IKERLANAK

# A 'MARGINALIST' MODEL OF NETWORK FORMATION

by

Norma Olaizola and Federico Valenciano

2016

Working Paper Series: IL. 99/16

Departamento de Fundamentos del Análisis Económico I

Ekonomi Analisiaren Oinarriak I Saila

University of the Basque Country

# A 'marginalist' model of network formation*

By Norma Olaizola† and Federico Valenciano‡

August 18, 2016

**Abstract**

We develop a network-formation model where the quality of a link depends on the amount invested in it and is determined by a link-formation "technology", an increasing strictly concave function which is the only exogenous ingredient in the model. The revenue from the investments in links is the information that the nodes receive through the network. Two approaches are considered. First, assuming that the investments in links are made by a planner, the basic question is that of the efficient investments, either relative to a given infrastructure (i.e. a set of feasible links) or in absolute terms. It is proved that efficient networks belong to a special class of weighted nested split graph networks. Second, assuming that links are the result of investments of the node-players involved, there is the question of stability in the underlying network-formation game, be it restricted to a given infrastructure or unrestricted. Necessary and sufficient conditions for stability of the complete and star networks, and nested split graph networks in general, are obtained.

*JEL* Classification Numbers: A14, C72, D85

*Key words:* Network formation, Efficiency, Stability, Nested split graphs, Core-periphery.

---

*We thank Matthew Jackson for his helpful comments. Part of this work was done while the first author was visiting the Department of Economics of Stanford University, whose hopitality and stimulant environment are gratefully acknowledged. We also thank participants of the 2nd Annual Conference on Network Science and Economics, and the Bilbao-Norwich-Málaga Workshop on Economics of Networks (BiNoMa), where preliminary versions of this paper were presented. 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 the 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

This work is a contribution to the literature on economic models of strategic network formation. In this line of work, an increasing flow of research has been contributed by game-theorists and economists in general after Myerson (1977) and Aumann and Myerson (1988).[1] In the wake of these pioneer works in the field, two seminal and most influential models of network formation are Jackson and Wolinsky's (1996) connections model and Bala and Goyal's (2000a) non cooperative model. In both models, networks are the result of creating links between pairs of individuals, be it by bilateral agreements in the first model or unilateral decisions in the second, and information flows through the resulting network. In both models, the cost of a link and its quality (i.e. its decay factor) are exogenously given, giving rise to two-parameter models. The simplicity of these basic models imposes some rigidity: either necessarily bilateral formation and compulsory equal share of the fixed cost of each link, or unilateral formation requiring full-covering of that fixed cost by its creator; and in both cases a fixed level of quality for the resulting link. These parameters, cost and quality of a link, are exogenously given. The point of this work is to provide and develop a more flexible model in both aspects: link-formation and link-performance.[2].

We develop a model of network formation where links are the result of investments, and the quality or strength of a link, i.e. the fidelity-level of its transmission, is never perfect and depends on the amount invested in it. A link-formation "technology" determines the quality of the resulting link as a function of the investment, and is the only exogenous ingredient in the model. Formally, a technology is assumed to be a continuously differentiable, increasing and strictly concave function whose range is $[0, 1)$, i.e. whatever the investment in a link, transmission is never perfect. The revenue from the investment in links is, as in the seminal models, the information that the nodes receive through the network that results. This model poses the following questions that are addressed.

A first approach assumes that the investments in links are made by a social planner. In this scenario there are two basic questions. First, given an "infrastructure" specified by an underlying graph of feasible links, there is the question of existence and determination of an optimal investment (i.e. maximizing the aggregate payoff of the nodes connected by that infrastructure) in these links using them all. A second question is that of efficiency or, equivalently, about which infrastructures are efficient in the sense of maximizing in absolute terms the aggregate payoff for a given technology for an optimal investment in their links. Necessary conditions for the existence of

[1] Goyal (2007), Jackson (2008) and Vega-Redondo (2007) are excellent monographs on social and economic networks. See also Bramoullé, Rogers and Galeotti Eds. (2015).

[2] In Olaizola and Valenciano (2015a), costs and flow levels for links singly-supported and doubly-supported differ and three exogenous parameters specify a synthesizing model. A subsequent paper (2015b) studies the impact of liberalizing cost-sharing in this model. In Section 2 other extensions of the seminal models are commented and compared with the one outlined in this introduction and developed in this paper.

an optimal investment in an infrastructure are established by requiring the marginal aggregate payoff of the investment in every link to be zero. A clear interpretation of the resulting conditions arises from the simplicity and naturalness of the model. As intuition suggests, the greater the amount of information that a link must convey, the greater the investment that efficiency imposes. This result is used to address the question of efficiency in general terms. First, it is proved constructively that any network is dominated by a weighted nested split graph (NSG) structure exhibiting a strong degree of organization beyond the hierarchical character of NSG-networks that we call "strongly nested split graph networks" (SNSG-networks). Then, by imposing optimality conditions we further refine this class and prove that efficient networks are either empty or optimal connected SNSG-networks. This class includes the all-encompassing star and complete networks exhibiting a maximal degree of symmetry as extreme cases.

A second approach assumes that links are the result of investments of the node-players involved, whose reward from forming them is the information they receive through the network that results. Links are formed according to a technology available to all players. In this game-theoretic scenario, there is the question of stability in the underlying network-formation game, be it restricted by an infrastructure or unrestricted. Stability in the sense of Nash equilibrium if coordination is not feasible, or pairwise stability if pairwise coordination is feasible. Necessary conditions for stability are established by imposing the marginal benefit of the investment of any player in each of his/her links to be zero. Again, the conditions resulting from this classical economic condition have a clear intuitive interpretation and allow to characterize stable complete and star networks, and to give necessary conditions for SNSG structures to be pairwise stable. It is shown the existence of pairwise stable SNSG-networks.

Note that the question of efficiency in the game-theoretic scenario is covered by the first approach, and yields the conclusion that efficient structures are not stable, and reciprocally. In addition to these basic issues, this model admits several variations and extensions that are briefly commented in the last section.

A remarkable outcome of this work is the emergence of nested split graph structures from the entirely homogeneous setting laid out by a simple model -at least at its formulation level-, where technology is the only exogenous ingredient. Core-periphery structures, consisting of a "core" of nodes highly or completely interconnected and a "periphery" of nodes not directly connected among them but connected to all or some in the core appear in many different contexts and also in economics.[3] For this reason, a lot of attention has been and continues to be paid in the economic literature to the question of how these structures emerge, and different models have been proposed to this end.[4] Nested split graph networks is a class of highly hierarchical structures covering a whole range of core-periphery degrees with different organizations of the connections of the

[3]See Csermely et al. (2013) for a review of core-periphery structures literature in different fields and a list of general open issues related to them.

[4]König, Tessone and Zenou (2014) includes an excellent review of this literature.

periphery with the core, ranging from the star to the complete network.[5] Since recently, these structures are also receiving attention in the economic literature. In 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.)

The paper is organized as follows. Section 2 briefly reviews some related literature. Section 3 introduces basic notation and terminology. In Section 4 the model is introduced. In Section 5 the question of efficiency, both constrained by an infrastructure and unconstrained, is addressed. Section 6 is devoted to stability. Finally, Section 7 summarizes briefly the results and suggests some possible extensions of the model. All proofs are relegated to the Appendix.

# 2 Related literature

Bloch and Dutta (2009) is possibly the closest model to the one introduced here in spite of the obvious differences. In their model as in ours the strength of a link depends on the investment of the two players involved. Nevertheless, the coincidences do not go beyond this. They assume that players have a fixed endowment, and the link strength is an *additively separable convex* function of individual investments. Not surprisingly the results are completely different. With respect to their assumption about non-decreasing returns to investments they claim that "*While this assumption may seem at odds with the classical literature on productive investments, we strongly believe that convexity is the right assumption to make when one discusses investments in communication links*." (Bloch and Dutta, 2009, p. 42). In this respect, we still believe that concavity, i.e. decreasing returns, is a reasonable assumption about link-formation technology.[6] In this point there is a similarity with several models in the networks literature, where the payoff function is often assumed concave, embodying a decreasing returns assumption. In Ballester et al. (2006), noncooperative games with linear-quadratic utilities where each player chooses his/her effort are interpreted as a network game in which payoffs are concave in own effort, global interaction effect is uniform and local complementarities are captured by a network. In Bramoullé and Kranton (2007), where the network is given and players choose their efforts, the payoff of each player is a strictly concave function of the efforts of the player and the player's neighbors. In Hojman and Szeidl's

[5] This is briefly discussed in Section 5.

[6] Our model seems a more natural extension of the basic models of Jackson and Wolinsky (1996) and Bala and Goyal (2000a). Bloch and Dutta (2009) claim that the literature on discrete link formation assumes an extreme form of convexity. But an echelon-function of the form

$$\phi(x) = \begin{cases} s, & \text{if } x \geq c, \\ 0, & \text{if } x < c, \end{cases}$$

is an equally extreme (if at all) form of concavity. In fact, such function can be seen as a limit case of one of the possible extensions of our model briefly commented in the last section.

(2008) model players choose their links, and a player's payoff is a strictly increasing and concave function of a weighted sum of the number players at different distances, with weights decreasing with distance. Thus the link-formation technology is the discrete one of the seminal models, w.r.t. which the crucial difference is in the payoff function, which embodies the decreasing returns assumption. The difference with our model is apparent, in it the non-discrete link-formation technology is the only exogenous ingredient, while the logic of the rest is as in the seminal models: *payoff* = *information* - *cost*. A comparison with Galeotti and Goyal (2010), who assume that returns from information (which can be acquired personally or through connections) are increasing and concave, while the costs of personally acquired information are linear, draws to a similar conclusion. In all these models the strict concavity assumption is placed in the payoff function, while in ours it is placed in the technology, which is the only exogenous element in the model.

On the other hand, weighted nested split graphs play a most relevant role in this work. These networks, as mentioned before, exhibit a high degree of hierarchy and have, in their non-weighted version, recently attracted the attention of economists.[7] In a recent paper König, Tessone and Zenou (2014) develop a dynamic network formation model to explain the observed nestedness in real-world networks and use stochastic stability to show the convergence to nested split graphs. More recently, Belhaj, Bervoets and Deroïan (2016), address the problem of a planner looking for the efficient network when agents play a network game with local complementarities choosing their effort levels, and links are costly. They show that the efficient networks in this sense are nested split graph networks, under different cost functions. The differences with the model developed here are clear. We address the question of efficiency and stability separately. In the first scenario, the planner chooses the weighted network, in the second, players are the only actors. A further development of our model would address the question of a planner fixing the infrastructure so as to maximize the social welfare in equilibrium.

However, in our opinion, beyond the obvious differences, the most remarkable difference with these more sophisticated models is the simplicity and naturalness of ours, directly elaborating from the basic model of Jackson and Wolinsky (1996) and Bala and Goyal (2000), by replacing a discreet cost function by a decreasing returns link-formation technology.

# 3 Preliminaries

An *undirected weighted 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 non-negative

[7]To the best of our knowledge, this is the first economic paper where *weighted* nested split graph networks appear. In fact, we were not aware of the existence of the notion of nested split graph network in the literature until Matthew Jackson mentioned it in a seminar at Stanford where the first author presented a preliminary version of this work.

real numbers and $g_{ii} = 0$. Alternatively, $g$ can, and often will, be interpreted as a map $g : N_2 \longrightarrow \mathbb{R}_+$, where $N_2$ denotes the set of all subsets of $N$ with cardinality 2, and $\mathbb{R}_+$ denotes the set of non-negative real numbers. In the sequel $\underline{ij}$ stands for $\{i, j\}$ and $g_{\underline{ij}}$ for $g(\{i, j\})$ for any $\{i, j\} \in N_2$.[8] 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* $g_{\underline{ij}}$ connects $i$ and $j$. $N^d(i, g) := \{j \in N : g_{\underline{ij}} > 0\}$ denotes the set of *neighbors* of node $i$, and its cardinality $\#N^d(i, g)$ is the *degree* of $i$. $N(i, g)$ denotes the set of nodes connected with $i$ by a *path*, i.e. a sequence of distinct nodes s.t. every two consecutive nodes are connected by a link. 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. A graph has a *cycle* if there are two nodes connected by a link and also by a path of length 2 or more (the *length* of a path is the number of links it contains, i.e. that of nodes minus 1).

Undirected graphs, weighted or not, underlie a variety of situations were actual links mean some sort of reciprocal connection or relationship. Such structures are commonly referred to as *networks*. As behind a network there always lies a graph as a most salient feature, 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.

The *empty network* is the 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 $\underline{ij} \in N_2$.[9] A *tree* is a connected network with no cycles. An *all-encompassing star* consists of a network with $n - 1$ links in which a node (the *center*) is connected with each of the remaining nodes by a link. A node is *peripheral* in a network if it is involved in one link only. An important class of networks consists of those whose underlying graph is a "nested split graph", which exhibit a strict hierarchical structure where nodes can be ranked by their number of neighbors.

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

$$\#N^d(i, g) \leq \#N^d(j, g) \Rightarrow N^d(i, g) \subseteq N^d(j, g) \cup \{j\}.$$

Networks whose underlying graph is nested split are referred to as (weighted or not) *NSG-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 that of those

[8] The convenience of the distinction between $ij$ and $\underline{ij}$, especially as subindices, will soon be apparent. With this convention $g_{\underline{ij}} = g_{\underline{ji}}$, while $g_{ij} \neq g_{ji}$ in general.

[9] Note that there exist infinite complete *weighted* networks, but only one non-weighted complete network.

in the preceding row. Nodes are then classified in classes, each of them containing the nodes with equal number of neighbors, referred to as *NSG-classes*[10].

# 4 The model

The main ingredient in the model that has been briefly sketched in the introduction is a link-formation technology.

**Definition 2** *A link-formation technology is a continuously differentiable map* $\delta : \mathbb{R}_+ \to [0, 1)$ *s.t.* $\delta(0) = 0$*, and satisfies the following conditions:*
$(C.1)$ $\delta'(c) > 0$*, for all* $c \geq 0$*, i.e. it is increasing.*
$(C.2)$ *It is strictly concave.*

The interpretation of this function and the assumptions are the following. If $c$ is the amount invested in a link to connect two *nodes* or *players,* $\delta(c)$ is interpreted as the level of fidelity of the transmission of information through the link.[11] More precisely, $\delta(c)$ is the fraction of information flowing through the link that remains intact.[12] Flow occurs only through links invested in ($\delta(0) = 0$), but a perfect fidelity in transmission between different nodes is never reached ($0 \leq \delta(c) < 1$). The smoothness of $\delta$ makes the use of differential calculus possible, which allows for a relatively simple formal "marginal analysis" without getting involved in sophisticated technical issues. Condition C.1 is quite plausible: the quality of a link is increasing with the investment in it. C.2 is a reasonable condition, at least "in the long run" given that $\delta'(c) > 0$ and the range of $\delta$.[13]

Based on this basic ingredient we consider two different scenarios.

**Scenario 1**: A set $N = \{1, 2, ..., n\}$ of nodes can be connected by links according to a link-formation technology. Investments are made by a central planner. A *link-investment vector* is an $n(n-1)/2$-vector, $\overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij} \in N_2}$, where $c_{\underline{ij}}$ denotes the investment in link $\underline{ij} \in N_2$ through which the fidelity-level is $\delta(c_{\underline{ij}})$. Then $\delta^{\overline{\mathbf{c}}} = (\delta(c_{\underline{ij}}))_{\underline{ij} \in N_2}$ denotes the resulting *weighted network*. For a given link-investment

[10]Isolated nodes, i.e. with no neighbors, form a class that plays no relevant role and is referred to as the *trivial class*.

[11]We often prefer the term 'node' to avoid a biased language. Moreover, in the first of the two scenarios that we presently describe it is more appropriate to speak of nodes given their passive role.

[12]Nevertheless, other interpretations are possible. For instance, as a degree of reliability, as in Bala and Goyal (2000b), or the "strength of a tie" (Granovetter, 1973), i.e. a measure of the quality/intensity/value of a relationship as e.g. in personal relationships, where 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, but we give preference here to the interpretation in terms of information.

[13]It would also be reasonable to assume $\delta$ to be convex up to a certain value of $c$, and concave beyond an inflection point. This and other variations of the model are considered in the concluding section.

vector $\overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij} \in N_2}$, a node $i$ receives from another node's value $v$ the fraction that reaches $i$ through the best possible route in the weighted network $\delta^{\overline{\mathbf{c}}}$.[14] Let $\mathcal{P}_{ij}(\delta^{\overline{\mathbf{c}}})$ denote the set of paths in $\delta^{\overline{\mathbf{c}}}$ connecting $i$ and $j$. For $p \in \mathcal{P}_{ij}(\delta^{\overline{\mathbf{c}}})$, let $\delta^{\overline{\mathbf{c}}}(p)$ denote the resulting fidelity-level determined by the product of the fidelity-levels through each link in that path. Then, $i$ values information originating from $j$ that arrives via $p$ by $v\delta^{\overline{\mathbf{c}}}(p)$. If information is routed via the best possible route from $j$ to $i$, then $i$'s valuation of the information originating from $j \neq i$ is

$$I_{ij}(\delta^{\overline{\mathbf{c}}}) = \max_{p \in \mathcal{P}_{ij}(\delta^{\overline{\mathbf{c}}})} v\delta^{\overline{\mathbf{c}}}(p) = v \max_{p \in \mathcal{P}_{ij}(\delta^{\overline{\mathbf{c}}})} \delta^{\overline{\mathbf{c}}}(p),$$

and $i$'s overall revenue from $\delta^{\overline{\mathbf{c}}}$ is

$$I_i(\delta^{\overline{\mathbf{c}}}) = \sum_{j \in N(i;\delta^{\overline{\mathbf{c}}})} I_{ij}(\delta^{\overline{\mathbf{c}}}).$$

The *value* of the network resulting from a link-investment vector $\overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij} \in N_2}$ is the aggregate payoff, i.e. the total value of the information received by the nodes minus the cost of the network:

$$v(\delta^{\overline{\mathbf{c}}}) := \sum_{i \in N} I_i(\delta^{\overline{\mathbf{c}}}) - \sum_{\underline{ij} \in N_2} c_{\underline{ij}}. \tag{1}$$

In this setting two questions are addressed. First, the problem of determining the link-investment vector that maximizes the aggregate value for a given "infrastructure" specified by an underlying graph of feasible links. Second, the characterization of efficient networks/link-investment vectors that maximize the aggregate value in absolute terms.

**Scenario 2**: Let $N$, $\delta$ and $v$ be as in Scenario 1, but now the nodes/players form the links by investing in them. The quality of a link depends on the total amount invested in it by the two players it connects, and it is assumed that a link-formation technology $\delta$ is available to all agents and determines the quality of a link as a function of the investment in it. An *investment profile* is specified by a matrix $\mathbf{c} = (c_{ij})_{i,j \in N}$, where $c_{ij} \geq 0$ (with $c_{ii} = 0$) is the investment of player $i$ in the link connecting $i$ and $j$, and determines a link-investment vector $\overline{\mathbf{c}}$

$$\mathbf{c} \to \overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij} \in N_2} \text{ s.t. } c_{\underline{ij}} := c_{ij} + c_{ji}.$$

The available link-formation technology, $\delta$, yields a *weighted network* for each investment profile $\mathbf{c}$. Namely, $\delta^{\mathbf{c}} := \delta^{\overline{\mathbf{c}}}$, where

$$\delta^{\mathbf{c}}_{\underline{ij}} = \delta^{\overline{\mathbf{c}}}_{\underline{ij}} = \delta(c_{\underline{ij}}) = \delta(c_{ij} + c_{ji}),$$

[14] We assume homogeneity in values. A more general model would assume this value to depend on the players.

whenever $i \neq j$. Thus, $i$'s *payoff* is the value of the information received by $i$ minus $i$'s investment:

$$\Pi_i^\delta(\mathbf{c}) = I_i(\delta^{\mathbf{c}}) - C_i(\mathbf{c}) = I_i(\delta^{\bar{\mathbf{c}}}) - \sum_{j \neq i} c_{ij}. \tag{2}$$

Note that a game in strategic form, where a *strategy* of a player is an $(n-1)$-vector of investments and the *payoff function* is given by (2), is implicitly defined. Therefore, the notion of Nash equilibrium can be applied: an investment profile is *Nash stable* if no player has an incentive to change his/her investments' vector. In this context, if pairwise coordination is feasible, the following is a natural adaptation of the notion of pairwise stability of Jackson and Wolinsky (1996) to this setting:

**Definition 3** *An investment profile* $\mathbf{c}$ *is pairwise stable if:* $(i)$ *no player can improve his/her payoff by changing the investment in any of his/her links, and* $(ii)$ *for any two players* $i$ *and* $j$*, and any* $\mathbf{c}'$ *s.t.* $c'_{kl} = c_{kl}$ *for all* $\underline{kl} \neq \underline{ij}$, $c'_{ik} = c_{ik}$ *for all* $k \neq j$*, and* $c'_{jk} = c_{jk}$ *for all* $k \neq i$*:*

$$\Pi_i^\delta(\mathbf{c}') > \Pi_i^\delta(\mathbf{c}) \quad \Rightarrow \quad \Pi_j^\delta(\mathbf{c}') < \Pi_j^\delta(\mathbf{c}).$$

A joint refinement of both stability notions consists of requiring both.[15]

**Definition 4** *An investment profile* $\mathbf{c}$ *is pairwise Nash stable if it is Nash stable and pairwise stable.*

The question of stability is addressed, first for investment profiles constrained by an infrastructure and then unconstrained in general terms.

We first consider Scenario 1 and address the question of efficiency in relative and absolute terms. Later, the question of stability is addressed in Scenario 2. All results assume a link-formation technology as specified in Definition 2. Thus we deal with a model with three "parameters": the number of nodes/players $n$, the value $v$ of the information at each node, and the link-formation technology represented by function $\delta$.[16]

[15] This strong version of pairwise stability was suggested by Jackson and Wolinsky (1996) and applied by Goyal and Joshi (2003) and Belleflamme and Bloch (2004) among others. See also Bloch and Jackson (2006) for a discussion of different notions of equilibrium in network formation and references therein.

[16] It can be assumed w.l.o.g. $v = 1$, which slightly simplifies the presentation. Nevertheless, it is preferable not to do so and keep explicit this otherwise hidden parameter. In Scenario 1 this value can be interpreted as a subjective evaluation of the planner w.r.t. which the efficiency objective is specified. Nevertheless, the reader may choose to ignore all occurrences of $v$ by assuming $v = 1$.

# 5 Scenario 1: Efficiency

Let $\overline{\mathbf{c}}$ and $\overline{\mathbf{c}}'$ be two link-investment vectors and $v(\delta^{\overline{\mathbf{c}}})$ and $v(\delta^{\overline{\mathbf{c}}'})$ their values as defined by (1). If $v(\delta^{\overline{\mathbf{c}}}) \geq v(\delta^{\overline{\mathbf{c}}'})$ we say that $\delta^{\overline{\mathbf{c}}}$ *dominates* $\delta^{\overline{\mathbf{c}}'}$ (or that $\overline{\mathbf{c}}$ dominates $\overline{\mathbf{c}}'$). Network $\delta^{\overline{\mathbf{c}}}$ (or link-investment vector $\overline{\mathbf{c}}$) is said to be *efficient* if it dominates any other.[17] Thus efficiency can be seen as the goal of a planner investing in links with the objective of maximizing the social welfare, i.e. the aggregate value received by the nodes minus the total cost of the network. We use the following notation: for all $i, j \in N, i \neq j$, $\overline{p}_{ij}$ denotes an optimal path connecting them (note it may be not unique), i.e. one for which the resulting fidelity-level is maximal:

$$\delta^{\overline{\mathbf{c}}}(\overline{p}_{ij}) = \max_{p \in \mathcal{P}_{ij}(\delta^{\overline{\mathbf{c}}})} \delta^{\overline{\mathbf{c}}}(p),$$

where $\delta^{\overline{\mathbf{c}}}(p)$ is the product of the fidelity-levels of the links forming path $p$ for the link-investment vector $\overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij} \in N_2}$, and the aggregate payoff for this link-investment vector is

$$v(\delta^{\overline{\mathbf{c}}}) = \sum_{i \in N} I_i(\delta^{\overline{\mathbf{c}}}) - \sum_{\underline{ij} \in N_2} c_{\underline{ij}} = 2v \sum_{\underline{ij} \in N_2} \delta^{\overline{\mathbf{c}}}(\overline{p}_{ij}) - \sum_{\underline{ij} \in N_2} c_{\underline{ij}}. \quad (3)$$

Note that in principle the last expression in (3) may not be unique. This occurs if for some pair of nodes the optimal path connecting them is not unique.

We make use of the following notation: if $\overline{p}_{kl}$ is an optimal path connecting nodes $k$ and $l$ for a link-investment vector $\overline{\mathbf{c}}$ that contains link $\underline{ij}$, $\delta(\overline{p}_{kl}^{\underline{ij}})$ denotes the product of fidelity-levels for the path that would result from replacing in $\overline{p}_{kl}$ the link $c_{\underline{ij}}$ by a perfect one, i.e. by $c'_{\underline{ij}} = 1$. In other terms:

$$\delta(\overline{p}_{kl}^{\underline{ij}}) = \frac{\delta^{\overline{\mathbf{c}}}(\overline{p}_{kl})}{\delta(c_{\underline{ij}})}.$$

Before addressing the question of efficiency in absolute terms, we address the problem of maximizing the aggregate payoff for a given "infrastructure" specified by an underlying graph of feasible links.

**Definition 5** *An infrastructure is a non-empty subset $S \subseteq N_2$ which specifies the set of links which must be invested in. And a link-investment vector $\overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij} \in N_2}$ is optimal for an infrastructure $S$ if for all $\underline{ij} \in N_2$, $c_{\underline{ij}} > 0$ if and only if $\underline{ij} \in S$, and maximizes the aggregate payoff given this constraint.*

Then the following result establishes necessary conditions for a link-investment vector to be optimal for an infrastructure.

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

**Lemma 1** *For a link-investment vector $\overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij} \in N_2}$ to be optimal for a given infrastructure $S \subseteq N_2$, the following conditions are necessary: $(i)$ For any two nodes connected in $S$ there is a unique optimal path connecting them, and $(ii)$ For all $\underline{ij} \in S$,*

$$\delta'(c_{\underline{ij}}) = \frac{1}{2v \sum_{\underline{kl} \in N_2 \ s.t. \ \underline{ij} \in \overline{p}_{kl}} \delta(\overline{p}_{kl}^{ij})}. \tag{4}$$

**Proof.** Let $\overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij} \in N_2}$ be a link-investment vector s.t. $c_{\underline{ij}} > 0$ if and only if $\underline{ij} \in S$. We prove first part $(ii)$. Assume $\overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij} \in N_2}$ to be optimal for $S$, and $\underline{ij} \in S$, i.e. $c_{\underline{ij}} > 0$. Then link $\underline{ij}$ is part of at least one optimal path in $\delta^{\overline{\mathbf{c}}}$, the one connecting $i$ and $j$, otherwise $\overline{\mathbf{c}}$ would not be optimal.[18] Then, in any of the possibly different but equivalent expressions of the right-hand side of (3), $\delta(c_{\underline{ij}})$ would appear at least once, i.e. $\delta^{\overline{\mathbf{c}}}(\overline{p}_{ij}) = \delta(c_{\underline{ij}})$ if $\overline{\mathbf{c}}$ is optimal, and possibly also in the product yielding $\delta^{\overline{\mathbf{c}}}(\overline{p}_{kl})$ for other pairs of nodes $k, l$. Fix any choice of these (possibly multiple) optimal paths for every two connected nodes, and let the aggregate payoff be given by the right-hand side of (3), which is an up to $n(n-1)/2$-variable function with partial derivatives. Even if a slight modification of some $c_{\underline{ij}} > 0$ might cause a path not to be optimal, a non-null partial derivative w.r.t. $c_{\underline{ij}}$ means that by slightly increasing (if it were $> 0$) or decreasing (if it were $< 0$) the investment in link $\underline{ij}$ the aggregate payoff through those the same paths, optimal or not but still available, would surely increase, which contradicts $\overline{\mathbf{c}}$'s optimality for $S$. Then the partial derivative of the right-hand side of (3) w.r.t. $c_{\underline{ij}}$ must be 0, i.e.

$$\frac{\partial}{\partial c_{\underline{ij}}}(2v \sum_{\underline{kl} \in N_2} \delta^{\overline{\mathbf{c}}}(\overline{p}_{kl}) - \sum_{\underline{kl} \in N_2} c_{\underline{kl}}) = 2v\delta'(c_{\underline{ij}}) \sum_{\underline{kl} \in N_2 \ s.t. \ \underline{ij} \in \overline{p}_{kl}} \delta(\overline{p}_{kl}^{ij}) - 1 = 0,$$

which yields (4).

$(i)$ Assume that two nodes $r$ and $s$ are connected by two different optimal paths in $\delta^{\overline{\mathbf{c}}}$. Then there is at least one link, say $\underline{ij}$, that is part of one of these paths but not of the other. Then the right-hand side of (4) admits at least *two* different expressions: one where the optimal path between any pair of nodes $k, l$ is $\overline{p}_{kl}$, and another one where it is $\overline{q}_{kl}$, and such that for any pair $k, l$ different from pair $r, s$, $\overline{p}_{kl} = \overline{q}_{kl}$, while for $r$ and $s$ the optimal path is different, i.e. $\overline{p}_{rs} \neq \overline{q}_{sr}$, and only the first one contains $\underline{ij}$. In that case,

$$\frac{1}{2v \sum_{\underline{kl} \in N_2 \ s.t. \ \underline{ij} \in \overline{p}_{kl}} \delta^{\overline{\mathbf{c}}}(\overline{p}_{kl}^{ij})} \neq \frac{1}{2v \sum_{\underline{kl} \in N_2 \ s.t. \ \underline{ij} \in \overline{q}_{kl}} \delta^{\overline{\mathbf{c}}'}(\overline{q}_{kl}^{ij})}$$

because

$$\sum_{\underline{kl} \in N_2 \ s.t. \ \underline{ij} \in \overline{p}_{kl}} \delta^{\overline{\mathbf{c}}}(\overline{p}_{kl}^{ij}) - \sum_{\underline{kl} \in N_2 \ s.t. \ \underline{ij} \in \overline{q}_{kl}} \delta^{\overline{\mathbf{c}}'}(\overline{q}_{kl}^{ij}) = \delta^{\overline{\mathbf{c}}}(\overline{p}_{rs}) > 0,$$

which leads to a contradiction because (4) yields two different values for $\delta'(c_{\underline{ij}})$. ■

[18] Note that $\delta(c_{\underline{ij}})$ actually appears in (3) only if $c_{\underline{ij}} > 0$.

**Comment:** Part $(i)$ establishes that in a network that results from an optimal link-investment vector for an infrastructure any two nodes "see" each other through a unique best path connecting them. This is a consequence of part $(ii)$, where (4) establishes a condition for a link-investment vector to be efficient which is the result of *requiring the marginal aggregate payoff of the investment in every link to be zero.* The intuition behind the resulting condition (4) is also clear. The denominator of the right-hand side of (4) is $2v$ times the sum of the fidelity-levels through all optimal paths of which link $\underline{ij}$ is part of (discounting link $\underline{ij}$, i.e. divided by $\delta(c_{\underline{ij}})$). In other words, *the actual flow of information that link* $\underline{ij}$ *conveys.* Thus this sum can thus be interpreted as a *measure of the importance of link* $\underline{ij}$ *in network* $\delta^{\overline{\mathbf{c}}}$. The interpretation of (4) is then clear: the greater this measure, the smaller $\delta'(c_{\underline{ij}})$ and consequently the greater $c_{\underline{ij}}$, that is, the greater the investment that optimality imposes in that link.

The following result establishes necessary conditions for a link-investment vector to be efficient in absolute terms.

**Proposition 1** *For a link-investment vector* $\overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij} \in N_2}$ *to be efficient the following conditions are necessary:* $(i)$ *For any two connected nodes there exists a unique optimal path connecting them.*
$(ii)$ *For each* $\underline{ij} \in N_2$ *s.t.* $c_{\underline{ij}} > 0$, *condition* (4) *holds.*
$(iii)$ *For each* $\underline{ij} \in N_2$ *s.t.* $c_{\underline{ij}} = 0$, *if* $\delta'(0) > 1/2v$,

$$2v\delta^{\overline{\mathbf{c}}}(\overline{p}_{ij}) \geq 2v\delta(c^{\sharp}) - c^{\sharp}, \tag{5}$$

*where*

$$c^{\sharp} = \arg\max(2v\delta(c) - c). \tag{6}$$

**Proof.** Assume $\overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij} \in N_2}$ to be efficient, then $\overline{\mathbf{c}}$ must be optimal for the infrastructure $S = \{\underline{ij} \in N_2 : c_{\underline{ij}} > 0\}$. Then $(i)$ and $(ii)$ follow immediately from Lemma 1.

$(iii)$ Assume $c_{\underline{ij}} = 0$. Then no investment in link $\underline{ij}$ can increase the aggregate payoff, that is, for all $c > 0$, $2v\delta^{\overline{\mathbf{c}}}(\overline{p}_{ij}) \geq 2v\delta(c) - c$, otherwise investing $c$ in link $\underline{ij}$ would surely increase the aggregate payoff. This yields condition (5). Note that condition (5) has a bite only if $\delta'(0) > 1/2v$, otherwise, by the assumptions about function $\delta$, $2v\delta(c) - c < 0$ for all $c > 0$. ■

**Comment:** Parts $(i)$ and $(ii)$ are immediate consequences of Lemma 1. Condition $(iii)$ is a necessary condition for the existence of efficient link-investment vectors where some link receives no investment: no investment in that link must be profitable for the two nodes because it would increase the aggregate payoff.

As a first step to identify the efficient structures, the following key result shows that any connected 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 we have chosen for them: "strongly NSG-graphs/networks" (SNSG-graphs/networks). As any undirected graph, a weighted NSG is completely specified by the triangular matrix above the main diagonal of 0-entries of its adjacency matrix for a certain order of the nodes. In terms of this triangular submatrix, SNSG-networks can be described as weighted NSGs where: (i) all positive entries in the first row are greater or equal than any other entries, (ii) each row consists of a non-decreasing sequence of positive entries followed by zeros, and (iii) from the second row downwards on, non-zero entries in the same column form a non-decreasing sequence.

Formally, we have the following definition.

**Definition 6** *A strongly nested split graph* (*SNSG*) *network is a weighted NSG-network $g$ such that, for a certain order of the nodes, satisfies the following conditions:* ($i$) *for all $i, j, k$ s.t. $1 < i < j < k$, $g_{\underline{1j}} \geq g_{\underline{ik}}$;* ($ii$) *for all $i, j, k$ s.t. $i < j < k$, $g_{\underline{ik}} > 0 \Rightarrow g_{\underline{ij}} \leq g_{\underline{ik}}$;* ($iii$) *for all $i > 1$, $g_{\underline{ij}} \leq g_{\underline{i+1,j}}$, whenever $g_{\underline{i+1,j}} \neq 0$ and $i + 1 < j$.*

Note that, as $\delta$ is strictly increasing, the adjacency matrix $(g_{ij})_{i,j \in N}$, with $g_{ij} = \delta(c_{\underline{ij}})$, and matrix $(c_{\underline{ij}})_{i,j \in N}$ have *identical* underlying NSG infrastructure, moreover one is SNSG if and only if the other is. The following result is crucial for the characterizing result.

**Lemma 2** *Any connected network $\delta^{\overline{\mathbf{c}}}$, where $\overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij} \in N_2}$, is dominated either by the empty network or by a connected SNSG-network.*

**Proof.** Let $\overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij} \in N_2}$ be a link-investment vector such that $\delta^{\overline{\mathbf{c}}}$ is connected, with $q$ links and positive aggregate payoff (if it were negative it would be dominated by the empty network). Starting from $\overline{\mathbf{c}}$, we describe an algorithm to construct a new link-investment vector that yields a dominant SNSG-network, $\overline{\mathbf{c}}' = (c'_{\underline{ij}})_{\underline{ij} \in N_2}$, as the final outcome of a sequence of link-investment vectors $\overline{\mathbf{c}}'_1, \overline{\mathbf{c}}'_2, \ldots$, each of them resulting from the preceding one by adding at most one link and perhaps reassigning those introduced so far after Step 1.

*Step 1*: Let $\overline{\mathbf{c}}'_1$ be the star that results from connecting node 1 with the other $n-1$ investing in each link exactly the same amount invested in each of the strongest $n-1$ links in $\overline{\mathbf{c}}$ and so that $c'_{\underline{12}} \leq c'_{\underline{13}} \leq \ldots \leq c'_{\underline{1n-1}} \leq c'_{\underline{1n}}$. And let $\overline{\mathbf{c}}_1$ be the result of eliminating in $\overline{\mathbf{c}}$ the $n-1$ strongest links.

*Step 2*: From now on proceed as follows with the current $\overline{\mathbf{c}}'_t$ to form $\overline{\mathbf{c}}'_{t+1}$: choose two of the nodes, $i$ and $j$, worst connected for $\overline{\mathbf{c}}_t$, in the first iteration nodes 2 and 3, and 2 and 4 in the second, and later on, to avoid ambiguity in case there are multiple equally worst connected pairs, replace any of them with the only condition of preserving the NSG underlying infrastructure,[19] and check whether the contribution to the value of the network of the connection of $i$ and $j$ via node 1 in $\overline{\mathbf{c}}'_t$ can or cannot be improved by connecting them with the worst available link, say $c_{\underline{kl}}$, in $\overline{\mathbf{c}}_t$. That is, check whether $2v\delta(c'_{\underline{1i}})\delta(c'_{\underline{1j}}) < 2v\delta(c_{\underline{kl}}) - c_{\underline{kl}}$. Then,

[19] That is, in terms of the triangular submatrix of the adjacency matrix above the main diagonal: give priority for the replacement to any one as much as possible to the left among those in the same row and to the uppermost among those in the same column.

- If it means an improvement, then connect them, i.e. make $c'_{\underline{ij}} := c_{\underline{kl}}$, $\overline{\mathbf{c}}'_{t+1} := \overline{\mathbf{c}}'_t + c'_{\underline{ij}}$ and $\overline{\mathbf{c}}_{t+1} := \overline{\mathbf{c}}_t - c_{\underline{kl}}$.

- Otherwise, by construction (available links in $\overline{\mathbf{c}}_t$ are picked in increasing strength order) $c_{\underline{kl}}$ is necessarily at least as good as the last link added; then replace it by $c_{\underline{kl}}$ and proceed similarly replacing the previous connection by the newly available link, and so on backwards.[20] This procedure leads to discarding in $\overline{\mathbf{c}}_t$ the weakest link of the added ones, currently connecting nodes 2 and 3. Let then $\overline{\mathbf{c}}_{t+1}$ be the result of this elimination in $\overline{\mathbf{c}}_t$, and let $\overline{\mathbf{c}}'_{t+1}$ be the updated network.

In all cases, go back to Step 2 unless $\overline{\mathbf{c}}_{t+1}$ is empty (i.e. there no remain any available links), then Stop.

Obviously the process ends in $1 + q - (n - 1) = q - n + 2$ iterations, when $\overline{\mathbf{c}}_t$ is the empty network and no links remain. Then, we show that if $\overline{\mathbf{c}}' = \overline{\mathbf{c}}'_{m-n+2}$, we have that $v(\delta^{\overline{\mathbf{c}}'}) \geq v(\delta^{\overline{\mathbf{c}}})$. As both $\delta^{\overline{\mathbf{c}}}$ and $\delta^{\overline{\mathbf{c}}'}$ are connected, $v(\delta^{\overline{\mathbf{c}}})$ and $v(\delta^{\overline{\mathbf{c}}'})$ are the sum of $n(n-1)/2$ terms, one for each pair of nodes. Each of these terms corresponds to one pair of nodes and gives the contribution to the aggregate payoff of the value that they receive from each other (minus the cost of the link if they are directly connected). In the first step, by organizing the $n-1$ strongest links as a star, they connect all nodes: $n-1$ pairs directly and all other pairs by the maximal number $((n-1)(n-2)/2)$ of two-link connections using these strongest links. From then on, by the way in which $\overline{\mathbf{c}}'$ has been formed, when an "available link" in $\overline{\mathbf{c}}$ (in $\overline{\mathbf{c}}_t$ at stage $t$) of cost $c$ is discarded this is because its direct contribution (i.e. term $2v\delta(c) - c$ in the sum that yields $v(\delta^{\overline{\mathbf{c}}})$) is *smaller or equal than* any term in the sum giving $v(\delta^{\overline{\mathbf{c}}'_t})$. As after Step 1, once the initial star has been formed, links added to form $\overline{\mathbf{c}}'$ are increasingly strong, one discarded link will never be missed. As a result, the aggregate value of the resulting network cannot be smaller than that of the initial one.

As links are added to the initial star (formed by the strongest links) in increasing strength order and always preserving an SNSG infrastructure, new links added corresponding to entries in the same row (column) in the triangular matrix are of increasing strength rightwards (downwards). Therefore the outcome is an SNSG-network. Thus any connected network which yields a positive aggregate payoff is dominated by such a structure. ■

**Comments:** (i) This lemma is the key for the characterizing result about efficiency. The idea of the constructive proof is based on the idea of rearranging the "available links" in any given network in a most efficient way. By its interest, we outline the proof formalized precisely in the Appendix. Given a link-investment vector, start by forming an all-encompassing star with the *strongest* $n-1$ links. If no links remain stop (this occurs if $m = n-1$, i.e. the starting network is a tree); otherwise, take the *weakest* of the remaining available links and connect the two worst connected peripheral nodes in this star with it if this improves the contribution to the value of the network of their

[20] Note for this we just need to keep track of the order in which new links have been added to the star formed at Step 1.

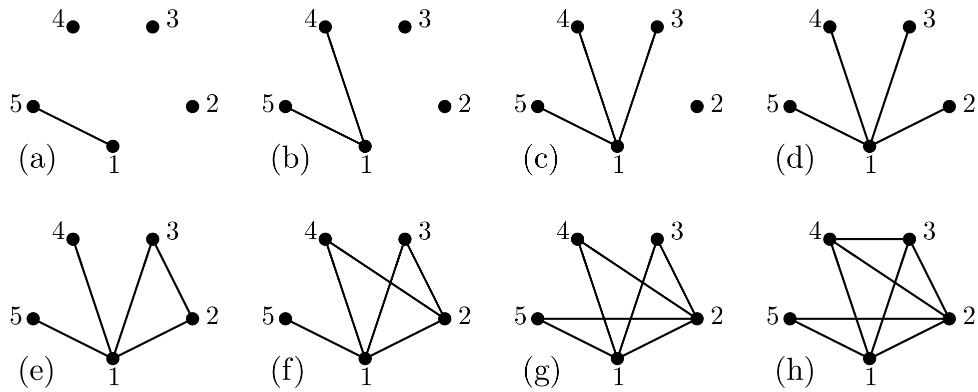

Figure 1: Construction of a dominant SNSG-network

connection, otherwise use it to improve the new connections made so far by replacing the last link added with it, and proceed similarly by replacing the previous connection by the newly available link, and so on backwards. Now repeat the procedure with a new of the weakest available links and a new pair of worst connected nodes in the network under construction up to no available link remains. At the end of this process, an SNSG-network which yields a greater or equal aggregate value as that of the initial network at less or equal cost arises. Figure 1 shows this process starting from a 6-node network with 8 links, assuming that at every stage the weakest available link actually improves the worst connection. First, form a star with the best 4 links in *decreasing* strength order: $c_{\underline{15}} \geq c_{\underline{14}} \geq c_{\underline{13}} \geq c_{\underline{12}}$ (stages (a)-(d)), then, using available links in *increasing* strength order, improve the weakest connection: $c_{\underline{23}} \leq c_{\underline{24}}$ (stages (e)-(f)). At stage (f), connect the pair worse connected, that is, 25 or 34. If it were the first, i.e. if $\delta(c_{\underline{15}})\delta(c_{\underline{12}}) \leq \delta(c_{\underline{13}})\delta(c_{\underline{14}})$, then connect them (stage (g)). Finally, connect the worse connected in (g), which are nodes 3 and 4 (stage (h)).

(ii) Therefore, for a certain relabeling of the nodes, a dominant SNSG-network can be seen as consisting of a star centered at node 1 (first row and column of the adjacency matrix), plus some additional links among spoke nodes of that star (nonzero entries on the northwest of the adjacency matrix).

(iii) Although Lemma 2 assumes $\delta^{\overline{\mathbf{c}}}$ to be connected, it is clear that the procedure described can be applied to any component of a non-connected network which yields a positive payoff, and yields an SNSG-network with the same number of nodes and a grater or equal aggregate payoff.

The following proposition narrows considerably the class of dominant weighted NSG-networks by imposing the optimality conditions established in Lemma 1 to the infrastructure underlying an SNSG-network. The conditions refer to the relabeling of the nodes for which the conditions of Definition 6 hold.

**Proposition 2** *Let $\delta^{\overline{\mathbf{c}}}$ be a connected SNSG-network, with $\overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij} \in N_2}$, as specified by Definition* 6*. The following conditions are necessary for $\delta^{\overline{\mathbf{c}}}$ to be efficient:*

*(i) There may exist links not involving node 1 only if $\delta'(0) > 1/2v$, and all such links and those connecting node 1 with nodes with as many neighbors as node 1 are invested in the same amount $c^\sharp$ s.t.*

$$\delta'(c^\sharp) = 1/2v. \tag{7}$$

*(ii) If there are $p > 1$ non-trivial NSG-classes $K_1, K_2, ..., K_p$ of cardinalities $k_1, k_2, ..., k_p$ with number of neighbors $n = n_1 > n_2 > ... > n_p$, the links connecting node 1 with those in each class $K_i$ ($i \in \{2, ..., p\}$) must receive the same investment $c_i$ s.t.*

$$\delta'(c_i) = \frac{1}{2v(1 + (k_i - 1)\delta(c_i) + \sum_{r:p-i+1<r<p(r\neq i)} k_r\delta(c_r))} \tag{8}$$

*if $i > p - i + 1$, while if $i \leq p - i + 1$,*

$$\delta'(c_i) = \frac{1}{2v(1 + \sum_{r:p-i+1<r<p} k_r\delta(c_r))}. \tag{9}$$

**Proof.** $(i)$ Let $\delta^{\overline{\mathbf{c}}}$, with $\overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij}\in N_2}$, be an SNSG-network. Assume $\overline{\mathbf{c}}$ is efficient. In $\delta^{\overline{\mathbf{c}}}$ every pair of nodes different from 1, directly connected or not, is connected by a 2-link path of the star centered at 1. By condition $(i)$ in Definition 6, the links forming this star that connects 1 with the other nodes are at least as strong as the remainder links. Efficiency requires that all links are used by optimal paths and each optimal path is unique. This entails that all links connecting pairs of nodes $i, j$ different from 1 cannot be part of an optimal path connecting any other pair of nodes, otherwise some link of the star would be superfluous. Then, for all these links, as $\delta(\overline{p}_{ij}) = \delta(c_{\underline{ij}})$, $\delta(\overline{p}^{ij}_{ij}) = 1$, and (4) yields $\delta'(c_{\underline{ij}}) = 1/2v$, i.e. $c_{\underline{ij}} = c^\sharp$. As to links connecting node 1 with nodes with as many neighbors as node 1, given that, as has just been proved, all links connecting nodes different from 1 are $c^\sharp$-links, they must be optimal paths themselves. Therefore the links connecting 1 with any of them must also be $c^\sharp$-links.

$(ii)$ Now consider the links forming the star (i.e. links $c_{\underline{1j}}$, $j \neq 1$). Given that these are the strongest links in $\delta^{\overline{\mathbf{c}}}$, only the weakest among them can be of strength $\delta(c^\sharp)$ (by $(i)$, at least those connecting 1 with nodes with as many neighbors as node 1). As to those of strength greater than $\delta(c^\sharp)$, by Proposition 1, all of them must be in some optimal path connecting pairs of players. In view of the NSG structure of $\delta^{\overline{\mathbf{c}}}$ and $\overline{\mathbf{c}}$, for all $j$ with fewer neighbors than 1, i.e. $j \in K_i$ for some $i > 1$ (note that when $j \leq n_i$, otherwise $j$ is in one of the $p - i + 1$ first NSG-classes):

$$N^d(j; \delta^{\overline{\mathbf{c}}}) = \begin{cases} \{1, 2, ..., n_i\} = (K_1 \cup K_2 \cup .. \cup K_{p-i+1}), \text{ if } j > n_i, \\ \{1, 2, ..., n_i\} \setminus \{j\} = (K_1 \cup K_2 \cup .. \cup K_{p-i+1}) \setminus \{j\}, \text{ if } j \leq n_i, \end{cases} \tag{10}$$

and $j$ is indirectly connected by a 2-path via node 1 with all nodes in $K_{p-i+2} \cup ... \cup K_p$. Therefore, by (4) in Lemma 1:

$$\delta'(c_{\underline{1j}}) = \frac{1}{2v(1 + \sum_{n_i<k\leq n(k\neq j)} \delta(c_{\underline{1k}}))}. \tag{11}$$

Note that "$k \neq j$" is superfluous in (11) if $j \leq n_i$. Then, starting by nodes with a minimal number of neighbors $j \in K_p$, (11) becomes:

$$\delta'(c_{\underline{1j}}) = \frac{1}{2v(1 + \sum_{k \in K_p \setminus \{j\}} \delta(c_{\underline{1k}}) + \sum_{r:1<r<p} \sum_{k \in K_r} \delta(c_{\underline{1k}}))}. \tag{12}$$

We prove now that for all $j, j' \in K_p$, $\delta'(c_{\underline{1j}}) = \delta'(c_{\underline{1j'}})$. Assume $\delta'(c_{\underline{1j}}) < \delta'(c_{\underline{1j'}})$ or, equivalently, $c_{\underline{1j}} > c_{\underline{1j'}}$, which implies $\delta(c_{\underline{1j}}) > \delta(c_{\underline{1j'}})$, and entails

$$\sum\nolimits_{k \in K_p \setminus \{j\}} \delta(c_{\underline{1k}}) < \sum\nolimits_{k \in K_p \setminus \{j'\}} \delta(c_{\underline{1k}}).$$

But then, by (12), it must be $\delta'(c_{\underline{1j}}) > \delta'(c_{\underline{1j'}})$, which contradicts $\delta'(c_{\underline{1j}}) < \delta'(c_{\underline{1j'}})$. Therefore, $\delta'(c_{\underline{1j}}) = \delta'(c_{\underline{1j'}})$ for all $j, j' \in K_p$. Then, (12) becomes for all $j \in K_p$ :

$$\delta'(c_{\underline{1j}}) = \frac{1}{2v(1 + (k_p - 1)\delta(c_{\underline{1j}}) + \sum_{r:1<r<p} \sum_{k \in K_r} \delta(c_{\underline{1k}}))}.$$

Now, by repeating a similar argument for nodes in classes with an increasing number of neighbors, we have the same conclusion: all nodes in the same class must receive the same investment. Thus all links $\underline{1j_i}$, with $j_i \in K_i$ ($i \in \{2, ..., p\}$) must receive the same investment $c_{\underline{1j_i}}$ s.t., if $i > p - i + 1$:

$$\delta'(c_{\underline{1j}_i}) = \frac{1}{2v(1 + (k_i - 1)\delta(c_i) + \sum_{r:p-i+1<r<p(r \neq i)} k_r \delta(c_{\underline{1j_r}}))},$$

while if $i \leq p - i + 1$:

$$\delta'(c_{\underline{1j_i}}) = \frac{1}{2v(1 + \sum_{r:p-i+1<r<p} k_r \delta(c_{\underline{1j_r}}))},$$

which yield (8) and (9). ■

**Comments:** (i) Note that condition (7) is equivalent to (6), which has a direct interpretation. In an efficient network every link must be part of an optimal path connecting two nodes. If the only optimal path a link is part of is the link itself, then the investment $c$ must maximize $2v\delta(c) - c$.

(ii) Proposition 2 follows from (4) in Lemma 1. As it establishes optimality conditions for an SNSG-network, we refer to any such network satisfying them as an *optimal SNSG-network*. These conditions narrow considerably the class of SNSGs that can be optimal. First, all links not involving central node 1 and those connecting node 1 with nodes with as many neighbors as node 1 must be $c^\sharp$-links. Second, there are as many levels of investments in stronger links involving node 1 as NSG-classes are. In terms of the adjacency matrix, the main features of the pattern that the link-investment matrix of an optimal SNSG-network must follow is represented in Figure 2, where there are 5 NSG-classes with fewer neighbors than node 1, and $c^{\sharp} < c < c' < c'' < c''' < c^{iv}$.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | $c^{\sharp}$ | $c$ | $c'$ | $c''$ | $c''$ | $c'''$ | $c'''$ | $c'''$ | $c^{iv}$ | $c^{iv}$ | $c^{iv}$ |
| 2 | $c^{\sharp}$ | 0 | $c^{\sharp}$ | $c^{\sharp}$ | $c^{\sharp}$ | $c^{\sharp}$ | $c^{\sharp}$ | $c^{\sharp}$ | $c^{\sharp}$ | $c^{\sharp}$ | $c^{\sharp}$ | $c^{\sharp}$ |
| 3 | $c$ | $c^{\sharp}$ | 0 | $c^{\sharp}$ | $c^{\sharp}$ | $c^{\sharp}$ | $c^{\sharp}$ | $c^{\sharp}$ | $c^{\sharp}$ | 0 | 0 | 0 |
| 4 | $c'$ | $c^{\sharp}$ | $c^{\sharp}$ | 0 | $c^{\sharp}$ | $c^{\sharp}$ | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | $c''$ | $c^{\sharp}$ | $c^{\sharp}$ | $c^{\sharp}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | $c''$ | $c^{\sharp}$ | $c^{\sharp}$ | $c^{\sharp}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | $c'''$ | $c^{\sharp}$ | $c^{\sharp}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | $c'''$ | $c^{\sharp}$ | $c^{\sharp}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | $c'''$ | $c^{\sharp}$ | $c^{\sharp}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | $c^{iv}$ | $c^{\sharp}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | $c^{iv}$ | $c^{\sharp}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | $c^{iv}$ | $c^{\sharp}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Figure 2: Link-investment of a connected optimal SNSG-network

(iii) As argued in the introduction, NSG-networks can be seen as "imperfect" core-periphery structures. To be precise, we use the following weak and strong versions of the core-periphery notion. A network is *core-periphery* if its nodes admit a partition into two sets of nodes, $C$ (core) and $P$ (periphery), s.t. ($i$) the nodes in $C$ form a complete subnetwork; ($ii$) every node in $P$ is connected with at least one node in $C$; ($iii$) every node in $C$ is connected with at least one node in $P$, and ($iv$) no pair of nodes in $P$ is connected by a link. A *perfect core-periphery* network is a network that admits a core-periphery partition s.t. conditions ($ii$) and ($iii$) are satisfied in this strongest form: every node in $P$ is connected with *all* nodes in $C$.[21]

Actual networks in real world seldom meet even the weak version, but often exhibit a conspicuous rough core-periphery aspect. It follows from their definition that for any connected optimal SNSG-network, nodes can be partitioned into two sets $C$ and $P$, where $C$ is the class of nodes with $n-1$ neighbors and $P = N \setminus C$. However, this partition may fail to be a core-periphery partition because, in general, it may be the case that some pair of nodes in $P$ are connected by a link. Alternatively, sometimes a different partition may avoid this.[22] The following example illustrates this and helps highlighting some features of optimal SNSG-networks.

**Example 1:** Figure 3 represents the link-investment matrix of an optimal SNSG-

[21] Goyal and Joshi (2003) identify core-periphery networks with these perfect core-periphery network, in other words, with NSG-networks where nodes are divided into *two* classes according to the number of neighbors.

[22] Sometimes at the cost of breaching condition ($iii$).

network with 8 nodes.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | $c^\sharp$ | $c^\sharp$ | $c$ | $c'$ | $c'$ | $c''$ | $c''$ |
| 2 | $c^\sharp$ | 0 | $c^\sharp$ | $c^\sharp$ | $c^\sharp$ | $c^\sharp$ | $c^\sharp$ | $c^\sharp$ |
| 3 | $c^\sharp$ | $c^\sharp$ | 0 | $c^\sharp$ | $c^\sharp$ | $c^\sharp$ | $c^\sharp$ | $c^\sharp$ |
| 4 | $c$ | $c^\sharp$ | $c^\sharp$ | 0 | $c^\sharp$ | $c^\sharp$ | 0 | 0 |
| 5 | $c'$ | $c^\sharp$ | $c^\sharp$ | $c^\sharp$ | 0 | 0 | 0 | 0 |
| 6 | $c'$ | $c^\sharp$ | $c^\sharp$ | $c^\sharp$ | 0 | 0 | 0 | 0 |
| 7 | $c''$ | $c^\sharp$ | $c^\sharp$ | 0 | 0 | 0 | 0 | 0 |
| 8 | $c''$ | $c^\sharp$ | $c^\sharp$ | 0 | 0 | 0 | 0 | 0 |

Figure 3: An optimal SNSG-network's link-investment

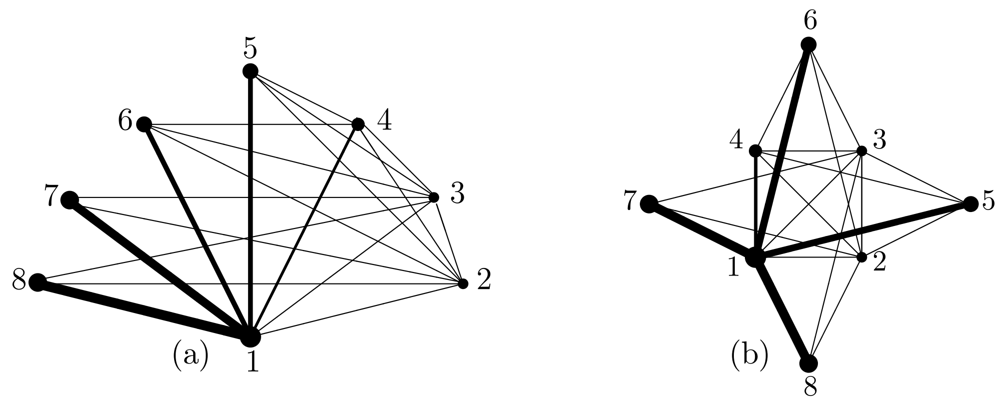


Figure 4: An 8-node optimal SNSG-network

Where by (8) and (9):

$$\delta'(c'') = \frac{1}{2v(1+\delta(c'')+2\delta(c')+\delta(c))}, \quad \delta'(c') = \frac{1}{2v(1+2\delta(c'')+\delta(c'))},$$

$$\delta'(c) = \frac{1}{2v(1+2\delta(c''))}, \quad \delta'(c^\sharp) = \frac{1}{2v}.$$

If $C = \{1, 2, 3\}$ and $P = \{4, 5, 6, 7, 8\}$, this is not a core-periphery partition because node 4 is connected with 5 and 6, and condition $(iv)$ is not satisfied. But if node 4 joins $C$, then core-periphery conditions $(i)$-$(iv)$ hold. Figure 4 displays the graph of the network in two different and equivalent pictures. Figure 4-(a) shows its SNSG nature, where the strength of a link is represented by its thickness: the thinnest ones are $c^\sharp$-links and links connecting central node 1 with nodes $5, 6, 7$ and 8 are of three different strengths. Figure 4-(b) shows its core-periphery nature (non-perfect because links $\underline{47}$ and $\underline{48}$ are missing).

A salient peculiarity of optimal SNSG-networks is the prominent role of a distinguished node, node 1, which is the node that receives the maximal amount of information.

The next proposition, whose easy proof is omitted, establishes further necessary conditions for an SNSG-network to be efficient.

**Proposition 3** *Let* $\mathbf{c}$ *be an investment profile s.t.* $\delta^{\mathbf{c}}$ *is an SNSG-network satisfying the necessary conditions for* $\delta^{\mathbf{c}}$ *to be efficient established in Proposition 2. Then the following conditions are also necessary to be efficient:*
$(i)$ *The less profitable link must be good enough to make it worth keeping it. That is,*

$$c^{\sharp} \leq 2v \min\{\delta(c^{\sharp}) - \delta(c_{\underline{1i}})\delta(c_{\underline{1j}}) : c_{\underline{ij}} \neq 0\}. \tag{13}$$

$(ii)$ *The weakest optimal indirect connection must be good enough to make its replacement by a* $c^{\sharp}$*-link not profitable. That is,*

$$c^{\sharp} \geq 2v \max\{\delta(c^{\sharp}) - \delta(c_{\underline{1i}})\delta(c_{\underline{1j}}) : c_{\underline{ij}} = 0\}. \tag{14}$$

In the seminal works of Jackson and Wolinsky (1996) and Bala and Goyal (2000) and also in different extensions and contexts, the complete network and the star emerge as efficient structures. Note that, weighted NSG-networks include as extreme cases the complete networks (all entries $> 0$) and the all-encompassing stars (non-zero entries in the first row and column, and no more non-zero entries, except those forming the main diagonal). Note also that, as is immediate to conclude, a weighted NSG is connected if and only if at least one node is connected with all the others. In the current model, from the results obtained so far, we derive necessary conditions for the efficiency of each of these structures. This will be of use to establish the existence of optimal SNSGs and characterize efficient networks.

**Proposition 4** *For a complete network to be efficient the following conditions are necessary:*
$(i)$ $\delta'(0) > 1/2v$ *and all links are invested in the same amount* $c^{\sharp} > 0$ *s.t.* $\delta'(c^{\sharp}) = 1/2v$.
$(ii)$ *The following relation holds:* $2v\delta(c^{\sharp})^2 \leq 2v\delta(c^{\sharp}) - c^{\sharp}$.

**Proof.** Assume $\overline{\mathbf{c}}$ to be efficient and complete, i.e. $c_{\underline{ij}} > 0$ for all $\underline{ij} \in N_2$. $(i)$ In view of Lemma 2 and Proposition 2, $\delta^{\overline{\mathbf{c}}}$ must be an optimal SNSG-network and all links connecting nodes different from the central node must be $c^{\sharp}$-links, s.t. $\delta'(c^{\sharp}) = 1/2v$. But as all nodes have the same number of neighbors, the same must be true for those involving the central node. Condition $\delta'(c^{\sharp}) = 1/2v$ is feasible only if $\delta'(0) > 1/2v$.

$(ii)$ Follows from Proposition 3-$(i)$. ■

We now establish necessary conditions for an all-encompassing star network to be efficient.

**Proposition 5** *For an all-encompassing star to be efficient the following conditions are necessary:*

*(i) All links are invested in the same amount $c_n^\times$ s.t.*

$$\delta'(c_n^\times) = \frac{1}{2v(1+(n-2)\delta\left(c_n^\times\right))}. \tag{15}$$

*(ii) Additionally, if $\delta'(0) > 1/2v$,*

$$2v\delta(c_n^\times)^2 \geq 2v\delta(c^\sharp) - c^\sharp, \tag{16}$$

*with $c^\sharp$ determined by (7).*

**Proof.** $(i)$ An all-encompassing star is an SNSG-network where the only links are those connecting a central node with all the others. Then only part $(ii)$ of Proposition 2 applies and yields (15).

$(ii)$ (16) follows from Proposition 3-$(ii)$, and holds trivially if $\delta'(0) \leq 1/2v$. ■

Note that, unlike $c^\sharp$, $c_n^\times$ depends on the number of nodes, hence the subindex.

The question of the feasibility of condition (15) arises. The following lemma shows that the existence of $c_n^\times$ such that (15) holds is guaranteed if $\delta'(0) > 1/2v$ whatever the number of nodes, and for $n$ big enough if $\delta'(0) \leq 1/2v$.

**Lemma 3** *Whatever the number of nodes, if $\delta'(0) > 1/2v$, then it is sure to exist $c_n^\times$ such that* (15) *holds, and also when $\delta'(0) \leq 1/2v$ for $n$ sufficiently large. On the contrary, for a fixed $n$, no such $c_n^\times$ exists if $\delta'(0) \leq \frac{1}{2v(1+(n-2)\delta(\infty))}$, where $\delta(\infty)$ denotes* $\lim_{c\to\infty} \delta(c)$.

**Proof.** Assume $\delta'(0) > 1/2v$, and let $\varphi$ be the function $\varphi(c) := \frac{1}{2v(1+(n-2)\delta(c))}$. We prove that $\varphi(c) = \delta'(c)$ holds necessarily for some $c > 0$. Note that, as $\delta(0) = 0$, $\varphi(0) = 1/2v < \delta'(0)$. On the other hand, $\varphi(c) = \frac{1}{2v(1+(n-2)\delta(c))} > \frac{1}{2v(1+(n-2)\delta(\infty))}$, for all $c > 0$. Thus, $\varphi(c)$ is a decreasing function whose value is always greater than $\frac{1}{2v(1+(n-2)\delta(\infty))}$ and it is $1/2v$ at 0, while $\delta'(c)$ is a decreasing function s.t. $\delta'(0) > 1/2v = \varphi(0)$. Moreover, by the assumptions on function $\delta$, $\lim_{c\to\infty} \delta'(c) = 0$. Therefore, as both functions are continuous, at some point necessarily $\varphi(c) = \delta'(c)$, i.e. (15) holds.

Now assume $\delta'(0) \leq 1/2v$. We prove that for $n$ sufficiently big there exists $c > 0$ s.t. $\varphi(c) < \delta'(c)$, for which it is sufficient to prove that for $n$ sufficiently big $\varphi(1) \leq \delta'(1)$. But it is easy to check that this is equivalent to

$$n - 2 \geq \frac{1 - 2v\delta'(1)}{2v\delta(1)\delta'(1)},$$

which is sure to hold for $n$ big enough.

Finally, if for a fixed $n$ inequality $\delta'(0) \leq \frac{1}{2v(1+(n-2)\delta(\infty))}$ holds, then the graphs of $\varphi(c)$ and $\delta'(c)$ do not intersect, because

$$\varphi(c) = \frac{1}{2v(1+(n-2)\delta(c))} > \frac{1}{2v(1+(n-2)\delta(\infty))} \geq \delta'(0) > \delta'(c)$$

for all $c$. Note that there is no contradiction with the preceding result: whatever the value of $\delta'(0)$, for $n$ big enough $\delta'(0) > \frac{1}{2v(1+(n-2)\delta(\infty))}$. ■

This permits to extend the existence result to *any* NSG infrastructure.[23]

**Proposition 6** *If $\delta'(0) > 1/2v$, there exists an optimal investment satisfying the conditions of Proposition* 2 *for any NSG infrastructure.*

**Proof.** If $\delta'(0) > 1/2v$, it is sure to exist $c^\sharp$ s.t. $\delta'(c^\sharp) = 1/2v$, and by Lemma 3 $c_n^\times$ s.t. (15). Let $S$ be an NSG infrastructure, with $p > 1$ non-trivial NSG-classes $K_1, K_2, ..., K_p$, of cardinalities $k_1, k_2, ..., k_p$, with number of neighbors $n = n_1 > n_2 > ... > n_p$. Then, let $\Omega = [c^\sharp, c_n^\times]^p$, and define map $\phi : \Omega \longrightarrow \Omega$,

$$(c_1, ..., c_p) \mapsto \phi(c_1, ..., c_p) = (\phi_i(c_1, ..., c_p))_{i=1,...,p}$$

as follows. First, define

$$D_i(c_1, ..., c_p) := \frac{1}{2v(1 + \sum_{r:p-i+1<r<p} k_r \delta(c_r))}$$

if $i \leq p - i + 1$ (otherwise use (8) similarly). In general, $\delta'(c_i) \neq D_i(c_1, ..., c_p)$, but $\delta'(c^\sharp) \geq \delta'(c_i) \geq \delta'(c_n^\times)$ is sure to hold. For $i = 1, ..., p$ define:

$$\phi_i(c_1, ..., c_p) := \frac{c_i + \delta'^{-1}(D_i(c_1, ..., c_p))}{2}.$$

Then $\phi : \Omega \longrightarrow \Omega$ is a continuous function and $\Omega$ is compact and convex. Brower's theorem guarantees the existence of a fixed point, or equivalently $c_1, ..., c_p$, s.t. $\delta'(c_i) = D_i(c_1, ..., c_p)$ for all $i$, i.e. (8) and (9) hold for all $i = 1, ..., p$. ■

Conditions in Proposition 4-($ii$) and Proposition 5-($ii$), necessary for the optimal complete network and the star to be efficient, are:

$$2v\delta(c^\sharp)^2 \leq 2v\delta(c^\sharp) - c^\sharp \leq 2v\delta(c_n^\times)^2.$$

The first inequality ensures that by deleting a link in an optimal complete network the aggregate payoff does not increase, while the second guarantees that connecting by a $c^\sharp$-link two nodes connected by a two $c_n^\times$-links path does not increase the aggregate payoff. Observe that these conditions determine an interval for $c^\sharp$

$$2v(\delta(c^\sharp) - \delta(c_n^\times)^2) \leq c^\sharp \leq 2v(\delta(c^\sharp) - \delta(c^\sharp)^2),$$

where *both* conditions hold. As it turns out, outside this interval the only non-empty efficient structures are either the optimal complete or the optimal all-encompassing star, while inside connected optimal SNSG-networks are the only possible efficient.

Then we have the following characterizing result.

[23] An infrastructure $S$ is a non-weighted non-directed graph, which as such can be NSG according to Definition 1.

**Proposition 7** *The only efficient structures are connected optimal SNSG-networks and the empty network. More precisely:*
*(i) If $\delta'(0) > 1/2v$ and $c^\sharp < 2v(\delta(c^\sharp) - \delta(c_n^\times)^2)$ the only efficient structure is the optimal complete network (described in Proposition 4).*
*(ii) If $\delta'(0) > 1/2v$ and $c^\sharp > 2v(\delta(c^\sharp) - \delta(c^\sharp)^2)$ the only efficient structure is the optimal all-encompassing star (described in Proposition 5).*
*(iii) If $\delta'(0) > 1/2v$ and $2v(\delta(c^\sharp) - \delta(c_n^\times)^2) \le c^\sharp \le 2v(\delta(c^\sharp) - \delta(c^\sharp)^2)$ the only efficient structures are connected optimal weighted SNSG-networks (described in Propositions 2 and 4).*
*(iv) If $\delta'(0) \le \frac{1}{2v(1+(n-2)\delta(\infty))}$ the only efficient structure is the empty network, while if $\frac{1}{2v(1+(n-2)\delta(\infty))} < \delta'(0) \le 1/2v$ the only efficient structure for $n$ big enough is the optimal all-encompassing star, where $\delta(\infty) = \lim_{c\to\infty} \delta(c)$.*

**Proof.** Let $\delta^{\overline{\mathbf{c}}}$ be any network, with $\overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij}\in N_2}$ nonempty and yielding a positive aggregate payoff. If $\delta^{\overline{\mathbf{c}}}$ is connected, in view of Lemma 2 and Propositions 2 and 3, it is dominated by an optimal SNSG-network. Otherwise, if $\delta^{\overline{\mathbf{c}}}$ is not connected, then as pointed out above, Lemma 2 and Propositions 2 and 3 can be applied to each of its components which yield a positive payoff. Let $S \subsetneq N$ be the set of nodes in the component for which the aggregate payoff of the dominant network resulting is the greatest. Then, either it contains some $c^\sharp$-link and then connecting by a $c^\sharp$-link a node in $S$ with a node in another component will increase the aggregate payoff; or it is an optimal star "$S$-encompassing" and adding spoke nodes can only increase it. Therefore, it must be connected.

Now, assuming $\delta'(0) > 1/2v$, there are the following possibilities:

$(i)$ If $c^\sharp < 2v\delta(c^\sharp) - 2v\delta(c_n^\times)^2$, i.e. $2v\delta(c_n^\times)^2 < 2v\delta(c^\sharp) - c^\sharp$, with $c_n^\times$ given by (15), then by connecting any two peripheral nodes in an all-encompassing star satisfying optimality condition (15) with a $c^\sharp$-link the aggregate payoff would increase. But then the same must occur for any two nodes not directly connected in $\delta^{\overline{\mathbf{c}}}$, because $\delta(c_n^\times) > \delta(c_{\underline{1i}})$, for all $i \neq 1$ s.t. $c_{\underline{1i}} > c^\sharp$, given that $\delta'(c_n^\times) < \delta'(c_{\underline{1i}})$. In other words, the only efficient structure is the optimal complete network .

$(ii)$ By Proposition 3-$(i)$, if any two nodes, $i$ and $j$, are connected by a $c^\sharp$-link, $c^\sharp \le 2v(\delta(c^\sharp) - \delta(c_{\underline{1i}})\delta(c_{1j})) \le 2v(\delta(c^\sharp) - \delta(c^\sharp)^2)$, which contradicts $c^\sharp > 2v(\delta(c^\sharp) - \delta(c^\sharp)^2)$. Therefore all existing links are invested in more than $c^\sharp$, which means that the $\delta^{\overline{\mathbf{c}}}$ is actually an all-encompassing star.

$(iii)$ If $2v(\delta(c^\sharp) - \delta(c_n^\times)^2) \le c^\sharp \le 2v(\delta(c^\sharp) - \delta(c^\sharp)^2)$, none of the two preceding conclusions applies, but some connected optimal SNSG-network is sure to be efficient. Which one depends on the technology.

$(iv)$ Now assume $\delta'(0) \le 1/2v$. By Lemma 3, if $\delta'(0) \le \frac{1}{2v(1+(n-2)\delta(\infty))}$ the only efficient structure is the empty network, while if $\frac{1}{2v(1+(n-2)\delta(\infty))} < \delta'(0) \le 1/2v$ the only efficient structure for $n$ big enough is the optimal all-encompassing star. ■

# 6 Scenario 2: Stability

Consider now the second scenario described in Section 4, where nodes are *players* who form links by investing in them. An *investment profile*, specified by a matrix $\mathbf{c} = (c_{ij})_{i,j\in N}$, where $c_{ij} \geq 0$ (with $c_{ii} = 0$) is the investment of player $i$ in the link connecting $i$ and $j$, determines a link-investment vector $\overline{\mathbf{c}}$

$$\mathbf{c} \rightarrow \overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij}\in N_2}, \text{ with } c_{\underline{ij}} := c_{ij} + c_{ji}.$$

The available link-formation technology, $\delta$, yields a *weighted network* for each investment profile $\mathbf{c}$. Namely, $\delta^{\mathbf{c}} := \delta^{\overline{\mathbf{c}}}$, where

$$\delta^{\mathbf{c}}_{\underline{ij}} = \delta^{\overline{\mathbf{c}}}_{\underline{ij}} = \delta(c_{\underline{ij}}) = \delta(c_{ij} + c_{ji}),$$

and payoffs are given by (2), i.e.

$$\Pi^{\delta}_i(\mathbf{c}) = \sum_{j\in N(i;\delta^{\overline{\mathbf{c}}})} I_{ij}(\delta^{\overline{\mathbf{c}}}) - \sum_{j\in N^d(i,\delta^{\mathbf{c}})} c_{ij} = v \sum_{j\in N(i;\delta^{\overline{\mathbf{c}}})} \delta^{\overline{\mathbf{c}}}(\overline{p}_{ij}) - \sum_{j\in N^d(i,\delta^{\mathbf{c}})} c_{ij}. \quad (17)$$

This scenario poses the question of stability. We address here the stability of the structures that have emerged as efficient: the empty network and the SNSG-networks, and among them the complete and the star networks. Although properly speaking one should refer to *stability of investment profiles*, we often express our results in terms of the resulting networks. Thus a Nash or pairwise "stable network" should be read as a *weighted network that results from a Nash or pairwise stable investment profile.* The following result establishes a necessary and sufficient condition for the *empty network* to be stable.

**Proposition 8** *The empty network is*
*(i) A Nash network if and only if* $\delta'(0) \leq 1/v$.
*(ii) A pairwise stable network if and only if* $\delta'(0) \leq 1/2v$.

**Proof.** $(i)$ Let $\delta^{\mathbf{0}}$ be the empty network, i.e. $c_{ij} = 0$ for all $i, j \in N$. In these conditions a player has an incentive to invest $c > 0$ in a link with another (or any number of them) only if $v\delta(c) - c > 0$. But by the assumptions on technology $\delta$, if $\delta'(0) \leq 1/v$ and $c > 0$, then $\delta(c) < c\delta'(0) \leq c/v$. Assume now that $\delta'(0) > 1/v$. Then, there exists $c > 0$ s.t. $\delta(c) > c/v$, and it is advantageous to invest $c$ in a link with another player. Therefore $(i)$ is proved.

$(ii)$ If pairwise coordination is feasible two players may form a link by jointly investing $c$ by investing $c/2$ each. If $\delta'(0) \leq 1/2v$, then $\delta(c) < c\delta'(0) \leq c/2v$. On the contrary, if $\delta'(0) > 1/2v$, then there exists $c > 0$ s.t. $\delta(c) > c/2v$, and it is advantageous to any two players to invest $c/2$ each in a link connecting them. ■

**Comments:** (i) Thus, it all depends on the technology $\delta$ and $v$, namely on the marginal fidelity-level at 0 investment: the empty network is Nash (pairwise) stable if and only if this marginal fidelity-level is equal or less than $1/v$ $(1/2v)$. Thus, the greater the value of the information at each player, the smaller this marginal fidelity-level must be for the empty network to be stable. When pairwise coordination is feasible stability is more demanding.

(ii) Therefore the empty network is Nash stable being surely inefficient (when $1/2v < \delta'(0) \leq 1/v$) or possibly inefficient (when $\delta'(0) \leq 1/2v$), and is pairwise stable only in the latter case.

The following result, similar to Lemma 1, establishes necessary conditions for an investment profile to be *stable for a given infrastructure* in the following sense.

**Definition 7** *An investment profile* $\mathbf{c} = (c_{ij})_{i,j \in N}$ *is Nash (pairwise) stable for an infrastructure* $S$ *if: (i)* $c_{\underline{ij}} > 0$ *if and only if* $\underline{ij} \in S$*, and (ii) it is Nash (pairwise) stable in the strategic network formation game that results when the strategies of players are constrained to invest only in links in* $S$*.*

**Lemma 4** *For an investment profile* $\mathbf{c} = (c_{ij})_{i,j \in N}$ *to be Nash or pairwise stable for a given infrastructure* $S \subseteq N_2$*, the following conditions are necessary: (i) If a player invests in two different links, the sets of players connected by optimal paths containing each of them are nonempty and disjoint.*

$(ii)$ For all $\underline{ij} \in S$ s.t. $c_{ij} > 0$ :

$$\delta'(c_{\underline{ij}}) = \frac{1}{v \sum_{k \in N(i;\delta^{\overline{\mathbf{c}}})\ s.t.\ \underline{ij} \in \overline{p}_{ik}} \delta(\overline{p}_{ik}^{ij})}. \tag{18}$$

$(iii)$ For all $\underline{ij} \in S$ s.t. $c_{ij} = 0$ :

$$\delta'(c_{\underline{ij}}) \leq \frac{1}{v \sum_{k \in N(i;\delta^{\overline{\mathbf{c}}})\ s.t.\ \underline{ij} \in \overline{p}_{ik}} \delta(\overline{p}_{ik}^{ij})}. \tag{19}$$

**Proof.** Let $\overline{\mathbf{c}} = (c_{\underline{ij}})_{\underline{ij} \in N_2}$ be the link-investment vector determined by investment profile $\mathbf{c}$, s.t. $c_{\underline{ij}} > 0$ if and only if $\underline{ij} \in S$. We prove first parts $(ii)$ and $(iii)$.

$(ii)$ Assume $\mathbf{c}$ to be Nash or pairwise stable for $S$, and $\underline{ij} \in S$. Then $i$ and/or $j$, at least one of them, say $i$, invests $c_{ij} > 0$. Then link $\underline{ij}$ is part of at least one optimal path in $\delta^{\overline{\mathbf{c}}}$ for $i$'s information, the one connecting $i$ and $j$, otherwise $i$ would withdraw support to it. Then, in any of the possibly different but equivalent expressions of the right-hand side of (17), $\delta(c_{\underline{ij}})$ would appear at least once, $\delta^{\overline{\mathbf{c}}}(\overline{p}_{ij}) = \delta(c_{\underline{ij}})$ if $\mathbf{c}$ is stable, and possibly also in the product yielding $\delta^{\overline{\mathbf{c}}}(\overline{p}_{ik})$ for other nodes $k$. Fix any choice of these (possibly multiple) optimal paths connecting $i$ with every other node with whom $i$ is connected and let $i$'s payoff be given by the right-hand side of (17). The right-hand side is an up to $n(n-1)$-variable function with partial derivatives. A non-null partial derivative w.r.t. $c_{ij}$ of this expression means that by slightly increasing (if it were $> 0$) or decreasing (if it were $< 0$) the investment of $i$ in link $\underline{ij}$ would increase $i$'s payoff (through the same available paths), which contradicts $\mathbf{c}$'s stability for $S$. Then the partial derivative of the right-hand side of (17) w.r.t. $c_{ij}$ must be 0, i.e., using the same notation as in Section 4,

$$\frac{\partial}{\partial c_{ij}}(v \sum_{k\in N(i;\delta^{\overline{\mathbf{c}}})} \delta^{\overline{\mathbf{c}}}(\overline{p}_{ik}) - \sum_{k\in N^d(i,\delta^{\mathbf{c}})} c_{ik}) = v\delta'(c_{\underline{ij}}) \sum_{k\in N(i;\delta^{\overline{\mathbf{c}}}) \ s.t. \ \underline{ij}\in\overline{p}_{ik}} \delta(\overline{p}_{ik}^{ij}) - 1 = 0,$$

which yields (18).[24]

($iii$) Assume now that $c_{\underline{ij}} > 0$ and $c_{ij} = 0$. A similar argument to the one used to prove part ($ii$) leads in this case to the conclusion that

$$v\delta'(c_{\underline{ij}}) \sum_{k\in N(i;\delta^{\overline{\mathbf{c}}}) \ s.t. \ \underline{ij}\in\overline{p}_{ik}} \delta(\overline{p}_{ik}^{ij}) - 1 \leq 0,$$

otherwise player $i$ would have an incentive to invest in link $\underline{ij}$, which yields (19).

($i$) Assume that player $i$ invests in links with two nodes $j$ and $k$, $c_{ij} > 0$ and $c_{ik} > 0$, and for some player $l$ there are two different optimal paths $\overline{p}_{il}$ and $\overline{p}'_{il}$ such that $\underline{ij} \in \overline{p}_{il}$ and $\underline{ik} \in \overline{p}'_{il}$. Then the right-hand side of (18) admits at least *two* different expressions where the optimal path connecting $i$ and any other player but $l$ is the same, but one uses $\overline{p}_{il}$ and the other uses $\overline{p}'_{il}$. In that case, (18) yields two different values for $\delta'(c_{\underline{ij}})$, which is a contradiction. ■

**Comment:** Part ($i$) establishes that what any player "sees" through different links in which he/she invests in a stable profile do not overlap: if $i$ sees $l$ through an optimal path that contains $\underline{ij}$, it cannot be the case that $i$ sees $l$ through another optimal path that contains $\underline{ik} \neq \underline{ij}$. Note the similarity and the *difference* with part ($i$) in Lemma 1. As it occurs in Lemma 1, ($i$) is a consequence of ($ii$), which here is the result of requiring the *marginal benefit of the investment of any player in each of his/her links to be zero*. The resulting condition (18) has also a clear interpretation. If player $i$ invests in a link with $j$, the denominator of the fraction in formula (18) that yields $\delta'(c_{\underline{ij}})$ is $v$ times the sum of the fidelity-levels through all optimal paths containing link $\underline{ij}$ (discounting that of link $\underline{ij}$) through which player $i$ receives information. In other words, the actual *amount of information that reaches $j$ on its optimal way to $i$*. Thus this sum is a measure of the importance of link $\underline{ij}$ *to player* $i$: the greater this amount, the smaller $\delta'(c_{\underline{ij}})$, i.e. the greater $c_{ij}$ and $\delta(c_{\underline{ij}})$. Note here the similarity and difference with the meaning of similar expression (4) for an efficient link-investment vector: there, it was the overall importance of the link for the flow of information, while here it is the importance for a player who invests in it.

As with efficiency, we have immediate consequences for stability in general:

[24] Just note that by the chain rule

$$\frac{\partial}{\partial c_{ij}}(\delta(c_{ij} + c_{ji})) = \delta'(c_{ij} + c_{ji}) \cdot 1 = \delta'(c_{\underline{ij}}).$$

**Proposition 9** *For an investment profile* $\mathbf{c} = (c_{ij})_{i,j \in N}$ *to be Nash or pairwise stable conditions* $(i)$-$(iii)$ *are necessary:* $(i)$ *If a player invests in two different links, the sets of players connected by optimal paths that contain each of them are nonempty and disjoint.*
$(ii)$ *For each* $\underline{ij} \in N_2$ *s.t.* $c_{ij} > 0$, *condition* (18) *holds.*
$(iii)$ *For each* $\underline{ij} \in N_2$ *s.t.* $c_{\underline{ij}} > 0$ *and* $c_{ij} = 0$, *condition* (19) *holds.*
$(iv)$ *To be Nash stable, it is also necessary that for each* $\underline{ij} \in N_2$ *s.t.* $c_{\underline{ij}} = 0$, *if* $\delta'(0) > 1/v$,

$$v\delta^{\overline{\mathbf{c}}}(\overline{p}_{ij}) \geq v\delta(c^{\#}) - c^{\#}, \tag{20}$$

*where* $c^{\#} = \arg\max_{c>0}(v\delta(c) - c)$.
$(v)$ *To be pairwise stable, condition* (20) *must be replaced by*

$$2v\delta^{\overline{\mathbf{c}}}(\overline{p}_{ij}) \geq 2v\delta(c^{\#}) - c^{\#}. \tag{21}$$

**Proof.** Assume $\mathbf{c} = (c_{ij})_{i,j \in N}$ to be Nash (pairwise) stable, then $\mathbf{c}$ must be Nash (pairwise) stable for the infrastructure $S = \{\underline{ij} \in N_2 : c_{\underline{ij}} > 0\}$. Then $(i)$, $(ii)$ and $(iii)$ follow immediately from Lemma 4.

$(iv)$ Assume $c_{\underline{ij}} = 0$. Then no investment in link $\underline{ij}$ from $i$ can increase $i$'s payoff, that is, for all $c > 0$, $v\delta^{\overline{\mathbf{c}}}(\overline{p}_{ij}) \geq v\delta(c) - c$, otherwise investing $c$ in link $\overline{ij}$ would surely increase $i$'s payoff. This yields condition (20). Note that if $\delta'(0) \leq 1/v$, then $v\delta(c) - c < 0$ for all $c > 0$.

$(v)$ If pairwise coordination is feasible and $2v\delta^{\overline{\mathbf{c}}}(\overline{p}_{ij}) < 2v\delta(c^{\#}) - c^{\#}$, players $i$ and $j$ have an incentive to invest $c^{\#}/2$ each on link $\underline{ij}$. ■

Parts $(i)$-$(iii)$ follow directly from Lemma 4, while parts $(iv)$ and $(v)$ refer to links not invested in and impose necessary conditions for Nash and pairwise stability, which amount to lack of incentives to invest in them when coordination is or not unfeasible.

A direct consequence of Proposition 9 is the following.

**Corollary 1** *If two players are connected by a link in the network resulting from a Nash or pairwise stable investment profile and do not benefit equally from the link, the investment in that link is made entirely by the player who benefits the most from the existence of the link.*

**Proof.** Let $\mathbf{c}$ be a Nash or pairwise stable investment profile and assume $c_{\underline{ij}} > 0$. Then if both invest in link $\underline{ij}$, condition (18) must hold for $i$ and $j$, which are compatible only if the denominator in the right-hand side of equation (18) are equal for $i$ and $j$. In other words, only if both players benefit equally from the existence of the link. If they benefit differently from it, both conditions are incompatible, and stability is possible only if the player who benefits the most covers entirely the investment. In this way both conditions (18) and (19) hold. ■

As it has been established in the preceding section, the only efficient networks are optimal connected SNSG-networks. This raises the question about the stability conditions for SNSG-networks or investment profiles for NSG infrastructures. The following lemma provides a shortcut to answer this question.[25] As has been commented, necessary conditions (4) for efficiency and those for stability (18) are similar but differ. Nevertheless, when the maximal distance between nodes is 2 condition (4) imposes for efficiency exactly the same that condition (18) imposes for stability *when parameter $v$ doubles its value*. So far, parameter $v$ has been excluded form the notation for unnecessary. Now we need to make it explicit and include it and write $\delta_v^{\mathbf{c}}$ or $\delta_v^{\overline{\mathbf{c}}}$ instead of $\delta^{\mathbf{c}}$ or $\delta^{\overline{\mathbf{c}}}$.

**Lemma 5** *Let* $\mathbf{c}$ *be an investment profile in an infrastructure* $S$ *where the maximal distance between any two nodes is* 2*; then,* $\delta_v^{\overline{\mathbf{c}}}$ *satisfies condition* (4) *of Lemma* 1 *for efficiency if and only if* $\delta_{2v}^{\mathbf{c}}$ *satisfies condition* (18) *of Lemma* 4 *for stability.*

**Proof.** Let $S$ be an infrastructure where the maximal distance between any two nodes is 2. Then if two nodes, $i, j$ are directly connected by a link, this link is the only optimal path it is part of. Thus, (4) yields $\delta'(c_{\underline{ij}}) = 1/2v$, while (18) yields $\delta'(c_{ij}) = 1/v$, so that both coincide if in the latter context the individual value *doubles* that in the first. Now consider a node $i$ at distance 2 of some node and any link $\underline{ij}$ part of an optimal path of length 2, then (4) yields

$$\delta'(c_{\underline{ij}}) = \frac{1}{2v\sum_{\underline{kl}\in N_2 \ s.t. \ \underline{ij}\in\overline{p}_{kl}} \delta(\overline{p}_{kl}^{ij})} = \frac{1}{2v(1+\sum_{l\in N^d(j,\delta_v^{\overline{\mathbf{c}}})\setminus\{i\}} \delta(c_{\underline{jl}}))};$$

while (18) *when parameter $v$ doubles* its value yields

$$\delta'(c_{\underline{ij}}) = \frac{1}{2v\sum_{k\in N(i;\delta^{\overline{\mathbf{c}}}) \ s.t. \ \underline{ij}\in\overline{p}_{ik}} \delta(\overline{p}_{ik}^{ij})} = \frac{1}{2v(1+\sum_{l\in N^d(j,\delta_v^{\overline{\mathbf{c}}})\setminus\{i\}} \delta(c_{\underline{jl}}))}.$$

■

As is well known, one of the properties of connected NSG networks is their "small world" character: no two connected nodes are at a distance greater than 2. Therefore, Lemma 5 can be applied to connected NSG-networks. Thus, from Proposition 2 and Lemma 5, along with Corollary 1, we have:

**Proposition 10** *Let* $\mathbf{c}$ *be an investment profile s.t.* $\delta^{\mathbf{c}}$ *is a connected SNSG-network. The following conditions are necessary for* $\delta^{\overline{\mathbf{c}}}$ *to be Nash or pairwise stable:*
(*i*) *There may exist links not involving node* 1 *only if* $\delta'(0) > 1/v$*, and all such links, and those connecting node* 1 *with nodes with as many neighbors as node* 1 *are invested in the same amount* $c^{\#}$ *s.t.*

$$\delta'(c^{\#}) = 1/v. \tag{22}$$

[25] The results about stability that follow, stemming from Lemma 5, can also be derived directly from Lemma 4.

*(ii) If there are $p > 1$ non-trivial NSG-classes $K_1, K_2, ..., K_p$ of cardinalities $k_1, k_2, ..., k_p$ with number of neighbors $n = n_1 > n_2 > ... > n_p$, the links connecting node 1 with those in each class $K_i$ ($i \in \{2, ..., p\}$) must receive the same investment $c_i$ s.t.*

$$\delta'(c_i) = \frac{1}{v(1 + (k_i - 1)\delta(c_i) + \sum_{r:p-i+1<r<p(r \neq i)} k_r \delta(c_r))}, \tag{23}$$

*if $i > p - i + 1$, while if $i \leq p - i + 1$,*

$$\delta'(c_i) = \frac{1}{v(1 + \sum_{r:p-i+1<r<p} k_r \delta(c_r))}. \tag{24}$$

*and must be entirely covered by the player in $K_i$.*

The following proposition, whose easy proof is omitted, establishes further necessary conditions for stability.

**Proposition 11** *Let $\mathbf{c}$ be an investment profile s.t. $\delta^{\mathbf{c}}$ is an SNSG-network satisfying the necessary conditions for $\delta^{\mathbf{c}}$ to be stable established in Proposition 10. Then the following conditions are also necessary to be Nash or pairwise stable:*
*(i) To be stable in either sense the less profitable link must be good enough to make it worth keeping it. That is, for all $i, j$ s.t. $c_{\underline{ij}} = c^{\#}$:*

$$c_{ij} \leq v(\delta(c^{\#}) - \delta(c_{\underline{1i}})\delta(c_{\underline{1j}})), \tag{25}$$

*which entails*

$$c^{\#} \leq 2v \min\{\delta(c^{\#}) - \delta(c_{\underline{1i}})\delta(c_{\underline{1j}}) : c_{\underline{ij}} = c^{\#}\}. \tag{26}$$

*(ii) To be Nash stable, the weakest optimal indirect connection must be good enough to make its replacement by a $c^{\#}$-link not profitable. That is,*

$$c^{\#} \geq v \max\{\delta(c^{\#}) - \delta(c_{\underline{1i}})\delta(c_{\underline{1j}}) : c_{\underline{ij}} = 0\}. \tag{27}$$

*(iii) To be pairwise stable, condition (27) must be replaced by*

$$c^{\#}/2 \geq v \max\{\delta(c^{\#}) - \delta(c_{\underline{1i}})\delta(c_{\underline{1j}}) : c_{\underline{ij}} = 0\}. \tag{28}$$

Now we turn our attention to complete and all-encompassing star networks. The latter are SNSG-networks, but the following result shows that the only stable complete network is also an SNSG-network.

**Proposition 12** *Let $\mathbf{c}$ be an investment profile such that $\delta^{\mathbf{c}}$ is complete, then $\delta^{\mathbf{c}}$ is a Nash/pairwise stable network if and only if the following conditions hold:*
*(i) $\delta'(0) > 1/v$ and all links receive the same joint investment $c^{\#} > 0$, such that $\delta'(c^{\#}) = 1/v$.*
*(ii) $c^{\#} \leq 2v(\delta(c^{\#}) - \delta(c^{\#})^2)$.*
*(iii) For all $i, j$ ($i \neq j$) : $c_{ij} \leq v(\delta(c^{\#}) - \delta(c^{\#})^2)$.*

**Proof.** (*Necessity*) Assume $\delta^{\mathbf{c}}$ is complete, i.e. $c_{ij} + c_{ji} > 0$ for all $i, j \in N$ $(i \neq j)$, and stable. Then from Proposition 9-$(i)$-$(ii)$ follows that each link itself is the only optimal path it belongs to and $\delta'(c_{ij} + c_{ij}) = 1/v$. Finally, for $\delta^{\mathbf{c}}$ to be pairwise or Nash stable no player must have an incentive to withdraw support to a link, i.e. for all $i, j$ $(i \neq j) : \delta(c^{\#})v - c_{ij} \geq \delta(c^{\#})^2 v$, which yields $(iii)$, which notice is compatible with $c_{ij} + c_{ji} = c^{\#}$ if and only if $(ii)$ holds.

(*Sufficiency*) If these conditions hold no player has an incentive to change his/her investments. And, as all pairs of players are directly connected, pairwise coordination does not give new options. ■

**Comments:** (i) Note that the only complete network that can satisfy these conditions is the extreme case of SNSG-network satisfying conditions established in Proposition 10, and that in that case it is *both* Nash and pairwise stable.

(ii) The first two conditions concern $\delta$, $c^{\#}$ and $v$, while only $(iii)$ concerns directly the way in which the cost of the link is shared by setting an upper bound to the investment of each player. That is, if conditions $(i)$-$(ii)$ hold, complete networks supported by any investment profile satisfying $(iii)$ are Nash stable. As $\delta'(0) > 1/v$ guarantees the existence of a unique $c^{\#}$ s.t. $\delta'(c^{\#}) = 1/v$, the stability of complete networks hinges upon condition $(iii)$.

As any star is an extreme case of SNSG-network, Lemma 5 can be applied and used to obtain the following characterization.

**Proposition 13** *Let* $\mathbf{c}$ *be an investment profile such that* $\delta^{\mathbf{c}}$ *is an all-encompassing star, then* $\delta^{\mathbf{c}}$ *is a Nash network if and only if* $(i)$ *and* $(ii)$ *hold:*
$(i)$ $\delta^{\mathbf{c}}$ *is a periphery-sponsored star where all peripheral players invest the same amount* $c_n^*$ *in the only link in which each of them is involved, such that*:

$$\delta'(c_n^*) = \frac{1}{v(1 + (n-2)\delta(c_n^*))}. \tag{29}$$

$(ii)$ *Additionally, if* $\delta'(0) > 1/v$,

$$v\delta(c_n^*)^2 \geq v\delta(c^{\#}) - c^{\#}. \tag{30}$$

$(iii)$ $\delta^{\mathbf{c}}$ *is a pairwise stable network if and only if* (29) *and*

$$2v\delta(c_n^*)^2 \geq \max_{c>0}(2v\delta(c) - c) = 2v\delta(c^{\#}) - c^{\#}. \tag{31}$$

**Proof.** (*Necessity*) Part $(i)$ follows directly from Proposition 10: (23) becomes equation (29), while (30) and (31) follow from (27) and (28) in Proposition 11.
(*Sufficiency*) If conditions $(i)$ and $(ii)$ hold, peripheral players's investments are optimal and no spoke node has an incentive to invest unilaterally in a link with another spoke node. Remains to be checked that it is worth for any peripheral player to invest $c_n^*$, i.e. that

$$c_n^* \leq \delta(c_n^*)(1 + (n-2)\delta(c_n^*))v.$$

But if (29) holds, this is equivalent to check that $c_n^* \leq \delta(c_n^*)/\delta'(c_n^*)$ or, equivalently that

$$\delta'(c_n^*) \leq \delta(c_n^*)/c_n^*$$

which follows from the smoothness and concavity of $\delta$. As to the center, as

$$(\delta(c_n^* + c) - \delta(c_n^*))/c < \delta'(c_n^*) < \delta'(c^\#) = 1/v,$$

for all $c > 0$, then $v\delta(c_n^* + c) - c \leq v\delta(c_n^*)$ and consequently the center has no incentive to invest in a link (or any number of them). Similarly, conditions $(i)$ and $(iii)$ preclude any improvement of spoke nodes and the center even if pairwise coordination is possible.

■

The following lemma, entirely similar to Lemma 3 and whose proof is omitted for this reason, gives a sufficient condition for (29) to hold for some $c_n^*$.

**Lemma 6** *Whatever the number of nodes, if $\delta'(0) > 1/v$, then it is sure to exist $c_n^*$ such that* (29) *holds, and also when $\delta'(0) \leq 1/v$ for $n$ sufficiently large. On the contrary, for a fixed $n$, no such $c_n^*$ exists if $\delta'(0) \leq \frac{1}{v(1+(n-2)\delta(\infty))}$, where $\delta(\infty)$ denotes* $\lim_{c\to\infty} \delta(c)$.

Now, the existence result of Proposition 6 can be easily translated to stability as follows.

**Proposition 14** *If $\delta'(0) > 1/v$, there exists an investment profile satisfying the conditions of Proposition* 10 *for any NSG infrastructure.*

The following example shows that there actually exist SNSG-networks pairwise stable different from the complete and the star.

**Example 2:** Assume $n = 12, v = 2$. Let $\mathbf{c}$ be an investment profile yielding an optimal SNSG-network where the central player 1 is directly connected with all the others but player 2 by $c_{\underline{1i}}$-links ($i \in K = \{3, ...12\}$), and player 2 is linked with all other players by $c^\#$-links. That is, in terms of Proposition 10 there is only one class $K$ (those nodes with 2 neighbors) apart from the NSG-class with 11 neighbors, formed by nodes 1 and 2. Assume $c^\# = 0.2$; $\delta(c^\#) = 0.75$, $\delta'(c^\#) = 0.5$, $c_{\underline{1i}} = 1$, and $\delta(c_{\underline{1i}}) = \delta(c_{11}^*) = 0.85$ and consequently (Proposition 10-$(ii)$)

$$\delta'(c_{\underline{1i}}) = \delta'(c_{11}^*) = \frac{1}{v(1 + (k-1)\delta(c_{11}^*))} = \frac{1}{2(1 + 9(0.85))} = 0.0578.$$

Further, players' investments are as follows: every player $i \in K$ fully covers the investments in its links with players 1 and 2, i.e. $c_{i2} = c^\#$ and $c_{i1} = c_{\underline{i1}} = 1$; and 1 and 2, share in any way the joint investment $c^\#$ in their link.
Conditions (28) and (26) of Proposition 11 hold, i.e.

$$2v(\delta(c^\#) - \delta(c_{\underline{i1}})^2) \leq c^\# \leq 2v(\delta(c^\#) - \delta(c^\#)\delta(c_{\underline{i1}})),$$

$$4(0.75-0.85^2)=0.11\leq 0.2\leq 0.45=4(0.75-0.75(0.85)).$$

Any player $i\in K$ is willing to pay for a $c^{\#}$-link with player 2, given that $c^{\#}$ is the optimal investment according to Proposition 10, and severing it would mean a loss:

$$c^{\#}=0.2\leq v(\delta(c^{\#})-\delta(c^{\#})\delta(c_{\underline{ip^*}}))=2(0.75-0.75(0.85))=0.225.$$

All links connecting player 1 with players in $K$ receive the investments necessary for equilibrium (Proposition 10-($ii$)), and no $i\in K$ has an incentive to sever his/her link with player 1 he/she is paying for because $\pi_i(\mathbf{c})\geq\pi_i(\mathbf{c}-\underline{i1})$ :

$$\begin{aligned}
\pi_i(\mathbf{c}) &= v(1+(k-1)\delta(c_{\underline{i1}}))\delta(c_{\underline{i1}})+v\delta(c^{\#})-c_{\underline{i1}}-c^{\#}\\
&= 2(1+9\,(0.85))0.85+2\,(0.75)-1-0.2=15.005\\
\pi_i(\mathbf{c}-\underline{i1}) &= v\left(1+p\delta(c^{\#})\right)\delta(c^{\#})-c^{\#}\\
&= 2(1+10\,(0.75)\,)0.75-0.2=12.55.
\end{aligned}$$

Whatever player 2 is paying for his/her link with player 1, there is no incentive for him/her to invest optimally in a link with player 1 so as to see everything through the resulting star, i.e. $c'_{\underline{21}}$ such that

$$\delta'(c'_{\underline{21}})=\frac{1}{2(1+10\,(0.85))}=0.05263,$$

if

$$(1+10\,(0.85))\delta(c'_{\underline{21}})-c'_{\underline{21}}\leq 11\,(0.75)-c^{\#}_{21}.$$

Now if $c'_{\underline{21}}=1.185$ and $\delta(c'_{\underline{21}})=0.86$, this expression yields:

$$9.5\,(0.86)-1.185\leq 8.25-c^{\#}_{21}\Leftrightarrow c^{\#}_{21}\leq 1.265,$$

which surely holds whatever player 2's share $c^{\#}_{21}$ of $c^{\#}=0.2$.
Investments $c^{\#}_{21}$ and $c^{\#}_{12}$ must also be s.t. $c^{\#}_{21}+c^{\#}_{12}=c^{\#}$ and by Proposition 11-($i$)

$$\max\{c^{\#}_{21},c^{\#}_{12}\}\leq v\delta(c^{\#})-\delta(c^{\#})\delta(c_{\underline{i1}})v=2(0.75)-0.75(0.85)2=0.225,$$

which surely holds whatever players 1 and 2's share of $c^{\#}=0.2$.
Finally, let us see that the slope of $\delta(c)$ at $c^{\#}$, $c_{\underline{i1}}$ and $c'_{\underline{21}}$ are consistent and compatible with the assumptions on the technology:[26]

[26] Given the strict concavity of $\delta$, for all $a,b$ $(0<a<b)$, conditions

$$\delta'(a)>\frac{\delta(b)-\delta(a)}{b-a}>\delta'(b)$$

must hold.

$$\begin{aligned}
\delta'(c^{\#}) &= 0.5 > \frac{\delta(c_{\underline{i1}}) - \delta(c^{\#})}{c_{\underline{i1}} - c^{\#}} = 0.125 > \delta'(c_{\underline{i1}}) = 0.0578, \\
\delta'(c_{\underline{i1}}) &= 0.0578 > \frac{\delta(c'_{\underline{21}}) - \delta(c_{\underline{i1}})}{c'_{\underline{21}} - c_{\underline{i1}}} = 0.05405\,4 > \delta'(c'_{\underline{21}}) = 0.05263, \\
\delta'(c'_{\underline{21}}) &= 0.05263 < 0.111\,68 = \frac{\delta(c_{\underline{21}}) - \delta(c^{\#})}{c_{\underline{21}} - c^{\#}}.
\end{aligned}$$

As to the payoffs we have:

$$\begin{aligned}
\pi_i(\mathbf{c}) &= v\delta(c_{\underline{i1}})(1 + 9\delta(c_{\underline{i1}})) + v\delta(c^{\#}) - c_{\underline{i1}} - c^{\#} = 15.005, \\
\pi_2(\mathbf{c}) &= 11v\delta(c^{\#}) - c_{21} = 16.5 - c_{21}, \\
\pi_1(\mathbf{c}) &= 10v\delta(c_{\underline{i1}}) + v\delta(c^{\#}) - c_{12} = 18.5 - c_{12},
\end{aligned}$$

where $c_{12} + c_{21} = c^{\#} = 0.2$, therefore, whatever the shares of players 1 and 2 of the cost of their link, we have

$$\pi_1(\mathbf{c}) > \pi_2(\mathbf{c}) > \pi_i(\mathbf{c}).$$

Thus player 1 receives the greatest payoff, followed by the other node in the core and those in the periphery. It is worth remarking that by adding any number of players connected with 1 and 2 and among themselves by $c^{\#}$-links the network remains pairwise stable.

# 7 Concluding remarks

We have developed a simple "marginalist" network formation model which is a natural extension of the seminal discrete models of Jackson and Wolinsky (1996) and Bala and Goyal (2000a). The basic logic is the same, *payoff* = *information* − *investment*, but based on a non-discrete decreasing returns link-formation technology, which is the only exogenous ingredient in the model.

The characterization of efficient networks is basically supported by crucial Lemma 2, proving constructively the dominance of *strongly* nested split graph networks, and Lemma 1 giving optimality conditions for investments in an infrastructure which allow to refine the conclusion of Lemma 2, and lead to Proposition 2, and finally to Proposition 7, establishing that efficient structures must be connected optimal SNSG-networks or empty. Similarly, Lemma 4, giving necessary conditions for Nash and pairwise stability for the game associated for an infrastructure, is crucial for the results relative to stability. Necessary and sufficient conditions for complete networks and stars to be stable in either sense, and necessary conditions for SNSG networks to be stable, which have been shown to be pairwise stable for certain technologies.

A comparison of the results obtained for efficiency and those for stability show a great parallelism. This is particularly so for the necessary conditions for optimality (Lemma 1) and for stability (Lemma 4), which entail further parallelisms. In this respect, the most remarkable is the fact that both imposing efficiency to an link-investment that yields an SNSG network (Proposition 2) and imposing stability (Proposition 10) lead to an entirely similar refinement of SNSG structures.[27] However, these conditions are *different*. In fact, from Lemmas 1 and 4, it follows that efficiency and stability are *incompatible* except in the case of the empty network. The reason is clear, a nonempty efficient network requires link-investments which are not stable because they give players the opportunity of free riding by taking advantage of the externalities. Both similarity and difference stem from the same source. Conditions for optimality and stability are based on the same economic principle: imposing zero marginal benefit, social benefit for efficiency, individual benefit for stability.

As already emphasized in the introduction, a remarkable result is the emergence of these weighted nested split graph structures from a simple model based on a single exogenous ingredient.

Apart from further exploring the model, there are a number of extensions worth investigating. These are some of them:

(i) Exploring the impact of assuming heterogeneity, technological and/or in individual values.

(ii) Exploring some variants of the technology function, as the following ones:

- Assuming $\delta(c) > 0$ only for $c > \overline{c}$, i.e. setting a "threshold" or minimal joint investment for a link to admit flow (translating the assumptions about $\delta$ to a map $\delta : [\overline{c}, \infty) \to [0, 1)$, with $\delta(c) = 0$ for $0 \leq c \leq \overline{c}$).

- Assuming $\delta$ convex up to an inflection point, then concave. This is intuitively appealing and would yield as limiting cases Jackson and Wolinsky's (1996) connections model and Bala and Goyal's (2000a) two-way flow model.

- Assuming $\delta$ a strictly concave continuously differentiable map $\delta : \mathbb{R}_+ \times \mathbb{R}_+ \to [0, 1)$ increasing in both arguments, i.e. both players' investments.

(iii) Enriching the model, the basic one or any of its extensions, introducing dynamics.

# References


[1] Aumann, R. and R. Myerson, 1988, Endogenous formation of links between players and coalitions: An application of the Shapley value. In *The Shapley Value*, Ed. A. Roth, Cambridge University Press, 175-191.

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


[27] Originally, these conditions were established independently. Only later, Lemma 5, used now to derive those about stability from those in optimality, gave the clue.

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

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

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

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

[7] Belleflamme, P., and F. Bloch, 2004, Market Sharing Agreements and Collusive Networks, *International Economic Review* 45, 387-411.

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

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

[10] Bramoullé, Y., B.W. Rogers and A. Galeotti, Eds., 2015, *The Oxford Handbook on the Economics of Networks*, Oxford University Press, Oxford.

[11] Bramoullé, Y. and R. Kranton, 2007, Public goods in networks, *Journal of Economic Theory* 135, 478-494.

[12] Csermely, P., A. London, L-Y. Wu and B. Uzzi, 2013, Structure and dynamics of core/periphery networks, *Journal of Complex Networks* 1, 93-123.

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

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

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

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

[17] Hojman, D.A., and A. Szeidl, 2008, Core and periphery in networks, *Journal of Economic Theory* 139, 295-309.

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

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

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

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

[22] Lagerås, A.N., and D. Seim, 2015, Strategic complementarities, network games and endogenous network formation, *International Journal of Game Theory* (forthcoming).

[23] Myerson, R., 1977, Graphs and cooperation in games, *Mathematics of Operations Research*, 2, 225-229.

[24] Olaizola, N., and F. Valenciano, 2015a, A unifying model of strategic network formation. *Ikerlanak Working Paper Series*: IL 85/15, Department of Foundations of Economic Analysis I, University os the Basque Country.

[25] Olaizola, N., and F. Valenciano, 2015b, The impact of liberalizing cost-sharing on basic models of network formation. *Ikerlanak Working Paper Series*: IL 91/15, Department of Foundations of Economic Analysis I, University os the Basque Country.

[26] Rogers, B. R., 2006, Learning and status in social networks, Ph.D. Thesis, CIT Pasadena, California.

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