Efficiency and stability in the connections model with heterogeneous node

Olaizola, Norma and Valenciano, Federico

University of the Basque Country UPV/EHU

17 May 2021

Efficiency and stability in the connections model with heterogeneous nodes*

By Norma Olaizola† and Federico Valenciano‡

May 17, 2021

Abstract

This paper studies the connections model (Jackson and Wolinsky, 1996) when nodes may have different values. It is shown that efficiency is reached by a strongly hierarchical structure that we call strong NSG-networks: Nested Split Graph networks where the hierarchy or ranking of nodes inherent in any such network is consistent with the rank of nodes according to their value, perhaps leaving some of the nodes with the lowest values disconnected. A simple algorithm is provided for calculating these efficient networks. We also introduce a natural extension of pairwise stability assuming that players are allowed to agree on how the cost of each link is split and prove that stability in this sense for connected strong NSG-networks entails efficiency.

JEL Classification Numbers: A14, C72, D85

Key words: Networks, Connections model, Heterogeneity, Efficiency, Stability.

*We thank the editor and two anonymous referees for their comments. This research is supported by the Spain's Ministerio de Ciencia, Innovación y Universidades under projects PID2019-106146GB-I00 and PID2019-107539GB-I00. Both authors also benefit from Basque Government Departamento de Educación, Política Lingüística y Cultura funding for Grupos Consolidados IT1367-19.

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

‡*BRiDGE* group (http://www.bridgebilbao.es), Departamento de Métodos Cuantitativos, Universidad del País Vasco UPV/EHU, Avenida Lehendakari Aguirre 83, 48015 Bilbao, Spain; federico.valenciano@ehu.es.

1 Introduction

This paper is a further step in a project that has been ongoing for some years now, building on the seminal papers of Jackson and Wolinsky (1996) and Bala and Goyal (2000).[1] The objective is to explore different extensions of some of the models introduced in those papers in order to test the robustness of some of their results when the setting is enriched in different directions. This paper focuses on an extension of the connections model of Jackson and Wolinsky (1996) by introducing heterogeneity in the values of nodes. In Jackson and Wolinsky's connections model, nodes can create links of a unique type, i.e. of a given decay factor, δ $(0 < \delta < 1)$, and a given cost $c > 0$, which must be equally shared by the two nodes involved. Each node is endowed with information of a given value $v > 0$, which is the same for all nodes, and can be partially accessed through the network.

In this paper we introduce node heterogeneity into that model, i.e. nodes are not assumed to be endowed with the same value, and study efficiency and stability in this setting. We first address and solve thoroughly the question of efficiency in the connections model with heterogeneous nodes, which to the best of our knowledge is an open issue. The noteworthy result is that when nodes have different values, efficiency is reached by a doubly hierarchical structure: The greater the value of a node the greater the number of neighbors, perhaps leaving some of the nodes with the smallest values disconnected. More precisely, efficiency is reached by a special type of nested split graph (NSG-) networks that we have called "*strong* nested split graph networks" or SNSG-networks, where the hierarchy or ranking of nodes inherent in any NSG-network is consistent with the rank of nodes according to their value. When preserving connectedness is a constraint, it is also proved that the optimal connected network is a connected SNSG-network. It is also proved that for a certain range of values of parameters c and δ, an SNSG-network, which is neither complete nor a star network, is sure to be efficient. In addition, a simple algorithm is provided for calculating efficient networks, given the values of the nodes and parameters c and δ. These results hold independently of how the network forms, be it designed by a planner or by node-players in a decentralized way under any assumptions whatsoever. They also solve thoroughly the question of efficiency in the connections model with heterogeneous nodes, an issue that has remained open for 25 years.

The question of stability in a decentralized context is addressed by introducing a natural extension of pairwise stability (Jackson and Wolinsky, 1996). Instead of assuming as Jackson and Wolinsky (1996) do, that the cost of each link must be equally shared by the two nodes connected by that link, we assume that upon forming a link players are able to agree upon how its cost is split, and adapt Jackson and Wolinsky's pairwise stability conditions to this scenario. Necessary and sufficient conditions for the stability in this sense of an SNSG-network are obtained, leading to the result that stability in this sense for connected SNSG-networks entails efficiency, and is generically

[1] Olaizola and Valenciano (2019, 2020a,b,c).

equivalent to it when no node has the central node with the highest value as its only neighbor. The relation between this notion and that of pairwise stability in the sense of Jackson and Wolinsky (1996) is clear: Classical pairwise stability is the particular case of pairwise stability under free cost sharing once cost sharing is no longer free and only a particular type of investment profiles is considered admissible (those where the cost of each link is split equally). Why this should be so? There are two possible interpretations or justifications of Jackson and Wolinsky's simplifying assumption, one technological and the other legal. The technological interpretation is not very convincing and is not usually found in real world examples, where free cost sharing is more plausible. If it were imposed by law, pairwise stable networks could become unstable because players may have incentives to make payments under the table. To compare the two notions of stability in the two senses, they are applied in both heterogeneous and homogeneous settings.

The paper is organized as follows. Section 2 reviews some related literature, Section 3 presents the model, Section 4 addresses efficiency, and Section 5 stability. Section 6 concludes, summing up and pointing out some lines for further work. All proofs are relegated to an Appendix.

2 Related literature

There is a considerable body of research on network formation.[2] In this brief review, we concentrate on papers that explore different extensions of the seminal models of Jackson and Wolinsky (1996) and Bala and Goyal (2000), where nodes receive utility from direct and indirect connections, particularly those dealing with different forms of heterogeneity.

Haller and Sarangi (2005) consider link heterogeneity in a model where links fail with different probabilities. Galeotti (2006) studies a one-way flow connections model in which players are heterogeneous with respect to values and the costs of forming links. Galeotti, Goyal and Kamphorst (2006) study an extension of the connections model allowing for heterogeneity in values and costs of forming links, although most of their results are without decay. They consider low levels of decay and focus on the case of two groups. Their principal finding is that high centrality and small average distances are salient properties of equilibrium networks while center-sponsorship is not a robust feature of equilibrium networks in the presence of decay. Kamphorst and Van Der Laan (2007) investigate a model of network formation where players are divided into groups and the costs of a link between any pair of players are increasing with the distance between the groups to which those players belong. Billand, Bravard and Sarangi (2012a) examine the existence of Nash networks in Bala and Goyal two-way flow model in the presence of partner heterogeneity. They show that Nash networks in

[2] Goyal (2007), Jackson (2008), and Vega-Redondo (2007) provide excellent overviews. See also Bramoullé, Galeotti and Rogers (2016).

pure strategies do not always exist in such model. They then impose restrictions on the payoff function to find conditions under which Nash networks always exist. Billand, Bravard and Sarangi (2012b) study Nash networks in the connections model under heterogeneity in links and values. They prove that under heterogeneity in values or decay involving only two degrees of freedom, all networks can be supported as Nash for some values of the parameters, but they show also that Nash networks may not always exist and that when heterogeneity is reduced both the earlier "anything goes" result and the non-existence problem disappear. Vandenbossche and Demuynck (2010) present a model of endogenous network formation with heterogeneous agents whose payoffs are determined by agent specific utility functions that depend on the number of direct links, and show that the cost of a link depends on the social distance and the value of the agents. Goyal (2018) discusses how heterogeneous networks give rise to very widely varying forms of behavior and potentially significant inequality, and how heterogeneous network structures are a natural outcome in a wide range of circumstances.

In a different approach, Souma, Fujiwara and Aoyama (2005) study the Japanese shareholding network at the end of March 2002 and suggest the existence of a hierarchical structure. They find that degree and company total assets correlate strongly. Goeree, Riedl and Ulle (2009) reports results from a laboratory experiment on network formation among heterogeneous agents. Their setting extends the Bala and Goyal two-way flow model, introducing agents with lower linking costs or higher benefits to others. They find that equilibrium predictions fail with homogeneous agents, while heterogeneity fosters network centrality, stability, and efficiency.

Nested split graph structures appear in the economic literature in Konig, Teasone and Zenou (2014), where, in a dynamic model in the wake of the highly influential paper by Ballester, Calvó-Armengol and Zenou (2006), they show that if individuals form links to maximize centrality, then the linking process leads to nested split graph networks in terms of stochastic stability.[3] To the best of our knowledge, Olaizola and Valenciano (2020a) contains the first result involving nested split graph networks in the much simpler setting of connections models. Here also these structures play a central role in a natural extension of Jackson and Wolinsky's (1996) seminal setting.

Last but not least, Bloch and Jackson (2006) consider a "linking game with transfers" which is very similar to the free cost sharing scenario that we consider here. In Section 5 we discuss in detail the relation between these two models as well as the relation between their pairwise Nash stability with transfers and our pairwise stability under free cost sharing.

[3] See Mariani et al. (2019) for an excellent overview focused on nestedness in networks, its emergence and implications in social, economic, and ecological contexts.

3 The model

$N = \{1, 2, ..., n\}$ $(n \geq 3)$ is a set of *nodes*. Each node $i \in N$ is endowed with one piece of information of *value* $v_i > 0$. Nodes can be connected by only one type of undirected link of cost $c > 0$ and strength or decay factor δ $(0 < \delta < 1)$. Once parameters c and δ are set, a network g can be specified by an $n \times n$ symmetric matrix where $g_{ij} = 1$ if there is an undirected link connecting i and j, and $g_{ij} = 0$ otherwise. If $g_{ij} = 1$, ij denotes the link between i and j, and we write $ij \in g$. If $g_{ij} = 0$, we write $ij \notin g$. Nevertheless, a network g can also be specified by the triangular matrix $\widehat{g} = (g_{ij})_{(i,j)\in\mathcal{T}}$, where $\mathcal{T} = \{(i,j) : 1 \leq i < n \;\&\; i < j \leq n\}$, with $n(n-1)/2$ entries, corresponding to the $n(n-1)/2$ pairs of nodes, that is

$$\widehat{g} = \left[\begin{array}{cccccccc}
g_{12} & g_{13} & g_{14} & \cdots & g_{1j} & \cdots & g_{1,n-1} & g_{1n} \\
 & g_{23} & g_{24} & \cdots & g_{2j} & \cdots & g_{2,n-1} & g_{2n} \\
 & & g_{34} & \cdots & g_{3j} & \cdots & g_{3,n-1} & g_{3n} \\
 & & & \ddots & \vdots & & & \vdots \\
 & & & & g_{j-1,j} & \cdots & \vdots & g_{j-1,n} \\
 & & & & & \ddots & & \vdots \\
 & & & & & & g_{n-2,n-1} & g_{n-2,n} \\
 & & & & & & & g_{n-1,n}
\end{array}\right].$$

Properly speaking, $\widehat{g} = (g_{ij})_{(i,j)\in\mathcal{T}}$ specifies the only possibly non-zero entries of an actual $(n-1) \times (n-1)$ triangular matrix, and contains the precise description of the network without redundancies. In order to deal with efficiency we opt for this representation instead of the $n \times n$ redundant symmetric matrix g, paying attention only to the entries above the main diagonal of zeros in that matrix. That is why in the next section we always show only the relevant entries, i.e. those g_{ij} with $(i,j) \in \mathcal{T}$.

Each node receives information from all nodes to which it is connected by a path. Two nodes are connected by a *path* if they are the first and last of a sequence of distinct nodes where every two consecutive nodes are connected by a link. The information received by node i in network g is

$$I_i(g) := \sum_{j\in N\setminus\{i\}} \delta^{d(i,j)} v_j,$$

where $d(i,j)$ is the geodesic distance between i and j, which is assumed to be ∞ if there is no path connecting them, so that $\delta^\infty = 0$. The *net value* of a network g, denoted by $v(g)$, is the sum of the information received by all the nodes minus its total cost, i.e.

$$v(g) := \sum_{(i,j)\in\mathcal{T}} \delta^{d(i,j)}(v_i + v_j) - C(g), \tag{1}$$

where $C(g)$ is the cost of the network, i.e. $\eta(g)c$, if $\eta(g)$ is the number of links of network g. Sums $v_i + v_j$ play a crucial role in what follows, particularly in the proofs, so it is convenient to use the following notation:

$$s_{ij} := v_i + v_j.$$

Then (1) can be rewritten like this:

$$v(g) := \sum_{(i,j)\in\mathcal{T}} \delta^{d(i,j)} s_{ij} - C(g). \tag{2}$$

A network is *connected* if any two nodes are connected by a path. A *component* of a network is a maximal connected subnetwork. A *trivial* component is one that consists of an isolated node.

4 Efficiency

A network g *dominates* a network g' if $v(g) \geq v(g')$. A network is *efficient* if it dominates any other. As shown here, efficiency may be reached by a network that leaves some nodes disconnected. A network is *efficient-constrained-to-keep-connectedness* if it is connected and dominates any other connected network.

Definition 1 *A nested split graph network (NSG-network) is a network g such that for all $(i,j) \in \mathcal{T}$,*

$$|N_i(g)| \leq |N_j(g)| \Rightarrow N_i(g) \subseteq N_j(g) \cup \{j\},$$

where $N_i(g)$ denotes the set of neighbors of i in g.

Thus, NSG-networks have hierarchical structures where nodes can be ranked by the number of their neighbors. Note that NSG-networks have at most one non-trivial component. Equivalently, an NSG-network can be defined as one where, for a certain numbering of the nodes,

$$g_{ij} = 1 \Rightarrow g_{kl} = 1, \text{ for all } (k,l) \in \mathcal{T} \text{ s.t. } k \leq i \text{ and } l \leq j.$$

NSG-networks of a special type play a central role in the results presented here, i.e. those that rank nodes consistently with their values in the following sense:

Definition 2 *A strong NSG-network (SNSG-network) is a nested split graph network g such that*

$$v_i > v_j \Rightarrow |N_i(g)| \geq |N_j(g)|. \tag{3}$$

Thus, *strong* NSG-networks are NSG-networks where the hierarchy or ranking of nodes according to their number of neighbors, inherent in any such network, is reinforced consistently by the rank of nodes according to their value.[4]

[4] An equivalent more condensed definition would be this one: *A strong NSG-network is a network g such that: $v_i > v_j \Rightarrow N_i(g) \cup \{i\} \supseteq N_j(g)$.*

Example 1. The two triangular matrices below specify two NSG-networks of 9 nodes, whatever their values. However, if they have different values, they are *strong* NSG-networks only if (3) holds.[5] This is so, for instance, if $v_1 \geq v_2 \geq ... \geq v_9$. Matrix $\widehat{g}$ corresponds to a connected NSG-network, where node 1 has 8 neighbors, nodes 2 and 3 have 6 neighbors, 4 has 5 neighbors, 5 and 6 have 4 neighbors, 7 has 3 neighbors, and nodes 8 and 9 have 1 neighbor. Matrix $\widehat{g}'$ corresponds to a non-connected NSG-network, where nodes 1 and 2 have 7 neighbors, 3 has 6 neighbors, 4 has 5 neighbors, 5 and 6 have 4 neighbors, 7 has 3 neighbors, 8 has 2 neighbors, and node 9 has no neighbors.

$$\widehat{g} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ {\scriptstyle (2)} & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ & {\scriptstyle (3)} & 1 & 1 & 1 & 1 & 0 & 0 \\ & & {\scriptstyle (4)} & 1 & 1 & 0 & 0 & 0 \\ & & & {\scriptstyle (5)} & 0 & 0 & 0 & 0 \\ & & & & {\scriptstyle (6)} & 0 & 0 & 0 \\ & & & & & {\scriptstyle (7)} & 0 & 0 \\ & & & & & & {\scriptstyle (8)} & 0 \end{bmatrix} \quad \widehat{g}' = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ {\scriptstyle (2)} & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ & {\scriptstyle (3)} & 1 & 1 & 1 & 1 & 0 & 0 \\ & & {\scriptstyle (4)} & 1 & 1 & 0 & 0 & 0 \\ & & & {\scriptstyle (5)} & 0 & 0 & 0 & 0 \\ & & & & {\scriptstyle (6)} & 0 & 0 & 0 \\ & & & & & {\scriptstyle (7)} & 0 & 0 \\ & & & & & & {\scriptstyle (8)} & 0 \end{bmatrix}$$

In what follows, whenever we refer to a network we assume nodes numbered so that $v_1 \geq v_2 \geq ... \geq v_n$.

Proposition 1 *Any connected network with a positive net value is dominated by a connected strong nested split graph network.*

The procedure described in the proof constructively produces a connected NSG-network g' that ranks nodes consistently with their values, i.e. a *strong* NSG-network which dominates the initial network g. The whole paper hinges upon this finding, as all the results presented here stem from this simple construction. Nevertheless, it may be the case that g' can be refined so that a network of the same type of still larger net value is obtained. The following result refines Proposition 1 by characterizing the connected SNSG-networks which maximize the net value, i.e. reach efficiency-constrained-to-keep-connectedness.

Proposition 2 *A connected strong nested split graph network g is efficient-constrained-to-keep-connectedness if and only if it yields a positive net value and the following conditions hold*

$$\max_{(j,k)\in\mathcal{T} \ \& \ g_{jk}=0} (v_j + v_k) \overset{(a)}{\leq} \frac{c}{\delta - \delta^2} \overset{(b)}{\leq} \min_{(j,k)\in\mathcal{T}, \ j\neq 1 \ \& \ g_{jk}=1} (v_j + v_k). \tag{4}$$

[5]The first row indicates the neighbors of node 1, and the last column those of node 9, while small numbers (2), (3),..., (8) below the diagonal enable the neighbors of the others to be counted. For instance, the five neighbors of node 4 in g correspond to the three 1-entries above (4) and the two 1-entries to its right.

Thus, Propositions 1 and 2 establish how to reach efficiency *under the constraint of not leaving any node disconnected*, i.e. under the constraint of *connectedness*. The following corollary plays a role for computing the efficient network.

Corollary 1 *If an n-node connected strong nested split graph network g is efficient-constrained-to-keep-connectedness, then the $(n-1)$-node network that results from g by eliminating the node with smallest value is efficient-constrained-to-keep-connectedness among networks with those $n-1$ nodes.*

However, if connectedness is not a constraint, efficiency in absolute terms may be reached while leaving some nodes disconnected.

Proposition 3 *Any network is dominated by a strong nested split graph network which is not necessarily connected s.t. condition* (4) *restricted to the only non-trivial component holds.*

This leads to the following characterization:

Proposition 4 *A connected strong nested split graph network g is efficient if and only if* (4) *and*

$$c \leq \delta(v_1 + v_n) + \delta^2 \sum_{k \in N \setminus \{1,n\}} (v_k + v_n) \quad \text{if} \;\; N_n(g) = \{1\}. \tag{5}$$

Two structures appear in the seminal connections model of Jackson and Wolinsky (1996) and some of its extensions as the only possibly non-empty efficient networks: The complete network and the all-encompassing star. The complete network and all-encompassing stars *centered on a node with the highest value* are extreme cases of SNSG-networks. Only condition (4-(b)) of Proposition 2 applies for a complete network, which is efficient if and only if

$$\frac{c}{\delta - \delta^2} \leq \min_{(j,k) \in \mathcal{T}} (v_j + v_k) = v_{n-1} + v_n. \tag{6}$$

Similarly, only conditions (4-(a)) of Proposition 2 and (5) of Proposition 4 apply to the all-encompassing star centered on node 1. Thus such a star is efficient if and only if

$$v_2 + v_3 = \max_{(j,k) \in \mathcal{T} \text{ s.t. } g_{jk}=0} (v_j + v_k) \leq \frac{c}{\delta - \delta^2}$$

and

$$c \leq \delta(v_1 + v_n) + \delta^2 \sum_{k \in N \setminus \{1,n\}} (v_k + v_n).$$

Therefore, *neither the complete network nor any all-encompassing star is efficient* when

$$v_{n-1} + v_n < \frac{c}{\delta - \delta^2} < v_2 + v_3. \tag{7}$$

Note also that when all nodes have the same value this interval *collapses to a point*. At one side the complete network is efficient and at the other the all-encompassing star, i.e. the well-known efficiency result in the homogeneous case in Jackson and Wolinsky (1996).

The following examples for a network with 4 nodes illustrate the results.

Example 2. Consider 4 nodes of values $v_1 = 4$, $v_2 = 3$, $v_3 = 2$ and $v_4 = 1$, and let $\delta = 0.5$ and $c = 1$. Then

$$\min_{(j,k)\in\mathcal{T}}(v_j + v_k) = 3 < \frac{c}{\delta - \delta^2} = 4 < 5 = \max_{j,k\in N\setminus\{i\}}(v_j + v_k).$$

There are 4 connected SNSG-networks, whose associated triangular matrices are

$$\widehat{g}_1 = \begin{bmatrix} 1 & 1 & 1 \\ & 0 & 0 \\ & & 0 \end{bmatrix}, \widehat{g}_2 = \begin{bmatrix} 1 & 1 & 1 \\ & 1 & 0 \\ & & 0 \end{bmatrix}, \widehat{g}_3 = \begin{bmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 0 \end{bmatrix}, \widehat{g}_4 = \begin{bmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{bmatrix},$$

which correspond to the four networks shown in Figure 1.

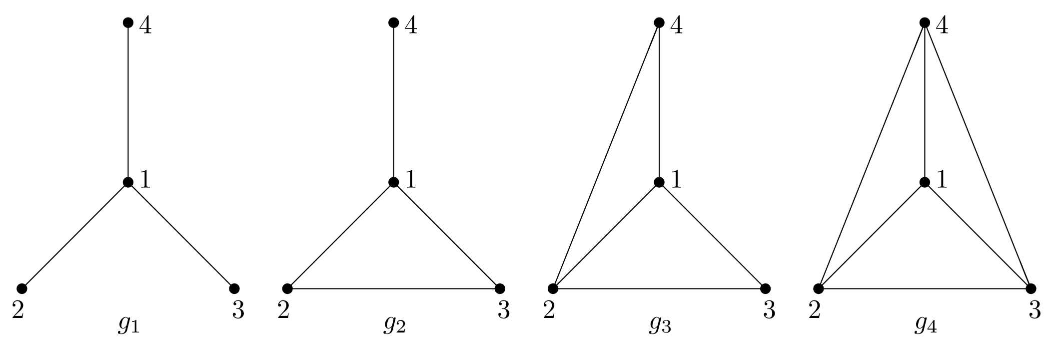

Figure 1: Connected SNSG-netwoks for $n = 4$

Both the complete network and the star yield a net value of 9, while the other two yield a net value of 9.25, thus *both g_2 and g_3 are efficient* because leaving disconnected node 4 does not improve that net value.

Example 3. If $\delta = 0.5$ and $c = 1$ as in Example 2, but the values of the nodes are $v_1 = 4$, $v_2 = 3$, $v_3 = 2.5$ and $v_4 = 0.5$, (7) holds and the net values of the complete network, the star network, and network g_3 yield the same net value as in the preceding example, but the net value of g_2 is 9.375, making it *the only efficient one* because leaving node 4 disconnected cannot improve that net value.

Example 4. If the values are $v_1 = v_2 = v_3 = 10$ and $0 < v_4 < 1$, with $\delta = 0.2$ and $c = 3$, then the net value of g_1 is $2.6 + 0.28v_4$, that of g_3 is $1.4 + 0.44v_4$, and that of g_4 is $0.6v_4$. Thus g_2, whose net value is $2.8 + 0.28v_4$, is the efficient network *constrained to keep connectedness*. However, if node 4 is left disconnected the complete network of nodes 1, 2 and 3 yields a net value of 3, greater than the net value of the three-node star, which yields 2.08. Therefore, if $v_4 < 5/7$, efficiency *in absolute terms* is reached by the network with a component formed by nodes 1, 2 and 3 fully connected and node 4 isolated.

In view of Propositions 2 and 3, efficiency is reached by an SNSG-network, perhaps leaving some of the less valuable nodes disconnected. By Corollary 1, one way to find the efficient SNSG-network is the following: First calculate an SNSG-network efficient-constrained-to-keep-connectedness and eliminate the weakest node as long as the net value of the resulting network increases. The following algorithm enables efficient-constrained-to-keep-connectedness SNSG-networks to be calculated.

Algorithm for obtaining an efficient network constrained to keep connectedness

Assume n nodes, $1, 2, ..., n$ of values $v_1 \geq v_2 \geq ... \geq v_n$. As before, denote $\mathcal{T} = \{(i,j) : 1 \leq i < n \ \& \ i < j \leq n\}$, and $\mathcal{S} = \{(1,j) : 2 \leq j < n\ \}$. For each $(i,j) \in \mathcal{T}$, denote $s_{ij} = v_i + v_j$, and order the pairs in $\mathcal{T}$ inversely to the values of the corresponding sums of weights. More precisely, let $\preceq$ be the linear order in $\mathcal{T}$ defined, for each $(i,j), (k,l) \in \mathcal{T}$, by[6]

$$(i,j) \preceq (k,l) \overset{DEF}{\Leftrightarrow} [\ (s_{ij} > s_{kl}) \text{ or } (s_{ij} = s_{kl} \ \& \ ((i \leq k \ \& \ j \leq l) \text{ or } (i < k \ \& \ j > l)))].$$

To keep it as simple as possible the algorithm is based on only one variable: A subset $\mathcal{G} \subseteq \mathcal{T}$, which implicitly specifies a network g s.t. $g_{ij} = 1$ if and only if $(i,j) \in \mathcal{G}$ at each stage. Then proceed as follows:

1. Form the all-encompassing star centered at node 1, i.e. make $\mathcal{G} := \mathcal{S}$.
2. Take $(i,j) := \min \mathcal{T} \backslash \mathcal{G}$.
3. If $\delta s_{ij} - c \geq \delta^2 s_{ij}$, make $\mathcal{G} := \mathcal{G} \cup \{(i,j)\}$, otherwise Stop.
4. If $\mathcal{T} \backslash \mathcal{G} = \varnothing$ Stop, otherwise go to 2.

The algorithm obviously ends after at most $\frac{(n-1)(n-2)}{2}$ (the cardinality of $\mathcal{T} \backslash \mathcal{S}$) iterations. Note that after every cycle $\mathcal{T} \backslash \mathcal{G}$ loses one element. Moreover, if the final output is $\mathcal{G}$, at the end the associated network g (s.t. $g_{ij} = 1$ if and only $(i,j) \in \mathcal{G}$) is efficient because after every cycle the current associated network is an NSG-network that ranks nodes consistently with their values and s.t. for all (i,j) s.t. $g_{ij} = 1$, $\delta s_{ij} - c \geq \delta^2 s_{ij}$, thus at the end

$$\frac{c}{\delta - \delta^2} \leq \min_{(j,k) \in \mathcal{T},\ j \neq 1 \ \& \ g_{jk} = 1} (v_j + v_k)$$

necessarily, which is inequality (4-(b)). And inequality (4-(a)) holds also because if the algorithm stops before $\mathcal{T} \backslash \mathcal{G}$ is empty, it does so because in the last repetition of step

[6]If $s_{ij} \neq s_{kl}$ whenever $(i,j) \neq (k,l)$, then $(i,j) \preceq (k,l) \overset{DEF}{\Leftrightarrow} (s_{ij} \geq s_{kl})$ is a linear order in $\mathcal{T}$. Otherwise, it is not antisymmetric. Hence the clause for the case when $s_{ij} = s_{kl}$, for breaking the ties.

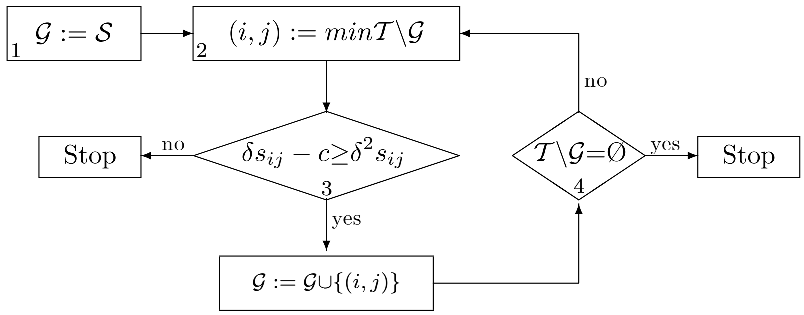

Figure 2: Flowchart of the algorithm

3 it happens that $\delta s_{ij} - c < \delta^2 s_{ij}$, but if this is so, for (i,j) then $\delta s_{kl} - c < \delta^2 s_{kl}$ holds also for any $(k,l) \succeq (i,j)$. In other words, inequality (4-(b)) also holds. Moreover, by construction, the resulting network is the only SNSG-network that satisfies (4) *unless* in the final steps, for the last two or more elements added to $\mathcal{G}$ before the Stop, it occurs that $\delta s_{ij} - c = \delta^2 s_{ij}$ (see Example 2), and the sum s_{ij} remains the same for them. In this case, the deletion of any such links does not change the net value.

This clarifies the question of uniqueness. In fact, uniqueness can be added to Propositions 2 and 4, if $\delta s_{ij} - c \neq \delta^2 s_{ij}$ for all $(i,j) \in \mathcal{T}$. Otherwise, as shown in the algorithm, a trivial form of multiplicity arises, and in some cases the SNSG-structure may be broken. Uniqueness can be recovered either by adding a condition of minimum cost, i.e. replacing condition $\delta s_{ij} - c \geq \delta^2 s_{ij}$ by condition $\delta s_{ij} - c > \delta^2 s_{ij}$ in the algorithm, or, by contrast, by maximizing the aggregate information received, which is in fact the outcome of the algorithm described.

5 Stability

To address the question of stability in a decentralized environment it is necessary to specify how nodes cover the cost of each link in which they are involved. We consider a natural extension of Jackson and Wolinsky's (1996) pairwise stability notion, where the cost of each link can be split in any way by the nodes connected by it. This contrasts with Jackson and Wolinsky (1996), where it is assumed that the cost must be split equally between the two nodes. An *investment profile* is a matrix of non-negative real numbers $\mathbf{c} = (c_{ij})_{i,j \in N}$ with $c_{ii} = 0$, where c_{ij} is the amount that node i invests in link ij. A link connecting i and j actually forms if and only if $c_{ij} + c_{ji} \geq c$. That is, an investment profile $\mathbf{c} = (c_{ij})_{i,j \in N}$ supports the network $g^{\mathbf{c}}$ s.t.

$$g^{\mathbf{c}}_{ij} = \begin{cases} 1, \text{ if } c_{ij} + c_{ji} \geq c, \\ 0, \text{ otherwise.} \end{cases}$$

The payoff of node-player i is

$$\pi_i(g^{\mathbf{c}}) := I_i(g^{\mathbf{c}}) - \sum_{j \in N_i(g^{\mathbf{c}})} c_{ij} = \sum_{j \in N \setminus \{i\}} \delta^{d(i,j)} v_j - \sum_{j \in N_i(g^{\mathbf{c}})} c_{ij}. \tag{8}$$

We drop $\mathbf{c}$ in $g^{\mathbf{c}}$ when the investment that supports network g is clear form the context. Given a network g, $g+ij$ denotes the network found by adding link ij to g, and $g-ij$ the network found by deleting link ij from g. Then, inspired by the pairwise stability notion of Jackson and Wolinsky (1996), we assume that an *investment profile* $(c_{ij})_{i,j\in N}$ *is pairwise stable* if: (i) No node can increase its payoff by decreasing the amount invested in any of its links; and (ii) If a new link increases the payoff of one of the nodes that it connects, then the payoff of the other node decreases, however its cost is split. An investment profile is *tight* if $c_{ij}+c_{ji}=c$ or $c_{ij}=c_{ji}=0$ for each $i,j\in N$. To be tight is an obvious necessary condition for an investment profile to be pairwise stable. Under this condition, decreasing the amount invested in a link entails its elimination. Then we have the following formal definition.

Definition 3 *An investment profile* $(c_{ij})_{i,j\in N}$ *that supports* g *is pairwise stable if it is tight and*
PSFC-(i) $ij \in g \Rightarrow I_i(g) - c_{ij} \geq I_i(g-ij)$, *and*
PSFC-(ii) $(ij \notin g \ \&\ 0 \leq k \leq c \ \&\ I_i(g+ij) - k > I_i(g)) \Rightarrow I_j(g+ij) - (c-k) < I_j(g)$.

We say that a *network is supportable in pairwise equilibrium (pairwise stable for brief) under free cost sharing* if there exists a pairwise stable investment profile that supports it. It may be the case that a network can be supported by different pairwise stable investment profiles. Jackson and Wolinsky's pairwise stability can be formulated like this:

Definition 4 (*Jackson and Wolinsky,* 1996) *A network* g *is pairwise stable if*
PS-(i) $ij \in g \Rightarrow I_i(g) - c/2 \geq I_i(g-ij)$, *and*
PS-(ii) $(ij \notin g \ \&\ I_i(g+ij) - c/2 > I_i(g)) \Rightarrow I_j(g+ij) - c/2 < I_j(g)$.

Note the similarity of definitions 3 and 4: Conditions *PS-*(i) and *PS-*(ii) for pairwise stability in the sense of Jackson and Wolinsky (1996) are the result of applying conditions *PSFC-*(i) and *PSFC-*(ii), *but constraining admissible investment profiles to be of the form*

$$c_{ij} = \begin{cases} c/2 & \text{if } g_{ij} = 1 \\ 0 & \text{if } g_{ij} = 0. \end{cases} \tag{9}$$

This considerably constrains the possibilities of making a link stable by imposing equal shares of the cost. By contrast, it broadens the stability of the gap of a non-existing link by limiting the ways in which a stable link can bridge it. Thus there is no implication in either direction between these two stability notions.[7]

[7]Alternatively, this free cost sharing setting can be interpreted as a specialization of Bloch and Jackson's (2006) *linking game with transfers*. To see this, let $u_i(g)$ be the payoff of node i in the

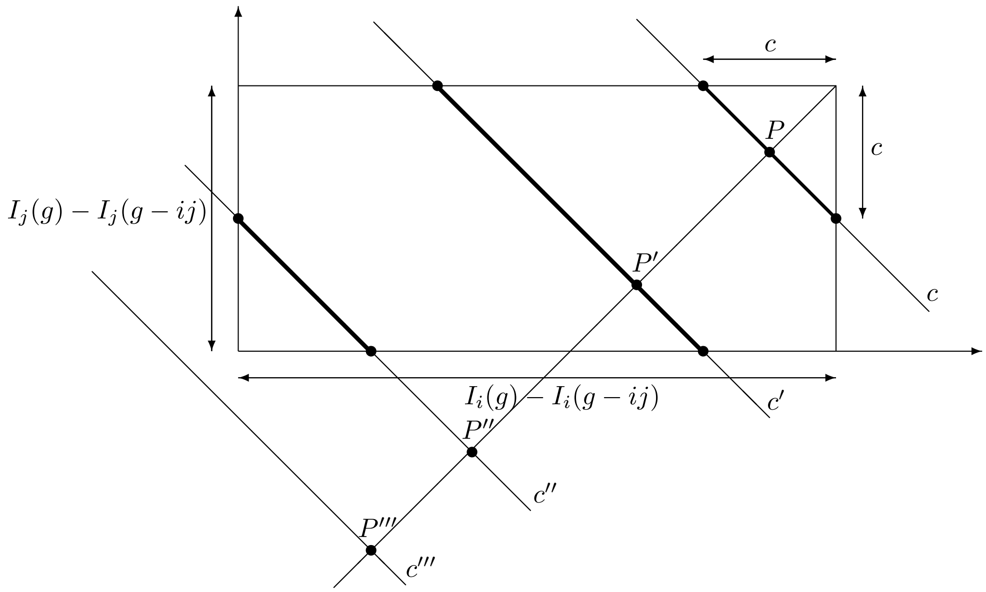

Figure 3: Pairwise stability vs. pairwise stability under cost sharing

Figure 3 illustrates this situation. The lengths of the sides of the rectangle are the increase of information that link ij means to i and to j. Straight lines of slope -1 correspond to different values of the cost of a link: c, c', c'', c'''. The points in the segments of the intersections of each of these lines with the rectangle correspond to the ways in which that cost can be shared and *PSFC*-(i) be satisfied; while the points P, P', P'' and P''' (intersections of the straight line of slope 1 crossing the northwest corner of the rectangle and those corresponding to the different costs) correspond to equal share of different costs. This makes clear the wider rank of costs compatible with *PSFC*-(i). Link $ij \in g$ can stabilized with an investment s.t. (9), when

$$c/2 \leq \min\{I_i(g) - I_i(g - ij), I_j(g) - I_j(g - ij)\},$$

while under free cost sharing it can be stabilized for the wider range:

$$c \leq I_i(g) - I_i(g - ij) + I_j(g) - I_j(g - ij).$$

symmetric connections model of Jackson and Wolinsky (1996), i.e. $u_i(g) = I_i(g) - \eta_i(g)c/2$, where $\eta_i(g)$ is the number of neighbors of node i. Assume that transfers in the sense of Bloch and Jackson's (2006) are allowed, with $t^i_{ij} := c/2 - c_{ji}$, and $g_{ij}(t) = 1$ iff $t^i_{ij} + t^j_{ij} = 0$ (i.e. $c_{ij} + c_{ji} = c$) and $g = g(t)$, the payoff of node i in the linking game with transfers is

$$\begin{aligned}\pi_i(g) &= u_i(g) - \sum_{j \in N_i(g(t))} t^i_{ij} = I_i(g) - \eta_i(g)c/2 - \sum_{j \in N_i(g(t))} t^i_{ij} \\ &= I_i(g) - \sum_{j \in N_i(g(t))} (c/2 - c_{ji}) = I_i(g) - \sum_{j \in N_i(g(t))} c_{ij},\end{aligned}$$

which is the payoff of node i given by (8). But this is consistent with our model only if $-c/2 \leq t^i_{ij} \leq c/2$, while Bloch and Jackson (2006) sets no bounds for the transfers and admit wasted transfers. Assuming these bounds and no wasted transfers, our pairwise stability under free cost sharing is very similar to their pairwise Nash equilibrium with transfers, although we do not assume Nash-stability.

For the same reason, a gap $ij \notin g$ is stable in the sense of *PS*-(ii) if

$$c/2 > \min\{I_i(g) - I_i(g - ij), I_j(g) - I_j(g - ij)\},$$

while it is stable in the sense of *PSFC*-(ii) only for the narrower range:

$$c > I_i(g + ij) - I_i(g) + I_j(+ij) - I_j(g).$$

Thus pairwise stability in the sense of Jackson and Wolinsky (1996) is not stronger nor weaker than pairwise stability under free cost sharing.[8]

We have the following characterization of connected strong nested split graph networks pairwise stable under free cost sharing.

Proposition 5 *A connected strong nested split graph network is pairwise stable under free cost sharing if and only if*

$$\max_{(j,k)\in\mathcal{T}\ s.t.\ g_{jk}=0}(v_j + v_k) \overset{(a)}{<} \frac{c}{\delta - \delta^2} \overset{(b)}{\leq} \min_{(j,k)\in\mathcal{T},\ j\neq 1\ \&\ g_{jk}=1}(v_j + v_k), \tag{10}$$

and

$$c \leq \delta(v_1 + v_n) + \delta^2 \sum_{k\in N\setminus\{1,n\}} v_k \quad \textit{if} \quad N_n(g) = \{1\}. \tag{11}$$

A comparison of conditions (4) and (5) in Proposition 4 with conditions (10) and (11), immediately yields the following.

Corollary 2 *A connected strong nested split graph network pairwise stable under free cost sharing is efficient.*

Note that the converse is not true in general. This can be seen clearly by comparing condition (5) and condition (11), which is obviously stronger[9]. The reason is that efficiency requires each link to contribute to the maximization of the aggregate payoff, while stability requires the nodes that support it to get the best from their investments. As the following example shows, this leads to possible instability of the links in an efficient network connecting some nodes with small values with the node with the highest value and the center of an SNSG-network if it is their only neighbor.

Example 5. Consider 4 nodes of values $v_1 = v_2 = v_3 = 10$ and $5/7 \leq v_4 < 1$, connected by network g_2 in Figure 1, and let $\delta = 0.2$ and $c = 3$. It is immediate to check that conditions (4) and (5) hold and consequently network g_2 is efficient. But g_2 is *not* supportable in pairwise equilibrium under free cost sharing. The reason is that

[8] At the end of this section we compare these two notions in the homogeneous setting of Jackson and Wolinsky (1996) and the heterogeneous one considered here.

[9] A less relevant trivial case may occur in the extreme case when (4-(a)) holds with equality and consequently (10-(a)) does not.

link 14 contributes to increase the net value of the network, but it is not supportable in equilibrium under free cost sharing. If nodes 1 and 4 invest in c_{14} and c_{41} so that $c_{14} + c_{41} = 3$, their increase of payoffs due to link 14 are

$$\Delta_4 = \pi_4(g_2) - \pi_4(g_2 - 14) = \delta 10 + 2\delta^2 10 - c_{41} = 2.8 - c_{41},$$

$$\Delta_1 = \pi_1(g_2) - \pi_1(g_2 - 14) = \delta v_4 - c_{14} = 0.2 v_4 - c_{14}.$$

But then

$$\Delta_1 + \Delta_4 = 2.8 + 0.2 v_4 - 3 = 0.2(v_4 - 1) < 0.$$

In other words, however nodes 1 and 4 share the cost of the link, if one gains the other loses, so g_2 is *not* supportable in pairwise equilibrium under free cost sharing. This shows that the reciprocal of Corollary 2 is not true. Moreover, the elimination of link 14 in g_2 yields a non-connected pairwise stable network under free cost sharing which is not efficient. This shows that the corollary does not hold for non-connected SNSG-networks.

For a connected SNSG-network with no nodes whose only neighbor is the center, condition (5) for efficiency and condition (11) for pairwise stability cease to apply. This leads to the following conclusion.

Corollary 3 *A connected strong nested split graph network with no nodes whose only neighbor is the center, generically, is efficient if and only if it is pairwise stable under free cost sharing.*

We conclude with a comparison of pairwise stability under free cost sharing and classical pairwise stability. First, comparing them when they are applied in the homogeneous setting of Jackson and Wolinsky (1996); then when they are applied with heterogeneous nodes.

Proposition 6 *In the connections model with heterogeneous nodes:*
(i) A network pairwise stable under free cost sharing has at most one non-trivial component.
(ii) The complete network is pairwise stable under free cost sharing if and only if

$$\frac{c}{\delta - \delta^2} \leq v_{n-1} + v_n. \tag{12}$$

(iii) The all-encompassing star centered at node 1 *is pairwise stable under free cost sharing if and only if*

$$(\delta - \delta^2)(v_2 + v_3) < c \leq \delta(v_1 + v_n) + \delta^2 \sum_{k \in N \setminus \{1,n\}} v_k. \tag{13}$$

In the homogeneous setting considered in Jackson and Wolinsky (1996), i.e. $v_1 = v_2 = ... = v_n = 1$, condition (12) becomes

$$c/2 \leq \delta - \delta^2,$$

while (13) becomes

$$\delta - \delta^2 < c/2 \leq \delta + \delta^2(n-2)/2.$$

A comparison of these two conditions with parts (*ii*) and (*iii*) of Proposition 2 in Jackson and Wolinsky (1996) relative to pairwise stability of these two networks shows the following. The first one coincides (note that they consider only "interior" conditions and their c is our $c/2$), i.e. the complete network is sustainable in pairwise equilibrium under free cost sharing if and only if it is pairwise stable. However, the second condition for the star differs considerably because the interval for the star to be pairwise stable is (Proposition 2 -(ii), Jackson and Wolinsky, 1996)

$$\delta - \delta^2 < c/2 \leq \delta.$$

That is, the range of values for c is considerably narrower. The reason for the difference is clear. Considering as admissible only those investment profiles where c_{ij} is 0 or $c/2$ considerably limits the stability of the star because the center is constrained to invest $c/2$ in order to sustain a link with a spoke node however many there may be, while under free cost sharing such a link may even be entirely supported by the spoke node in equilibrium. Moreover, this is sure to be so however small the value of the nodes is if there are enough of them. For the same reason, part (*iv*) of Proposition 2 in Jackson and Wolinsky (1996) does not hold for pairwise stability under free cost sharing, while part (*i*) does.

To complete the comparison with pairwise stability in the sense of Jackson and Wolinsky (1996) we have the following result, parallel to Proposition 5, which characterizes pairwise stable connected strong nested split graph networks (the proof, simple and similar, is omitted).

Proposition 7 *A connected strong nested split graph network is pairwise stable if and only if*

$$\max_{(j,k)\in\mathcal{T}\ s.t.\ g_{jk}=0} \min\{2v_j, 2v_k\} \overset{(a)}{<} \frac{c}{\delta - \delta^2} \overset{(b)}{\leq} \min_{(j,k)\in\mathcal{T},\ j\neq 1\ \&\ g_{jk}=1} \{2v_j, 2v_k\}, \tag{14}$$

and

$$c \leq 2\delta v_n if N_n(g) = \{1\}. \tag{15}$$

A comparison of the intervals determined by (10) and by (14) shows that, in general, there is no inclusion between them in either direction. As to conditions (11) and (15), the second condition for pairwise stability in the sense of Jackson and Wolinsky (1996) is stronger, considerably constraining the possibility of stability of connected strong nested split graph networks with nodes of small value connected only to the central node of greatest value.

6 Concluding remarks

We study a natural extension of the connections model of Jackson and Wolinsky (1996), introducing heterogeneity in the values of nodes. We characterize the type of structures that enable efficiency to be reached, be it under the constraint of keeping connectedness or unconstrained in absolute terms. These structures, which we call strong nested split graph networks, are highly hierarchical. They are nested split graph networks in which the ranking of nodes according to the number of neighbors is consistent with their ranking according to their values, possibly leaving some nodes with the smallest values disconnected. These structures include complete and star networks, but also a full range of intermediate cases, which are illustrated with examples.[10] A simple algorithm for obtaining the efficient networks is also provided.

We also study a notion of pairwise stability under free cost sharing. If players are able to coordinate to form a link it is only natural to assume that they are also able to agreeing on how its cost is shared. This moderate dose of 'cooperativeness' leads to a crisp result: Under free cost sharing, connected strong nested split graph networks pairwise stable are efficient. Moreover, a connected strong nested split graph network were no node has the center as its only neighbor is pairwise stable under free cost sharing generically if and only if it is efficient. We also discuss the relation between this notion and that of pairwise stability in the sense of Jackson and Wolinsky (1996), and show that this is the particular case of pairwise stability under free cost sharing once cost sharing is no longer free, but only a particular type of investment profiles is considered admissible.

If the multiplicity of investment profiles that support a pairwise stable network under free cost sharing is considered undesirable, a more reasonable prescription than equal splitting of the cost of each link would be to split it in proportion to the increase in information that the link provides for each of the two nodes that it connects. This would produce links which are stable under free cost sharing.[11]

This study is complementary of Olaizola and Valenciano (2020a,b,c), which deal with extensions of the connections model where nodes are homogeneous, but links are not only heterogeneous but of endogenous strength, based on a technology, so that the strength of a link depends on the investment in it. A clear line for further work is to combine the two approaches: Heterogeneity of nodes and heterogeneity of endogenous links based on a technology as in Olaizola and Valenciano (2020c).

Appendix

[10] In contrast with what happens when nodes are homogeneous but links are heterogeneous and endogenous (Olaizola and Valenciano, 2020c), where the complete network and the all-encompassing star are the only possible nonempty efficient networks.

[11] In graphical terms, the interpretation of this prescription over Figure 3 is given by the intersection of the diagonal of the rectangle with positive slope with the segment that represents stable ways of sharing the cost. Note that this can be interpreted as the Kalai and Smorodinsky (1975) solution of the bargaining problem every two players face in order to share the cost of a link.

Proposition 1:

Proof. Let g be a connected network, with n nodes s.t. $v_1 \geq v_2 \geq ... \geq v_n$, and k links of strength $\delta > 0$ and such that $v(g) > 0$. The proof is constructive and consists of producing an SNSG-network g' s.t. $v(g') \geq v(g)$. Arrange all $n(n-1)/2$ sums $s_{ij} = v_i + v_j$ in a triangular matrix T of $n-1$ rows, i.e.

$$T = \begin{bmatrix} s_{12} & s_{13} & s_{14} & \cdots & s_{1j} & \cdots & s_{1n-1} & s_{1n} \\ & s_{23} & s_{24} & \cdots & s_{2j} & \cdots & s_{2n-1} & s_{2n} \\ & & s_{34} & \cdots & s_{3j} & \cdots & s_{3n-1} & s_{3n} \\ & & & \ddots & \vdots & & \vdots & \vdots \\ & & & & s_{j-1,j} & \cdots & & s_{j-1n} \\ & & & & & \ddots & & \vdots \\ & & & & & & s_{n-2,n-1} & s_{n-2,n} \\ & & & & & & & s_{n-1,n} \end{bmatrix}.$$

That is, $T = (s_{ij})_{(i,j)\in\mathcal{T}}$, where $\mathcal{T} = \{(i,j) : 1 \leq i < n \;\&\; i < j \leq n\}$. In what follows we use the following notation. Given a subset $\mathcal{C} \subseteq \mathcal{T}$,

$$\Sigma(\mathcal{C}) := \sum\nolimits_{(i,j)\in\mathcal{C}} s_{ij}.$$

Note that $\Sigma(\mathcal{T}\backslash\mathcal{C}) = \Sigma(\mathcal{T}) - \Sigma(\mathcal{C})$. The connectedness of g implies that k, the number of links of g, is at least $n-1$, i.e. $k \geq n-1$. Decompose the set of $n(n-1)/2$ sums into two subsets, H and L, where H contains the k largest sums and L the smallest $(n(n-1)/2) - k$ sums, and let S be the sums in the first row, i.e. $S = \{s_{1i} : i = 2, ..., n\}$. Then denote by $\mathcal{H}$, $\mathcal{L}$ and $\mathcal{S}$ the subsets of $\mathcal{T}$ corresponding to H, L and S. Then proceed as follows to construct a connected network g' s.t. $v(g') \geq v(g)$. Form an all-encompassing star centered on node 1 (the node with the highest value) with $n-1$ of the links, i.e. connect all pairs of nodes in $\mathcal{S}$ directly, and also connect the pairs of nodes corresponding to the largest $k - (n-1)$ sums in $H\backslash S$ directly.

Let g' be the resulting network. We prove that $v(g') \geq v(g)$. Decompose the total information received by all nodes (i.e. net value + cost) of the initial network g into two parts: Part A, the value generated by the direct connections (k summands corresponding to a set of pairs $\mathcal{K} \subseteq \mathcal{T}$); and Part B, the value generated by the indirect connections ($(n(n-1)/2) - k$ summands corresponding to pairs in $\mathcal{T}\backslash\mathcal{K}$). That is,

$$A = \delta \sum\nolimits_{(i,j)\in\mathcal{K}} s_{ij} = \delta\Sigma(\mathcal{K})$$

and

$$B = \sum\nolimits_{(i,j)\in\mathcal{T}\backslash\mathcal{K}} \delta^{d(i,j)} s_{ij} \leq \delta^2\Sigma(\mathcal{T}\backslash\mathcal{K}).$$

Decompose the total information received by all nodes of g' in the same way, i.e. the value generated by the direct connections (k summands corresponding to a set of pairs $\mathcal{K}' \subseteq \mathcal{T}$)

$$A' = \delta \sum\nolimits_{(i,j)\in\mathcal{K}'} s_{ij} = \delta\Sigma(\mathcal{K}'),$$

and the value generated by the indirect connections ($(n(n-1)/2) - k$ summands corresponding to pairs in $\mathcal{T} \backslash \mathcal{K}'$)

$$B' = \delta^2 \sum\nolimits_{(i,j) \in \mathcal{T} \backslash \mathcal{K}'} s_{ij} = \delta^2 \Sigma(\mathcal{T} \backslash \mathcal{K}').$$

We show that $A' + B' \geq A + B$.

Case 1: $\mathcal{S} \subseteq \mathcal{H}$. In this case $\mathcal{K}' = \mathcal{H}$. Then,

$$A' + B' = \delta \Sigma(\mathcal{H}) + \delta^2 \Sigma(\mathcal{T} \backslash \mathcal{H}) = \delta \Sigma(\mathcal{H}) + \delta^2 (\Sigma(\mathcal{T}) - \Sigma(\mathcal{H})),$$

$$A + B \leq \delta \Sigma(\mathcal{K}) + \delta^2 \Sigma(\mathcal{T} \backslash \mathcal{K}) = \delta \Sigma(\mathcal{K}) + \delta^2 (\Sigma(\mathcal{T}) - \Sigma(\mathcal{K})).$$

Then, given that $\Sigma(\mathcal{H}) \geq \Sigma(\mathcal{K})$,

$$\begin{aligned} (A' + B') - (A + B) &\geq (\delta \Sigma(\mathcal{H}) - \delta^2 \Sigma(\mathcal{H})) - (\delta \Sigma(\mathcal{K}) - \delta^2 \Sigma(\mathcal{K})) \\ &= (\delta - \delta^2)(\Sigma(\mathcal{H}) - \Sigma(\mathcal{K})) \geq 0. \end{aligned}$$

Finally, the cost of both networks is the same so necessarily $v(g') \geq v(g)$.

Case 2: $\mathcal{S} \subsetneq \mathcal{H}$, i.e. $v_1 + v_j \notin H$ for all $j > r$, for some r s.t. $3 \leq r \leq n$.[12] In this case, the direct connection of node 1 with all $j > r$ yields $n - r$ summands of the form δs_{1j}, with $(1, j) \in \mathcal{L} \cap \mathcal{S}$. That is, unlike the preceding case, some of the links corresponding to summands in A' (these $n - r$ exactly) in g' connect nodes which yield sums not in H, and by the choice of g' only these summands in A' are not in H. Nevertheless, given that g is connected, for any such $(1, j) \in \mathcal{L} \cap \mathcal{S}$ s.t. $g'_{1j} = 1$ there must be some i s.t. $g_{ij} = 1$ whose contribution cannot be greater because $s_{ij} \leq s_{1j}$. Therefore,

$$\Sigma(\mathcal{K}') \geq \Sigma(\mathcal{K}),$$

which, as in the first case, entails $v(g') \geq v(g)$.

It only remains to be shown that g' is a strong nested split graph network. Note first that $H = \{s_{ij} : (i, j) \in \mathcal{H}\}$ contains the k largest sums, and

$$s_{ij} \in H \Rightarrow s_{kl} \in H, \text{ for all } (k, l) \in \mathcal{T} \text{ s.t. } k \leq i \ \& \ l \leq j.$$

In Case 2, the $n - r$ links that connect each of the worst $n - r$ pairs of the k best are eliminated, but network g' remains connected by connecting node 1 with those $n - r$ nodes j s.t. $s_{1j} \notin H$ directly. Thus, in both cases g' is a connected nested split graph network. Moreover, the rank of the nodes according to degree or number of neighbors is consistent with their rank according to their values, i.e. $v_i > v_j \Rightarrow |N_i(g')| \geq |N_j(g')|$. Thus, g' is a connected *strong* nested split graph network. ■

Proposition 2:

Proof. By Proposition 1, an efficient-constrained-to-keep-connectedness network can be found among those connected SNSG-networks whenever the empty network is not

[12] Notice that $v_1 + v_2$ and $v_1 + v_3$ necessarily belong to H.

efficient. Let g be a connected SNSG-network. Then, if $g_{jk} = 1$ for any $(j,k) \in \mathcal{T}$, with $j \neq 1$, and $\delta^2 s_{jk} > \delta s_{jk} - c$, the elimination of link jk increases the net value because all other pairs of nodes continue to see each other as before the elimination. Therefore, a necessary condition for g to be efficient-constrained-to-keep-connectedness is that $s_{ij} \geq \frac{c}{\delta - \delta^2}$ for all $(j,k) \in \mathcal{T}$, with $j \neq 1$. This yields the necessity of (4-(b)). Now if $g_{jk} = 0$ for any $(j,k) \in \mathcal{T}$, and $\delta^2 s_{jk} < \delta s_{jk} - c$, connecting j and k by a link would increase the net value of the network. This yields the necessity of (4-(a)).

These conditions are also sufficient because any other SNSG-network can be reached from this one by adding and/or deleting links, which in either case cannot increase the net value of the network if (4) holds. ■

Corollary 1:
Proof. Observe that if (4) holds and the node with the smallest value is eliminated, condition (4) implies the corresponding similar condition for the $(n-1)$-node networks with the remaining $n-1$ nodes with their values. Moreover, note that the interval determined by condition (4) may only increase after the elimination of the node with smallest value, because the right-hand side of (4-(b)) cannot decrease after that elimination. ■

Proposition 3:
Proof. Let g be any network s.t. $v(g) > 0$. If g is connected the conclusion follows from Proposition 2. Assume that g is not connected. If g has only one non-trivial component, apply the construction in the proof of Propositions 1 and 2 to that component. If it has 2 or more non-trivial components, apply the same procedure to the subnetwork h formed by the union of all non-trivial components. Then the proof of Proposition 1 adapts easily to form a connected dominant SNSG-network with the nodes in h. There are only two differences. First, in Case 2, i.e. if $\mathcal{S} \subsetneq \mathcal{H}$, i.e. $v_{1i} = v_1 + v_i \notin H$ for all $i > r$, for some r s.t. $3 \leq r \leq n$, whenever $(1,j) \in \mathcal{L} \cap \mathcal{S}$, as j belongs to one of the non-trivial components, there must be some i s.t. $g_{ij} = 1$ whose contribution cannot be greater because $s_{ij} \leq s_{1j}$. The second difference is that $\sum_{(i,j) \in \mathcal{T} \backslash \mathcal{K}} \delta^{d(i,j)} s_{ij}$ contains zeros (as $\delta^{d(i,j)} = 0$ whenever i and j are not in the same component), but this implies

$$v(g) < \delta \Sigma(\mathcal{K}) + \delta^2 \Sigma(\mathcal{T} \backslash \mathcal{K}) = \delta \Sigma(\mathcal{K}) + \delta^2 (\Sigma(\mathcal{T}) - \Sigma(\mathcal{K})).$$

The rest follows the same steps. Let h' be the optimal connected network s.t. (3) and (4) with nodes in h. Then the network consisting of a unique non-trivial component h' plus the remaining isolated nodes in g (if g has isolated nodes) dominates g.

Finally, if $v(g) \leq 0$, then g is dominated by the empty network, which is a trivial strong nested split graph network because it satisfies trivially all conditions for an SNSG-network. ■

Proposition 4:

Proof. (*Necessity*) Assume g is efficient. Then it is obviously efficient-constrained-to-keep-connectedness, which implies (4) by Proposition 2. If $N_n(g) = \{1\}$ and

$$c > \delta(v_1 + v_n) + \delta^2 \sum_{k \in N \setminus \{1,n\}} (v_k + v_n),$$

the elimination of link $1n$ would increase the net value, contradicting the efficiency of g. Thus (5) must hold.

(*Sufficiency*) Let g be a connected SNSG-network s.t. (4) and (5). By Proposition 2, g is efficient-constrained-to-keep-connectedness. Given the structure of g, it is immediate to check that the smaller the value of a node, the smaller its contribution to the net value of the network. If no node has node 1 as its only neighbor then the contribution of node n is sure to be positive because then

$$\delta s_{1n} - c \geq \delta s_{2n} - c \geq \delta^2 s_{2n} > 0.$$

In that case, g is efficient. If there are nodes whose only neighbor is node 1, then the smallest contribution is that of node n, which is

$$\delta(v_1 + v_n) + \delta^2 \sum_{k \in N \setminus \{1,n\}} (v_k + v_n) - c \geq 0,$$

by (5). Thus, the elimination of link $1n$ or any other link of node 1 cannot increase the net value of the network. ■

Proposition 5:
Proof. Let g be a connected SNSG-network and $(c_{ij})_{i,j \in N}$ an investment profile that supports g. Assume $g_{jk} = 0$, then *PSFC*-(*ii*) requires that whenever $0 \leq c_{kj} \leq c$

$$\delta v_j - c_{kj} > \delta^2 v_j \;\Rightarrow\; \delta v_k - (c - c_{kj}) < \delta^2 v_k.$$

In other words, there is no c_{kj} $(0 \leq c_{kj} \leq c)$ s.t.

$$\delta v_j - c_{kj} > \delta^2 v_j \;\;\&\;\; \delta v_k - (c - c_{kj}) \geq \delta^2 v_k,$$

which is so if and only if

$$\delta(v_j + v_k) - c < \delta^2(v_j + v_k).$$

Thus, there is no incentive for any pair of nodes not connected directly to form a link if and only if

$$\frac{c}{\delta - \delta^2} > \max_{(j,k) \in \mathcal{T} \text{ s.t. } g_{jk}=0} (v_j + v_k),$$

which is (10-(a)).

Now assume $(j,k) \in \mathcal{T}$ s.t. $j \neq 1$ and $g_{jk} = 1$. For investment profile $(c_{ij})_{i,j \in N}$ to support g in pairwise equilibrium under free cost sharing it is necessary that $c_{jk}+c_{kj} = c$ and

$$\delta v_j - c_{kj} \geq \delta^2 v_j \quad \text{and} \quad \delta v_k - c_{jk} \geq \delta^2 v_k,$$

otherwise at least one of the nodes j and k would have an incentive to withdraw support for the link. But such a pair c_{jk} and c_{kj} do exist if and only if

$$\delta(v_j + v_k) - c \geq \delta^2(v_j + v_k).$$

Therefore, there exists stable support of a link $(j,k) \in \mathcal{T}$ s.t. $j \neq 1$, if and only if

$$\frac{c}{\delta - \delta^2} \leq \min_{(j,k) \in \mathcal{T},\ j \neq 1\ \&\ g_{jk}=1} (v_j + v_k),$$

which is (10-(b)).

The conditions for the links of node 1 remain to be checked. If $j \neq 1$ has any other neighbors, then by the SNSG-structure of g, node 2 is sure to be one of them. Link $2j$ is stable if and only if $\delta v_2 - c_{j2} \geq \delta^2 v_2$ and $\delta v_j - c_{2j} \geq \delta^2 v_j$, and $c_{j2} + c_{2j} = c$. Such c_{j2} and c_{2j} are certain to exist because of (10-(b)). But $v_1 \geq v_2$, so this implies that there exist also c_{j1} and c_{1j} s.t. $\delta v_1 - c_{j1} \geq \delta^2 v_1$ and $\delta v_j - c_{1j} \geq \delta^2 v_j$, and $c_{j1} + c_{1j} = c$. Finally, if $N_j(g) = \{1\}$, link $1j$ is stable if and only if

$$c_{j1} \leq \delta v_1 + \delta^2 \sum_{k \in N \setminus \{1,j\}} v_k \quad \text{and} \quad c_{1j} \leq \delta v_j,$$

and such pair, c_{j1} and c_{1j}, exists if and only if

$$c \leq \delta v_1 + \delta v_j + \delta^2 \sum_{k \in N \setminus \{1,j\}} v_k.$$

Finally, this condition holds for all j s.t. $N_j(g) = \{1\}$ if and only if it holds for $j = n$, i.e. the node with smallest value, which yields (11). ■

Corollary 3:
Proof. Let g be a connected SNSG-network with no nodes whose only neighbor is the center. Then condition (10) becomes necessary and sufficient for pairwise stability under free cost sharing, and the interval determined by this condition differs from that determined by (4), which becomes necessary and sufficient condition for efficiency, in just one point: the lower bound, included in the latter but not in the first one. ■

Proposition 6:
Proof. (i) The proof is similar to that of part (i) of Proposition 2 in Jackson and Wolinsky (1996). Just choose ij and kl in different components and s.t. $c_{ij} \geq c/2$ and $c_{lk} \geq c/2$.

(ii) For the complete network only condition (10-(b)) of Proposition 5 applies, which becomes (12).

(iii) Only an all-encompassing star centered at a node of greatest value is an SNSG-network. In this case only conditions (10-(a)) and (11) of Proposition 5 apply, which yield (13). ■

References

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

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

[3] Billand, P., C. Bravard and S. Sarangi, 2012a, Existence of Nash networks and partner heterogeneity, *Mathematical Social Sciences* 64, 152–158.

[4] Billand, P., C. Bravard and S. Sarangi, 2012b, On the interaction between heterogeneity and decay in two-way flow models, *Theory and Decision* 73, 525-538.

[5] Bloch, F. and M.O. Jackson, 2006, Definitions of equilibrium in network formation games, *International Journal of Game Theory* 34, 305-318.

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

[7] Galeotti, A., 2006, One-way Flow Networks: The Role of Heterogeneity, *Economic Theory* 29(1), 163–179.

[8] Galeotti, A., S. Goyal and J. Kamphorst, 2006, Network Formation with Heterogeneous Players, *Games and Economic Behavior* 54(2), 353–372.

[9] Goeree, J. K., A. Riedl and A. Ulle, 2009, In search of stars: Network formation among heterogeneous agents, *Games and Economic Behavior* 67(2), 445-466.

[10] Goyal, S., 2007, *Connections. An Introduction to the Economics of Networks*. Princeton University Press. Princeton.

[11] Goyal, S., 2018, Heterogeneity and Networks, in *Handbook of Computational Economics,* Vol. 4, 687-712.

[12] Haller, H. and S. Sarangi, 2005, Nash networks with heterogeneous links, *Mathematical Social Sciences* 50(2), 18-201.

[13] Jackson, M., 2008, *Social and Economic Networks*. Princeton University Press, London. Princeton.

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

[15] Kalai, E., and M. Smorodinsky, 1975, Other solutions to Nash's bargaining problem, *Econometrica*, 43 (3), 513-518.

[16] Kamphorst, J. and G. Van Der Laan, 2007, Network formation under heterogeneous costs: The multiple group model, *International Game Theory Review* 09, No. 04, 599-635.

[17] Koenig, M., C. J. Tessone and Y. Zenou, 2014, Nestedness in Networks: A Theoretical Model and Some Applications, *Theoretical Economics* 9, 695-752.

[18] Mariani, M.S., Z.-M. Ren, J. Bascompte and C. J. Tessone, 2019, Nestedness in complex networks: Observation, emergence, and implications, *Physics Reports* 813, 1-90.

[19] Olaizola, N., and F. Valenciano, 2019, A unifying model of network formation, *International Journal of Game Theory* 47, 1033-1063.

[20] Olaizola, N., and F. Valenciano, 2020a, Dominance of weighted nested split graph networks in connections models, *International Journal of Game Theory* 49, 75-96.

[21] Olaizola, N., and F. Valenciano, 2020b, Characterization of efficient networks in a generalized connections model with endogenous link strength, *SERIEs*, 11(3), 341-367.

[22] Olaizola, N., and F. Valenciano, 2020c, Efficiency and stability in a decreasing returns connections model, Working paper, Ikerlanak IL 114/20.

[23] Souma, W., Y. Fujiwara and H. Aoyama, 2005, Heterogeneous Economic Networks, in *The Complex Networks of Economic Interactions. Essays in Agent-Based Economics and Econophysics*, Editors A. Namatame, T. Kaizouji and Y. Aruka. Spinger.

[24] Vandenbossche, J., and T. Demuynck, 2013, Network formation with heterogeneous agents and absolute friction, *Computational Economics* 42, 23-45.

[25] Vega-Redondo, F., 2007, *Complex Social Networks*. Econometric Society Monographs, Cambridge University Press.