Available online at www.sciencedirect.com
SCIENCE DIRECT®
Mathematical Social Sciences 50 (2005) 181–201
mathematical social sciences
www.elsevier.com/locate/econbase

Nash networks with heterogeneous links

Hans Haller[a,*], Sudipta Sarangi[b]

[a]Department of Economics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0316, USA
[b]Department of Economics, Louisiana State University, Baton Rouge, LA 70803-6306, USA

Received 1 May 2004; received in revised form 1 December 2004; accepted 1 February 2005
Available online 25 May 2005

Abstract

A non-cooperative model of network formation is developed. Link formation is one-sided. Information flow is two-way. The model builds on the work of Bala and Goyal who permit links to fail with a certain common probability. In our model the probability of failure can be different for different links. The set of networks which are Nash for suitably chosen model parameters consists of all essential networks. We specifically investigate Nash networks that are connected, super-connected, or stars. Efficiency, Pareto-optimality, and existence issues are discussed through examples. Three alternative model specifications are explored to address potential shortcomings.

Keywords: Network formation; Non-cooperative games

JEL Classification: C72; D85

1. Introduction

Social networks have played a vital role in the diffusion of information across society in settings as diverse as referral networks for jobs (Granovetter, 1974) and in assessing quality of products ranging from cars to computers (Rogers and Kincaid, 1981). Information in most societies can either be obtained in the market-place or through a non-market environment like a social network. For instance, in developed countries credit agencies provide credit ratings for borrowers, while in many developing countries credit

*Corresponding author. Tel.: +1 540 231 7591; fax: +1 540 231 5097.
E-mail address: haller@vt.edu (H. Haller).

0165-4896/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.mathsocsci.2005.02.003

worthiness is assessed through a social network organized along ethnic lines. The *Internet* provides ample testimony to the fact that information dissemination affects all aspects of economic activity. Nowadays fashions and fads emerging in one country are easily communicated almost instantaneously across the world. Financial troubles in one country now have devastating consequences for other economies as the contagion moves across boundaries with relative ease. Yet, the East Asian financial crisis also demonstrated that economies where information networks were relatively primitive remained largely insulated from the crisis. In other words both the structure and the technology of information dispersion are important determinants of its consequences.

In this paper we develop a non-cooperative model of network formation which is a generalization of Bala and Goyal (2000b).[1] The resulting networks serve as a mechanism for information transmission. Specifically, the structural aspects of information dissemination are modelled by means of a social network. The role of technology is studied by examining the reliability of the network. Agents in our model are endowed with some information which can be accessed by other agents forming links with them. Link formation is costly and links transmit information randomly. The cost of establishing a link is incurred only by the agent who initiates it, and the initiating agent has access to the other agent's information with a certain probability. In addition, he has access to the information from all the links of the other agent. Thus each link can generate substantial positive externalities of a non-rival nature in the network. Moreover, the flow of benefits through a link occurs both ways. It can differ across agents, since the strength of ties varies across agents (although all links cost the same) so that links fail with possibly different probabilities. This reflects the fact that in reality, communication often embodies a degree of costly uncertainty. We frequently have to ask someone to reiterate what they tell us, explain it again and even seek second opinions.

We introduce heterogeneity by allowing for the probability of link failure to differ across links.[2] This distinctive feature reflects the nature of the transmission technology or the quality of information. The generalization provides a richer model in terms of answering theoretical as well as practical questions: connectedness and super-connectedness,[3] selection of central agents in star networks, efficiency, and Pareto-optimality. Besides imparting greater realism to the model, the introduction of heterogeneous links allows us to check the robustness of the conclusions obtained in Bala and Goyal (2000b). Whereas their findings still hold under certain conditions, link heterogeneity gives rise to a much greater variety of equilibrium outcomes. The set of possible strict Nash equilibrium networks (in short strict Nash networks) consists of all essential networks, i.e. networks

[1] The non-cooperative or strategic approach to network formation constitutes one of three major strands of literature. For a comprehensive review of the literature on pairwise stability see Jackson (2005). For the literature related to cooperative games see Slikker and van den Nouweland (2001).

[2] For other forms of heterogeneity, see Johnson and Gilles (2000) and Galeotti et al. (in press).

[3] A network is connected if information can flow (directly or indirectly) between each pair of agents. A connected network is minimally connected if it ceases to be connected after the deletion of any link. Super-connected networks (in our weak sense) are the connected networks which are not minimally connected, that is they remain connected after deletion of a particular link. Super-connected networks in the strong sense of Bala and Goyal remain connected after the deletion of any link.

without duplication of direct links. It includes as proper subsets the set of possible strict Nash equilibrium outcomes of Bala and Goyal (2000b) and the set of possible strict Nash equilibrium outcomes in the model with general cost and value heterogeneity of Galeotti et al. (in press). This result obtains even when there are only two different failure probabilities, which shows that the nature of heterogeneity is as important as the degree of heterogeneity.

Bala and Goyal show for both their models that Nash networks must be either connected or empty. With heterogeneous links, this dichotomy need no longer hold. It still holds when the probabilities of success are not very different from each other. Another central finding of Bala and Goyal is that compared to information decay, imperfect reliability has very different effects on network formation. With information decay, minimally connected networks (notably the star) are Nash for a wide range of cost and decay parameters, independently of the size of society. In contrast, with imperfect reliability and small link costs, minimally connected networks tend to be replaced by super-connected networks (connected networks with redundant links) as the player set increases. However, with link heterogeneity neither connectedness nor super-connectedness need arise asymptotically. Still, sufficient conditions for asymptotic super-connectedness can be given. Furthermore, in order for star networks to be Nash, they have to satisfy additional conditions. Interestingly enough, heterogeneity helps resolve another ambiguity associated with the homogeneous model: Owing to the additional equilibrium conditions, the coordination problem inherent in selecting the center of a star is mitigated to a certain degree.

In addition, we find that Nash networks may be nested and Pareto-ranked. We also demonstrate by example that inefficient Nash networks can be Pareto-optimal. A further example shows that Nash networks do not always exist with non-uniform link success probabilities. Criticisms of the non-cooperative approach to network formation are addressed as well. We extend the model to allow for beneficial duplication of direct links and to analyze Nash networks with incomplete information. Finally, the implications of endogenous success probabilities for Nash networks are discussed and explored.

In Section 2, we introduce the basic setup of the paper. In Section 3, we present general results and observations on Nash networks. Alternative formulations of the model are considered in Section 4. Section 5 concludes.

2. The model

Let $N=\{1,\ldots,n\}$denote the set of agents with $n\geq 3$. For ordered pairs $(i,j)\in N\times N$, the shorthand notation ij is used and non-ordered pairs $\{i,j\}$ are denoted by $[ij]$. The symbol $\subset$ for set inclusion permits equality. Each agent in the model has some information of value to the other agents. An agent can access more information by forming links with other agents. Link formation occurs simultaneously and is costly.

Each agent's strategy is a vector $g_i=(g_{i1},\ldots,g_{ii-1}, g_{ii+1},\ldots,g_{in})$ where $i\in N$ and $g_{ij}\in\{0,1\}$ for each $j\in N\backslash\{i\}$. The value $g_{ij}=1$ means that agents i and j have a link initiated by i, whereas $g_{ij}=0$ means that i does not initiate the link. This does not preclude the possibility of agent j initiating a link with i. We shall only consider pure strategies. The set of all pure strategies of agent i is denoted by $\mathcal{G}_i$. Given that agent i can initiate a link with each of the remaining $n-1$ agents, the number of strategies available to agent i is $|\mathcal{G}_i| = 2^{n-1}$. The strategy space of all agents is given by $\mathcal{G} = \mathcal{G}_1 \times \cdots \times \mathcal{G}_n$. A strategy profile $g=(g_1,\ldots, g_n)$ can be represented as a **directed graph** or **network**. Notice that there is a one-to-one correspondence between the set of all directed networks with n vertices or nodes and the set of strategies $\mathcal{G}$. The link g_{ij} will be represented pictorially by an edge starting at j with the arrowhead pointing towards i to indicate that agent i has initiated the link and therefore incurs the cost of this link. Let $\mu_i^d(g_i)=|\{k \in N:g_{ik}=1\}|$ denote the number of links in g initiated by i which is used in the determination of i's costs.

Next we define the closure of g which is instrumental for computing benefits. Pictorially the closure of a network is equivalent to replacing each directed edge of g by a non-directed one.

Definition 1. The **closure** of g is a non-directed network denoted by $h=cl(g)$ and defined as $cl(g)=\{ij \in N \times N:i \neq j$, and $g_{ij}=1$ or $g_{ji}=1\}$.

Benefits. A link between agents i and j potentially allows for **two-way (symmetric) flow of information**. Therefore, the benefits from network g are derived from its closure $h=cl(g)$—which is given by $h_{ij}=h_{ji}=\max\{g_{ij},g_{ji}\}$. For agents $i \neq j$, the non-ordered pair $[ij]$ represents the simultaneous occurrence of ij and ji. If $h_{ij}=0$, then $[ij]$ does not permit any information flow. If $h_{ij}=1$, then $[ij]$ succeeds (allows information flow) with probability $p_{ij} \in (0,1)$ and fails (does not permit information flow) with probability $1-p_{ij}$, where p_{ij} is not necessarily equal to p_{ik} for $j \neq k$. It is assumed, however, that $p_{ij}=p_{ji}$. Let $P=[p_{ij}]$ denote the symmetric $n \times n$-matrix of exogenous success probabilities, where we set $p_{ii}=1, i \in N$, for the insignificant diagonal elements. Furthermore, the successes of direct links between different pairs of agents are assumed to be independent events. Thus, h may be regarded as a random network with possibly different probabilities of realization for different edges. We call a non-directed network h' a realization of h ($h' \subset h$) if it satisfies $h'_{ij} \leq h_{ij}$ for all i,j with $i \neq j$. The notation $[ij] \in h'$ signifies that the undirected link $[ij]$ belongs to h', that is $h'_{ij}=h'_{ji}=1$.

Definition 2. For $h' \subset h$, a **path** of length m from an agent i to j, where $i \neq j$, is a finite sequence $i_0,i_1,\ldots,i_m$ of pairwise distinct agents such that $i_0=i$, $i_m=j$, and $h'_{i_k i_{k+1}}=1$ for $k=0,\ldots,m-1$.

We say that player i **observes** player j in the realization h', if there exists a path from i to j in h'. Invoking the independence assumption, the probability of the network h' being realized, given h is

$$\lambda(h'|h) = \prod_{[ij] \in h'} p_{ij} \prod_{[ij] \notin h'} \left(1-p_{ij}\right).$$

Let $\mu_i(h')$ be the number of players that i observes in the realization h'. Each observed agent yields a benefit $V>0$ to agent i, and without loss of generality we set $V=1$.

Given a strategy profile g agent i's expected benefit from the random network h is given by the following benefit function $B_i(h)$:

$$B_i(h) = \sum_{h' \subset h} \lambda(h'|h)\mu_i(h').$$

The probability of network h' being realized is given by $\lambda(h'|h)$, in which case agent i gets access to the information of $\mu_i(h')$ agents in total. Summing up over all possible realizations $h' \subset h$ yields total benefits. Note that every agent's benefit function is non-decreasing with respect to the inclusion ordering for sets of undirected links. That is, the addition of any links never decreases the agent's benefit.

Payoffs. We assume that each link formed by agent i costs $c>0$. Agent i's expected payoff from the strategy profile g is

$$\Pi_i(g) = B_i(cl(g)) - \mu_i^d(g_i)c. \tag{1}$$

Given a network $g \in \mathcal{G}$, let g_{-i} denote the network that remains when all of agent i's links have been removed. Clearly $g=g_i \oplus g_{-i}$ where the symbol $\oplus$ indicates that g is formed by the union of links in g_i and g_{-i}.

Definition 3. A strategy g_i is a **best response** of agent i to g_{-i} if

$$\Pi_i(g_i \oplus g_{-i}) \geq \Pi_i\left(g_i' \oplus g_{-i}\right) \text{ for all } g_i' \in \mathcal{G}_i.$$

Let $\mathrm{BR}_i(g_{-i})$ denote the set of agent i's best responses to g_{-i}. A network $g=(g_1,\ldots,g_n)$ is said to be a **Nash network** if $g_i \in \mathrm{BR}_i(g_{-i})$ for each i, that is if g is a Nash equilibrium of the strategic game with normal form $\left(N, \left(\mathcal{G}_i\right)_{i\in N}, (\Pi_i)_{i\in N}\right)$. A strict Nash network is one where agents are playing strict best responses.

Agent i's benefit from the direct link ij to agent j is at most $p_{ij}(n-1)$. Set $p_0=p_0(c,n)=c(n-1)^{-1}$. If $p_{ij}<p_0$, it never benefits agent i to initiate a link to j, regardless of how reliably j is linked to other agents and, therefore, $g_{ij}=0$ in any Nash equilibrium g.

We now introduce some definitions of a more graph-theoretic nature. A network g is said to be **connected** if there is a path in $h=cl(g)$, between any two agents i and j. A connected network g is said to be **super-connected**, if there exist links after whose deletion the network is still connected.[4] A connected network g is **minimally connected** if it is no longer connected after the deletion of *any* link. A network g is called complete, if all links exist in $cl(g)$. A network with no links is called an **empty network**.

Definition 4. A set $C \subset N$ is called a **component** of g if there exists a path in $cl(g)$ between any two different agents i and j in C and there is no strict superset C' of C for which this holds true.

For each network g, the components of g form a partition of the player set (node set, vertex set) N into non-empty subsets. Each isolated point $i \in N$ in g, that is a player or node i with $g_{ij}=g_{ji}=0$ for all $j \neq i$, gives rise to a singleton component $\{i\}$. In particular, the components of the empty network are the sets $\{i\}, i \in N$. N is the only component of g if and only if g is connected.

If C is a component of the network g, we denote by g^C the network induced by g on the set of agents C, that is $g_{ij}^C=g_{ij}$ for $i,j \in C, i \neq j$. A network g is **minimal**, if g^C is minimally connected for every component C of g. Minimally connected networks are both connected and minimal.

[4] Bala and Goyal (2000b) call a network super-connected if it remains connected after the deletion of an arbitrary single link.

The commonly used welfare measure is defined as the sum of the payoffs of all the agents. Formally, let $W:\mathcal{G}\rightarrow\mathbb{R}$ be defined as

$$W(g) = \sum_{i=1}^{n} \Pi_i(g) \textit{ for } g\in\mathcal{G}.$$

Definition 5. A network g is **efficient** if $W(g)\geq W(g')$ for all $g'\in\mathcal{G}$.

Furthermore, a network g is **Pareto-optimal**, if there does not exist another network g' such that $\Pi_i(g')\geq\Pi_i(g)$ for all i and $\Pi_i(g')>\Pi_i(g)$ for some i. Obviously, every efficient network is Pareto-optimal. However, we will show that not every Pareto-optimal network is efficient. In fact, we provide an example of a Pareto-optimal Nash network which is inefficient, while the unique efficient network is not Nash.

We finally introduce the notion of an essential network. A network $g\in\mathcal{G}$ is **essential** if $g_{ij}=1$ implies $g_{ji}=0$. Note that if $c>0$ and $g\in\mathcal{G}$ is a Nash network or an efficient network, then it must be essential. This follows from the fact that the information flow is two-way and independent of which agent invests in forming the link, that is $h_{ij}=\max\{g_{ij},g_{ji}\}$. Minimal networks are also essential.

3. Nash networks

In this section we look at Nash networks, beginning with an attempt to delineate the scope of possible Nash network architectures when the parameters of the model can be freely chosen.

3.1. Scope of Nash network architectures

We argue in the introduction that link heterogeneity in our strategic model of network formation gives rise to a broad spectrum of possible Nash network architectures—which is substantiated in the following theorem:

Theorem 1. *Let g be an essential network. Then there exist a link cost $c>0$ and an array $P=[p_{ij}]$ of link success probabilities such that g is a strict Nash network in the corresponding network formation game.*

Proof. Let g be an essential network. Let $p=1/(4n)$, $c=p/3$, $q=c/n$. We are going to construct a symmetric $n\times n$-matrix $P=[p_{kj}]$ of success probabilities as follows. If $i\neq j$ and i and j are directly linked, i.e. $g_{ij}=1$ or $g_{ji}=1$, set $p_{ij}=p$. Otherwise, set $p_{ij}=q$. Now consider $i\neq j$. If $g_{ij}=0$, then either $g_{ji}=1$ or $g_{ji}=0$. In the first case, agent i would receive zero benefit but incur a positive cost when forming the link ij; hence $g_{ij}=0$ is the unique optimal choice for i given g_{-i}. In case $g_{ji}=0$, $p_{ij}=q=c/n<p_0(c,n)$; hence regardless of other links, not initiating the link ij is uniquely optimal for agent i. If $g_{ij}=1$, then by essentiality of g, $g_{ji}=0$. Further $p_{ij}=p$. Without the link ij, the probability that information flows between i and j via other links is at most $\sum_{k\neq i,j}p_{ik}\leq(n-2)p<n/(4n)<1/2$. Hence regardless of the other links in g, the benefit of player i from initiating the link ij is at least

$p_{ij}(1-1/2)=p/2$ which by exceeds $c=p/3$, the cost of the link. Therefore, initiating the link ij is uniquely optimal for player i. This shows that for each i, g_i is the unique best response against g_{-i}, that is g is a strict Nash network. □

Discussion: This result that "everything goes" except inessential networks, prompts two immediate questions. One question concerns the cause for the richness of Nash network architectures. Arguably, the many degrees of freedom in choosing model parameters could be responsible for the broad spectrum of possible Nash network architectures. The model parameters reside in $\mathbb{R}_+ \times (0,1)^{n(n-1)/2}$ so that there are $1+n(n-1)/2$ exogenous continuous parameters or degrees of freedom in choosing parameter constellations. However, this *prima facie* very compelling explanation proves less convincing upon closer inspection. First, the above proof shows that one can severely restrict the parameter choices and still reach the conclusion of Theorem 1. It suffices to pick two probabilities p and q plus a cost parameter c such that $0<q<c<p<1$. In addition, one has to make a discrete choice of which probabilities p_{ij} equal p, with the remaining ones equal to q. Hence parameters can be restricted to the union of $2^{n(n-1)/2}$ convex sets of dimension 3. Second, a comparison with the findings of Galeotti et al. (in press) also shows that counting degrees of freedom can be misleading. There are $2n(n-1)$ exogenous continuous parameters in their most general model with cost and value heterogeneity. The networks which can arise as strict Nash networks for some parameter constellation are precisely the minimal networks. Hence, whereas their model has more degrees of freedom than ours it supports fewer network architectures.

The other question concerns the impact of parameter restrictions on the scope of Nash network architectures. In the context of Galeotti et al. (in press), assuming value homogeneity while maintaining cost heterogeneity, which reduces the number of exogenous parameters to $n(n-1)$, does not affect the scope of strict Nash network architectures. Specific further reductions of the degrees of freedom can have drastic effects. In our context, a restriction to essentially 3 parameters does not alter the conclusion of Theorem 1, as we have argued in the last paragraph. But the restriction to two parameters, a cost parameter c and a probability p, as in Bala and Goyal (2000b), does have a significant impact: Nash networks are either empty or connected which rules out many essential network architectures. However, the conclusion that Nash networks are either empty or connected is not confined to the two-parameter case. Proposition 1 asserts that for the conclusion to hold, success probabilities can differ across links but cannot be too dispersed.[5] Finally, additional restrictions yield special connected Nash network architectures: Bala and Goyal and we identify sufficient conditions for the existence of specific star Nash networks. Bala and Goyal find that the condition $p(1-p^{n/2})>c$ is sufficient for the super-connectedness of Nash networks. The examples in Section 3.2 indicate how a variety of asymptotic behaviors can be generated and others can be ruled out by very specific assumptions on link success probabilities.

To conclude, the role of model parameters proves as important as their number. Minimal heterogeneity of links: the presence of weak links with success probability q and

[5] Asymptotically, as n tends to infinity, the hypothesis of Proposition 1 collapses to the assumption of homogeneous links.

strong links with success probability $p > q$: is enough to sustain a broader spectrum of strict Nash network architectures than the model of Bala and Goyal (2000b) on the one hand and the model of Galeotti et al. (in press) on the other hand.

3.2. Connectivity and super-connectivity

With homogeneous links, Nash networks are either connected or empty (Bala and Goyal, 2000b). With heterogeneous links, this dichotomy does not always hold.

Proposition 1. *If* $p_{ij} \geq \frac{1}{1+c/n^2} p_{mk}$ *for any* $i \neq j$ *and* $m \neq k$*, then every Nash network is either empty or connected.*

Proof. Consider a Nash network g. Suppose g is neither empty nor connected. Then there exist three agents i,j, and k such that i and j belong to one connected component of $cl(g)$, C_1 and k belongs to a different connected component of $cl(g)$,C_2. Then either $g_{ij}=1$ or $g_{ji}=1$, whereas $g_{mk}=g_{km}=0$ for all $m \in C_1$. Without loss of generality assume $g_{ij}=1$. Then the incremental benefit to i of having the direct link to j is $b_1 \geq c$. Let g' denote the network which one obtains, if in g all direct links with i as a vertex are severed. The incremental expected benefit to i of forming the link ij in g' is $b_2 \geq b_1 \geq c$ and can be written as $b_2 = p_{ij}(1+V_j)$ where V_j is j's expected benefit from all the links j has in addition to ij.

Now consider a link from k to j, given $g' \oplus g_{ij}$. This link is worth $b_3 = p_{kj}(p_{ij}+1+V_j)$ to k. A link from k to j, given g; is worth $b_4 \geq b_3$ to k. We claim that $b_3 > b_2$, i.e.

$$p_{kj} > p_{ij} \frac{1+V_j}{1+V_j+p_{ij}}$$

Since g is Nash and $g_{ij}=1$, we know $p_{ij} \geq p_0 > c/n$. By assumption, $p_{kj} \geq \frac{1}{1+c/n^2} p_{ij}$. Therefore,

$$p_{kj} > \frac{1}{1+p_{ij}/n} p_{ij} = p_{ij} \frac{1+n-1}{1+n-1+p_{ij}} \geq p_{ij} \frac{1+V_j}{1+V_j+P_{ij}}$$

where we use the fact that V_j is bounded above by $n-1$. This shows the claim that $b_4 \geq b_3 > b_2 \geq b_1 \geq c$. Initiating the link kj is better for k than not initiating It. This contradicts that g is Nash. Hence to the contrary, g has to be either empty or connected. □

This result says that if the probabilities are not too widely dispersed, then the empty versus connected dichotomy still holds. If, however, the probabilities are widely dispersed, then a host of possibilities can arise and a single dichotomous characterization is no longer adequate. Example 1 (see also Fig. 1) presents a Nash network which is neither empty nor connected. Theorem 1 determines the entire range of possibilities.

Asymptotic properties: Link heterogeneity gives rise to a host of novel asymptotic properties when the number of agents tends to infinity. Asymptotic (super-)connectivity means that for sufficiently large player sets all Nash networks are (super-)connected.

I. As noted, in Footnote 5, the hypothesis of Proposition 1 collapses to the assumption of homogeneous links as n tends to infinity.

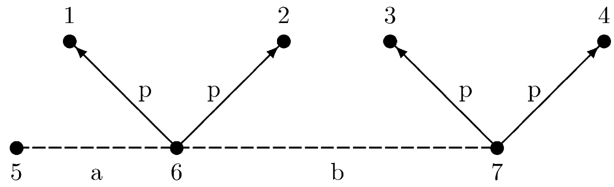

Fig. 1. Inefficient and Pareto Nash network.

II. Bala and Goyal (2000b) show that with homogeneous links and imperfect reliability, Nash networks tend to become super-connected as the size of the society increases. This result warrants several qualifications. The first one concerns an obvious trade-off even in the case of homogeneous links. While it is correct that for any given $p>c$, super-connectivity obtains asymptotically, the minimum number of players it takes to get connectivity goes to infinity as $p \rightarrow 0$. Let n^* be any number of agents. If $p<p_0(c,n^*)$, then it takes at least n^*+1 agents to obtain even a connected Nash network.

III. In our model asymptotic connectivity need no longer obtain, eliminating any scope for super-connectivity. Consider an infinite sequence of agents $i=1,2,\ldots,n,\ldots$ and a sequence of probabilities $p_2,p_3,\ldots$ such that $p_{ij}=p_{ji}=p_j$ for all $i<j$. Then the sequence $p_k,k\geq 2$, can be constructed in such a way that the empty network is the only Nash network for any agent set $I_n=\{1,\ldots,n\},n\geq 2$. Namely, let $c>0$ and set $p_k=p_0(c,k)/2$ for $k=2,3,\ldots$.

IV. With heterogeneous links, asymptotic super-connectivity still obtains, if there exist q_0 and q_1 such that $0<q_0\leq p_{ij}\leq q_1<1$ for all ij and the analogue of the sufficient condition $p(1-p^{n/2})>c$ in Bala and Goyal (2000b) holds. Inspecting the argument in the proof of Proposition 3.3 in Bala and Goyal (2000b, pp. 219–221) shows that if

$$\frac{1-q_1}{1-q_0}q_0\left(1-q_0^{n/2}\right)>c,$$

then every Nash network is super-connected in the strong sense of Bala and Goyal. Hence if $(1-q_1)q_0/(1-q_0)>c$, then for sufficiently large n, all Nash networks are super-connected. For example, $c=0.2$, $q_0=0.4$, and $q_1=0.6$ will do.

V. The lack of a common positive lower bound for the success probabilities does not necessarily rule out asymptotic super-connectivity if some success probabilities can be zero. A positive example is given by $c=3/4$, $p_{ij}=p_{ji}=p_j=j^{-1/2}$ for $i=1,2,j\geq 3$, and $p_{ij}=0$ otherwise. Then for $n>16$ and agent set I_n, the network g given by $g_{ij}=1$ if and only if $3\leq i\leq n$ and $j=1,2$, is strict Nash and super-connected. To show that g is strict Nash, observe that any link ij with $g_{ij}=0$ would have zero benefit, because it duplicates existing links or satisfies $p_{ij}=0$. Moreover,

$$\sum_{k=3}^{n-1}p_{1k}=\sum_{k=3}^{n-1}p_{2k}>\int_3^n s^{-1/2}ds=2\left[n^{1/2}-3^{1/2}\right]>n^{1/2}.$$

Consider $i\in\{3,\ldots,n\}$. Given g_{-i}, i receives benefits $\Delta_{i1}=p_{i1}(1-p_{i1}+\sum_{k=3}^{n}p_{1k})>i^{-1/2}n^{1/2}\geq 1$ from forming the single link i_1. Agent i receives additional benefits $\Delta_{i2}=(1-p_{i1})p_{i2}\left(1-p_{i2}+\sum_{k=3}^{n}p_{2k}\right)>(1-p_{i1})p_{i2}n^{1/2}$ from forming the extra link $i2$.

Specifically, $\Delta_{32}>\left(1-1/\sqrt{3}\right)/\sqrt{3}n^{1/2}>0.97$. For $3<i\le n, 1/2\ge i^{-1/2}>n^{-1/2}$. Hence $1/4\ge(1-p_{i1})p_{i2}>(1-p_{n1})p_{n2}$ and $\Delta_{i2}>(1-p_{n1})p_{n2}n^{1/2}=(1-n^{-1/2})n^{-1/2}n^{1/2}>3/4$. Hence for any $i\in\{3,\ldots,n\},\Delta_{i1}>c,\Delta_{i2}>c$, no links except $i1$ and ij are beneficial for i, and thus g_i, i.e. initiating both links $i1$ and $i2$, is the unique best response against g_{-i}. For $i=1,2$, given g_{-i}, any link initiated by i would have zero benefit to i. Hence g_i, i.e. not initiating any links, is the unique best response against g_{-i}. Therefore, g is a strict Nash network. To verify that g is super-connected, it suffices to demonstrate that each pair $i\neq j$ is connected via at least two disjoint paths in $h=cl(g)$. For instance $i=1,j=2$ are connected via the paths 1,3,2 and 1,4,2. Next $i=1,j\ge 3$ are connected via $1,j$ and $1,j+1,2,j$, for example. Also $i=2,j\ge 3$ are connected via $2,j$ and $2,j+1,1,j$, for example. Finally, $i>j\ge 3$ are connected via $i,1,j$ and $i,2,j$.

VI. Other possibilities exist as well. For instance, super-connectivity may be established at some point, but connectivity may break down when more agents are added, reemerging later, etc. Thus the Bala and Goyal (2000b) result is altered significantly in our model. To illustrate this possibility, consider $c>0$, $p\in(0,1)$ and an integer $\ell>2$ such that $p(1-p^{\ell/2})>c$. For each integer $j\ge 2$, there exists a unique integer $m(j)\ge 0$ such that $j\in[m(j)\ell+1,m(j)\ell+\ell]$. For two agents i and j with $i<j$, define $p_{ij}=p_{ji}=p$ if $m(j)$ is even and $p_{ij}=p_{ji}=p_0(c,j+1)$ if $m(j)$ is odd. Then for $n\ge 3$ and agent set I_n, all Nash networks are super-connected in case $m(n)$ is even and n is an isolated point in all Nash networks in case $m(n)$ is odd.

VII. Suppose we allow some success probabilities to be zero. Then asymptotically, several connected components can persist in each Nash network, with super-connectivity within each component. For a trivial example, let $1>p>c>0$ and set $p_{ij}=p_{ji}=p$ if $i\neq j,i+j$ even and $p_{ij}=p_{ji}=0$ if $i\neq j,i+j$ odd.

3.3. The polar cases

Let $P=[p_{ij}]$ denote the matrix of link success probabilities for all agents $(i,j)\in N\times N$, where $p_{ij}\in(0,1)$.

Proposition 2. *For any P, there exists c(P)>0 such that each essential complete network is (strict) Nash for all $c\in(0,c(P))$. The empty network is strict Nash for $c>\max\{p_{ij}\}$.*

Proof. Let $g=g_i\oplus g_{-i}$ be any essential complete network. Consider an arbitrary agent i with one or more links in his strategy g_i. Let $\mathcal{G}_i'=\left\{g_i'\in\mathcal{G}_i : g_{ij}'\le g_{ij}\ for\ all\ j\neq i\right\}$. Clearly, if $c=0$ then for agent i, g_i is a strict best response in $\mathcal{G}_i'$ against g_{-i}. By continuity, there exists $c_i(P,g_{-i})>0$ so that g_i is a strict best response in $\mathcal{G}_i'$ against g_{-i} for all $c\in(0,c_i(P,g_{-i}))$. Suppose $c\in(0,c_i(P,g_{-i}))$. If $g_i^*\in\mathcal{G}_i\setminus\mathcal{G}_i'$ then $g_{ij}^*=g_{ji}=1$ for some $j\neq i$ and there exists a better response $g_i'\in\mathcal{G}_i'$ without redundant costly links. Since g_i is at least as good a response as g_i', it is also a better response than g_i^*. Hence for $c\in(0,c_i(P,g_{-i}))$, g_i is a strict best response in $\mathcal{G}_i$ against g_{-i}. Now let $c(P)$ be the minimum of $c_i(P,g_{-i})$ over all conceivable combinations of i and g_{-i}. The first part of the claim follows from this.

For the second part, if $c>\max\{p_{ij}\}$ and no other agent forms a link, then it will not be worthwhile for agent i to form a link. Hence the empty network is strict Nash as asserted. □

3.4. Star networks

Star networks are among the most widely studied network architectures. They are characterized by one central agent through whom all the other players access each other. There are three possible types of star networks. In the inward pointing (center-sponsored) star the central agent establishes links to all other agents and incurs the cost of the entire network. An outward pointing (periphery-sponsored) star has a central agent with whom all the other $n-1$ players form links. A mixed star is a combination of the inward and outward pointing stars. Here we will focus on the periphery-sponsored star and the proofs provided below can be easily adapted to the other types of stars.

While the method of computing Nash networks does not change with the introduction of heterogeneous links, the process of identifying the parameter ranges for specific Nash network architectures is now more complicated. Without loss of generality assume that player n is the central agent in the star. Define M to be the set of all the agents except n or $M=N\backslash\{n\}$ and let $K_m=M\backslash\{m\}$ be the set M without agent m. Also let $J_{m\wedge k}=K_m\backslash\{k\}$ denote a set K_m without agent k and $\sum_{m\wedge k}=\sum_{j\in J_{m\wedge k}}p_{jn}$.We shall use the following *star condition*:

For all $m\in M,k\in K_m$: Either $p_{mn}>p_{mk}$ or $[p_{mn}<p_{mk},p_{mn}>p_{mk}p_{kn}$, and $(p_{mn}-p_{mk}p_{kn})\sum_{m\wedge k}>(p_{mk}-p_{mn})+p_{kn}(p_{mk}-p_{mn})]$.

Proposition 3. *Given $c\in(0,1)$; there exists a threshold probability $\delta\in(0,1)$ such that the outward pointing star is Nash if :*

1. *$p_{ij}\in(\delta,1)$ for all pairs ij.*
2. *The star condition holds.*

Proof. Consider the outward pointing star with n as the central agent. Choose the threshold probability $\delta\in(c,1)$ to satisfy the inequality

$$\max_{m\in M}\left[(1-p_{nm})+\left((n-2)-p_{nm}\sum_{k\in K_m}p_{nk}\right)\right]<c \tag{2}$$

if $p_{ij}\in(\delta,1)$ for all ij. Next we know that n has no links to sever, and does not want to add a link since $g_{mn}=1$ for all $m\in M$ and the flow of benefits is two-way. Now consider an agent $m\neq n$ who might wish to sever the link with n and instead link with some other $k\in K_m$. Player m's payoff from the outward pointing star is $\Pi_m(g^{ot})=p_{mn}+p_{mn}\sum_{k\in K_m}p_{kn}-c$. His payoff from deviating is $\Pi_m(g^{ot}-g_{mn}+g_{mk})=p_{mk}+p_{mk}p_{kn}+p_{mk}p_{kn}\sum_{m\wedge k}-c$. We get $\Pi_m(g^{ot})-\Pi_m(g^{ot}-g_{mn}+g_{mk})=(p_{mn}-p_{mk})+p_{kn}(p_{mn}-p_{mk})+(p_{mn}-p_{mk}p_{kn})\sum_{m\wedge k}$. This is clearly positive when $p_{mn}>p_{mk}$ for all $m\in M$.

However, when the inequality is reversed, we need the second part of the star condition. Essentially, the difference between the benefits from accessing agents $j\in J_{m\wedge k}$ through n instead of the indirect link through k, in this case, should exceed net benefits from agents n and k when agent m establishes a link with k instead of the central agent. Note that player m can only sever one link in an outward pointing star and hence we need not consider any more instances of link substitution by player m.

Next we need to check that no $m \in M$ wants to form a link with any $k \in K_m$. Note that payoffs with this additional link are bounded above by $(n-1)-2c$. Taking the difference between $\Pi_m(g^{ot}+g_{mk})$ and $\Pi_m(g^{ot})$ we get $[(1-p_{mn})+(1-p_{1n}p_{mn})+\ldots+(1-p_{m-1n}p_{mn})+(1-p_{m+1n}p_{mn})+\ldots+(1-p_{n-1n}p_{mn})]<c$ as the condition that the additional link is lowering m's payoff. Verifying that this is satisfied for all $m \in M$, gives us: $\max_{m \in M} [(1-p_{mn}) + (1-p_{1n}p_{mn}) + \ldots + (1-p_{m-1n}p_{mn}) + (1-p_{m+1n}p_{mn})+\ldots+ (1-p_{n-1n}p_{mn})] <$ c which is equivalent to (2). Since we use the upper bound on the payoffs to show that it is not worthwhile to add even one extra link by any player $m \in M$, this obviates the need to check that a player may want to add more than one link. □

The introduction of heterogeneous links alters things significantly. While part of the difference involves more complex conditions for establishing any star network, heterogeneity comes with its own reward. A different probability for the success of each link resolves the coordination problem implicit in the Bala and Goyal framework. Under homogeneous reliability, the role of the central agent in a star (Nash) network can be assigned to any player. With heterogeneous links, however, there are some natural candidates for the central agent position as illustrated by the following example. Let $p_{nm}=1$ for all $m \in M$ and $p_{jk}=q<1$ for all $jk \in M \times M$. Also let $c \in (0,1)$ be the cost of a single link. First we show that an outward pointing star with agent n at the center is Nash. We know that agent n does not wish to make any links and cannot break any links. Also $\Pi_m(g^{ot})=(n-1)-c$. Agent m will not wish to break her link to the center since this will give a payoff of zero. Suppose she considers breaking her link to the center and establishing a link with some peripheral agent k. Then her payoffs are $\Pi_m(g^{ot}-g_{mn}+g_{mk})=(n-1)q-c$ which is always lower than the Nash payoff. Finally suppose a peripheral agent considers adding an extra link. Her benefits in the network $g^{ot}+g_{mk}$ are still bounded above by $(n-1)$. Since $\Pi_m(g^{ot}+g_{mk})=(n-1)-2c<\Pi_m(g^{ot})=(n-1)-c, g^{ot}$ is Nash. Next consider an outward pointing star where the central position is occupied by some agent $m \neq n$. It is easy to check that every non-central agent $k \neq m,n$ is better off by breaking their link to m and connecting to n. Hence such an outward pointing star can never be Nash. This makes agent n a natural candidate for the central agent position.

Notice also that the determination of δ involves the different probabilities of all other links, making it quite complicated. Further, the benefits from deviation are also altered now. In the Bala and Goyal framework, no agent in the outward pointing star will ever deviate by severing a link with the central agent. In our model, links to the central agent will be severed unless the probabilities in the relevant range satisfy some additional conditions.

In our framework the inward pointing star is Nash in the above specified range of costs if the central agent's worst link yields higher benefits than c and (**2**) is satisfied. The mixed star can be supported as Nash when conditions required by the inward and the outward pointing star are satisfied for the relevant agents.

Next let us consider the case where $c > 1$. Here $c > p_{ij}$ for all links ij.

Proposition 4. *Given $c \in (1, n-1)$ there exists a threshold probability $\delta < 1$ such that for $p_{ij} \in (\delta, 1)$ the outward pointing star is Nash.*

Proof. Let agent n be the center with whom all the other players establish links. Since $c \in (1, n-1)$ we can choose $\delta \in (0,1)$ such that if $p_{ij} \in (\delta, 1)$ for all ij, then (**2**) holds and $\min_{m \in M}\left[p_{mn}\left(1 + \sum_{k \in K_m} p_{kn}\right)\right] > c$. Then no $m \in M$ wants to sever his link with n. The remainder of the proof is similar to the proof of Proposition 3. □

Once again it is possible to identify a natural candidate for the role of the central player. The argument for the case $c \in (0,1)$ holds for this case, too. Also note that the inward pointing and mixed star will never be Nash within this range of costs.

3.5. Efficiency issues

Efficiency is a key issue in Jackson and Wolinsky (1996), Bala and Goyal (2000a,b), and Johnson and Gilles (2000). The thrust of these studies has been on the conflict between stability and efficiency. When costs are very high or very low, or when links are highly reliable, there is virtually no conflict between pairwise stability and efficiency in the symmetric connections model of Jackson and Wolinsky (1996) and between Nash networks and efficiency in the Bala and Goyal (2000b) framework. This observation still holds in our context. However, there is a conflict between Nash networks and efficiency for intermediate ranges of costs and link reliability, even with the same probability of link failure for all links. In particular, Nash networks may be under-connected relative to the social optimum as the subsequent example shows.

With the exception of Jackson (2005), the literature on networks has neglected Pareto-efficient or Pareto-optimal networks, focusing instead on efficiency. Our above definition of Pareto-optimality is standard and based on the individual payoffs $\Pi_i(g), i \in N$ achievable in the game, whereas efficiency is based on the sum of the payoffs across players, $W(g)$. More generally, in the setting of Jackson and Wolinsky (1996), Jackson (2005), and the related literature, Pareto-optimality is based on the individual payoffs $Y_i(g,v), i \in N$, determined by the allocation rule $Y = (Y_1, \ldots, Y_n)$ and the value function v while efficiency is defined in terms of the value $v(g)$. Jackson (2005) reports that in a number of popular models with component-balanced and anonymous allocation rules, including some variants of the symmetric connections model with decay, some or all pairwise stable networks are not only inefficient but also Parete-inefficient. He finds that for some combinations of value functions and component-balanced and anonymous allocation rules, efficient networks are a proper subset of Pareto efficient networks: Namely, there exist pairwise stable and Pareto-efficient networks and no pairwise stable and efficient networks. We complement his study by examining Pareto-efficiency (Pareto-optimality) in the context of strategic reliability.

We add two important observations not made before. First, it is possible that Nash networks are nested and Pareto-ranked. Second, at least in our context, the following can coexist: a Nash network which is not efficient, but Pareto-optimal and a unique efficient network which is not Nash and does not weakly Pareto-dominate the Nash network. The first observation is supported by the following example: $c=1$, $n=4$ and $p_{ij}=0.51$ for all ij. In this case, both the empty network and the outward pointing star with center 4 are Nash networks. The "outward pointing star" consisting of the links 14, 24 and 34 contains and strictly Pareto-dominates the empty network. Moreover, the empty network is under-connected. Our second observation is based on the following example, depicted in Fig. 1.

Example 1. $c=1$, $n=7$. $p_{16}=p_{26}=p_{37}=p_{47}=p=0.6181$, $p_{56}=a=0.2$, $p_{67}=b=0.3$, and corresponding probabilities for the symmetric links. All other links have probabilities $p_{ij}<p_0$. Now g given by $g_{16}=g_{26}=g_{37}=g_{47}=1$ and $g_{ij}=0$ otherwise is a Nash network. Indeed, p is barely large enough to make this a Nash network. The critical value for p satisfies $p(1+p)=1$ with solution $0.6180\ldots$ But g is not efficient. Linking also 5 with 6 and 6 with 7 provides added benefits which total 4.16.

Hence the total benefits exceed the cost of establishing these two additional links and the resulting network $g'=g\oplus 1_{56}\oplus 1_{67}$ would be efficient, where 1_{ij} denotes the network which has the single link ij. But neither 6 nor 7 benefits enough from the additional link between them to cover the cost of the link. Thus, the enlarged efficient network is not Nash. Since one of the two agents must incur the entire cost of the new link, the enlarged efficient network cannot weakly Pareto-dominate g. In fact, g is Pareto-optimal while inefficient. Reconciling efficiency and Pareto-optimality would require the possibility of cost sharing and side payments. Let us consider two such possibilities.

First, suppose that the cost of a link ij is equally split between i and j regardless of who initiates the link. Then g' becomes a Nash network, whereas the corresponding undirected graph is not pairwise stable in the sense of Jackson and Wolinsky (1996); the pairwise stable network is the undirected graph corresponding to $g\oplus 1_{67}$. Second, suppose that the cost of a link–but not its benefits–is equally split among all agents. Under this cost allocation rule, g' is a Nash network and the corresponding undirected graph is pairwise stable. Since $c=1$, and given g, the benefit from link 67 is at least 0.67 for both 6 and 7, the link ij or $[ij]$, exists in a Nash network or the pairwise stable network, respectively, under both cost allocation rules. The reason why the cost allocation rule matters is that given $g\oplus 1_{67}$, the link 56 or [56] creates a benefit of 0.58 for agent 5 and a combined benefit of 0.78 for agents 5 and 6. It creates a social benefit of 1.16 so that $g\oplus 1_{67}$ is more efficient than g, but still inefficient.

3.6. Existence of Nash equilibria

The existence of Nash networks has not been systematically explored in the literature, in contrast to the work on pairwise stability. Jackson and Watts (2001) give necessary as well as sufficient conditions for the existence of pairwise stable networks. Jackson and Watts (2002) provide an example for non-existence of pairwise stable networks. Jackson (2005) shows existence of such networks for several prominent allocation rules. Bala and Goyal (2000a) outline a constructive proof of the existence of Nash networks under perfect reliability. In addition, the literature contains assertions that for certain parameter ranges, the model admits Nash networks (or pairwise stable networks, respectively) with specific properties. If the various regions happen to cover the entire parameter space, then as a by-product, existence has been shown for the particular model. But if the various regions do not cover the entire parameter space, existence of Nash networks remains an open question for some parameter constellations. We show that when links have different success probabilities, a Nash network may not exist, i.e. link heterogeneity can lead to non-existence of Nash equilibria.[6]

Example 2. Let there be a total of 83 agents labelled $i=0, 1, 2, 3, 301,\ldots,363, 4, 401,\ldots,$ 415. Set $p=p_{12}=p_{21}=0.4$; $q=p_{23}=p_{32}=0.01473$; $r=p_{34}=p_{43}=1/32$; $s=p_{14}=p_{41}=1/16$; and $t=p_{20}=p_{02}=1$. Further put $p_{3j}=p_{j3}=1$ for $j=301,\ldots,363$; $p_{4j}=p_{j4}=1$ for $j=401,\ldots,415$; and $p_{ij}=p_{ji}=0$ for all remaining ij. Finally, choose $c=0.95$. Then the following links will always be established: 02 or 20; $3j$ or $j3$ for $j=301,\ldots,363$ and $4j$ or $j4$ for $j=401,\ldots,415$. Obviously, none of the links ij with $p_{ij}=0$ will be established. Moreover, 1 will always establish the link 14, 4 will always establish the link 43, 2 will never establish the link 21 and 3 will never establish the link 32. Now the existence of a Nash network can be decided by assessing the benefits from links 12 and 23 to players 1 and 2, respectively, given that all other links have been established or not according to our foregoing account. We obtain:

- Without 23, player 1 strictly prefers not to establish 12.
- With 23, player 1 strictly prefers to establish 12.
- Without 12, the benefit to player 2 from link 23 is 0.95011 and establishing 23 is a strict best response.
- With 12, player 2's benefit from link 23 is reduced by $81pqrs=0.00093$ (due to redundancies) and not establishing 23 is a strict best response.

Hence there are no mutual best responses regarding establishment of 12 and 23. Consequently, a Nash network does not exist.

To understand why the particular choice of q has player 2 switch back and forth, replace q by a $\tilde{q}$ such that without 12, player 2 is indifferent between having and not having the link 23, i.e. $\tilde{q}\ (64+r16+rs)=c$. This yields =0.014728236. Then with 12, player 2 would not want the link because of redundancies. By continuity, q slightly larger than $\tilde{q}$ produces the best response properties exhibited above.

4. Alternative model specifications

In this section we will consider three alternative specifications of our current model. The first variation introduces greater realism in the formation of networks by allowing agents to duplicate existing links, thereby raising link success probabilities. The second specification considers network formation under incomplete information. Finally, Nash networks with endogenous link success probabilities are examined. In Haller and Sarangi (2003), we discuss a fourth alternative where link formation requires the consent of the other agent; see also Gilles and Sarangi (2003).

[6] So far the literature on strategic network formation has not considered equilibria in mixed strategies—which would overcome the existence problem in finite games.

4.1. An alternative formulation of network reliability

The payoff function in the previous section and in subsequent subsections assumes that if $h_{ij}=h_{ji}=1$, then $[ij]$ succeeds (allows two-way information flow) with probability $p_{ij}\in(0, 1)$, regardless whether i, j or both initiated a link. Here we contend that if both agents involved in a relationship form a link each, it will succeed with a higher probability. A double link, i.e. $g_{ij}=1$ and $g_{ji}=1$, opens up a second independent two-way connection between agents i and j. Hence a direct link between i and j is now in effect with probability $r_{ij}=1-(1-p_{ij})^2$ if $g_{ij}+g_{ji}=2$ and with probability p_{ij} if $g_{ij}+g_{ji}=1$. We retain all other assumptions of the previous formulation. The next example demonstrates the different possibilities that can arise in this model.

Example 3. Let $n=3$. First consider the following probabilities and costs: $p_{12}=0.8$, $p_{23}=0.9$, $p_{13}=0$ and $c=0.7$. Let $g^1\equiv(g_{21}=g_{23}=1)$. It is easy to verify that g^1 is Nash since player 2 does not wish to break any links. If we now allow for double links, g^1 continues to be Nash. Moreover, in any network involving double links, one of the agents contributing to a double link is better off breaking it. Next consider the case where $p_{12}=0.5$, $p_{23}=0.3$, $p_{13}=0$ and $c=0.2$. Once again when no double links are permitted the network g^1 is Nash. But in contrast to the previous case, when double links are allowed, the equilibrium is given by $g_{12}=g_{21}=1=g_{23}=g_{32}$. Finally, consider the case where $p_{12}=0.5$, $p_{23}=0.3$, $p_{13}=0$ and $c=0.22$. Again, g^1 is still Nash under the single link model. Once we allow for double links however, the Nash network is given by $g_{12}=g_{21}=g_{32}=1$.

The incremental reliability of a double link $r_{ij}-p_{ij}=p_{ij}-p_{ij}^2$ is maximized at $p_{ij}=0.5$. Hence as shown in Example 3 double links are most likely to occur when the probabilities are medium range. In the third numerical specification, creation of the link 32 is beneficial for 3, but reduces the benefit of the prior link 23 for 2 from $p_{23}=0.3$ to 0.21 which is less than its cost. Hence in response to the establishment of link 32, the prior link 23 is dropped. From the example, it should also be evident that this formulation can lead to super-connected networks of a different sort—where agents may reinforce existing higher probability links instead of creating new links.

Observe that when double links are possible some new terminology is warranted. For instance, using our old definitions it is impossible to describe the third equilibrium architecture in Example 3. We define a **degree two star** as a star network where peripheral agents and the central agent have double links. Clearly, the definition of a mixed star also takes a new meaning now, since a mixed star can include both double links and single links that are initiated by either the central or a peripheral agent. Similarly, we can define a **degree two connected network** as a connected network where all links are double links. In a mixed connected network, both double and single links exist. Also, the **degree two complete network** is the complete network with double links. Next we show that a version of Proposition 1 holds under this alternative formulation.

Remark 1 (*Proposition 1′*). If $p_{ij}\geq\frac{1}{1+c/n^2}r_{mk}$ for any $i\neq j$ and $m\neq k$, then every Nash network is either empty or connected.

The proof resembles the previous proof, while allowing for the possibility that there is a double link between i and j. The condition on success probabilities implies for uniform success probability p that $p \geq 1 - c/n^2$ which in turn implies $p - p^2 < p_0(n, c)$. This means that double links are not worth initiating and, consequently, connected Nash networks are essential although double links are allowed. In contrast, Example 3 illustrates that for specific parameter constellations, Nash networks are connected with some or all links duplicated. Further, a version of Proposition 2 continues to hold. Let again $P = [p_{ij}]$ denote the matrix of link success probabilities for all agents $(i, j) \in N \times N$ where $p_{ij} \in (0, 1)$.

Remark 2 (*Proposition 2′*). For any P, there exists $c(P) > 0$ such that the degree two complete network is (strict) Nash for all $c \in (0, c(P))$. The empty network is strict Nash for $c > \max\{p_{ij}\}$.

The proof is very similar to the proof of Proposition 2. For $n = 3$ and a uniform success probability $p \in [1/4, 1)$, one can show that if the degree two complete network is Nash, then the essential complete networks are not Nash. For instance, for $c = 0.25$ and $p = 0.71727$, the wheel with links 12, 23, 31 is Nash and, consequently, the degree two complete network cannot be Nash. In general, it is an open question whether other complete Nash networks can coexist with the degree two complete network. Example 3 offers three cases where the only Nash network is connected but not complete, where all, some, or none of the links are double links.

The consequences of effective double links are now explored by reexamining Proposition 3. The main impact of the double link model is on the threshold probability value δ, altering the range of costs and probabilities under which the outward pointing star can be supported as Nash. The payoff function used earlier for determining the payoff from an additional link assumed that payoffs have an upper bound of $n - 1 - \alpha c$ where α denotes the number of links formed. By computing the precise value of the payoffs from additional links instead of using the upper bound, we show how things change in the current model. We find that the resulting new threshold $\tilde{\delta} = \max_{m \in M} \max_{i \neq m} \delta_i^m$ will suffice. Each δ_i^m is a threshold value related to the specific link mi.

Proposition 5. *Suppose that the links $g_{ij} = 1$ and $g_{ji} = 1$ are independent, and $c \in (0, 1)$. Then there exists a threshold probability $\tilde{\delta} \in (0, 1)$ such that the outward pointing star is Nash if:*

1. *$p_{ij} \in (\tilde{\delta}, 1)$ for all pairs ij.*
2. *The star condition holds.*

Proof. See Appendix of Haller and Sarangi (2003). □

In our previous formulation, δ can also be obtained as the maximum of link-specific thresholds. The latter tends to be lower if duplication has no benefits. Thus, in general $\tilde{\delta} > \delta$. To see why this is the case let $p_{ij} = p$ for all $ij \in N \times N$. Then $p > \delta$ is equivalent to (i) $(1 - p) + (1 - (n - 2)p^2) < c$. On the other hand, $p > \tilde{\delta}$ is equivalent to (ii) $p(1 - p^2) + (n - 2)(1 - p)p^2 < c$ plus (iii) $p(1 - p) + (1 - p)(n - 2)p^2 < c$. Now $p = 1 - c$ satisfies (i) if $n \geq 2 + 1/(1 - c)^2$, but violates (ii) and (iii) for sufficiently large n. Hence given c, one obtains $\tilde{\delta} > \delta$ for sufficiently large n. Such is the case for $c = 1/2$, $n = 10$. Moreover, based on the example provided above, it is also obvious that a whole range of mixed stars may arise as equilibria.

4.2. Nash networks under incomplete information

In our incomplete information formulation each agent $i \in N$ has knowledge of the probability of success of all her direct links. However, she is not aware of the probability of success of indirect links, i.e. agent i knows the value of p_{ij}, but is unaware of the value of p_{jk}, where $i \neq j$, k. The assumption that $p_{ij} = p_{ji}$ is still retained.

In order to solve for the equilibrium networks, each agent i must now have some beliefs about indirect links. We assume each agent postulates that, on average, every other agent's world is identical to her own. Thus, agent i assigns a value of $p_i = \frac{1}{n-1} \sum_{i \neq m} p_{im}$ to all indirect links p_{jk} for $i \neq j$, k. This has some immediate consequences for the payoff function and star networks. Consider some agent $m \in M$. She now believes that her payoff from the outward pointing star is given by $\Pi_m(g^{ot}) = p_{nm} + |K_m| p_{nm} p_m - c$ which is clearly different from her actual payoff. We obtain the *modified star condition* by replacing each term p_{nk} with p_m. Let δ be the probability threshold of Proposition 3.

Proposition 6. *Given each agent's beliefs about her indirect links, the outward pointing star is Nash if:*

1. $p_{ij} \in (\delta, 1)$ *for all pairs ij.*
2. *The modified star condition holds.*
3. $(n-2)(1-p_{nm}p_m) < ((n-2) - p_{nm} \sum_{k \in K_m} p_{nk})$ *for all* $m \in M$.

Proof. The proof is similar to that of Proposition 3 and hence omitted. □

This formulation provides us with some interesting insights about the role of the indirect links and the vulnerability of Nash networks. Assume that the actual probabilities satisfy the hypothesis of Proposition 3. It is possible that $1 - p_{nm} + (n-2)(1 - p_{nm}p_m) > c > 1 - p_{nm} + ((n-2) - p_{nm} \sum_{k \in K_m p_{nk}})$, in which case agents will create new links destroying the star architecture. Consequently, the realized network yields lower payoffs than the star network. Thus, the introduction of incomplete information can easily lead to network failure, in the sense that the outcome is less efficient than it would be otherwise. We now examine additional consequences of this formulation through an explicit example.

Example 4. Consider a network with $n = 6$. Suppose agents 1 to 4 are linked in a star formation with agent 4 being the central agent, i.e. $g_{14} = g_{24} = g_{34} = 1$. Further $g_{56} = 1$ and we will examine what happens to the link g_{45} under complete and incomplete information. Let $c = 1/12$, $p_{14} = p_{24} = p_{34} = p = 4/10$, $p_{56} = r = 1/2$ and $p_{45} = q = 1/24$. The probabilities of all other links are assumed to be zero.

Under these objective probabilities $q(1+r) < c$ and hence agent 4 will never initiate the link with agent 5. However, agent 5 will initiate this link since $q(1+3p) > c$. The resulting connected network is Nash since all other links yield no benefits. Note that for our current formulation with incomplete information, $p_5 = \bar{r} = 1/5(r+q)$ and $p_4 = \bar{p} = 1/5(3p+q)$. Under these beliefs about the probabilities of the indirect links, agent 4 will never establish the link since $q(1+\bar{r}) < c$. Similarly, agent 5 will not establish the link since $q(1+3\bar{p}) < c$. With incomplete information the above disconnected network with $g_{45} = 0$ is a Nash network.

Thus incomplete information may destroy a crucial link and give rise to two connected components.

In this section we have demonstrated that incomplete information as modelled here can either lead to new links yielding lower payoffs or destroy crucial links in the network. There are other interesting alternatives to introduce incomplete information into a strategic model of network formation. Specifically, McBride (2002) considers a model with fully reliable links, but with incomplete information about the existence of certain indirect links and incomplete information about the benefits which accrue to a player via each of her direct links.

4.3. Endogenous link probabilities

We now consider the possibility that the addition of a link renders all adjacent links less reliable. In other words, any given node might become less effective in providing information via its direct links, if it gets accessed through one more direct link. Consider the following example.

Example 5. *Negative link externalities.* Let $c=0$. For $i\in N$ and $g\in\mathcal{G}$, set $n_i(g)=|\{k\in N\backslash\{i\}:g_{ik}=1 \text{ or } g_{ki}=1\}|$ as the number of agents to whom i has direct links in g. For any two agents i and j and any network g, the endogenous probability of success of link ij is given as

$$p_{ij}(g)=\begin{cases}\dfrac{1}{n_i(g)}\cdot\dfrac{1}{n_j(g)}, & \text{if } g_{ij}+g_{ji}>0;\\ 0, & \text{if } g_{ij}+g_{ji}=0.\end{cases}$$

First consider the case $n=3$. In an essential complete network with $p=1/4$, each player i receives payoff $\Pi_i(g)=19/32$. Suppose that i is one of the players who have initiated a link ij. After severance of that link, the two remaining links have each success probability 1/2 and i's payoff becomes 3/4. Thus with endogenous success probabilities and zero or negligible costs, complete networks need no longer be Nash—in stark contrast to Proposition 2. Moreover, for $n\geq 4$, wheels with simple links, line networks with simple links, and stars are not Nash under the current assumptions. Finally, it also turns out that a network g is Nash if each component C either satisfies $|C|=3$ and is incomplete (not a wheel) or satisfies $|C|=2$.

Next let us consider a more general model of endogenous success probabilities encompassing capacity constraints where an agent i cannot have more than L_i links. We assume $P_{ij}=P_{ji}\in(0,\ 1]$ and a non-increasing function $\alpha_i:\mathbb{N}\to[0,1]$ $\alpha_i(1)=1$ for each agent. Then

$$p_{ij}(g)=\begin{cases}\alpha_i(n_i(g))\alpha_j\big(n_j(g)\big)P_{ij}, & \text{if } g_{ij}+g_{ji}>0;\\ 0, & \text{if } g_{ij}+g_{ji}=0.\end{cases}$$

In Example 5, $\alpha_i(z)=1/z$. Now set $\alpha_i(z)=1$ for $z\leq L_i$ and $\alpha_i(z)=0$ for $z>L_i$. Then clearly no star network can be Nash if $(n-1)\alpha_i(n-1)<c^{1/2}$ for all i. Next suppose $c\in(0,$ $1)$ and $P_{ij}=p\in(0,\ 1)$. The existence of star Nash networks now depends on the severity of

the capacity constraints, while the subsequent example illustrates that novel and interesting Nash networks may also arise due to negative externalities.

Proposition 7. *There exists $p(c)\in(c, 1)$ such that for $p\in(p(c), 1)$: (i) If for all i, $L_i<n-1$, then no star network is Nash, and (ii) If for some i, $L_i\geq n-1$, then all stars with center i are Nash.*

Proof. By Proposition 3.2(b) of Bala and Goyal (2000b), there exists $p(c)\in(c, 1)$ such that for $p\in(p(c), 1)$, all star networks are Nash in the absence of capacity constraints. If $L_i<n-1$ and i is the center of a star, then $p_{ij}=0$ for all $j\neq i$ and an agent is better off severing a link to or from i. Hence (i). If $L_i\geq n-1$ and i is the center of a star, then $p_{ij}=p$ for all $j\neq i$ and the star remains Nash after the imposition of the capacity constraints. Hence (ii). □

Example 6. Let $n=4$, $c=0.35$, $P_{1j}=P_{j1}=0.8$ for $j\neq 1$, and $P_{ij}=0.5$ for all other links. With exogenous probabilities $p_{ij}=P_{ij}$, all stars with center 1 are Nash, whereas the wheel with links 12, 23, 34, 41 is not Nash because 1 gains from establishing the extra diagonal link 13. Now assume endogenous probabilities and $\alpha_i(1)=\alpha_i(2)=1$ for all i. Then for sufficiently small values of $\alpha_i(3)$ and $\alpha_i(4)$, the stars cease to be Nash and the above wheel becomes Nash.

The approach taken here has provided several immediate and important insights. The asymptotic behavior of the network (as n increases) crucially depends, among things, on the properties of the functions $(n-1)\alpha_i(n-1)$. While super-connectivity may or may not occur connected multi-hub systems can also emerge. The present approach can be generalized in several ways. The multiplicative form $\alpha_i\alpha_j$ incorporates both some degree of substitutability and some degree of complementarity between nodes. The additive form $\alpha_i+\alpha_j$ after suitable normalization, would reflect perfect substitutability and the form max $\{\alpha_i,\alpha_j\}$ perfect complementarity. This idea has been explored in Brueckner (2003) through effort choice in link formation. Finally, Goyal and Joshi (2003) in a model with full reliability, focus on how the marginal benefit of an extra link is affected by the number of links of the agent or the number of links of other agents. In our context, with imperfect reliability the effects of additional links by other agents can be of either sign, both with exogenous and endogenous success probabilities. Therefore, some but by no means all numerical specifications of our model will fit into the Goyal and Joshi classification.

5. Concluding remarks

The assumption of link heterogeneity in the form of imperfect reliability in social networks provides a richer set of results than the homogeneous setting. In particular, Bala and Goyal's work on Nash networks shows that results under imperfect reliability are quite different from those in a deterministic setting. With the introduction of heterogeneity this clear distinction no longer prevails and our findings encompass results of both types of models. For instance, decay models (with perfect reliability) compute benefits by considering only the shortest path between agents. Extra indirect links do not contribute to benefits. Given a resulting minimally connected Nash network g of such a model, there exists a parameter specification of our model that also gives rise to g as a Nash network. On the other hand, super-connected Nash networks can occur as well.

Acknowledgements

This is an abbreviated and revised version of Haller and Sarangi (2003). We would like to thank two referees for beneficial critique and very careful reading of the paper. We are grateful to Rob Gilles, Sanjeev Goyal, Mark Stegeman and Kaz Miyagiwa for helpful suggestions and to Richard Baron, Jacques Durieu, Christoph Hofmann, and Philippe Solal for thoughtful comments. The paper has also benefited from comments of the participants in the SITE 2000 workshop at Stanford University and seminar audiences at Bielefeld, Bonn, Delhi School of Economics, Heidelberg, Karlsruhe, LSE and University of Arizona. Sudipta Sarangi gratefully acknowledges the hospitality of *DIW* Berlin where a part of this research was carried out. The usual disclaimer applies.

References

Bala, V., Goyal, S., 2000a. A non-cooperative model of network formation. Econometrica 68, 1181–1229.

Bala, V., Goyal, S., 2000b. A strategic analysis of network reliability. Review of Economic Design 5, 205–228.

Brueckner, J., 2003. Friendship networks. Mimeo, University of Illinios Urbana-Champaign.

Galeotti, A., Goyal, S., Kamphorst, J., in press. Network formation with heterogeneous players. Games and Economic Behavior.

Gilles, R.P., Sarangi, S., 2003. The role of trust in costly network formation. Center Discussion Paper Series 2003-53, Tilburg University.

Goyal, S., Joshi, S., 2003. Unequal connections. Mimeo.

Granovetter, M., 1974. Getting a Job: A Study of Contacts and Careers. Harvard University Press, Cambridge MA.

Haller, H., Sarangi, S., 2003. Nash networks with heterogeneous agents. LSU Working Paper 2003-06, available at http://ideas.repec.org/p/lsu/lsuwpp/2003-06.html.

Jackson, M.O., 2005. A survey of models of network formation: stability and efficiency. Chapter 1. In: Demange, G., Wooders, M. (Eds.), Group Formation in Economics; Networks, Clubs and Coalitions. Cambridge University Press, Cambridge.

Jackson, M.O., Watts, A., 2001. The existence of pairwise stable networks. Seoul Journal of Economics 14, 299–321.

Jackson, M.O., Watts, A., 2002. The evolution of social and economic networks. Journal of Economic Theory 106, 265–295.

Jackson, M.O., Wolinsky, A., 1996. A strategic model of economic and social networks. Journal of Economic Theory 71, 44–74.

Johnson, C., Gilles, R.P., 2000. Spatial social networks. Review of Economic Design 5, 273–299.

McBride, M., 2002. Position-specific information in social networks. Mimeo, Department of Economics, UC Irvine.

Rogers, E., Kincaid, D.L., 1981. Communication Networks: Towards a New Paradigm for Research. Free Press, New York.

Slikker, M., van den Nouweland, A., 2001. Social and Economic Networks in Cooperative Game Theory. Kluwer Academic Publishers, Boston.