IKERLANAK

EFFICIENCY OF WEIGHTED NETWORKS

by

Norma Olaizola and Federico Valenciano

2017

Working Paper Series: IL. 104/17

Departamento de Fundamentos del Análisis Económico I

Ekonomi Analisiaren Oinarriak I Saila

UPV EHU

University of the Basque Country

Efficiency of weighted networks[*]

By Norma Olaizola[‡] and Federico Valenciano[§]

November 14, 2017

Abstract

In this paper, we address the question of the efficiency of weighted networks in a setting where nodes derive utility from their direct and indirect connections. Under rather general conditions, based on a set of assumptions about the value that connections in a weighted network generate, and about link-formation technology, we prove that any network is dominated by a special type of nested split graph *weighted* network. These conditions include some of the models in the literature, which are seen as particular cases of this general model.

JEL Classification Numbers: A14, C72, D85

Key words: Weighted network, Efficiency, Nested split graph

---

[*]This paper is part of an unpublished working paper circulated with the title "A marginalist model of network formation" (Olaizola and Valenciano, 2016) that we decided to split into two separated papers. This one contains a result relative to efficiency, valid in rather general conditions. A second paper will present and study the specific marginalist model.

[†]We thank Coralio Ballester, Francis Bloch, Sebastian Bervoets, Penélope Hernández and Matthew Jackson for their helpful comments. Part of the work presented here was done while the first author was visiting the Department of Economics of Stanford University, whose hospitality and stimulating environment are gratefully acknowledged. We also thank participants in the 2nd Annual Conference on Network Science and Economics, the Bilbao-Norwich-Málaga Workshop on Economics of Networks (BiNoMa), and the 2nd Queen Mary Network Workshop. This research is supported by the Spanish Ministerio de Economía y Competitividad under projects ECO2015-66027-P and ECO2015-67519-P (MINECO/FEDER). Both authors also benefit from Basque Government Departamento de Educación, Política Lingüística y Cultura funding for Grupos Consolidados IT869-13 and IT568-13.

[‡]*BRiDGE* group (http://www.bridgebilbao.es), Departamento de Fundamentos del Análisis Económico I, Universidad del País Vasco, Avenida Lehendakari Aguirre 83, 48015 Bilbao, Spain; norma.olaizola@ehu.es.

[§]*BRiDGE* group (http://www.bridgebilbao.es), Departamento de Economía Aplicada IV, Universidad del País Vasco, Avenida Lehendakari Aguirre 83, 48015 Bilbao, Spain; federico.valenciano@ehu.es.

1 Introduction

In this paper, we address the question of efficiency of weighted networks in a setting where nodes derive utility from their direct and indirect connections[1]. Under a set of assumptions about the value that connections in a weighted network generate, and a simple notion of link-formation technology, we prove that any network is dominated by a special type of nested split graph *weighted* network[2]. These conditions include some of the models in the literature, which can be seen as particular cases of this general model. The problem addressed admits two interpretations. The most obvious and natural interpretation is as the goal of a planner looking for the efficient network, in the sense of maximizing social welfare or aggregate utility. Nevertheless, it also applies to the question of efficient structures that can arise from decentralized interaction in some connections models. The proof of this result is constructive and based on an algorithm that solves an interesting optimization problem: How to make the best of a given set of available links of different strengths. It is proved that a special class of weighted nested split graph networks form the optimizing structures. Thus, the algorithm enables a dominant weighted nested split graph network to be constructed from any initial network by optimizing the use of its links.

Nested split graph networks form a proper subclass of core-periphery networks which is highly hierarchical, covering a whole range of core-periphery degrees with different organizations of the connections between the periphery and the core, ranging from the star to the complete network[3]. As pointed out by Bramoullé and Kranton (2016), recent theoretical work has drawn attention to these structures in economic literature. In the words of Michael D. König, "*The wider applicability of nested split networks suggests that a network formation process that generates these graphs* (...) *may be of general relevance for understanding economic and social networks.*" (König, 2009, p. 69.)

Most recent work in network formation adopts a setting where agents choose a level of effort and the utility of each agent is determined by his/her effort and that of his neighbors. A very influential seminal paper in this line is Ballester, Calvó-Armengol and Zenou (2006). In contrast with this approach, the setting considered here corresponds to connections models after Jackson and Wolinsky (1996), where agents derive utility from their direct *and indirect* connections by investing in links, which are costly, but according to a link-formation technology which is not necessarily discreet as in the seminal model. It is in the first line of work that nested split graph networks emerge in the literature (König, Tessone and Zenou (2014), see also König, Battiston, Napoletano and Schweitzer (2012)). However, to the best of our knowledge ours is the first result on nested split graph (weighted) networks in a connections setting.

[1] Seminal papers in this approach are Jackson and Wolinsky (1996) and Bala and Goyal (2000a).

[2] To the best of our knowledge, this is the first economic paper where *weighted* nested split graph networks appear.

[3] König, Tessone and Zenou (2014) includes an excellent review of the related economic literature.

König, Tessone and Zenou (2014) develop a dynamic network formation model to explain the nestedness observed in real-world networks, using stochastic stability to show the convergence to nested split graphs. They also systematically study some properties of this class of networks. Hierarchical, although not nested split graph, structures where agents are ranked according the number of neighbors also emerge in equilibrium in Baetz (2014), assuming one-sided network formation, strategic complementarities and a concave value function. Lagerås and Seim (2016) apply a game with local complementarities to nested split graphs and provide a dynamic model which, under certain conditions, converges to the absorbing class of nested split graph networks. Hiller (2017) assumes strategic complementarities and positive local externalities, but assumes convexity of the value function. In this model agents choose an effort level and the set of agents with whom the agent wants to be linked, but only links desired by both agents form and in this case both share cost equally. Kinateder and Merlino (2016) study a local public good game with heterogeneous agents and endogenous network, where the equilibrium networks are empty or nested split graph networks when agents differ in their valuation of the local public good.

The question of efficient networks is addressed in the connections setting by Jackson and Wolinsky (1996) and Bala and Goyal (2000a) in their seminal models, and in Olaizola and Valenciano (2017) in a model bridging them. In all these cases, depending on the parameters of the model, only the empty, complete, and star networks can be efficient. The question of efficiency is also addressed in other settings, e.g. in Goyal and Moraga-González (2001), Goyal and Joshi (2003), Galeotti and Goyal (2010), Westbrock (2010), König, Battiston, Napoletano and Schweitzer (2012), and Billand, Bravard, Durieu and Sarangi (2015). The closest work in terms of its goal and result, although also in a completely different setting, is Belhaj, Bervoets and Deroïan (2016). They address the problem of a planner looking for the efficient network when agents play a network game with local complementarities choosing their effort levels, links are costly and the total cost is an increasing function of the sum of the individual linking types. Assuming that players choose their equilibrium levels, they show that efficient networks are nested split graph networks under different cost functions. To that end they use a reallocation procedure or "switch" guided by the Bonacich (1987) measure which, when feasible, improves the aggregate utility. They prove that any non nested split network admits an improving switch. The setting and the logic of the proof are completely different, which makes the coincidence more interesting, i.e. the optimality of nested split graph networks in two different environments.

The paper is organized as follows. Section 2 contains the notation and basic definitions. The model, i.e. the set of assumptions under which the result is to be established, is presented in Section 3. Section 4 addresses the question of efficiency, i.e. the problem of a planner who chooses the weighted network so as to maximize the aggregate payoff, and present the result. The detailed proof is relegated to the Appendix.

2 Preliminaries

An *undirected weighted N-graph* consists of a set of *nodes* $N = \{1, 2, ..., n\}$ with $n \geq 3$ and a set of *links* specified by a symmetric *adjacency matrix* $g = (g_{ij})_{i,j \in N}$ of real numbers $g_{ij} \in [0, 1)$ and $g_{ii} = 0$. Alternatively, g can, and often will, be interpreted as a map $g : N_2 \to [0, 1)$ (i.e. $g \in [0, 1)^{N_2}$), where N_2 denotes the set of all subsets of N with cardinality 2[4]. In what follows, $\underline{ij}$ stands for $\{i, j\}$ and $g_{\underline{ij}}$ for $g(\{i, j\})$ for any $\{i, j\} \in N_2$. When $g_{\underline{ij}}$ only takes the values 0 or 1, g is said to be *non-weighted*. When $g_{\underline{ij}} > 0$, it is said that a *link of weight or strength* $g_{\underline{ij}}$ connects i and j. $N^1(i, g) := \{j \in N : g_{\underline{ij}} > 0\}$ denotes the set of *neighbors* of node i, and its cardinality $|N^1(i, g)|$ is the *degree* of i. A *path p from i to j* is a sequence of distinct nodes s.t. the first one is i, the last one is j, every two consecutive nodes are connected by a link, and the path from j to i consisting of the same nodes in inverse order is denoted by p^{-1}. $N(i, g)$ denotes the set of nodes connected to i by a *path*. $\mathcal{P}_{ij}(g)$ denotes the set of paths in g from j to i. If i and j are two consecutive nodes in a path p, we write $\underline{ij} \in p$. The *length* of a path is the number of its nodes minus 1, i.e. the number of links that form it. The *distance* between two nodes connected by a path is the length of the shortest path that connects them. For any integer $k \geq 2$, $N^k(i, g)$ denotes the set of nodes connected with i at a distance of k or less. If $g_{\underline{ij}} > 0$, $g - \underline{ij}$ denotes the graph that results from eliminating link $\underline{ij}$, i.e. $g - \underline{ij} = g'$ s.t. $g'_{\underline{ij}} = 0$ and $g'_{\underline{kl}} = g_{\underline{kl}}$ for all $\underline{kl} \neq \underline{ij}$. A graph is *connected* if any two nodes are connected by a path. A *component* of a graph is a maximal connected subgraph.

Undirected graphs, weighted or not, underlie a variety of situations where actual links mean some sort of reciprocal connection or relationship. Such structures are commonly referred to as *networks*. Behind a network there is always a graph as a highly salient feature, so we transfer the notions introduced so far for graphs to networks, identifying them with the underlying graph, and refer the new ones directly to networks.

An *empty network* is one for which $g_{\underline{ij}} = 0$ for all $\underline{ij} \in N_2$. A *complete network* is one where $g_{\underline{ij}} > 0$ for all $\underline{ij} \in N_2$. An *all-encompassing star* consists of a network with $n - 1$ links in which one node (the *center*) is connected to each of the remaining nodes by a link.

An important class of networks is that in which the underlying graph is a "nested split graph". These networks exhibit a strict hierarchical structure where nodes can be ranked by their number of neighbors[5].

Definition 1 *A nested split graph (NSG) is an undirected (weighted or not) graph such that*

$$\left|N^1(i, g)\right| \leq \left|N^1(j, g)\right| \Rightarrow N^1(i, g) \subseteq N^1(j, g) \cup \{j\}.$$

[4] Alternatively, a weighted N-graph can be defined as a map $g : N_2 \to \mathbb{R}_+$, where $\mathbb{R}_+$ denotes the set of non-negative real numbers. But, w.l.o.g. we prefer to assume $g : N_2 \to [0, 1)$.

[5] König, Tessone and Zenou (2014) contains an interesting study of the topological properties of these networks.

In terms of the adjacency matrix, they have a simple structure. It is a symmetric matrix such that for a certain renumbering of the nodes, each row consists of a sequence of non-zero entries (apart from those in the main diagonal) followed by zeros, and the number of nonzero entries in each row is not greater than in the preceding row. Nodes are then classified in classes, each containing the nodes with the same number of neighbors, referred to as *NSG-classes* (isolated nodes, i.e. with no neighbors, form a class that plays no relevant role and is referred to as the *trivial class*). Fig.1 shows the adjacency matrix (a) of a nested split weighted graph with 8 nodes and 5 *NSG*-classes, while matrix (b) is that of a non-weighted nested split graph with the same non-zero entries[6].

	1	2	3	4	5	6	7	8
1	0	.4	.5	.1	.7	.9	.2	.4
2	.4	0	.2	.4	.5	.1	.8	0
3	.5	.2	0	.2	.6	0	0	0
4	.1	.4	.2	0	.4	0	0	0
5	.7	.5	.6	.4	0	0	0	0
6	.9	.1	0	0	0	0	0	0
7	.2	.8	0	0	0	0	0	0
8	.4	0	0	0	0	0	0	0

(a)

	1	2	3	4	5	6	7	8
1	0	1	1	1	1	1	1	1
2	1	0	1	1	1	1	1	0
3	1	1	0	1	1	0	0	0
4	1	1	1	0	1	0	0	0
5	1	1	1	1	0	0	0	0
6	1	1	0	0	0	0	0	0
7	1	1	0	0	0	0	0	0
8	1	0	0	0	0	0	0	0

(b)

Figure 1: Adjacency matrices of nested split graphs: (a) weighted, (b) non-weighted

3 The model

We address the question of efficiency of weighted networks in a generalized connections model setting. The problem can be stated in general terms, to be further specified, as follows: A planner invests in links between nodes in a given set in order to form an efficient network, i.e. a network that maximizes the aggregate value generated minus its cost, assuming that nodes derive their utilities from their direct and indirect connections through the network. Even if the network is the result of decentralized actions of the node-players, as might be the case, the question of efficiency can be interpreted as the goal of a planner with the ability to play the role of all nodes

[6] The example illustrates a common mistake of defining NSG by condition:

$$\left|N^1(i,g)\right| \leq \left|N^1(j,g)\right| \Rightarrow N^1(i,g) \subseteq N^1(j,g),$$

which does not hold, for instance, for $i = 3$ and $j = 1$, by a simple reason: $1 \in N^1(3,g))$ but $1 \notin N^1(1,g)$. Similarly, it is also incorrect to define NSG by condition

$$\left|N^1(i,g)\right| \leq \left|N^1(j,g)\right| \Rightarrow N^1(i,g) \cup \{i\} \subseteq N^1(j,g) \cup \{j\},$$

which in the example fails to hold, for instance, for $i = 7$ and $j = 3$.

simultaneously trying to maximize the aggregate payoff. It is assumed that forming weighted links entails a cost, and that a weighted network generates an aggregate value from the initial worth attached to each node[7]. This requires further specification.

Different connections models in the literature make different assumptions about the way in which a weighted network generates value by connecting nodes. Here we make several assumptions about it, including different alternatives for some of them, which enable us to establish a general result.

We assume that a weighted N-network g generates value under the following assumptions:

Assumption 1: Each node $i \in N$ is endowed with initial *worth* $v > 0$, which is partially perceived, received or benefitted from, by any other node $j \in N$ connected with i directly or indirectly by a path in g[8].

Assumption 2: For each $j \neq i$, if $p \in \mathcal{P}_{ij}(g)$, the *strength* of the connection between i and j *via path* p in g, denoted $\delta_i^j(p)_g$ (when no ambiguity arises, we just write $\delta_i^j(p)$), is a number in the interval $(0,1)$ which is a function of the vector of strengths of the links that form the path such that:

Assumption 2-*i*: If path p consists of a single link, i.e. $p = ij$, then $\delta_i^j(p) = g_{\underline{ij}}$.

Assumption 2-*ii*: For any path p from j to i, $\delta_i^j(p) = \delta_j^i(p^{-1})$.

Assumption 2-*iii*: For any two networks g and g', and any path $p \in \mathcal{P}_{ij}(g) \cap \mathcal{P}_{ij}(g')$ s.t. for all $kl \in p$, $g_{\underline{kl}} \geq g'_{\underline{kl}}$, $\delta_i^j(p)_g \geq \delta_i^j(p)_{g'}$ holds.

Assumption 2-*iv*: If $p = i_1 i_2 ... i_l$ and $p' = i_1 i_2 ... i_l i_{l+1} ... i_m$, with $i = i_1$, $j = i_l$ and $k = i_m$, then $\delta_i^j(p) \geq \delta_i^k(p')$.

By *Assumption* 2-*i*, when p is just a link its strength is the strength of the link. By the condition in *Assumption* 2-*ii*, the strength of the connection via any path is the same in both directions. In view of this, in what follows for any path p connecting two nodes i and j it is possible to drop i and j in $\delta_i^j(p)$ and write just $\delta(p)$[9]. *Assumption* 2-*iii* postulates a sort of "strength monotonicity": The strength of a path is non decreasing w.r.t. the increase of strengths of its links. And *Assumption* 2-*iv* postulates a sort of "length monotonicity" on any path $p \in \mathcal{P}_{ij}(g)$, the strength of the connection of i with the furthest node in that path *through that path* is *weaker* than the connection with closer nodes in that path through the corresponding subpath. Note that nothing is assumed about nodes on *different* paths. Links are the building blocks of a network. Their weight is their strength (*Assumption* 2-*i*), which admits different interpretations in different models. It can be the fraction of information/worth flowing through the link that remains intact, but other interpretations are possible. For instance they can be seen as the "strength of a tie" (Granovetter, 1973), i.e. a measure of the quality/intensity/value of a relationship, e.g. in personal relationships, where

[7] In fact, these ingredients are common in a range of models in the literature as commented below.

[8] More generally, it can be assumed that this amount v_i^j depends on the pair of nodes involved. Here we assume *homogeneity* in the sense that $v_i^j = v_j^i$, for all i, j.

[9] This intuitive simplification, i.e. undirectedness of links extends to paths, is consistent with the undirectedness of links.

the quality/strength of a link is a function of the investments of each of the two people involved. A link can also be a means for the flow of other goods[10]. We do not commit ourselves to any particular interpretation as our point is to establish a general result based on a set of assumptions.

Assumption 3: For all $p \in \mathcal{P}_{ij}(g)$, define $v_i^j(p) := v\delta(p)$, which is interpreted as the value that player i receives from j's worth endowment v via p (in fact, what i receives from j's worth in the network consisting of the links that form path p only). Then the value that i receives from j's worth in network g, denoted by $v_i^j(g)$, is the *maximal* value of $v_i^j(p)$, i.e.

$$v_i^j(g) = \max_{p \in \mathcal{P}_{ij}(g)} v_i^j(p) = v \max_{p \in \mathcal{P}_{ij}(g)} \delta(p).$$

It is understood that $v_i^j(g) = 0$ if $\mathcal{P}_{ij}(g) = \varnothing$, i.e. no path connects i and j.

Assumption 4: Now the *value* that a node i receives from a network g can be defined. Again, several options are possible. An option is the sum of the values received from *all* other nodes:

$$V_i(g) = \sum_{j \in N(i;g)} v_i^j(g). \tag{1}$$

Alternatively, if it is assumed than only nodes at k or less distance are received,

$$V_i(g) = \sum_{j \in N^k(i;g)} v_i^j(g). \tag{2}$$

Or, assuming a threshold of sensitivity, when only nodes with which the strength of the connection is at least $\sigma > 0$ are received,

$$V_i(g) = \sum_{j \in N_\sigma(i;g)} v_i^j(g). \tag{3}$$

where $N_\sigma(i;g) = \{j \in N : \delta_i^j(g) \geq \sigma\}$.

Note that which path is best, i.e. which one maximizes $\delta(p)$ and consequently $v_i^j(p)$, *depends on the definition of* $\delta_i^j : \mathcal{P}_{ij}(g) \longrightarrow [0,1)$, which it only assumed to satisfy the conditions in *Assumption* 2-*i*-*iv*. The main result is established under these conditions about function δ_i^j, without specifying it, but there are two main cases in the literature.

Accumulative friction or decay: The strength of a link is interpreted as the fraction of value that crosses it. Thus the strength of a path is given by the *product* of strengths of the links that form it:

$$\delta_i^j(p) = \prod_{\underline{kl} \in p} g_{\underline{kl}}, \tag{4}$$

and the best path is that with the smallest friction.

[10] Alternatively, the strength of a link can be interpreted as a degree of reliability, i.e. the probability of successful transmission, as in Bala and Goyal (2000b), but this interpretation is not consistent with the assumptions made here.

Weakest link strength: The strength of a path is the strength of the *weakest of the links* that form the path:

$$\delta_i^j(p) = \min\{g_{\underline{kl}} : \underline{kl} \in p\}. \tag{5}$$

It is immediately apparent that in both cases the conditions in *Assumption* 2-*i*-*iv* are satisfied.

In any case, the *aggregate value* of the network is:

$$V(g) = \sum_{i \in N} V_i(g).$$

So far the specification of the *value generated by a weighted network* through its connections from the nodes' worth. But links are costly, and we assume that a *link-formation technology* rules the formation of links according to the following definition.

Definition 2 *A link-formation technology is a non decreasing map* $\delta : \mathbb{R}_+ \to [0, 1)$ *s.t.* $\delta(0) = 0$.

The interpretation of this function is the following. If c is the amount invested in a link to connect two *nodes,* $\delta(c)$ is the *strength* of the link formed. Only links invested in form ($\delta(0) = 0$), but perfection is never reached ($\delta(c) < 1$).

We assume that a planner invests in links, according to a given technology, connecting pairs of nodes in N, and each of them is endowed with an initial worth v (*Assumption* 1). The network resulting from a link-investment vector $\underline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij} \in N_2}$ is a weighted network $g^{\underline{\mathbf{c}}} = (g^{\underline{\mathbf{c}}}_{\underline{ij}})_{\underline{ij} \in N_2}$, with $g^{\underline{\mathbf{c}}}_{\underline{ij}} = \delta(c_{\underline{ij}})$, that generates value in accordance with *Assumptions* 2-4. The *net value* of the network resulting from link-investment vector $\underline{\mathbf{c}}$ is the aggregate value generated by the network $g^{\underline{\mathbf{c}}}$ minus its cost:

$$v(g^{\underline{\mathbf{c}}}) := V(g^{\underline{\mathbf{c}}}) - \sum_{\underline{ij} \in N_2} c_{\underline{ij}} = \sum_{i \in N} V_i(g^{\underline{\mathbf{c}}}) - \sum_{\underline{ij} \in N_2} c_{\underline{ij}}, \tag{6}$$

where $V_i(g^{\underline{\mathbf{c}}})$ is given by (1), (2) or (3).

Let $\underline{\mathbf{c}}$ and $\underline{\mathbf{c}}'$ be two link-investment vectors, $g^{\underline{\mathbf{c}}}$ and $g^{\underline{\mathbf{c}}'}$ the resulting networks, and let $v(g^{\underline{\mathbf{c}}})$ and $v(g^{\underline{\mathbf{c}}'})$ be their net values as defined by (6). If $v(g^{\underline{\mathbf{c}}}) \geq v(g^{\underline{\mathbf{c}}'})$ then it is said that $g^{\underline{\mathbf{c}}}$ *dominates* $g^{\underline{\mathbf{c}}'}$ (or that $\underline{\mathbf{c}}$ *dominates* $\underline{\mathbf{c}}'$). Thus a network $g^{\underline{\mathbf{c}}}$ (or link-investment vector $\underline{\mathbf{c}}$) is *efficient* if it dominates any other[11].

Before proceeding to the result, we review some models in the literature that meet the conditions assumed.

In the seminal model of Jackson and Wolinsky (1996), the formation of a link of strength $\overline{\delta}$ ($0 < \overline{\delta} < 1$) requires an investment of a fixed amount $\overline{c} > 0$ by each of the two nodes involved. The question of efficiency is equivalent to the question of an efficient investment by a planner if the available technology is

$$\delta(c) := \begin{cases} \overline{\delta}, \text{ if } c \geq 2\overline{c} \\ 0, \text{ if } c < 2\overline{c}. \end{cases}$$

[11] This is the "strong efficiency" notion introduced by Jackson and Wolinsky (1996).

They assume $v = 1$ and accumulative friction.

Bala and Goyal (2000a): In this model links can be formed unilaterally at a fixed cost $\bar{c}$. Thus everything is similar to the previous model from the efficiency point of view, but the technology available to a planner is now $\delta(c) = \bar{\delta}$, if $c \geq \bar{c}$.

Olaizola and Valenciano (2017) is an extension of these two models where there are two types of links: Weak links, only supported by one of the two players, and strong links, supported by both players. The friction is greater through weak links. The question of efficiency in this case is equivalent to that of an efficient investment for the technology:

$$\delta(c) = \begin{cases} \bar{\delta}, \text{ if } c \geq 2\bar{c} \\ \bar{\alpha}, \text{ if } \bar{c} \leq c < 2\bar{c} \\ 0, \text{ if } c < \bar{c}, \end{cases}$$

where $0 < \bar{\alpha} < \bar{\delta} < 1$. Friction is assumed to be accumulative.

In Olaizola and Valenciano (2016) the strength of a link is a non decreasing function of the total investment $\delta : \mathbb{R}_+ \to [0, 1)$ s.t. $\delta(0) = 0$, continuously differentiable and strictly concave, and friction accumulative.

It is easy to check that in all the previous models *Assumptions* 1-4 hold.

In Bloch and Dutta (2009), if node/player i invests c_{ij} in link $\underline{ij}$, then the strength of the resulting link is $\phi(c_{ij}) + \phi(c_{ji})$, where ϕ is a non decreasing convex function s.t. $\phi(0) = 0$ and $\phi(c) < 1/2$, for all $c > 0$. Under these conditions, a planner able to play the roles of all players simultaneously would be using the technology:

$$\delta(c) = \max_{\substack{c_{ij}+c_{ji} \leq c \\ c_{ij}, c_{ji} \geq 0}} (\phi(c_{ij}) + \phi(c_{ji})) = \phi(c).$$

Where the last equality follows from ϕ's convexity. In fact, Bloch and Dutta (2009) does not specify the domain of function ϕ, which should be an interval of the form $[0, k)$ or $[0, k]$, while according to Definition 1 a technology is defined as a map $\delta : \mathbb{R}_+ \to [0, 1)$. Nevertheless, as $\delta = \phi$, which is non-decreasing and s.t. $\phi(0) = 0$ and $\phi(c) < 1/2 < 1$ on its domain ($\subsetneq \mathbb{R}_+$), this small difference is not an obstacle to applying the conclusion of our result because the conditions in *Assumptions* 1-4 hold. They consider accumulative friction and also that of the weakest link strength.

As mentioned in the Introduction, models following Ballester, Calvó-Armengol and Zenou (2006) do not fit into the conditions assumed here.

4 Efficiency of weighted networks

We now address the problem of a planner investing in links with the objective of maximizing social welfare, i.e. the aggregate value received by the nodes minus the total cost of the network, under the assumptions formulated in the previous section: A link-formation technology as specified in Definition 2 and network-generation of value in accordance with *Assumptions* 1-4.

We use the following notation. Given a link-investment vector $\underline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij} \in N_2}$, for all $i, j \in N, i \neq j$, $\overline{p}_{ij}$ denotes an *optimal* path connecting them in $g^{\underline{\mathbf{c}}}$ (note that it may be not unique), i.e. one for which the resulting strength is maximal, that is

$$\overline{p}_{ij} \in \arg \max_{p \in \mathcal{P}_{ij}(g^{\underline{\mathbf{c}}})} \delta(p) \qquad \text{i.e.} \quad \delta(\overline{p}_{ij}) = \max_{p \in \mathcal{P}_{ij}(g^{\underline{\mathbf{c}}})} \delta(p).$$

The net value of the network for this link-investment vector is then

$$v(g^{\underline{\mathbf{c}}}) = 2v \sum_{\underline{ij} \in M} \delta(\overline{p}_{ij}) - \sum_{\underline{ij} \in N_2} c_{\underline{ij}}, \tag{7}$$

where $M = N(i; g)$ if (1), $M = N^k(i; g)$ if (2), and $M = N_\sigma(i; g)$ if (3).

Note that in principle the last expression in (7) may not be unique. This occurs if the optimal path connecting a pair of nodes is not unique.

The following result shows that *under Assumptions* 1-4 *in all their variants*, and *for any link-formation technology*, any weighted network is dominated by a particular type of weighted NSG-network which exhibits stronger hierarchical features beyond those specified by Definition 1, hence the name that we have given to them: "Strongly NSG-graphs/networks" (SNSG-graphs/networks). A weighted NSG-network g, like any undirected graph, is completely specified by the triangular matrix $T(g)$ above the main diagonal of 0-entries of its adjacency matrix for a certain order of the nodes, $T(g) := (g_{ij})_{i<j}$.

Formally, the definition is as follows.

Definition 3 *A strongly nested split graph* (*SNSG*) *network is a weighted NSG-network* g *such that, for a certain renumbering of the nodes, in* $T(g)$: (*i*) *each row consists of a non-decreasing sequence of positive entries followed by zeros;* (*ii*) *all positive entries in the first row are greater than or equal to any other entries; and* (*iii*) *from the second row downwards, non-zero entries in the same column form a non-decreasing sequence.*

Fig. 2 shows the triangular matrix $T(g)$ for the adjacency matrix of an NSG and an SNSG. Both exhibit the same pattern of the non-zero entries of an NSG, but only for (b) do the conditions (i-iii) of Definition 2 hold, thus only the second is *strongly* NSG.

	1	2	3	4	5	6	7	8
1	·	.4	.5	.1	.7	.9	.2	.4
2	·	·	.2	.4	.5	.1	.8	0
3	·	·	·	.5	.2	8	0	0
4	·	·	·	·	.4	0	0	0
5	·	·	·	·	·	0	0	0
6	·	·	·	·	·	·	0	0
7	·	·	·	·	·	·	·	0
8	·	·	·	·	·	·	·	·

(*a*)

	1	2	3	4	5	6	7	8
1	·	.4	.5	.6	.7	.8	.8	.9
2	·	·	.1	.2	.3	.3	.4	0
3	·	·	·	.3	.35	4	0	0
4	·	·	·	·	.4	0	0	0
5	·	·	·	·	·	0	0	0
6	·	·	·	·	·	·	0	0
7	·	·	·	·	·	·	·	0
8	·	·	·	·	·	·	·	·

(*b*)

Figure 2: Adjacency matrices of an NSG and an SNSG

This gives the following result:

Proposition 1 *Under Assumptions* 1-4*, for any link-formation technology* δ*, any network* $g^{\underline{c}}$ *is dominated either by the empty network or by a connected SNSG-network.*

This means that efficient networks are to be found among the SNSG-networks. The constructive proof is based on the idea of rearranging the "available links" in any given network g_0 which yields a positive aggregate payoff (i.e. which is not dominated by the empty network) in the most efficient way by producing a dominant SNSG network g' s.t. $v(g') \geq v(g_0)$. A simple algorithm does the trick. The algorithm and the proof are formalized precisely in the Appendix, but given their interest we outline them briefly here[12].

Let g_0 be any given network s.t. $v(g_0) > 0$, consisting of d links. If $d \leq n-1$ it is shown that an all-encompassing star (a particular case of SNSG-network) dominates g_0. Otherwise, if $d > n-1$, *then apply the following algorithm* (see flowchart in Fig. 3): Initialize: $g := g_0$ and g' as the empty network.

Step 1: Form an all-encompassing star g' with the *strongest* $n-1$ links in g and update g by eliminating the links used to form the star g' and go to *Step* 2.

Step 2 (*Discarding procedure*): Update g by disposing of all links in g that, if used to replace the link that connects the two worst connected nodes (2 and 3) in the star g', do *not* increase the net value of the network. If no links remain (i.e. $g = \varnothing$) *Stop.* Otherwise *go to Step* 3.

Step 3 (*Improving procedure*): Take the *weakest* of the available links in g, denoted by $w(g)$, (labeled as "*Pick* $w(g)$" in the diagram of the algorithm in Fig. 3) and connect with it the two worst connected nodes (2 and 3) in star g', update g and g' (labeled as "$g' := g' + w(g)$" in the diagram). Now repeat the following procedure until no available

[12] In the proof, the algorithm is described in terms of the link-investment vectors underlying networks, rather than in terms of networks themselves. Nevertheless, the idea can easily be conveyed in terms of networks directly as we do here.

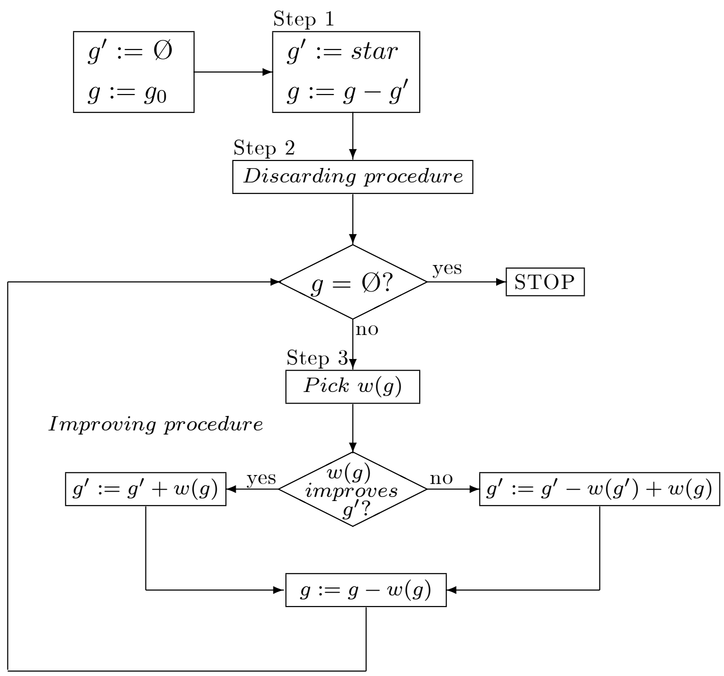

Figure 3: Flowchart of the algorithm

links are left in g: Take the weakest of the remaining available links in g, $w(g)$, and connect with it the two worst indirectly connected nodes if this improves the value of the network g' (labeled as $g' := g' + w(g)$ in the diagram). Otherwise, use it to replace the last link added to g' and proceed backwards replacing each previously added link by the one that was added immediately afterwards and dispose of the "oldest" link added, i.e. $w(g') = g'_{23}$ (labeled as $g' := g' - w(g') + w(g)$ in the diagram[13]). At the end of this process, i.e. when g is empty, g' is an SNSG-network by construction which, based on *Assumptions* 1-4, is proved to yield a net value greater than or equal to that of the initial network at a lower or equal cost.

Figure 4 shows this process, once the star has been formed in *Step* 1 with the four strongest links, starting from a 5-node network with 8 links, assuming that at every stage, but perhaps at the last one, the weakest available link actually improves the net value of the current network by connecting the two worst connected nodes. Figure 5 shows the same process in terms of the adjacency matrix. Assume that the investments in the 8 links of the initial network are $c_1 \geq c_2 \geq c_3 \geq c_4 \geq c_5 \geq c_6 \geq c_7 \geq c_8$. First (*Step* 1), form a star with the best 4 links in $\underline{\mathbf{c}}$ in *increasing* order of strength with the 4 strongest links: $(c_1, c_2, c_3$ and $c_4)$ (stage (a) in Figures 4 and 5); then, using available

[13]Note that this procedure leads to discarding the worst link in g', i.e. $w(g') = g'_{23}$, and adds to g' the worst in g, $w(g)$ (rearranging all added links in the way described).

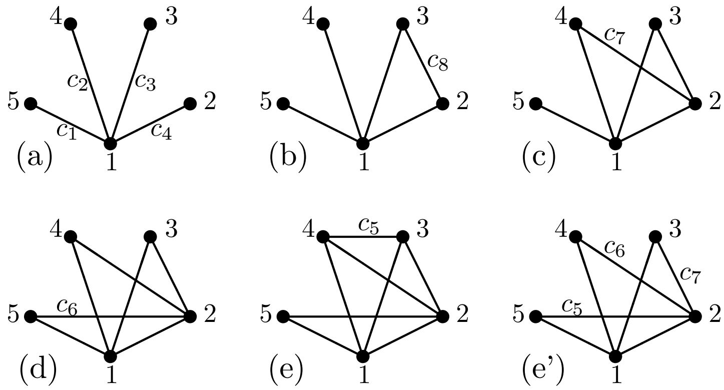

Figure 4: Construction of a dominant SNSG-network

links in *increasing* order of strength, improve the weakest connection if this increases the net value of the network, i.e. 23 in the first iteration, with a c_8-link, and then 24 with a c_7-link, (stages (b)-(c) in Figures 4 and 5). Then, connect the pair worst connected after stage (c), i.e. 25 or 34. If it is 25, i.e. if $\delta(512) \leq \delta(314)$, then connect them with a c_6-link (stage (d)). Finally, connect the worst connected nodes after stage (d), which are nodes 3 and 4 with a c_5-link (stage (e)), if this increases the net value of the network. Otherwise, if connecting 3 and 4 with a c_5-link does *not* increase the net value of the network, replace c_6 by c_5, c_7 by c_6 and c_8 by c_7, and dispose of c_8 (stage (e')). Observe that, by construction, *at every stage* the current network in formation is an SNSG-network.

$$
\underset{(a)}{\begin{array}{c|ccccc} & 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 0 & c_4 & c_3 & c_2 & c_1 \\ 2 & c_4 & 0 & 0 & 0 & 0 \\ 3 & c_3 & 0 & 0 & 0 & 0 \\ 4 & c_2 & 0 & 0 & 0 & 0 \\ 5 & c_1 & 0 & 0 & 0 & 0 \end{array}}
\rightarrow
\underset{(b)}{\begin{array}{c|ccccc} & 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 0 & c_4 & c_3 & c_2 & c_1 \\ 2 & c_4 & 0 & c_8 & 0 & 0 \\ 3 & c_3 & c_8 & 0 & 0 & 0 \\ 4 & c_2 & 0 & 0 & 0 & 0 \\ 5 & c_1 & 0 & 0 & 0 & 0 \end{array}}
\rightarrow
\underset{(c)}{\begin{array}{c|ccccc} & 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 0 & c_4 & c_3 & c_2 & c_1 \\ 2 & c_4 & 0 & c_8 & c_7 & 0 \\ 3 & c_3 & c_8 & 0 & 0 & 0 \\ 4 & c_2 & c_7 & 0 & 0 & 0 \\ 5 & c_1 & 0 & 0 & 0 & 0 \end{array}}
$$

$$
\rightarrow
\underset{(d)}{\begin{array}{c|ccccc} & 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 0 & c_4 & c_3 & c_2 & c_1 \\ 2 & c_4 & 0 & c_8 & c_7 & c_6 \\ 3 & c_3 & c_8 & 0 & 0 & 0 \\ 4 & c_2 & c_7 & 0 & 0 & 0 \\ 5 & c_1 & c_6 & 0 & 0 & 0 \end{array}}
\rightarrow
\underset{(e)}{\begin{array}{c|ccccc} & 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 0 & c_4 & c_3 & c_2 & c_1 \\ 2 & c_4 & 0 & c_8 & c_7 & c_6 \\ 3 & c_3 & c_8 & 0 & c_5 & 0 \\ 4 & c_2 & c_7 & c_5 & 0 & 0 \\ 5 & c_1 & c_6 & 0 & 0 & 0 \end{array}}
\text{ or }
\underset{(e')}{\begin{array}{c|ccccc} & 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 0 & c_4 & c_3 & c_2 & c_1 \\ 2 & c_4 & 0 & c_7 & c_6 & c_5 \\ 3 & c_3 & c_7 & 0 & 0 & 0 \\ 4 & c_2 & c_6 & 0 & 0 & 0 \\ 5 & c_1 & c_5 & 0 & 0 & 0 \end{array}}
$$

Figure 5: Construction of a dominant SNSG-network

Remark: For a certain relabeling of the nodes, an SNSG-network can be seen as consisting of a star centered on node 1 (first row and column of the adjacency matrix) plus some additional links between spoke nodes of that star worst connected through it (nonzero entries on the northwest of the adjacency matrix) according to the pattern specified in Definition 3. But the constructive proof of Proposition 1 shows that a dominant SNSG exhibits further features beyond that of a pure SNSG structure in general. Namely, apart from the pattern described in Definition 3, it must hold $2v\delta(c_{\underline{ij}}) - c_{\underline{ij}} > 2v\delta(i1j)$, for all $i, j \in \{2, 3, ...n\}$ s.t. $c_{\underline{ij}} > 0$, i.e. their direct connection should improve the contribution to the net value of the network of their indirect connection through 1. For instance, under accumulative friction, for an investment matrix SNSG to be efficient $2v\delta(c_{\underline{ij}}) - c_{\underline{ij}} > 2v\delta(c_{\underline{1i}})\delta(c_{\underline{1j}})$ must hold for all $c_{\underline{ij}} > 0$, with $i, j \neq 1$.

5 Appendix

Proof of Proposition 1: Let δ be a link formation technology according to Definition 2, and let $\underline{\mathbf{c}}_0 = (c_{\underline{ij}})_{\underline{ij} \in N_2}$ be a link-investment vector s.t. $g^{\underline{\mathbf{c}}_0}$ has d links and positive net value (if it were negative it would be dominated by the empty network). If $d \leq n-1$ rearrange the d available links as a star. By *Assumption* 2-*iv*, the resulting star-network dominates $g^{\underline{\mathbf{c}}_0}$. If $d = n-1$, the star is all-encompassing and consequently a connected SNSG-network, and we are done. Otherwise, if $d < n-1$, form an all-encompassing star as many links as necessary by adding to the d-link star, i.e. $n-1-d$, using links with the same investment as any of those in the d-link star whose marginal contribution is positive (note that there must be at least one).

Assume now that $d > n-1$. Without loss of generality, it can be assumed that in $g^{\underline{\mathbf{c}}_0}$ no link invested in is superfluous, i.e. $\delta(\overline{p}_{ij}) = \delta(c_{\underline{ij}}) = g_{\underline{ij}} > 0$ whenever $c_{\underline{ij}} > 0$. By rearranging the links that form the network, i.e. by reassigning the amounts invested in each of the d links in a different set of d links, only the sum $\sum_{\underline{ij} \in M} \delta(\overline{p}_{ij})$ in (7) would change. As shown below, by reassigning the amounts invested in the links of the network, perhaps even disposing of some of them, it is always possible to obtain an SNSG-network that dominates $g^{\underline{\mathbf{c}}_0}$. More precisely, starting from an arbitrary link-investment $\underline{\mathbf{c}} := \underline{\mathbf{c}}_0$, with more than $n-1$ positive entries, we describe an algorithm for constructing a new link-investment vector, $\underline{\mathbf{c}}' = (c'_{\underline{ij}})_{\underline{ij} \in N_2}$, that yields an SNSG-network $g^{\underline{\mathbf{c}}'}$ that dominates $g^{\underline{\mathbf{c}}_0}$ as the final outcome of a sequence of link-investment vectors, each of them resulting from the preceding one by adding at most one link and perhaps reassigning those introduced so far after Step 1 and disposing of the cheapest of them.

Step 1: Let $\underline{\mathbf{c}}'$ be the link-investment that yields the all-encompassing star that results from connecting node 1 with the other $n-1$ by investing in each link exactly the same amount invested in each of the *strongest* $n-1$ links in $g^{\underline{\mathbf{c}}_0}$, so that $c'_{\underline{12}} \leq c'_{\underline{13}} \leq ... \leq c'_{\underline{1n-1}} \leq c'_{\underline{1n}}$. Update $\underline{\mathbf{c}}$ by eliminating in $\underline{\mathbf{c}}_0$ the investments corresponding to the $n-1$ strongest links, and *go to Step* 2.

$$\begin{array}{c|ccccc} & 1 & 2 & 3 & \dots & n \\ \hline 1 & 0 & c'_{\underline{12}} & c'_{\underline{13}} & \dots & c'_{\underline{1n}} \\ 2 & c'_{\underline{12}} & 0 & 0 & \dots & 0 \\ 3 & c'_{\underline{13}} & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ n & c'_{\underline{1n}} & 0 & 0 & \dots & 0 \end{array} \quad \rightarrow \quad \begin{array}{c|ccccc} & 1 & 2 & 3 & \dots & n \\ \hline 1 & 0 & c'_{\underline{12}} & c'_{\underline{13}} & \dots & c'_{\underline{1n}} \\ 2 & c'_{\underline{12}} & 0 & c'_{\underline{23}} & \dots & 0 \\ 3 & c'_{\underline{13}} & c'_{\underline{23}} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ n & c'_{\underline{1n}} & 0 & 0 & \dots & 0 \end{array}$$

Figure 6: Step 1 and first iteration of Step 3

Step 2 (*Discarding procedure*): Update $\underline{\mathbf{c}}$ by setting to 0 all entries in $\underline{\mathbf{c}}$ whose investment in the worst connected nodes in the star $\underline{\mathbf{c}}'$ (nodes 2 and 3) would *not* increase the net value of the network. If no links remain (i.e. $\underline{\mathbf{c}} = \mathbf{0}$) *Stop*. Otherwise *go to Step* 3.

Step 3 (*Improving procedure*): Now proceed as follows. Choose two of the nodes, i and j, which are the worst connected in $g^{\underline{\mathbf{c}}'}$, in the first iteration nodes 2 and 3 (Figure 3). Then, to avoid ambiguity if there are multiple equally worst connected pairs, choose any of them with the sole condition of preserving the NSG underlying structure if a link to connect them is added[14], and check whether the contribution to the net value of the network of the connection of i and j via node 1 in $g^{\underline{\mathbf{c}}'}$ can be improved by connecting them with the *worst* available link, i.e. investing the smallest entry, say $c_{\underline{kl}}$, in $\underline{\mathbf{c}}$. That is, check whether

$$2v\delta(\overline{p}'_{ij}) < 2v\delta(c_{\underline{kl}}) - c_{\underline{kl}}.$$

- If there is an improvement (which is certain to be in the first iteration), then connect them, i.e. update $\underline{\mathbf{c}}'$ by adding entry $c'_{\underline{ij}} := c_{\underline{kl}}$, and update $\underline{\mathbf{c}}$ by eliminating its smallest entry, i.e. setting its $\underline{kl}$-entry to 0.

- Otherwise, by construction (available links in $\underline{\mathbf{c}}$ are picked in *increasing* order of strength) $c_{\underline{kl}}$ is necessarily at least as good as any link previously added. Then replace the investment in the last link added by $c_{\underline{kl}}$ and proceed similarly, replacing the previously added connection by the newly available link, and so on backwards.[15] This procedure leads to the weakest of the added links, currently connecting nodes 2 and 3, being discarded. Update $\underline{\mathbf{c}}$ after this elimination, and let $\underline{\mathbf{c}}'$ be the updated investment vector.

[14] That is, in terms of the triangular submatrix of the adjacency matrix above the main diagonal, give priority to a new investment in a link ij as far as possible to the left among those in the same row and to the uppermost among those in the same column.

[15] Note that for this we just need to keep track of the order in which new links were added to the star formed in *Step* 1. If the investments in the links added so far to the initial star currently were $c'_{i_1} \leq c'_{i_2} \leq ... \leq c'_{i_r}$, with $c'_{i_1} = c'_{23}$, then upgrade the investments in the links added previously by replacing c'_{i_r} by $c_{\underline{kl}}$, $c'_{i_{r-1}}$ by c'_{i_r}, $c'_{i_{r-2}}$ by $c'_{i_{r-1}}$, etc., and dispose of the weakest link, corresponding to $c'_{i_1} = c'_{23}$.

In all cases, go back to *Step* 3 unless $\underline{\mathbf{c}}$ is empty (i.e. no available links remain), if $\underline{\mathbf{c}}$ is empty then *Stop*.

Obviously the process ends in finite iterations of *Step* 3, when $\underline{\mathbf{c}}$ is the zero-investment vector and no available links remain.

We show now that $v(g^{\underline{\mathbf{c}}'}) \geq v(g^{\underline{\mathbf{c}}})$. Both $v(g^{\underline{\mathbf{c}}})$ and $v(g^{\underline{\mathbf{c}}'})$ are the sum of at most $n(n-1)/2$. Each of these terms corresponds to one pair of nodes whose connection contributes to the net value of the network, and gives its contribution, i.e. the value received from each other (minus the cost of the link if they are directly connected).

Let D (R) and D' (R') denote the sets of pairs of nodes connected directly (indirectly) in $g^{\underline{\mathbf{c}}}$ and $g^{\underline{\mathbf{c}}'}$, and let d (r), d' (r') be their cardinalities. Then, decomposing the contribution to the net value of pairs directly and indirectly connected, (7) can be rewritten for $g^{\underline{\mathbf{c}}}$ and $g^{\underline{\mathbf{c}}'}$ as

$$v(g^{\underline{\mathbf{c}}}) = \sum_{\underline{ij} \in D} 2v\delta(c_{\underline{ij}}) + 2v \sum_{\underline{ij} \in M(R)} \delta(\overline{p}_{ij}) - \sum_{\underline{ij} \in D} c_{\underline{ij}},$$

$$v(g^{\underline{\mathbf{c}}'}) = \sum_{\underline{ij} \in D'} 2v\delta(c'_{\underline{ij}}) + 2v \sum_{\underline{ij} \in M(R')} \delta(\overline{p}'_{ij}) - \sum_{\underline{ij} \in D'} c'_{\underline{ij}},$$

where $M(R)$ is R if (1), $R \cap N^k(i;g)$ if (2), or $R \cap N_\sigma(i;g)$ if (3). The first term of the sum that yields $v(g^{\underline{\mathbf{c}}})$ is the sum of d terms $\neq 0$, corresponding to its d links, while in the first term of the sum that yields $v(g^{\underline{\mathbf{c}}'})$ there may be *fewer* terms if any link has been discarded in *Step* 2 because it does not improve any indirect connection through node 1, or in *Step* 3 when the weakest link in g' connecting 2 and 3 is replaced and discarded. That is, $d' \leq d$. Moreover, for each term in this sum for $g^{\underline{\mathbf{c}}'}$, a term with exactly the same value occurs in the sum for $g^{\underline{\mathbf{c}}}$. Note that if a link that received an investment $c > 0$ in $g^{\underline{\mathbf{c}}}$ is discarded in *Step* 2, then $2v\delta(c) - c \leq 2v\delta(\overline{p}'_{ij})$ *for all* $i, j = \{2, .., n\}$, where $\overline{p}'_{ij}$ is the 2-link path $i1j$. The same occurs if it was disposed of in *Step* 3. Therefore any term in the sum that yields $v(g^{\underline{\mathbf{c}}})$ corresponding to these c's is outweighed by *all* indirect connections in $g^{\underline{\mathbf{c}}'}$, therefore their sum is outweighed by the sum of the same number $(d - d')$ of *weakest* indirect connections in $g^{\underline{\mathbf{c}}'}$. Then, if B' denotes the set of best $\frac{n(n-1)}{2} - (d - d')$ indirect connections in $g^{\underline{\mathbf{c}}'}$, as $\sum_{\underline{ij} \in N_2} c_{\underline{ij}} \geq \sum_{\underline{ij} \in N_2} c'_{\underline{ij}}$, we have

$$\begin{aligned} v(g^{\underline{\mathbf{c}}'}) - v(g^{\underline{\mathbf{c}}}) &\geq \sum_{\underline{ij} \in M(B')} 2v\delta(\overline{p}'_{ij}) - \sum_{\underline{ij} \in M(R)} 2v\delta(\overline{p}_{ij}) + \sum_{\underline{ij} \in N_2} c_{\underline{ij}} - \sum_{\underline{ij} \in N_2} c'_{\underline{ij}} \\ &\geq \sum_{\underline{ij} \in M(B')} 2v\delta(\overline{p}'_{ij}) - \sum_{\underline{ij} \in M(R)} 2v\delta(\overline{p}_{ij}) \geq 0. \end{aligned}$$

The last inequality holds because B' contains the *best* $\frac{n(n-1)}{2} - (d - d')$ paths of length 2 that can be formed with links in the collection $\{\delta(c_{\underline{ij}})\}_{\underline{ij} \in N_2}$, in other words, the strongest $\frac{n(n-1)}{2} - (d - d')$ indirect connections that can be formed with the available links. However any indirect connection in $g^{\underline{\mathbf{c}}}$ consists of a path of length at least 2, and consequently of strength no greater than the strength of the path formed by its first two links (by "length monotonicity" *Assumption* 2-*iv*). But by *Assumption* 2-*iii* ("strength monotonicity"), these in turn cannot outweigh those in B'. Now if $M(B') \subsetneq B'$ and/or $M(R) \subsetneq R$, again *Assumption* 2-*iii*-*iv* ensures the last inequality. Finally, as links are added to the initial star (formed by the strongest links) in increasing order of strength and always preserving an SNSG structure, new links added corresponding to entries in the same row (column) in the triangular matrix are of increasing strength rightward (downward). Therefore the outcome is a connected SNSG-network. Thus, any network which yields a positive aggregate payoff is dominated by a connected SNSG-network.■

References

[1] Baetz, O., 2015, Social activity and network formation, *Theoretical Economics* 10, 315-340.

[2] Bala, V., and S. Goyal, 2000a, A non-cooperative model of network formation, *Econometrica* 68, 1181-1229.

[3] Bala, V., and S. Goyal, 2000b, A strategic analysis of network reliability, *Review of Economic Design* 5, 205-228.

[4] Ballester, C., A. Calvó-Armengol and Y. Zenou, 2006, Who's who in networks. Wanted: The key player, *Econometrica* 74, 1403–1417.

[5] Belhaj, M., S. Bervoets and F. Deroïan, 2016, Efficient networks in games with local complementarities, *Theoretical Economics* 11, 357-380.

[6] Billand, P., C. Bravard, J. Durieu and S. Sarangi, 2015, Efficient networks for a class of games with global spillovers, *Journal of Mathematical Economics* 61, 203-310.

[7] Bloch, F., and B. Dutta, 2009, Communication networks with endogenous link strength, *Games and Economic Behavior* 66, 39-56.

[8] Bonacich, P., 1987, Power and centrality: A family of measures, *American Journal of Sociology* 92, 1170-1182.

[9] Bramoullé, Y. and R. Kranton, 2016, Games played in networks, *in* Bramoullé, Rogers and Galeotti, Eds., 2016, 83-112.

[10] Bramoullé, Y., B.W. Rogers and A. Galeotti, Eds., 2016, *The Oxford Handbook on the Economics of Networks*, Oxford University Press, Oxford.

[11] Galeotti, A. and S. Goyal, 2010, The Law of the Few, *American Economic Review* 100, 1468-1492.

[12] Goyal, S. and S. Joshi, 2003, Networks of collaboration in oligopoly, *Games and Economic Behavior* 43, 57-85.

[13] Goyal, S. and J.L. Moraga-González, 2001, R&D Networks, *RAND Journal of Economics* 32, 686-707.

[14] Granovetter, M.S., 1973, The strength of weak ties, *American Journal of Sociology* 78(6), 1360-1380.

[15] Hiller, T., 2017, Peer Effects in Endogenous Networks, *Games and Economic Behavior* 105, 349-367.

[16] Jackson, M., and A. Wolinsky, 1996, A strategic model of social and economic networks, *Journal of Economic Theory* 71, 44-74.

[17] Kinateder, M. and L.P. Merlino, 2016, Public Goods in Endogenous networks, Working Paper.

[18] König, M. D., 2009, *Dynamic and R&D Networks. The efficiency and evolution of interfirm collaboration networks.* Thesis Dissertation, University of Zurich.

[19] König, M. D., S. Battiston, M. Napoletano and F. Schweitzer, 2012, The efficiency and stability of R&D networks, *Games and Economic Behavior* 75, 694-713.

[20] König, M. D., C. J. Tessone and Y. Zenou, 2014, Nestedness in networks: A theoretical model and some applications, *Theoretical Economics* 9, 695-752.

[21] Lagerås, A.N., and D. Seim, 2016, Strategic complementarities, network games and endogenous network formation, *International Journal of Game Theory* 3, 497-509.

[22] Olaizola, N., and F. Valenciano, 2016, A marginalist model of network formation. *Ikerlanak Working Paper Series*: IL 99/16, Department of Foundations of Economic Analysis I, University os the Basque Country.

[23] Olaizola, N., and F. Valenciano, 2017, A unifying model of strategic network formation. *International Journal of Game Theory* (forthcoming), https://doi.org/10.1007/s00182-017-0592-8.

[24] Westbrock, B., 2010, Natural concentration in industrial research collaboration, *RAND Journal of Economics* 41, 351-371.