \mathcal{E}_+$—in which Equation (6) holds as a strict inequality.

As in the case of the pure activity game on exogenous networks, we call equilibria $(\mathbf{x}^*, g^*)$ with $\mathbf{x}^* = \mathbf{0}$ zero activity equilibria and equilibria with $\mathbf{x}^* > \mathbf{0}$ positive activity equilibria. We subdivide positive activity equilibria in single-level equilibria with $x_i^* = x_j^*$ for all $i, j \in N$ and multi-level equilibria in which there exist $i, j \in N$ with $x_i^* \neq x_j^*$.

[19] More precisely, we say that a pair $(\mathbf{x}^*, g^*)$ constitutes a *generic* equilibrium iff there exists an open interval $(\underline{k}, \bar{k})$ such that $(\mathbf{x}^*, g^*)$ constitutes an equilibrium for all $k \in (\underline{k}, \bar{k})$.

3 Exogenous Networks

In this section, we discuss pure activity games on exogenous interaction networks. Without loss of generality, we assume that the interaction network $\bar{g}$ is connected. Otherwise, the analysis of each component of a disconnected network follows the same lines as below.

The first proposition addresses the existence of Nash equilibria. For brevity, we call the minimum degree of a node in $\bar{g}$ (that is the smallest number of neighbours any player has) $n_{min} = \min_{i \in N} n_{i,\bar{g}}$ and the maximum degree $n_{max} = \max_{i \in N} n_{i,\bar{g}}$.[20]

Proposition 1. *For any connected network $\bar{g} \in \bar{G}^n$:*
(i) *A zero activity equilibrium with $\mathbf{x}^* = \mathbf{0}$ exists.*
(ii) *A positive activity equilibrium with $\mathbf{x}^* > \mathbf{0}$ exists and is uniquely defined*
    (a) ***if*** $f'(0) > \frac{1}{n_{min}}$ *and* $\lim_{x \to \infty} f'(x) < \frac{1}{n_{max}}$.
    (b) ***only if*** $f'(0) > \frac{1}{n_{max}}$ *and* $\lim_{x \to \infty} f'(x) < \frac{1}{n_{min}}$.
(iii) *There are no other equilibria.*

The proposition tells us that there are at most two equilibria: one in which every player is inactive and another one in which every player has strictly positive activity. In fact, the existence of the positive activity equilibrium depends on the slope of the best response function (relative to the properties of the network). On the one hand, if the best response function is not steep enough at the origin, some agent prefers to reduce his activity at any positive activity vector. On the other hand, if the best response is not sufficiently concave, some agent always prefers to increase his activity.

It is noteworthy that social activity games on exogenous networks are closely related to supermodular games:

Remark 1. π_i has positive cross partial derivatives at best response activity by Assumption 1, i.e. $\frac{\partial^2 \pi_i}{\partial x_i \partial x_j}\Big|_{x_i = x_i^*(\mathbf{x}_{-i}, \bar{g})} \geq 0$ for all $j \neq i \in N$.

In an orthodox supermodular game, π_i would have to be supermodular in $(x_i, \mathbf{x}_{-i})$, i.e. $\frac{\partial^2 \pi_i}{\partial x_i \partial x_j} \geq 0$ for all $j \neq i \in N$, and X would have to be a closed interval in $\mathbb{R}$. Thus similar concepts as used in the literature on supermodular

[20] Note that $n - 1 \geq n_{i,\bar{g}} \geq 1$ as $\bar{g}$ is connected.

games are applicable by restricting the strategy space conveniently.[21] We discuss some resulting properties of the positive activity equilibrium in the supplementary material in Appendix B.2.[22]

Consider now a possibly *disconnected* interaction network $\bar{g} \in \bar{G}^n$. Following the insights from Proposition 1, we strengthen our assumption on f' for the remainder of the paper to guarantee existence of a positive activity equilibrium in any (non-singleton) *component* of $\bar{g}$:

Assumption 3. $f'(0) > 1$ *and* $\lim_{x \to \infty} f'(x) < \frac{1}{n-1}$.

Hence for any network $\bar{g} \in \bar{G}^n$, a particular type of activity vector $\mathbf{x}^*_{\bar{g}}$ in which isolated players choose zero activity and non-isolated players choose strictly positive *equilibrium* activity is well defined, that is:

Definition 1. For any $\bar{g} \in \bar{G}^n$, let $\mathbf{x}^*_{\bar{g}}$ be such that $x^*_i > 0$ iff $n_{i,\bar{g}} \geq 1$ and $x^*_i = f(\sum_{j \in N_{i,\bar{g}}} x^*_j)$ for all $i \in N$.

Note that if $\bar{g}$ is connected then $\mathbf{x}^*_{\bar{g}}$ coincides with the strictly positive activity equilibrium vector discussed in the proposition.

4 Endogenous Network Formation

In the social activity game with endogenous network formation, each player simultaneously chooses some level of activity and forms links. Players' actions can be summarised in a pair $(\mathbf{x}, g)$ and an equilibrium is denoted by $(\mathbf{x}^*, g^*)$.

As discussed in Section 2, an equilibrium of this game has to fulfill two different conditions: First, each player's activity has to be optimal given the activity of his neighbours in the induced network, see Equation (5). Second,

[21] The main results of the literature were developed in the classic papers Topkis (1979), Vives (1990), and Milgrom and Roberts (1990). For a more recent overview see Vives (2005).

[22] In particular, we show that the equilibrium is stable, displays multiplicative effects so that every player gains from the introduction of an additional link, and that individual equilibrium activity is smaller than the cumulative neighbour activity. Note that our results for endogenous network formation determine activity patterns on particular exogenous networks. However, a general characterisation of activity patterns on arbitrary exogenous networks (e.g. by connecting individual equilibrium activity with some measure of centrality) is left open for future research.

each player's link sponsorship has to be optimal given the actions of all other players (even when he simultaneously adjusts his own activity) as stated in Equation (6).

The second condition can be split into three separate (sub-)statements due to concavity of h—i.e. due to the diminishing marginal utility of optimising agents: the most profitable potential linking deviation is either to add a *single* link to the most active non-neighbour or delete the *single* least valuable self-sponsored link or to do both simultaneously. Thus we finally arrive at four simple equilibrium conditions:

Lemma 1. *A pair* $(\mathbf{x}^*, g^*)$ *constitutes a Nash equilibrium of the endogenous network formation game for linking costs* $k \in (0, \infty)$ *iff for all* $i \in N$

i. $x_i^* = f(\sum_{j \in N_{i,\bar{g}^*}} x_j^*)$,

ii. $h(\sum_{j \in N_{i,\bar{g}^*}} x_j^* + \bar{x}_i) - h(\sum_{j \in N_{i,\bar{g}^*}} x_j^*) \leq k$,

iii. $h(\sum_{j \in N_{i,\bar{g}^*}} x_j^*) - h(\sum_{j \in N_{i,\bar{g}^*}} x_j^* - \underline{x}_i) \geq k$,

iv. $\underline{x}_i \geq \bar{x}_i$,

where $\bar{x}_i = \max_{j:\, \bar{g}_{ij}^*=0}\{x_j^*, 0\}$ *and* $\underline{x}_i = \min_{j:\, g_{ij}^*=1}\{x_j^*, \infty\}$.[23]

Recall that $\mathbf{x}_{\bar{g}}^*$ denotes a uniquely defined vector of activity in which non-isolated players (and only those) choose strictly positive equilibrium activity by Definition 1 from Section 3. As linking costs are strictly positive, it is easy to see that in equilibrium a player chooses zero activity if and only if he is isolated, i.e. $\mathcal{E} \subseteq \{(\mathbf{x}_{\bar{g}}^*, g) : g \in G\}$. Furthermore, isolated and non-isolated players cannot co-exist in generic equilibria.[24] The following proposition provides a formal statement and elaborates on the implied partition of $\mathcal{E}_+$:

[23] For completeness, we define $h(-\infty) = -\infty$.

[24] In such an equilibrium, isolated players would have to be indifferent between their isolation and forming a single link to the / a most active player (with simultaneous activity adjustment). These equilibria are not robust to small changes of k and thus excluded from $\mathcal{E}_+$. We fully characterise them in the supplementary material in Appendix B.3. Note that $\mathcal{E}_+$ also rules out some non-generic *positive activity* equilibria from $\mathcal{E}$.

Proposition 2. *$\mathcal{E} \subseteq \{(\mathbf{x}^*_{\bar{g}}, g) : g \in G\}$ and $\mathcal{E}_+$ is partitioned into two sets:*
(i) *A zero activity equilibrium with $\mathbf{x}^* = \mathbf{0}$ and empty network $\bar{g}^*_{emp}$*
(ii) *Positive activity equilibria with $\mathbf{x}^* > \mathbf{0}$ and $n_{i,\bar{g}^*} \geq 1, \forall i \in N$*

As in the pure activity game on exogenous networks, all players are either active or all are inactive in generic equilibria. However, so far we have not discussed for which (interval of) linking costs the respective equilibria exist. The next proposition provides some general insights, where $\bar{\mathbf{k}}_{max} = \lim_{x\to\infty} h(f(x))$ is the supremum value of a single link:

Proposition 3. *The zero activity equilibrium exists for any k.*
There are thresholds $0 < \underline{\mathbf{k}}_n \leq \bar{\mathbf{k}}_n < \bar{\mathbf{k}}_{max}$ such that
(i) *a positive activity equilibrium exists iff $k \leq \bar{\mathbf{k}}_n$.*
(ii) *there is a unique positive activity equilibrium with $\bar{g}^* = \bar{g}_{com}$ if $k < \underline{\mathbf{k}}_n$.*

The proposition sets the agenda for the remainder of the paper as it suggests two different sorts of questions: The first sort of questions concern the existence of different types of positive activity equilibria. For instance, we would like to find out which interaction networks—apart from the complete and the empty network—can arise for linking costs $k \in [\underline{\mathbf{k}}_n, \bar{\mathbf{k}}_n]$ in equilibrium. Likewise, we would like to determine the distribution of activity in such equilibria—e.g. the number of different activity levels that can be sustained. The remainder of this section is dedicated to this agenda. Section 4.1 discusses equilibria with a single positive level of activity whereas Section 4.2 addresses equilibria that display multiple positive activity levels and Section 4.3 studies *strict* equilibria.

The second sort of questions concern "prominence" of different equilibria—that is finding out which positive activity equilibria are supported by large intervals of linking costs. If, for instance, $\bar{\mathbf{k}}_n \approx 0$ then the zero activity equilibrium is the unique equilibrium for most k. Conversely, if $\bar{\mathbf{k}}_n \gg 0$ and $\underline{\mathbf{k}}_n \approx \bar{\mathbf{k}}_n$ then an equilibrium with a complete network exists for most k and is the unique positive activity equilibrium. We discuss these questions in Section 5 in the context of large societies and show how individual preferences determine which positive activity equilibrium prevails.

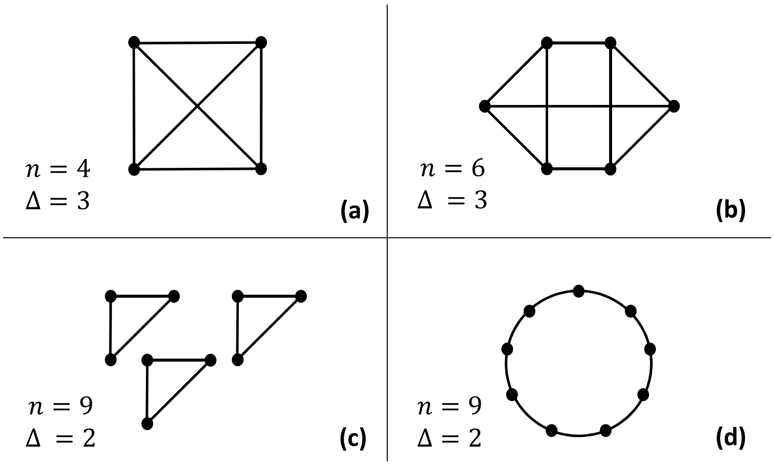

Figure 1: Regular Networks

4.1 Single-Level Equilibria

In this section, we analyse single-level equilibria in which each agent chooses symmetric equilibrium activity $x_i^* = x > 0$.

For the discussion of our results we need to introduce the notion of a (non-empty) "regular network": a network $\bar{g}$ is called Δ-regular iff each player i has common degree $n_{i,\bar{g}} = \Delta \geq 1$. Some regular networks are depicted in Figure 1. We define $\bar{G}_{reg}$ as the set of all regular graphs and $\bar{G}_{reg}^{\Delta} \subset \bar{G}_{reg}$ as the set of all Δ-regular graphs.[25] The following results hold:

Proposition 4.
(i) $(\mathbf{x}_{\bar{g}}^*, g) \in \mathcal{E}_+$ *and displays a single positive level of activity iff* $\bar{g} \in \bar{G}_{reg}$.[26]
(ii) *If* $\bar{g} \in \bar{G}_{reg}^{\Delta}$ *then* $x_{i,\bar{g}}^* = x_\Delta$ *independent of* n *with* $x_\Delta > x_{\Delta-1}$.[27]

The proposition tells us, first, that every regular network is an equilibrium outcome for some interval of linking costs. Second, every single-level equilibrium has a regular interaction network. Third, the common individual

[25] It is a well established result in graph theory that a Δ -regular graph of size n exists if Δn is even due to the existence condition in Erdős and Gallai (1960).

[26] The proof in fact also shows that any single level is generic, i.e. there are no non-generic single-level equilibria in $\mathcal{E}$.

[27] For the comparative statics we implicitly assume $n \geq \Delta + 1$ and $x_0 = 0$.

activity is solely determined by, and increases in, the degree Δ of the regular network—independent of the population size n.

This last finding stems from the *locality of social interactions* on the network. Consider for instance regular networks with $\Delta = 2$, i.e. circles: As each player interacts with exactly two other players, the (common) individual equilibrium activity is the same in networks with three circles of three players (Figure 1 (c)), one circle of nine players (Figure 1 (d)), and one circle of nine hundred players.

Two further regularities follow immediately: As all players access the same cumulative neighbour activity in equilibria with Δ-regular networks, they have common equilibrium (gross) payoff $\pi^* = h(\Delta x_\Delta)$. For the same reason, all equilibria with Δ-regular networks have a common interval of equilibrium supporting linking costs $[\underline{k}_\Delta, \bar{k}_\Delta]$—independent of the population size n—as long as $\Delta < n - 1$.[28] If $\Delta = n - 1$ instead, then the network is complete (e.g. Figure 1 (a)); as no player can add a link even if desired, the equilibrium exists for all $k \in (0, \bar{k}_\Delta]$ instead.

Finally, note that our results in Proposition 4 crucially depend on the concavity of the value function. Lemma 5 in Hiller (2010) essentially shows that, under a convex value function, players choose symmetric equilibrium activity if and only if the network is either empty or complete.

4.2 Multi-Level Equilibria

In this section, we examine (generic) multi-level equilibria. In these equilibria, all agents choose positive activity and there are at least two agents i and j with $x_i^* \neq x_j^*$.

We split this section into three different parts: first, we discuss general properties that hold in any multi-level equilibrium. Subsequently, we analyse two prominent types of networks as examples more in detail: we show that every biregular bipartite network exists in equilibrium, whereas core-periphery structures only arise under particular specifications of the model.

[28]It depends on the concrete payoff function how the bounds $\underline{k}_\Delta$ and $\bar{k}_\Delta$ change as a function of Δ as shown by example of the baseline model in the supplementary material in Appendix B.6.

4.2.1 General Properties

To begin with, it is important to find out whether (and how many) different levels of activity can be sustained in equilibrium in principle. A detailed discussion of so-called complete multipartite networks in Appendix B.4 supports the following general insight:

Result 1. *There are multi-level equilibria which display a substantial number of distinct levels of activity.*

Next, we consider the direction of link sponsorship in multi-level equilibria. The lemma below tells us that for any equilibrium in which a highly active player sponsors a link to a less active player, there exists another equilibrium in which the less active player sponsors the link, i.e. the link sponsorship is switched. Furthermore, the latter equilibrium is sustainable under a greater—and sometimes *strictly* greater—interval of linking costs than the former.

Lemma 2. *Let* $(\mathbf{x}^*, g^*) \in \mathcal{E}$ *and consider* g' *with* (i) $g'_{ij} = 1$ *and* $g'_{ji} = 0$ *if* $\bar{g}^*_{ij} = 1$ *and* $x^*_i < x^*_j$*;* (ii) $g'_{ij} = g^*_{ij}$ *otherwise. Then* $(\mathbf{x}^*, g')$ *is an equilibrium for a weakly greater interval of* k*.*

The lemma has strong implications: when we (dis-)prove existence of equilibria with certain interaction networks, we can assume that there is "upward linking", i.e. no link is sponsored by an adjacent player with strictly higher activity. There is quite a neat interpretation for those equilibria: links are always sponsored by the adjacent player with higher valuation.

Under endogenous network formation, we can show a strong connection between an agent's activity and his position in the network. First note that as utility increases in the activity of neighbours, players with higher activity are more "popular" friends. As a result, higher activity players have more neighbours in any equilibrium. In the case of Facebook, a strong positive correlation between social degree and user activity has indeed been confirmed empirically (see Wilson *et al.* (2009)).

Although high activity players have many neighbours, they tend to initiate few links themselves: as they have already many friends, their marginal value for any particular link—and thus their willingness to pay for links—is small. Thereby, high activity or “popular” player benefit twice from their attractiveness: they not only have many friends but also tend to pay little for link formation.

If we assume upward linking explicitly, there are two more regularities: first, an agent who sponsors a link to another agent with some particular level of activity also forms links to all agents with higher activity. Second, agents who access the same cumulative neighbour activity through incoming links behave similarly: they sponsor the same number of links and choose the same level of activity. The next proposition describes these properties formally:

Proposition 5. *Let* $(\mathbf{x}^*, g^*) \in \mathcal{E}_+$ *such that* $g^*_{ij} = 1$ *implies* $x^*_i \leq x^*_j$.
(i) $g_{ij} = 1$ *and* $x^*_{j'} > x^*_j$ *imply* $g_{ij'} = 1$.
(ii) $x^*_i > x^*_j$ *implies* $\eta_i \leq \eta_j$ *and* $n_i > n_j$.
(iii) $\sum_{l:\, g_{li}=1} x^*_l = \sum_{l:\, g_{lj}=1} x^*_l$ *iff* $\eta_i = \eta_j$ *and* $x^*_i = x^*_j$.

Finally, casual observation shows that online social networks are (normally) not minimally connected and contain cycles.[29] Similarly, we can show that “sparse” interaction networks arise in our model in equilibrium only under special conditions. First, if there are more than two levels of activity then the *network* is connected that is there is a path between any two nodes. Second, the only minimally connected *components* are stars or lines with four players.[30]

Proposition 6. *Let* $(\mathbf{x}^*, g^*) \in \mathcal{E}_+$ *with* $\mathbf{x}^* > \mathbf{0}$.
(i) *If* $\mathbf{x}^*$ *displays more than two levels of activity then* $\bar{g}^*$ *is connected.*
(ii) *Any minimally connected component of* $\bar{g}^*$ *is either a star or a four-line.*

4.2.2 Example: Biregular Bipartite Networks

An interaction network is a *bipartite graph* if the set of agents N can be partitioned into two independent sets $P_{1,2}$ (so-called *partite sets* or *parts*).

[29] A connected network (or a component of a network) is called *minimally connected* if it does not contain any cycles, i.e. the removal of any link disconnects the network (component).

[30] A *line* is a minimally connected network (component) with maximal degree two.

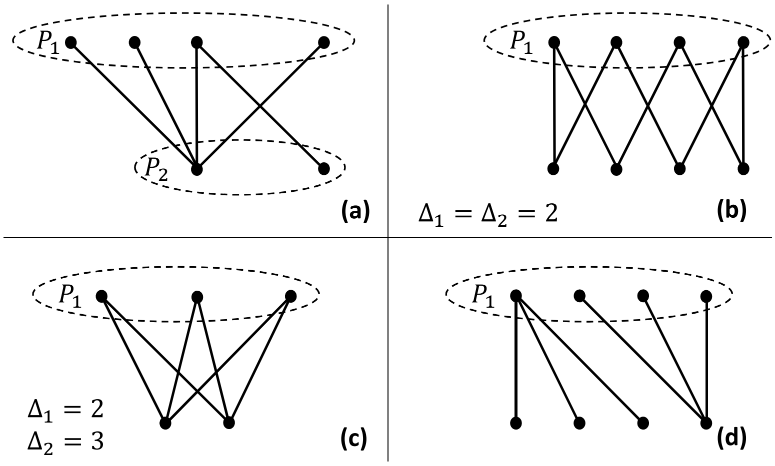

Figure 2: Bipartite Networks

Some bipartite graphs are depicted in Figure 2.

A bipartite graph is *biregular* if there exists a bipartition such that players within each partite set have common degree. We implicitly assume that the common degree differs between both parts that is $n_{i,\bar{g}} = \Delta_j$ for all $i \in P_j$ and $1 \leq \Delta_1 < \Delta_2$.[31] We collect these biregular bipartite graphs in the set $\bar{G}_{bp}^{\Delta_{1,2}}$.

Biregular bipartite graphs include for instance star networks and complete bipartite graphs. The bipartite graph in Figure 2 (c) is biregular. Bipartite network (d) is biregular as well, which can be seen by redefining the partite sets, whereas (b) is regular and (a) is neither regular nor biregular. The following results apply:

Proposition 7. (i) $\bar{G}_{bp}^{\Delta_{1,2}} \subset \bar{G}_{\mathcal{E}_+}$.
(ii) *If* $\bar{g}^* \in \bar{G}_{bp}^{\Delta_{1,2}}$ *then* $x_i^* = x_j$ *for all* $i \in P_j$ *and* $\Delta_1 < \Delta_2$ *implies*

$$1 < \frac{x_2}{x_1} < \frac{\Delta_2}{\Delta_1} = \frac{|P_1|}{|P_2|}.$$

The proposition states that every biregular bipartite graph can be sustained in equilibrium for some interval of linking costs (part (i)). Players from the same partite set choose common equilibrium activity, which is higher in

[31]If $\Delta_1 = \Delta_2$, then $\bar{g}$ is regular and thus, by Proposition 4, there is a single-level equilibrium with $\bar{g}^* = \bar{g}$.

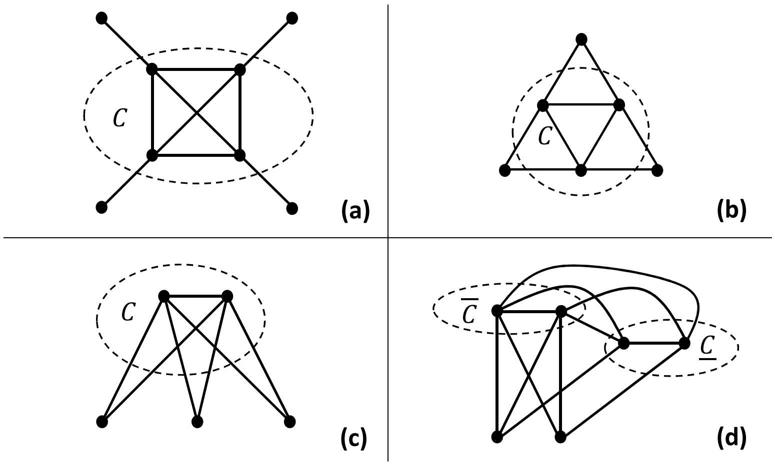

Figure 3: Core-Periphery Networks

the smaller set (part (ii)).

Moreover it is noteworthy that any collection of similar biregular bipartite graphs constitutes a biregular bipartite graph itself. That is, not only a single star with $n-1$ spokes is an equilibrium network for some linking costs but also any collection of stars with s spokes such that $s+1$ divides n.

4.2.3 Example: Core-Periphery Networks

An interaction network is a *core-periphery network* if players can be partitioned into two sets, core players in $\mathcal{C}$ and periphery players in $\mathcal{P}$, and the following three conditions are met:

(i) Core players share links with all other core players, $\bar{g}_{ij} = 1, \ \forall i, j \in \mathcal{C}$.
(ii) Periphery players do not share links among themselves, $\bar{g}_{ij} = 0, \ \forall i, j \in \mathcal{P}$.
(iii) Core players have a link to a per. player, $\forall i \in \mathcal{C}$, $\exists j \in \mathcal{P}$ s.t. $\bar{g}_{ij} = 1$.[32]

We collect all core-periphery networks in the set $\bar{G}_{cp}$. The subset $\bar{G}_{cp}^{com} \subset \bar{G}_{cp}$ contains the special class of *complete* core-periphery network in which

[32]The last condition is w.l.o.g. and purely introduced for presentation purposes. It guarantees that $|\mathcal{C}|$ is well defined for any core-periphery network: otherwise, for instance, network (c) in Figure 3 could be regarded as a core-periphery network with either $|\mathcal{C}| = 2$ or with $|\mathcal{C}| = 3$.

all periphery players share a link with all core players. Some examples are depicted in Figure 3, where network (c) is the only complete core-periphery network.

The following proposition summarises our main findings:

Proposition 8. *Let* $|\mathcal{P}| \in \{3, \ldots, n-2\}$.[33] *Then:*
(i) $\bar{g} \in \bar{G}_{cp} \cap \bar{G}_{\mathcal{E}_+}$ implies $\bar{g} \in \bar{G}_{cp}^{com}$.
(ii) *For any* $\bar{g} \in \bar{G}_{cp}^{com}$, $\bar{g} \in \bar{G}_{\mathcal{E}_+}$ *iff* $|h''|$ *is (locally) sufficiently close to 0.*[34]
(iii) *If* $\bar{g}^* \in \bar{G}_{cp}^{com}$ *then* $x_i^* = x_p$ *and* $x_j^* = x_c$ *for all* $i \in \mathcal{P}$ *and* $j \in C$ *with*

$$x_p < x_c < |\mathcal{P}| x_p.$$

The intuition for the results is quite simple: Core players have higher equilibrium activity in any (exogenous) core-periphery interaction network due to complementarity. Thereby, the marginal value of a link is relatively small for core players. Any linking costs preventing them from deleting their marginal link to another core player makes periphery players willing to add a link to a core player. Therefore, *non-complete* core-periphery networks do not exist in equilibrium (part (i)).

In contrast, in a *complete* core-periphery network, periphery players already share a link with all (more active) core players and could only add a link to a (less active) periphery player. That alleviates the tension between correct incentives for core and periphery players so that those networks exist in equilibrium if h is sufficiently linear (part (ii)).

[33] If $|\mathcal{P}| = 1$ the network is complete which is analysed in Proposition 4. If $|\mathcal{P}| = n-1$ then the network is a star which is covered by Proposition 7. Finally, the special case $|\mathcal{P}| = 2$ is covered in Appendix B.5.

[34] The reader might worry that by changing the value function h, the best response function f and, thereby, the equilibrium activity is altered. However, any payoff function π can be transformed to a payoff function $\hat{\pi}$—by adding an appropriate function $v(\sum_{j \in N_{i,\bar{g}}} x_j)$ which does not depend on x_i—such that $\hat{f} = f$ and $\hat{h}$ becomes (locally) arbitrarily close to a linear function.

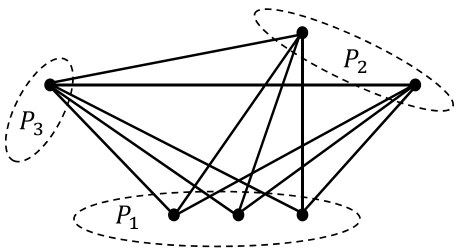

Figure 4: Complete Multipartite Network

4.3 Strict Nash Equilibria

We now apply an equilibrium refinement and study strict Nash equilibria $\mathcal{E}_{++} \subset \mathcal{E}_{+}$. As an example, recall from Section 4.2.2 that biregular bipartite graphs can be sustained in *generic* equilibria, i.e. $\bar{G}_{bp}^{\Delta_{1,2}} \subset \bar{G}_{\mathcal{E}_{+}}$. However, the equilibrium might not be *strict*: If a player (only) shares links with a subset of players from the other partite set, he can relocate one of his self-sponsored links to a non-neighbour (in that partite set) without effecting his payoff.

In order to find the set of networks which can be sustained in *strict* equilibria, some additional notation is necessary: *Multipartite networks* are a generalisation of bipartite networks, in which the set of agents N can be partitioned into any number of independent sets (so-called *partite sets* or *parts*). In a *complete* multipartite network, every agent shares a link with all agents outside its own part. An example of such a network can be seen in Figure 4. Also recall that a clique is a group of fully connected players.

Proposition 9. *Any network $\bar{g}^* \in \bar{G}_{\mathcal{E}_{++}}$ which is supported by a strict Nash equilibrium takes either of the following forms:*
(i) *a complete multipartite graph of different-sized partite sets, or*
(ii) *variations of these graphs in which one partite set is replaced by a clique.*

The proposition allows—among others—for the empty network, the complete network, complete core-periphery networks and complete bipartite graphs in *strict* equilibria. Conversely, (non-complete) regular graphs and (non-complete) biregular bipartite graphs are not supported in *strict* equilibria—even though they are supported in *generic* equilibria as shown in Sections 4.1 and 4.2.

Due to complementarity, the activity levels in strict equilibria can easily be ranked: If it exists, agents in the clique have the highest level of activity as they share a link with every other player. Other than that, agents in the same partite set have symmetric activity which is anti-proportional to the number of agents in their part.

5 Large Societies

While the last section was concerned with *existence* of different types of positive activity equilibria, this section focuses on the *prominence of different equilibria* in large societies—that is finding equilibria that can be sustained for large intervals of linking costs. We are going to show two main results:

First, we show in Section 5.1 that a positive activity equilibrium exists in large societies for any linking costs smaller than the supremum value of a single link (as long as individual activity is not bounded from above).

Second, in Section 5.2, we demonstrate by example of the *baseline model* that it depends on the player's preferences *which* type of positive activity equilibrium prevails in large societies and *which* ceases to exist.

5.1 Positive Activity Equilibria in Large Societies

From our previous discussion, we already know that positive activity equilibria exist in large societies *for some linking costs*: First, equilibria with (non-complete) Δ-regular interaction networks exist for linking costs in the interval $[\underline{k}_\Delta, \bar{k}_\Delta]$ independent of population size n (see Section 4.1).[35] Second, biregular bipartite networks of any size are supported by a non-degenerate interval of linking costs (see Proposition 7). However, we are going to show in this section that equilibria with a star or a complete network are special insofar as one of them exists in large societies *independent of the linking costs.*

For that purpose, we extend our assumptions on f and h as follows:

[35]Assuming a (non-complete) Δ-regular network with n players exists, see also Footnote 25.

Assumption 4. (i) $\lim_{x\to\infty} f'(x) = 0$*;* (ii) $\lim_{x\to\infty} f(x) = \infty$; (iii) $h(x + f(x)) - h(x)$ *and* $h(x) - h(x - f(x))$ *have finitely many (local) extrema in* $\mathbb{R}_+$.

The assumption is, for instance, met by the baseline model in Example 2. We now provide a brief intuition for the effects of each part of the assumption: By the first part, Proposition 1 implies that the equilibrium activity vector $\mathbf{x}^*_{\bar{g}}$, in which every non-isolated player chooses strictly positive activity, exists for any $\bar{g} \in \bar{G}$—independent of the number of players n. Thus it becomes meaningful to study the game with endogenous network formation in the limit as n goes to infinity.

The second part implies that individual activity—and thereby the activity that can be accessed through a single link—is unbounded from above. Otherwise, the complete network would not be an equilibrium for large linking costs and, for small linking costs, spokes in a star network would like to form links among themselves—even as the societies becomes large.

Finally, the third part is a technical assumption which ensures that the limit intervals of linking costs (for which the star or the complete network can be sustained in equilibrium) are well defined.[36]

Recall from Proposition 3 that $\bar{\mathbf{k}}_{max} = \lim_{x\to\infty} h(f(x))$ is the supremum value of a single link so that for $k \geq \bar{\mathbf{k}}_{max}$ the unique equilibrium displays zero activity and has an empty network. The following result holds true under Assumption 4:

Proposition 10. *A positive activity equilibrium with either the star or the complete network exists for any* $k \in (0, \bar{\mathbf{k}}_{max})$ *for sufficiently large* n.

The proposition implies that $\bar{\mathbf{k}}_n$—the upper bound of linking costs for positive activity equilibria (see Proposition 3)—converges to $\bar{\mathbf{k}}_{max}$ in n. If, additionally, the value function h is unbounded from above then $\bar{\mathbf{k}}_{max} = \infty$ and an equilibrium with *either* the star *or* the complete network exists for

[36] A weaker version of the third part states that both functions either converge to a finite limit for $x \to \infty$ or diverge to $\pm\infty$. Then a weaker version of Proposition 10 holds: there is at most one $k \in (0, \bar{\mathbf{k}}_{max})$ for which neither the star nor the complete network are an equilibrium for sufficiently large n.

any linking costs $k \in (0, \infty)$ in sufficiently large societies. As we demonstrate in the next section, it depends on the players' preferences *which* of the two equilibria prevails and which ceases to exist in large societies.

There is a deeper reason for our findings in this section. As we are going to argue, the complete and the star network are conceptional counterparts. Take the complete network first. No player can add any links, even if desired, and therefore there is no lower bound of linking costs under which such an equilibrium is supported. Conversely, as the star is minimally connected, there are very strong incentives to keep this single link. For this reason, the upper bound of linking costs supporting a star in equilibrium converges to $\bar{\mathbf{k}}_{max}$ as the society becomes large—independent of the details of the model.

5.2 Individual Preferences and the Social Network

In this section, we show how individual preferences shape the social network in large societies. We focus on the baseline model (Example 2) which meets Assumption 4. Recall that agents have in this model a Cobb-Douglas utility function and linear activity costs, that is

$$\pi_i(\mathbf{x}, g) = 2\sqrt{(\sum_{j \in N_{i,\bar{g}}} x_j)^q x_i} - cx_i,$$

where $q \in (0, 1)$. Both the best response function f and the value function h are then proportional to the concave power function $(\sum_{j \in N_i} x_j)^q$. If q is small, the power function takes the form of a step function and players are relatively satiated even from little neighbours' activity. If q is closer to one, the power function becomes more linear and agents get less easily satiated.

To present our results, we denote by K_∞ the limiting set of linking costs which support a certain type of network in large societies.[37] For short, we say that a type of network is *robust* if $K_\infty = (0, \infty)$ and is *not robust* if $K_\infty = \emptyset$. Details and proofs of all results presented in this section as well as further

[37] Formally, consider an infinite sequence of networks with increasing number of players $\bar{\mathbf{g}} = (\bar{\mathbf{g}}_2 \in \bar{G}^2, \bar{\mathbf{g}}_3 \in \bar{G}^3, \dots)$. Then $k \in K_\infty(\bar{\mathbf{g}})$ if and only if there exist a threshold $\hat{n}(k) \geq 2$ such that, for all $n' > \hat{n}(k)$, there exists an equilibrium $(\mathbf{x}^*, g^*)$ with $\bar{g}^* = \bar{\mathbf{g}}_{n'}$ for linking costs k.

q	empty	complete	Star
$\in (\frac{1}{2}, 1)$	$(0, \infty)$	$(0, \infty)$	$\emptyset$
$= \frac{1}{2}$	$(0, \infty)$	$(0, \frac{1}{2}c^{-3}]$	$[\frac{1}{2}c^{-3}, \infty)$
$\in (0, \frac{1}{2})$	$(0, \infty)$	$\emptyset$	$(0, \infty)$

Table 1: Limit supporting sets K_∞ of selected networks (baseline model)

discussion are relegated to the supplementary material in Appendix B.6.

Table 1 summarises the limit supporting sets K_∞ of selected network-types in the baseline model. The empty network in which all players are inactive is robust independently of q in accordance with Proposition 3. Furthermore, a positive activity equilibrium with either the star or the complete network always exists for sufficiently large n as argued in Proposition 10. In particular, if $q > \frac{1}{2}$ then the complete network is robust and if $q < \frac{1}{2}$ then the star is robust.

These findings indicate how different individual user preferences can shape (online) social networks. If users are foremost interested in participation per se (e.g. have some people that follow them on the social network or provide some content) but get quickly satiated, having several links is a waste and the star is a robust equilibrium. Conversely, if social interaction plays a more prominent role (and own activity is responsive to neighbour activity even at high levels) then users want to form links to many other users and the complete network is robust.

Although our model is rather simple and static, empirical findings for online social networks point in a similar direction. User interaction in the network plays a much more central role in pure online social networks such as Facebook than in content distribution sites with (less elaborate) social network components such as YouTube. Most activities on Facebook (including messaging, tagging people in photos, and many apps) are direct social interaction between Facebook-friends. In contrast, many people who watch videos on YouTube do not even own a channel and hence are not part of the social network. Wilson *et al.* (2009) show that pure social networks typically exhibit much higher social connectivity and their degree distributions first-order stochastically dominate those of content distribution sites with social network components.

6 Conclusion

This paper discusses games of social activity with strategic complementarity on exogenous networks as well as under endogenous network formation. Our assumptions resemble key features of online social networks and we are able to replicate several empirical findings.

On any exogenous network, there is a zero activity equilibrium as well as a positive activity equilibrium (under mild conditions).

Under endogenous network formation, either all players are isolated and have zero activity or all players have at least one neighbour and strictly positive activity. In the latter case, they have a common level of activity iff the network is regular.

In general, there are equilibria with many different levels of activity. We have highlighted several regularities: as in empirical findings, the degree of an agent and his equilibrium activity are positively correlated. Furthermore, "popular" players with many friends tend to sponsor few links themselves and equilibrium networks are—with few exceptions—not minimally connected. In strict equilibria, interaction networks take a distinct form: they are either complete multipartite graphs or particular variations of them.

Finally, we have shown how individual preferences shape the social structure in large societies. Our results suggest that we should expect denser social networks for services with a more elaborate range of features in which users get less easily satiated. Empirical findings indeed show that pure social network sites such as Facebook exhibit a significantly higher connectivity than content distribution sites with social network components such as YouTube.

Our research is extendable in several directions. First, our results on endogenous networks only characterise activity patterns on selected (exogenous) networks. It would be desirable to characterise activity on arbitrary exogenous networks more comprehensively. Second, we have left welfare considerations out of this paper due to the amount of material. Third, we only consider one-shot interactions whereas real world social networks evolve dynamically. Finally, agents are homogenous in our model. Introducing heterogeneity might lead to further insights and a better description of real world social networks.

A Proofs from the Main Body

Proof of ***Proposition 1***. ***Part (i):*** Pick some player i randomly and assume that $\forall j \in N \setminus \{i\}$, $x_j^* = 0$. Then the best response in Equation (2) implies $x_i^* = 0$ for any g; thus $x_i^* = 0$, $\forall i$ is an equilibrium for any g.

Part (ii): We are first going to show that there is at most one equilibrium with $\mathbf{x}^* > \mathbf{0}$ and then discuss the conditions for its existence separately.

We define the following function $f_{\bar{g}} : X^n \to X^n$ as

$$f_{\bar{g}}(\mathbf{x}) = \begin{pmatrix} f(\sum_{j \in N_{1,\bar{g}}} x_j) \\ \vdots \\ f(\sum_{j \in N_{n,\bar{g}}} x_j) \end{pmatrix};$$

$f_{\bar{g},i}(\mathbf{x})$ denotes the ith entry of $f_{\bar{g}}(\mathbf{x})$.

From the equilibrium conditions in Equation (4), it is clear that $\mathbf{x}^*$ is an equilibrium of the social activity game on exogenous network $\bar{g}$ iff $\mathbf{x}^*$ is a fixed point of $f_{\bar{g}}$ that is $f_{\bar{g}}(\mathbf{x}^*) = \mathbf{x}^*$.

We are going to use three important properties of $f_{\bar{g}}$. First, $f_{\bar{g}}$ is non-decreasing that is $\mathbf{x} \geq \mathbf{y}$ implies $f_{\bar{g}}(\mathbf{x}) \geq f_{\bar{g}}(\mathbf{y})$ as f is strictly increasing and so $f(\sum_{j \in N_{i,\bar{g}}} x_j) \geq f(\sum_{j \in N_{i,\bar{g}}} y_j)$, $\forall i$. Second, $\mathbf{x} > \mathbf{0}$ implies $f_{\bar{g},i}(\mathbf{x}) > 0$ as g is connected so that each player i has at least one neighbour. Third, $f_{\bar{g}}(\lambda \mathbf{x}) > \lambda f_{\bar{g}}(\mathbf{x})$ for $\lambda \in (0, 1)$ and $\mathbf{x} > \mathbf{0}$ as f is strictly concave and g connected i.e. $f(\sum_{j \in N_{i,\bar{g}}} \lambda x_j) > \lambda f(\sum_{j \in N_{i,\bar{g}}} x_j)$, $\forall i$.

Let $u(\mathbf{x}) = f_{\bar{g}}(\mathbf{x}) - \mathbf{x}$. Using notation from Kennan (2001), u is strictly R-quasiconcave: $u(\mathbf{x}) = \mathbf{0}$, $\mathbf{x} > \mathbf{0}$, and $\lambda \in (0, 1)$ implies $u(\lambda \mathbf{x}) = f_{\bar{g}}(\lambda \mathbf{x}) - \lambda \mathbf{x} > \lambda (f_{\bar{g}}(\mathbf{x}) - \mathbf{x}) = \mathbf{0}$. u is also quasi-increasing: $x_i = y_i$ and $x_j \geq y_j$ for all j implies $u_i(\mathbf{x}) = f_{\bar{g},i}(\mathbf{x}) - x_i \geq f_{\bar{g},i}(\mathbf{y}) - y_i = u_i(\mathbf{y})$. Thus the conditions of theorem 3.1 in Kennan (2001) are met and $f_{\bar{g}}$ has at most one fixed point with $\mathbf{x}^* > \mathbf{0}$.

Part (ii) (a): First, we show that $f'(0) > \frac{1}{n_{min}}$ implies existence of $\mathbf{a}$ with $f_{\bar{g}}(\mathbf{a}) > \mathbf{a} > \mathbf{0}$. Let $\mathbf{a} = (\epsilon, \ldots, \epsilon)^t$. As $\bar{g}$ is connected and f strictly increasing, $f_{\bar{g},i}(\mathbf{a}) = f(n_{i,\bar{g}}\epsilon) \geq f(n_{min}\epsilon)$. Consequently, if there is some ϵ with $f(n_{min}\epsilon) > \epsilon$ then $\mathbf{a}$ exists. Consider the Taylor expansion of f around 0 evaluated at $n_{min}\epsilon$

$$f(n_{min}\epsilon) = f(0) + f'(0)n_{min}\epsilon + \frac{f''(0)}{2!}(n_{min}\epsilon)^2 + \ldots.$$

As $f(0) = 0$ and higher order terms vanish for small ϵ, $f'(0) > \frac{1}{n_{min}}$ implies $f(n_{min}\epsilon) - \epsilon = (f'(0)n_{min} - 1)\epsilon + \frac{f''(0)}{2!}n_{min}^2\epsilon^2 + \cdots > 0$ for ϵ sufficiently small.

Second, we show that $\lim_{x\to\infty} f'(x) < \frac{1}{n_{max}}$ implies existence of $\mathbf{b} = (b, \ldots, b)^t > \mathbf{a}$ with $f_{\bar{g}}(\mathbf{b}) < \mathbf{b}$. Consider the function $g(x) = f(n_{max}x) - x$. Note that $g(x)$ is eventually decreasing as

$$\lim_{x\to\infty} g'(x) = \lim_{x\to\infty} n_{max}f'(n_{max}x) - 1 < 0.$$

As $g(x)$ is also concave, there exists some x_0 with $g(x) < 0$ for all $x \geq x_0$ or, equivalently, b with $f(n_{max}b) < b$. Then $f_{\bar{g}}(\mathbf{b}) < \mathbf{b}$ as f strictly increases and each $i \in N$ has at most n_{max} neighbours in $\bar{g}$.

Third, as $f_{\bar{g}}$ is non-decreasing, $\mathbf{a} < f_{\bar{g}}(\mathbf{x}) < \mathbf{b}$ for $\mathbf{x} \in (\mathbf{a}, \mathbf{b})$. Therefore, $f_{\bar{g}} : [\mathbf{a}, \mathbf{b}] \to [\mathbf{a}, \mathbf{b}]$, Tarski's fixed point theorem (Tarski (1955)) is applicable and $f_{\bar{g}}$ has at least one fixed point $\mathbf{x}^* \in [\mathbf{a}, \mathbf{b}]$. Recall that we already know from part two that there is at most one fixed point $\mathbf{x}^* > \mathbf{0}$ so $\mathbf{x}^*$ is unique.

Part (ii) (b): First, assume by contradiction $f'(0) \leq \frac{1}{n_{max}}$ and the positive fixed point $\mathbf{x}^* > 0$ exists. Then for all $i \in N$

$$x_i^* = f_{\bar{g},i}(\mathbf{x}^*) < f'(0)(\sum_{j\in N_{i,\bar{g}}} x_j^*) \leq \frac{1}{n_{max}}(\sum_{j\in N_{i,\bar{g}}} x_j^*),$$

where the second inequality follows as $f(x) < f(0) + f'(0)x$ for $x > 0$ by strict concavity. Summing both sides over all $i \in N$ implies

$$\sum_{j \in N} x_j^* < \frac{1}{n_{max}} \sum_{j \in N} n_j x_j^* \leq \sum_{j \in N} x_j^*,$$

a contradiction.

Second, assume by contradiction $\lim_{x \to \infty} f'(x) \geq \frac{1}{n_{min}}$ and the positive fixed point $\mathbf{x}^* > 0$ exists. Then for all $i \in N$

$$x_i^* = f_{\bar{g},i}(\mathbf{x}^*) > \lim_{x \to \infty} f'(x)(\sum_{j \in N_{i,\bar{g}}} x_j^*) \geq \frac{1}{n_{min}}(\sum_{j \in N_{i,\bar{g}}} x_j^*),$$

where the second inequality follows as $f(x) > f(0) + \lim_{x \to \infty} f'(x)x$ for $x > 0$ by strict concavity. Summing both sides over all $i \in N$ leads to a similar contradiction as above.

Part (iii): Assume by contradiction that there is some equilibrium $\mathbf{x}^* = f_{\bar{g}}(\mathbf{x}^*)$ such that $\exists\, i, j$ with $x_i^* > 0$ but $x_j^* = 0$. Fix i with $x_i^* > 0$ and pick some $l \in N_{i,\bar{g}}$ arbitrarily. As $\mathbf{x}^*$ is an equilibrium, $x_l^* = f(\sum_{j \in N_{l,\bar{g}}} x_j^*)$. From $f(\sum_{j \in N_{l,\bar{g}}} x_j^*) \geq f(x_i^*) > 0$ we conclude $x_l^* > 0$. By repeating the argument and as $\bar{g}$ is connected by assumption, $x_w^* > 0$ for all $w \in N$, a contradiction. □

Proof of ***Remark 1****.* Assumption 1 states that there is a unique best response $x_i^*(\mathbf{x}_{-i}, \bar{g})$ which is an increasing, concave function of i's cumulative neighbour activity. Thus x_i^* maximises π_i and $\frac{\partial x_i^*}{\partial x_j} \geq 0$.

The necessary condition for a maximum implies $\frac{\partial \pi_i}{\partial x_i}\big|_{x_i = x_i^*(\mathbf{x}_{-i}, \bar{g})} = 0$. Using the implicit function theorem and rewriting we get

$$\frac{\partial^2 \pi_i}{\partial x_i \partial x_j}\Big|_{x_i = x_i^*(\mathbf{x}_{-i}, \bar{g})} = -\frac{\partial^2 \pi_i}{\partial x_i \partial x_i}\Big|_{x_i = x_i^*(\mathbf{x}_{-i}, \bar{g})} \frac{\partial x_i^*}{\partial x_j}.$$

As x_i^* is a maximum $\frac{\partial^2 \pi_i}{\partial x_i \partial x_i}\big|_{x_i = x_i^*(\mathbf{x}_{-i}, \bar{g})} < 0$ and the claim follows. □

Proof of ***Lemma 1.*** Condition one is simply the optimal activity condition from Equation (5).

Conditions two to four are clearly implied by the linking condition in Equation (6). Conversely, they jointly imply Equation (6) for strictly concave h as violation of condition (6) implies a violation of either condition two, three or four. To see that, assume some linking vector $g_i' \in \mathfrak{g}_i$ is a profitable deviation for player i in some candidate equilibrium.

If player i's deviation g_i' strictly increases the set of players to whom i supports links, then only adding a link to the most active player, $\arg\max_{j:\,\bar{g}^*_{ij}=0}\{x^*_j\}$, must also be a profitable deviation by concavity of h.

If player i's deviation g_i' strictly decreases the set of players to whom i supports links, then only deleting the link to the least active player, $\arg\min_{j:\,g^*_{ij}=1}\{x^*_j\}$, must also be a profitable deviation by concavity of h.

Finally, consider a profitable deviation g_i' that demands to delete some links currently supported and form some new ones instead. Then either condition four is violated or the activity of all players to whom i forms a new link is strictly lower than the activity of all players to whom i deletes his link. As g_i' is by definition a profitable deviation, there is in such a case another (strictly better) deviation g_i'': if the net number of added links under g_i' is positive, g_i'' demands to form some new links without deleting any old links. If the net number of added links is negative, g_i'' demands to delete some links without adding any. Both these cases are discussed above. □

Proof of ***Proposition 2.*** To show $\mathcal{E} \subseteq \{(\mathbf{x}^*_{\bar{g}}, g) : g \in G\}$, consider some arbitrary equilibrium $(\mathbf{x}^*, g) \in \mathcal{E}$ and assume by contradiction that $\mathbf{x}^* \neq \mathbf{x}^*_{\bar{g}}$. By the first condition of Lemma 1 and the definition of $\mathbf{x}^*_{\bar{g}}$, this implies that there is some player with $x^*_i = 0$ and $n_{i,\bar{g}} \geq 1$. As $x^*_i = f(\sum_{j \in N_{i,\bar{g}}} x^*_j)$ and $x^*_j \geq 0$, every player in i's component has zero activity. As $n_{i,\bar{g}} \geq 1$ there is at least one link in i's component. As $k > 0$, deleting this link then strictly increases the sponsoring player's net payoff but leaves his gross payoff unaltered, a contradiction.

By the reasoning above, there can arise three different cases: either all players are isolated, i.e. $n_{i,\bar{g}} = 0$ for all $i \in N$, and choose $x_i^* = 0$ (case 1) or no player is isolated, i.e. $n_{i,\bar{g}} \geq 1$ for all $i \in N$, and choose $x_i^* > 0$ (case 2) or there are some isolated players and some non-isolated players (case 3).

We now briefly show that case 3 only exists in non-generic equilibria. Define $\tilde{x} = h^{-1}(k)$. As some players sponsor links $\max x_i^* \geq \tilde{x}$ because no link is sponsored to any player with $x_i^* < \tilde{x}$ by Lemma 1 and concavity of h. Conversely, as some players are isolated $\max x_i^* \leq \tilde{x}$ because isolated players strictly gain from linking to a player with $x_i^* > \tilde{x}$ (and adjusting their activity). Thus $\max x_i^* = \tilde{x}$ and links are only formed to players with $x_i^* = \tilde{x}$. Fix the interaction network $\bar{g}$ and thus the equilibrium activity. If k decreases by any small ϵ, isolated players strictly prefer to form a link to a maximum activity player (and adjust their activity). Conversely, if k increases by ϵ, players who sponsor a link prefer to delete their link(s) and become inactive. □

Proof of ***Proposition 3***. The zero activity equilibrium with an empty network exists for any linking costs as all conditions in Lemma 1 are met: it is optimal for an isolated player to choose zero activity and forming a link to any non-active player is costly as $k > 0$.

To prove the remaining claims, let $\mathbb{Q} \subset \mathcal{E}$ be the set of all positive activity equilibria existing under some linking costs k. As $\mathcal{E} \subseteq \{(\mathbf{x}_{\bar{g}}^*, g) : g \in G\}$ by Proposition 2, G is finite for fixed n and there is a unique $\mathbf{x}_{\bar{g}}^*$ for any $g \in G$, we conclude $\mathbb{Q}$ is finite.

For any element of $\mathbb{Q}$ we can find the maximal linking costs $\bar{k}_g$ such that condition 3 of Lemma 1 is fulfilled for every edge in g^*, i.e. nobody wants to delete any links. This number is finite and positive as each player has finite positive activity and the graph is not empty. As $\mathbb{Q}$ is a finite set, we can find the desired threshold $\bar{\mathbf{k}}_n = \max\{\bar{k}_g\}_{(\mathbf{x}_{\bar{g}^*}^*, g^*) \in \mathbb{Q}}$. Also note that $\bar{\mathbf{k}}_n < \bar{\mathbf{k}}_{max} = \lim_{x \to \infty} h(f(x))$, as this is the supremum utility increase due to any link.

Likewise, for any element of $\mathbb{Q}$ in which $\bar{g}^* \neq \bar{g}_{com}$, there is a positive finite number $\underline{k}_g > 0$, which is defined as the minimal linking costs such that condition 2 of Lemma 1 is fulfilled for every edge in g, i.e. nobody wants to add any links. Note that no links can be added in the complete network $\bar{g}_{com}$. As $\mathbb{Q}$ is finite we can find $\underline{\mathbf{k}}_n = \min\{\underline{k}_g\}_{(\mathbf{x}^*_{\bar{g}^*}, g^*) \in \mathbb{Q}:\, \bar{g}^* \neq \bar{g}_{com}}$, that is the unique positive minimal linking cost such that a positive activity equilibrium with a non-complete network exists. □

Proof of ***Proposition 4***. We first show the "only if" claim of part (i), then prove part (ii) and finally the "if" claim of part (i).

"Only if" claim of part (i): Assume that $(\mathbf{x}^*_{\bar{g}}, g)$ is a single-level equilibrium and let $x^* > 0$ be the common (positive) level of activity. Condition one for optimal activity from Lemma 1 tells us that $x^*_i = x^* = f(n_{i,\bar{g}} x^*)$ for any player i in equilibrium. As f is strictly increasing, every player must then have a common number of links $n_{i,\bar{g}} = \Delta \in \{1, 2, \ldots, n-1\}$—i.e. $\bar{g} \in \bar{G}_{reg}$.

Part (ii): Let $\bar{g} \in \bar{G}^{\Delta}_{reg}$ be a Δ-regular interaction network. We want to determine the corresponding positive activity equilibrium vector $\mathbf{x}^*_{\bar{g}}$. As f is concave and by Assumption 3, there is a unique $x_\Delta > 0$ that solves $x_\Delta = f(\Delta x_\Delta)$ for any $\Delta \in \{1, \ldots, n-1\}$. Thus $\mathbf{x}_\Delta = (x_\Delta, \ldots, x_\Delta)^t$ is a positive activity equilibrium vector and from uniqueness in Proposition 1 $\mathbf{x}_\Delta = \mathbf{x}^*_{\bar{g}}$.

For the comparative statics results, first note that the definition $x_\Delta = f(\Delta x_\Delta)$ is independent of the size of the society. Second, assuming that $\Delta \geq 1$ is continuous and differentiating both sides gives

$$\frac{\partial x_\Delta}{\partial \Delta} = \frac{x_\Delta}{\frac{1}{f'(\Delta x_\Delta)} - \Delta} > 0$$

as $x_\Delta = f(\Delta x_\Delta) > f(0) + \Delta x_\Delta f'(\Delta x_\Delta) = \Delta x_\Delta f'(\Delta x_\Delta)$.

"If" claim of part (i): Assume that g_Δ is an arbitrary *directed* graph with $\bar{g}_\Delta \in \bar{G}^{\Delta}_{reg}$. As $\mathbf{x}_\Delta = \mathbf{x}^*_{\bar{g}}$ from the reasoning above, it suffices to show $(\mathbf{x}_\Delta, g_\Delta) \in \mathcal{E}_+$, i.e. that equilibrium conditions 2 to 4 from Lemma 1 are met for some non-degenerate interval $[\underline{k}_\Delta, \bar{k}_\Delta]$. Condition two and three translate into the upper bound of linking costs $\bar{k}_\Delta$ and the lower bound $\underline{k}_\Delta$as shown below

$$\begin{aligned}
\underline{k}_\Delta &= \begin{cases} h((\Delta+1)x_\Delta) - h(\Delta x_\Delta) & \Delta < n-1 \\ 0 & \Delta = n-1 \end{cases} \\
\bar{k}_\Delta &= h(\Delta x_\Delta) - h((\Delta-1)x_\Delta).
\end{aligned}$$

Condition four is automatically met. The claim follows as $\underline{k}_\Delta < \bar{k}_\Delta$ since the function $h(\Delta a) - h((\Delta-1)a)$ is strictly decreasing in Δ for any $a > 0$ by concavity of h. □

Proof of ***Lemma 2****.* First note that g' differs from g^* only insofar as links between two players with unequal activity are now sponsored by the lower activity player; in particular $\bar{g}' = \bar{g}^*$ and $N_{i,\bar{g}'} = N_{i,\bar{g}^*}$. Therefore, $\mathbf{x}^*$ is an equilibrium activity vector on $\bar{g}'$—that is $(\mathbf{x}^*, g')$ meets the first equilibrium condition of Lemma 1—and the cumulative neighbour activity accessed by any player is unaltered, that is $\sum_{j \in N_{i,\bar{g}'}} x_j^* = \sum_{j \in N_{i,\bar{g}^*}} x_j^*$.

As $\bar{g}' = \bar{g}^*$ and every player has unaltered activity x_i^* in g', no player wants to form any additional link(s) in g' under any linking costs k under which he refrains from adding a link in g^*; hence, in particular, $(\mathbf{x}^*, g')$ meets the second equilibrium condition from Lemma 1.

Consider any link which is sponsored by the same player in g' as in g^*, i.e. $g'_{ij} = g^*_{ij} = 1$. For the same reason as above, no player wants to delete (or switch) such a link for any linking costs k under which he refrains from deleting (or switching) it in g^*. Thus equilibrium conditions three and four from Lemma 1 are met for these links.

Instead, consider some players m and l who share a link with altered sponsorship, i.e. $g'_{lm} = g^*_{ml} = 1$. As the link between l and m is sponsored by l in g' but by m in g^*, $x_l^* < x_m^*$ by the assumptions of the lemma. To finalise the proof, we are going to show *(i)* that l refrains from deleting his sponsored link to m in g' for a greater interval of linking costs k than the interval under which m refrains from deleting his sponsored link to l in g^* (equilibrium condition three). Subsequently, we are going to show *(ii)* that l does not want to switch his link (equilibrium condition four).

(i) As $(\mathbf{x}^*, g^*) \in \mathcal{E}$ and $g^*_{ml} = 1$, equilibrium condition three of Lemma 1 implies

$$h(\sum_{j \in N_m} x^*_j) - h(\sum_{j \in N_m \setminus \{l\}} x^*_j) \geq k.$$

Thus l refrains from deleting his sponsored link to m in g' for a greater interval of linking costs k if

$$h(\sum_{j \in N_l} x^*_j) - h(\sum_{j \in N_l \setminus \{m\}} x^*_j) > h(\sum_{j \in N_m} x^*_j) - h(\sum_{j \in N_m \setminus \{l\}} x^*_j)$$

Since activity $x^*_l < x^*_m$ and as f is strictly increasing, condition one of Lemma 1 implies that cumulative neighbourhood activity of l is smaller than m's, $\sum_{j \in N_l} x^*_j < \sum_{j \in N_m} x^*_j$ under both g^* and g'. Thus $x^*_l < x^*_m$ implies $\sum_{j \in N_l \setminus \{m\}} x^*_j < \sum_{j \in N_m \setminus \{l\}} x^*_j$. As h is strictly increasing and concave, and from the findings above

$$\begin{aligned} & h(\sum_{j \in N_l} x^*_j) - h(\sum_{j \in N_l \setminus \{m\}} x^*_j) \\ > \; & h(\sum_{j \in N_l \setminus \{m\}} x^*_j + x^*_l) - h(\sum_{j \in N_l \setminus \{m\}} x^*_j) \\ > \; & h(\sum_{j \in N_m \setminus \{l\}} x^*_j + x^*_l) - h(\sum_{j \in N_m \setminus \{l\}} x^*_j). \end{aligned}$$

(ii) Finally, let us turn to condition four. If this condition does not hold for $(\mathbf{x}^*, g')$ then there must be some player $o \in N$ with $x^*_o > x^*_m$ and $\bar{g}'_{lo} = \bar{g}^*_{lo} = 0$ so that l prefers to form a link to o instead of m. As shown below, this leads to a contradiction as $(\mathbf{x}^*, g^*)$ would not be an equilibrium if such a player o exists.

Assume $(\mathbf{x}^*, g^*) \in \mathcal{E}$ but, by contradiction, player o as described above exists. On the one side, since player m supports a link to l in g^* and $x^*_o > x^*_l$

we infer

$$
\begin{aligned}
k &\leq h(\sum_{j \in N_m \setminus \{l\}} x_j^* + x_l^*) - h(\sum_{j \in N_m \setminus \{l\}} x_j^*) \\
&< h(\sum_{j \in N_m \setminus \{l\}} x_j^* + x_o^*) - h(\sum_{j \in N_m \setminus \{l\}} x_j^*).
\end{aligned}
$$

Since player l does not form a link to o in g^* (and o is currently no neighbour of l) we also know

$$k \geq h(\sum_{j \in N_l} x_j^* + x_o^*) - h(\sum_{j \in N_l} x_j^*).$$

As h is strictly concave, these two inequalities imply $\sum_{j \in N_l} x_j^* > \sum_{j \in N_m \setminus \{l\}} x_j^*$ or

$$x_l^* > \sum_{j \in N_m} x_j^* - \sum_{j \in N_l} x_j^*. \tag{7}$$

On the other side, we also know that player l only supports links to players with activity $x_i^* \geq x_o^*$ in g^* as he does not support a link to o. Conversely, as m supports a link to l in g^*, he must have a link to all player i with $x_i^* > x_l^*$—i.e. in particular to all players with $x_i^* \geq x_o^*$. Thus every player to whom l forms a link is a neighbour of m. Furthermore, every player that forms a link to l must have a link to all players with $x_i^* > x_l^*$—i.e. in particular to m.

These two findings and the fact that o is a neighbour of m but not of l imply

$$\sum_{j \in N_m} x_j^* - \sum_{j \in N_l} x_j^* \geq x_o^* + x_l^* - x_m^* > x_l^*,$$

in contradiction to Equation (7). □

Proof of **Proposition** ***5***. ***Part (i):*** Assume $g_{ij} = 1$ in equilibrium and, by contradiction, that there exists some $j' \in N$ with $x_{j'}^* > x_j^*$ but $g_{ij'} = 0$. As i sponsors a link to j and there is upwards linking, $x_j^* \geq x_i^*$. This implies $x_{j'}^* > x_i^*$ so that $g_{j'i} = 0$ i.e. j' does not sponsor a link to i. But then i

can increase his payoff by deleting the link to j and linking to j' instead, a contradiction.

Part (ii): Assume by contradiction $x_i^* > x_j^*$ but $\eta_i > \eta_j$ in some $(\mathbf{x}^*, g^*) \in \mathcal{E}_+$ with upward linking. Let $\underline{x} = \min\{x_l : g_{il} = 1\}$, i.e. minimum level of activity accessed by i through a self-sponsored link, and let $\kappa_i = |\{l \in N : x_l = \underline{x} \wedge g_{il} = 1\}|$ be the number of players with activity $\underline{x}$ to whom i sponsors a link. Note that $\underline{x} \geq x_i^*$ as there is upwards linking so that $\underline{x} > x_j^*$ and no player with activity $\underline{x}$ or higher sponsors a link to j. The reasoning above from the proof of the first statement—i.e. each player preferring to link to higher activity players—and $\eta_j < \eta_i$ implies $\kappa_j < \kappa_i$ so that j sponsors less links to players with activity $\underline{x}$.

On the one hand, as i sponsors a link to somebody with activity $\underline{x}$ in equilibrium and j refrains from adding a link to one more person with activity $\underline{x}$ (which would be possible as $\kappa_j < \kappa_i$ and players with activity $\underline{x}$ do not sponsor links to j), Lemma 1 tells us:

$$h(\sum_{l \in N_i} x_l^*) - h(\sum_{l \in N_i} x_l^* - \underline{x}) \geq k \geq h(\sum_{l \in N_j} x_l^* + \underline{x}) - h(\sum_{l \in N_j} x_l^*). \tag{8}$$

As h is strictly concave, this can only hold true if

$$\sum_{l \in N_i} x_l^* \leq \sum_{l \in N_j} x_l^* + \underline{x}. \tag{9}$$

On the other hand, as $x_i^* > x_j^*$, any player l who sponsors a link to j has to sponsor a link to i as well by the first statement of this lemma so that i receives more cumulative neighbour activity than j through incoming links. And as $\kappa_j < \kappa_i$, i also receives at least $(\kappa_i - \kappa_j)\underline{x}$ more cumulative neighbour activity than j through self-sponsored links. Together this implies

$$\sum_{l \in N_i} x_l^* \geq \sum_{l \in N_j} x_l^* + (\kappa_i - \kappa_j)\underline{x}. \tag{10}$$

Combining Equations (9) and (10), we get

$$\sum_{l\in N_j} x_l^* + (\kappa_i - \kappa_j)\underline{x} \leq \sum_{l\in N_i} x_l^* \leq \sum_{l\in N_j} x_l^* + \underline{x},$$

which can only hold true if $\kappa_i - \kappa_j = 1$ and $\sum_{l\in N_i} x_l^* = \sum_{l\in N_j} x_l^* + \underline{x}$. The latter finding together with Equation (8) implies

$$h(\sum_{l\in N_i} x_l^*) - h(\sum_{l\in N_i} x_l^* - \underline{x}) = k = h(\sum_{l\in N_j} x_l^* + \underline{x}) - h(\sum_{l\in N_j} x_l^*).$$

Then for any small $\epsilon > 0$, under linking costs $k' = k + \epsilon$, player i would prefer to delete a link and for $k' = k - \epsilon$, player j would prefer to add a link. Thus the equilibrium is non-generic, a contradiction. Thus $x_i^* > x_j^*$ implies $\eta_i \leq \eta_j$.

We now show $x_i^* > x_j^*$ also implies $n_i > n_j$. Let N_j^{in} be the set of players who sponsor links to j. By part (i) and $x_i^* > x_j^*$, everybody who sponsors a link to j sponsors a link to i so that $N_j^{in} \subseteq N_i^{in}$ is also the set of players who sponsor links to i as well as to j. Let $A = N_j^{in} \cup \{i, j\}$. Then $N_j \setminus A$ is the set of players to whom j sponsors links apart from i with $|N_j \setminus A| = \eta_j - \bar{g}_{ij}$. Similarly, $N_i \setminus A$ is the set of i's neighbours apart from j who do not sponsor links to i as well as j. As we delete in both cases the same *number* of neighbours we have

$$|N_i \setminus A| = n_i - (n_j - |N_j \setminus A|) = n_i - n_j + \eta_j - \bar{g}_{ij}.$$

From optimal activity, $x_i^* > x_j^*$ implies

$$\sum_{l\in N_i\setminus A} x_l > \sum_{l\in N_j\setminus A} x_l. \tag{11}$$

Consider the set $N \setminus A$ and relabel players such that $\tilde{x}_1 \geq \cdots \geq \tilde{x}_{|N\setminus A|}$. By part (i), j sponsors links to some set of most active players that is

$$\sum_{l\in N_j\setminus A} x_l = \sum_{l=1}^{\eta_j - \bar{g}_{ij}} \tilde{x}_l,$$

whereas the remaining neighbour activity that i accesses is bounded from above by

$$\sum_{l \in N_i \setminus A} x_l \leq \sum_{l=1}^{n_i - n_j + \eta_j - \bar{g}_{ij}} \tilde{x}_l .$$

Together with Equation (11), $n_i > n_j$ follows.

"If" claim of part (iii): As there is upward linking and every player prefers to link to more active players, $\eta_i = \eta_j$ implies $\sum_{l:\, g_{il}=1} x_l^* = \sum_{l:\, g_{jl}=1} x_l^*$. $x_i^* = x_j^*$ implies $\sum_{l \in N_i} x_l^* = \sum_{l \in N_j} x_l^*$ from optimal activity. Thus the claim holds.

"Only if" claim of part (iii): Assume $\sum_{l:\, g_{li}=1} x_l^* = \sum_{l:\, g_{lj}=1} x_l^*$ and by contradiction $x_i^* > x_j^*$. The second part of the lemma then implies $\eta_i \leq \eta_j$. As i only forms links to players with $x_{j'}^* \geq x_i^* > x_j^*$ due to upward linking, players j' to whom i sponsors links do not sponsor a link to j. Thus j could copy i's linking decisions and strictly increase his payoff, a contradiction.

$\sum_{l:\, g_{li}=1} x_l^* = \sum_{l:\, g_{lj}=1} x_l^*$ and $x_i^* = x_j^*$ together imply $\sum_{l:\, g_{il}=1} x_l^* = \sum_{l:\, g_{jl}=1} x_l^*$. As there is upward linking and by part one $\eta_i = \eta_j$ follows. □

Proof of ***Proposition 6****.* ***Part (i):*** Consider any generic equilibrium with at least three distinct levels of positive activity. From Proposition 2, we know there are no isolated players in such an equilibrium and, applying Lemma 2, we can assume that no link is sponsored by an adjacent player with strictly higher activity.

First note that any least activity player who has no incoming links has to sponsor at least one link to a highest activity player. We distinguish two cases.

If there is some least activity player who sponsors a link to another least activity player then such a player sponsors by Proposition 5 a link to all higher activity players. Thus $\bar{g}^*$ is connected.

Assume on the contrary that there are no links between least activity players. As there is upward linking and they do not link among themselves, they have no incoming links and by Proposition 5 sponsor a common number of links $\tilde{\eta} \geq 1$. Any player with the second lowest level of activity has at least one incoming link as he would otherwise, by Proposition 5 again, have similar activity as lowest activity players. Furthermore, any player who sponsors a link to a second lowest activity player, sponsors links to all higher activity players. As there are at least three levels of activity, the network is connected.

Part (ii): Consider any minimally connected component of $\bar{g}^*$. If the longest path in the component is two or three, then the component is a star. If it is four (and as it is minimally connected), the component is a core-periphery network—as introduced in Section 4.2.3—with two core players and with periphery players who only link to one of the two core players. Proposition 8 from that section and Proposition 14 from Appendix B.5 imply that the line with four players is the only network with these properties that exists in equilibrium.

Thus consider a minimally connected component with a longest path of at least five. Consider any player i with $n_i = 1$ and let j be the single player he is linked to. As the component consists of at least five players $n_j \geq 2$, i.e. j has another neighbour. Assume by contradiction $x_i^* \geq x_j^*$. Then optimal activity and strictly increasing f imply

$$\begin{aligned} f(x_j^*) &\geq f(x_i^* + \sum_{j' \in N_j \setminus \{i\}} x_{j'}^*) \\ \Rightarrow x_j^* &> x_i^*, \end{aligned}$$

a contradiction. Applying Lemma 2, we can assume that all players with $n_i = 1$ sponsor their link. Then optimal linking implies that they sponsor their single link to some player with highest activity $\bar{x}$ (in the entire network) and, thus, all players with $n_i = 1$ have common activity $\underline{x} < \bar{x}$.

Consider some player with $n_i = 1$ from whom a path of (at least) five players extends. We call these players $a, \ldots, e$. From the reasoning above, $x_a^* = \underline{x}$ and $x_b^* = \bar{x}$. Consider player c. As he is linked to b as well as to d his cumulative neighbour activity is higher than a's so that $x_c^* > \underline{x}$.

Assume that $x_c^* < \bar{x}$. Then d cannot sponsor a link to c as he would strictly improve his payoff by switching the link to b (whom he is currently not linked to as the component is minimally connected). But then (by upwards linking) c sponsors links to b and d so that $\eta_c \geq 2$, $\eta_a = 1$ but $x_c^* > x_a^*$, a contradiction to Proposition 5. As $\bar{x}$ is the maximal level of activity, we conclude $x_c^* = \bar{x}$.

Similar reasoning shows that $x_d^* = \bar{x}$. As $x_{b,c,d}^* = \bar{x}$ a stronger version of Lemma 2 holds:[38] if any such equilibrium exist, then there is an equilibrium with any arbitrary direction of link sponsorship between b and c or c and d. In particular, there is an equilibrium in which c sponsors both links, a similar contradiction to Proposition 5 arises as above. □

Proof of ***Proposition 7***. We prove part (ii) first and then turn to part (i).

Part (ii): Fix any $\bar{g}^* \in \bar{G}_{bp}^{\Delta_{1,2}}$ arbitrarily. We conjecture that all players within partite set j choose symmetric positive activity x_j in equilibrium. Under the conjecture the best response activity of all players in P_1 implies $x_i^* = x_1 = f(\Delta_1 x_2)$ and the best response of players in P_2 implies $x_i^* = x_2 = f(\Delta_2 x_1)$. The arguments presented in Proposition 1 and Assumption 3 guarantee that there is a positive solution $x_{1,2}$ that solves both equations simultaneously. That is we have found a positive activity equilibrium vector $\mathbf{x}^*$ for $\bar{g}^*$. As there is a unique positive activity vector by Proposition 1 for every component of $\bar{g}^*$ and $\mathbf{x}^* > 0$ in any generic equilibrium with non-empty interaction network by Proposition 2, the conjecture is confirmed.

Assume by contradiction $x_1 \geq x_2$. As f is strictly increasing, optimal activity then implies $\Delta_1 x_2 \geq \Delta_2 x_1$. Since $\Delta_1 < \Delta_2$ it follows $x_2 > x_1$, a contradiction.

By $x_1 < x_2$, optimal activity then implies $\Delta_1 x_2 < \Delta_2 x_1$. Finally, $\Delta_2|P_2| = \Delta_1|P_1|$ as the total number of links in both independent sets has to be equal.

Part (i): Pick any $\hat{g} \in \bar{G}_{bp}^{\Delta_{1,2}}$ arbitrarily. We need to show that there is some g^* with $\bar{g}^* = \hat{g}$ and $\mathbf{x}^*$ such that $(\mathbf{x}^*, g^*) \in \mathcal{E}_+$.

[38]Note that as these players have maximal activity, equilibrium condition four is automatically met when link sponsoring between any two players is switched. Then the stronger version follows immediately from the proof of the lemma.

From *part (ii), we* know that players in P_1 choose symmetric activity x_1 whereas players in P_2 choose symmetric activity $x_2 > x_1$ in any $(\mathbf{x}^*, g^*) \in \mathcal{E}_+$ with $\bar{g}^* = \hat{g}$. Fix some g^* with $\bar{g}^* = \hat{g}$ such that all links are sponsored by players in P_1. To prove the statement we need to show that equilibrium conditions 2 to 4 from Lemma 1 hold for $k \in [\underline{k}, \bar{k}] \neq \emptyset$ given the equilibrium activity pattern described above.

It is obvious that equilibrium condition 4 is met as $x_1 < x_2$, all links are sponsored by players in P_1 and are directed towards players in P_2. As all links are sponsored by players in P_1, equilibrium condition 3 of the lemma preventing link deletion writes as

$$h(\Delta_1 x_2) - h(\Delta_1 x_2 - x_2) \geq k. \tag{12}$$

Equilibrium condition 2 of the lemma preventing link adding can be split into three conditions, namely that players in P_2 do not gain from adding a link to another player in P_2 (which they strictly prefer to adding a link to some player in P_1) and that players in P_1 do not gain from adding a link to either a player in P_1 or—if possible—in P_2. However, as h is strictly increasing and concave and $\Delta_1 x_2 < \Delta_2 x_1$, the following two inequalities hold

$$\begin{aligned} h(\Delta_2 x_1 + x_2) - h(\Delta_2 x_1) &< h(\Delta_1 x_2 + x_2) - h(\Delta_1 x_2) \\ h(\Delta_1 x_2 + x_1) - h(\Delta_1 x_2) &< h(\Delta_1 x_2 + x_2) - h(\Delta_1 x_2), \end{aligned}$$

which shows that players in P_1 have the strongest incentive to add a link if there is some (hypothetical) suitable player in P_2 available. Therefore, the condition below preventing players in P_1 from adding a link to a player in P_2, is sufficient for equilibrium condition 2:

$$h(\Delta_1 x_2 + x_2) - h(\Delta_1 x_2) \leq k. \tag{13}$$

As h is strictly concave, we know that

$$h(\Delta_1 x_2 + x_2) - h(\Delta_1 x_2) < h(\Delta_1 x_2) - h(\Delta_1 x_2 - x_2)$$

so that Equations (12) and (13) are simultaneously met for $k \in [\underline{k}, \bar{k}] \neq \emptyset$ which finalises the proof. $\square$

*Proof to **Proposition 8**.* We are going to prove *part (i)* first, then *part (iii)* and finally turn to *part (ii)*.

Part (i): We are going to show that $(\mathbf{x}^*, g^*) \in \mathcal{E}_+$ with $\bar{g}^* \in \bar{G}_{cp} \backslash \bar{G}_{cp}^{com} = \bar{G}_{cp}^{\neg com}$ and $\min\{|\mathcal{C}|, |\mathcal{P}|\} > 1$ implies $|\mathcal{P}| = 2$.

Throughout the proof, we denote by $\underline{\mathcal{C}} = \{c \in \mathcal{C} : x_c^* \leq x_{c'}^*, \forall c' \in \mathcal{C}\}$ the set of least active core players with typical element $\underline{c} \in \underline{\mathcal{C}}$ and $\bar{\mathcal{P}} = \{p \in \mathcal{P} : x_p^* \geq x_{p'}^*, \forall p' \in \mathcal{P}\}$ the set of most active periphery players with typical element $\bar{p} \in \bar{\mathcal{P}}$. The following claim holds:

Claim. If $(\mathbf{x}^*, g^*) \in \mathcal{E}_+$ with $\bar{g}^* \in \bar{G}_{cp}^{\neg com}$ then $x_{\bar{p}} < x_{\underline{c}}$.

Proof. Assume by contradiction that $x_{\bar{p}} \geq x_{\underline{c}}$ and $\bar{p}$ and $\underline{c}$ do not share a link. As $\underline{c}$ is linked to some (other) periphery player $\tilde{p}$ by assumption, we get

$$f(\sum_{c \in \mathcal{C}} x_c - x_{\underline{c}}) \geq x_{\bar{p}} \geq x_{\underline{c}} \geq f(\sum_{c \in \mathcal{C}} x_c - x_{\underline{c}} + x_{\tilde{p}}),$$

which implies $x_{\tilde{p}} \leq 0$, a contradiction.

Assume instead $\bar{p}$ and $\underline{c}$ share a link. On the one hand, if $x_{\bar{p}} > x_{\underline{c}}$ then

$$f(\sum_{c \in \mathcal{C}} x_c) \geq x_{\bar{p}} > x_{\underline{c}} \geq f(\sum_{c \in \mathcal{C}} x_c - x_{\underline{c}} + x_{\bar{p}}),$$

which implies $x_{\underline{c}} > x_{\bar{p}}$, a contradiction. On the other hand, $x_{\underline{c}} = x_{\bar{p}}$ can only hold true if $\bar{p}$ is the only periphery player linked to $\underline{c}$ and is himself linked to all players in $\mathcal{C}$. As $\bar{p}$ and $\underline{c}$ are linked to the same set of people we can assume without any loss of generality that $g_{\bar{p}\underline{c}}^* = 1$, i.e. $\bar{p}$ sponsors the link. As $N_{\bar{p}} = \mathcal{C}$ and $\bar{p}$ sponsors a link to $\underline{c}$ with lowest core activity $x_{\underline{c}}^*$ any other player in the periphery $p' \in \mathcal{P}$ has to have $N_{p'} = \mathcal{C}$ in any generic equilibrium. Thus, the core-periphery network is complete, $\bar{g}^* \in \bar{G}_{cp}^{com}$, a contradiction. $\square$

As $x_{\bar{p}} < x_{\underline{c}}$, we can apply Lemma 2 and assume that periphery players sponsor all their links. As periphery players do not link among themselves, Proposition 5 then implies that all periphery players sponsor a common number of links η_p in any generic equilibrium and have a common level of activity $x^*_{p'} = x^*_p$, $\forall\, p, p' \in \mathcal{P}$. As $\bar{g}^* \in \bar{G}^{\neg com}_{cp}$ we know that $\eta_p < |\mathcal{C}|$ (recall that $|\mathcal{C}| > 1$ by assumption). Then there cannot be a single least active core player as no periphery player would sponsor a link to him if a higher activity core player is available.

In conclusion, $1 < |\underline{\mathcal{C}}| \leq |\mathcal{C}|$ so that there are at least two players in $|\underline{\mathcal{C}}|$. In equilibrium, periphery players sponsor links to all core players with $x^*_c > x^*_{\underline{c}}$ and a *strict subset* of players in $\underline{\mathcal{C}}$, i.e. $\eta_p \in (|N^{\bar{c}}|, |\mathcal{C}|)$. Therefore, all players in $N^{\bar{c}} = \mathcal{C} \setminus \underline{\mathcal{C}}$ have the same common activity $x^*_{\bar{c}}$. As all players in $\underline{\mathcal{C}}$ have the same common activity $x^*_{\underline{c}}$, they access the same cumulative neighbour activity, i.e. they have links a common number of neighbours from the periphery $\kappa_{\underline{c}} \geq 1$.

As $|\underline{\mathcal{C}}| \geq 2$, there must be a player in $\underline{\mathcal{C}}$ sponsoring a link to another player in $\underline{\mathcal{C}}$ and thus the equilibrium conditions in Lemma 1 imply

$$h(\sum_{j \in \mathcal{C}} x^*_j - x^*_{\underline{c}} + \kappa_{\underline{c}} x^*_p) - h(\sum_{j \in \mathcal{C}} x^*_j - 2x^*_{\underline{c}} + \kappa_{\underline{c}} x^*_p) \geq k,$$

whereas any periphery player $p \in \mathcal{P}$ refrains from adding a link to another player with $x^*_{\underline{c}}$ which implies

$$h(\sum_{j \in \mathcal{C} \cap N_p} x^*_j + x^*_{\underline{c}}) - h(\sum_{j \in \mathcal{C} \cap N_p} x^*_j) \leq k.$$

Since h is strictly concave, both statements can only be met simultaneously if

$$\sum_{j \in \mathcal{C}} x^*_j - x^*_{\underline{c}} + \kappa_{\underline{c}} x^*_p \leq \sum_{j \in \mathcal{C} \cap N_p} x^*_j + x^*_{\underline{c}}. \tag{14}$$

As $\kappa_{\underline{c}} x^*_p > 0$ the inequality above implies $\sum_{j \in \mathcal{C}} x^*_j - \sum_{j \in \mathcal{C} \cap N_p} x^*_j < 2x^*_{\underline{c}}$ so that $\eta_p > |\mathcal{C}| - 2$. As $\eta_p < |\mathcal{C}|$ by assumption, every periphery player sponsors sponsors exactly $\eta_p = |\mathcal{C}| - 1$ links so that $\sum_{j \in \mathcal{C} \cap N_p} x^*_j + x^*_{\underline{c}} = \sum_{j \in \mathcal{C}} x^*_j$. Then

Equation (14) reads as $x_p^* \leq \frac{1}{\kappa_{\underline{c}}} x_{\underline{c}}^*$.

Assume by contradiction $\kappa_{\underline{c}} > 1$. Optimal activity in Lemma 1, concavity of f and $f(0) = 0$ imply

$$x_p^* = f(\sum_{j \in \mathcal{C}} x_j^* - x_{\underline{c}}^*) \leq \frac{1}{\kappa_{\underline{c}}} x_{\underline{c}}^* < f(\frac{1}{\kappa_{\underline{c}}}(\sum_{j \in \mathcal{C}} x_j^* - x_{\underline{c}}^* + \kappa_{\underline{c}} x_p^*)).$$

As f is strictly increasing (first step), $\sum_{j \in \mathcal{C}} x_j^* \geq |\mathcal{C}| x_{\underline{c}}^*$ (third step), and $\kappa_{\underline{c}} x_p^* \leq x_{\underline{c}}^*$ (fourth step), we can reformulate that condition

$$\begin{aligned} \sum_{j \in \mathcal{C}} x_j^* - x_{\underline{c}}^* &< \frac{1}{\kappa_{\underline{c}}}(\sum_{j \in \mathcal{C}} x_j^* - x_{\underline{c}}^*) + x_p^* \\ \Leftrightarrow (\sum_{j \in \mathcal{C}} x_j^* - x_{\underline{c}}^*)(1 - \frac{1}{\kappa_{\underline{c}}}) &< x_p^* \\ \Rightarrow (|\mathcal{C}| - 1) x_{\underline{c}}^* (1 - \frac{1}{\kappa_{\underline{c}}}) &< x_p^* \\ \Rightarrow (|\mathcal{C}| - 1)(\kappa_{\underline{c}} - 1) &< 1 \end{aligned}$$

which can only hold true for $\kappa_{\underline{c}} = 1$ as $|\mathcal{C}| > 1$ by assumption.

As $\kappa_{\underline{c}} = 1$, every player in $\underline{\mathcal{C}}$ shares a link with exactly one periphery player. As $\eta_p = |\mathcal{C}| - 1$ every periphery player shares a link with all but one player in $\underline{\mathcal{C}}$. Together with $|\underline{\mathcal{C}}| \geq 2$ and $|\mathcal{P}| \geq 2$, we conclude $|\underline{\mathcal{C}}| = |\mathcal{P}| = 2$.

Part (iii): Assume there is some equilibrium with a complete core-periphery interaction graph $\bar{g}^* \in \bar{G}_{cp}^{com}$. We want to determine the positive activity equilibrium vector, which exists by Assumption 3. As all periphery players are linked to (and only to) all core players, they choose common activity $x_i^* = x_p$.

Similarly, all core players choose common activity $x_i^* = x_c$: assume by contradiction $x_{c'}^* < x_{c''}^*$. Then optimal activity implies

$$f(\sum_{j \in N \setminus \{c', c''\}} x_j^* + x_{c''}^*) < f(\sum_{j \in N \setminus \{c', c''\}} x_j^* + x_{c'}^*),$$

which implies $x_{c''}^* < x_{c'}^*$, a contradiction.

These common equilibrium activity levels can be ranked: first, assume by contradiction that $x_c \leq x_p$. Then optimal activity and increasing f imply

$$(|\mathcal{C}|-1)x_c + |\mathcal{P}|x_p \leq |\mathcal{C}|x_c,$$

which implies $x_p < x_c$ as $|\mathcal{P}| > 1$, a contradiction.

Second, assume by contradiction that $|\mathcal{P}|x_p \leq x_c$. From equilibrium condition 1 of Lemma 1, as $|\mathcal{P}| > 1$, and f strictly increasing and concave it follows

$$\begin{aligned} |\mathcal{P}|x_p &\leq x_c \\ \Leftrightarrow |\mathcal{P}|f(|\mathcal{C}|x_c) &\leq f(|\mathcal{P}|x_p + (|\mathcal{C}|-1)x_c) \\ \Rightarrow |\mathcal{P}||\mathcal{C}|x_c &< |\mathcal{P}|x_p + (|\mathcal{C}|-1)x_c \\ \Leftrightarrow \{(|\mathcal{P}|-1)|\mathcal{C}|+1\}x_c &< |\mathcal{P}|x_p, \end{aligned}$$

which by $(|\mathcal{P}|-1)|\mathcal{C}| > 0$ implies $x_c < |\mathcal{P}|x_p$, a contradiction.

Part (ii): Pick any $\hat{g} \in \bar{G}_{cp}^{com}$ arbitrarily. From *part (iii),* we know that core players choose symmetric activity x_c whereas periphery players choose symmetric activity $x_p < x_c$ in any $(\mathbf{x}^*, g^*) \in \mathcal{E}_+$ with $\bar{g}^* = \hat{g}$, i.e. when equilibrium conditions one in Lemma 1 is met. Furthermore, condition four of the lemma is then met trivially in any core-periphery network.

We now show that (the remaining) equilibrium conditions two and three from Lemma 1 are met for a non-degenerate interval of linking costs $k \in [\underline{k}, \bar{k}] \neq \emptyset$ (and when $|\mathcal{P}| > 2$) iff h is sufficiently linear:

As $x_p < x_c$ and by Lemma 2, incentives to deviate are minimised when periphery players sponsor all their links. Then the third condition of the lemma demands that neither a core player (15) nor a periphery player (16) is allowed to gain from deleting a link to a core player and hence

$$h(|\mathcal{P}|x_p + (|\mathcal{C}|-1)x_c) - h(|\mathcal{P}|x_p + (|\mathcal{C}|-2)x_c) \geq k \qquad (15)$$

$$h(|\mathcal{C}|x_c) - h((|\mathcal{C}|-1)x_c) \geq k. \qquad (16)$$

The first inequality is sufficient for the second since $x_c < |\mathcal{P}|x_p$ by *part (iii)* and h concave. In other words, if core players prefer to sustain their marginal link to other core players, then periphery players also prefer to sustain their marginal link.

The second equilibrium condition of Lemma 1 (preventing adding of links) is met trivially for core players as they are already linked to all remaining players. Thus we only need to assure that a periphery player does not gain from adding a link to another periphery player i.e.

$$h(|\mathcal{C}|x_c + x_p) - h(|\mathcal{C}|x_c) \leq k. \tag{17}$$

In conclusion, for any $\hat{g} \in \bar{G}^{com}_{cp}$ there is $(\mathbf{x}^*, g^*) \in \mathcal{E}_+$ with $\bar{g}^* = \hat{g}$ iff Equations (15) and (17) are fulfilled simultaneously for a non-degenerate interval of linking costs. That is

$$\begin{aligned} & h(|\mathcal{P}|x_p + (|\mathcal{C}|-1)x_c) - h(|\mathcal{P}|x_p + (|\mathcal{C}|-2)x_c) \\ > \; & h(|\mathcal{C}|x_c + x_p) - h(|\mathcal{C}|x_c). \end{aligned} \tag{18}$$

For $|\mathcal{P}| > 2$, the inequality holds true if h is (locally) sufficiently linear as $x_c > x_p$; however, if h is (locally) sufficiently concave, the inequality is not satisfied since—as shown below—$|\mathcal{P}|x_p + (|\mathcal{C}|-2)x_c > |\mathcal{C}|x_c$.

To see that, assume $|\mathcal{P}| > 2$ and by contradiction $\frac{|\mathcal{P}|}{2}x_p \leq x_c$. Then from optimal activity and the properties of f and as $|\mathcal{C}| \geq 2$ we get

$$\begin{aligned} \frac{|\mathcal{P}|}{2}x_p &\leq x_c \\ \Leftrightarrow \frac{|\mathcal{P}|}{2}f(|\mathcal{C}|x_c) &\leq f(|\mathcal{P}|x_p + (|\mathcal{C}|-1)x_c) \\ \Rightarrow \frac{|\mathcal{P}|}{2}|\mathcal{C}|x_c &< |\mathcal{P}|x_p + (|\mathcal{C}|-1)x_c \\ \Leftrightarrow (\frac{|\mathcal{P}|-2}{4}|\mathcal{C}| + \frac{1}{2})x_c &< \frac{|\mathcal{P}|}{2}x_p \\ \Rightarrow x_c &< \frac{|\mathcal{P}|}{2}x_p \end{aligned}$$

□

Proof of ***Proposition 9****.* Consider any equilibrium $(\mathbf{x}^*, g^*) \in \mathcal{E}_{++}$. We can partition the set of players N into $\bar{l}$ sets of equi-activity players N^l, i.e. $\cup_{l=1}^{\bar{l}} N^l = N$ and $x_i^* = x^l$, $\forall i \in N^l$. We order these sets by activity so that $x^1 < x^2 < \cdots < x^{\bar{l}}$.

For our proof, two *observations* are decisive: *(i)* If a player $i \in N^i$ sponsors a link to some player $j \in N^j$ in a *strict* equilibrium, then he must share a link with all other players $j' \in N^j$. Otherwise player i can—without changing his payoff—delete his link to j and create a new link to some $j' \in N^j$ (with whom he currently does not share a link) instead and the equilibrium is not strict.

(ii) If player i shares a link with all other players $j \in N \setminus \{i\}$ then his activity is maximal, i.e. $i \in N^{\bar{l}}$ and $x_i^* = x^{\bar{l}}$. By contradiction assume there exists $j \in N$ with $x_j^* > x_i^*$. Then from optimal activity we have

$$\begin{aligned} x_i^* = f\big(\sum_{s \in N_i} x_s^*\big) \quad &< \quad f\big(\sum_{s \in N_j} x_s^*\big) = x_j^* \\ \Rightarrow \sum_{s \in N} x_s^* - x_i^* \quad &< \quad \sum_{s \in N_j} x_s^* \leq \sum_{s \in N} x_s^* - x_j^* \\ \Rightarrow x_j^* \quad &< \quad x_i^*. \end{aligned}$$

For the proof, we apply Lemma 2 and assume that links are sponsored by an adjacent player with weakly smaller activity.

Consider players in N^1 first. There are three possible configurations: *(a)* If players in N^1 do not sponsors any links then they have maximal activity $x^1 = x^{\bar{l}}$: Due to complementarity, any hypothetical higher activity players would have to share some links among themselves, in contradiction to Proposition 5 part (ii). Thus the network is empty, i.e. a "complete" onepartite graph.

(b) Or some player $i \in N^1$ sponors a link to another player in N^1. Then he must (in equilibrium) share a link with all higher activity players in $N \setminus N^1$. By the two observations above, $\bar{l} = 1$ follows and the network is complete. The complete network is a variation of a "complete" onepartite network as described by the proposition.

(c) Or players in N^1 (only) sponsor links to higher activity players (and N^1 is an independent set). Assume by contradiction that they did not sponsor any links to players in N^2. Then Proposition 5 parts (ii) and (iii) implies $x^1 = x^2$ which is not possible by definition. Therefore and as players in N^1 have to sponsor the same number of links (again by Proposition 5), from Observation (i) above and as players prefer to sponsor links to higher activity players, players in N^1 have to sponsor links to all higher activity players $\cup_{l=2}^{\bar{l}} N^l$.

Repeating a similar argument for N^2 to $N^{\bar{l}}$ concludes the proof. □

*Proof of **Proposition 10**.* First consider equilibria with a complete interaction network and $n \geq 2$ players. The (positive) equilibrium activity solves $x_n = f((n-1)x_n) > 0$. Furthermore, x_n is increasing in n: assuming that $n \geq 2$ is continuous, differentiating both sides gives

$$\frac{\partial x_n}{\partial n} = \frac{x_n}{\frac{1}{f'((n-1)x_n)} - (n-1)} > 0$$

where the denominator is positive since $x_n = f((n-1)x_n) > f'((n-1)x_n)(n-1)x_n$ by concavity of f and as $f(0) = 0$.

There is no lower bound of linking costs in a complete network as no player can add a link even if desired; the upper bound of linking costs preventing players from deleting their marginal link is given by

$$\begin{aligned} \bar{k}_n^{com} &= h((n-1)x_n) - h((n-2)x_n) \\ &= h((n-1)x_n) - h((n-1)x_n - f((n-1)x_n)). \end{aligned}$$

Finally as $(n-1)x_n$ increases in n and goes to infinity

$$\bar{k}_\infty^{com} = \lim_{n\to\infty} \bar{k}_n^{com} = \lim_{x\to\infty} h(x) - h(x - f(x)),$$

which is well defined as $h(x) - h(x - f(x))$ is eventually monotone by Assumption 4 and the limit of a monotone function exists. Thus an equilibrium with a complete network exists for any $k \in (0, \bar{k}^{\infty}_{com})$ for finite but sufficiently large n.

Consider instead equilibria with a (single) star network with $n-1$ spokes. The centers (positive) equilibrium activity x_c solves $x_c = f((n-1)x_s)$ and the spokes activity x_s is determined by $x_s = f(x_c)$. Similar arguments about complementarity as above imply that x_s as well as x_c increase in n. Also note that for $n \geq 3$, the center's activity is higher than the activity of the spokes. Then Lemma 2 implies that incentives to deviate from linking are minimised if links are sponsored by spokes—which we assume henceforth. As spokes have a single link, there is the following upper bound on linking costs that prevent them from deleting their link

$$\begin{aligned} \bar{k}^{star}_n &= h(x_c) - h(0) \\ &= h(f((n-1)x_s)). \end{aligned}$$

As $(n-1)x_s$ goes to infinity in n

$$\bar{k}^{star}_{\infty} = \lim_{n\to\infty} \bar{k}^{star}_n = \lim_{x\to\infty} h(f(x)) = \bar{\mathbf{k}}_{max}.$$

Conversely, there is a lower bound of linking costs preventing spokes from linking to other spokes

$$\begin{aligned} \underline{k}^{star}_n &= h(x_c + x_s) - h(x_c) \\ &= h(x_c + f(x_c)) - h(x_c). \end{aligned}$$

As f is unbounded from above, we know that the centers activity x_c converges to infinity in n and thus

$$\underline{k}^{star}_{\infty} = \lim_{n\to\infty} \underline{k}^{star} = \lim_{x\to\infty} h(x + f(x)) - h(x),$$

where the limit exists again due to Assumption 4. Thus the star network exists for any $k \in (\underline{k}^{star}_{\infty}, \bar{\mathbf{k}}_{max})$ for finite but sufficiently large n.

Finally, we need to compare $\underline{k}^{star}_{\infty}$ and $\bar{k}^{com}_{\infty}$. As h is strictly concave we know for each finite $x > 0$

$$0 < h(x + f(x)) - h(x) < h(x) - h(x - f(x)).$$

Thus in the limit for $x \to \infty$ it follows $\underline{k}_\infty^{star} \leq \bar{k}_\infty^{com}$. If $\underline{k}_\infty^{star} < \bar{k}_\infty^{com}$ or $\underline{k}_\infty^{star} = \bar{k}_\infty^{com} \in \{0, \bar{\mathbf{k}}_{max}\}$ the claim holds trivially. Thus assume that $\underline{k}_\infty^{star} = \bar{k}_\infty^{com} \in (0, \bar{\mathbf{k}}_{max})$ i.e. both functions converge to the *same* finite limit. As both functions are eventually monotone and as there is a positive gap between them for any finite x, it can then not be the case that $h(x + f(x)) - h(x)$ eventually decreases while $h(x) - h(x - f(x))$ eventually increases. If the former eventually increases, the star is an equilibrium for $k \in [\underline{k}_\infty^{star}, \bar{\mathbf{k}}_{max})$ for sufficiently high n, if the latter eventually decreases the complete network is an equilibrium for $k \in (0, \bar{k}_\infty^{com}]$ for sufficiently high n. Thus either a star or a complete network are an equilibrium for all $k \in (0, \bar{\mathbf{k}}_{max})$ for sufficiently high n. □

References

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

Ballester, C., Calvó-Armengol, A. and Zenou, Y. (2006). Who's who in networks. wanted: The key player. *Econometrica*, **74** (5), 1403–1417.

Bramoullé, Y. and Kranton, R. (2007). Public goods in networks. *Journal of Economic Theory*, **135** (1), 478–494.

—, — and D'Amours, M. (2012). Strategic interaction and networks, unpublished.

—, López-Pintado, D., Goyal, S. and Vega-Redondo, F. (2004). Network formation and anti-coordination games. *International Journal of Game Theory*, **33** (1), 1–19.

Chui, M., Manyika, J., Bughin, J., Dobbs, R., Roxburgh, C., Sarrazin, H., Sands, G. and Westergren, M. (2012). The social economy: Unlocking value and productivity through social technologies.

Dembosky, A. and Demos, T. (2012). Facebook valued at $104bn in ipo.

Erdős, P. and Gallai, T. (1960). Graphs with prescribed degree of vertices (hungarian). *Mat. Lapok*, **11**, 264–274.

Galeotti, A. and Goyal, S. (2010). The law of the few. *American Economic Review*, **100** (4), 1468–92.

—, —, Jackson, M. O., Vega-Redondo, F. and Yariv, L. (2010). Network games. *Review of Economic Studies*, **77** (1), 218–244.

Hiller, T. (2010). Peer Effects and Endogenous Network Formation. *SSRN eLibrary*, **http://ssrn.com/abstract=1581510**.

Kennan, J. (2001). Uniqueness of positive fixed points for increasing concave functions on rn: An elementary result. *Review of Economic Dynamics*, **4** (4), 893–899.

Lagerås, A. N. and Seim, D. (2012). Strategic complementarities, network games and endogenous network formation, http://people.su.se/ dase3344/lagerasseim111102.pdf.

Milgrom, P. and Roberts, J. (1990). Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica*, **58** (6), pp. 1255–1277.

Tarski, A. (1955). A lattice-theoretical fixpoint theorem and its applications. *Pacific Journal of Mathematics*, **5** (2), 285–309.

Topkis, D. M. (1979). Equilibrium points in nonzero-sum n-person submodular games. *SIAM Journal on Control and Optimization*, **17** (6), 773–787.

Ugander, J., Karrer, B., Backstrom, L. and Marlow, C. (2011). The anatomy of the facebook social graph. *CoRR*, **abs/1111.4503**.

Vives, X. (1990). Nash equilibrium with strategic complementarities. *Journal of Mathematical Economics*, **19** (3), 305–321.

— (2005). Complementarities and games: New developments. *Journal of Economic Literature*, **43** (2), 437–479.

Wilson, C., Boe, B., Sala, A., Puttaswamy, K. P. and Zhao, B. Y. (2009). User interactions in social networks and their implications. In *Proceedings of the 4th ACM European conference on Computer systems*, EuroSys '09, New York, NY, USA: ACM, pp. 205–218.

Zhang, Y., Park, J. and van der Schaar, M. (2011). Production and network formation games with content heterogeneity. *CoRR*, **http://arxiv.org/abs/1109.4201v1**.

B Supplementary Material for "Social Activity and Network Formation"

B.1 A Family of Social Activity Models

In this appendix, we present a family of social activity models which fulfil the properties assumed in the main part of the paper. Consider the following gross payoffs before linking costs are realised

$$\pi_i(\mathbf{x}, \bar{g}) = \lambda\big[v(\sum_{j \in N_{i,\bar{g}}} x_j)x_i\big]^{\frac{\tau}{\lambda}} - (cx_i)^\tau, \tag{19}$$

where $c \in (0, \infty)$ denotes marginal costs from activity and $\tau > 0$ so that these costs increase in activity. The value of neighbour activity $v(\cdot)$ is strictly increasing, concave and normalised, i.e. $v(0) = 0$ and $v' > 0$ and $v'' < 0$. For consistency with the general model, we additionally assume that v^τ is strictly concave and $\lambda \geq 2$ (where $\lambda = 2$ comes at no loss of generality as shown below).

There is a *unique* optimal level of activity for agent i given *any* activity $\mathbf{x}_{-i}$ and *any* network $\bar{g}$ which solves the first order condition of the gross payoff in Equation (19):

$$x_i^*(\mathbf{x}_{-i}, \bar{g}) = \frac{v(\sum_{j \in N_{i,\bar{g}}} x_j)^{\frac{1}{\lambda-1}}}{c^{\frac{\lambda}{\lambda-1}}} \equiv f(\sum_{j \in N_{i,\bar{g}}} x_j), \, \forall i \in N. \tag{20}$$

Note that $v(\cdot)^{\frac{1}{\lambda-1}}$ is strictly increasing and concave by the assumptions on v and as $\frac{1}{\lambda-1} \in (0, 1)$. Thus all assumptions on f from the main part are met.

For $\lambda = 2$, which is without loss of generality as shown below, $v'(0) > c^2$ and $\lim_{x\to\infty} v'(x) < \frac{c^2}{n-1}$ ensures existence of a positive activity equilibrium on any non-singleton component of $\bar{g} \in \bar{G}^n$ by Proposition 1.

Plugging the best response activity back into the gross payoff π_i we get the maximised gross payoff as

$$\pi_i^*(x_i^*, \mathbf{x}_{-i}, \bar{g}) = (\lambda - 1)\frac{v(\sum_{j\in N_{i,\bar{g}}} x_j)^{\frac{\tau}{\lambda-1}}}{c^{\frac{\tau}{\lambda-1}}} \equiv h(\sum_{j\in N_{i,\bar{g}}} x_j),\ \forall i \in N, \tag{21}$$

where $v^{\frac{\tau}{\lambda-1}}$ is strictly increasing and concave as $\frac{1}{\lambda-1} \in (0,1)$ and v^τ is strictly increasing and concave. Thus the assumptions on h from the main part of the paper are met.

Finally, we quickly show that assuming $\lambda = 2$ comes without any loss of generality for the model. Fix any marginal activity costs c, linking costs k, τ, and some $\lambda > 2$. Consider the following transformation of the model: $\hat{v} \equiv v^{\frac{1}{\lambda-1}}$, $\hat{c} = c^{\frac{\lambda}{2(\lambda-1)}}$, $\hat{k} = \frac{k}{\lambda-1}c^{\frac{\tau(2-\lambda)}{2(\lambda-1)}}$, $\hat{\lambda} = 2$, and unaltered $\hat{\tau} = \tau$. As we show below, a pair $(\mathbf{x}^*, g^*)$ constitutes a Nash equilibrium in the original model if and only if it constitutes a Nash equilibrium in the transformed model (compare with the equilibrium conditions in Lemma 1).

The optimal activity under the transformed model and the original model is identical as

$$\frac{\hat{v}(\sum_{j\in N_{i,\bar{g}}} x_j)}{\hat{c}^2} = \frac{v(\sum_{j\in N_{i,\bar{g}}} x_j)^{\frac{1}{\lambda-1}}}{c^{\frac{\lambda}{\lambda-1}}}.$$

Furthermore, optimal linking behaviour is also identical as $\hat{c}^{\hat{\tau}}\hat{k} = \frac{k}{\lambda-1}c^{\frac{\tau}{\lambda-1}}$ so that

$$\begin{aligned} &\frac{\hat{v}(\sum_{j\in N_{i,\bar{g}}} x_j)^{\hat{\tau}}}{\hat{c}^{\hat{\tau}}} \gtreqless \hat{k} \\ \Leftrightarrow\quad &(\lambda-1)\frac{v(\sum_{j\in N_{i,\bar{g}}} x_j)^{\frac{\tau}{\lambda-1}}}{c^{\frac{\tau}{\lambda-1}}} \gtreqless k. \end{aligned}$$

B.2 Properties of Positive Activity Equilibria on Exogenous Networks

In this section we discuss some properties of positive activity equilibria on exogenous and connected networks (or on any non-empty component of a disconnected network) and elaborate on comparative statics as the network density increases:

Proposition 11. *Let $\bar{g}$ be connected and $\mathbf{x}^* > \mathbf{0}$:*
(i) $\mathbf{x}^*$ *is globally asympt. stable for $\mathbf{x} > \mathbf{0}$ under best response dynamics.*
(ii) $\bar{g}_{ij} = 0$ *implies* $\mathbf{x}^*_{\overline{g \oplus ij}} > \mathbf{x}^*_{\bar{g}}$.
(iii) $n > 2$ *implies* $\hat{x} < x^*_i < \sum_{j \in N_i} x^*_j$ *where* $\hat{x} = f(\hat{x}) > 0$.

Proof of ***Proposition 11****.* Throughout the proof we carry over definitions from the proof of Proposition 1.

Part (i): The proof follows similar reasoning as applied in theorem 8 of Milgrom and Roberts (1990). Fix some $\tilde{\mathbf{x}} > \mathbf{0}$. $\mathbf{x}^*$ is globally asymptotically stable under best response dynamics if iterative application of $f_{\bar{g}}$ on $\tilde{\mathbf{x}}$ converges to $\mathbf{x}^*$.

By Assumption 3, there is a unique positive fixed point $\mathbf{x}^* > 0$. Using the same notation as in the proof of Proposition 1, we can find $\mathbf{a}$, $\mathbf{b}$ such that $\mathbf{x}^*, \tilde{\mathbf{x}} \in (\mathbf{a}, \mathbf{b})$ and restricted $f_{\bar{g}} : [\mathbf{a}, \mathbf{b}] \to (\mathbf{a}, \mathbf{b})$ as $\mathbf{a}$ can be arbitrarily close to $\mathbf{0}$ and $\mathbf{b}$ can be arbitrarily large.

Consider three sequences starting at $\alpha^1 \in \{\mathbf{a}, \mathbf{b}, \tilde{\mathbf{x}}\}$ and $\alpha^s = f_{\bar{g}}(\alpha^{s-1})$. As $f_{\bar{g}}$ is non-decreasing and $f_{\bar{g}}(\mathbf{a}) > \mathbf{a}$, the sequence $\mathbf{a}^s$ is non-decreasing. As it is also bounded above by $\mathbf{b}$, it converges to some $\mathbf{x}^0 = \sup(\mathbf{a}^s)$. As both sides of $\mathbf{a}^s = f_{\bar{g}}(\mathbf{a}^{s-1})$ converge to $\mathbf{x}^0$, $\mathbf{x}^0$ is a fixed point of $f_{\bar{g}}$ in $(\mathbf{a}, \mathbf{b})$ and thus $\mathbf{x}^0 = \mathbf{x}^*$. For similar reasons, $\mathbf{b}^s$ is a non-increasing sequence converging to $\mathbf{x}^*$. Finally, $\tilde{\mathbf{x}}^s \in [\mathbf{a}^s, \mathbf{b}^s]$ at the sth elements of the three series as $f_{\bar{g}}$ is non-decreasing so that the sequence $\tilde{\mathbf{x}}^s$ converges to $\mathbf{x}^*$ as well.

For the proofs of the other two parts, we need to provide a stronger finding to facilitate *strict* comparative statics:

Claim 1. Consider some $\mathbf{x}^1 > \mathbf{0}$ with $f_{\bar{g}}(\mathbf{x}^1) \geq \mathbf{x}^1$ and at least one strict entry $f_{\bar{g},i}(\mathbf{x}^1) > x^1_i$. Then the sequence $\mathbf{x}^s$ with $f_{\bar{g}}(\mathbf{x}^s) = \mathbf{x}^{s-1}$ converges to the unique positive fixed point $\mathbf{x}^* > \mathbf{x}^1$.

Proof. From the reasoning above, it is clear that the sequence $\mathbf{x}^s$ converges to the unique positive fixed point $\mathbf{x}^* = f_{\bar{g}}(\mathbf{x}^*)$. As $f_{\bar{g}}(\mathbf{x}^1) \geq \mathbf{x}^1$ and $f_{\bar{g}}$ is non-decreasing, the sequence $\mathbf{x}^s$ is also non-decreasing and thus $\mathbf{x}^* \geq \mathbf{x}^1$. Assume by contradiction that there exists $i \in N$ with $x^*_i = x^1_i$. As $\mathbf{x}^* \geq \mathbf{x}^1$ and by the fixed point finding $x_i^* = f_{\bar{g},i}(\mathbf{x}^*)$ it follows immediately that $x_j^* = x_j^1$ for all neighbours $j \in N_{i,\bar{g}}$. Iterating the argument and as $\bar{g}$ is connected this implies $\mathbf{x}^* = \mathbf{x}^1$. As $\mathbf{x}^*$ is a fixed point, $f_{\bar{g}}(\mathbf{x}^1) = \mathbf{x}^1$, a contradiction to the assumptions of the claim. □

Part (ii): Consider some network with $g_{ij} = 0$. Since $f_{\bar{g}}(\mathbf{x}_{\bar{g}}^*) = \mathbf{x}_{\bar{g}}^* > \mathbf{0}$ it follows $f_{\overline{g \oplus ij}}(\mathbf{x}_{\bar{g}}^*) \geq f_{\bar{g}}(\mathbf{x}_{\bar{g}}^*) = \mathbf{x}_{\bar{g}}^*$ and $f_{\overline{g \oplus ij},i}(\mathbf{x}_{\bar{g}}^*) > f_{\bar{g},i}(\mathbf{x}_{\bar{g}}^*)$. By Claim 1, the sequence $\mathbf{x}^s = f_{\overline{g \oplus ij}}(\mathbf{x}^{s-1})$ with $\mathbf{x}^1 = \mathbf{x}_{\bar{g}}^*$ then converges to $\mathbf{x}_{\overline{g \oplus ij}}^* > \mathbf{x}^1 = \mathbf{x}_{\bar{g}}^*$.

Part (iii): Note that there is a unique positive $\hat{x}$ for which $\hat{x} = f(\hat{x})$ as $f(0) = 0$, $f'(0) > 1$, f strictly concave and $\lim_{x \to \infty} f'(x) < 1$. Furthermore, for $y > 0$ and $y \gtreqless \hat{x}$ we have $y \gtreqless f(y)$. Consider any connected graph g with $n > 2$ and assume $\mathbf{x}^1 = \hat{\mathbf{x}}$. There is at least one player who has more than one neighbour in $\bar{g}$ and for whom $f_{\bar{g},i}(\mathbf{x}^1) > x_i^1$ by definition of $\hat{x}$. All other players have only one neighbour, i.e. $f_{\bar{g},i}(\mathbf{x}^1) = x_i^1$. By Claim 1, this implies $\mathbf{x}^* > \mathbf{x}^1 = \hat{\mathbf{x}}$ and so $x_i^* < \sum_{j \in N_i} x_j^*$ for all i. □

The proposition illuminates some properties of the positive activity equilibrium: the first finding tells us that the positive activity equilibrium is stable under best response dynamics even if the society is severely shocked. Conversely, it is easy to see that the zero activity equilibrium is unstable: if two adjacent players—i.e. two neighbours—simultaneously make an error and choose some non-zero activity, then best response dynamics inevitably lead to the positive activity equilibrium.

The second finding shows that increasing the density of a connected network by adding a single link is not just beneficial for adjacent players who are directly effected from additional positive externalities. By complementarity, their strategic reaction to higher marginal benefits is increasing activity on their own which is in turn beneficial for all their neighbours. Thus there is a multiplicative effect and ultimately all players increase their activity and benefit from the new link.

The third finding re-emphasises that diminishing activity incentives—sufficient concavity of f—is crucial to keep total network activity in balance. Although there is complementarity, players' activity is strictly lower than their cumulative neighbour activity in any positive activity equilibrium with more than two players.

B.3 The Partition of $\mathcal{E}$

In this section, we characterise the equilibrium-types that exist in $\mathcal{E}$ with a special focus on some non-generic equilibria which are excluded from $\mathcal{E}_+$.

Proposition 12. *$\mathcal{E} \subseteq \{(\mathbf{x}^*_{\bar{g}}, g)\}_{g \in G}$ so that $\mathcal{E}$ is partitioned into three sets:*

i. A zero activity equilibrium with $\mathbf{x}^* = \mathbf{0}$ and empty network $\bar{g}^*_{emp}$

ii. Positive activity equilibria with $\mathbf{x}^* > \mathbf{0}$ and $n_{i,\bar{g}^*} \geq 1, \forall i \in N$

iii. Non-generic equilibria in which—for $\hat{x} = f(\hat{x})$ and $\tilde{x} = h^{-1}(k)$—either

(a) $x^*_i \in \{0, \tilde{x}\}$, $\tilde{x} = \hat{x}$, i.e. $k = h(\hat{x})$; the network consists of isolated players with $x^*_i = 0$ and pairs of players with $x^*_i = \tilde{x}$ each; or

(b) $x^*_i \in \{0, f(\tilde{x}), \tilde{x}\}$, $\tilde{x} > \hat{x}$, i.e. $k > h(\hat{x})$; the network consists of isolated players with $x^*_i = 0$ and stars in which the center has activity $x^*_c = \tilde{x} = f(r x^*_s)$ and all $r > 1$ spokes have activity $x^*_s = f(\tilde{x})$.[39]

*Proof of **Proposition 12**.* The proposition is an extension of Proposition 2. Thus we only need to discuss the non-generic equilibria in case 3 from its proof in more detail. Recall that there are isolated as well as non-isolated players in these equilibria and define $\tilde{x} = h^{-1}(k)$. As some players sponsor links $\max x^*_i \geq \tilde{x}$ because no link is sponsored to any player with $x^*_i < \tilde{x}$ by Lemma 1 and by concavity of h. Conversely, as some players are isolated $\max x^*_i \leq \tilde{x}$ because isolated players strictly gain from linking to a player with $x^*_i > \tilde{x}$. Thus $\max x^*_i = \tilde{x}$ and the following statements are true: links are only formed to players with $x^*_i = \tilde{x}$, every player sponsors at most one

[39] For instance, $\hat{x} = 1$ and $\tilde{x} = k^2$ in the baseline model with $c = 1$ and $q = \frac{1}{2}$. Then a non-generic equilibrium with isolated players and stars—in which the center has activity k^2 and the $r > 1$ spokes activity k—exists iff linking costs $k = \sqrt[3]{r}$, e.g. for linking costs $k = 2$ stars with eight spokes and any number of isolated players constitute an equilibrium.

link and if he sponsors a link, he has no incoming links. We consider two sub-cases in turn.

For part *(iii) (a)*, assume that there exists some player i with $x_i^* = \tilde{x}$ who sponsors a link to some other player j (with $x_j^* = \tilde{x}$) in equilibrium. Then $\tilde{x} = \hat{x}$ where $\hat{x} = f(\hat{x})$ from optimal activity. As f strictly increases, no player has more than one (incoming) link and there is a unique positive level of activity—i.e. the network consists of pairs of players with $x_i^* = \hat{x} = \tilde{x}$ each and isolated players with $x_i^* = 0$.

For part *(iii) (b)*, assume that no player with $x_i^* = \tilde{x}$ sponsors any links. Then only players with $x_s^* \in (0, \tilde{x})$ sponsor links to players with $x_i^* = \tilde{x}$. As all those players sponsor only one link and have no incoming links, they must choose a symmetric level of activity solving $x_s^* = f(\tilde{x}) = f(h^{-1}(k))$ by the equilibrium condition. Consequently, there are three different levels of activity in such an equilibrium $x_i^* \in \{0, f(\tilde{x}), \tilde{x}\}$ and the equilibrium network consists of any number of isolated players with zero activity and stars—in which the center's activity is $x_c^* = \tilde{x}$ and the spokes' activity is $x_s^* = f(\tilde{x})$. Furthermore, $x_s^* = f(\tilde{x}) < \tilde{x} = x_c^*$ implies $\tilde{x} > \hat{x} = f(\hat{x})$ as f is concave and fixed point $\hat{x}$ exists by Assumption 3. $\tilde{x} > \hat{x}$ implies $k = h(\tilde{x}) > h(\hat{x})$. Finally, if each star consists of a center and $r > 1$ spokes, optimality of center activity implies $x_c^* = \tilde{x} = f(rx_s^*)$. □

The two non-generic equilibria are special cases in which the value from linking to (and only to) one of the most active players is exactly equal to the linking costs, i.e. the most active players choose activity $\tilde{x} = h^{-1}(k)$. Therefore, isolated players are indifferent between their isolation and forming a single link to one of the most active players (with simultaneous activity adjustment to $f(\tilde{x})$). Although these non-generic equilibria are contained in $\mathcal{E}$ they are excluded from the set $\mathcal{E}_+$: for any $k' < k$, isolated players would strictly prefer to form a link to some player with $\tilde{x}$ and for any $k' > k$, players who support a link prefer to delete it.

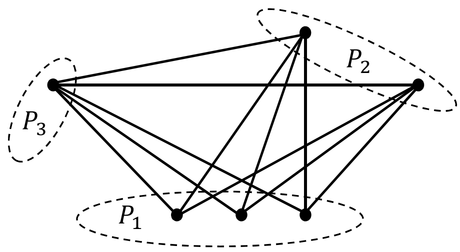

Figure 5: Complete Multipartite Network

B.4 Complete Multipartite Networks

Multipartite networks are a generalisation of bipartite networks as discussed in Section 4.2.2: the set of agents N can be partitioned into any number of independent sets (so-called *partite sets* or *parts*). In a *complete* multipartite network, every agent shares a link with all agents outside his own part. An example of such a network can be seen in Figure 5.

We assume in this section additionally that there are at least two parts and all of them are of different size, that is the parts can be labeled in descending order

$$|P_1| > |P_2| > \cdots > |P_{\bar{l}}| \geq 1.$$

We collect (undirected) networks which fulfil these properties in the set $\bar{G}^{com}_{mp}$.

There are some regularities for the equilibrium activity on any (exogenously given) complete multipartite network: First, as all players from a partite set share the same neighbours, they choose common equilibrium activity. Second, as players from smaller partite sets share links with a larger fraction of the entire society and due to strategic complementarity, they have higher equilibrium activity.

Next, consider some directed network g' with $\bar{g}' \in \bar{G}^{com}_{mp}$. We apply Lemma 2 and assume that all links are sponsored by the adjacent player from the larger partite set (who has smaller equilibrium activity) so that incentives to deviate are minimised. We call the lower (respectively upper) bound of linking costs such that no player from the lth partite set changes his linking behaviour in g' as $\underline{k}^l$ (respectively $\bar{k}^l$).

We are able to show that there is always a non-zero interval of linking costs $[\underline{k}^l, \bar{k}^l]$ such that players from partite set P_l do not deviate from their linking. However, $(\mathbf{x}^*_{\bar{g}'}, g')$ is only an equilibrium if *all* the intervals $[\underline{k}^l, \bar{k}^l]$ have a non-zero intersect, that is $\cap_{l=1}^{\bar{l}}[\underline{k}^l, \bar{k}^l] \neq \emptyset$. Unfortunately, we are not able to determine in general when this holds (apart from the special case of a complete bipartite network). The proposition below states our partial results formally:

Proposition 13. *Let $\bar{g} \in \bar{G}^{com}_{mp}$.*
(i) *Every $i \in P_l$ chooses common $x^*_i = x^l$ in equilibrium with $x^l < x^{l+1}$.*
(ii) *If $g_{ij} = 1$ implies $x^*_i < x^*_j$ then $\underline{k}^l < \bar{k}^l$ and $\underline{k}^{l+1} < \bar{k}^l$.*

Proof of ***Proposition 13****.* ***Part (i):*** Players from the same partite set are linked to the same set of players, thus access the same cumulative neighbour activity and hence choose some common level of activity by condition one of Lemma 1.

Assume $x^l \geq x^{l+1}$. As f is strictly increasing, optimal activity implies $|P_{l+1}|x^{l+1} \geq |P_l|x^l$. As $|P_{l+1}| < |P_l|$, this in turn implies $x^{l+1} > x^l$, a contradiction.

Part (ii): Let y^l be the total cumulative neighbour activity accessed by any player from the lth partite set. As players from lower partite sets have less equilibrium activity by *part (i)* and by upwards linking, the remaining equilibrium conditions in Lemma 1 demand that—for any $l \in \{1, \dots, \bar{l}\}$—players in the lth partite set refrain from adding a single link to another player in the lth partite set and (if $l < \bar{l}$) from deleting a single link to a player in the $l+1$th independent set. Thus

$$\begin{aligned} k &\geq h(y^l + x^l) - h(y^l) \equiv \underline{k}^l \\ k &\leq h(y^l) - h(y^l - x^{l+1}) \equiv \bar{k}^l. \end{aligned}$$

First, $\underline{k}^l < \bar{k}^l$ as by concavity of h

$$h(y^l + x^l) - h(y^l) < h(y^l + x^{l+1}) - h(y^l) < h(y^l) - h(y^l - x^{l+1}).$$

Second, note that as $x^l < x^{l+1}$, optimal activity implies $y^l < y^{l+1}$. Hence

q	$\|P_1\|$	$\Delta\|P_l\|$	$\bar{l}_{max}$
0.4	21	1	7
	21	3	5
	100	3	5
0.5	21	1	21
0.7	21	1	3
	21	3	3
	21	5	2

Table 2: Equilibria with Complete Multipartite Interaction Networks

$\underline{k}^{l+1} < \bar{k}^l$ since

$$h(y^{l+1} + x^{l+1}) - h(y^{l+1}) < h(y^l + x^{l+1}) - h(y^l) < h(y^l) - h(y^l - x^{l+1}).$$

□

As we are unable to provide general theoretical results regarding equilibrium existence (i.e. determine when $\cap_{l=1}^{\bar{l}}[\underline{k}^l, \bar{k}^l] \neq \emptyset$ under upward linking), table 2 instead presents some numerical analyses for the baseline model.[40] Fixing different values of q (the power of the root function), the size of P_1, and a (constant) rate of decrease between succeeding partitions $\Delta|P_l| = |P_l| - |P_{l+1}|$, the table shows the maximal number of partite sets $\bar{l}_{max}$ that can be sustained in equilibrium—that is the maximal number of activity levels achievable in a complete multipartite network under these conditions.

We draw two major insights from the numerical analyses: first, there can be a considerable number of partite sets in equilibrium. In other words, multi-level equilibria can display a substantial number of distinct levels of activity. For example if $q = 0.5$ and $|P_1| = 21$, a complete 21-partite network (with 21 distinct levels of activity) is an equilibrium for some linking costs.[41]

Second, whether a particular complete multipartite network can be sustained in equilibrium depends on the details of the model, that is on the

[40] The baseline model is covered in more detail in Section 5.2 and Appendix B.6. For the numerical analysis, we normalise the marginal cost of activity $c = 1$.

[41] In fact, we conjecture that for $q = 0.5$ complete multipartite networks with any number of partitions can be sustained in equilibrium for $k = 0.5$.

shape of f and h. For instance, there is an equilibrium (for some linking costs) inducing a complete five-partite network with $|P_1| = 21$ and $\Delta|P_l| = 3$ if $q = 0.4$ but not if $q = 0.7$.

B.5 Core-Periphery Networks with two Periphery players

For the discussion in this section, we need to specify some networks: The set $\bar{G}_{cp}^{\neg com}$ contains all non-complete core-periphery networks and networks in the special subclass $\bar{G}_{cp}^{\neg com*}$ consist of two periphery players, an arbitrary number of "high activity" and two "low activity" core players, i.e. $|\mathcal{P}| = 2$ and $\mathcal{C} = \bar{\mathcal{C}} \cup \underline{\mathcal{C}}$ with $|\underline{\mathcal{C}}| = 2$. Additionally, both periphery players are linked to all "high activity" core players as well as one and different "low activity" core players. An example with two high activity core players is depicted in Figure 3 (d) in Section 4.2.3.

The proposition below summarises our findings for core-periphery networks with two periphery players:

Proposition 14. *Let* $|\mathcal{P}| = 2$. *Then:*
(i) $\bar{G}_{cp}^{com} \subset \bar{G}_{\mathcal{E}_+}$
(ii) $\bar{g} \in \bar{G}_{cp}^{\neg com} \cap \bar{G}_{\mathcal{E}_+}$ *implies* $\bar{g} \in \bar{G}_{cp}^{\neg com*}$.
(iii) *For any* $\bar{g} \in \bar{G}_{cp}^{\neg com*}$, $\bar{g} \in \bar{G}_{\mathcal{E}_+}$ *iff* $|h''|$ *is (loc.) suff. small or* $|\bar{\mathcal{C}}| \in \{0, 1\}$.
(iv) *If* $\bar{g}^* \in \bar{G}_{cp}^{\neg com*}$ *then players within* $\mathcal{P}$, $\underline{\mathcal{C}}$, *and* $\bar{\mathcal{C}}$ *have common activity with*

$$x_p < x_{\underline{c}} < x_{\bar{c}} < x_p + x_{\underline{c}}.$$

Proof of ***Proposition 14***. We are going to prove *part (i)*, *part (ii)*, and *part (iv)* first and turn to *part (iii)* subsequently.

Part (i): The claim follows directly from the proof of part (ii) of Proposition 8 since Equation (18) is always met for $|\mathcal{P}| = 2$ as h is concave, $x_p < x_c$, and $2x_p + (|\mathcal{C}| - 2)x_c < |\mathcal{C}|x_c$.

Part (ii): The claim follows directly from the proof of part (i) of Proposition 8.

Part (iv): Assume there is some equilibrium with an incomplete core-periphery interaction graph $\bar{g}^* \in \bar{G}_{cp}^{\neg com*}$. We want to determine the positive activity equilibrium vector, which exists by Assumption 3. As all players in $\bar{\mathcal{C}}$ link to all other players, they choose common activity $x_{\bar{c}}$ by the same argument as presented in the previous proposition.

Furthermore, both periphery players have common equilibrium activity x_p and both $\underline{\mathcal{C}}$ -player have common equilibrium activity $x_{\underline{c}}$. Assume not. Then fixing the activity of $\bar{\mathcal{C}}$-player and switching the activity of both periphery players as well as both $\underline{\mathcal{C}}$-player gives a second distinct positive activity equilibrium vector by symmetry of the network; a contradiction to the uniqueness result in Proposition 1.

These common equilibrium activity levels can be ranked: first, $x_p < \min\{x_{\underline{c}}, x_{\bar{c}}\}$ as shown in the proof of part (i) of Proposition 8. Second, assume by contradiction that $x_{\underline{c}} \geq x_{\bar{c}}$. Optimal activity and increasing f implies

$$\begin{aligned} |\bar{\mathcal{C}}|x_{\bar{c}} + x_{\underline{c}} + x_p &\geq (|\bar{\mathcal{C}}| - 1)x_{\bar{c}} + 2x_{\underline{c}} + 2x_p \\ \Leftrightarrow x_{\bar{c}} &\geq x_{\underline{c}} + x_p, \end{aligned}$$

a contradiction as $x_p > 0$. By the same reasoning, $x_{\underline{c}} < x_{\bar{c}}$ implies $x_{\bar{c}} < x_{\underline{c}} + x_p$.

Part (iii): Pick any $\hat{g} \in \bar{G}_{cp}^{\neg com*}$ arbitrarily. From *part (iii)*, we know that both periphery, both $\underline{\mathcal{C}}$, and all $\bar{\mathcal{C}}$-player choose common activity with $x_p < x_{\underline{c}} < x_{\bar{c}}$ in any $(\mathbf{x}^*, g^*) \in \mathcal{E}_+$ with $\bar{g}^* = \hat{g}$, i.e. when equilibrium conditions one in Lemma 1 is met. Equilibrium condition four of the lemma is then met trivially in any core-periphery network. We now discuss the remaining two conditions in turn.

As $x_p < x_{\underline{c}} < x_{\bar{c}}$ and by Lemma 2, incentives to deviate are minimised if periphery players sponsor all their links and $\underline{\mathcal{C}}$ players sponsor their links to $\bar{\mathcal{C}}$ players. Thus the third equilibrium condition preventing deletion of links (to lowest activity sponsored friends) writes out for $\bar{\mathcal{C}}$-player (22; if $|\bar{\mathcal{C}}| \geq 2$),

$\underline{\mathcal{C}}$-players (23), and periphery players (24) as follows:

$$h(2x_p + 2x_{\underline{c}} + (|\bar{\mathcal{C}}| - 1)x_{\bar{c}}) - h(2x_p + 2x_{\underline{c}} + (|\bar{\mathcal{C}}| - 2)x_{\bar{c}}) \geq k \quad (22)$$
$$h(x_p + x_{\underline{c}} + |\bar{\mathcal{C}}|x_{\bar{c}}) - h(x_p + |\bar{\mathcal{C}}|x_{\bar{c}}) \geq k \quad (23)$$
$$h(x_{\underline{c}} + |\bar{\mathcal{C}}|x_{\bar{c}}) - h(|\bar{\mathcal{C}}|x_{\bar{c}}) \geq k. \quad (24)$$

As h is concave, Equation (23) is sufficient for Equation (24).

The second equilibrium condition of the lemma preventing adding of links is met trivially for $\bar{\mathcal{C}}$ players as they are already linked to all remaining players. Thus we only need to assure that a periphery player does not gain from adding a link to a $\underline{\mathcal{C}}$ player and vice versa:

$$h(2x_{\underline{c}} + |\bar{\mathcal{C}}|x_{\bar{c}}) - h(x_{\underline{c}} + |\bar{\mathcal{C}}|x_{\bar{c}}) \leq k \quad (25)$$
$$h(x_{\underline{c}} + 2x_p + |\bar{\mathcal{C}}|x_{\bar{c}}) - h(x_{\underline{c}} + x_p + |\bar{\mathcal{C}}|x_{\bar{c}}) \leq k. \quad (26)$$

Similarly to above, Equation (25) is sufficient for Equation (26) as h is increasing and concave and $x_p < x_{\underline{c}}$—that is periphery players have a stronger incentive to add a link than $\underline{\mathcal{C}}$ players.

In conclusion, for any $\hat{g} \in \bar{G}_{cp}^{\neg com*}$ there is $(\mathbf{x}^*, g^*) \in \mathcal{E}_+$ with $\bar{g}^* = \hat{g}$ iff Equations (23), (25), and—for $|\mathcal{C}| \geq 2$—Equation (22) are fulfilled simultaneously for a non-degenerate interval of linking costs. That is

$$\begin{aligned} h(2x_{\underline{c}} + |\bar{\mathcal{C}}|x_{\bar{c}}) - h(x_{\underline{c}} + |\bar{\mathcal{C}}|x_{\bar{c}}) &< h(x_p + x_{\underline{c}} + |\bar{\mathcal{C}}|x_{\bar{c}}) - h(x_p + |\bar{\mathcal{C}}|x_{\bar{c}}) \text{ and} \\ h(2x_{\underline{c}} + |\bar{\mathcal{C}}|x_{\bar{c}}) - h(x_{\underline{c}} + |\bar{\mathcal{C}}|x_{\bar{c}}) &< h(2x_p + 2x_{\underline{c}} + (|\bar{\mathcal{C}}| - 1)x_{\bar{c}}) \\ &\quad -h(2x_p + 2x_{\underline{c}} + (|\bar{\mathcal{C}}| - 2)x_{\bar{c}}). \end{aligned}$$

As h is concave and $x_p < x_{\underline{c}}$ the first equation is trivially true and hence if $\bar{\mathcal{C}} \in \{0, 1\}$ then $\hat{g} \in \bar{G}_{\mathcal{E}_+}$. If h is (locally) sufficiently linear, the second equation holds true since $x_{\underline{c}} < x_{\bar{c}}$. Conversely, if h is (locally) sufficiently concave, the second equation is not satisfied since—as shown below—$x_{\underline{c}} + |\bar{\mathcal{C}}|x_{\bar{c}} < 2x_p + 2x_{\underline{c}} + (|\bar{\mathcal{C}}| - 2)x_{\bar{c}}$ for $|\bar{\mathcal{C}}| \geq 2$.

To see that, note that it is equivalent to show $2x_{\bar{c}} < 3x_p < 2x_p + x_{\underline{c}}$ and assume by contradiction $\frac{3}{2}x_p \leq x_{\bar{c}}$. Then from optimal activity and the properties of f we get

$$\begin{aligned}
\frac{3}{2}f(x_{\underline{c}} + |\bar{\mathcal{C}}|x_{\bar{c}}) &\leq f(2x_p + 2x_{\underline{c}} + (|\bar{\mathcal{C}}| - 1)x_{\bar{c}}) \\
\Rightarrow \frac{3}{2}(x_{\underline{c}} + |\bar{\mathcal{C}}|x_{\bar{c}}) &< 2x_p + 2x_{\underline{c}} + (|\bar{\mathcal{C}}| - 1)x_{\bar{c}} \\
\Leftrightarrow (|\bar{\mathcal{C}}| + 2)x_{\bar{c}} &< 4x_p + x_{\underline{c}} \\
\Rightarrow (|\bar{\mathcal{C}}| + 1)x_{\bar{c}} &< 4x_p \\
\Rightarrow x_{\bar{c}} &< \frac{4}{3}x_p < \frac{3}{2}x_p.
\end{aligned}$$

□

The proposition shows that equilibrium core-periphery networks with two periphery players are either complete (part (i)) or from the special class of incomplete core-periphery networks $\bar{G}_{cp}^{\neg com*}$ (part (ii)).

Finally, equilibria with an incomplete core-periphery network from $\bar{G}_{cp}^{\neg com*}$ display three levels of activity: both periphery players have low activity. The two core players who are only linked to one periphery player have medium activity whereas the remaining core players have high activity.

B.6 The Baseline Model and Large Societies

Recall that the gross payoff function in the baseline model was given in Equation (3) as

$$\pi_i(\mathbf{x}, g) = 2\sqrt{(\sum_{j \in N_{i,\bar{g}}} x_j)^q x_i} - cx_i,$$

where $q \in (0, 1)$. It follows straight forwardly that f and h then take the form of root functions, that is

$$x_i^* = \frac{(\sum_{j \in N_{i,\bar{g}}} x_j)^q}{c^2} \equiv f(\sum_{j \in N_{i,\bar{g}}} x_j) \quad (27)$$

$$\pi_i^*(x_i^*, \mathbf{x}_{-i}, \bar{g}) = \frac{(\sum_{j \in N_{i,\bar{g}}} x_j)^q}{c} \equiv h(\sum_{j \in N_{i,\bar{g}}} x_j). \quad (28)$$

If the exponent q is close to zero, then agents are already satiated from little neighbours' activity. The best response function f and the value function h become closer to a step function, i.e. optimal activity jumps up from zero abruptly if some neighbours start to be active but does not change much as cumulative neighbour activity increases any further.

If the exponent q is close to one instead, agents get less easily satiated and f as well as h are relatively linear. In other words, complementarity imposes a relatively proportional positive reaction to increasing cumulative neighbour activity regardless of the current level.

As $f'(x) = qc^{-2}x^{q-1}$, Proposition 1 implies that a positive activity equilibrium exists on any (exogenously given) $\bar{g} \in \bar{G}$—independent of the number of players n—and thus it becomes meaningful to consider the game with endogenous network formation in the limit case when n gets large.

In the analysis, we focus on three types of interaction networks: Section B.6.1 covers the complete and (non-complete) regular interaction networks, whereas Section B.6.2 covers the star network.

B.6.1 Regular and Complete Networks

Recall that Proposition 3 tells us that all regular interaction networks arise in equilibrium for a non-degenerate interval of linking costs. Furthermore, in any equilibrium that induces a Δ-regular interaction network there is a single specific positive level of activity x_Δ. In particular, this level of activity x_Δ is independent of the size of the society n (as long as a Δ-regular network with n agents exists) and the specifics of the network structure (as long as it is Δ-regular).

In the baseline model we can find an explicitly expression for x_Δ. Solving Equation (27), gives $x_\Delta = \left(\frac{\Delta q}{c^2}\right)^{\frac{1}{1-q}}$.

As x_Δ is independent of n, we also know that the (non-degenerate) interval of supporting linking costs $[\underline{k}_\Delta, \bar{k}_\Delta]$ for an equilibrium with a Δ-regular interaction network does not change in n as long as $n - 1 > \Delta$. In the special case $n - 1 = \Delta$, the interaction network is complete and no player can add any links even if desired. Thus such an equilibrium is supported for all

$k \in (0, \bar{k}_\Delta]$.

The lemma below discusses how the lower bound $\underline{k}_\Delta$ and upper bound $\bar{k}_\Delta$ of linking costs supporting equilibria with Δ-regular interaction networks change as the network becomes denser i.e. Δ increases:

Lemma 3. *Consider the baseline model and let* $\Delta \in \{1, \ldots, n-3\}$ *with* $n \ggg 1$. *Then:*

q	$\underline{k}_{\Delta+1} - \underline{k}_\Delta$	$\lim_{\Delta\to\infty} \underline{k}_\Delta$	$\bar{k}_{\Delta+1} - \bar{k}_\Delta$	$\lim_{\Delta\to\infty} \bar{k}_\Delta$
$> \frac{1}{2}$	> 0	∞		∞
$= \frac{1}{2}$	> 0	$\frac{1}{2}c^{-3}$	< 0	$\frac{1}{2}c^{-3}$
$< \frac{1}{2}$		0	< 0	0

Table 3: Analysis of $\underline{k}_\Delta$ and $\bar{k}_\Delta$

*Proof of **Lemma 3**.* From the discussion in the main text above, we know that $x_\Delta = (\frac{\Delta^q}{c^2})^{\frac{1}{1-q}}$. Plugging that into $\bar{k}_\Delta$ and $\underline{k}_\Delta$ as given in the proof to Proposition 4 and substituting for h from Equation (28) gives the following expression for upper and lower bound of linking costs for equilibria with Δ-regular network in the baseline model:

$$\begin{aligned} \bar{k}_\Delta &= \frac{\Delta^{\frac{2q-1}{1-q}} \Delta^{1-q} (\Delta^q - (\Delta - 1)^q)}{c^{\frac{1+q}{1-q}}} \\ \underline{k}_\Delta &= \frac{\Delta^{\frac{2q-1}{1-q}} \Delta^{1-q} ((\Delta + 1)^q - \Delta^q)}{c^{\frac{1+q}{1-q}}}. \end{aligned}$$

For the remainder of the proof, we treat $\Delta \geq 1$ as a continuous variable.

The factor $\Delta^{\frac{2q-1}{1-q}}$ is common to both bounds and we consider its behaviour first. The result is summarised in Table 4 below:

q	$\frac{2q-1}{1-q}$	$\frac{\partial}{\partial \Delta} \Delta^{\frac{2q-1}{1-q}}$	$\lim_{\Delta\to\infty} \Delta^{\frac{2q-1}{1-q}}$
$> \frac{1}{2}$	> 0	> 0	∞
$= \frac{1}{2}$	$= 0$	$= 0$	1
$< \frac{1}{2}$	< 0	< 0	0

Table 4: Analysis of $\Delta^{\frac{2q-1}{1-q}}$

Next, we discuss the behaviour of $\bar{k}_\Delta$. Using a Taylor expansion around Δ and evaluating it at $\Delta - 1$ gives

$$(\Delta - 1)^q = \Delta^q - q\Delta^{q-1} - \frac{q(1-q)}{2!}\Delta^{q-2}\dots.$$

Using this finding gives

$$\Delta^{1-q}(\Delta^q - (\Delta - 1)^q) = q + \frac{q(1-q)}{2!}\Delta^{-1} + \frac{q(1-q)(2-q)}{3!}\Delta^{-2} + \dots.$$

Clearly, this expression decreases in Δ and converges to q from above. Together with the results from the analysis of $\Delta^{\frac{2q-1}{1-q}}$ we can determine the behaviour of $\bar{k}_\Delta$ as shown in Table 5 below:

q	$\frac{\partial}{\partial\Delta}\bar{k}_\Delta$	$\lim_{\Delta\to\infty}\bar{k}_\Delta$
$> \frac{1}{2}$		∞
$= \frac{1}{2}$	< 0	$\frac{1}{2}c^{-3}$
$< \frac{1}{2}$	< 0	0

Table 5: Analysis of $\bar{k}_\Delta$

The analysis of $\underline{k}_\Delta$ follows similar arguments. Using a Taylor expansion around Δ and evaluating it at $\Delta + 1$ gives

$$(\Delta + 1)^q = \Delta^q + q\Delta^{q-1} - \frac{q(1-q)}{2!}\Delta^{q-2} + \frac{q(1-q)(2-q)}{3!}\Delta^{q-3} - \dots.$$

Using this finding gives

$$\Delta^{1-q}((\Delta + 1)^q - \Delta^q) = q - \frac{q(1-q)}{2!}\Delta^{-1} + \frac{q(1-q)(2-q)}{3!}\Delta^{-2} - \dots.$$

Clearly, this expression converges to q. The expression is strictly increasing as the absolute value of each negative term is strictly greater than the following positive term. We can determine the behaviour of $\underline{k}_\Delta$ as shown in Table 6 below:

q	$\frac{\partial}{\partial \Delta}\underline{k}_\Delta$	$\lim_{\Delta\to\infty}\underline{k}_\Delta$
$> \frac{1}{2}$	> 0	∞
$= \frac{1}{2}$	> 0	$\frac{1}{2}c^{-3}$
$< \frac{1}{2}$		0

Table 6: Analysis of $\underline{k}_\Delta$

□

The behaviour of equilibrium supporting linking costs depends quite drastically on exponent q and thus ultimately on the form of the best response function and the value function. The band of supporting linking costs increases and converges to infinity in the network's density if $q > \frac{1}{2}$. Conversely, it decreases and converges to zero if $q < \frac{1}{2}$. Below, we will explain the intuition for these results. We first discuss the effects of increasing network density onto equilibrium activity and then onto the value of the marginal link (and both bounds of linking costs) for different q.

If $q > \frac{1}{2}$, the equilibrium activity $x^*_\Delta = (\frac{\Delta^q}{c^2})^{\frac{1}{1-q}}$ increases convexly as the equilibrium network becomes denser—i.e. Δ increases. Conversely, if $q < \frac{1}{2}$, then equilibrium activity grows concavely as the equilibrium network becomes denser.

The network density influences the value of the marginal link through two competing channels. As *cumulative* neighbour activity increases, the marginal gross payoff from neighbour activity decreases which reduces the incentive to support the marginal link. However, as the *individual* activity increases, the additional activity accessed through the marginal link becomes higher, which increases the incentive to support the marginal link.

If $q > \frac{1}{2}$ and activity grows convexly in the density of the equilibrium network, the latter channel is more prevalent and the value of the marginal link rises. Consequently, the upper bound of linking costs for which the marginal current link can be supported increases. At the same time, the lower bound preventing players from forming additional links has to increase as well.

If $q < \frac{1}{2}$ and activity grows concavely instead, the opposite holds true: the former channel is more prevalent and the value of the marginal link is depressed. Thus the upper bound of linking costs preventing players from destroying their marginal link has to lessen and the lower bound preventing them from forming additional links diminishes as well.

The next proposition illuminates the effects of these result in terms of the limit sets of supporting linking costs:[42]

Proposition 15. *The following table shows the limit sets of supporting linking costs K_∞ of complete and (non-compl.) regular interaction networks in the baseline model:*

q	complete	(non-compl.) regular
$> 2-\sqrt{2}$	$(0,\infty)$	$\subsetneq [(2^q-1)c^{\frac{1+q}{q-1}},\infty)$
$\in (\frac{1}{2}, 2-\sqrt{2}]$	$(0,\infty)$	$\subseteq [(2^q-1)c^{\frac{1+q}{q-1}},\infty)$
$= \frac{1}{2}$	$(0,\frac{1}{2}c^{-3}]$	$[(\sqrt{2}-1)c^{-3}, c^{-3}]$
$< \frac{1}{2}$	$\emptyset$	$(0, c^{\frac{1+q}{q-1}}]$

Table 7: Limit supporting sets K_∞ of selected networks (baseline model)

*Proof of **Proposition 15**.* First, we discuss non-complete regular networks. In the limit for large n, a non-complete Δ-regular network is supported in equilibrium iff $k \in [\underline{k}_\Delta, \bar{k}_\Delta]$ as $\Delta < n-1$. The limit supporting set of non-complete regular networks is the union of these intervals, i.e. $K_\infty = \cup_{\Delta=1}^{\infty}[\underline{k}_\Delta, \bar{k}_\Delta]$. Applying Lemma 3, we discuss different values of q in turn.

Let $q \in (0, \frac{1}{2})$. Then $\bar{k}_\Delta$ decreases in Δ and converges to zero. If the upper bound decreases slowly enough in comparison to the lower bound, some equilibrium exists for all $k \in (0, \bar{k}_1]$. In particular, if $\underline{k}_\Delta \leq \bar{k}_{\Delta+1}$ for all $\Delta \geq 1$ then an equilibrium with a non-complete network exists for all $k \leq \bar{k}_1 = c^{-\frac{1+q}{1-q}}$. Using the definition of $\underline{k}_\Delta$ and $\bar{k}_{\Delta+1}$ from the proof of Lemma 3, one can easily show that $\underline{k}_\Delta < \bar{k}_{\Delta+1}$ holds as $\Delta^{\frac{q^2}{1-q}} < (\Delta+1)^{\frac{q^2}{1-q}}$.

Let $q = \frac{1}{2}$. As the upper bound decreases in Δ and the lower bound increases, there is some equilibrium with iff $k \in [\underline{k}_1, \bar{k}_1] = [(\sqrt{2}-1)c^{-3}, c^{-3}]$.

[42]We employ a slight abuse of notation: as discussed in section 4.1, the interval of linking costs $[\underline{k}_\Delta, \bar{k}_\Delta]$ supporting equilibria with a non-complete Δ-regular network is independent of n (if a Δ-regular network with n players exists). Thus we define the limit supporting set K_∞ of non-complete regular networks as $\cup_{\Delta=1}^{\infty}[\underline{k}_\Delta, \bar{k}_\Delta]$.

Let $q \in (\frac{1}{2}, 1)$. As $\underline{k}_\Delta$ strictly increases in Δ and goes to infinity $K_\infty \subseteq [\underline{k}_1, \infty) = [(2^q - 1)c^{\frac{1+q}{q-1}}, \infty)$. However, if the intervals $[\underline{k}_\Delta, \bar{k}_\Delta]$ and $[\underline{k}_{\Delta+1}, \bar{k}_{\Delta+1}]$ are not overlapping then there are some $k \in [(2^q - 1)c^{\frac{1+q}{q-1}}, \infty)$ for which no equilibrium with a non-complete regular network exists. The intervals are not overlapping if $\underline{k}_{\Delta+1} > \bar{k}_\Delta$ for some $\Delta \in \{1, \dots n-2\}$ or equivalently

$$(\Delta+1)^{\frac{q^2}{1-q}}((\Delta+2)^q - (\Delta+1)^q) > \Delta^{\frac{q^2}{1-q}}(\Delta^q - (\Delta-1)^q).$$

We are going to show that the ratio $\frac{\underline{k}_{\Delta+1}}{\bar{k}_\Delta}$ converges to 1 from above if $q > 2-\sqrt{2}$ and thus $[\underline{k}_\Delta, \bar{k}_\Delta]$ and $[\underline{k}_{\Delta+1}, \bar{k}_{\Delta+1}]$ are not overlapping for Δ sufficiently large. Expending the ratio by $(\Delta+2)^{1-q}$ and Δ^{1-q} and collecting terms gives

$$\frac{\underline{k}_{\Delta+1}}{\bar{k}_\Delta} = \frac{(1+\frac{1}{\Delta})^{\frac{q^2}{1-q}}(\Delta+2)^{1-q}((\Delta+2)^q - (\Delta+1)^q)}{(1+\frac{2}{\Delta})^{1-q}\Delta^{1-q}(\Delta^q - (\Delta-1)^q)}.$$

Using similar Taylor expansions as in the proof of Lemma 3, we can write the ratio as

$$\frac{(1+\frac{q^2}{1-q}\frac{1}{\Delta} + \frac{q^2}{1-q}(\frac{q^2}{1-q}-1)\frac{1}{2!\Delta^2} + \dots)(1+\frac{1-q}{2!(\Delta+2)} + \frac{(1-q)(2-q)}{3!(\Delta+2)^2} + \dots)}{(1+(1-q)\frac{2}{\Delta} + (1-q)(-q)\frac{4}{2!\Delta^2} + \dots)(1+\frac{1-q}{2!\Delta} + \frac{(1-q)(2-q)}{3!\Delta^2} + \dots)}.$$

From this expression it is clear that the ratio converges to 1 for $\Delta \to \infty$. For large Δ, terms of order $\frac{1}{\Delta}$ determine the behaviour of the ratio and as $\Delta \approx \Delta + 2$, the ratio converges to 1 from above if $\frac{q^2}{1-q} > 2(1-q)$ or $q < \frac{\sqrt{2}}{1+\sqrt{2}} = 2 - \sqrt{2}$.

Second, we discuss equilibria with a complete network. For any fixed n the interval of supporting linking costs is $(0, \bar{k}_\Delta]$ for $\Delta = n-1$. Thus the limit sets of supporting linking costs K_∞ for large n follow directly from the limits of $\bar{k}_\Delta$ *in* Δ as provided in Lemma 3 and from $\bar{k}_\Delta$ strictly decreasing for $q = \frac{1}{2}$. □

If agents get satiated from little neighbours' activity, single-level equilibria can only be sustained for small linking costs: as the network becomes denser, equilibrium activity increases concavely and the value from the marginal link decreases.

In contrast, if agents do not get satiated easily then equilibrium activity (and the value of the marginal link) increases convexly as the network becomes denser. As a consequence, a single-level equilibria with a complete network exists for any linking costs (in sufficiently large societies).

B.6.2 Star Networks

In this section, we consider a star network with one center and $n-1$ spokes. The equilibrium conditions in Lemma 1 imply the following activity for the center x_c^* and any spoke x_s^*:

$$x_c^* = \frac{(n-1)^{\frac{q}{1-q^2}}}{c^{\frac{2}{1-q}}}$$
$$x_s^* = \frac{(n-1)^{\frac{q^2}{1-q^2}}}{c^{\frac{2}{1-q}}}.$$

As $q < 1$, we can immediately see that the center's activity is higher than any spoke's activity for $n > 2$. Thus, we know from Lemma 2 that the incentive to sponsor a link is maximised if the spoke sponsors it. Therefore, we focus subsequently on periphery sponsored stars and $n > 2$.

Let $\underline{k}_n^s$ be the lower and $\bar{k}_n^s$ be the upper bound of linking costs supporting a periphery sponsored star with $n-1$ spokes. The following lemma applies:

Lemma 4. *$\underline{k}_n^s$ and $\bar{k}_n^s$ behave as summarised below:*

q	$\underline{k}_{n+1}^s - \underline{k}_n^s$	$\lim_{n\to\infty} \underline{k}_n^s$	$\bar{k}_{n+1}^s - \bar{k}_n^s$	$\lim_{n\to\infty} \bar{k}_n^s$
$> \frac{1}{2}$	> 0	∞	> 0	∞
$= \frac{1}{2}$	> 0	$\frac{1}{2}c^{-3}$	> 0	∞
$< \frac{1}{2}$		0	> 0	∞

Table 8: Analysis of $\underline{k}_n^s$ and $\bar{k}_n^s$

*Proof of **Lemma 4**.* We start our analysis with the upper bound $\bar{k}_n^s$ which prevents spokes from deleting their link to the center. From Lemma 1 and using the definition of the baseline model, the upper bound writes as

$$\bar{k}_n^s = h(x_c^*) - h(0) = \frac{(n-1)^{\frac{q^2}{1-q^2}}}{c^{\frac{1+q}{1-q}}},$$

which converges to infinity in n.

Next notice that the center is already linked to all other players. Thus linking costs must (only) be high enough to prevent spokes from linking to other spokes and from Lemma 1 we know

$$\begin{aligned}
\underline{k}_n^s &= h(x_c^* + x_s^*) - h(x_c^*) \\
&= \frac{\left((n-1)^{\frac{q}{1-q^2}} + (n-1)^{\frac{q^2}{1-q^2}}\right)^q - (n-1)^{\frac{q^2}{1-q^2}}}{c^{\frac{1+q}{1-q}}} \\
&\propto \frac{\left(1 + (n-1)^{-\frac{q}{1+q}}\right)^q - 1}{(n-1)^{-\frac{q^2}{1-q^2}}}.
\end{aligned}$$

As both, denominator and numerator of the last expression converge to zero for $n \to \infty$, we can apply l'Hôpital's rule and get

$$\lim_{n\to\infty} \underline{k}_n^s = \lim_{n\to\infty} \{ \frac{1-q}{c^{\frac{1+q}{1-q}}} \; \frac{1}{\left(1+(n-1)^{-\frac{q}{1+q}}\right)^{1-q}} \; (n-1)^{\frac{q(2q-1)}{1-q^2}} \}.$$

The middle term strictly increases and converges to one from below. As

$$\frac{q(2q-1)}{1-q^2} \gtreqless 0 \Leftrightarrow q \gtreqless \frac{1}{2},$$

the last term strictly increases and converges to infinity for $q > \frac{1}{2}$; it strictly decreases and converges to zero for $q < \frac{1}{2}$. Together with the earlier findings, the claim follows. □

The behaviour of the upper bound and the lower bound have direct implications for the limit sets of supporting linking costs of star networks in large societies:

Proposition 16. *The following table shows the limit sets of supporting linking costs K_∞ of the star network in the baseline model:*

q	Star
$\in (\frac{1}{2}, 1)$	$\emptyset$
$= \frac{1}{2}$	$[\frac{1}{2}c^{-3}, \infty)$
$\in (0, \frac{1}{2})$	$(0, \infty)$

Table 9: Limit set K_∞ of the star network (baseline model)