View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Munich Personal RePEc Archive

MPRA
Munich Personal RePEc Archive

A noncooperative model of network formation with decreasing productivity

Banchongsan Charoensook

ALHOSN University, Faculty of Business

10. February 2012

Online at https://mpra.ub.uni-muenchen.de/36570/
MPRA Paper No. 36570, posted 2. March 2012 20:17 UTC

Network Formation with Productivity as Decay

Banchongsan Charoensook *

First Version: February 2012

March 2, 2012

Abstract

This paper develops a model of noncooperative network formation. Link formation is two-sided. Information flow is two-way. The paper is built upon Bala and Goyal (2000). A unique assumption is that the value of information decays as it flows through each agent, and the decay is increasing and concave in the number of his links. Thus, an agent may choose to avoid accessing an agent who possess many links since he is aware of the decay incurred through this agent. This avoidance leads to two particular results in the analysis of Nash networks: (1) Nash networks are not always connected; (2) Nash networks do not exist under some parameters. Since disconnectedness is reminiscent of a common feature of real-world network, the model may explain why real-world networks may exhibit this feature even when there is no heterogeneity among agents. Discussion on this insight is provided.

*ALHOSN University, Uniterd Arab Emirates. E-mail: b.charoensook@alhosnu.ae. This paper is a chapter of my doctoral dissertation, which was supervised by Dino Gerardi. I thank him for his guidance and generous supports

1 Introduction

This paper presents a model of network formation game that is built upon the two-way flow model of Bala and Goyal (2000), henceforth BG. A unique assumption is that *an increase in link establishment damages the quality of information that flows in the network*. Under this assumption, each agent knows that his decision to add a link causes a decline in the value of information flow. This decline is a disbenefit both to himself and other agents in the network. Hence, on top of link formation cost, there are additional disbenefits associated with link formation. This paper aims to understand how this assumption may affect link-formation decision of agents and hence the shape of networks. To this end, we characterize the shapes of equilibrium networks and analyze why they differ from those in the literature. Finally, using the analyses the paper discuss how the model may explain some features of real-world networks.

We argue that our assumption is realistic and hence worth studying, particularly in the context of information network. Consider a firm in which employees'task is to communicate with each other. In this network, there may be a center-like agent whose role is to collect and distribute informations of other agents. Therefore, the quality of information flow depends on the center's communicating performance. Thus, it is likely that his performance declines as contacts between him and other agents increase. In such context, each agent has to take into account that contacting the center damages the information flow. Hence, the value of information he receives may not worth the efforts to contact. When this is the case, he may avoid contacting the center by contacting another agent or staying disconnected from this network. At the same time, the center may decline the contact initiated by an agent if the decline in his communicating performance does not allow him to reap much benefit from the information that flows through him.This problem is known as *network congestion* in the context of communication networks, where we call the first case. However, how this realism affects agents' linking decision has not been investigated in the literature in network formation to our knowledge. Thus, our attempt to address this uninvestigated issue is the central contribution of this paper.

With this situation in mind, we address this network congestion issue by making the following modification to the two-way flow model of BG. Whenever information passes through an agent, a quantity of information loss is incurred. The value of the remaining information is decreasing and strictly concave in the amount of agent's links [1]. In contrast, in the original model of BG, the quantity of information loss depends solely on the distance between agents. Specifically, the original setting of BG is as follows. Each agent possesses a unique private piece of information that is nonrival. He can choose to sponsor costly links to any agents without their agreements. All links together form the network. If there is a link or a series of links between two agents, they are obliged to share their private informations. Thus, the decision of

[1] The justification of this stylized assumption is relegated to the model section.

agent to form a link represents his decision to make his private information available to other agents in exchange of receiving their informations, and concurrently his willingness to be an information transmitting device.

Since we model this network congestion in a stylized way, we provide two justifications. First, this model makes observing the effects of congestion avoidance easier. The original BG model and our model permit each agent to access others without their agreements. This implies that each agent decides on his own as to how to avoid the congestion he finds in the network, hence easing the observation. This advantage is also facilitated by the assumption that agents'information are nonrival [2]. Second, because links are formed in a noncooperative way, Nash equilibrium in pure strategies can be applied as the solution cencept. This eases the analysis.

Despite, the one-sided access assumption has a major disadvantage. It implies that agent cannot defend against an access by another agent, even when the access lowers his payoff. This implication is not realistic. For example, in a file sharing network, one agent may decline an access by another agent if the access lowers his internet speed. Hence, our model does not provide an insight to this side of reality.

Based on the observation from the main results, two insights into the effects of network congestion on the structure of real-world network can be learnt. First, when network congestion is present, an equilibrium network may be fragmented, consisting of subnetworks disconnected from each other. The intuition is that agent in one network may avoid entering another due to the congestion. This may explain why empirical literature finds that disconnected networks are common in the real world. Second, with network congestion, moving from a smaller network to a larger one (a network with more agents) does not imply that the moving agent will improve his payoffs. The intuition is that agents in a larger network may be more congested (having more links), causing information to flow better in a smaller network. This may explain why real-world networks often consist of fragmented communities of notably different sizes. For example, in a friendship network, some students may prefer to keep their friendship within a small group rather than joining the crowd because they enjoy a stronger friendship that provides a higher benefit flow. While our paper models network congestion in a stylized way, this analysis may provide an alternative explanation of 'social isolates' observed in the real world, suggesting that the underlying cause of such network feature may not necessarily be heterogeneity among agents.These insights can be observed in our first proposition, which finds that no nash network is connected under some restriction on the congestion parameter. This disconnectedness is a sharp contrast to the result in the original model of BG that all Nash Networks are connected.

[2] if assumed otherwise, it may be difficult to distinguish whether agent decides not to access another as a result of the congestion or the rival nature of information.

Beside the above disconnectedness, two results are also different from BG's. First, Nash Network in pure strategies do not exist under some parameters. This result is shown by an example. Second, no stars are Nash except center-sponsored star [3].

Our paper contributes to the literature in network formation. This literature is pioneered by the work of Jackson. and Wolinsky (1996)[4]. Their model assumes that two agents must share a mutual consent in order that a link is established. A seminal work that contrasts to this model is that of BG, in which one-sided link formation is assumed. Among existing extensions of BG, the model of Caffarelli (2009) has in mind a situation in which managing too many links simultaneously leads to information congestion. It assumes that the cost of link maintenance increases in relation to the quantity of informations received. Hence, accessing an agent does not directly damage the quality of information flow at the accessed agent. Our model differs in that network congestion is reflected directly in the increasing information loss both the agent being accessed and the accessing agent. This allows us to better observe the effects of congestion avoidance. Beside this difference, Caffarelli (2009) assumes that information sharing is not two-way, in that the the agent who forms link does not share his information with his partner.

2 The Model

$N = \{1, ..., n\}$ is a set of agents and i and j are typical members of this set. Each agent possesses a unique private piece of information that is valuable both to himself and anyone who has an entry to it. Whenever i and j together share their informations, i has an entry to the information of j and vice versa. However, the information transmission is made possible only if a *pairwise link* between them is established.

Link establishment is costly and one-sided. i can spend the cost c to establish a link with j without j's consent. Therefore, a strategy of i is a set $g_i = (g_{i1}, ..., g_{ii-1}, g_{ii+1}, ..., g_{in})$ where $g_{ij} \in \{0, 1\}$ and $g_{ij} = 1$ if and only if i forms$j$ by paying c. In this case, we say that i *accesses* j. We restrict our analysis to pure strategies throughout the paper. Let $g = (g_1, ..., g_n)$ be a strategy profile. The strategy space i is $\mathcal{G}_i$ and the set of all pure strategy profiles is $\mathcal{G} = \{\times \mathcal{G}_i\}_{i=1}^{n}$.

To visualize how information flows among agents, a strategy profile g can be represented by a *network*. Pictorially, a network consists of a set of nodes and a set of arrows pointing from one node to another. To enable the network representation of a strategy profile, a one-to-one correspondence between the set of all directed networks with n nodes and the set of strategy ptofiles $\mathcal{G}$ is contructed by the following rule. In a network g, we enumerate the node from 1 to

[3] a star is a network in which there is a unique center-like agent who connects to all other agents. But all other agents have no links with each other. A center-sponsored star is a star that the center sponsors the link to everyone.

[4] Jackson (2007) provides an overview of network formation literature

n and let there be an arrow pointing from i to j if and only if i accesses j in the strategy profile g. Therefore, we use the term network and strategy profile interchangably onwards. Figure 1 depicts an example of a network

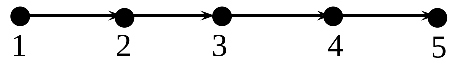


Figure 1: A network with five agents. $n = 5, g_1 = \{1,0,0,0\}, g_2 = \{0,1,0,0\}, g_3 = \{0,0,1,0\}, g_4 = \{0,0,0,1\}, g_5 = \{0,0,0,0\}$

In a network, a link of i can be sponsored by himself or the other agent. Thus, to distinguish the sponsorship, let $N(i;g) = \{k \in N | g_{ik} = 1\} \cup \{i\}$ be the set of all agents whom i accesses and $\mu_i(g) \equiv |N(i;g)| - 1$ be the number of links that i establishes. Notice that $i \in N(i;g)$ because i can access his own information [5]. We indicate whether there is a link between i and j by the term $\bar{g}_{ij} = max\{g_{ij}, g_{ji}\}$. Hence, $\bar{g}_{ij} = 1$ if and on if there is a link between i and j. In this case, we say that *i links with j*. Similarly, we define $\bar{N}(i;g) = \{k \in N | \bar{g}_{ik} = 1\} \cup \{i\}$ and $\bar{\mu}_i(g) \equiv |\bar{N}(i;g)| - 1$. An agent in $N(i;g)$ and $\bar{N}(i;g)$ is called *directed neighbor* and *neighbor* of i respectively.

With these notations, we turn to describe how information flows in a network. Apart from the direct transmission via a single link, the information also flow indirectly via a series of links. Formally, an *ij-path* is a sequence $\bar{g}_{i,j_1}, \bar{g}_{j_1 j_2}, ..., \bar{g}_{j_m j}$ whose each element is 1, and is denoted by $P_{ij}(g)$. The set of all $P_{ij}(g)$ is $\mathcal{P}_{ij}(g)$. If an *ij-path* exists, we say that *i observes j*. Notice that the existence of P_{ij} guarantees the existence of $P_{ji}(g) = \bar{g}_{jj_m}, \bar{g}_{j_m j_m - 1}, ..., \bar{g}_{j_1 i}$.

Ideally, if information is trasmitted and received perfectly by i, it gives i the payoff of 1. However, throughout the transmission this value may decay. In this paper, we assume that the decay is incurred *nodewise*. As the information traverses through agent i, the *productivity* of i, $\sigma(i;g)$, is the percentage rate at which the value is preserved. Hence, if the information is transmitted through a path $\bar{g}_{i,j_1}, \bar{g}_{j_1 j_2}, ..., \bar{g}_{j_m j}$, the value of j's information that i receives is $\sigma(i;g)\,\sigma(j_1;g)\,\sigma(j_2;g)\,...\sigma(j;g)$ and is denoted by V_{ij}. Figure 2 illustrates how the values of information of other agents flow to agent 1 in the network of Figure 1.

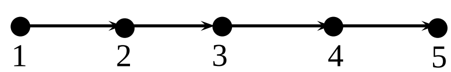


Figure 2: In the above network, $V_{12} = \sigma_1(1)\,\sigma_2(2) = \sigma_2, V_{13} = \sigma_1(1)\,\sigma_2(2)\,\sigma_2(3) = \sigma_2^2, V_{14} = \sigma_1(1)\,\sigma_2(2)\,\sigma_2(3)\,\sigma_2(4) = \sigma_2^3, V_{15} = \sigma_1(1)\,\sigma_2(2)\,\sigma_2(3)\,\sigma_2(4)\,\sigma_1(5) = \sigma_2^3$

If there are multiple ij-paths, the value of j's information to i is given by the *optimal paths*. Formally, let $\mathcal{P}_{ij}(g) = \left\{P_{ij}^1(g), P_{ij}^2(g), ..., P_{ij}^L(g)\right\}$ be the set of all paths through which i observes j in a network g. The value of the information of j that i obtains in this network is $\bar{V}_{ij}(g) = \max\limits_{k \in 1,..,L} V(P_{ij}^k; g)$. We call a path that solves $\max_{k \in 1,..,L} V(P_{ij}^k; g)$ an *optimal ij-path*. The set

[5] This assumption follows the convention set by the original model of BG.

of all optimal paths is $\bar{\mathcal{P}}_{ij}(g)$. If there is no path through which i observes j, we let $\bar{V}_{ij}(g) = 0$. For i's own information, we let $V_{ii} = \sigma(i;g)$. That is, the decrease in his productivity also decreases his own value of information [6].

Having defined the value of information, we are now ready to define the payoff of player i from the strategy profile g in a game with n players. It is:

$$U(i;g) = \sum_{j \in N} \bar{V}_{ij}(g) - c \cdot \mu_i(g)$$

The first term on the right-hand side is the total value of information i receives in g or *the revenue of* i in g and is denoted by $Rev(i;g)$.

We point out a difference between our model and BG's before adjourning this subsection. This difference is in how information decays. In BG, the decay factor is assumed to be *linkwise* and *geometric*. For example, let λ be this decay. If an ij-path consists of m links, then the information of j decays to λ^m when it arrives to i. Hence, the aggregated decay of a path depends solely on its length. In contrast, the decay in our model is defined *nodewise*, $\sigma(i;g)$. Therefore, two ij-paths with the same length may not provide the same value.

2.1 Assumptions on decay

Our key assumption is that the pdoductivity $\sigma(i;g)$ depends solely on the number of i's links. To formalize this idea, let $\sigma(i;g)$ be a function of $\bar{\mu}_i$, ie., $\sigma(i;g) : \mathbb{N} \to [0,1]$, and the value of $\sigma(i;g)$ at $\bar{\mu}_i$ is $\sigma_{\bar{\mu}_i}(i;g)$. When omission is possible we simply write $\sigma(i;g)$. Throughout the paper, the following assumptions on $\sigma(i;g)$ are assumed.

Assumption 1 (Concave Decreasing Decay). *σ is decreasing and strictly concave in $\bar{\mu}_i$. Moreover,*

1. $\sigma(i;g) = \sigma(j;g)$ *if* $\bar{\mu}_i = \bar{\mu}_j$
2. $\sigma_1 = 1$
3. *There exists a positive number* $K > 1$ *such that* $\sigma_x = 0$ *for any* $x \geq K$.

Let us justify these assumptions. The first assumption implies that agents are homogeneous in productivity. Second, 2 implies that perfect communication when agent has one link. Finally, the strict concavity implies that the decline in productivity increases at an increasing rate. While there is no theoretical support, this assumption can be justified by some realistic scenarios. For

[6] For example, if agent's own information is an unread newspaper article, then he has to put efforts to understand it in order to receive the benefit. Hence, his productivity affects also his own value of information. This assumption is also established in existing literature. Feri and Melendez-Jimenez (2009) assumes likewise.

example, suppose that an agent stores all pieces of information in one place, then due to the limitedness of space the chance that two pieces of information get mixed up, causing more difficulties in communicating accurately is likely to increase at an increasing rate. Another example is when each piece of information is very similar to one another, then the chance that an agent does not know which is which is likely to increase at an increasing rate.

2.2 Network-related Definitions

This subsection introduces some properties of networks and definitions of special networks that are frequently referred in the analysis. Our first definition captures the concept that two networks can have the same structure even if links are not sponsored by the same agent.

Definition (Network Architecture). *Let g^1 and g^2 be networks and N^1 and N^2 be their set of players, g^1 and g^2 are said to have the same architecture if there exists a permutation $\pi : N^1 \to N^2$ such that $g_{ij} \in g^1$ if and only if $g_{\pi(i)\pi(j)} \in g^2$*

Definition (subnetwork). *Let g^1 and g^2 be networks and N^1 and N^2 be their set of agents, g^1 is a subnetwork of g_2 if $N^1 \subset N^2$ and $g^1 \subset g^2$*

Definition (Connected network). *a network g is connected if an ij-path exists for any $i, j \in N$.*

Definition (component). *Let g^1 be a subnetwork of g^2 and N^1 and N^2 be their set of agents, g^1 is a component of g^2 if*

1. *g^1 is connected*
2. *there exists no ij-path in g_2 if $i \in N_1$ and $j \in N_2$*

The above definition reflects that a network can be fragmented, having different subnetworks that are disconnected from each other.

Definition (Minimal Network). *a network is minimal if every ij-path is unique.*

With the introduction of these properties, the followings are the definitions of some network architectures.

Definition (Line). *A network g is a line if there exists a permutation of m agents $j_1, ..., j_m$ such that $\bar{g} = \{\bar{g}_{j_1,j_2}, \bar{g}_{j_2,j_3}, ..., \bar{g}_{j_m-1 j_m}\}$. Moreover, if there are m agents in a line, we say that its length is $m - 1$.*

Definition (Empty network). *A network is an empty network if every agent in it has no link.*

Definition (Star). *A network g is a star if*

1. *there exists exactly one agent i^c that has a link with every other agent;*

2. *for any two agents that are not i^c, there is no link between them.*

Moreover, if i^c accesses everyone, we say that the star is a center-sponsored star.

2.3 Nash Networks

Given a network g, if we remove all the links that i establishes, the network that remains is a collection of the strategies of all other agents except i. Denote this remaining network by g_{-i}. We write $g = g_i \oplus g_{-i}$ to emphasize that a network g can be formed by the union of g_{-i} and g_i. We use these notations to introduce to following definitions.

Definition (Best response)**.** *A strategy g_i is a* best response *of i to g_{-i} if*

$$U\left(i; g_i \oplus g_{-i}\right) \geq U\left(i; g_i' \oplus g_{-i}\right), \quad \textit{for all } g_i' \in G_i$$

Definition (Nash network)**.** *A network g is a* Nash network *if g_i is a best response to g_{-i} for every agent $i \in N$.*

We remark the following relation between Nash Network and its architecture. If a network is Nash, so are all networks that have the same architecture. This relation is used tostate the results in the next section.

3 Main Results

Our goal is to identify Nash Networks and their properties. For $\sigma_2 \leq \frac{1}{2}$, Nash network exists regardless to the cost range and number of players. Proposition 1 also provides a full equilibrium characterziation. In contrast, for $\sigma_2 > \frac{1}{2}$, Nash Network always does not exist. We provide an example that shows the nonexistence and study some properties of minimal Nash network in Proposition 2, 3 and 4.

Proposition 1. *1. If $\sigma_2 \leq \frac{1}{2}$, Nash network exists for any cost c and number of players n. Moreover, any component in it is one of the following three types.*

- *a three-agent line whose central agent does not establish a link, ie., network (a) in Figure 3*
- *a two-agent line, ie., network (b) in Figure 3*
- *an isolated agent, ie., network (c) in Figure 3*

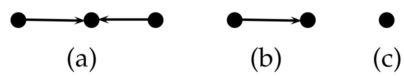


Figure 3: Three types of components in a Nash network, given that $\sigma_2 \leq \frac{1}{2}$

2. Using the network (a), (b) and (c) in Figure 3, the set of nash networks for each set of parameters c and σ_2 is given below.

- *If $c > 1$ and $\sigma_2 = \frac{1}{2}$, then the empty network is the unique Nash.*
- *If $c \leq 1$ and $\sigma_2 = \frac{1}{2}$, then the set of Nash networks consists of networks that have the following architectures:*
  - *the empty network*
  - *the network with the following properties: (1) each component is either (a) or (b) or (c); and (2) at most one component is (c)*
- *If $c > 1$ and $\sigma_2 < \frac{1}{2}$, then the empty network is the unique Nash.*
- *If $c = 1$ and $\sigma_2 < \frac{1}{2}$, then the set of Nash networks consists of all networks that have the following architectures:*
  - *the empty network*
  - *the network with the following properties: (1) each component is either (b) or (c); and (2) at most one component is (c)*
- *If $2\sigma_2 < c < 1$, then the set of Nash networks consists of the networks whose architecture has the following properties: (1) each component is either (b) or (c); and (2) at most one component is (c)*
- *If $c \leq 2\sigma_2 < 1$, then the set of Nash networks consists of the networks that have the following architectures:*
  - *at most one component is (c), the rest of the components are (b)*
  - *each component is (a) or (b)*

A noticable feature of Nash Networks in Proposition 1 is that none of them are connected, given that $n > 3$. This is a sharp contrast to Proposition 5.3 and its generalization by Jaegher and Kamphorst (2008) that show that every Nash Networks is connected. What drives this contrast? The result in BG relies on the following intuition: if i believes that the component he is accessing provides more benefit than another component that j is accessing, then j's deviation is to leave his component and enter the component of i. By this deviation, j receives at least as much as the payoff of i. However, under the concave decreasing decay assumption, this

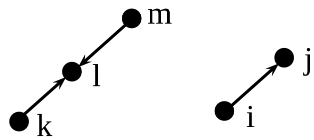


Figure 4: A nash network with five agents for $c = 2\sigma_2 < 1$

reasoning is not valid. Whenever j enters the component of i, he is increasing the congestion at the node being accessed. The congestion may cause the information flow to be very low. Hence, there is no guarantee that his payoff will improve. The following example clarifies this intuition by showing what happens when the constant decay assumption in BG is replaced by our assumption.

Example 1. *Consider the Nash network for $c = 2\sigma_2 < 1$ in Figure 4. Suppose that a decay on each link is a constant $\lambda < 1$ as in BG. k's gain from accessing l is $\lambda + \lambda^2$ while i's gain from accessing j is λ. Therefore, i is better off imitating the strategy of k by accessing l instead of j. This imitating deviation of i gives him the gain of $\lambda + 2\lambda^2$ and rules out this network to be Nash.*

We now replace the constant λ by our assumption that $2\sigma_2 = c < 1$ and show that the same imitating strategy is no longer a positive deviation, causing this network to remain a candidate for Nash network. Define the gain from accessing an agent as the total information value that arrives from that agent. As opposed to the above case, the gain of i from accessing j, $\sigma(j) = 1$, is higher than the gain of k from accessing l, $2\sigma_2 = \sigma_2(l) + \sigma_2(l)\sigma_1(m)$. Moreover, because $\sigma_3 = 0$, i's payoff becomes 0 if he imitates the strategy of k by accessing l. Therefore, such imitating strategy does not improve i's payoff. In addition, neither accessing k or m will improve his payoff because $\sigma_2 < \frac{1}{2}$.

On the other hand, can k improve his payoff by imitating the strategy of i? If k accesses j instead of l, his gain is $2\sigma_2 = \sigma_2(j) + \sigma_2(j)\sigma_1(i)$. Because the gain from accessing j is equal to the gain from accessing l, this imitating strategy is not a positive deviation of k.

Contrary to Proposition 1, if $\sigma_2 > \frac{1}{2}$, Nash network *does not exist* for some parameters c and n. An example which shows the nonexistence is given below. While the proof is relegated to the appendix, an intution is hereby provided. The non-existence stems from that $\sigma(i; g)$ of any agent i changes in a discrete way. In other words, whenever an agent deviates from his strategy, the productivities of involved agents decrease or increase discretely. Hence, for some σ, c and n, it may turn out that there exists an agent who finds a positive deviation in any given network.

Example 2. *Given that $\frac{1}{\sqrt{2}} > \sigma_2 > \frac{1}{2}$, $\sigma_3 = 0$, and $c = 0.98$, no network with 5 agents is Nash.*

The second proposition below gives a noticable property that two single-neighbor agents never access the same center-like agent in a minimal Nash network. This result is driven by congestion avoidance: a single-neighbor agent avoids the center and access another single-neighbor agent instead. Such avoidance is profitable because $\sigma_2 > \frac{1}{2}$ guarantees a sufficiently low congestion at the agent being accessed. The consequence is that every link between a single-neighbor agent and the center is sponsored by the center in equilibrium. Formally, call a single-neighbor agent, *end node*, the agent who is his neighbor *parent*, and an agent who has no link *isolated node*. We use these terms to state Proposition 2.

Proposition 2. *Given that* $\sigma_2 > \frac{1}{2}$ *and* $n > 3$ [7]*. In a minimal Nash network g, let j be an end node and i be his parent,*

1. *if j accesses i, j is the only end node of i;*
2. *if i accesses j, i accesses all his end nodes. Morover, if he has more than one end node, then he accesses all his neighbors, including ones that are not end nodes.*

The same intuition above is also applied to prove Proposition 3. Informally, If there are two centers, i and k, who access their end nodes, i will find that accessing an end node of k is more profitable than accessing his own end node. The reason is that accessing his end node provides him only one piece of information, while accessing an end node of k provide multiple pieces, yet with a higher congestion. Such deviation becomes profitable if the higher connection at k is not too high, which is guaranteed by $\sigma_2 > \frac{1}{2}$.

Proposition 3. *Given that* $\sigma_2 > \frac{1}{2}$*, if g is a minimal Nash network, g has at most one component that contains a parent who accesses all his neighbors.*

The power of Proposition 2 and 3 is that they rule out many minimal networks to be Nash regardless to σ, c and n. For example, a network that consists of two disconnected center-sponsored stars cannot be Nash because one center will find a positive deviation by accessing an end node of the other center. Figure 5 illustrates some networks that are ruled out.

As a result of Proposition 2, the only star that remains a candidate for Nash Network is the center-sponsored star. This result differs from Proposition 5.3 in BG which shows that all kinds of stars are Nash under some range of decay. Proposition 4 below provides a neccessary and sufficient condition for a center-sponsored star to be Nash. This condition involves a restriction on σ and n. Figure 6 gives an example of parameters that satisfie this restriction.

Proposition 4. *Given that* $n > 3$*, let g be a star network*

1. *If g is not a center-sponsored star, g is not nash*

[7] if $n \leq 3$, this proposition does not apply. Every component of nash network is either a line or empty. The proof is trivial and is omitted

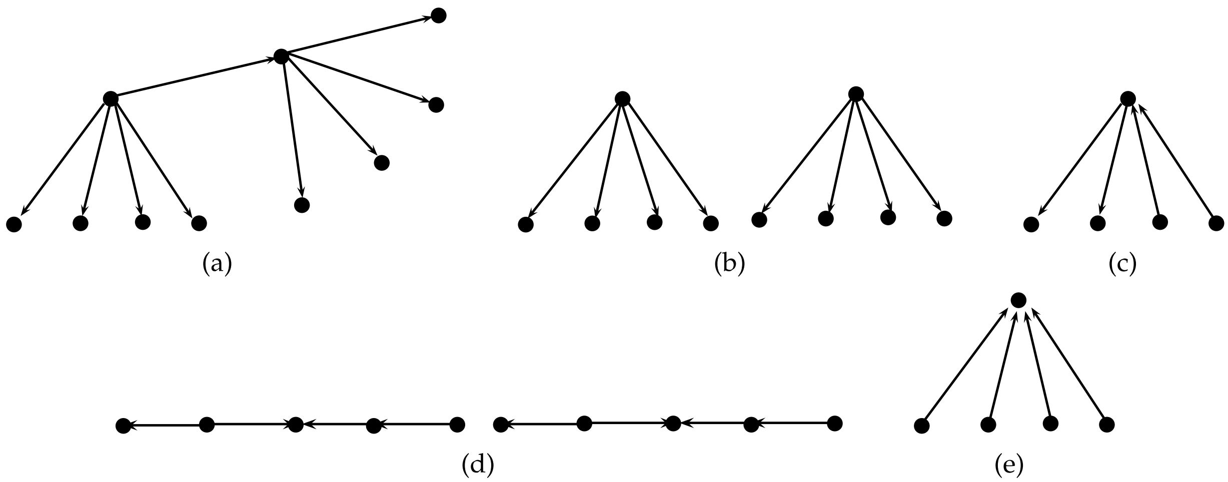


Figure 5: Networks that are ruled out to be Nash according to Propositon 2 and 3

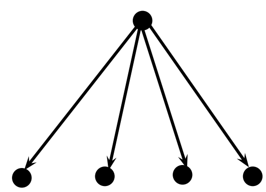


Figure 6: Nash Network with $n = 5$, $c = 0.5$, $\sigma_5 = 0.9$, $\sigma_4 = 0.969$ $\sigma_3 = 0.9799$ $\sigma_2 = 0.99$ $\sigma_1 = 1$

2. *If g is a center-sponsored star, the following inequalities are neccessary and sufficient condtions for g to be nash:*

$$\sigma_{n-1} - (n-1)\left(\sigma_{n-2} - \sigma_{n-1}\right) \geq c \geq (\sigma_2)^2 - \sigma_{n-1} - (1 - \sigma_2)\left(1 + (n-2)\,\sigma_{n-1}\right)$$

Beside center-sponsored star, line is also a network architecture that is Nash for some σ, c and n. Figure 7 exemplifies two lines that are Nash.

(a) $n = 6, c = 0.2, \sigma_2 = 0.97, \sigma_3 = 0.2, \sigma_4 = \sigma_5 = \sigma_6 = 0$

(b) $n = 4, c = 0.3, \sigma_2 = 0.9, \sigma_3 = 0.5, \sigma_4 = 0$

Figure 7: Two lines that are Nash

4 Discussions

This section observes two features of Nash Networks in this model that differ from those in BG. These features results from that the constant decay assumption of BG is replaced by concave decreasing productivity. Intuitions and discussions on how these features are likely to be exhibited in real-world settings are also provided.

4.1 Network congestion may lead equilibrium networks to be disconnected

The first observation comes from that all Nash networks for $\sigma_2 \leq \frac{1}{2}$ are disconnected (Proposition 1). The intuition, which is made clear by Example 1, can be summarized as follows. While establishing a link to an agent is a way to access a component, it also increases the congestion at the agent being accessed. This congestion may cause much loss in the information transmitted via the agent. When such congestion, or inefficieny in information transmission, is sufficiently high, agent may be better off avoiding the congestion altogether and remaining disconnected from the component.

How does this observation help understand real-world phenomena? Our model may serve as a hypothesis that explains why empirical evidences find that real-world networks are often disconnected [8]. For example, if a society is considered as a network in which information is exchanged among agents, it is likely that the society is fragmented into small communities if agents find that avoiding connecting to each other is a way to reduce inefficiency in information flow.

4.2 Connecting to a larger component does not imply a higher gain

Our second observation is that a smaller component may provide a higher gain to their members than a larger one. The observation comes from that many nash networks in Proposition 1 consists of components whose size, or the number of agents, are not equal. Consider, for example, the equilibrium network in Example 1. Observe that i chooses to access an isolated agent j rather than someone in the larger component. If i accesses j, j's productivity is σ_1. If i accesses someone in the larger component, the productivity of the accessed agent is at most σ_2. Hence, if σ_2 is sufficiently lower than σ_1, then entering a larger component gives i a lower gain.

This observation may explain why there are agents who prefer to reside in a smaller community rather than a larger one in a real-world social network. When link is a source of inefficiency, a smaller community that has less connections may provide a higher benefit to the participating members such that they do not want to join a crowded community. In other words, agents may face a tradeoff between quantity of information and quality of information when network congestion is present. While a larger community may have more information, the quality of information may be deterred if agents possess too many connections. A friendship network among students may serve as an example of this hypothesis. Some students may choose to maintain their friendships within a smaller group and avoid contacting the crowd because they enjoy a stronger tie of friendship.

[8] For instance, Ennett and Bauman (2000) observes that a common feature of friendship networks is that there are agents who are social isolate, disconnecting themselves from the principal component. R. Kumar and Tomkins (2010) also finds that some online social networks contain isolated communities.

5 Conclusion

This paper provides a stylized model with two key assumptions. First, link can be formed without a mutual consent between agents. Second, link addition increases the congestion, or more information loss, at the accessed agent and the agent who accesses. The model allows us to see how an agent may avoid accessing other agents due to an increasing congestion. The two key assumptions lead to equilibrium networks that are disconnected. Moreover, nonexistence of equilibrium network in pure strategies arises under some parameters. These two features are different from the results in the original setting of Bala and Goyal (2000) from which this model is developed.

References

BALA, V., AND S. GOYAL (2000): "A Noncooperative Model of Network Formation," *Econometrica*, 68(5), 1181–1230.

CAFFARELLI, F. (2009): "Networks with decreasing returns to linking," Economics Working Papers ECO2004/18, Bank of Italy.

ENNETT, S., AND K. BAUMAN (2000): *Increasing Prevention Effectiveness*chap. Adolescent Social Networks: Friendship Cliques, Social Isolates, and Drug Use Risk, pp. 83 – 92. Tanglewood Research, Inc.

FERI, F., AND M. A. MELENDEZ-JIMENEZ (2009): "Coordination in Evolving Networks with Endogenous Decay," Working Papers 2009-19, Faculty of Economics and Statistics, University of Innsbruck.

JACKSON, M. (2007): *The Missing Links: Formation and Decay of Economic Networks*chap. 2: The Study of Social Networks in Economics, pp. 19–40. Russell Sage Foundation.

JACKSON., M., AND A. WOLINSKY (1996): "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, 71(1), 44–74.

JAEGHER, K. D., AND J. KAMPHORST (2008): "Network formation with decreasing marginal benefits of information," Working Papers 08-16, Utrecht School of Economics.

R. KUMAR, J. N., AND A. TOMKINS (2010): *Link Mining: Models, Algorithms, and Applications*chap. Structure and Evolution of Online Social Networks, p. 611âĂŞ617. Sprin.

A Appendix

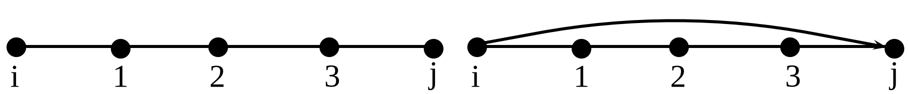
This appendix covers all the proofs in this paper. We begin by introducing some concepts and a lemma.

A.1 The Concepts of Marginal Cost and Marginal Revenue

Let us introduce two concepts that will be used to simplify the equilibrium analyses. Given a network g, by identifying that an agent i has a profitable deviation by adding or destroying just one link, we can show that g is not in equilibrium. Most of our analyses rest upon this proving technique. Let $g + ij = g \cup g_{ij}$ and $g - ij = g \backslash g_{ij}$ be the networks corresponding to the addition and deletion of g_{ij} respectively.

When i adds g_{ij} , this new link may give an entry to some new agents that are not i's neighbor in g. At the same time, via g_{ij}, i may find a new path through which he observes the neighbors that pre-exist in g. For some neighbors, a new path may yield a higher payoff than the optimal one that i uses when he is in g. Therefore, as g changes to $g + ij$, the optimal path to reach some pre-existing neighbors may also change. We denote the set of such new agents and the set of such pre-existing neighbors of i by $N_i\left(g, g+ij\right) = \{j \in N | j \in N_i\left(g+ij\right) \wedge j \notin N_i\left(g\right)\}$ and $M_i\left(g, g+ij\right) = \left\{j \in N_i\left(g\right) | \bar{\mathcal{P}}_{ij}\left(g\right) \cap \bar{\mathcal{P}}_{ij}\left(g+ij\right) = \emptyset\right\}$, where $\bar{\mathcal{P}}_{ij}\left(g\right)$ is defined as in section 2, respectively. The example below illustrates how $N_i\left(g, g+ij\right)$ and $M_i\left(g, g+ij\right)$ are identified.

Example 3. *Consider Figure 8.* $N_i\left(g, g+ij\right) = \emptyset$*, because i finds no new neighbor in* $g+ij$*.* $M_i\left(g, g+ij\right) = \{j, 3\}$*, because* $\bar{P}_{ij}\left(g\right) = \left\{\bar{g}_{i1}, \bar{g}_{12}, \bar{g}_{23}, \bar{g}_{3j}\right\}$ *but* $\bar{P}_{ij}\left(g+ij\right) = \left\{\bar{g}_{ij}\right\}$ *and* $\bar{P}_{i3}\left(g\right) = \left\{\bar{g}_{i1}, \bar{g}_{12}, \bar{g}_{23}\right\}$ *but* $\bar{P}_{ij}\left(g+ij\right) = \left\{\bar{g}_{ij}, \bar{g}_{j3}\right\}$



(a) Network g (b) Network $g+ij$

Figure 8: Example 3

With these notations, we are ready to introduce the concept of *marginal revenue*, which sums together the additional benefit of i when he adds g_{ij} to g

Definition (Marginal Revenue)**.** *Let* $N_i\left(g, g+ij\right)$ *and* $M_i\left(g, g+ij\right)$ *be defined as above, the marginal revenue of i when adding* g_{ij} *to g is*

$$MR\left(i; g_{ij} \to g\right) = \sum_{k \in N_i\left(g, g+ij\right)} \bar{V}_{ik}\left(g+ij\right) + \sum_{k \in M_i\left(g, g+ij\right)} \left(V_{ik}\left(g+ij\right) - V_{ik}\left(g\right)\right)$$

The first summation answers how much i gains from having entries to the new neighbors. The second summation quantifies how much i gains from having new optimal paths in $g+ij$ as compared to the old one in g. Observe that $V_{ik}(g+ij) - V_{ik}$ can be negative due to the decreasing productivity of i. Hence, $MR\left(i; g_{ij} \to g\right)$ can be negative.

Example 4. *In Figure 8, because* $M_i(g, g+ij) = \{j, 3\}$ *and* $N_i(g, g+ij) = \emptyset$, $MR(g, g+ij) = \Big(\sigma_2(i; g+ij)\,\sigma_2(j; g+ij) - \sigma_1(i; g)\,\sigma_2(1; g)\,\sigma_2(2; g)\,\sigma_2(3; g)\,\sigma_1(j; g)\Big) + \Big(\sigma_2(i; g+ij)\,\sigma_2(j; g+ij)\,\sigma_2(3; g+ij) - \sigma_1(i; g)\,\sigma_2(1; g)\,\sigma_2(2; g)\,\sigma_2(3; g)\Big) = \left(\sigma_2^2 - \sigma_2^3\right) + \left(\sigma_2^3 - \sigma_2^3\right)$

We now introduce the concept of marginal cost. As the network changes from g to $g+ij$, there are neighbors of i that he still uses the same optimal path to reach them. The set of these neighbors is $N_i(g) \backslash M_i(g, g+ij)$. Inspite of the same optimal paths, the information's value that reach to i decreases due to the decreasing productivity of i that results from adding the new link. The decreasing value, together with link establishment cost c, are the marginal cost formalized below.

Definition (Marginal Cost)**.** *Let* $M_i(g, g+ij)$ *be defined as above, the marginal cost of* i *from adding* g_{ij} *to* g *is* $MC\left(i; g_{ij} \to g\right) = c + \sum_{N_i(g)\backslash M(g,g+ij)} \left(\bar{V}_{ij}(g) - \left(\bar{V}_{ij}(g+ij)\right)\right)$

Example 5. *In Figure 8, because* $M_i(g, g+ij) = \{j, 1\}$ *and* $N_i(g, g+ij) = \emptyset$, $MC\left(i; g_{ij} \to g\right) = \left(\sigma_1(i; g) - \sigma_2(i; g+ij)\right) + \left(\sigma_1(i; g)\,\sigma_2(3; g) - \sigma_2(i; g+ij)\,\sigma_2(3; g+ij)\right) + \left(\sigma_1(i; g)\,\sigma_2(3; g)\,\sigma_2(2; g) - \sigma_2(i; g+ij)\,\sigma_2(3; g+ij)\,\sigma_2(2; g+ij)\right)$

Remark. *While MC is defined as above, the fact that i's productivity decreases is not expressed explicitly in the definition. By simple algebraic rearrangement, it can be expressed as* $MC\left(i; g_{ij} \to g\right) = \left(\sigma_{\mu_i(g)} - \sigma_{\mu_i(g)+1}\right)\left(\sum_{k\in N_i(g)\backslash M_i(g,g+ij)} \bar{V}_{ik}(g) \cdot \frac{1}{\sigma_{\mu_i(g)}}\right)$*. One can think of* $\bar{V}_{ik}(g) \cdot \frac{1}{\sigma_{\mu_i(g)}}$ *as the value of information of* k *that arrives to* i*, without taking* i*'s own productivity into consideration. Hence, this term remains the same, while the only change as* g *becomes* $g+ij$*, which is the productivity of* i*, is expressed as* $\sigma_{\mu_i(g)} - \sigma_{\mu_i+1(g)}$.

To use the MR and MC in the equilibrium analysis, we finally needs to show that i wants to deviate from the network g by adding g_{ij} if he finds that $MR\left(i; g_{ij} \to g\right)$ exceeds $MC\left(i; g_{ij} \to g\right)$. The following lemma serves this purpose.

Lemma 1. $U(i; g+ij) - U(i; g) = MR\left(i; g_{ij} \to g\right) - MC\left(i; g_{ij} \to g\right)$

Proof. *By the definition of* $N_i(g, g+ij)$ *and* $M_i(g, g+ij)$, $N_i(g+ij) = N_i(g) \sqcup N_i(g, g+ij)$ *and* $N_i(g) = \left(N_i(g) \backslash M_i(g, g+ij)\right) \sqcup M_i(g, g+ij)$*. Therefore,* $U_i(g+ij) = U_i(g) + K$ *if and only if*

$$\sum_{k\in M_i(g,g+ij)} \bar{V}_{ik}(g+ij) + \sum_{k\in N_i(g)\setminus M_i(g,g+ij)} \bar{V}_{ik}(g+ij) + \sum_{k\in N_i(g,g+ij)} \bar{V}_{ik}(g+ij)$$
$$+ \mu_i^d(g+ij)\cdot c = \sum_{k\in M_i(g,g+ij)} \bar{V}_{ik}(g) + \sum_{k\in N_i(g)\setminus M_i(g,g+ij)} \bar{V}_{ik}(g) + \mu_i^d(g)\cdot c + K$$

Rearranging the inequality,

$$\sum_{k\in M_i(g,g+ij)} \left(\bar{V}_{ik}(g+ij) - \bar{V}_{ik}(g)\right) + \sum_{k\in N_i(g,g+ij)} \bar{V}_{ik}(g+ij)$$
$$= c + \sum_{k\in N_i(g)\setminus M(g,g+ij)} \left(\bar{V}_{ik}(g) - \bar{V}_{ik}(g+ij)\right) + K$$

Observe that the left-hand side is $MR\left(i; g_{ij} \to g\right)$ *and the right-hand side is* $MC\left(i; g_{ij} \to g\right) + K$. *Therefore,* $U(i; g+ij) - U(i; g) = MR\left(i; g_{ij} \to g\right) - MC\left(i; g_{ij} \to g\right)$ ☐

Remark. *While the above lemma shows whether i has a profitable deviation by adding* g_{ij} *to g, it can also be used to show whether i has a profitable deviation by deleting* g_{ij} *in g. To do so, replace the term* $MC\left(i; g_{ij} \to g\right)$ *and* $MR\left(i; g_{ij} \to g\right)$ *with* $MC\left(i; g_{ij} \to (g-ij)\right)$ *and* $MR\left(i; g_{ij} \to (g-ij)\right)$ *in the above proof. This gives another version of the lemma:* $U(i; g) - U(i; g-ij) = MR\left(i; g_{ij} \to g-ij\right) - MC\left(i; g_{ij} \to g-ij\right)$.

A.2 Proofs of the propositions

Proof (Proof of Proposition 1)**.** *The strategy of this proof is to first eliminate all networks that have an agent with a positive deviation, and then identify which of the non-eliminated networks are Nash given c and* σ_2. *For convenience, type A, B and C component in the proof refers to the network (a), (b) and (c) in Figure 3.*

Step 1: If a network g ***has an agent who has more than two links,*** g ***is not nash.*** *Let this agent be i. Observe that* $\sigma_3 = \sigma_4 = ... = 0$ *because* $\sigma_2 \le \frac{1}{2}$ *and* σ *is strictly concave. Therefore, i's productivity* $\sigma(i; g) = \sigma_{\mu_i(g)}$ *is* 0 *because he has more than two links. Because* $\sigma(i; g) = 0$ *and i is on every ij-path,* $V\left(P_{ij}; g\right) = \prod_{k\in N\left(P_{ij}\right)} \sigma(k; g) = 0$ *for any j that is a neighbor of i. As a result, the revenue of i,* $\sum_{j\in N} \bar{V}_{ij}(g)$*, is* 0. *Therefore, if i accesses some agent in this network, he will want to remove it because the link establishment is costly but his revenue is* 0. *Moreover, if i is accessed by a neighbor j, j will also want to remove the link for the same reason. Because of these deviations, g is not nash.*

Step 2: For any network g ***that is not ruled out by Step 1, if it contains a component that is not one of the three networks in Figure 3, it is not nash.*** *Such component is either a cycle or a line that is neither (a) nor (b) in Figure 3. In what follows, we show that there is an agent who wants to deviate from such component.*

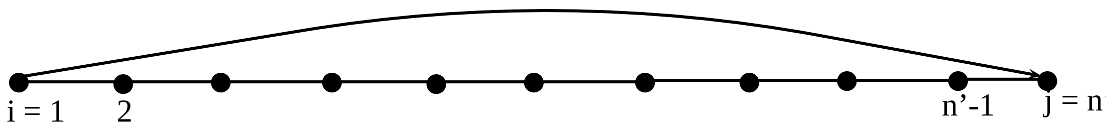


Figure 9: A cycle with n' agents, enumerated from left to right.

Step 2.1: Suppose that this component is a cycle with n' agents. *Denote this cycle agents by g^{cy}. In g^{cy}, consider an agent i that establishes at least one link g_{ij}. Observe that if i removes g_{ij} the cycle becomes a line. Denote this line by g^{li}. We will show that the payoff of i in g^{li} is greater than his payoff in g^{cy}.*

Consider the value of information that flows to i in g^{cy}. Enumerate the agents in g^{cy} according to Figure 9. For any k, there are two ik-paths. One is by using the same path that i uses when he is in the line, ie. when there is no g_{ij}. The other one is to retrieve the information of k via the new link g_{ij}. Let V_{ik}^{old} and V_{ik}^{new} denote the value of information from the two paths. Because every node in the cycle has productivity σ_2, $V_{ik}^{old} = (\sigma_2)^k$ and $V_{ik}^{new} = (\sigma_2)^{n-k+2}$. Therefore, $MR\left(i; g_{ij} \to g - ij\right)$ is $\sum_{k>\frac{n+2}{2}+2}^{n} \bar{V}_{ik}\left(g^{cy}\right) = \sum_{k>\frac{n+2}{2}}^{n} V_{ik}^{new}$ because $V_{ik}^{new} > V_{ik}^{old}$ for $n \geq k > \frac{n+2}{2}$. We now turn to find $MC\left(i; g_{ij} \to g - ij\right)$. Because $V_{ik}^{new} \leq V_{ik}^{old}$ for $k \leq \frac{n+2}{2}$, $MC\left(i; g_{ij} \to g - ij\right) = c + (1-\sigma_2)\left(\frac{1}{\sigma_2}\right)\left(\sum_{k\leq\frac{n+2}{2}} V_{ik}^{old}\right)$. Hence, $MC\left(i; g_{ij} \to g - ij\right) > MR\left(i; g_{ij} \to g - ij\right)$ because $\sigma_2 \leq \frac{1}{2}$. Applying Lemma 1 to this inequality, we conclude that i is strictly better off by deleting g_{ij}. Therefore, i has a positive deviation.

Step 2.2: Suppose that this component is a line network that is neither (a) nor (b) in Figure 3. *Denote this component by g^{line}. Consider all agents that have exactly two links. Among them, there exists at least one agent who establish one or two links. In what follows we aim to show that such agent is strictly better off by deleting exactly one link, breaking the line into two disconnected lines.*

We let this agent be i and a link that he establishes be g_{ij}. Without g_{ij}, i is disconnected from the line that contains j, denote the component that contains j by g^j and the other component, which contains i, by g^i. Suppose that there are n' agents in g^j, the total benefit from estaliblishing g_{ij} is $MR\left(i; g_{ij} \to g^{line} - ij\right) = \sigma_2\left(i; g^{line}\right)\sigma_1\left(j; g^{line}\right) = \sigma_2$ if $n' = 1$ and $MR\left(i; g_{ij} \to g^{line} - ij\right) = \sum_{k=1}^{n'} \bar{V}_{l,k} = \sum_{k=1}^{n'-1} \sigma_2^k \sigma_2\left(i; g^{line}\right) + \sigma_2^{n'-1}\sigma_2\left(i; g^{line}\right)$. In relation to $MR\left(i; g_{ij} \to g^{line} - ij\right)$, the total cost for establishing the link is $MC\left(i; g_{ij} \to g^{line} - ij\right)$, which we now identify a lower bound. Beside the cost c, i's productivity drops from $\sigma_1 = 1$ to σ_2 if he establishes g_{ij}. Therefore, the lower bound $MC\left(i; g_{ij} \to g^{line} - ij\right)$ is $\underline{MC} = c + (\sigma_1 - \sigma_2) = c + (1-\sigma_2)$. Because $\sigma_2 \leq \frac{1}{2}$, $\underline{MC} > \frac{1}{2}$ but $MR\left(i; g_{ij} \to g^{line} - ij\right) \leq \frac{1}{2}$. Therefore, $MR\left(i; g_{ij} \to g^{line} - ij\right) < MC\left(i; g_{ij} \to g^{line} - ij\right)$. Applying Lemma 1 to this inequality, we conclude that $U\left(i; g^{line} - ij\right) < U\left(i; g^{line}\right) = U(i; g)$. Hence, i is strictly better off by deleting g_{ij}.

Step 3: Equilibrium characterizations. *By step 2, a nash network consists of components that is either type A, B or C. Thus, this step's goal is to shown exactly which combination of these three kinds of components composes a Nash Network for a given c and σ_2. To achieve this goal, it is neccessary to* *identify exactly which combination gives an existence of an agent who has a positive deviation and which does not. Because checking all such deviations can be cumbersome, our strategy is the following. We first identify all kinds of deviations that are never positive, regardless of c and σ_2. This allows us to pin down only some deviations that need to be checked by quantifying the payoff of the deviating agent.*

First, choose any network whose each component is either type A, B or C. We will show that any deviation such that the deviating agent has more than one link is never positive. Consider a deviation in which the deviating agent has two links. Because all other agents in the network have at most two links, the deviation makes him a part of a line or a cycle. If he is in a cycle, in step 2 we show that the agent's payoff is strictly higher if he removes the link and remain in a line. Similarly, if he is in a line, in step 2 we show that his payoff is stricly higher if he removes the link, breaking the line into two disconnected ones. Therefore, applying step 2, any deviation such that the deviating agent has two links is never positive. Now consider, instead, a deviation such that the deviating agent has more than two links. In step 1, we have shown that such deviation yields the zero revenue to himself because his productivity is 0 *while he has to pay for the link establishment cost. Therefore, any deviation such that the deviating agent has more than one link is never positive.*

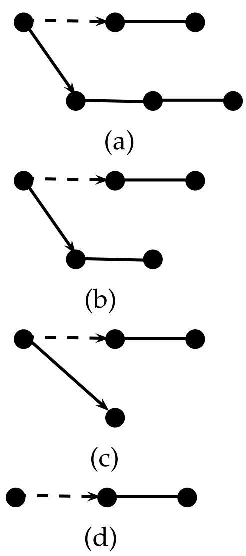
we now further eliminate some combinations that are not Nash for some given parameters. For $c < 1$, any combination that has more than one component that is type C is not nash. Type C component is an isolated agent. Let the isolated agents be i and j. If i access j, his payoff is $1 + (1 - c)$. If i does not access j, he remains isolated and his payoff is 1. *Therefore, if $c < 1$ and there are more than one isolated agents (two type C components), it is not nash because one isolated agent will deviate by accessing another. Hence, there is at most one type C component in a Nash network if $c < 1$.*

Using these results, we list all the networks that remain candidates for Nash Networks below. All of them are combinations whose each component is type A, B or C

1. *all components are C (only for $c \geq 1$)*
2. *at least one A, at least one B, exactly one C*
3. *at least one A, at least one B, no C*
4. *all A, no B, exactly one C*
5. *all A, no B, no C*
6. *all B, no C*
7. *all B, exactly one C*

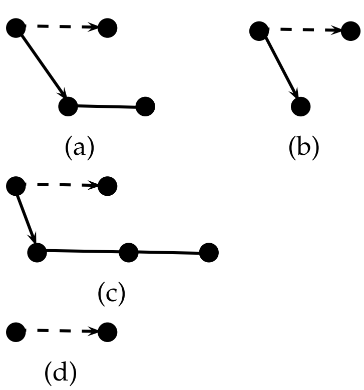
We now list all possible deviations in these combinations. Exploiting the above result that a deviation such that the deviating agent has more than one link is never positive, Figure 10, 11, 12 show all possible *one-link deviations which orginate from an agent in type A, B and C component respectively. In Figure 10, 11, The first three deviations are such that the agent removes himself from the component and access another component. The dotted line shows the link that is removed and the arrow-headed link shows the decision to access another component. The last deviation is the deviation such that the agent removes himself from the component and becomes isolated. In Figure 12, because type C component is an isolated agent, there are three possible deviations - accessing type A, B and C component.*

Finally, because a combination in the list is Nash if and only if there is no positive deviation, whether the combination is nash depends on whether the existence of any deviation that gives a higher payoff than the payoff from the no-deviating strategy, given c and σ_2. To identify the existence, tables in Figure 10, 11, and 12 compare the payoff of the deviating agent when he does not deviate with the payoff when he deviates. Therefore, for a given a set of parameters c and σ_2, applying this table to identify whether each combination in the above list is Nash gives us the set of Nash Networks, which completes the proof.



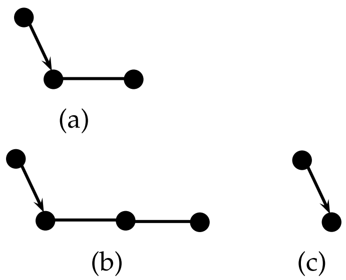
Deviation	from-deviation payoff	no-deviation payoff
(a)	$2\sigma_2^2 + \sigma_2 + 1 - c$	$2\sigma_2 + 1 - c$
(b)	$2\sigma_2 + 1 - c$	$2\sigma_2 + 1 - c$
(c)	$2 - c$	$2\sigma_2 + 1 - c$
(d)	1	$2\sigma_2 + 1 - c$

Figure 10: Deviations in type A component



Deviation	from-deviation payoff	no-deviation payoff
(a)	$2\sigma_2 + 1 - c$	$2 - c$
(b)	$2 - c$	$2 - c$
(c)	$2\sigma_2^2 + \sigma_2 + 1 - c$	$2 - c$
(d)	1	$2 - c$

Figure 11: Deviations in type B component



Deviation	from-deviation payoff	no-deviation payoff
(a)	$2\sigma_2 + 1 - c$	1
(b)	$2\sigma_2^2 + \sigma_2 + 1 - c$	1
(c)	$2 - c$	1

Figure 12: Deviations in type C component

Proof (Proof of Proposition 2)**.** *The proof is divided into two parts according to the two properties in the proposition.*

Part 1: If j accesses i, j is the only end node of i. *We prove by the following contradiction. Suppose there exists another end node k of i in a minimal network g, this end node has a postive deviation by removing the link with i and establish a link with the end node j instead. Formally, let g' be the network with such deviation of k and g^c be the component that k and j reside. Observe that k spends only one unit of link establishment cost c both in g and g'. Therefore, to show that $U(k;g) < U(k;g')$, it suffices to show that $Rev(k;g) < Rev(k;g')$ in what follows.*

Consider $Rev(k;g)$. We can decompose $Rev(k;g)$ in the following way. While all the information of every agent in the component g^c flows to k via g_{ik}, we can partition all agents in g^c in to three sets and will decompose $Rev(k;g)$ accordingly. Let N_1, N_2 and N_3 be the three sets. Let N_1 consists only the agent j, N_2 consists of all end nodes of i except j and k, and N_3 consists of the rest of the agents in g^c except k. Moreover, let $V_k(N_p;g)$ and $\bar{V}_k(N_p;g')$, be the sum of $\bar{V}_{kl}(k;g)$ and $V_{kl}(k;g)$, where $l \in \{N_p\}$. We express $V_k(N_1;g)$ and $V_k(N_2;g)$ as follows.

$$V_k(N_1;g) = \sigma_1(j)\,\sigma_{\mu_i(g)}\sigma_1(k)(i)$$

$$V_k(N_2;g) = (\mu_i(g) - 3)\,\sigma_1\sigma_{\mu_i(g)}\sigma_1(k)$$

Therefore,

$$Rev(k;g) = \sigma_{\mu_i(g)} + (\mu_i(g) - 3)\,\sigma_{\mu_i(g)} + V_k(N_3;g)$$

Similarly, we express $V_k(N_1;g')$ and $V_k(N_2;g')$ as follows.

$$V_k(N_1;g') = \sigma_2(j)\,\sigma_1(k)$$

To express $V_k(N_2;g')$, we use the fact that $\sigma(i;g') = \sigma_{\mu_i(g)-1}$ because k removes the link he has with i in g.

$$\begin{aligned} V_k(N_2;g') &= \sigma_2(j)\,(\mu_i(g) - 3)\,\sigma(i;g')\,\sigma_1(k) \\ &= \sigma_2(j)\,(\mu_i(g) - 3)\,\sigma_{\mu_i(g)-1}\sigma_1(k) \end{aligned}$$

Therefore,

$$Rev\left(k;g'\right) = \sigma_2 + \sigma_2\sigma_{\mu_i(g)-1}\left(\mu_i\left(g\right) - 3\right) + V_i\left(N_3;g'\right)$$

In the expression of $Rev\left(k;g'\right)$ *and* $Rev\left(k;g\right)$ *above, the only terms left unexpressed are* $V_k\left(N_3;g'\right)$ *and* $V_k\left(N_3;g\right)$*. Therefore, in order to check whether* $Rev\left(k;g\right) < Rev\left(k;g'\right)$ *and finish the proof, we need to find a relation between* $V_k\left(N_3;g'\right)$ *and* $V_k\left(N_3;g\right)$*. This relation is expressed below.*

$$V_k\left(N_3;g'\right) = V_k\left(N_3;g\right)\frac{1}{\sigma_{\mu_i}\left(g\right)}\sigma_{\mu_i(g)-1}\sigma_2\left(j\right) \tag{1}$$

We now explain how the relation is derived. Consider how information of an agent in the set N_3 *flows to k in g and* g'*. The flows of information, in both networks, have to pass through i. The only two differences in the flow, however, are the followings. While the information flows to i and finally reaches k in g, it flows to i, then j and finally k in* g'*. Hence, the term* $\sigma_2\left(j;g'\right)$ *apprears on the right hand side of the expression. Moreover, the productivity of i in g is* $\sigma_{\mu_i(g)}$ *while it is* $\sigma_{\mu_i(g)-1}$ *in* g' *because k removes his link that he establishes with i in g. This change in productivty of i appears as the term* $\frac{1}{\sigma_{\mu_i}(g)}\sigma_{\mu_i(g)-1}$*. By these two differences in the flow of information of agents in the set* N_3*, we have the above relation.*

Having found the above relation, to finish the proof it remains to show that $Rev\left(k;g'\right) > Rev\left(k;g\right)$*.*

$$Rev\left(k;g'\right) = \sigma_2 + \sigma_2\sigma_{\mu_i(g)-1}\left(\mu_i\left(g\right) - 3\right) + V_i\left(N_3;g\right)\frac{1}{\sigma_{\mu_i}\left(g\right)}\sigma_{\mu_i(g)-1}\sigma_2 \tag{2a}$$

$$Rev\left(k;g\right) = \sigma_{\mu_i(g)} + \left(\mu_i\left(g\right) - 3\right)\sigma_{\mu_i(g)} + V_i\left(N_3;g\right) \tag{2b}$$

Therefore, $Rev\left(k;g'\right) > Rev\left(k;g\right)$ *if and only if*

$$\begin{aligned}&\sigma_2 + \sigma_2\sigma_{(\mu_i(g)-1)}\left(\mu_i\left(g\right) - 3\right) + V_i\left(N_3;g\right)\frac{1}{\sigma_{\mu_i}\left(g\right)}\sigma_{\mu_i(g)-1}\sigma_2 \\ &\quad > \sigma_{\mu_i(g)} + \left(\mu_i\left(g\right) - 3\right)\sigma_{\mu_i(g)} + V_i\left(N_3;g\right)\end{aligned} \tag{2c}$$

Rearranging the inequality,

$$\begin{aligned}&V_i\left(N_3;g\right)\left(\frac{1}{\sigma_{\mu_i}\left(g\right)}\sigma_{\mu_i(g)-1}\sigma_2 - 1\right) + \left(\sigma_2 - \sigma_{\mu_i(g)}\right) + \left(\mu_i\left(g\right) - 3\right)\left(\sigma_{\mu_i(g)-1}\sigma_2 - \sigma_{\mu_i(g)}\right) \\ &\quad > 0\end{aligned} \tag{2d}$$

Observe that the second term on the left-hand side is positive. Moreover, The first and the third are also postive because $\frac{\sigma_{\mu_i(g)-1}}{\sigma_{\mu_i(g)}} > \frac{\sigma_2}{\sigma_1}$ *due to the assumption that* σ *is concave. Because all the terms on the left-hand side is positive, the inequality above is valid and* $Rev\left(k;g'\right) > Rev\left(k;g\right)$*.* □

Part 2: If i accesses j, i accesses all his end nodes. Morover, if he has more than one end node, he accesses all his neighbors, including ones that are not end nodes.

For the first part, that i accesses all his end nodes, the proof is identical to the proof of Part 1. That is, if k is an end node that accesses i in g, k has a positive deviation by removing the link with i and accesses j instead.

We now consider the second part, that if he has more than one end node, he accesses all his neighbors. The proof follows the same strategy as the proof of Part 1. Formally, let o be a neighbor that is not an end node of i. Suppose, by contradiction, that o accesses i in the network g. Our goal is to show that o has a positive deviation by deleting the link with i and accesses j. Let g' *be the network with such deviation of o. As with the proof of Part 1, by Lemma 1 it suffices to show that* $Rev\left(o;g\right) < Rev\left(o;g'\right)$ *in what follows.*

Let N_1 *be the set of all agents, except o and j, in the component that o belongs. Let* $V_o\left(N_1;g\right) = \sum_{p\in N_1}\bar{V}_{op}\left(g\right)$ *and* $V_o\left(N_1;g'\right) = \sum_{p\in N_1}\bar{V}_{op}\left(g'\right)$*. We express* $Rev\left(o;g'\right)$ *and* $Rev\left(o;g\right)$ *below.*

$$Rev\left(o;g\right) = \sigma_1\left(j\right)\sigma_{\mu_i(g)}\left(i\right)\sigma\left(o;g\right) + V_o\left(N_1;g\right) + \sigma\left(o;g\right)$$

$$Rev\left(o;g'\right) = \sigma_2\left(j\right)\sigma_{\mu_i(g)-1}\left(i\right)\sigma\left(o;g'\right) + V_o\left(N_1;g'\right) + \sigma\left(o;g\right)$$

, where that $\sigma\left(o;g\right) = \sigma\left(o;g'\right)$ *because o deviates from g to* g' *by accessing j instead of i.*

Hence, to show that $Rev\left(o;g\right) < Rev\left(o;g'\right)$ *, it remains to find a relation between* $V_o\left(N_1;g'\right)$ *and* $V_o\left(N_1;g\right)$*. This relation is expressed below using the same reasoning for the Equation 1 of Part 1.*

$$V_o\left(N_1;g'\right) = V_o\left(N_1;g\right)\frac{1}{\sigma_{\mu_i}\left(g\right)}\sigma_{\mu_i(g)-1}\sigma_2\left(j\right)$$

Having found this relation, we apply a calculation similar to those in the equations 2 to show that $Rev\left(o;g\right) < Rev\left(o;g'\right)$*, which finishes the proof.* □

Proof (Proof of Proposition 3)**.** *We prove by contradiction. Let* g' *and* g'' *be components that contain a parent who supports all his neighbors. Let i and k be such parents in* g' *and* g'' *respectively. Let j and l be end nodes of i and k. Our proof aims to show that i has an incentive to remove the link with j and accesses l instead. Our proof is composed of three cases: (1) k has only one neighbor, (2) k has exactly two neighbors, (3) k has more than two neighbors.*

Case 1 *Let* $g = g_i \oplus g_{-i}$*. We introduce another strategy of i,* $\tilde{g}_i = \left\{g_i \backslash \left\{g_{ij}\right\}\right\} \cup \left\{g_{il}\right\}$*, which is* g_i *with the only modification that i accesses l instead of j. Let* $\tilde{g} = \tilde{g}_i \oplus g_{-i}$*. With these notations, in what follows we show that* g_i *is not a best response of i to* g_{-i} *because* $U_i\left(\tilde{g}_i \oplus g_{-i}\right) > U_i\left(g_i \oplus g_{-i}\right)$*.*

To show this inequality, by lemma 1, it suffices to show that $MR\left(i; g_{il} \to (g-ij)\right) >$ $MR\left(i; g_{ij} \to (g-ij)\right)$*. This is because*

$$\begin{aligned} MC\left((g-ij)+il\right) &= MC\left((g-ij)+ij\right) \\ &= c + \left(\sigma_{\mu_i(g-ij)} - \sigma_{\mu_i(g-ij)+1}\right)\left(1 + \sum_{q \in N_i(g'-ij)\setminus i} \bar{V}_{iq}\left(g'-ij\right)\frac{1}{\sigma_{\mu_i(g-ij)}}\right). \end{aligned}$$

That is, since the margical cost of accessing l and j are equal given the network $g-ij$*, we can compare only the benefits that i obtains from accessing l and j.* $MR\left(i; g_{il} \to (g-ij)\right)$ *and* $MR\left(i; g_{ij} \to (g-ij)\right)$ *are expressed below.*

$$\begin{aligned} MR\left(i; g_{ij} \to (g-ij)\right) &= \sigma\left(i; g\right)\sigma_1\left(j; g\right) \\ &= \sigma\left(i; g\right) \end{aligned}$$

$$\begin{aligned} MR\left(i; g_{il} \to (g-ij)\right) &= \sigma\left(i; (g-ij)+il\right)\sigma_2\left(l; (g-ij)+il\right) \\ &\quad + \sigma\left(i; (g-ij)+il\right)\sigma_2\left(l; (g-ij)+il\right)\sigma_1\left(k; (g-ij)+il\right) \\ &= \sigma\left(i; (g-ij)+il\right)\left(2\cdot\sigma_2\right) \end{aligned}$$

Because $\sigma\left(i; g\right) = \sigma\left(i; (g-ij)+il\right)$, $MR\left(i; g_{il} \to (g-ij)\right) > MR\left(i; g_{ij} \to (g-ij)\right)$

Case 2 *Let m be a neighbor of k and* $m \neq l$*. As in case 1,* $MC\left(i; g_{il} \to (g-ij)\right) = MC\left(i; g_{ij} \to (g-ij)\right)$, $MR\left(i; g_{ij} \to (g-ij)\right) = \sigma\left(i; g\right)\sigma_1\left(i; g\right)$*. Therefore, it remains to quantify* $MR\left(i; g_{il} \to (g-ij)\right)$ *and verify that* $MR\left(i; g_{il} \to (g-ij)\right) > MR\left(i; g_{ij} \to (g-ij)\right)$*.*

We now quantify a lower bound of $MR\left(i; g_{il} \to (g-ij)\right)$*. Consider the decision of i to access l when he faces* $g-ij$*. In* $g-ij$*, i is not in component* g'' *but l is. Therefore, as he accesses l, he uses l to retrieve information of every agent in* g''*. Hence,* $MR\left(i; g_{il} \to (g-ij)\right)$ *can be expressed as*

$$\begin{aligned} &MR\left(i; g_{il} \to (g-ij)\right) = \\ &\quad \sigma\left(i; \tilde{g}\right)\sigma\left(l; \tilde{g}\right) + \sigma\left(i; \tilde{g}\right)\sum_{q \in g''; q \neq l} \bar{V}_{lq}\left(\tilde{g}\right) \end{aligned} \tag{3}$$

The first term is the value of l's own information that arrives to i. The second term is the value of information of every node in g''*, apart from l, that travels to i. Observe that the term* $\sum_{q \in g''; q \neq l} \bar{V}_{lq}\left(\tilde{g}\right)$ *is* $\sum_{q \in g''; q \neq l} V\left(\bar{P}_{lq}\left(\tilde{g}\right)\right)$*, where* $\bar{P}_{lq}$ *is an optimal lq-path, which is unique because* $\tilde{g}$ *is minimally connected. Consider how information flows in* $\bar{P}_{lq}\left(\tilde{g}\right)$ *for* $q \in g''$*. Information of any q flows to k, then k passes to l via the link* g_{kl}*. Therefore, the term* $\sum_{q \in g''; q \neq l} \bar{V}_{lq}\left(\tilde{g}\right)$ *can be expressed as*

$$\sum_{q\in g'';q\neq l}\bar{V}_{lq}\left(\tilde{g}\right)=$$
$$\sigma\left(l;\tilde{g}\right)+\sigma\left(l;\tilde{g}\right)\sum_{q\in g'';q\neq l}\bar{V}_{kq}\left(\tilde{g}\right) \tag{4}$$

Therefore, to quantify a lower bound of $MR\left(i;g_{il}\rightarrow\left(g-ij\right)\right)$*, we turn to find a lower bound of* $\sum_{q\in g'';q\neq l}\bar{V}_{kq}\left(\tilde{g}\right)$*. Observe that* $\bar{V}_{kq}\left(\tilde{g}\right)=\bar{V}_{kq}\left(g\right)$ *because the addition of link* g_{il} *changes only the relation between i and l, and l is not in any optimal kq-path. To identify* $\bar{V}_{kq}\left(g\right)$*, consider the payoff of k in g that is expressed below,*

$$\begin{aligned}U_k\left(g\right)&=Rev\left(k;g\right)-\bar{\mu}_k\left(g\right)\cdot c\\&=\sum_{q\in g}\bar{V}_{kq}\left(g\right)-\bar{\mu}_k\left(g\right)\cdot c\end{aligned} \tag{5a}$$

Because k has a link only with those in g''*, substitute g with* g''*,*

$$\begin{aligned}U_k\left(g\right)&=\sum_{q\in g''}\bar{V}_{kq}\left(g''\right)-\bar{\mu}_k\left(g''\right)\cdot c\\&=\sum_{q\in g'';q\neq l}\bar{V}_{kq}\left(g\right)+\sigma\left(l;g''\right)\sigma\left(k;g''\right)-\bar{\mu}_k\left(g''\right)\cdot c\\&=\sum_{q\in g'';q\neq l,k}\bar{V}_{kq}\left(g\right)+\sigma\left(l;g''\right)\sigma\left(k;g''\right)+\sigma\left(k;g''\right)-\bar{\mu}_k\left(g''\right)\cdot c\end{aligned} \tag{5b}$$

Because we assume that $\mu_k\left(g''\right)=2$*, k establishes exactly two links with l and m. Therefore, for any agent* $q\neq k,l$*, k receives q's information via* g_{km}*. The total value of information that flows via* g_{km} *is* $\sum_{q\in g'';q\neq l}\bar{V}_{kq}\left(g\right)$*, which appears in the last line above. Moreover, this strategy* g_k *is k's best response to* g_k*. That he maintains exactly two links, together with that this strategy is his best response, are used to to identify the following lower bound of* $\sum_{q\in g'';q\neq l}\bar{V}_{kq}\left(g\right)$*.*

$$U_k\left(g\right)>U_k\left(g-km\right) \tag{6a}$$

$$U_k\left(\left(g-km\right)+km\right)>U_k\left(g-km\right) \tag{6b}$$

$$MR\left(k;g_{km}\rightarrow\left(g-km\right)\right)>MC\left(k;g_{km}\rightarrow\left(g-km\right)\right) \tag{6c}$$

$$\sum_{q\in g'';q\neq l,k}\bar{V}_{kq}\left(g\right)>c+\left(\sigma_1-\sigma_2\right)\left(\sigma_1\left(k;g-km\right)+\sigma_1\left(k;g-km\right)\sigma_1\left(l;g-km\right)\right) \tag{6d}$$

$$\sum_{q\in g'';q\neq l,k}\bar{V}_{kq}\left(g\right)>c+\left(1-\sigma_2\right)\left(2\right) \tag{6e}$$

Thus, we have found the lower bound of $\sum_{q\in g'';q\neq l,k}\bar{V}_{kq}(g)$ in the last line of the above expression. To finish the proof, we use it to show that $MR(i;g_{il}\to(g-ij)) > MC\left(i;g_{ij}\to(g-ij)\right)$

$$MR(i;g_{il}\to(g-ij)) > MR\left(i;g_{ij}\to(g-ij)\right) \tag{7}$$

$$\textit{if and only if} \tag{8}$$

$$\sigma(i;\tilde{g})\,\sigma(l;\tilde{g}) + \sigma(i;\tilde{g})\sum_{q\in g'';q\neq l}\bar{V}_{lq}(\tilde{g}) > \sigma(i;g)\,\sigma(j;g) \tag{9}$$

$$\textit{Because } \sigma(i;\tilde{g}) = \sigma(i;g)\,, \textit{and } \sigma(j;g) = 1$$

$$\sigma(l;\tilde{g}) + \sum_{q\in g'';q\neq l}\bar{V}_{lq}(\tilde{g}) > 1 \tag{10}$$

We now show that the last expression above is valid.

$$\sigma(l;\tilde{g}) + \sum_{q\in g'';q\neq l}\bar{V}_{lq}(\tilde{g}) > 1 \tag{11}$$

$$\sigma(l;\tilde{g}) + \sigma(l;\tilde{g})\,\sigma(k;\tilde{g}) + \sigma(l;\tilde{g})\sum_{q\in g'';q\neq l,k}\bar{V}_{kq}(\tilde{g}) > 1 \tag{12}$$

$$\sigma_2 + \sigma_2\sigma_2 + \sigma_2\sigma_2\sum_{q\in g'';q\neq l,k}\bar{V}_{kq}(\tilde{g}) > 1 \tag{13}$$

The last inequality is satisfied because $\sigma_2 > \frac{1}{2}$ and that $\sum_{q\in g'';q\neq l,k}\bar{V}_{kq}(g) = \sum_{q\in g'';q\neq l,k}\bar{V}_{kq}(\tilde{g}) > c + (1-\sigma_2)(2)$. Therefore, $MR(i;g_{il}\to(g-ij)) > MR\left(j;g_{ij}\to(g-ij)\right)$ ☐

case 3 *The proof follows the same analogy as that of Case 2. The only difference is that the term $\sum_{q\in g'';q\neq l,k}\bar{V}_{kq}(g)$ in inequality 6e is changed because $\bar{\mu}_k(g) = \mu_k(g) > 2$. Despite this change, in what follows we show that inequality 6e, $\sum_{q\in g'';q\neq l,k}\bar{V}_{kq}(g) > c + (1-\sigma_2)$, remains valid so that the rest of the inequalities in the proof of Case 2 are still satisfied. In what follows, parent k is assumed to have at most one neighbor m that is not an end node.*

We provide the proof that guarantees the existence of such parent in any minimally connected network g before proceeding. Suppose, by contradiction, that each parent has at least two neighbors that are not an end node. Remove all end nodes in the network and all links with them. We have a modified network $\hat{g}$. Observe that $\hat{g}$ is minimally connected because g is minimally connected. Observe also that $\hat{g}$ has no end node because each node in it has at least two links. However, if there is no end node, $\hat{g}$ is not *minimally connected* [9]. *A contradiction.*

Because m is the only neighbor that is not an end node of k, k observes any non-neighbor q in g'' via the link g_{km}. Formally, let N_{km} be the set of all such agents and $N^c_{km} = \{q \in g' | q \notin N_{km} \wedge q \neq k, l\}$. We can express $\sum_{q \in g''; q \neq l,k} \bar{V}_{kq}(g)$ as

$$\sum_{q \in g''; q \neq l,k} \bar{V}_{kq}(g) = \sum_{q \in N_{km}} \bar{V}_{kq}(g) + \sum_{q \in N^c_{km}} \bar{V}_{kq}(g) \tag{14}$$

We now show that $\sum_{q \in N_{km}} \bar{V}_{kq}(g) > c + (1 - \sigma_2)(2)$ to finish the proof. Because $\sum_{q \in N_{km}} \bar{V}_{kq}(g)$ is exactly $MR(k; g_{km} \to (g - km))$ and that k's strategy to maintain the links with all his neighbors is his best response in g, it follows that

$$MR(k; g_{km} \to (g - km)) > MC(k; g_{km} \to (g - km)) \tag{15}$$

$$\sum_{q \in N_{km}} \bar{V}_{kq}(g) > c + \left(\sigma_{\mu_k(g)} - \sigma_{\mu_k(g)+1}\right)(1 + (\mu_k(g) - 2)) \tag{16}$$

Moreover,

$$c + \left(\sigma_{\mu_k(g)} - \sigma_{\mu_k(g)+1}\right)(1 + (\mu_k(g) - 2)) > c + (1 - \sigma_2)(2) \tag{17}$$

The last inequality is satisfied because $\mu_k(g) \geq 3$ and $\left(\sigma_{\mu_k(g)} - \sigma_{\mu_k(g)+1}\right) > (\sigma_1 - \sigma_2)$ due to the concavity of σ, we conclude that $\sum_{q \in N_{km}} \bar{V}_{kq}(g) > c + (1 - \sigma_2)(2)$. □

Proof (Proof of Proposition 4)**.** *The first part is a corollary of the second proposition. For the second part, let i be the center who estaliblishes all the links and j be a neighbor of i and g^* be a center-sponsored star. Our proof consists of two parts. The first one shows that the inequality on the left is a neccessary and sufficient condition for i to maintain all his links in g. The second one shows that the inequality on the right is a neccessary and sufficient condition for j to sponsors no link in g. Both parts rely on the strict concavity of σ.*

[9] any minimal non-empty network has at least one end node. The proof is as follows. Suppose not. Let every node has at least two links but the network remains minimal. Choose any p and q that are connected. Because q is has at least two links, q is connected to q' who is not p, otherwise the network is not minimal. Moreover, q' also is connected to q'' who is neither q' nor p nor q for the same reason. This induction repetes infinitely because every node has at least two links. It follows that this network has infinite nodes. A contradiction.

Part 1: The neccessary and sufficient condition for i to maintain all his links in g^* is $\sigma_{n-1} - (n-1)(\sigma_{n-2} - \sigma_{n-1}) \geq c$. *We first show the neccessary condition. If i's best response is to maintain all his links, then he does not want to deviate from g^* by deleting one link, i.e., $U_i(g^*) - U_i(g^* - ij) \geq 0$ or $MR(i; g_{ij} \to (g^* - ij)) - MC(i; g_{ij} \to (g^* - ij)) \geq 0$ by Lemma 1. Because $MR(i; g_{ij} \to (g^* - ij)) = \sigma_n(i)\sigma_1(j)$ and $MC(i; g_{ij} \to (g^* - ij)) = c + (\sigma_{n-2} - \sigma_{n-1})(n-1)$, $\sigma_{n-1} - (n-1)(\sigma_{n-2} - \sigma_{n-1}) > c$ is the neccessary condition.*

To show the sufficient condition, our goal is to show that if the sufficient condition $\sigma_{n-1} = MR(i; g_{ij} \to (g^* - ij)) \geq c + (n-1)(\sigma_{n-2} - \sigma_{n-1}) = MC(i; g_{ij} \to g^*)$ holds, i's best response to g^*_{-i} is to sponsor $n-1$ links to all agents. Let $g^k = g^*_{-i} \cup \{g_{i,j_1}, g_{i,j_2}, ..., g_{i,j_k}\}$ be a situation in which i responses g^*_{-i} by sponsoring k links. To finish the proof, we now show that $U_i(g^* = g^{n-1}) \geq U_i(g^{n-2}) \geq \ldots \geq U_i(g^*_{-i})$, given the sufficient condition.

By lemma 1, this is equivalent to prove that $MR(i; g_{ik} \to g^{k-1}) \geq MC(i; g_{ik} \to g^{k-1})$ for all k. Observe that $MR(i; g_{ik} \to g^{k-1}) = \sigma_k(i)\sigma_1$ and $MC(i; g_{ik} \to g^{k-1}) = c + (\sigma_k - \sigma_{k-1})(1 + (k-1))$. Moreover, $MR(i; g_{ik} \to g^{k-1})$ is strictly decreasing in k while $MC(i; g_{ik} \to g^{k-1})$ is strictly increasing in k because $\sigma_k < \sigma_{k-1}$ and $(\sigma_k - \sigma_{k-1}) > (\sigma_{k-1} - \sigma_{k-2})$ due to the strict concavity of σ. Therefore, if $MR(i; g_{ij} \to (g - ij)) \geq MC(i; g_{ij} \to (g - ij))$, then $MR(i; g_{ik} \to (g^{k-1})) = \sigma_k \geq MR(i; g_{in-1} \to (g^{n-2})) \geq MC(i; g_{in-1} \to (g^{n-2})) \geq MC(i; g_{ik} \to (g^{k-1}))$; $\forall k \leq n-2$. Hence, $U_i(g^* = g^{n-1}) \geq U_i(g^{n-2}) \geq \ldots > U_i(g^*_{-i})$.

Part 2: The neccessary and sufficient condition for j to maintain all his links in g^* is $c \geq (\sigma_2)^2 - \sigma_{n-1} - (1 - \sigma_2)(1 + (n-2)\sigma_{n-1})$. *Both the neccessary and sufficient condition will be proved using the analogy of Part 1. The only difference is that this inequality motivates j to maintain no link, while $c \leq \sigma_{n-1} - (n-1)(\sigma_{n-2} - \sigma_{n-1})$ motivates i to maintain all the links.*

For the neccessary condition, let $j' \neq j$ be a neighbor of i. We will show that the inequality that causes j to add no link to j' is exactly $c \geq (\sigma_2)^2 - \sigma_{n-1} - (1 - \sigma_2)(1 + (n-2)\sigma_{n-1})$, making this inequality a neccessary condition. By Lemma 1, $U_i(g^* + jj') - U_i(g^*) \leq 0$ if and only if $MR\left(j; g_{jj'} \to g^*\right) \leq MC\left(j; g_{jj'} \to g^*\right)$. Observe that $MR\left(j; g_{jj'} \to g^*\right) = \sigma_2(j; g^* + jj')\sigma_2(j'; g^* + jj') - \sigma_{n-1}(i; g^*)$ $\sigma_1(i; g^* + jj')$ because j benefits from obtaining the information of j' via the new link instead of obtaining it indirectly via i. Moreover, $MC\left(j; g_{jj'} \to g^*\right) = c + (\sigma_1(i; g^*) - \sigma_2(i; g^* + jj'))(1 + \sigma_{n-1}(n-2))$, where the term $\sigma_{n-1}(n-2)$ is what j receives via i except the information of j'. Therefore, $MR\left(j; g_{jj'} \to g^*\right) - MC\left(j; g_{jj'} \to g^*\right) = (n-1)(\sigma_{n-2} - \sigma_{n-1}) + c - \sigma_{n-1} \leq 0$ is the neccessary condition.

Next, let us consider the sufficient condition. In g^*, j has to access all the nodes via i. Thus, his deviation is to add some links with other neighbors in order that he can reach them directly. We will find that such deviation is never profitable if the sufficient condition holds. Let $g^m = g^* \cup g_{j,j_1}, g_{j,j_2}, ..., g_{j,j_m}; n-2 \geq m \geq 1$ be a situation in which j adds m links to the nodes that are not i. Now consider the following four terms

- $MR\left(j; g_{jm} \to g^m\right) = \sigma_2\sigma_{m+1} - \sigma_{n-1}\sigma_m$
- $MR\left(j; g_{jj'} \to g^*\right) = \sigma_2 - \sigma_{n-1}$
- $MC\left(j; g_{jm} \to g^m\right) = (\sigma_m - \sigma_{m+1})\left(1 + \sigma_{n-1} + \sigma_{n-1}(n-3-m) + \sigma_2 \cdot m\right)$
- $MC\left(j; g_{jj'} \to g^*\right) = (\sigma_1 - \sigma_2)\left(1 + \sigma_{n-1} + \sigma_{n-1}(n-3)\right)$

Because $\sigma_m < \sigma_{m-1}$ and $(\sigma_1 - \sigma_2) < (\sigma_m - \sigma_{m+1})$, $MR\left(j; g_{jj'} \to g^*\right) > MR\left(j; g_{jm} \to g^m\right)$ and $MC\left(j; g_{jm} \to g^m\right) > MC\left(j; g_{jj'} \to g^*\right)$ for all m. Therefore, the sufficient condition, which is $MR\left(j; g_{jj'} \to g^*\right) \leq MC\left(j; g_{jj'} \to g^*\right)$, implies that $MR\left(j; g_{jm} \to g^m\right) < MR\left(j; g_{jj'} \to g^*\right) < MC\left(j; g_{jj'} \to g^*\right) < MC\left(j; g_{jm} \to g^m\right)$ for all m. By lemma 1, these inequalities imply that $U_j\left(g^*\right) \geq U_j\left(g^* + jj' = g^1\right) > U_j\left(g^2\right) > \cdots > U_j\left(g^{n-2}\right)$. Therefore, j's best response to g_i^* is to maintain no link. □

A.3 Proof of Example 2

Proof (Proof of Example 2)**.** *Any network that has a node with three or more neighbors is not nash because $\sigma_3 = 0$. Therefore, onwards we consider only networks whose every node has at most two neighbors.*

The first kind of networks we eliminate as candidates for nash networks is any network that contains a component that is a cycle of less than 5 players. Let g_{ij} be a link in such cycle. If i deletes g_{ij}, the cycle component becomes a line. Because $MR\left(i; g_{ij} \to (g - ij)\right) < 0$ for any a cycle of less than 5 players, i's payoff in the line is higher than his payoff in the cycle. Therefore, he has an incentive to destroy g_{ij}.

Next, among the networks that remain to be considered, we partition them into eight sets of networks according to the following criterion. For $g \neq g'$, $\bar{g} = \bar{g}'$ if and only if they belong to the same group. Figure 13 depicts all of them. Except group (a), positive deviation in each group can be easily identified. Table 1 summarizes the deviations.

Finally, for the first set, we list all networks in this set in Figure 14 and point out a positive deviation in each of them in 2

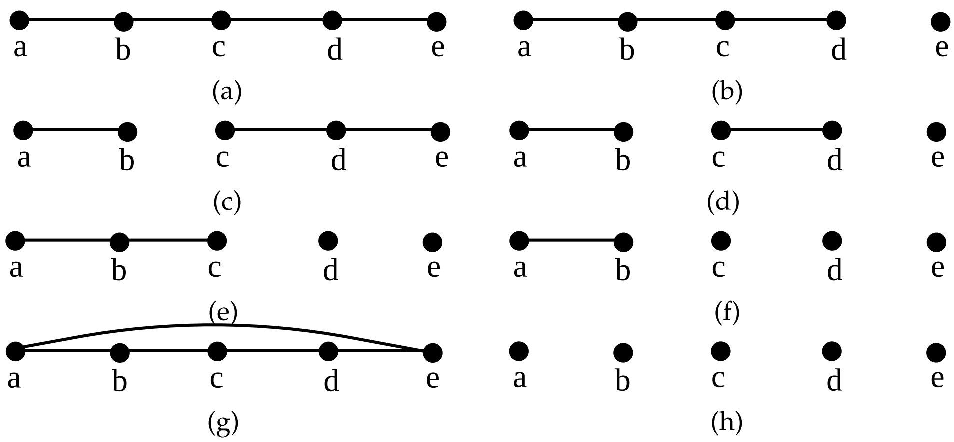


Figure 13: We divide networks into 8 groups according to the following criterion. $\bar{g} = \bar{g}'$ if and only if they belong to the same group.

Network	Deviating Agent	Deviation
(b)	e	access d
(c)	a (b)	destroy g_{ab} (g_{ba}) and make g_{ac} (g_{bc})
(d)	e	access d
(e)	e	access d
(f)	e	access d
(g)	e (a)	destroy g_{ea} (g_{ae})
(h)	e	access d

Table 1: Positive deviations found in each group of networks in 13

Network	Deviating Agent	Deviation
(a)	d	destroy g_{de}
(b)	c	destroy g_{cd}
(c)	c	destroy g_{de}
(d)	d	destroy g_{de}
(e)	d	destroy g_{de}
(f)	d	destroy g_{de}
(g)	d	destroy g_{de}
(h)	b	destroy g_{bc}
(i)	b	destroy g_{bc}
(j)	d	destroy g_{de}
(k)	b	destroy g_{bc}
(l)	c	destroy g_{cb}
(m)	b	destroy g_{ba}
(n)	d	destroy g_{dc}
(o)	b	destroy g_{ba}
(p)	a	destroy g_{ba}

Table 2: Positive deviation found in each network in group (a)

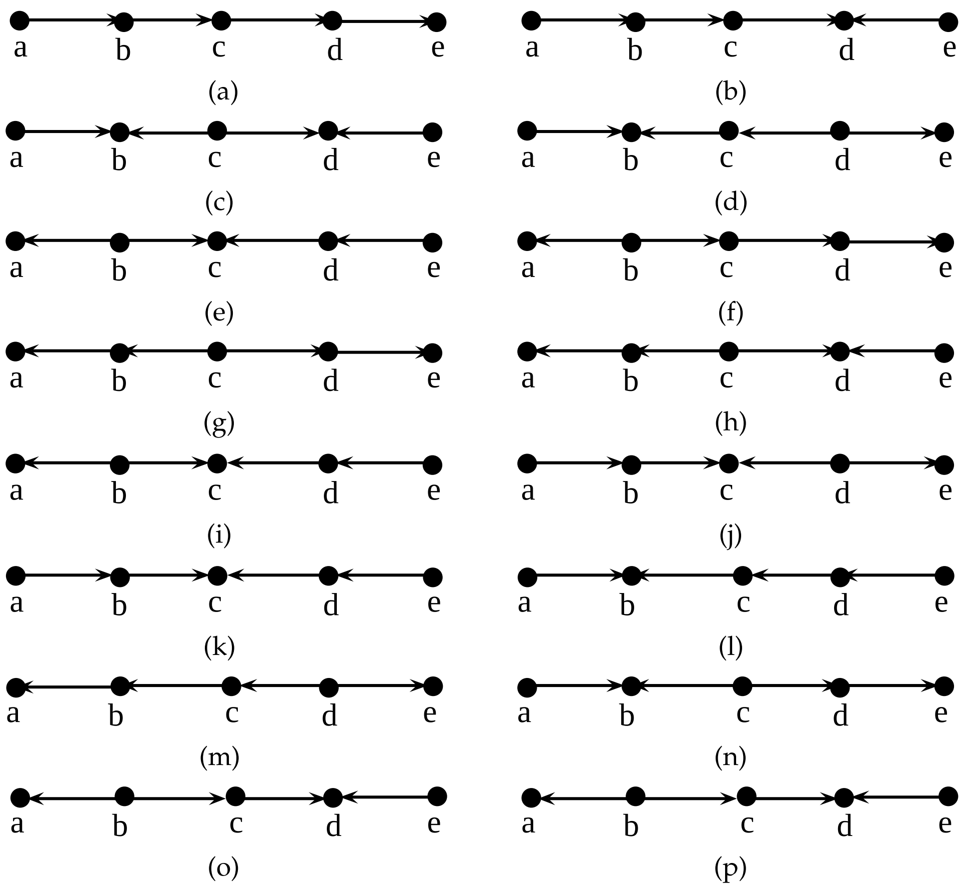


Figure 14: All possible 16 networks in group (a)