Communication Networks with Endogenous Link Strength

Francis Bloch and Bhaskar Dutta

No 723

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

THE UNIVERSITY OF WARWICK

Communication Networks with Endogeneous Link Strength*

Francis Bloch, Bhaskar Dutta †

April 1, 2005

Abstract

This paper analyzes the formation of communication networks when players choose endogenously their investment on communication links. We consider two alternative definitions of network reliability; product reliability, where the decay of information depends on the product of the strength of communication links, and min reliability where the speed of connection is affected by the weakest communication link. When investments are separable, the architecture of the efficient network depends crucially on the shape of the transformation function linking investments to the quality of communication links. With increasing marginal returns to investment, the efficient network is a star ; with decreasing marginal returns, the conflict between maximization of direct and indirect benefits prevents a complete characterization of efficient networks. However, with min reliability, the efficient network must be a tree. Furthermore, in the particular case of linear transformation functions, in an efficient network, all links must have equal strength. When investments are perfect complements, the results change drastically: under product reliability, the efficient network must contain a cycle, and is in fact a circle for small societies. With min reliability, the efficient network is either a circle or a line.

*We thank C.Blackorby and M. Jackson for helpful discussions. We have benefited from the comments of participants in seminars and conferences at Guanajuato...

†Bloch is at GREQAM, Universite d'Aix-Marseille, 2 rue de Charite, 13002 Marseille, France. Dutta is in the Department of Economics, University of Warwick, Coventry CV4 7AL, England.

As in classical models of network formation, efficient networks may not be supported by private invesment decisions. We provide examples to show that the star may not be stable when the transformation functions is strictly convex. We also note that with perfect substitutes and perfect complements (when the efficient network displays a very symmetric structure), the efficient network can indeed be supported by private investments when the society is large.

JEL Classification Numbers: D85, C70

Keywords: communication networks, network reliability

1 Introduction

Networks play a fundamental role in the diffusion of information. Both formal communication networks (like the telephone and the internet) and informal social networks are extensively used to transmit information about job opportunities, technological innovations, new ideas, social and cultural events.[1] In recent years, economists have studied how networks are formed by self-interested agents, paying particular attention to communication networks (Bala and Goyal (2000a) and Jackson and Wolinsky (1996)'s connections model). For the most part, this literature assumes that all communication links have the same quality. However, the strength of links in actual communication networks varies widely. In formal communication networks, the reliability of a communication link depends on the physical characteristics of the connection, which varies across the network. In informal social networks, the strength of a social link depends on the frequency and the length of social interactions which also display a wide variation across a given network. In this paper, our objective is to extend the analysis of the formation of formal and informal communication networks by allowing agents to choose endogenously the strength of communication links.[2]

In communication networks, agents derive a positive utility from the agents to whom they are connected.[3] When communication links have a fixed value, agents incur a fixed cost per link, and the formation of communication networks results from the trade-off between connection benefits and the fixed cost of communication links. Bala and Goyal (2000a) and Jackson and Wolinsky (1996) characterize stable and efficient networks as a function of the fixed cost. Interestingly, social and private incentives are not necessarily aligned, and there exist values of the fixed cost for which stable and efficient networks differ.

In this paper, we take a very different perspective on the cost of link

[1] The use of social networks in job referrals has been studied,a mong others, by Granovetter (1974), Boorman (1975), Montgomery (1991), Calvo-Armengol and Jackson (1994). A classic reference on the role of social networks in the diffusion of innovation is Coleman, Katz and Menzel (1966).

[2] See Goyal(2004), which emphasizes the importance of studying networks where the srength of links can be chosen endogenously.

[3] For most of their analysis, Bala and Goyal (2000a) suppose that every connection (direct or indirect) provides the same benefits. Jackson and Wolinsky (1996) assume that the benefit of a connection is inversely related to the shortest distance between agents in the network.

formation. We suppose that agents are endowed with a fixed endowment X that they allocate across different connections. In formal communication networks, X should be interpreted as a total budget that the agent can invest on communication links ; in informal social networks, X represents the fixed amount of time that an agent can spend on different social links. The strength of a communication link is an increasing function of the investments of the two parties, and is normalized to lie in the interval $[0, 1]$. In this setting, agents face the following choice: they can either spend their endowment on a small number of strong links, or on a large number of weak links.

In most of the analysis, we assume that investments are *separable*, so that every agent can unilaterally contribute to the formation of the communication link.[4] A special case of this formulation is the *perfect substitutes case* where the quality of a link is the sum of individual investments. In general, we distinguish between two situations: either the quality of a link is a convex transformation or a concave transformation of the investment. Convex transformations correspond to situations where the marginal returns to investments in physical communication links or social links are increasing ; concave transformations to the case where marginal returns are decreasing. In our view, there is no reason to select one of the two alternatives, and both situations are likely to arise in different settings. At the end of the paper, we contrast the case of separable investments with the case of *perfect complements*. When investments are perfect complements, as in Jackson and Wolinsky (1996), both agents must contribute to the formation of a communication link, and the quality of the link depends on the minimal contribution of the two agents.

Once the strength of communication links is computed, we can specify the utility that agents derive from the communication network. When two agents are connected in the network, we compute the *reliability* of all paths connecting the two agents, and assume that they will communicate through the path of maximal reliability. We consider two alternative definitions of path reliability, corresponding to two different notions of reliability of communication links. In the *product reliability* model, the reliability of a path is given by the product of the strength of communication links along the path. This specification seems adequate to describe communication in social networks, where the diffusion of information depends on the strength of so-

[4] In this sense, the model is close to Bala and Goyal (2000a) who also assume that a bilateral link can be formed by a single agent.

cial links (the frequency and length of social interactions), and the value of a communication path is a function of the strength of all social links along the path. In the *min reliability* model, the reliability of a path is given by the strength of the weakest link along the path. This model corresponds to some formal communication networks (like the internet) where the speed of a connection between two agents depends on the bottleneck of the network, and hence the value of a connection path is a function of the weakest link along the path.[5]

The utility that one agent gets from another is simply the reliability of the optimal path connecting the two agents. Notice that if the two agents are directly connected to one another, then the optimal path may be the shortest path - namely the one consisting of the direct link. In this case, we will say that the agents derive *direct* benefits fron one another. If two connected agents do not serve direct benefits from one another, then we will say that they obtain *indirect* benefits.

The main results of our analysis concern the characterization of strongly *efficient* networks, which maximize the sum of utilities of all agents. With separable investments, we point out a crucial difference between the cases of increasing and decreasing marginal returns. With increasing returns, direct benefits are maximized when all agents concentrate their investment on a single link. This concentration of investments results in very strong links and points towards efficient network architectures which are minimally connected. We obtain a complete characterization of efficient networks when the transformation function is convex. In the product reliability model, where the length of indirect connections matters, the efficient network architecture is a star which minimizes the distance between agents. In the min reliability model, where distance is irrelevant, different network architectures can be supported as efficient networks, but in all of them, all agents but one concentrate their investment on a single link. In the particular case of perfect

[5] The difference between the product reliability and min reliability models can also be interpreted in terms of decay of information. In the product reliability model, messages are transformed along the communication network (parts of the messages can be lost, or changed), and the rate of transformation depends on the quality of the communication link. The total decay of information can thus be computed as the product of the decay along all communication links. In the min reliability model, there is no decay of information and the only issue is the speed of connection. The strength of a link is an index of the capacity of that link, and the speed of a connection only depends on the minimal capacity in the network.

substitutes, we furthermore show that all links must have equal strength, so that the hub of the star distributes equally its investment on all peripheral agents in the product reliability model, and any tree with links of equal strength can be supported as an efficient network in the min reliability model.

When returns to investment are decreasing, there is a conflict between the maximization of direct and indirect benefits The conflict arises because direct benefits are maximized when all agents distribute their investments on links with all other players, resulting in very weak communication links. Notice that in a complete graph where all links have the same strength, there are no indirect benefits at all. In general, weak communication links make indirect connections very poor, and the total value of the network depends on a complex trade-off between the strength of direct links and the value of indirect connections. In the product reliability model, we construct an example to show that different network architectures (including the complete network) can be efficient. In the min reliability model, efficient networks must be minimally connected, and we prove that the star is never efficient when the transformation function is strictly concave.

When investments are perfect complements, we obtain very different results. In the product reliability model, we show that trees cannot be efficient, and that the circle is efficient for small numbers of agents. However, an example shows that the efficient architecture is much more complex than the circle for larger numbers of agents. In the min-reliability case, the efficient architecture is either the circle or the line, which maximize the strengths of links in the network.

In the cases where we obtain a full characterization of efficient networks, we also investigate whether efficient networks can be supported by individual choices of agents, considering both the notion of Nash stability (Nash equilibria of a noncooperative game of link formation), and a version of the notion of pairwise stability introduced by Jackson and Wolinsky (1996). The tension between efficiency and stability noted in models with fixed quality of links also appears when agents choose the strength of links. We show that efficient networks may fail to be stable, except in two notable cases: perfect substitutes and perfect complements with min reliability. In these cases, the efficient network is perfectly symmetric, and any reallocation of investments results in asymmetric links which necessarily reduce the value of the weakest link in the network.

Related Work

We consider our analysis as a first step in the study of network formation with weighted links. In recent years, other studies have looked at networks with links of varying strength, and we now relate our contribution to these alternative studies.

We have already mentioned the papers of Bala and Goyal (2000a) and Jackson and Wolinsky (1996). We will discuss the relationship between these papers and ours later after describing our model formally. Bala and Goyal (2000b) model reliability as the exogenous probability that a link forms once agents have paid the cost of establishing the link. Communication networks are thus random graphs, and expected benefits can be computed as a polynomial function of the exogenous probability of link success. This formulation of network reliability in a probabilistic environment is very different from our modeling of the quality of links affecting the decay of information in a deterministic setting. In a specific model of strategic alliances among firms, Goyal and Moraga Gonzales (2001) consider a two-stage model where firms first form links and then decide their R&D investment in every bilateral relationship. The strength of a link (measured by the investments in R&D by both partners in the alliance) is thus determined endogenously. The analysis is conducted with a specific cost and demand formulation, and the main focus of the analysis is on regular networks (where all firms have the same number of links), and the effect of the number of links on the R&D investments. Durieu, Haller and Solal (2004) construct a model of nonspecific networking. Agents choose a single investment, which applies to the links with all other agents. This formulation seems adequate for settings where agents cannot discriminate among other agents in the society, but does not capture situations where agents choose to form bilateral links. Finally, Brueckner (2003) considers a model of friendship networks which is closely related to ours. Agents choose to invest in relationships, and the value of indirect benefits is given by the product of the strength of links, as in our model of min reliability. For most of his analysis, Brueckner (2003) concentrates on *three* player networks, and studies the effect of the network structure on the investment choices in the complete and star networks. In this paper, we consider an arbitrary number of agents, and simultaneously solve for the formation of the network and the choice of investments.

The rest of the paper is organized as follows. We introduce the model and notations in Section 2. In Section 3, we study networks with separable investments. Section 4 is devoted to the analysis of networks with perfect complements. Section 5 contains our Conclusions.

2 Model and Notations

Let $N = \{1, 2, ..., n\}$ be a set of individuals. Individuals derive benefits from links to other individuals. These benefits may be the pleasure from friendship, or the utility from (non-rival) information possessed by other individuals, and so on. In order to fix ideas, we will henceforth interpret benefits as coming from information possessed by other individuals. Each individual has a total resource (time, money) of $X > 0$, and has to decide on how to allocate X in establishing links with others.

Let x_i^j denote the amount of resource invested by player i in the relationship with j. Then, the *strength* of the relationship between i and j is a function of x_i^j and x_j^i. Let s_{ij} denote this strength as a function of the amounts x_i^j and x_j^i. Let this functional dependence be denoted as

$$s_{ij} = f(x_i^j, x_j^i)$$

with $s_{ij} \in [0, 1]$.

Alternative assumptions can be made about the functional form of f. Throughout much of the paper, we will focus on the following case.

Assumption 1 (Separable Investments) : For each $i, j \in N$, $f(x_i^j, x_j^i) = \phi(x_i^j) + \phi(x_j^i)$, where ϕ is strictly increasing, with $\phi(0) = 0$ and $\phi(X) \leq 1/2$.

With separable investments, every agent can individually form a link, and there is no complementarity in the investments by the two parties. A special case is when the function ϕ is linear, $\phi(x) = \alpha x$ for all x. In this case, the contributions of any i and j to the link ij are *perfect substitutes*. When the function ϕ is convex, returns to investments are *increasing*, and agents have an incentive to invest in a single link ; when the function ϕ is concave, returns to investments are *decreasing*, and agents have an incentive to distribute their investments equally across all links.

An alternative assumption is the following.

Assumption 2 (Perfect Complements): For each $i, j \in N$, $f(x_i^j, x_j^i) = \min(x_i^j, x_j^i)$.

Assumption 2 says that the contributions of i and j are *perfect complements*.

We say that individuals i and j are linked if and only if $s_{ij} > 0$. Each pattern of allocations of X, that is the vector $\mathbf{x} \equiv (x_i^j)_{\{i,j \in N, i \neq j\}}$ results in a weighted graph, which we denote by $g(\mathbf{x})$. [6] We say that $ij \in g(\mathbf{x})$ if $x_i^j + x_j^i > 0$.

Given any g, a path between individuals i and j is a sequence $i^0 = i, i^1, ..., i^m, ..., i^M = j$ such that $i^{m-1}i^m \in g$ for all m. Two individuals are *connected* if there exists a path between them. Connectedness defines an equivalence relation, and we can partition the set of individuals according to this relation. Blocks of that partition are called components, and we let $\mathcal{N}(g)$ denote the set of components of the graph g.

Suppose i and j are connected. Then, the benefit that i derives from j depends on the *reliability* with which i can access j's information.

For any pair of individuals i and j and graph g let $P(i,j)$ denote the set of paths linking i to j. Define

$$\begin{aligned} p^*(i,j) &= \arg \max_{p(i,j) \in P(i,j)} s_{ii^1} \ldots s_{i^{m-1}i^m} \ldots s_{i^{M-1}j}, \\ R^p(i,j) &= \max_{p(i,j) \in P(i,j)} s_{ii^1} \ldots s_{i^{m-1}i^m} \ldots s_{i^{M-1}j}. \end{aligned}$$

So, for any two individuals i and j, $p^*(i,j)$ denotes the path which has the highest reliability, as measured by the products of strengths of the links on that path. The reliability of that path is denoted $R^p(i,j)$.

As an alternative, we will also define reliability of any path in terms of the strength of the weakest link in that chain. Hence, we would have

$$R^m(i,j) = \max_{p(i,j) \in P(i,j)} \min_{s_{i^{m-1}i^m} \in p(i,j)} s_{i^{m-1}i^m}$$

Henceforth, we will refer to R^p and R^m as *Product-Reliability* and *Min-Reliability* respectively.

The utility that individual i gets from j is then $R^p(i,j)$ or $R^m(i,j)$ depending on the notion of reliability.

Notice that the optimal path between i and j could be the path $\{i,j\}$ itself. In this case, we will say that i and j derive *direct* benefits from each other. Of course, the direct benefit equals s_{ij}.

Individual i's payoff from a graph g is then given by

$$U_i(g) = \sum_{j \in N \setminus i} R^p(i,j) \quad \text{or} \quad U_i(g) = \sum_{j \in N \setminus i} R^m(i,j)$$

[6] To simplify notation, we will sometimes ignore the dependence of g on the specific pattern of allocations.

The value of a weighted graph g is given by $V(g) = \sum_i U_i(g)$.

Definition 1 *A graph g is efficient if $V(g) \geq V(g')$ for all g'.*

Our model is related to the original connections model of Jackson and Wolinsky (1996) and to the two-way flow model of Bala and Goyal (2000). However, there are significant differences. First, Jackson and Wolinsky (1996) and Bala and Goyal (2000a) assumed that a link between i and j either exists or not. That is, either $s_{ij} = 1$ or $s_{ij} = 0$. In contrast, we allow s_{ij} to take on any value between 0 and 1. Second, Jackson and Wolinsky (1996) assume that i and j each pay an exogenously given cost c if they form the link ij. Bala and Goyal (2000a), by contrast, suppose that the fixed cost is only paid by one of the agents. As we have described earlier, there are no exogenous costs of forming a specific link in our model. Instead, individuals have a "budget constraint" and only face opportunity costs for investing in one relation rather than in another. Lastly, the reliability function in Jackson and Wolinsky (1996) and Bala and Goyal (2000a) is also different from ours - they assume that the indirect benefit that i derives from j is δ^{d-1}, where d is the geodesic distance between i and j, while $\delta \in (0, 1)$ is a parameter.

Jackson and Wolinsky (1996) initiated the analysis of the potential conflict between efficiency and "stability" of endogenously formed networks. We now describe the concepts of stability that will be used in this paper.

Given any pattern of investments $\mathbf{x}$, and individual i, $(\mathbf{x_{-i}}, x'_i)$ denotes the vector where i deviates from x_i to x'_i. Similarly, $(\mathbf{x_{-i,j}}, x'_{i,j})$ denotes the vector where i and j have jointly deviated from (x_i, x_j) to (x'_i, x'_j).

Definition 2 *A graph $g(\mathbf{x})$ is Nash stable if there is no individual i and x'_i such that $U_i(g(\mathbf{x_{-i}}, x'_i)) > U_i(g(\mathbf{x}))$.*

So, a graph g induced by a vector $\mathbf{x}$ is Nash stable if no individual can change her pattern of investment in the different links and obtain a higher utility.

Definition 3 *A graph $g(\mathbf{x})$ is Strongly Pairwise Stable if there is no pair of individuals (i, j) and joint deviation (x'_i, x'_j) such that*

$$U_i(g(\mathbf{x_{-i,j}}, x'_{i,j})) + U_j(g(\mathbf{x_{-i,j}}, x'_{i,j})) > U_i(g(\mathbf{x})) + U_j(g(\mathbf{x}))$$

So, a graph is strongly pairwise stable if no pair of individuals can be jointly better off by changing their pattern of investment. Notice that we define a pair to be better off if the *sum* of their utilities is higher after the deviation. This leads to a stronger definition of stability than a corresponding definition with the requirement that both individuals be strictly better off after the deviation. The current definition implicitly assumes that individuals can make side payments to one another. The availability of side payments is consistent with our definition of efficiency - a graph is defined to be efficient if it maximises the *sum* of utilities of all individuals.[7]

Jackson and Wolinsky (1996) define a weaker notion of stability - *pairwise stability*. They basically restrict deviations by assuming that only one link at a time can be changed.[8]

Some specific network architectures will be important in subsequent sections. We define these below.

Definition 4 *A graph g is a star if there is some $i \in N$ such that $g = \{ik|k \in N, k \neq i\}$.*

The distinguished individual i figuring in the definition will be referred to as the "hub".

The degree of a node i in graph g is $\#\{j|ij \in g\}$. That is, the degree of a node equals the number of its neighbours.

Definition 5 *A k-regular graph is a graph where every node has degree k. A circle is the unique 2-regular graph.*

Definition 6 *A graph g is a line if there is a labelling of individuals $i_1, \ldots, i_N$ such that $g = \{i_k i_{k+1}|k = 1, \ldots, N-1\}$.*

[7]While we allow for side-payments among agents, we also assume that initial endowments are not transferable across players. If X is a fixed allocation of time, this is clearly understood. If X is a fixed budget, we assume that individuals can only make side-payments ex post, once the communication network has formed, but cannot trade their initial endowments.

[8]Our current definition corresponds to the definition of pairwise stability used by Dutta and Mutuswami (1997).See also Gilles and Sarangi (2004) and Bloch and Jackson (2004) for a comparison of different stability concepts.

3 Separable Investments

3.1 Product Reliability

In this subsection, we discuss the nature of efficient graphs with separable investments under product reliability. It will become clear in what follows that the efficient architecture involves a trade-off between the graphs which maximise direct and indirect benefits. It is easy to see that the network architectures which maximize direct benefits depend crucially on the concavity or convexity of the function ϕ. If the function ϕ is concave, direct benefits are maximized when all agents invest an equal amount $X/(n-1)$ on every link. Hence, direct benefits are maximized in the complete graph with symmetric value on every link. If the function ϕ is convex, direct benefits are maximized when every agent invests in a single link. This implies that there are at most n links in the optimal graph (but there can be as few as $n/2$ and the graph can be disconnected), and that every link has value either $\phi(X)$ or $2\phi(X)$. In both cases, graphs which maximize direct benefits may fail to maximize indirect benefits: In the concave case, in the complete graph, indirect benefits are equal to zero ; in the convex case, the concentration of resources on a single link may reduce aggregate indirect benefits. The conflict between maximization of direct and indirect benefits disappears when the function ϕ is linear. With perfect substitutes, direct benefits are maximal for any allocation of resources, and the characterization of the efficient network only involves the maximization of indirect benefits.This enables us to get particularly sharp characterizations of efficient graphs in this case.

Notice that a star connecting everyone has two important properties. First, it is a tree, and so minimizes the number of direct connections amongst all connected graphs. This increases the average strength of direct links. Also, the fewer the number of direct links, the greater is the scope for indirect benefits. Second, it also minimizes the distance between all nodes which do not have a direct connection. Given the product form of the reliability function, the latter also increases the scope of higher indirect benefits. Thus, for all these reasons, a star should be the most desirable architecture when returns to investments are increasing. The following theorem verifies this intuition.

Theorem 1 *Suppose Assumption 1 holds and that ϕ is convex. Then,* *(i) The unique efficient graph is a star with every peripheral agent investing fully in the arc with the hub.*

(ii) Morover, if ϕ is linear, the hub invests $\frac{X}{n-1}$ on every link with a peripheral agent and all links have the same strength.

Proof. : We prove the first part of the theorem in three steps. Consider any feasible component h of g of size m.

Step 1: We construct a star $\hat{S}$ with higher aggregate utility than h.

Step 2: If $\hat{S}$ is infeasible, then we construct a feasible star S^*which has higher aggregate utility than $\hat{S}$.

Step 3: If the graph g contains different components, we construct a single connected star which has higher aggregate utility than the sum of the stars S^*

Proof of Step 1: Let h have $K \geq m-1$ links. We label the strength of these links z_k and assume, without loss of generality, that

$$z_1 \geq z_2 \geq \ldots \geq z_K$$

We construct the star $\hat{S}$ with hub m as follows. For each $i \neq m$, define

$$\bar{x}_i^m = \min(\phi^{-1}(z_i), X), \bar{x}_m^i = \phi^{-1}(z_i) - \bar{x}_i^m$$

If g is a tree, then $K = m-1$. Let $\hat{x}_i^j = \bar{x}_i^j$ for all i, j, and $\hat{S} \equiv h(\hat{\mathbf{x}})$.

If g is not a tree, then $K > m-1$, and the investments involved in $\{z_m, \ldots, z_K\}$ have not been distributed.

Since $z_i \geq z_{i+1}$, $\bar{x}_i^m \geq \bar{x}_{i+1}^m$. If $\bar{x}_i^m = X$ for all $i = 1, \ldots, m-1$, then from convexity of ϕ, $\sum\limits_{i=1}^{m-1} \bar{x}_m^i < X$. In this case, let $\bar{x}_i^m = \hat{x}_i^m$ for all $i = 1, \ldots, m-1$, and $\hat{x}_m^1 = X - \sum\limits_{k=2}^{m-1} \bar{x}_m^k > \bar{x}_m^1$, and $\hat{x}_m^k = \bar{x}_m^k$ for $k = 2, \ldots, m-1$. Let $\hat{S} \equiv h(\hat{\mathbf{x}})$.[9]

Otherwise, let $\tilde{k} \geq 1$ be the smallest integer such that $\bar{x}_{\tilde{k}}^m < X$. Define $\hat{x}_i^m = \bar{x}_i^m = X$ for all $i < \tilde{k}$. Then, distribute $\{z_m, \ldots, z_K\}$ sequentially

[9]Notice that in this case $\hat{S}$ is a feasible graph since $\sum\limits_{j=1, j\neq i}^{m} x_i^j = X$ for all i.

amongst $\{\tilde{k}, \ldots, m-1\}$ as follows:

$$\begin{aligned}
\hat{x}^m_{\tilde{k}} &= \min(X, \bar{x}^m_{\tilde{k}} + \sum_{j=m}^{K} \phi^{-1}(z_j)) \\
\text{for } k > \tilde{k}, \hat{x}^m_k &= \min(X, \bar{x}^m_k + \sum_{j=m}^{K} \phi^{-1}(z_j) - \sum_{k'=\tilde{k}}^{k-1} (\hat{x}^m_{k'} - \bar{x}^m_{k'}))
\end{aligned}$$

Finally, if this procedure has not exhausted $\sum_{j=m}^{K} \phi^{-1}(z_j)$, then distribute the excess sequentially to $\hat{x}^1_m, \hat{x}^2_m, \ldots$, while satisfying the constraint that none of them exceeds X. Again, define $\hat{S} \equiv h(\hat{\mathbf{x}})$.

Suppose h is a tree. Then, from convexity of ϕ, we know that

$$\phi(\hat{x}^m_i) + \phi(\hat{x}^i_m) \geq z_i \text{ for all } i = 1, \ldots, m-1 \tag{1}$$

This implies that direct benefits are at least as high in $\hat{S}$ as in h. A similar argument ensures that direct benefits are at least as high in $\hat{S}$ as in h even when the latter is not a tree.

Now, we check that indirect benefits are strictly higher in $\hat{S}$. We distinguish between two cases.

Case 1: h is a tree.

Let $D = \{ij | ij \notin h\}$. For each pair i, j in D, let z_{k_i} and z_{k_j} denote the strengths of the *first* and *last* links in the path $p^*(i,j)$. Notice that while the choice of which link is first and which is last is arbitrary, the *pair* (z_{k_i}, z_{k_j}) is uniquely defined. Clearly, for all $i, j \in D$,

$$R^p(i,j) \leq z_{k_i} z_{k_j}$$

Moreover, since h is not a star, the inequality must be strict for some pair i, j. Hence, letting I denote the sum of indirect benefits in h and $\widehat{I}$ the corresponding sum in $\widehat{S}$, we have

$$I < 2 \sum_{i \neq j, i,j=1}^{m-1} z_i z_j \leq \widehat{I}$$

where the last inequality follows from equation 1.

Case 2: h is not a tree.

Let h have $K > m-1$ links. Again, let D be the pairs which derive indirect benefits from each other in h, and z_{k_i}, z_{k_j} the strengths of the first and last links in $p^*(i,j)$.[10] Similarly, let I and $\widehat{I}$ denote the sum of indirect benefits in h and $\hat{S}$ respectively. Now, the number of pairs of agents deriving indirect benefits in h is strictly less than $\frac{(m-1)(m-2)}{2}$, which is the number of pairs deriving indirect benefits in $\hat{S}$. Recalling that $z_i \geq z_{i+1}$ for all $i = 1, \ldots, K$,

$$I < \sum_{i \neq j, i,j=1}^{m-1} z_i z_j \leq \widehat{I}$$

This concludes the proof of Step 1.

Proof of Step 2: Notice that by construction, $\hat{x}_i^m \leq X$ for all $i = 1, \ldots, m-1$. So, either $\hat{S}$ is feasible, or $\sum \hat{x}_m^i > X$. We now show that if $\hat{S}$ is not feasible, then we can construct a star S^* with hub m which has higher total utility compared to $\hat{S}$.

Although $\hat{S}$ is not feasible,

$$\sum_{i=1}^{m-1} \left(\hat{x}_i^m + \hat{x}_m^i \right) \leq mX \tag{2}$$

Also,

$$\hat{x}_{m-1}^m < X$$

This follows because $\hat{x}_{m-1}^m \leq \hat{x}_i^m$ for all $i < m-1$, and from equation 2.

Consider the star S^* where

(i) $x_i^{m*} = \hat{x}_i^m$ for all $i \neq m-1$ and $x_m^{i*} = \hat{x}_m^i$ for all $i \neq 1$.

(ii) $x_{m-1}^{m*} = \min(X, \hat{x}_{m-1}^m + \hat{x}_m^1)$, and $x_m^{1*} = \hat{x}_{m-1}^m + \hat{x}_m^1 - x_{m-1}^{m*}$.

Then,

$$\phi(x_{m-1}^{m*}) + \phi(x_m^{1*}) \geq \phi(\hat{x}_{m-1}^m) + \phi(\hat{x}_m^1) \tag{3}$$

with strict inequality holding if ϕ is strictly convex. So, the sum of direct benefits in S^* is at least as large as in $\hat{S}$.

We now show that the sum of indirect benefits in S^* is also at least as high as in $\hat{S}$. Let I^* and $\hat{I}$ represent the indirect benefits from S^* and $\hat{S}$

[10] Since h is no longer a tree, the path $p^*(i,j)$ is not uniquely defined. The choice of best path is immaterial in what follows.

respectively. Since $\hat{x}^m_{m-1} < X$, $\hat{x}^{m-1}_m = 0$. Also, $\hat{x}^m_1 = X$ since $\hat{x}^1_m > 0$. Hence,

$$I^* - \hat{I} \geq \phi(X)[\phi(x^{m*}_{m-1}) - \phi(\hat{x}^m_{m-1})] + \phi(x^{1*}_m)\phi(x^{m*}_{m-1}) - \phi(\hat{x}^1_m)\phi(\hat{x}^m_{m-1})$$

Now, if $x^{m*}_{m-1} < X$, then $x^{1*}_m = 0$ and $\phi(x^{m*}_{m-1}) \geq \phi(\hat{x}^m_{m-1}) + \phi(\hat{x}^1_m)$. Since $\phi(X) \geq \phi(\hat{x}^1_m)$, this implies that $I^* \geq \hat{I}$.

Suppose $x^{m*}_{m-1} = X$. Then, $I^* \geq \hat{I}$ if

$$\phi(X)(\phi(X) + \phi(x^{1*}_m)) \geq (\phi(X) + \phi(\hat{x}^1_m))\phi(\hat{x}^m_{m-1}) \tag{4}$$

Convexity of ϕ ensures that

$$\phi(X) + \phi(x^{1*}_m) \geq \phi(\hat{x}^1_m) + \phi(\hat{x}^m_{m-1})$$

Since $X \geq \hat{x}^1_m$, this ensures that equation 4 is satisfied.

Note that $\sum_{i=1}^{m-1}(x^{i*}_m - \hat{x}^i_m) < 0$. If S is not feasible, clearly we can proceed in this way by increasing x^m_{m-2}, x^m_{m-3}, etc. until a feasible star is obtained.

This completes the proof of Step 2.

Proof of Step 3: Consider two feasible stars S^*_1 and S^*_2 of sizes m_1 and m_2. Construct a new star $S^{**}$of size $m_1 + m_2$ centered around the hub of S^*_2 with the following investments:

$$\begin{aligned} x^{m_2**}_i &= X \text{ for all } i < m_2 \\ x^{i**}_{m_2} &= x^{i*}_{m_2} \end{aligned}$$

As ϕ is increasing and convex, the aggregate direct benefits in star S^{**} are at least as high as the sum of aggregate benefits in the two stars S^*_1 and S^*_2. Consider then indirect benefits in the new star, I^{**} and the sum of indirect benefits in the two stars, $I^*_1 + I^*_2$. Indirect benefits inside the star S^*_2 have not changed, and peripheral nodes have gained access to new indirect connections. The difference in aggregate indirect benefits for agents in the star S^*_1 is given by

$$
\begin{aligned}
I_1^{**} - I_1^* &= \sum_{i \in S_1 \backslash \{m_1\}, j \in S_2 \backslash \{m_2\}} \phi(X)(\phi(X) + \phi(x_{m_2}^{j*})) + m_1(m_1-1)\phi(X)^2 \\
&\quad - \sum_{i \in S_1 \backslash \{m_1\}, j \in S_1 \backslash \{m_1\}} (\phi(x_i^{m_1*}) + \phi(x_{m_1}^{i*}))(\phi(x_j^{m_1*}) + \phi(x_{m_1}^{j*})) \\
&> m_1(m_1-1)\phi(X)^2 - \sum_{i \in S_1 \backslash \{m_1\}, j \in S_1 \backslash \{m_1\}} (\phi(X) + \phi(x_{m_1}^{i*}))(\phi(X) + \phi(x_{m_1}^{j*})) \\
&= 2(m_1-1)\phi(X)^2 - 2(m-2)\phi(X) \sum_{i \in S_1 \backslash \{m_1\}} (\phi(x_i^{m_1*}) - 2 \sum_{i,j \in S_1 \backslash \{m_1\}} \phi(x_{m_1}^{i*})\phi(x_{m_1}^{j*}).
\end{aligned}
$$

By convexity of ϕ,

$$
\sum_{i \in S_1 \backslash \{m_1\}} (\phi(x_i^{m_1*}) \leq \phi(X)
$$

and

$$
\sum_{i,j \in S_1 \backslash \{m_1\}} \phi(x_{m_1}^{i*})\phi(x_{m_1}^{j*}) < (\sum_{i \in S_1 \backslash \{m_1\}} (\phi(x_i^{m_1*}))^2 \leq \phi(X)^2
$$

so that

$$
I^{**} > I_1^* + I_2^*.
$$

As the same argument can be repated with any pair of stars, this argument completes step 3 and the proof of the first part of the Theorem.

(ii)Suppose $\phi(x) = \alpha x$ for all x. Let S^* be the star with hub at n, where all arcs have strength $X + \frac{X}{n-1}$. Let S be any other star with hub n where s_{in} may not be equal to s_{jn} for $i \neq j$, but where $\sum_{i=1}^{n-1}(x_i^n + x_n^i) = nX$. It is straightforward to check that total direct benefits are maximised at both S^* and S. We now show that the sum of *indirect* benefits in S^* is greater than that in S.

Without loss of generality let s_{1n} and s_{2n} denote the weakest and strongest links in S. Consider the effect of increasing investment on s_{1n} by ε and simultaneously decreasing investment on s_{2n} by ε.

The effect on the overall value can be computed as

$$
\Delta V = 2[\varepsilon(s_{2n} - s_{1n}) - \varepsilon^2]
$$

Hence, for ε small enough, $\Delta V > 0$ and so local changes in the direction of equalization are profitable. But this implies that the symmetric star has higher value than the asymmetric star. ■

Theorem 1 shows that with increasing returns the efficient architecture is a star where all peripheral nodes invest fully in the link with the hub, but does not characterize the allocation of investments by the hub. Except in the case where investments are perfect substitutes, there is a conflict between maximization of direct and indirect benefits. In order to maximize direct benefits, the hub should invest fully in one link with a peripheral node. This however lowers aggregate benefits for all other peripheral nodes, and may result in a lower total value of the network. The optimal distribution of investments of the hub thus depends on the curvature of the function ϕ. To illustrate this point, we consider in the following example a quadratic transformation function, and characterize the optimal allocation of investments in the hub as a function of the parameters of the utility function.

Example 1 *Let* $\phi(x) = \frac{1}{2}[\lambda x + (1-\lambda)x^2]$*, for some* $\lambda \in [0,1]$ *and assume* $X = 1$

For $\lambda = 1$, the transformation function is linear, and for $\lambda = 0$, the function is quadratic. Lower values of the parameter λ correspond to higher degrees of convexity of the transformation function. Consider a star where the hub, denoted n, allocates his investments on the different peripheral nodes, $x_n^1, ..., x_n^{n-1}$ and each peripheral node invests fully on the link to the hub. The value of the star is given by:

$$\begin{aligned} V &= \frac{n(n-1)}{2} + 2n(1-\lambda) + (\lambda^2 - 2n(1-\lambda)) \sum_i \sum_j x_n^i x_n^j \\ &\quad + \lambda(1-\lambda) \sum_i \sum_j x_n^i x_n^j (x_n^i + x_n^j) + (1-\lambda)^2 \sum_i \sum_j (x_n^i x_n^j)^2. \end{aligned}$$

If x_n^k is given for all $k \neq i, j$, $(x_n^i + x_n^j) = 1 - \sum_{k \neq i,j} x_n^k$ is fixed, and V can be written as a quadratic convex function of the product $x_n^i x_n^j$. As V is a convex function, it reaches its maximum either at $x_n^i x_n^j = 0$ or $x_n^i x_n^j = (x_n^i + x_n^j)^2/4$. We thus conclude that, at the optimum, if the hub invests a positive amount on two nodes x_n^i and x_n^j, then it must invest the same amount on these two nodes. In other words, at the optimum, the hub chooses the number k of links on which it invests, and invests an equal amount $1/k$ on each of those links. The following table lists, for $n = 5$ and different values of λ, the number of links on which the hub invests at the optimum.

λ	optimal number of links
0	1
0.9	1
0.91	2
0.92	3
0.93	4
1	4

Table 1: optimal allocation of the hub's investment with a quadratic function

Table 1 shows that for a large fraction of the paramater range, the hub optimally invests in a single link. As the degree of convexity of the transformation function goes down, the number of links on which the hub optimally invests increases, and when the function is linear, the hub invests equally on all links, as we showed in Theorem 1.

When the transformation function ϕ is concave, the optimal network architecture depends on the specification of the transformation function. This is easily illustrated in the simplest case with three agents. Depending on the function ϕ, either the complete symmetric network or the star network where the hub invests equally on the two links with peripheral agents can be efficient. More precisely, the complete network is efficient if and only if:

$$4\phi(X/2) - 2\phi(X) \geq (\phi(X) + \phi(X/2))^2.$$

Hence, the complete network will be efficient when the degree of concavity of the transformation function is high, and the value of a connection is low. With more agents, a number of network architectures can emerge as efficient architectures, depending on the degree of concavity of the transformation function, as illustrated in the following example.

Example 2 *Suppose that* $\phi(x) = \alpha \log(1 + x)$ *,* $X = 1$ *and* $\alpha \leq 1/(2 \log 2)$*.. Let* $n = 4$*.*

We compare the value of benefits for three network architectures: the symmetric complete graph C, the symmetric cycle Y, and the symmetric star S.[11] We compute:

[11] Given that ϕ is concave, these allocations of investments are optimal for the given network architectures.

$$\begin{aligned}
V^C &= 24\alpha \log(4/3), \\
V^Y &= 16\alpha \log(3/2) + 16\alpha^2 (\log(3/2))^2, \\
V^S &= 6\alpha(\log 2 + \log(4/3)) + 6\alpha^2(\log 2 + \log(4/3))^2
\end{aligned}$$

Let $\alpha_1 = \frac{3\log(4/3) - 2\log(3/2)}{2(\log(3/2))^2} \approx 0.158$ and $\alpha_2 = \frac{8\log(3/2) - 6(\log 2 + \log(4/3))}{6(\log 2 + \log(4/3))^2 - 8(\log(3/2))^2} \approx 0.185$. We can easily check that the complete network dominates the two other architectures for $\alpha \leq \alpha_1$, that the symmetric star dominates the two other architectures for $\alpha \geq \alpha_2$ and that the circle dominates the two other architectures for $\alpha_1 \leq \alpha \leq \alpha_2$. While this example does not explicitly characterize the optimal architecture, it shows that for intermediate values of concavity, neither the complete graph nor the star are optimal.

We now turn our attention to stable networks. Since Jackson and Wolinsky (1996), it is well known that stable and efficient networks may not coincide when agents only choose the set of links that they form. This conflict between private and social incentives to build links derives from two sources. First, in Jackson and Wolinsky (1996), agents face a coordination problem in the formation of links, as both agents must agree for the link to be formed.[12] Second, in most formulations of network benefits, the formation of a link induces externalities on the other agents, and these externalities cannot be internalized in standard models of network formation.[13]

When investments are separable, every agent can independently contribute to the formation of the link, and the coordination problem is not too severe. However, externalities still play a crucial role, and may prevent the private formation of efficient networks. To understand this fact, consider a situation without externalities, where every agent only derives utility from the agents she is directly connected to.It is easy to see that in that case, the set of efficient and Nash stable networks coincide. If ϕ is concave, it is a dominant strategy for every agent to allocate his resources symmetrically on all links ; if ϕ is convex, it is a dominant strategy for all the agents to concentrate all their investments on a single link.

When agents derive utility from indirect connections, externalities become important, and the model with endogenous link strength does not behave

[12] It is precisely this coordination problem which led Jackson and Wolinsky to define the concept of pairwise stability, where pairs of agents can coordinate their moves to form a new link.

[13] See Bloch and Jackson (2004) for the study of a model with transfers, where agents can internalize some of those externalities.

differently from standard models of link formation. There is generally a conflict between stable and efficient networks. Consider again Example 1. For λ high enough, the efficient graph is a star where the hub *does not* specialize completely in one of the links. But, then such a graph cannot be *Nash stable* because the agent at the hub is strictly better off by investing X in just one of the links - the hub does not derive any indirect benefits and so is better off by maximising her own direct benefit.

Now, consider the case when $\phi(x) = \alpha x$ for all x. Suppose $n = 3$. In the symmetric star, the hub (say individual 1) derives benefits of $3X$ and will never want to deviate from the star. Consider peripheral individuals 2 and 3. It is easy to check that the best joint deviation is for both 2 and 3 to set $x_2^3 = x_3^2 = X$. Then, 2 and 3 each get additional *direct* benefits of X. However, they lose indirect benefits from each other. This loss equals $\frac{9}{4}X^2$. Hence, the star is stable iff $X \geq \frac{4}{9}$.

Suppose $n > 3$. Again, let 1 be the hub. As before, individual 1 has no profitable deviation. Consider 2 and 3, and potentially mutually profitable deviations where $x_2^3 = \delta_2, x_2^1 = X - \delta_2$ and $x_3^2 = \delta_3, x_3^1 = X - \delta_3$, while other investments are as in the symmetric star. Without loss of generality, let $\delta_2 \leq \delta_3$.

Individual 3 gains an additional direct benefit of δ_2 from 2, but loses the indirect benefit of $(X + \frac{X}{n-1})^2$ from 2. In addition, 2 also suffers a loss in indirect benefit from each of the other peripheral nodes. The exact loss depends on whether $p^*(3, i)$ for $i > 3$ includes 31 or $(32, 21)$.[14] Hence, the loss in indirect benefit from each of the other nodes $4, \ldots, n$ is at least $\delta_2(X + \frac{X}{n-1})$. So, the deviation is profitable if

$$\delta_2 > \delta_2(n-3)(X + \frac{X}{n-1}) + (X + \frac{X}{n-1})^2$$

If this inequality holds for some $\delta_2 < X$, it must also hold for $\delta_2 = X$. Substituting this value of δ_2, and simplifying, we get

$$X < \frac{(n-1)^2}{n(n^2 - 3n + 3)}$$

Putting these observations together, we get the following.

[14]The latter is a possibility since we have assumed $\delta_3 \geq \delta_2$.

Theorem 2 *Suppose that Assumption 1 holds.*

(i) There exists a convex function ϕ such that no efficient graph is Nash stable.

(ii) Suppose ϕ is linear. Then, the symmetric star, which is the unique efficient graph, is strongly pairwise stable iff $X \geq \frac{(n-1)^2}{n(n^2-3n+3)}$.

Theorem 2 thus illustrates the classic conflict between efficiency and stability in our model. In the case of perfect substitutes, we note that, for $n \geq 3$, $\frac{(n-1)^2}{n(n^2-3n+3)}$ is decreasing in n so that, as n increases, the lower bound on X decreases. This is easily interpreted. If two nodes divert investment from their links with the hub, they lose indirect benefits from other peripheral nodes. This loss increases with the number of peripheral nodes. As a consequence, it is easier to sustain the efficient network as a stable network when the number of agents in the society increases.

3.2 Min Reliability

In this subsection, we first describe efficient networks when the reliability of a path is measured by the strength of its weakest link. We then go on to check whether efficient networks are stable.

Our first result shows that the efficient network must be a tree. With increasing returns, as in the case of product reliability, the star where all peripheral agents invest fully in their relation with the hub is efficient. With perfect substitutes, the hub invests equally in all its relations with the peripheral agents. With decreasing returns, the conflict between maximization of direct and indirect benefits again prevents a complete characterization of the efficient networks.

Theorem 3 *Suppose that Assumption 1 holds.*

(i) For any strictly increasing function ϕ, the efficient network is a tree.

(ii) If ϕ is convex, any efficient network is equivalent to a star where all peripheral agents invest fully in their relation with the hub.

(iii) Moreover, if ϕ is linear, any efficient network is equivalent to a star where the hub invests equally on all its links with peripheral agents.

Proof. (i) Consider a graph g with a cycle, and pick one of the links of the cycle with minimal strength. Label i and j the two nodes connected by this link, and assume without loss of generality that $x_i^j > 0$. Let k be a node directly connected to i such that $p^*(i,k) = s_{ik}$ (the value of the direct connection between i and k is at least as high as the value of any indirect path between i and k). Consider then a new graph architecture $\overline{g}$ where $\overline{x}_i^j = 0$ and $\overline{x}_i^k = x_i^k + x_i^j > x_i^k$. We claim that aggregate benefits are higher in $\overline{g}$ than in g. For any nodes l, m such that $ij \notin p^*(l,m)$, the value of the connection is at least as high under $\overline{g}$ than under g. Furthermore, as $\phi(\overline{x}_i^k) > \phi(x_i^k)$, the value of the connection between i and k is strictly greater in the new graph $\overline{g}$. Next, consider l, m such that $ij \in p^*(l,m)$. Let l and m now be connected by using the path $i, i-1, ..., j+1$.Because ij is a link of minimal strength in the cycle, the new path must have value at least as high as ij. This argument shows that the value of the connection between l and m in the graph $\overline{g}$ is at least as high as the value of the connection in the graph g.

(ii) The proof follows the same lines as the proof of Theorem 1

Step 1: Start with a tree g of size m and label the link strengths:

$$z_1 \geq z_2 \geq \ldots \geq z_{m-1}$$

Construct the star S as in the proof of Theorem 1. Direct benefits are at least as large in S as in g. The sum of indirect benefits in the star S is easily calculated.

$$I = 2\sum_{i=1}^{m-2}(m-i-1)z_{m-i}$$

To check this, note that the link z_{m-1} is the minimum in all comparisons, and so the sum of indirect benefits obtained between $m-1$ and other nodes is $2(m-2)z_{m-1}$. Similarly, z_{m-2} is the minimum in $(m-3)$ comparisons and so on. Note that the arc z_{m-1} must also be involved in at least $(m-2)$ comparisons in the graph g (the minimum being attained if z_{m-1} connects some terminal node). Similarly, the arcs $\{z_{m-2}, z_{m-1}\}$ must be the minimum in at least $(2m-5)$ connections, and so on. This establishes that the sum of direct and indirect benefits is at least as high in S as in g.

Suppose that S is feasible. If $z_{m-1} < \phi(X)$, then clearly there is a feasible star whose total utility is strictly higher than that of S and hence that of g. If $z_{m-1} = \phi(X)$, then either direct benefits are strictly higher in S or the original g was such that $m-1$ nodes specialized fully in their investment.

Step 2: Suppose that S is not feasible. Then, $z_{m-1} < \phi(X)$. That is, $x^m_{m-1} = \phi^{-1}(z_{m-1}) < X$. Let

$$i = \min\{k | x^m_k < X\}$$

and

$$j = \max\{k | x^k_m > 0\}$$

Construct the star $\bar{S}$ where all arcs except the following are the same as in S.

$$\bar{x_i}^m = \min(X, x^m_i + x^j_m), \bar{x}^j_m = x^j_m - \bar{x}^m_i + x^m_i$$

Strict convexity ensures that direct benefits are at least as high since

$$\phi(\bar{x}^m_i) + \phi(\bar{x}^j_m) \geq \phi(x^m_i) + \phi(x^j_m) \tag{5}$$

Note that the following is true for all $k, l = 1, \ldots, m-1$:

$$(x^m_k + x^k_m) \geq (x^m_l + x^l_m) \rightarrow (\bar{x}^m_k + \bar{x}^k_m) \geq (\bar{x}^m_l + \bar{x}^l_m)$$

That is the ordering of nodes $1, 2 \ldots, m-1$ in terms of (decreasing) order of strength of links is the same in the two stars. Since $z_j > z_i$, the reallocation must have increased indirect benefits too.

Step 3: The argument can be repeated for all components of the graph. We now show that if the graph contains two stars S_1 and S_2, it is dominated by the graph where the two stars are merged into a single stars, as in the proof of Theorem 1. By merging the two stars centered around m_1 and m_2 into a single star with hub m_2, direct benefits have increased. Furthermore, indirect benefits for players in star S_2 have strictly increased. For a peripheral agent i in star S_1, indirect benefits were equal to

$$\begin{aligned} I_i &= \sum_{j \in S_1 \backslash m_1} \phi(X) + \min\{\phi(x^i_{m_1}), \phi((x^j_{m_1})\} \\ &= (m_1 - 1)\phi(X) + \sum_{j \in S_1 \backslash m_1} \min\{\phi(x^i_{m_1}), \phi((x^j_{m_1})\}. \end{aligned}$$

In the new star, indirect benefits are given by:

$$I^*_i = (m_1 + m_2 - 1)\phi(X).$$

As $m_2 \geq 1$ and $\phi(X) \geq \sum_{j \in S_1 \backslash m_1} \min\{\phi(x^i_{m_1}), \phi((x^j_{m_1})\}, I^*_i \geq I_i$ and indirect benefits cannot have decreased.

(iii) Suppose that ϕ is linear, and that the hub invests different amounts on its links with peripehral nodes. Let i be a node for which x_n^i is maximal, and j a node for which x_n^j is minimal. Consider a reallocation of investments, $\widetilde{x_n^i} = x_n^i - \varepsilon$, $\widetilde{x_n^j} = x_n^j + \varepsilon$. This reallocation does not affect direct benefits. For ε small enough, it only reduces indirect benefits between i and other players who are connected to the hub by links of maximal strength by ε. Indirect benefits between player j and all these players have been increased by ε, and indirect benefits between players $i$and j have strictly increased. Hence, this reallocation of investments has strictly increased the value of the graph, showing that the hub must put equal weight on all links with peripheral agents. ■

Theorem 3 shows that every pair of agents is connected by a single path when reliability is measured by the strength of the weakest link. This result is easily explained: if two agents are connected by two paths, a reallocation of investments towards one of the two paths will necessarily increase the minimal value of the connection, and result in higher benefits. When the function ϕ is convex, Theorem 3 shows that a star is efficient, but does not determine the allocation of the investments of the hub. As in the case of product reliability, the following example shows that the hub will either choose to invest fully in one link, or to allocate its investment across different links, depending on the convexity of the function ϕ.

Example 3 *Let $\phi(x) = \frac{1}{2}x^\alpha$ with $\alpha \geq 1$ and $X = 1$.*

For $n = 3$, we compute the value of the star when the hub allocates x on its link with one of the nodes, and $(1 - x)$ on the other:

$$V = \frac{3}{2} + \frac{1}{2}\{x^\alpha + (1-x)^\alpha + \min\{x, 1-x\}^\alpha\}.$$

It is easy to check that V is a convex function of x on the two intervals $[0, 1/2]$ and $[1/2, 1]$, so that the maximum is either obtained for $x = 0$ or $x = 1/2$ or $x = 1$. A simple computation shows that if $\alpha \leq \frac{\log 3}{\log 2}$, the maximum is attained at $x = 1/2$ and if $\alpha \geq \frac{\log 3}{\log 2}$, the maximum is attained either at $x = 0$ or $x = 1$.

Contrary to the case of product reliability, the star need not be the only efficient architecture. Because distances do not matter with min reliability, a number of alternative network architectures can give rise to the same value as the star. Suppose for example that a star where the hub invests fully in one of the links is efficient. Then a line, where the second node invests fully in its connection with the first agent gives rise to the same aggregate benefits as the star. In the case of perfect substitutes, a star where the hub invests equally in all its links is efficient. Then, as the following proposition shows, *any* tree where all links have equal strength is an efficient network.

Proposition 1 *Suppose that Assumption 1 holds. If ϕ is linear, then any tree can be supported as an efficient network.*

Proof. We propose an algorithm which shows that for any tree, there exists a feasible allocation of investments such that all links have equal strength.Pick any node x_0 as the root of the tree. Define then a binary predecessor relation corresponding to this root : $i \prec j$ if and only if $i \in P(j, x_0)$ where $P(j, x_0)$ is the unique path from j to x_0. One may also define the immediate predecessor of a node j as the node i such that $i \prec j$ and if $k \prec j$ and $k \neq i$ then $k \prec i$. Given that the network is a tree, this unique immediate predecessor is well defined. Now, to any node x in the tree, attach an integer $\kappa(x)$ which corresponds to the number of nodes which have x as a predecessor, i.e., $\kappa(x) = \#\{i, x \prec i\}$. Clearly, if x is a leaf of the tree, $\kappa(x) = 0$ and for the root, $\kappa(x_0) = n-1$. For any node x, let $I(x) = \{y_1, ..., y_m, ..., y_M\}$ denote the set of nodes which admit x as an immediate predecessor. Finally, consider the following allocation of resources for node x: invest $X(\kappa(y_m)+1)/(n-1)$ on each point $y_m \in I(x)$ and invest the remainder, $X(n-1-\kappa(x))/(n-1)$ on the relation with the unique immediate predecessor of x. Clearly, this allocation of resources satisfies individual budget balance. Now, consider any link between x and y and assume without loss of generality that x is the immediate predecessor of y. Then the value of the link is $X(\kappa(y) + 1 + n - 1 - \kappa(y))/(n-1) = nX/(n-1)$ which is independent of x and y. Hence, this allocation of resources results in all links having equal strength. ■

Turning to the case of decreasing returns, we face the same conflict between maximization of direct and indirect benefits as in the case of product reliability. The following Proposition provides a very partial characterization of efficient networks in that case. It implies that a star can never be efficient.

Proposition 2 *Suppose that Assumption 1 holds , the function ϕ is strictly concave and $n \geq 4$. If a network is efficient, then no agent can put positive investment on two or more terminal nodes.*

Proof. Consider a graph where some agent i is connected to two terminal agents j and k. Because $n \geq 4$, agent i is also connected to some other agent l in the network from whom i derives direct benefits. Without loss of generality, suppose that $x_i^j \geq x_i^k > 0$ so that $s_{ij} = X + x_i^j \geq s_{ik} = X + x_i^k$.

Consider first the case where $x_i^j + x_i^k = X$. We then have $s_{il} \leq X < s_{ij}$. Consider then a reallocation $\widehat{x}_i^l = x_i^l + \varepsilon = \varepsilon$, $\widehat{x}_i^j = x_i^j - \varepsilon$. By concavity, for ε small enough,

$$\phi(\widehat{x}_i^l) + \phi(\widehat{x}_i^j) > \phi(x_i^l) + \phi(x_i^j),$$

so this allocation results in an increase in direct benefits. Furthermore, consider indirect benefits. This reallocation has reduced all indirect benefits from player j through player i where s_{ij} is the minimum link value. (So in particular, this does not change indirect benefits using the link il). For all these indirect benefits, the loss is equal to $\phi(\widehat{x}_i^j) - \phi(x_i^j)$. On the other hand, this reallocation has increased at least all indirect connections from player l through player i where s_{il} is the mimimum link value. For all these indirect benefits, the increase is equal to $\phi(\widehat{x}_i^l) - \phi(x_i^l)$. Clearly, the number of indirect connections increased by the reallocation is at least as large as the number of indirect connections reduced by the reallocation, and as $\phi(\widehat{x}_i^l) + \phi(\widehat{x}_i^j) > \phi(x_i^l) + \phi(x_i^j)$ the total value of indirect connections has increased. This shows that the initial graph cannot be efficient.

Consider now the case where $x_i^j + x_i^k < X$ and consider the following reallocation $\widehat{x}_i^j = x_i^j + x_i^k$, $\widehat{x}_j^i = X - x_i^k$, $\widehat{x}_j^k = x_i^k, \widehat{x}_k^j = X$. After this reallocation, the links become s_{ij}, s_{jk} and the value of the terminal link, s_{jk} is the same as the value of the terminal link s_{ik}. On the other hand, we claim that the value of link s_{ij} has strictly increased. In fact,

$$\widehat{s}_{ij} = \phi(\widehat{x}_i^j) + \phi(\widehat{x}_j^i) = \phi(x_i^j + x_i^k) + \phi(X - x_i^k)$$

whereas

$$s_{ij} = \phi(x_i^j) + \phi(X).$$

As long as $x_i^j + x_i^k < X$, $x_i^j \geq x_i^k > 0$ and , and by strict concavity of the function ϕ,

$$\phi(x_i^j + x_i^k) + \phi(X - x_i^k) > \phi(x_i^j) + \phi(X).$$

This argument shows that this reallocation has strictly increased direct benefits. But the only indirect benefits which have been affected by this reallocation are the indirect benefits flowing from connections to j and k. Now as $\widehat{s_{ij}} > s_{ij} \geq s_{ik} = \widehat{s_{ik}}$, the value of indirect connections to player k are at least as large after the reallocation. Similarly, as $\widehat{s}_{ij} > s_{ij}$, the value of indirect connections to player j cannot have beend reduced by the reallocation. This shows that the initial network cannot be efficient. ■

Corollary 1 *Suppose Assumption 1 holds, and ϕ is strictly concave. If $n \geq 4$, then an efficient graph cannot be a star.*

Proof. Suppose g is a star. Then, if g is efficient, the hub must be investing fully on one link. Then, the aggregate utility from g is equivalent to that of a line g' where one of the links, say ij has strength $s_{ij} = 2\phi(X)$, while all other links have strength $\phi(X)$.

Without loss of generality, let i not be a peripheral node in g', and suppose $ik \in g$. Let i split her investment between j and k. From strict convexity, direct benefits go up. Also, since g' is a line where all links other than ij have strength $\phi(X)$, there is no loss in indirect benefits. Hence, g' (and so g) cannot be efficient. ■

We now examine the stability of efficient networks when the reliability function is R^m. If ϕ is strictly convex, Example 3 shows that the efficient network architecture may fail to be Nash stable. As in the case of product reliability, the maximization of indirect benefits requires that the hub divides its investment over two nodes for α low enough. However, the hub of the star has an incentive to concentrate its investment on a single link in order to maximize his private benefits. In the case of perfect substitutes, no agent has an incentive to deviate from an equal allocation of link strengths, as this would reduce the value of the weakest link. Hence, we obtain:

Theorem 4 *Suppose that Assumption 1 holds.*

(i) There exists a strictly convex ϕ such that the efficient network is not Nash stable.

(ii)If ϕ is linear, then, any tree with equal strength on all links is strongly pairwise stable.

Proof. We only prove (ii). Every link in a symmetric tree has strength $\frac{n}{n-1}X$. Since distance between nodes does not matter under R^m, every node derives a benefit of $\frac{n}{n-1}X$ from every other node. It is easy to check that no deviation by a pair can improve both individuals' payoffs. ■

4 Perfect Complements

We now consider the case of perfect complements. In this case, both agents need to invest for the link to be formed. This corresponds to friendship networks where both agents must exert effort for a communication link to be established. We first obtain a partial characterization result for product reliability, and then give a complete description of efficient networks for min reliability.

4.1 Product reliability

When reliability is measured by the product of the strengths of communication links, we first show that, as opposed to the case of separable investments, the efficient network must contain a cycle.

Proposition 3 *Suppose Assumption 2 holds. If g is efficient, it cannot have any component with three or more nodes which is a tree.*

Proof. Suppose g is efficient and has a component with three or more nodes, where two nodes have degree one. Denote these nodes by i and j and their immediate predecessors by k and l respectively. Because the component is connected, the degrees of k and l are necessarily greater than one. But this implies that $x_k^i < X$ and $x_l^j < X$. Furthermore because $\sum_{m \in N \setminus \{i\}} x_k^m \leq X, x_k^m \leq X - x_k^i$ for all node $m \neq i$ to which k is connected. Now, this implies that the value of the indirect connection between i and j in the graph is strictly smaller than $\min\{X - x_k^i, X - x_l^j\})$. Furthermore, in an efficient graph, $x_i^k = x_k^i$ and $x_j^l = x_l^j$ so that individual i can invest $X - x_k^i$ in the direct link with j and individual j can invest $X - x_l^j$ in the link with i. But, because the value of the indirect link is smaller than $\min\{X - x_k^i, X - x_l^j\}$, the investment in the direct link strictly increases the value of the graph, yielding a contradiction.

The proof of the theorem is completed with the observation that in a tree, at least two nodes have degree one. ∎

Unfortunately, we have been unable to characterize completely efficient networks when inputs are perfect complements. For low values of n, we can show that the circle is the unique efficient network structure.

Proposition 4 *Suppose Assumption 2 holds. For $3 \leq n \leq 7$, the circle where every link has value $X/2$ is the unique efficient network.*

Proof. For $n = 3$, the circle is the only connected graph which is not a tree. Now, notice that direct benefits are equal to nX and hence are maximal in the circle. For $n = 4, 5$, we show that the circle also maximizes the value of indirect benefits. Notice first that the value of an indirect connection is always bounded above by $(X/2)^2$ as the middle player must allocate X over at least two links. For $n = 4$ and $n = 5$ all indirect connections in the circle are of length 2 and have value $(X/2)^2$. Hence, the circle achieves the highest sum of indirect links and is efficient. It is easy to check that any other allocation of investments results in a lower value of indirect links, so the circle with links of equal strength is uniquely efficient.

Suppose now that $n = 6, 7$. The indirect benefit for any node in the circle is

$$I = \frac{X^2}{2} + \frac{X^3}{4}$$

Consider any other graph g. If this graph is to "dominate" the cycle, then at least one node (say i^*) has to derive an indirect benefit exceeding I. For each k, check that the circle maximises indirect benefits from nodes at a distance of k. So, if i is to derive a larger indirect benefit in g, it must have more than two nodes at a distance of 2.[15]

It is tedious to show that the maximum indirect benefit that i^* can derive occurs when i^* has two neighbours, j_1, j_2, with each neighbour of i^* having three neighbours including i^* itself. Moreover, the optimum pattern of allocation from the point of view of i^* is

$$x_i^{j_1} = x_i^{j_2} = x_{j_2}^i = x_{j_2}^i = \frac{1}{2}$$

This yields i^* a total indirect benefit of $\frac{X^2}{2} < I$. ■

For larger numbers of players, the above argument does not hold. The circle may be dominated by denser graphs, where the number of indirect connections is lower but the distance of indirect connections is lower as well. The following example shows that the circle can indeed be dominated by another network architecture (the "Petersen graph")[16] for $n = 10$.

[15] Since $n \leq 7$, the maximum distance between any two nodes in the circle is 3.

[16] See Holton and Sheehan(1993).

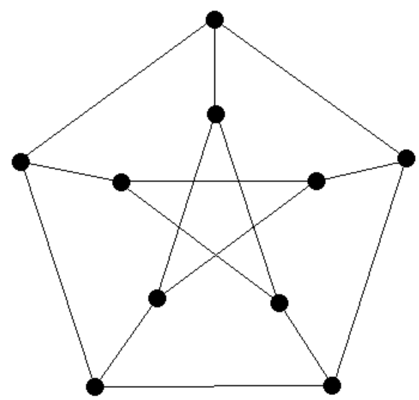

Figure 1: The Petersen graph

Remark 1 *Let $n = 10$. The circle may be dominated by the Petersen graph.*

The Petersen graph is a regular graph of degree 3 such that, for any node i, and any neighbors j and k of i, the set of direct neighbors of j and k (other than i) is disjoint. Consider the Petersen graph where every link has value $X/3$. The direct benefits are maximized, and each player has 6 indirect connections of length 2 and value $(X/3)^2$. Hence, the value of indirect connections for any player is $6X^2/9 = 2X^2/3$. In the circle, each player has 2 indirect connections of length 2, 2 indirect connections of length 3, 2 indirect connections of length 4 and 1 indirect connection of length 5, so the total value of indirect connections is given by

$$IC = \frac{X^2}{2} + \frac{X^3}{4} + \frac{X^4}{8} + \frac{X^5}{32}.$$

For small values of X, it is obvious that

$$IC < \frac{2X^2}{3}$$

so that the Petersen graph dominates the circle.

The symmetric cycle is always stable, as the following Proposition shows.

Proposition 5 *Let Assumption 2 hold. Then, the symmetric cycle is both Nash stable and strongly pairwise stable.*

Proof. : It is straightforward to check that the symmetric cycle is Nash stable. We just show that the symmetric cycle is strongly pairwise stable.

In the symmetric cycle, each i gets a direct benefit of X. No pattern of investment can result in higher direct benefits. So, we check whether a deviation by i and j can improve their indirect benefits.

Suppose i and j are neighbours in the cycle. Consider the effect on i of increasing investment to $\frac{X}{2}+y$ by both i and j on the link ij, and decreasing their investments on their other neighbours by y. The change in indirect benefit for i from j's other neighbour is $(\frac{X}{2}+y)(\frac{X}{2}-y)-(\frac{X}{2})^2 < 0$. A similar calculation shows that i also loses from nodes which are further away.

Suppose i and j ae not neighbours in the cycle. Let i and j mutually invest y each on the link ij and simultaneously decrease investment on their previous neighbours by $\frac{y}{2}$. It is easy to check that this is the best possible deviation.

Clearly, this can only increase indirect benefit for i if there is some k such that the distance between i and k is now lower. This means that i accesses k through j. Let k be a neighbour of j. Then, the indirect benefit for i from k is

$$I = y(\frac{X-y}{2}) = \frac{Xy}{2} - \frac{y^2}{2}$$

Now, i has reduced the strength of links with each of its previous neighbours by $\frac{y}{2}$. Also, since k is not at a distance of 2 from i in the cycle, there must be some node m, distinct from k which is at a distance of 2 from i. The loss in indirect benefit for i from m is

$$I' = (\frac{X-y}{2})\frac{X}{2} - \frac{X^2}{4} = \frac{Xy}{2}$$

Hence, the indirect beenfit for i from k is lower than the loss in indirect benefit from m.

Repeating this argument, it can be shown that i's total indirect benefit will actually go down as a result of the deviation. ■

4.2 Min Reliability

Efficient networks architectures are easily characterized with min reliability.

Proposition 6 *If Assumption 2 holds, the efficient graphs are the symmetric line and the cycle.*

Proof. Consider any node i, and suppose that j is a neighbour of i. The maximum indirect benefit that i can get from any node using a path involving ij is $\min(x_i^j, X - x_i^j)$, since $x_j^i + x_j^k \leq X$ for all $k \neq i, j$. Hence, for any node, the maximal indirect benefit from any other node is $\frac{X}{2}$, which is obtained by equalizing the value of every link at $X/2$. Any other architecture must involve some link of value smaller than $X/2$ and hence decrease the value. Note that the symmetric cycle yields every node a total benefit of $(n-1)\frac{X}{2}$. Hence, these must be efficient architectures.

Clearly, no node with degree greater than 2 can attain this value. Also, no disconnected graph can be efficient. Hence, the line and cycle must be the *only* efficient graphs. ■

Because distances do not matter with min reliability, in the symmetric line and cycle, each node derives a benefit of $\frac{X}{2}$ from every other node. It is trivial to check that no deviation by a pair (or coalition) can improve mutual payoffs. Hence, we have the following result.

Proposition 7 *Suppose Assumption 2 holds. Then, the symmetric line and the cycle are both Nash stable and strongly pairwise stable.*

5 Conclusion

In this paper, we analyze the formation of communication networks when players choose endogenously their investment on communication links. We consider two alternative definitions of network reliability; product reliability, where the decay of information depends on the product of the strength of communication links, and min reliability where the speed of connection is affected by the weakest communication link. When investments are separable, we show that the architecture of the efficient network depends crucially on the shape of the transformation function linking investments to the quality of communication links. With increasing marginal returns to investment, the efficient network is a star ; with decreasing marginal returns, the conflict between maximization of direct and indirect benefits prevents us from obtaining a complete characterization of efficient networks. However, we show that with min reliability, the efficient network must be a tree. Furthermore, in the particular case of linear transformation functions, we show that in an efficient network, all links must have equal strength. When investments are perfect complements, the results change drastically: under product reliability, the efficient network must contain a cycle, and is in fact a circle for small societies. With min reliability, the efficient network is either a circle or a line.

As in classical models of network formation, because of widespread externalities, efficient networks may not be supported by private investment decisions. We provide examples to show that the star may not be stable when the transformation functions is strictly convex. We also note that with perfect substitutes and perfect complements (when the efficient network displays a very symmetric structure), the efficient network can indeed be supported by private investments when the society is large.

In our view, this paper provides a first step in the study of networks where agents endogenously choose the quality of the links they form. An ambitious objective would be to revisit the recent literature on strategic network formation assuming that networks are represented by weighted graphs, but the complexity of the analysis in the simple case of communication networks indicates that a general study of weighted networks may be intractable. Instead, we would like to suggest three possible directions fo further research. First, we would like to extend the analysis to situations where agents face investment costs, rather than opportunity costs due to a fixed budget. Second, we plan to analyze network formation for other specifications of indirect benefits, assuming for example that information decays completely for paths of length greater than two. Finally, we need to understand better the relation between investments and network structures, and wish to study in more detail the pattern of link investments for fixed network architectures, and the comparative statics of investments with respect to changes in the network structure.

6 References

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

Bala, V. and S. Goyal (2000b) "A strategic analysis of network reliability," *Review of Economic Design* **5**, 205-228.

Bloch, F. and M. O. Jackson (2004) "The formation of networks with transfers among players", mimeo, GREQAM and Cal Tech.

Boorman, S. (1975) " A Combinatorial Optimisation Model for Transmission of Job Information through Contact Networks", *Bell Journal of Economics* **6**, 216-249.

BRUECKNER, J. (2003) "Friendship networks", mimeo., University of Illinois at Urbana Champaign.

CALVO-ARMENGOL, A. AND M.JACKSON (2004) "The Effects of Social Networks on Employment and Inequality" (2004) *American Economic Review* **94**, 426-454.

COLEMAN, J., E. KATZ AND H. MENZEL (1966) *Medical Innovation: A Diffusion Study*, Bobbs Hill, New York.

DURIEU, J., H. HALLER AND P. SOLAL (2004) "Nonspecific networking", mimeo., Virginia Tech.and University of Saint-Etienne.

DUTTA, B. AND S. MUTUSWAMI (1997) "Stable networks", *Journal of Economic Theory* **76**, 322-344.

GILLES, R. AND S. SARANGI (2004) "Social network formation with consent", mimeo., Virginia Tech and University of Louisiana.

GOYAL, S. (2004) "Strong and Weak Links", mimeo., University of Essex.

GOYAL, S. AND J. MORAGA GONZALES (2001) "R&D networks", *RAND Journal of Economics* **32**, 686-707.

GRANOVETTER, M. (1974) *Gretting a job: A Study of Contacts and Careers*, Harvard University Press, Cambridge.

HOLTON, D.A. AND J. SHEEHAN (1993) *The Petersen Graph*, Cambridge University Press.

JACKSON, M. AND A. WOLINSKY (1996) "A strategic model of economic and social networks," *Journal of Economic Theory* **71**, 44-74.

MONTGOMERY, I. (1991) "Social Networks and Labour Market Outcomes" (1991) *American Economic Review* **81**, 1408-1418.