Designing efficient networks sequentially

Yang Sun[*] Wei Zhao[†] Junjie Zhou[‡]

November 23, 2021
Preliminary Version!

Abstract

We study the problem of designing efficient network sequentially. In each period, the planner connects two unlinked agents in the network formed in previous period, then the agents play a game with local complementarity under the newly formed network. The planner benefits from the entire discounted stream of equilibrium welfare. We show that, forming a nested split graph in each period is an optimal strategy for the planner for any specific values of discount factors. Moreover, when the planner heavily discounts future welfare, the optimal strategy induces a quasi-complete graph in each period regardless of the strength of complementary effect. Our paper therefore provides a micro-foundation for quasi-complete network since it is formed under greedy algorithm. We also discuss the robustness of these results under non-linear best response and heterogeneous agents.
JEL Classification: D85; C72.

Keywords: Katz-Bonacich centrality; Network formation; Efficiency; Nested split graph; Quasi-complete graph; Greedy algorithm.

1 Introduction

Structure of interactions among economic subjects, represented as a network, crucially determines economic behavior as well as individual's well-being.[1] Since network structure affects economic consequences, it is crucial to know which network induces the best outcome. To be concrete, consider the infrastructure network, such as highways and railroads, connecting city nodes. A vast empirical studies highlighted the irreplaceable contribution of infrastructure network on the

[*]Department of Economics, Sichuan University, China. *Email*: sunyang789987@gmail.com
[†]Department of Economics and Decision Sciences, HEC Paris, France. *Email*: wei.zhao1@hec.edu
[‡]Department of Economics, National University of Singapore, Singapore. *Email*: zhoujj03001@gmail.com
[1]See, for example Calvó-Armengol et al. (2009) on pupil school performance; David and Dina (2004) on brand choice, Bramoullé and Kranton (2007); Allouch (2017) on public goods provision. For recent surveys, see Bramoullé et al. (2016); Jackson et al. (2017); Elliott et al. (2019).

performance of cities.[2] A natural problem for the government is to design routes to maximize performances of cities subject to limited construction budget. This network design problem has two features. First, the network may be formed sequentially due to lack of organizational resource: roads can not be built simultaneously due to limited facilities and labors; Second, the government assigns different weights over the flow of welfare at each period. For instance, myopic ruling governor only cares about performances in his official career, and does not value the benefit of the project after leaving office.

To address this network design problem, we examine a dynamic model of network formation. In each period during network formation process, multiple agents and a single social planner play a two-stage game. In the first stage, the planner connects two unlinked agents in the network formed in the last period. In the second stage, the agents play a game introduced by Ballester et al. (2006) under the newly formed network. The social planner behaves in a farsighted manner and aims at maximizing the discounted sum of social welfare at each period.

Our model allows for various discount factors corresponding to different formation periods. This helps us study the formed network given different levels of farsightedness for the planner. Specifically, we say that the social planner is farsighted if she only cares about the social welfare under the network formed in the last period. If the planner heavily discounts the social welfare in future periods and in convergence, the planner only cares about the social welfare at the current period, then we say that the planner is myopic.

Our main results are in two folds. First, given any discount factor profile for the planner, the formed network will be nested split graph (NSG) at each period. This result shows that NSG is a robust prediction even if the network is formed sequentially and the planner may have different levels of farsightedness. We further show that, this result still holds under convex best response. Second, we show that the formed network is quasi-complete (QC) if the planner is myopic. This result offers a micro-foundation for QC network since it is formed under greedy algorithm.

As immediate corollaries of the main results, we first refine the prediction of globally efficient network, drawn by previous literature, by excluding one subclass of NSGs. Second, we also provide a sufficient condition under which greedy algorithm leads to global optimum.

We discussed the robustness of our main results. First, we can extend our results to the scenario where the social planner's aims at maximizing aggregate efforts. Our results are still true if the social planner strategically picks an agent, who then links to a new neighbour to maximize his current period utility. Moreover, the dynamic model of formation also induces an NSG in each period when the agents have different intrinsic marginal utilities.

A vast of literature analyze network formation problem from different perspectives. Many early works focus on the stability and efficiency of networks given the outcomes of individuals' interactions (for example Jackson and Wolinsky 1996, Jackson and Watts 2002, Dutta and Mutuswami 1997 and Bala and Goyal 2000). The follow ups can be generally categorized in two main classes: 1. action

[2]For example, Duranton et al. (2014) estimate the effect of highways on the city exports. Banerjee et al. (2020) show the positive effect of transportation network on per capital GDP levels across sectors. Atack et al. (2010) demonstrate the positive effects of railroad on urbanization and population growth in the US.

choices endogenously determined by network games, including strategic substitution (Galeotti and Goyal 2010, Billand et al. 2015 and van Leeuwen et al. 2019) and strategic complementarity (Belhaj et al. 2016, Baetz 2015, Hiller 2017 and Li 2020); 2. incorporating dynamics into the formation process (Watts 2001, Dutta et al. 2005 and Song and van der Schaar 2020). Combining these two strands, this paper studies dynamic network formation with endogenous choice of actions. One closely related work is König et al. (2014) where, at each period, a randomly selected agent builds up a new link or eliminate an existing link to maximize his equilibrium payoff at the current period (i.e. being myopic). There are two differences between our paper and theirs. First, we assume that it is the social planner who builds up a new link in each period to maximize social welfare. Second, we allow for different degree of farsightedness for the social planner.

We use the seminal work of Ballester et al. (2006) to model (local) strategic complementarities among agents. A central feature of this model is that the agent's equilibrium action and utility are proportional to the Katz-Bonacich centrality and square of this centrality respectively. This framework was later found to have broad applications for network economics such as price discrimination (Bloch and Quérou 2013), collaboration among firms (König et al. 2019), geographical space (Helsley and Zenou 2014).

In static setting, Belhaj et al. (2016) study optimal networks in which the agents play the same complementary network game. They show that, given a fixed number of links, the optimal network which maximizes both social welfare and aggregate effort is nested split graph (NSG). Li (2020) generalizes this result by allowing for non-linear best response, but he studies the optimal weighted and directed networks. Departing from these works focusing on global efficiency, this paper presents a dynamic formation model which degenerates to (global) efficient problem when the social planner is farsighted. In particular, our results contribute to Belhaj et al. (2016) in three dimensions. First, we show that one subclass of NSGs are strictly dominated. Second, we show that NSG is a robust prediction of formed even if the network is formed sequentially and the planner has different levels of farsightedness. Third, we extend this result to convex best response function. Besides, we also give a counter-example to show that the link reallocation approach developed by Belhaj et al. (2016) does not apply to the case of concave best response function.

The remainder of this paper is organized as follows. Section 2 introduces the framework. We present our main results in Section 3 and discuss the robustness of our results in Section 4. Section 5 concludes. To facilitate reading, all the proofs are relegated to Appendix.

2 The Model

In this section, we introduce a dynamic model of network formation with the objective of equilibrium welfare maximization. Each formation period can be viewed as a two-stage game played by a social planner and a set of agents. In the first stage, the social planner connects two unlinked agents in the network formed in previous periods. In the second stage, all agents play a network game following Ballester et al. (2006) with local complementarities. The optimal link chosen by the social planner in each period is set to maximize the discounted sum of entire stream of equilibrium welfare.

2.1 The network game

The network formed by the set $N = \{1, 2, ..., n\}$ of agents is represented by an adjacency matrix $\boldsymbol{G} = (g_{ij})_{n\times n}$. Assume $g_{ij} = g_{ji} \in \{0, 1\}$ and $g_{ii} = 0$ for all $i \in N$. We use notation $ij \in \boldsymbol{G}$ to indicate that agents i and j are connected and $ij \notin \boldsymbol{G}$ otherwise. Each agent $i \in N$ chooses an effort $a_i \in \mathbb{R}_+$, and let $\boldsymbol{a} \in \mathbb{R}_+^n$ be the profile of agent efforts. Given the network $\boldsymbol{G}$, and effort profile $\boldsymbol{a}$, agent i's payoff is given as follows:

$$u_i(\boldsymbol{a}, \boldsymbol{G}) = a_i - \frac{1}{2}a_i^2 + \phi \sum_{j\in N} g_{ij} a_i a_j. \tag{1}$$

This specific utility form follows from Ballester et al. (2006), where a_i and $\frac{1}{2}a_i^2$ denote intrinsic utility and cost of effort respectively, and the last term $\phi \sum\limits_{j\in N} g_{ij} a_i a_j$ reflects network externalities. The scalar parameter ϕ controls the strength of network effect. We assume $\phi > 0$ so the game exhibits strategic complementarity. We use $\Gamma(\boldsymbol{G})$ to denote the network game represented above.

Denote by $\lambda_{\max}(\boldsymbol{G})$ the largest eigenvalue of matrix $\boldsymbol{G}$. As shown in Ballester et al. (2006), if $0 < \phi < \frac{1}{\lambda_{\max}(\boldsymbol{G})}$, $\Gamma(\boldsymbol{G})$ has a unique Nash equilibrium given by

$$\boldsymbol{a}^*(\boldsymbol{G}) = (\boldsymbol{I} - \phi\boldsymbol{G})^{-1}\boldsymbol{1},$$

where $\boldsymbol{I}$ is the identity matrix and $\boldsymbol{1} = (1, 1, ..., 1)'$ is the vector of 1s (with suitable dimension). That is, each agent's equilibrium effort coincides with the well-known measure of centralities in the network – (unweighted) Katz-Bonacich centrality, which summarizes the total number of paths starting from this node with length discount ϕ in the network.[3] Therefore, the aggregate equilibrium efforts

$$a^*(\boldsymbol{G}) = \sum_{i\in N} a_i^*(\boldsymbol{G}) = \sum_{k=0} \phi^k \boldsymbol{1}'\boldsymbol{G}^k\boldsymbol{1}. \tag{2}$$

Moreover, the equilibrium welfare of game $\Gamma(\boldsymbol{G})$ is given by

$$u^*(\boldsymbol{G}) = \sum_{i\in N} u_i^*(\boldsymbol{G}),$$

where agent i's utility at equilibrium $u_i^*(\boldsymbol{G}) = \frac{1}{2}(a_i^*(\boldsymbol{G}))^2$.[4] The equilibrium welfare can be rewrit-

[3]The path counting explanation of Katz-Bonacich centrality follows from the following identity of the Leontief inverse matrix:

$$(\boldsymbol{I} - \phi\boldsymbol{G})^{-1} = \boldsymbol{I} + \phi\boldsymbol{G} + \phi^2\boldsymbol{G}^2 + ...$$

[4]See Ballester et al. (2006) for details.

ten as weighted sum of aggregate paths with length discount:

$$u^*(\boldsymbol{G}) = \frac{1}{2}\sum_{k=0}(k+1)\,\phi^k \mathbf{1}'\boldsymbol{G}^k\mathbf{1}.^5 \tag{3}$$

Such an elegant relationship between equilibrium outcomes and aggregate paths in the network is the starting point of our analysis.

2.2 The network formation process

The network is designed by the social planner over T periods (indexed by $t = 1, ..., T$, where $T \leq \frac{n(n-1)}{2}$). At each period t, the game is divided into two stages: in the first stage, the social planner intervenes network $\boldsymbol{G}(t-1)$ by adding a new link between two unconnected agents $ij \notin \boldsymbol{G}(t-1)$; then the newly formed network becomes $\boldsymbol{G}(t) = \boldsymbol{G}(t-1) + \boldsymbol{E}_{ij}$, where $\boldsymbol{E}_{ij}$ denotes the matrix with 1 on (i, j) and (j, i) entries, 0 on all the other entries; in the second stage, the agents play network game $\Gamma(\boldsymbol{G}(t))$ and choose equilibrium effort level $\boldsymbol{a}^*(\boldsymbol{G}(t))$. The social planner benefits from social welfare at the end of period.

Before proceeding, we impose the following assumptions.

Assumption 1. The network formation process starts with empty network, i.e., $\boldsymbol{G}(0) = \mathbf{0}$, where $\mathbf{0}$ is the matrix with 0s. Thus, the index t in $\boldsymbol{G}(t)$ also indicates the total number of links in $\boldsymbol{G}(t)$.

Assumption 2. The strength of network effect ϕ is smaller than $\frac{1}{\lambda_{\max}(\boldsymbol{G}(T))}$ to ensure the uniqueness of Nash equilibrium at each period. In particular, by Perron-Frobenius theorem $\lambda_{\max}(\boldsymbol{G}(T)) \leq \max_{i\in N}\left\{\sum_j g_{ij}\right\}$, and therefore Assumption 2 holds when $\phi(n-1) < 1$.

We say network $\hat{\boldsymbol{G}}$ succeeds network $\boldsymbol{G}$ if $\hat{\boldsymbol{G}}$ can be obtained by building up a new link from $\boldsymbol{G}$.

Definition 1. *For any two networks $\hat{\boldsymbol{G}}$ and $\boldsymbol{G}$, $\hat{\boldsymbol{G}}$ is said to succeed $\boldsymbol{G}$, if there exists a pair of agents $\{i, j\}$ such that $ij \notin \boldsymbol{G}$ and $\hat{\boldsymbol{G}} = \boldsymbol{G} + \boldsymbol{E}_{ij}$. Let $\mathbb{S}(\boldsymbol{G}) \equiv \left\{\hat{\boldsymbol{G}} | \hat{\boldsymbol{G}} = \boldsymbol{G} + \boldsymbol{E}_{ij} \text{ for some } ij \notin \boldsymbol{G}\right\}$ denote the set of networks which succeeds $\boldsymbol{G}$.*

A strategy of the social planner, $\boldsymbol{s}$, is a sequence of successive networks $(\boldsymbol{G}(1), ..., \boldsymbol{G}(T))$, which specifies network $\boldsymbol{G}(t)$ in period t. Denote the set of strategies by

$$S \equiv \{(\boldsymbol{G}(1), ..., \boldsymbol{G}(T))\,|\boldsymbol{G}(t) \in \mathbb{S}(\boldsymbol{G}(t-1)), \forall t = 1, ..., T\}.$$

Given a strategy of planner $\boldsymbol{s} = (\boldsymbol{G}(t))_{t=1}^T \in S$, the planner's payoff is determined by entire stream

[5] More specifically, we have

$$u^*(\boldsymbol{G}) = \frac{1}{2}\mathbf{1}'(\boldsymbol{I}-\phi\boldsymbol{G})^{-1}(\boldsymbol{I}-\phi\boldsymbol{G})^{-1}\mathbf{1} = \frac{1}{2}\mathbf{1}'\left(\boldsymbol{I}+\phi\,\boldsymbol{G}+\phi^2\boldsymbol{G}^2+...\right)^2\mathbf{1} = \frac{1}{2}\mathbf{1}'\left(\sum_{k=0}(k+1)\,\phi^k\boldsymbol{G}^k\right)\mathbf{1}.$$

of equilibrium welfare::

$$v_{\boldsymbol{\delta}}(\boldsymbol{s}) = \sum_{t=1}^{T} \delta_t u^* (\boldsymbol{G}(t)),$$

where $\boldsymbol{\delta} = (\delta_t)_{t=1}^T$ is the profile of discount factors in which $\delta_t \in [0,1]$ is the discount factor of period t. The network design problem is then described as

$$\max_{\boldsymbol{s} \in S} v_{\boldsymbol{\delta}}(\boldsymbol{s}). \tag{4}$$

The profile of discount factors $\boldsymbol{\delta} = (\delta_t)_{t=1}^T$ describes various levels of social planner's farsightedness. In particular, for some $\epsilon > 0$, when the discount factor profile is $(\delta_t)_{t=1}^T = (\epsilon^t)_{t=1}^T$, then the future benefits are more heavily discounted as ϵ decreases. Therefore, the social planner is more myopic when ϵ decreases. We leave the formal discussion of myopia in Section 3.2. Here, we formally define far-sightedness.

Definition 2. *The social planner is far-sighted if*

$$\delta_t = \begin{cases} 0 & \text{if } 1 \le t \le T-1 \\ 1 & \text{if } t = T \end{cases}$$

When the planner is far-sighted, then Problem (4) can be simplified as designing an efficient network with total number of T links, which can be formulated as

$$\max_{\boldsymbol{G} \text{ s.t. } \mathbf{1}'\boldsymbol{G}\mathbf{1} = 2T} u^* (\boldsymbol{G}). \tag{5}$$

This network design problem has been studied by Belhaj et al. (2016).

2.3 Notations

We end this section by introducing some special network structures that will play an essential role in following analysis. Denote $N_i(\boldsymbol{G}) = \{j : g_{ij} = 1\}$ the set of i's neighbors in network $\boldsymbol{G}$.

Definition 3. *A network $\boldsymbol{G}$ is called a nested split graph (NSG) if for each $i \neq j$, either $N_i(\boldsymbol{G}) \setminus \{j\} \subseteq N_j(\boldsymbol{G}) \setminus \{i\}$ or $N_j(\boldsymbol{G}) \setminus \{i\} \subseteq N_i(\boldsymbol{G}) \setminus \{j\}$.*

For the convention of our proof, we introduce Definition 3 among several equivalent definitions of NSG. For any positive integer k, we use $\mathcal{G}(k)$ to denote the set of networks with k links, and $\mathcal{NSG}(k)$ to denote the set of NSGs with k links. NSG is a large family of networks including various network structures. Figure 1 presents $\mathcal{NSG}(8)$ when the network is formed by $n = 7$ agents.

Definition 4. *A network $\boldsymbol{G} \in \mathcal{NSG}(t)$ is a quasi-complete graph, denoted by $\boldsymbol{QC}(t)$ if it contains a complete subgraph formed by p nodes with $\frac{p(p-1)}{2} \le t < \frac{p(p+1)}{2}$, and the remaining $t - \frac{p(p-t)}{2}$ links are set between one other node and nodes in the complete subgraph.*

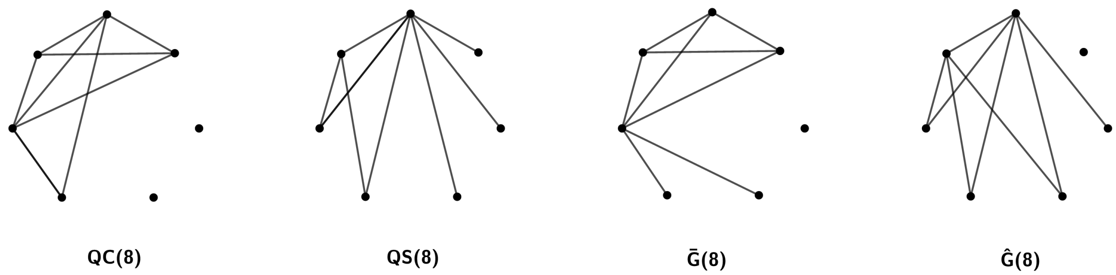


Figure 1: Nested split graphs $\mathcal{NSG}(8)$

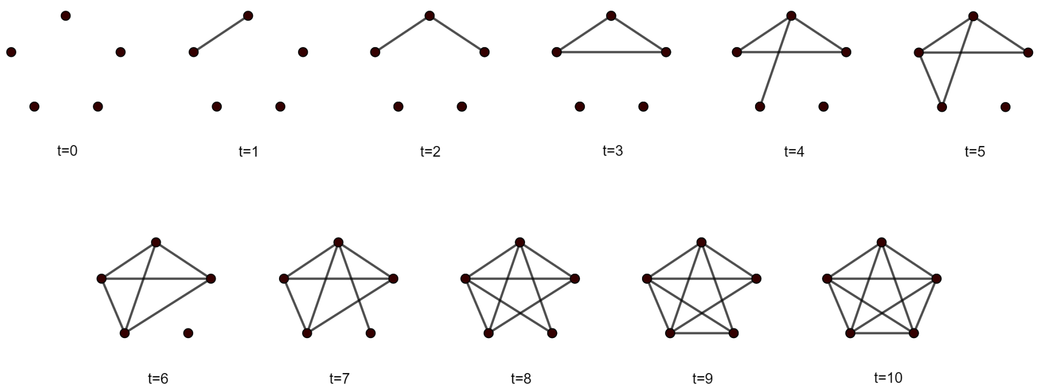


Figure 2: Quasi-complete networks.

Quasi-complete graph is a subclass of NSGs. Given the total number of links, quasi-complete graph has the largest possible complete subgraph.[6] Figure 2 presents quasi-complete graphs with 5 nodes and various number of links. It is worth noting that, given a fixed number of links, a quasi-complete graph is unique up to permutation. Figure 2 also illustrates a dynamic network formation process with 10 periods.

Definition 5. *A network $\boldsymbol{G} \in \mathcal{NSG}(t)$ is a quasi-star graph, denoted by $\boldsymbol{QS}(t)$, if it has a set of p central nodes with $n-1$ links, and the remaining $t - p(n-1)$ links are set so as to construct another central node.*

Quasi-star is another prominent subclass of NSGs which is also unique up to permutation fixing the total number of links. Quasi-star has as many nodes with degree $n-1$ as possible. Figure 3 illustrates quasi-star graphs with 5 nodes and various number of links. In particular, $\boldsymbol{QS}(t)$ is the graph complement of $\boldsymbol{QC}(t')$ where $t' = \frac{n(n-1)}{2} - t$. That is, the sum of the two adjacency matrices $\boldsymbol{QS}(t) + \boldsymbol{QC}(t')$ is equal to the adjacency matrix of complete graph after permutation. Moreover, $\boldsymbol{QS}(t) = \boldsymbol{QC}(t)$ if and only if $t \in \left\{0, 1, 2, \frac{n(n-1)}{2} - 2, \frac{n(n-1)}{2} - 1, \frac{n(n-1)}{2}\right\}$.[7]

[6] A quasi-complete graph with t links, where $\frac{p(p-1)}{2} \leq t < \frac{p(p+1)}{2}$, contains a unique maximal, and thus maximum, clique of size p.

[7] The necessity is straightforward. Suppose there exists an other t such that $\boldsymbol{QS}(t) = \boldsymbol{QC}(t)$. Assume there are

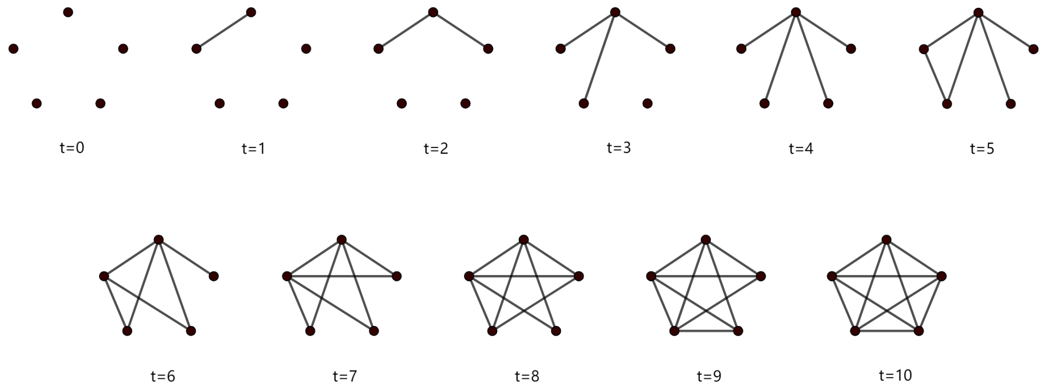


Figure 3: Quasi-star networks.

3 Optimal networks

We present our main results in this section. First, we show that, under the prescribed formation process, forming an NSG is optimal for the social planner at each period. This result is shown to be robust against convex best response. Then, when the social planner is myopic, his optimal strategy induces a quasi-complete graph in each period. Finally, we compare the myopic and global optimum and provide a sufficient condition to ensure that they are identical.

3.1 The optimal formation process

The following Lemma is crucial for our analysis.

Lemma 1. *Given a network $\boldsymbol{G}$ and two distinct nodes i, j such that $N_j(\boldsymbol{G}) \setminus \{i\} \neq N_i(\boldsymbol{G}) \setminus \{j\}$. Defined by $L = \{l \in N \setminus \{i,j\} \,|\, g_{il} = 0$ and $g_{jl} = 1\}$ the set of j's neighbors who are not neighbors of i, and $\hat{\boldsymbol{G}} = \boldsymbol{G} + \sum_{l \in L} \boldsymbol{E}_{il} - \sum_{l \in L} \boldsymbol{E}_{jl}$ is the network obtained by neighborhood switch from j to i. Then, we have*

$$\mathbf{1}'\boldsymbol{G}^k\mathbf{1} < \mathbf{1}'\hat{\boldsymbol{G}}^k\mathbf{1} \text{ for any integer } k \geq 2.$$

The link reallocation, which switches all j's neighbors (but not i's) to i, does not alter the total number of links. Thus, the aggregate paths of length 1 does not change after the link reallocation. Lemma 1 shows that, such link reallocation increases aggregate paths of length larger than 1, and therefore, both equilibrium activities and welfare raise. Belhaj et al. (2016)'s Lemma 1 argues that such reallocation increases welfare if $a_i^*(\boldsymbol{G}) \geq a_j^*(\boldsymbol{G})$.[8] Our Lemma 1 generalizes Belhaj et al. (2016)'s Lemma 1 in two directions. First, for any objective function which is monotone in aggregate paths of any length, such link reallocation leads to an improvement. For instance, besides

k isolated agents in $\boldsymbol{QC}(t)$. Then we have $\frac{(n-k-1)(n-k-2)}{2} + 1 \leq t \leq \frac{(n-k)(n-k-1)}{2}$. If there are k isolated agents in $\boldsymbol{QS}(t)$, we have $t = n-k-1$. It is easy to derive the contradiction when $t \in \left(2, \frac{n(n-1)}{2} - 2\right)$.

[8] Belhaj et al. (2016)'s Lemma 1 states that, in our notations, consider a network $\boldsymbol{G}$ and two distinct nodes i, j

equilibrium activities and welfare, aggregate diffusion centrality, proposed by Banerjee et al. (2013), also increases. Second, we argue that link reallocation from agent with lower equilibrium effort to agent with higher effort will also improve the social welfare.

Theorem 1. *Given any profile of discount factors* $\boldsymbol{\delta} = (\delta_t)_{t=1}^{T} \in [0,1]^T$*, there exists a solution* $\boldsymbol{s}^* = (\boldsymbol{G}^*(t))_{t=1}^{T}$ *of Problem* (4) *such that* $\boldsymbol{G}^*(t) \in \mathcal{NSG}(t)$ *for any* $t \leq T$*. Moreover, if* $\boldsymbol{s}^* = (\boldsymbol{G}^*(t))_{t=1}^{T}$ *is a solution of Problem* (4)*, then* $\boldsymbol{G}^*(t) \in \mathcal{NSG}(t)$ *for any* t *such that* $\delta_t \in (0,1]$.

We prove this theorem by showing that, for any strategy $\boldsymbol{s} = (\boldsymbol{G}(t))_{t=1}^{T}$ such that $\boldsymbol{G}(t) \notin \mathcal{NSG}(t)$ for some $1 \leq t \leq T$, there exists a weakly dominant strategy $\hat{\boldsymbol{s}} = (\boldsymbol{G}(t))_{t=1}^{T}$ such that $\widehat{\boldsymbol{G}(t)} \in \mathcal{NSG}(t)$. The proof is constructive based on Algorithm 1.

Algorithm 1. *For any strategy* $\boldsymbol{s} = (\boldsymbol{G}(t))_{t=1}^{T}$*, define algorithm as follows,*

Step 1*: Check whether* $G(t) \in \mathcal{NSG}(t)$ *for all* $1 \leq t \leq T$*,*

- *If it is true, then the algorithm stops.*
- *If it is wrong, then*
  - ***Step 2****: Find* $\tilde{t}$ *such that* $\boldsymbol{G}(t) \in \mathcal{NSG}(t)$, $\forall t \leq \tilde{t}-1$ *and* $\boldsymbol{G}(\tilde{t}) \notin \mathcal{NSG}(\tilde{t})$*. Find the pair of nodes* $i, j \in N$ *which violates the nestedness at* $\tilde{t}$ *(suppose* i*'s degree is larger than* j*'s).*
  - ***Step 3****: Construct another strategy* $\hat{\boldsymbol{s}} = (\hat{\boldsymbol{G}}(t))_{t=1}^{T}$ *according to the following rules,*
    - *If* $t < \tilde{t}$*, let* $\hat{\boldsymbol{G}}(t) = \boldsymbol{G}(t)$*.*
    - *If* $t \geq \tilde{t}$ *and* $\boldsymbol{G}(t+1) = \boldsymbol{G}(t) + \boldsymbol{E}_{jl}$ *for some* $l \notin \{i,j\}$*, then*
      - *If* $\hat{g}_{il}(t) = 0$*, let* $\hat{\boldsymbol{G}}(t+1) = \hat{\boldsymbol{G}}(t) + \boldsymbol{E}_{il}$ *(e.g.,* $t = 4, 5$ *in Figure 4);*
      - *If* $\hat{g}_{il}(t) = 1$*, let* $\hat{\boldsymbol{G}}(t+1) = \hat{\boldsymbol{G}}(t) + \boldsymbol{E}_{jl}$*.*
    - *If* $t \geq \tilde{t}$ *and* $\boldsymbol{G}(t+1) = \boldsymbol{G}(t) + \boldsymbol{E}_{il}$ *for some* $l \notin \{i,j\}$*, then*
      - *If* $\hat{g}_{il}(t) = 0$*, let* $\hat{\boldsymbol{G}}(t+1) = \hat{\boldsymbol{G}}(t) + \boldsymbol{E}_{il}$*;*
      - *If* $\hat{g}_{il}(t) = 1$*, let* $\hat{\boldsymbol{G}}(t+1) = \hat{\boldsymbol{G}}(t) + \boldsymbol{E}_{jl}$ *(e.g.,* $t = 7$ *in Figure 4).*
    - *If* $t \geq \tilde{t}$ *and* $\boldsymbol{G}(t+1) = \boldsymbol{G}(t) + \boldsymbol{E}_{lk}$ *for some* $l, k \notin \{i,j\}$ *or* $(l.k) = (i,j)$*, then let* $\hat{\boldsymbol{G}}(t+1) = \hat{\boldsymbol{G}}(t) + \boldsymbol{E}_{lk}$ *(e.g.,* $t = 6$ *in Figure 4).*
  - ***Step 4****: Set* $\boldsymbol{s} = \hat{\boldsymbol{s}}$ *and return to* ***Step 1****.*

We use the following six-node example (Figure 4) to illustrate the idea of the algorithm. Suppose a strategy $\boldsymbol{s} = (\boldsymbol{G}(t))_{t=1}^{8}$ induces non-NSGs at some periods. For instance, $\tilde{t}$ is the first

such that $a_i^*(\boldsymbol{G}) \geq a_j^*(\boldsymbol{G})$. Then, for the set of nodes $L = \{l \in N \setminus \{i,j\} \,|\, g_{il} = 0 \text{ and } g_{ij} = 1\}$, we have

$$u^*(\boldsymbol{G}) < u^*\left(\boldsymbol{G} + \sum_{l \in L} \boldsymbol{E}_{il} - \sum_{l \in L} \boldsymbol{E}_{jl}\right).$$

However, the network induced by switching all j's neighbors to i and the one by switching all i's neighbors to j are isomorphic. Therefore, the requirement of $a_i^*(\boldsymbol{G}) \geq a_j^*(\boldsymbol{G})$ in Belhaj et al. (2016)'s Lemma 1 is redundant.

time that $\boldsymbol{s}$ forms a non-NSG in which the neighbors of j (lower degree node) are not nested by that of i (higher degree node). Then we construct another strategy $\hat{\boldsymbol{s}} = \left(\hat{\boldsymbol{G}}(t)\right)_{t=1}^{8}$ based on Step 3 of Algorithm 1. It is easy to verify that the newly constructed strategy $\hat{\boldsymbol{s}}$ is plausible, i.e., $\hat{\boldsymbol{G}}(t+1) \in \mathbb{S}\left(\hat{\boldsymbol{G}}(t)\right)$, $\forall 1 \leq t \leq 6$. Furthermore, for each t, there exists a set of nodes $L(t) := \{l \in N : \hat{g}_{il}(t) > g_{il}(t)\} = \{l \in N : \hat{g}_{jl}(t) > g_{jl}(t)\}$ such that $\hat{\boldsymbol{G}}(t) = \boldsymbol{G}(t) - \sum_{l \in L(t)} \boldsymbol{E}_{jl} + \sum_{l \in L(t)} \boldsymbol{E}_{il}$. In Figure 4, the set $L(t)$ of nodes is shown by dotted circles. Therefore, Lemma 1 implies that the original strategy $\boldsymbol{s}$ is weakly dominated by $\hat{\boldsymbol{s}}$. Note that $\hat{\boldsymbol{s}}$ induces NSGs for any $t \leq 4$. However, it does not induce an NSG at $t = 6$. We then iterate Step 2 and Step 3 to construct a new strategy which forms NSGs for any $t \leq 6$. Note that at every iteration, the newly constructed strategy forms NSGs for at least one more period, i.e. for any $t \leq \tilde{t}$. Therefore the algorithm terminates in at most $T - \tilde{t}$ iterations. The terminal strategy induces NSGs at each period and weakly dominates the original one.

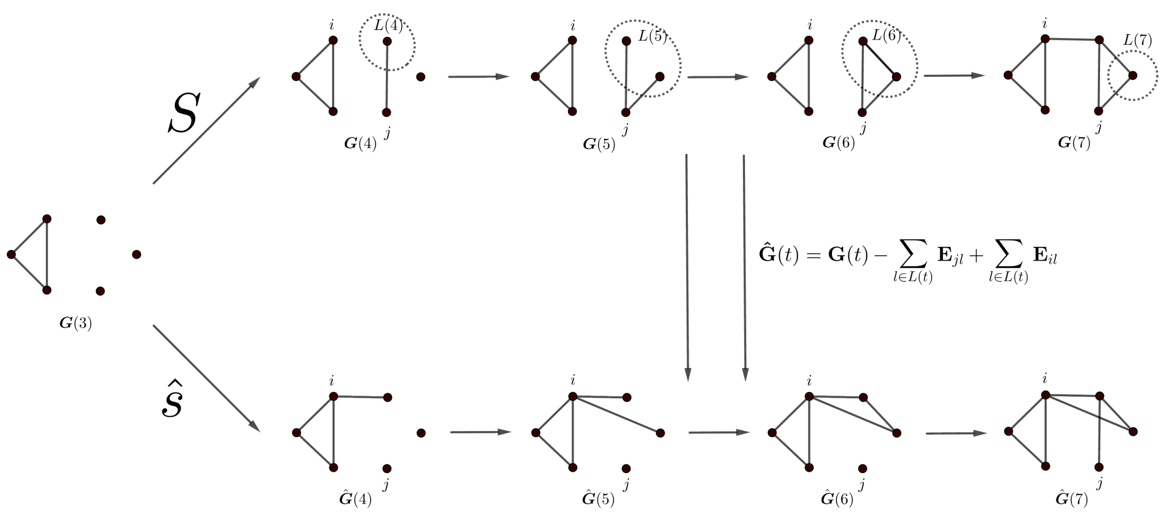


Figure 4: An illustrative example for the proof of Theorem 1

Theorem 1 demonstrates the importance of NSGs in the network formation process. Building up a link to form an NSG is an optimal strategy for the social planner regardless of the level of farsightedness. When $v_{\boldsymbol{\delta}}(s) = u^*(\boldsymbol{G}(T))$, Theorem 1 degenerates to Belhaj et al. (2016) which states that given a total number of T links, the globally efficient network is an NSG.

It is worth noting that, the optimal strategy $\boldsymbol{s}^*$ crucially depends on the strength of network effect: different network effect ϕ may result in different optimal strategies, even though the profile of discount factors $(\delta_t)_{t=1}^{T}$ remains the same. Example 1 highlights impact of strength of network effect on the network formation process.

Example 1. *Consider the planner's problem (4). Suppose $n = 7$, $T = 8$ and the planner is far-sighted, i.e. $v_{\boldsymbol{\delta}}(s) = u^*(\boldsymbol{G}(T))$. According to Theorem 1, the formed network must be one of the four NSGs listed in Figure 1. Table 1 lists equilibrium welfare induced by these four NSGs with various values of ϕ*

Comparison welfare with different ϕ

Strength of network effect	$\boldsymbol{QC}(8)$	$\boldsymbol{QS}(8)$	$\bar{\boldsymbol{G}}(8)$	$\hat{\boldsymbol{G}}(8)$
$\phi = 0.01$	*3.6685*	*3.6687**	*3.6684*	*3.6681*
$\phi = 0.1$	*6.5104**	*6.4047*	*6.4259*	*6.41*
$\phi = 0.2$	*22.6716**	*18.0464*	*19.5074*	*19.9595*

When the strength of network effect is sufficient small, i.e. $\phi = 0.01$*, the globally efficient network is quasi-star. However, the efficient network becomes quasi-complete when* ϕ *is large, i.e.,* $\phi = 0.1$ *or* 0.2*. Therefore, the planner's optimal strategy differs as the strength* ϕ *of network effect varies.*

3.1.1 Convex best response

Theorem 1 is also robust against convex best response. Suppose agent i's payoff is given by

$$u_i(\boldsymbol{a}; \boldsymbol{G}) = a_i + \psi(\sum_{j \in N} g_{ij} a_j) a_i - \frac{1}{2} a_i^2, \tag{6}$$

where function $\psi(\cdot)$ is weakly convex and increasing. We can then derive agent i's best response function as

$$\alpha_i^*(a_{-i}; \boldsymbol{G}) = 1 + \psi(\sum_{j \in N} g_{ij} a_j).$$

Moreover, we assume $\psi'(\cdot) < \frac{1}{\lambda_{\max}(\boldsymbol{G})}$ to ensure that the network game has a unique Nash equilibrium. Let $\boldsymbol{a}^*(\boldsymbol{G}) \equiv (a_i^*(\boldsymbol{G}))_{i \in N}$ be the Nash equilibrium, then agent i's welfare also equals to half of the square of his own equilibrium effort, i.e. $u_i^*(\boldsymbol{G}) = \frac{1}{2}(a_i^*(\boldsymbol{G}))^2$. Lemma 2 is a non-linear extension of Lemma 1.

Lemma 2. *Suppose each agent i's utility is given by (6). Consider two distinct agents* i, j *such that* $N_j(\boldsymbol{G}) \setminus \{i\} \neq N_i(\boldsymbol{G}) \setminus \{j\}$*. Let* $L = \{l \in N \setminus \{i, j\} \,|g_{il} = 0$ *and* $g_{jl} = 1\}$ *and* $\hat{\boldsymbol{G}} = \boldsymbol{G} + \sum_{l \in L} \boldsymbol{E}_{il} - \sum_{l \in L} \boldsymbol{E}_{jl}$*. Then we have*

$$\begin{aligned} a^*(\boldsymbol{G}) &< a^*\left(\hat{\boldsymbol{G}}\right), \text{ and} \\ u^*(\boldsymbol{G}) &< u^*\left(\hat{\boldsymbol{G}}\right). \end{aligned}$$

The link reallocation, $\sum_{l \in L} \boldsymbol{E}_{il} - \sum_{l \in L} \boldsymbol{E}_{jl}$, increases the gap between agent i's and j's equilibrium efforts. The assumption of convex best response function ensures that the increment of agent i's equilibrium efforts exceeds the decreasing of agent j's efforts, due to the reallocation. Denote $k \in N$ as one of nodes who is linked to both node i and j. The convexity of best response function guarantees that the equilibrium effort of node k increases post reallocation.

Based on Lemma 2, we can then adopt Algorithm 1 to find a dominant strategy over any strategy which induces non-NSG at some periods. Therefore, Theorem 1 still holds with convex best response.

Proposition 1. *When each agent i's utility is given by (6), all the statements in Theorem 1 still hold.*

The following example illustrates that Lemma 2 fails when $\psi(\cdot)$ is concave.

Example 2. *Suppose each agent i's utility is given by (6), where $\psi(x) = (0.1x)^{\frac{1}{2}}$ is concave.*[9] *Then, the best response of agent i,*

$$\alpha_i^*(a_{-i}; \boldsymbol{G}) = 1 + \left(0.1 \sum_{j \in N} g_{ij} a_j\right)^{\frac{1}{2}}.$$

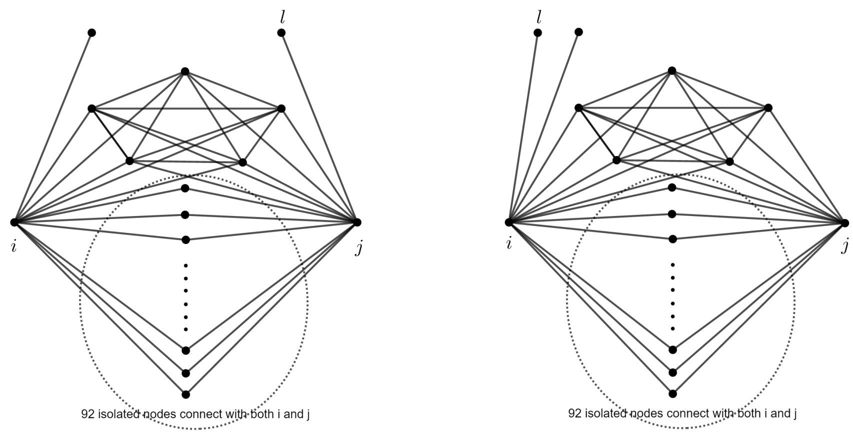
Consider network $\boldsymbol{G}$ formed by 101 nodes. In the network, nodes i, j are disconnected and there is only one node, l, belongs to j's neighbors but not i's. All the other nodes connect with both i and j. In particular, there is a clique of size 5 in network $\boldsymbol{G}$. Shifting link jl to il increases i's effort and meanwhile decreases j's. Due to concavity of the best response function, the increment of i's effort is less than the decreasing of j. Specifically, i's equilibrium effort increases from 2.0998 in network $\boldsymbol{G}$ to 2.1647 in $\boldsymbol{G} + \boldsymbol{E}_{il} - \boldsymbol{E}_{jl}$ (net increase 0.0649) while j decreases from 2.0998 to 2.0313 (net decrease 0.0685). Therefore, the efforts of all nodes connecting with both i and j are decreasing. Such decreasing is amplified through dense intra-connection of these nodes (the clique): the equilibrium effort of agents in the clique decreases from 2.1273 to 2.1271 and that of the 92 agents decreases from 1.6480 to 1.6478. The two nodes only connecting with i will increases their effort corresponding to the link shifting ($a_l^*(\boldsymbol{G}) = 1.4582$ and $a_l^*\left(\hat{\boldsymbol{G}}\right) = 1.4653$), while the total decreasing of efforts from other nodes covers the effort increments of these two nodes. As a result, the link shifting decreases aggregate equilibrium activities. Figure 2 summarizes the numerical results.

$\boldsymbol{G}$ *v.s.* $\hat{\boldsymbol{G}}$

	$\boldsymbol{G}$	$\hat{\boldsymbol{G}}$
Aggregate activities	*169.3722**	*169.3570*
Welfare	*285.5736**	*285.5231*

Example 2 illustrates the failure of Lemma 2 when $\psi(\cdot)$ is concave. It implies that the link re-allocation approach developed by Belhaj et al. (2016) does not hold when best response function is concave. However, it is not sufficient to conclude that Theorem 1 does not hold. The question of designing efficient discrete network under concave best response is an open question and left for future research.

[9] This utility form encounters the failure of sufficient condition of unique equilibrium: $\psi'(\cdot) < \frac{1}{\lambda_{\max}(\boldsymbol{G})}$. We have verified that decreasing the strength of network effect from 0.1 to 0.0001, which guarantees the uniqueness of equilibrium, does not change the comparison results between $\boldsymbol{G}$ and $\hat{\boldsymbol{G}}$.



$\boldsymbol{G}$ $\qquad\qquad$ $\boldsymbol{G} - \boldsymbol{E}_{jl} + \boldsymbol{E}_{il}$

Figure 5: Lemma 2 fails when the best response is concave

3.2 Myopic optimum

Example 1 highlights the impact of the strength of network effect on the network formation process. However, the discount factor profiles also play an important role in determining the optimal strategy. In this subsection, we study the optimal strategy when the social planner becomes sufficiently myopic.

The greedy algorithm is commonly adopted to pin down the solution to Problem (5). We first consider its underlying strategy.

Definition 6. *A strategy* $\mathring{s} = \left(\mathring{\boldsymbol{G}}(t)\right)_{t=1}^{T}$ *is greedy strategy of Problem (4) if*

$$\mathring{\boldsymbol{G}}(t) \in \arg \max_{\boldsymbol{G} \in \mathbb{S}\left(\mathring{\boldsymbol{G}}(t-1)\right)} u^*(\boldsymbol{G}) \text{ for any } t.$$

The greedy strategy builds up a link between an unlinked pair in the network formed at the last period to maximize the social welfare at the current period. The following lemma shows that the greedy strategy is optimal for a sufficiently myopic social planner.

Lemma 3. *There exists an* $\bar{\varepsilon} > 0$ *such that when* $\delta_t > 0$ *and* $\frac{\delta_{t+1}}{\delta_t} < \bar{\varepsilon}$ *for any* $t \leq T$*, then the solution of Problem* (4) *is greedy strategy, i.e.,* $s^* = \mathring{s}$ *if* $\delta_t > 0$ *and* $\frac{\delta_{t+1}}{\delta_t} < \bar{\varepsilon}, \forall t \leq T$*.*

Lemma 3 is intuitive. Whenever the planner heavily discounts future payoffs, then at each period, she will choose the strategy to maximize her payoff at the current period. We therefore define that a social planner is **myopic** if her discount factor profile satisfies $\delta_t > 0$ and $\frac{\delta_{t+1}}{\delta_t} < \bar{\varepsilon}$ for any t. Our next results fully characterizes the optimal strategy for a myopic social planner, i.e., the greedy strategy.

Theorem 2. *The greedy strategy induces a quasi-complete graph in each period, i.e.,* $\mathring{\boldsymbol{G}}(t) = \boldsymbol{QC}(t)$ *for any* t.

In the proof of Theorem 2, we further discriminate among NSGs. We use mathematical induction and assume that a quasi-complete network is formed at period $t-1$. By Theorem 1, we restrict our attention to the set of strategies which induce an NSG at period t. There exists at most two different NSGs succeeding a quasi-complete graph. Figure 6 illustrates building up a link in a quasi-complete network with 5 agents at $t = 5$. There are only two ways of forming an NSG, i.e. quasi-complete network $\mathring{\boldsymbol{G}} = \boldsymbol{G} + \boldsymbol{E}_{34}$ and the other NSG $\bar{\boldsymbol{G}} = \boldsymbol{G} + \boldsymbol{E}_{15}$.

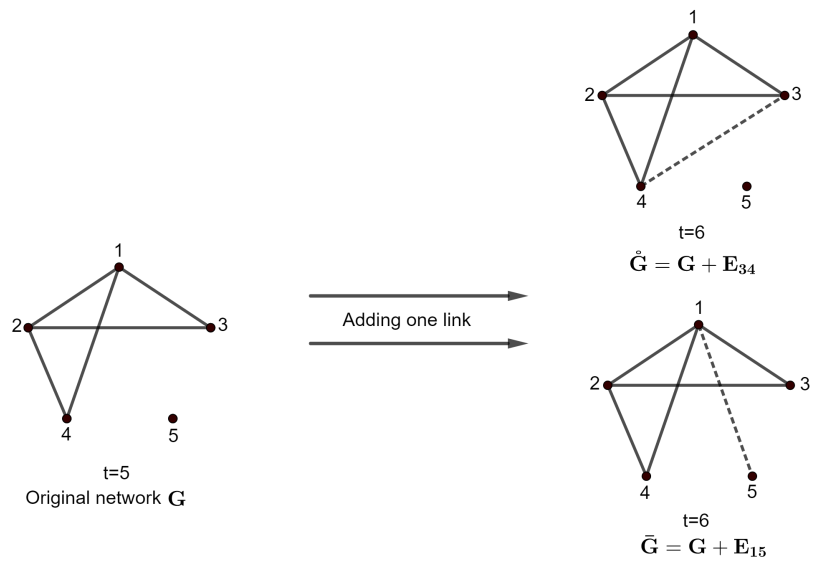


Figure 6: Two different NSGs succeeding a quasi-complete graph

We then compare these two classes of NSGs through inductively showing that aggregate walks of length $k \geq 2$ of quasi-complete network strictly dominates that of the other NSG. We put this argument in the following Lemma.

Lemma 4. *For any integer* $k \geq 2$ *and* $t \geq 0$, $\mathbf{1}'\boldsymbol{QC}(t+1)^k\,\mathbf{1} > \mathbf{1}'\bar{\boldsymbol{G}}(t+1)^k\,\mathbf{1}$.

Theorem 2 demonstrates that the myopically efficient formation process is unique up to permutation, regardless of the strength of network effect. It provides a micro-foundation for quasi-complete graphs as they are formed under greedy algorithm. Figure 2 exactly illustrates the dynamic formation process in 10 periods when the planner is myopic. Theorem 2, together with Lemma 3, highlights the role of discount factors on the optimal strategy, i.e. when the social planner is myopic, the formed network under the optimal strategy will always be quasi-complete irrespective of the strength of network effect.

3.2.1 Greedy algorithm and global efficiency

Belhaj et al. (2016) show that, given a fixed number of total links, the globally efficient network is an NSG. Different strength of network effect corresponds to different NSG as the globally efficient network. Can each NSG be globally efficient at some strength of network effect? One by-product of Theorem 2 is to further refine the prediction of globally efficient network. Given a total number of T links, let $\dot{\boldsymbol{G}}(T)$ denote the globally efficient network, i.e., $\dot{\boldsymbol{G}}(T) \in \arg\max_{\boldsymbol{G}\in\mathcal{G}(T)} u^*(\boldsymbol{G})$.

Corollary 1. *The globally efficient network* $\dot{\boldsymbol{G}}(T) \in \mathcal{NSG}(T) \setminus \bar{\boldsymbol{G}}(T)$.

Theorem 2 implies that, when greedy algorithm does result in global optimum, then the globally efficient network must be quasi-complete. However, as illustrated by Example 1, the globally efficient network may not be quasi-complete, and thus, the greedy algorithm fails to reach global optimum. The next proposition is an immediate corollary of Bernardo M. Ábrego (2009)'s Theorem 2.8. It shows that the greedy algorithm leads to global optimum when the number of links is exactly 3 or the strength of network effect is small enough and the number of links is large enough as well.

Proposition 2. *When the number of agents* $n \geq 3$*, then we have that*

1. *If the total number of links* $t = 3$*, then* $\dot{\boldsymbol{G}}(t) = \boldsymbol{QC}(t)$*;*
2. *If the total number of links* $t \in (3, \frac{n^2-3n}{4})$*, then there exists* $\hat{\phi} > 0$ *such that for any* $\phi \in (0, \hat{\phi})$*, then* $\dot{\boldsymbol{G}}(t) = \boldsymbol{QS}(t)$*;*
3. *If the total number of links* $t \in \left(\frac{n^2+n}{4}, \frac{n(n-1)}{2}\right]$*, then there exists* $\bar{\phi} > 0$ *such that for any* $\phi \in (0, \bar{\phi})$*, then* $\dot{\boldsymbol{G}}(t) = \boldsymbol{QC}(t)$*, i.e. Greedy algorithm leads to global optimum.*

Our next example illustrates Proposition 2 by comparing quasi-star and quasi-complete network with various number of links.

Example 3. *Suppose that there are* $n = 6$ *agents and the discount factor is given by* $\phi = 0.1$*. If the total number of links is given by* $t = 3$*, then quasi-complete network strictly dominates the quasi-star network in terms of aggregate equilibrium welfare; However if total number of links* t *is increased to* 4 *(which is the unique integer belongs to* $(3, \frac{n^2-3n}{4})$*), then the welfare induced by quasi-star network is strictly higher than that of quasi-complete network, which implies the divergence between globally efficient network and what is led by greedy strategy; Finally, if we increase the total number of links* t *to* 11 *(which is an integer belongs to* $\left(\frac{n^2+n}{4}, \frac{n(n-1)}{2}\right]$*), then quasi-complete network again dominates the quasi-star network. Figure 7 illustrates the net welfare induced by quasi-complete and quasi-star.*

The final remark is that the optimal strategy is not an array of globally efficient networks, i.e.,

$$\boldsymbol{s}^* := (\boldsymbol{G}^*(t))_{1\leq t\leq T} \neq (\dot{\boldsymbol{G}}(t))_{1\leq t\leq T}.$$

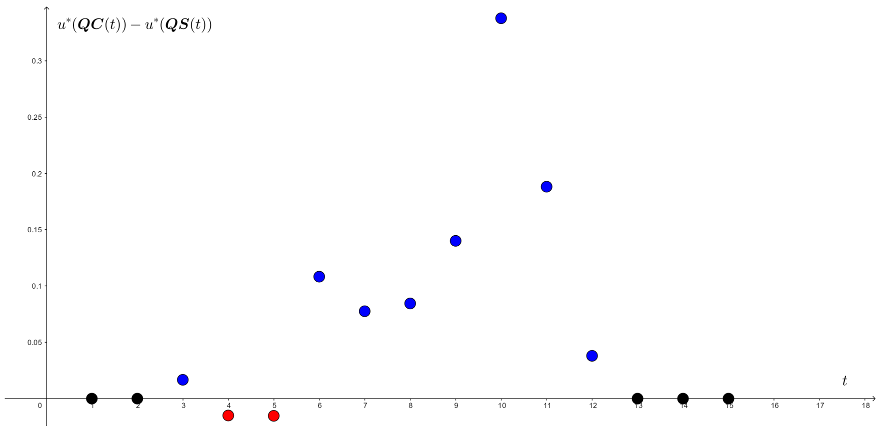


Figure 7: Quasi-complete v.s. Quasi-star: $u^*(\boldsymbol{QC}(t)) - u^*(\boldsymbol{QS}(t))$

The main reason is that the strategy $(\dot{\boldsymbol{G}}(t))_{1\le t\le T}$ is infeasible or there exists some $1 \le t \le T-1$ such that $\dot{\boldsymbol{G}}(t+1)$ does NOT succeed $\dot{\boldsymbol{G}}(t)$. For instance, we show that the myopic social planner should form quasi-complete network at each period. However, in Example 1, when $\phi = 0.01$, we have $\dot{\boldsymbol{G}}(8) = \boldsymbol{QS(8)} \neq \boldsymbol{QC(8)}$. If $(\dot{\boldsymbol{G}}(t))_{1\le t\le T}$ is feasible, then it must also be optimal, which contradicts the optimality of greedy strategy.

4 Robustness of results

We end this paper by discussing three changes in the model's specifications.

Activity concern. Our proof concerning the maximization of welfare naturally covers that of equilibrium activities since, as shown by equation (2), the activity concern is proportional to aggregate discounted paths in the network. Lemma 1 implies that forming an NSG in each period induces a stream of highest equilibrium activities; Lemma 4 demonstrates that $\boldsymbol{QC}(t)$ dominates $\bar{\boldsymbol{G}}(t)$ in terms of equilibrium activities, and therefore the myopically optimal network is quasi-complete.

.

Heterogeneous agents. Suppose the utility function of agent i is

$$u_i(\boldsymbol{a},\boldsymbol{G}) = \theta_i a_i - \frac{1}{2}a_i^2 + \phi\sum_{j\in N} g_{ij}a_ia_j,$$

where the intrinsic marginal utilities $\boldsymbol{\theta} = (\theta_i)_{i\in N}$ of agents are distinct. Then, structure of network $\boldsymbol{G}$ and intrinsic marginal utilities $\boldsymbol{\theta}$ jointly determine the equilibrium which is given by $\boldsymbol{a}^*(\boldsymbol{G},\boldsymbol{\theta}) = (\boldsymbol{I} - \phi\boldsymbol{G})^{-1}\boldsymbol{\theta}$. To analyze the dynamic model of network formation with agent heterogeneity, we need to modify Lemma 1 as follows:

Lemma 5. *Given a network $\boldsymbol{G}$ and two distinct nodes i, j such that $\theta_j \le \theta_i$ and $N_j(\boldsymbol{G}) \subsetneq N_i(\boldsymbol{G})$.*

Then, for the set of nodes $L = \{l \in N \backslash \{i, j\} \,|g_{il} = 0$ *and* $g_{ij} = 1\}$ *we have*

$$\mathbf{1}'\boldsymbol{G}^k\boldsymbol{\theta} < \mathbf{1}'\left(\boldsymbol{G} + \sum_{l\in L}\boldsymbol{E}_{il} - \sum_{l\in L}\boldsymbol{E}_{jl}\right)^k \boldsymbol{\theta} \textit{ for any integer } k \geq 2.$$

According to Lemma 5, the social planner tends to accumulate the agents with high θ before connecting the agents with polarized θ. Naturally, the optimal strategy forms networks in which the neighbors of agent with low θ is nested by that with high θ. As a result, the formed network is also an NSG in each period. Theorem 1 still holds.

Note that, Theorem 2 fails when agents are heterogeneous: if an agent's θ is much larger than others', then the optimal network formed in period $t = n - 1$, $\boldsymbol{G}^*(n-1)$, is a star network with the central agent being the high-θ one.

Delegation of link formation. In the description of network formation process, the social planner builds up links sequentially to maximize his objective (4). We introduce an alternative dynamic model of network formation in which the social planner strategically picks an agent (rather than link) and the agent chooses to form a link that maximizes her utility.

Formally, the network is formed over T periods. At each period t, the game is divided into three stages: in the first stage, the social planner selects an agent i; in the second stage, agent i connects with a non-neighbor j to maximize her utility, i.e., agent i connects with $j \in \arg\max_{k,\text{ s.t., } ik\notin \boldsymbol{G}(t-1)} u_i^*(\boldsymbol{G}(t-1) + \boldsymbol{E}_{ik})$; in the last stage, the agents play network game $\Gamma(\boldsymbol{G}(t))$, where $\boldsymbol{G}(t) = \boldsymbol{G}(t-1) + \boldsymbol{E}_{ik}$ and then the social planner benefits from equilibrium welfare. This dynamic model of network formation is introduced in the spirit of König et al. (2014) which analyzes network formation with decentralized concerns. In König et al. (2014)'s model, a randomly picked agent chooses to form a link that increases her utility the most in each period. Departure from their model, the agent is strategically picked by the social planner here.

In this model, a strategy of the planner is a sequence of players $p = (p_t)_{t=1}^T$, where p_t denotes the nominated agent in period t. Any strategy p also induces a sequence of networks $(\boldsymbol{G}(t))_{t=1}^T$, where $\boldsymbol{G}(t) = \boldsymbol{G}(t-1) + \boldsymbol{E}_{p_t j}$ and $j \in \arg\max_{k,\text{ s.t., } p_t k\notin \boldsymbol{G}(t-1)} u_i^*(\boldsymbol{G}(t-1) + \boldsymbol{E}_{ik})$. Denote P as the set of strategies. Then, the network design problem is described as

$$\begin{aligned} & \max_{p\in P}\sum_{t=1}^{T}\delta_t u^*(\boldsymbol{G}(t)) \\ \text{s.t. } \boldsymbol{G}(t) &= \boldsymbol{G}(t-1) + \boldsymbol{E}_{p_t j}, \\ \text{where } j &\in \arg\max_{k,\text{ s.t., } p_t k\notin \boldsymbol{G}(t-1)} u_i^*(\boldsymbol{G}(t-1) + \boldsymbol{E}_{ik}). \end{aligned} \tag{7}$$

The following proposition states that this dynamic model of network formation in which the planner delegates an agent to form a new link is (outcome) equivalent to that in which the planner builds up new links directly.

Proposition 3. *Given discount factors $(\delta_t)_{t=1}^{T}$, a sequence of networks $(\boldsymbol{G}^*(t))_{t=1}^{T}$ is induced by the solution of Problem (7) if and only if $(\boldsymbol{G}^*(t))_{t=1}^{T}$ is a solution of Problem (4).*

5 Conclusion

This paper has studied a network formation process for linear complementary interaction. We shown that, starting from empty network, forming an NSG in each period maximizes discounted sum of equilibrium welfare across all periods. In particular, the myopic optimal strategy of the planner induces quasi-complete graph in each period. This work complements classical results focusing on global efficient networks by precluding a subclass of NSG from efficiency. Moreover, the parity and disparity between myopic and global optima has been illustrated. When the strength of complementary effect is small and the total number of links is large, both myopically and globally efficient structure is quasi-complete.

6 Appendix

Proof of Lemma 1: We prove an equivalent statement here: Given a network $\boldsymbol{G}$ and two distinct nodes i, j such that $N_j(\boldsymbol{G}) \setminus \{i\} \subsetneq N_i(\boldsymbol{G}) \setminus \{j\}$. Then, for any set of nodes $L = \{l_1, ..., l_k\} \subseteq N \setminus \{i, j\}$ such that $L \cap N_i(\boldsymbol{G}) = \emptyset$ we have

$$\mathbf{1}'\left(\boldsymbol{G}+\sum_{l\in L}\boldsymbol{E}_{il}\right)^k \mathbf{1} > \mathbf{1}'\left(\boldsymbol{G}+\sum_{l\in L}\boldsymbol{E}_{jl}\right)^k \mathbf{1} \text{ for any integer } k \geq 2.$$

This statement implies Lemma 1 directly.

We use $\hat{\boldsymbol{G}}$ and $\check{\boldsymbol{G}}$ to denote $\boldsymbol{G}+\sum_{l\in L}\boldsymbol{E}_{il}$ and $\boldsymbol{G}+\sum_{l\in L}\boldsymbol{E}_{jl}$ respectively. When $k = 2$, $\mathbf{1}'\boldsymbol{G}^2\,\mathbf{1} = \sum_{m\in N} e_m^2(\boldsymbol{G})$, where $e_m(\boldsymbol{G}) = |N_m(\boldsymbol{G})|$ is degree of agent m. Apparently, $e_m\left(\hat{\boldsymbol{G}}\right) = e_m\left(\check{\boldsymbol{G}}\right)$ for any $m \notin \{i, j\}$. Moreover, $e_i\left(\hat{\boldsymbol{G}}\right) - e_i\left(\check{\boldsymbol{G}}\right) = e_j\left(\hat{\boldsymbol{G}}\right) - e_j\left(\check{\boldsymbol{G}}\right)$ and $e_i\left(\hat{\boldsymbol{G}}\right) + e_i\left(\check{\boldsymbol{G}}\right) > e_j\left(\hat{\boldsymbol{G}}\right) + e_j\left(\check{\boldsymbol{G}}\right)$. Therefore, we have $\sum_{m\in N} e_m^2\left(\hat{\boldsymbol{G}}\right) > \sum_{m\in N} e_m^2\left(\check{\boldsymbol{G}}\right)$.

Denote vector $\boldsymbol{a}^k = \left(\hat{\boldsymbol{G}}\right)^k \mathbf{1}$ and $\boldsymbol{b}^k = \left(\check{\boldsymbol{G}}\right)^k \mathbf{1}$. We inductively show that $\boldsymbol{a}_m^k \geq \boldsymbol{b}_m^k$ for any $m \neq j$ and $\boldsymbol{a}_i^k + \boldsymbol{a}_j^k > \boldsymbol{b}_i^k + \boldsymbol{b}_j^k$.

Suppose the statement is true for all $k' \le k$. Then for any $m \notin \{i,j\}$,

$$\begin{aligned}
\boldsymbol{a}_m^{k+1} &= \left[\left(\boldsymbol{G}+\sum_{l\in L}\boldsymbol{E}_{il}\right)^{k+1}\mathbf{1}\right]_m \\
&= \left[\left(\boldsymbol{G}+\sum_{l\in L}\boldsymbol{E}_{il}\right)\boldsymbol{a}^k\right]_m \\
&= \sum_{p\in N} g_{mp}\boldsymbol{a}_p^k \\
&= \sum_{p\notin\{i,j\}} g_{mp}\boldsymbol{a}_p^k + g_{mi}\boldsymbol{a}_i^k + g_{mj}\boldsymbol{a}_j^k
\end{aligned}$$

Moreover, by $\boldsymbol{a}_i^k + \boldsymbol{a}_j^k > \boldsymbol{b}_i^k + \boldsymbol{b}_j^k$, we have $\boldsymbol{a}_i^k - \boldsymbol{b}_i^k > \boldsymbol{b}_j^k - \boldsymbol{a}_j^k$. Thus, $g_{mi}\left(\boldsymbol{a}_i^k - \boldsymbol{b}_i^k\right) \ge g_{mj}\left(\boldsymbol{b}_j^k - \boldsymbol{a}_j^k\right)$ since $g_{mi} \ge g_{mj}$ and $\boldsymbol{a}_i^k - \boldsymbol{b}_i^k > 0$. Therefore,

$$\begin{aligned}
\boldsymbol{a}_m^{k+1} &= \sum_{p\notin\{i,j\}} g_{mp}\boldsymbol{a}_p^k + g_{mi}\boldsymbol{a}_i^k + g_{mj}\boldsymbol{a}_j^k \\
&\ge \sum_{p\notin\{i,j\}} g_{mp}\boldsymbol{b}_p^k + g_{mi}\boldsymbol{b}_i^k + g_{mj}\boldsymbol{b}_j^k = \boldsymbol{b}_m^{k+1}
\end{aligned}$$

for $m \notin \{i,j\}$. Note that, for some node m such that $g_{mi} = 1$ and $g_{mj} = 0$, we have $\boldsymbol{a}_m^{k+1} > \boldsymbol{b}_m^{k+1}$. Now we compare $\boldsymbol{a}_i^{k+1}$ and $\boldsymbol{b}_i^{k+1}$,

$$\begin{aligned}
\boldsymbol{a}_i^{k+1} - \boldsymbol{b}_i^{k+1} &= \sum_{p\in N} g_{ip}\boldsymbol{a}_p^k + \sum_{l\in L}\boldsymbol{a}_l^k - \sum_{p\in N} g_{ip}\boldsymbol{b}_p^k \\
&\ge g_{ij}\left(\boldsymbol{a}_j^k - \boldsymbol{b}_j^k\right) + \sum_{l\in L}\boldsymbol{a}_l^k \\
&= g_{ij}\left(\sum_{p\in N} g_{jp}\boldsymbol{a}_p^{k-1} - \sum_{p\in N} g_{jp}\boldsymbol{b}_p^{k-1} - \sum_{l\in L}\boldsymbol{b}_l^{k-1}\right) + \sum_{l\in L}\boldsymbol{a}_l^k \\
&\ge g_{ij}\left(g_{ij}\left(\boldsymbol{a}_i^{k-1} - \boldsymbol{b}_i^{k-1}\right) - \sum_{l\in L}\boldsymbol{b}_l^{k-1}\right) + \sum_{l\in L}\boldsymbol{a}_l^k \\
&\ge \sum_{l\in L}\boldsymbol{a}_l^k - \sum_{l\in L}\boldsymbol{b}_l^{k-1} \\
&\ge \sum_{l\in L}\boldsymbol{a}_l^k - \sum_{l\in L}\boldsymbol{b}_l^k \ge 0.
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\boldsymbol{a}_i^{k+1} + \boldsymbol{a}_j^{k+1} &= \sum_{p\in N} g_{ip}\boldsymbol{a}_p^k + \sum_{l\in L}\boldsymbol{a}_l^k + \sum_{p\in N} g_{jp}\boldsymbol{a}_p^k \\
&= 2\sum_{p\notin\{i,j\}} (g_{ip}+g_{jp})\,\boldsymbol{a}_p^k + \sum_{l\in L}\boldsymbol{a}_l^k + g_{ij}\boldsymbol{a}_j^k + g_{ij}\boldsymbol{a}_i^k \\
&> 2\sum_{p\notin\{i,j\}} (g_{ip}+g_{jp})\,\boldsymbol{b}_p^k + \sum_{l\in L}\boldsymbol{b}_l^k + g_{ij}\boldsymbol{b}_j^k + g_{ij}\boldsymbol{b}_i^k \\
&= \boldsymbol{b}_i^{k+1} + \boldsymbol{b}_j^{k+1}.
\end{aligned}$$

Note that, the strict inequality holds since there exists some node $p \notin \{i,j\}$ such that $\boldsymbol{a}_p^{k+1} > \boldsymbol{b}_p^{k+1}$. $\square$

Proof of Theorem 1: We prove this result by contradiction. Consider a strategy $\tilde{\boldsymbol{s}} = \left(\tilde{\boldsymbol{G}}(t)\right)_{t=1}^{T}$ which induces non-NSGs in some period. Let t^* be the first time that $\tilde{s}$ induces a non-NSG, i.e., $\tilde{\boldsymbol{G}}(t^*) \notin \mathcal{G}^{NSG}(t^*)$ and $\tilde{\boldsymbol{G}}(t) \in \mathcal{G}^{NSG}(t)$ for any $t < t^*$. We can construct another strategy $\hat{\boldsymbol{s}} = \left(\hat{\boldsymbol{G}}(t)\right)_{t=1}^{T}$ which induces an NSG at period t^* and dominates $\tilde{s}$.

Since $\tilde{\boldsymbol{G}}(t^*)$ is not nested split, we can find two agents i, j such that i's degree is larger than j but j's neighbors are not all i's, i.e., $|N_i\left(\tilde{\boldsymbol{G}}(t^*)\right)| > |N_j\left(\tilde{\boldsymbol{G}}(t^*)\right)|$ while $N_i\left(\tilde{\boldsymbol{G}}(t^*)\right) \cap N_j\left(\tilde{\boldsymbol{G}}(t^*)\right) \neq N_j\left(\tilde{\boldsymbol{G}}(t^*)\right)$.

Construct $\hat{\boldsymbol{s}} = \left(\hat{\boldsymbol{G}}(t)\right)_{t=1}^{T}$ such that $\hat{\boldsymbol{G}}(t) = \tilde{\boldsymbol{G}}(t^*)$ for all $t < t^*$. At period $t \geq t^*$, let $\hat{\boldsymbol{G}}(t)$ be:

1. If $\tilde{\boldsymbol{G}}(t+1) = \tilde{\boldsymbol{G}}(t) + \boldsymbol{E}_{jl}$ for some $l \notin \{i,j\}$, then
   (a) If $\hat{g}_{il}(t) = 0$, let $\hat{\boldsymbol{G}}(t+1) = \hat{\boldsymbol{G}}(t) + \boldsymbol{E}_{il}$;
   (b) If $\hat{g}_{il}(t) = 1$, let $\hat{\boldsymbol{G}}(t+1) = \hat{\boldsymbol{G}}(t) + \boldsymbol{E}_{jl}$.
2. If $\tilde{\boldsymbol{G}}(t+1) = \tilde{\boldsymbol{G}}(t) + \boldsymbol{E}_{il}$ for some $l \notin \{i,j\}$, then
   (a) If $\hat{g}_{il}(t) = 0$, let $\hat{\boldsymbol{G}}(t+1) = \hat{\boldsymbol{G}}(t) + \boldsymbol{E}_{il}$;
   (b) If $\hat{g}_{il}(t) = 1$, let $\hat{\boldsymbol{G}}(t+1) = \hat{\boldsymbol{G}}(t) + \boldsymbol{E}_{jl}$.
3. If $\tilde{\boldsymbol{G}}(t+1) = \tilde{\boldsymbol{G}}(t) + \boldsymbol{E}_{lk}$ for some $l, k \notin \{i,j\}$, then let $\hat{\boldsymbol{G}}(t+1) = \hat{\boldsymbol{G}}(t) + \boldsymbol{E}_{lk}$.

First, we show that the newly constructed strategy $\hat{\boldsymbol{s}}$ is always plausible. Since the link is always reallocated between (i,l) and (j,l) for some $l \notin \{i,j\}$, $\tilde{g}_{lk} = \hat{g}_{l,k}$ for any $l, k \notin \{i,j\}$. This implies that the third step in the algorithm above always holds. We prove the remaining argument by contradiction. Now suppose that $\tilde{\boldsymbol{G}}(t) = \tilde{\boldsymbol{G}}(t-1) + \boldsymbol{E}_{jl}$ and $\hat{g}_{il}(t-1) = 1$. If $\hat{g}_{jl}(t-1) = 1$, together with $\tilde{g}_{jl}(t-1) = 0$, then there exists some step $t^* \leq t' \leq t-2$ such that $\tilde{\boldsymbol{G}}(t') = \tilde{\boldsymbol{G}}(t'-1) + \boldsymbol{E}_{il}$ and $\hat{g}_{il}(t'-1) = 1$. $\hat{g}_{il}(t'-1) = 1$ and $\tilde{g}_{il}(t'-1) = 0$ together imply that there exists some $t^* \leq t'' \leq t'-1$ such that $\tilde{\boldsymbol{G}}(t'') = \tilde{\boldsymbol{G}}(t''-1) + \boldsymbol{E}_{jl}$ and $\hat{g}_{il}(t''-1) = 0$, which therefore contradicts to the plausibility of $\tilde{\boldsymbol{s}}$. Finally, suppose that at step $t^* \leq t \leq T$, $\tilde{\boldsymbol{G}}(t) = \tilde{\boldsymbol{G}}(t-1) + \boldsymbol{E}_{il}$ for some $l \notin \{i,j\}$, $\hat{g}_{il}(t-1) = \hat{g}_{jl}(t-1) = 1$. Since $\tilde{g}_{il}(t-1) = 0$ and $\hat{g}_{il}(t) = 1$, then there exists some $t^* \leq t' \leq t-1$ such that $\tilde{\boldsymbol{G}}(t') = \tilde{\boldsymbol{G}}(t'-1) + \boldsymbol{E}_{jl}$ and $\hat{g}_{il}(t'-1) = 0$. Combining $\hat{g}_{il}(t'-1) = 0$ and $\tilde{g}_{jl}(t'-1) = 0$, we then have that $\hat{g}_{jl}(t'-1) = 0$. However, combining $\hat{g}_{jl}(t-1) = 1$ and $\hat{g}_{jl}(t'-1) = 0$, one then have that there exists some $t'+1 \leq t'' \leq t-1$ such that either $\tilde{\boldsymbol{G}}(t'') = \tilde{\boldsymbol{G}}(t''-1) + \boldsymbol{E}_{jl}$ or $\tilde{\boldsymbol{G}}(t'') = \tilde{\boldsymbol{G}}(t''-1) + \boldsymbol{E}_{il}$. However, both cases violate the plausibility of $\tilde{\boldsymbol{s}}$.

Moreover, for each period t, define a set of agents

$$L = \{l \in N : \hat{g}_{il}(t) > \tilde{g}_{il}(t)\} = \{l \in N : \hat{g}_{jl}(t) > \tilde{g}_{jl}(t)\}.$$

Note that, by this construction $\hat{\boldsymbol{G}}(t) = \tilde{\boldsymbol{G}}(t) - \sum_{l \in L} \boldsymbol{E}_{jl} + \sum_{l \in L} \boldsymbol{E}_{il}$. Thus, by Lemma 1, $u^*\left(\hat{\boldsymbol{G}}(t)\right) > u^*\left(\tilde{\boldsymbol{G}}(t)\right)$ for any t. We can then iterate this procedure to produce a weakly better strategy $s^* = (\boldsymbol{G}^*(t))_{t=1}^{T}$ until $\boldsymbol{G}^*(t) \in \mathcal{G}^{NSG}(t)$ for all t.

Note that, when $\delta_{t^*} > 0$, the constructed strategy $\hat{s}$ is strictly better than $\tilde{s}$. Therefore, the optimal strategy s^* must induce an NSG at period t such that $\delta_t > 0$. Suppose not. Let $\bar{t} = \min\{t | \boldsymbol{G}^*(t) \notin \mathcal{NSG}(t)\}$ and $\underline{t} = \min\{t | \boldsymbol{G}^*(t) \notin \mathcal{NSG}(t) \text{ and } \delta_t > 0\}$. Apparently, $\bar{t} \leq \underline{t}$. By Algorithm 1, we can construct $\hat{s}^* = \left(\hat{\boldsymbol{G}}^*(t)\right)_{t=1}^{T}$ such that $\hat{\boldsymbol{G}}^*(t) \in \mathcal{NSG}(t)$ for any $t < \underline{t}$ and $\hat{s}^*$ induces same payoff with s^* to the social planner since $\delta_t = 0$ for $\bar{t} \leq t < \underline{t}$. Then, $\underline{t} = \min\left\{t | \hat{\boldsymbol{G}}^*(t) \notin \mathcal{NSG}(t)\right\}$. By Algorithm 1, we can find another strategy induces strictly higher payoff than $\hat{s}^*$ which contradict with the fact that s^* is optimal. □

Proof of Lemma 2: Similar with Proof of Lemma 1, we prove an equivalent statement here: Given a network $\boldsymbol{G}$ and two distinct nodes i, j such that $N_j(\boldsymbol{G}) \setminus \{i\} \subsetneq N_i(\boldsymbol{G}) \setminus \{j\}$. Then, for any set of nodes $L = \{l_1, ..., l_k\} \subseteq N \setminus \{i,j\}$ such that $L \cap N_i(\boldsymbol{G}) = \emptyset$ we have

$$\begin{aligned} a^*\left(\boldsymbol{G} + \sum_{l \in L} \boldsymbol{E}_{il}\right) &> a^*\left(\boldsymbol{G} + \sum_{l \in L} \boldsymbol{E}_{jl}\right) \text{ and} \\ u^*\left(\boldsymbol{G} + \sum_{l \in L} \boldsymbol{E}_{il}\right) &> u^*\left(\boldsymbol{G} + \sum_{l \in L} \boldsymbol{E}_{jl}\right). \end{aligned}$$

Define $x_k^{(a)}$ iteratively as

$$x_k^{(0)} = 0;\, x_k^{(a+1)} = \psi\left(\sum_{k' \in N_k\left(\boldsymbol{G} + \sum_{l \in L} \boldsymbol{E}_{il}\right)} x_{k'}^{(a)}\right).$$

Similarly, define $y_k^{(a)}$ iteratively as

$$y_k^{(0)} = 0;\; y_k^{(a+1)} = \psi\left(\sum_{k' \in N_k\left(\boldsymbol{G} + \sum_{l\in L} \boldsymbol{E}_{jl}\right)} y_{k'}^{(a)}\right).$$

We use mathematical induction to show that for all $a \geq 0$, the following four arguments hold:

1. $x_k^{(a)} \geq y_k^{(a)}$, $\forall k \neq j$.
2. $x_i^{(a)} \geq y_j^{(a)}$;
3. $x_i^{(a)} + x_j^{(a)} \geq y_i^{(a)} + y_j^{(a)}$;
4. $x_k^{(a)}$ and $y_k^{(a)}$ is increasing in a for any k.

When $a = 0$ and 1, the four arguments trivially hold. Assume that these four arguments hold for any $a \leq a^*$ where $a^* \geq 1$.

First, we show that $x_k^{(a^*+1)} \geq y_k^{(a^*+1)}$ for any $k \notin \{i,j\} \cup L$.

$$\begin{aligned} x_k^{(a^*+1)} &= \psi\left(\sum_{k' \notin \{i,j\}} g_{kk'} x_{k'}^{(a^*)} + g_{ki} x_i^{(a^*)} + g_{kj} x_j^{(a^*)}\right) \\ &\geq \psi\left(\sum_{k' \notin \{i,j\}} g_{kk'} y_{k'}^{(a^*)} + g_{ki} x_i^{(a^*)} + g_{kj} x_j^{(a^*)}\right) \\ &\geq \psi\left(\sum_{k' \notin \{i,j\}} g_{kk'} y_{k'}^{(a^*)} + g_{ki} y_i^{(a^*)} + g_{kj} y_j^{(a^*)}\right) = y_k^{(a^*+1)}. \end{aligned}$$

The first inequality follows from the fact that $x_k^{(a^*)} \geq y_k^{(a^*)}$, $\forall k \notin \{i,j\}$. The second inequality follows from the that that $g_{ki} \geq g_{kj}$, $x_i^{(a^*)} \geq y_j^{(a^*)}$ and $x_i^{(a^*)} + x_j^{(a^*)} \geq y_i^{(a^*)} + y_j^{(a^*)}$.

Note that, for some node k such that $g_{ki} = 1$ and $g_{kj} = 0$, we have $x_k^{(a^*+1)} > y_k^{(a^*+1)}$.

Second, we show that, for any $l \in L$, $x_l^{(a^*+1)} \geq y_l^{(a^*+1)}$. The inequality holds since

$$x_l^{(a^*+1)} = \psi\left(\sum_{k'} g_{lk'} x_{k'}^{(a^*)} + x_i^{(a^*)}\right) \geq \psi\left(\sum_{k'} g_{lk'} y_{k'}^{(a^*)} + y_i^{(a^*)}\right) = y_l^{(a^*+1)}.$$

Third, we show that $x_i^{(a^*+1)} \geq y_i^{(a^*+1)}$. Note that

$$x_i^{(a^*+1)} = \psi\left(g_{ij}x_j^{(a^*)} + \sum_{k\neq j} g_{ik}x_k^{(a^*)} + \sum_{l\in L} x_l^{(a^*)}\right),$$

the inequality trivially hold when $g_{ij}=0$ since $x_k^{(a^*)} \geq y_k^{(a^*)}$, $\forall k \neq j$. When $g_{ij}=1$, we have

$$\begin{aligned}
x_i^{(a^*+1)} &= \psi\left(x_j^{(a^*)} + \sum_{k\neq j} g_{ik}x_k^{(a^*)} + \sum_{l\in L} x_l^{(a^*)}\right) \\
&= \psi\left(\psi\left(x_i^{(a^*-1)} + \sum_{k\neq i} g_{jk}x_k^{(a^*-1)}\right) + \sum_{k\neq j} g_{ik}x_k^{(a^*)} + \sum_{l\in L} x_l^{(a^*)}\right) \\
&\geq \psi\left(\psi\left(x_i^{(a^*-1)} + \sum_{k\neq i} g_{jk}x_k^{(a^*-1)}\right) + \sum_{k\neq j} g_{ik}x_k^{(a^*)} + \sum_{l\in L} x_l^{(a^*-1)}\right) \\
&\geq \psi\left(\psi\left(y_i^{(a^*-1)} + \sum_{k\neq i} g_{jk}y_k^{(a^*-1)}\right) + \sum_{k\neq j} g_{ik}y_k^{(a^*)} + \sum_{l\in L} y_l^{(a^*-1)}\right) \\
&\geq \psi\left(\psi\left(y_i^{(a^*-1)} + \sum_{k\neq i} g_{jk}y_k^{(a^*-1)} + \sum_{l\in L} y_l^{(a^*-1)}\right) + \sum_{k\neq j} g_{ik}y_k^{(a^*)}\right) = y_i^{(a^*+1)}.
\end{aligned}$$

The first inequality follows from the fact that $x_k^{(a)}$ is increasing in a. The second inequality follows from $x_k^{(a)} \geq y_k^{(a)}$, $\forall k \neq j$ and $a \leq a^*$. The third inequality follows from the fact that $\psi'(\cdot) \leq 1$, since $\psi'(\cdot) \leq \frac{1}{\lambda_{\max}\left(\boldsymbol{G}+\sum_{l\in L}\boldsymbol{E}_{il}\right)}$ and the largest eigenvalue of a non-empty network is no less than 1.

Then, we show that $x_i^{(a^*+1)} \geq y_j^{(a^*+1)}$. Similarly, the argument trivially holds when $g_{ij}=0$. We focus on the case of $g_{ij}=1$.

$$
\begin{aligned}
x_i^{(a^*+1)} &= \psi\left(x_j^{(a^*)}+\sum_{k\neq j} g_{ik}x_k^{(a^*)}+\sum_{l\in L} x_l^{(a^*)}\right)\\
&= \psi\left(\psi\left(x_i^{(a^*-1)}+\sum_{k\neq i} g_{jk}x_k^{(a^*-1)}\right)+\sum_{k\neq j} g_{ik}x_k^{(a^*)}+\sum_{l\in L} x_l^{(a^*)}\right)\\
&\geq \psi\left(\psi\left(y_j^{(a^*-1)}+\sum_{k\neq i} g_{jk}x_k^{(a^*-1)}\right)+\sum_{k\neq j} g_{ik}x_k^{(a^*)}+\sum_{l\in L} x_l^{(a^*)}\right)\\
&\geq \psi\left(\psi\left(y_j^{(a^*-1)}+\sum_{k\neq i} g_{jk}y_k^{(a^*-1)}\right)+\sum_{k\neq j} g_{ik}y_k^{(a^*)}+\sum_{l\in L} y_l^{(a^*)}\right)\\
&= \psi\left(\psi\left(y_j^{(a^*-1)}+\sum_{k\neq i} g_{jk}y_k^{(a^*-1)}\right)+\sum_{k\neq j,k\in N_i(\boldsymbol{G})\setminus N_j(\boldsymbol{G})} y_k^{(a^*)}+\sum_{k\neq i} g_{jk}y_k^{(a^*)}+\sum_{l\in L} y_l^{(a^*)}\right)\\
&\geq \psi\left(\psi\left(y_j^{(a^*-1)}+\sum_{k\neq i} g_{jk}y_k^{(a^*-1)}\right)+\sum_{k\neq j,k\in N_i(\boldsymbol{G})\setminus N_j(\boldsymbol{G})} y_k^{(a^*-1)}+\sum_{k\neq i} g_{jk}y_k^{(a^*)}+\sum_{l\in L} y_l^{(a^*)}\right)\\
&\geq \psi\left(\psi\left(y_j^{(a^*-1)}+\sum_{k\neq i} g_{jk}y_k^{(a^*-1)}+\sum_{k\neq j,k\in N_i(\boldsymbol{G})\setminus N_j(\boldsymbol{G})} y_k^{(a^*-1)}\right)+\sum_{k\neq i} g_{jk}y_k^{(a^*)}+\sum_{l\in L} y_l^{(a^*)}\right)\\
&= \psi\left(\psi\left(y_j^{(a^*-1)}+\sum_{k\neq j} g_{ik}y_k^{(a^*-1)}\right)+\sum_{k\neq i} g_{jk}y_k^{(a^*)}+\sum_{l\in L} y_l^{(a^*)}\right)=y_j^{(a^*+1)}.
\end{aligned}
$$

The last inequality comes from $\psi'(\cdot)<1$ and the others follow the inductive conditions.

Finally, we show that $x_i^{(a^*+1)}+x_j^{(a^*+1)}\geq y_i^{(a^*+1)}+y_j^{(a^*+1)}$. Note that, by the convexity of $\psi(\cdot)$, for any four real numbers a,b,c,d, we have $\psi(a)+\psi(b)\geq\psi(c)+\psi(d)$ if $a+b\geq c+d$ and $\max\{a,b\}\geq\max\{c,d\}$. Therefore, to prove the inequality, we we only need to show the following two arguments:

$$
\begin{aligned}
& \left(x_j^{(a^*)}+\sum_{k\neq j} g_{ik}x_k^{(a^*)}+\sum_{l\in L} x_l^{(a^*)}\right)+\left(x_i^{(a^*)}+\sum_{k\neq i} g_{jk}x_k^{(a^*)}\right)\\
\geq\ & \left(y_j^{(a^*)}+\sum_{k\neq j} g_{ik}y_k^{(a^*)}\right)+\left(y_i^{(a^*)}+\sum_{k\neq i} g_{jk}y_k^{(a^*)}+\sum_{l\in L} y_l^{(a^*)}\right);
\end{aligned}
$$

$$
\begin{aligned}
& \max\left\{ x_j^{(a^*)} + \sum_{k\neq j} g_{ik} x_k^{(a^*)} + \sum_{l\in L} x_l^{(a^*)},\ x_i^{(a^*)} + \sum_{k\neq i} g_{jk} x_k^{(a^*)} \right\} \\
\geq \ & \max\left\{ y_j^{(a^*)} + \sum_{k\neq j} g_{ik} y_k^{(a^*)},\ y_i^{(a^*)} + \sum_{k\neq i} g_{jk} y_k^{(a^*)} + \sum_{l\in L} y_l^{(a^*)} \right\};
\end{aligned}
$$

The first argument holds since

$$
\begin{aligned}
& x_j^{(a^*)} + \sum_{k\neq j} g_{ik} x_k^{(a^*)} + \sum_{l\in L} x_l^{(a^*)} + x_i^{(a^*)} + \sum_{k\neq i} g_{jk} x_k^{(a^*)} \\
\geq \ & y_i^{(a^*)} + y_j^{(a^*)} + \sum_{k\neq j} g_{ik} y_k^{(a^*)} + \sum_{l\in L} y_l^{(a^*)} + \sum_{k\neq i} g_{jk}^{(a^*)} y_k^{(a^*)}
\end{aligned}
$$

by the inductive conditions. To prove the second argument, note that we have shown $x_i^{(a^*+1)} \geq \max\left\{ y_i^{(a^*+1)}, y_j^{(a^*+1)} \right\}$. Therefore,

$$
\begin{aligned}
& \max\left\{ x_j^{(a^*)} + \sum_{k\neq j} g_{ik} x_k^{(a^*)} + \sum_{l\in L} x_l^{(a^*)},\ x_i^{(a^*)} + \sum_{k\neq i} g_{jk} x_k^{(a^*)} \right\} \\
\geq \ & x_j^{(a^*)} + \sum_{k\neq j} g_{ik} x_k^{(a^*)} + \sum_{l\in L} x_l^{(a^*)} \\
\geq \ & \max\left\{ y_j^{(a^*)} + \sum_{k\neq j} g_{ik} y_k^{(a^*)},\ y_i^{(a^*)} + \sum_{k\neq i} g_{jk} y_k^{(a^*)} + \sum_{l\in L} y_l^{(a^*)} \right\}.
\end{aligned}
$$

By the four inductive arguments, $\sum_k x_k^{(a)} > \sum_k y_k^{(a)}$ for any a (the inequality is strict since for some node k such that $g_{ki} = 1$ and $g_{kj} = 0$, we have $x_k^{(a^*+1)} > y_k^{(a^*+1)}$). Moreover, the equilibrium choice of each agent k in network $\boldsymbol{G}+\sum_{l\in L} \boldsymbol{E}_{il}$ is limit of $x_k^{(a)}$, the aggregate activities induced by $\boldsymbol{G}+\sum_{l\in L} \boldsymbol{E}_{il}$ is larger than that by $\boldsymbol{G}+\sum_{l\in L} \boldsymbol{E}_{jl}$

To prove $u^*\left(\boldsymbol{G}+\sum_{l\in L} \boldsymbol{E}_{il}\right) > u^*\left(\boldsymbol{G}+\sum_{l\in L} \boldsymbol{E}_{jl}\right)$, we need the following four arguments

1. $x_k^{(a)} \geq y_k^{(a)}, \forall k \neq j$;
2. $x_i^{(a)} \geq y_j^{(a)}$;
3. $\left(x_i^{(a)}\right)^2 + \left(x_j^{(a)}\right)^2 \geq \left(y_i^{(a)}\right)^2 + \left(y_j^{(a)}\right)^2$;
4. $x_k^{(a)}$ and $y_k^{(a)}$ is increasing in a for any k.

Note that, the third argument hold since the square function is convex, and as we have shown $x_i^{(a)} + x_j^{(a)} \geq y_i^{(a)} + y_j^{(a)}$, $x_i^{(a)} \geq \max\left\{y_i^{(a)}, y_j^{(a)}\right\}$. $\square$

Proof of Proposition 1: Based on Lemma 2, the remaining proof is same with that of Theorem 1. $\square$

Proof of Lemma 3: Let $\triangle u(t) \equiv \min_{\boldsymbol{G} \in \mathbb{S}(\boldsymbol{QC}(t-1)) \backslash \{\boldsymbol{QC}(t)\}} (u^*(\boldsymbol{QC}(t)) - u^*(\boldsymbol{G}))$ be the welfare gap between first and second best in period t, where $\triangle u(t) = u^*(\boldsymbol{QC}(t))$ if $\mathcal{S}(\boldsymbol{QC}(t-1)) = \{\boldsymbol{QC}(t)\}$. Denote $\underline{\triangle u} = \min_{1 \leq t \leq T} \triangle u(t)$ as the least welfare gap during the formation process. The adjacency matrix of complete graph is denoted by $\boldsymbol{C}$. Note that, $u^*(\boldsymbol{C})$ is the maximum welfare can be generated. Therefore, if $\frac{\delta_{t+1}}{\delta_t} < \varepsilon$ for any t, $\sum_{t=1}^{T} \varepsilon^t u^*(\boldsymbol{C})$ is the largest welfare can be obtained during the network formation process. Furthermore, we have $\sum_{t=1}^{T} \varepsilon^t u^*(\boldsymbol{C}) \leq T\varepsilon u^*(\boldsymbol{C})$ when $\varepsilon \leq 1$. As a result, when $T\varepsilon u^*(\boldsymbol{C}) \leq \underline{\triangle u}$, the social planner has no incentive to choose the suboptimum at any period. That is, when $\frac{\delta_{t+1}}{\delta_t} < \varepsilon$ for any t, where $\varepsilon \leq \frac{\underline{\triangle u}}{Tu^*(\boldsymbol{C})}$, the social planner is myopic. $\square$

Proof of Theorem 2: We introduce some notations firstly. Consider a quasi-complete network $(N, \boldsymbol{G})$ with t links. Suppose network $\boldsymbol{G}$ contains a complete subgraph formed by $p \geq 3$ nodes and $\frac{p(p-1)}{2} + 1 \leq t \leq \frac{p(p+1)}{2}$. That is, the first p agents form a complete network, the $p+1$-th player connects with first $t - \frac{p(p-1)}{2}$ players and the last $n - (p+1)$ players are isolated. We use $\boldsymbol{QC}(t+1) = \boldsymbol{G} + \boldsymbol{E}_{p+1, t+1-\frac{p(p-1)}{2}}$ and $\bar{\boldsymbol{G}}(t+1) = \boldsymbol{G} + \boldsymbol{E}_{p+2,1}$ to denote the two nested split networks obtained by adding a single link from $\boldsymbol{G}$. Thus, $\left[2, t - \frac{p(p-1)}{2}\right] \cup \left[t + 2 - \frac{p(p-1)}{2}, p\right]$ is the set of agents those whom are in the complete subnetwork and whose neighbors are the same in $\boldsymbol{QC}(t+1)$ and $\bar{\boldsymbol{G}}(t+1)$.

Theorem 2 directly follows Lemma 4.

In the proof, we omit the term of the total links t and use $\boldsymbol{QC}$ and $\bar{\boldsymbol{G}}$ to denote the two NSGs. The proof consists of two major steps.

Step 1: $\boldsymbol{E}_{p+1,i} \cdot \boldsymbol{QC}^k \cdot \mathbf{1} \geq \boldsymbol{E}_{p+1,i} \cdot \bar{\boldsymbol{G}}^k \cdot \mathbf{1}$, for any agent $i \in \left[2, t - \frac{p(p-1)}{2}\right] \cup \left[t + 2 - \frac{p(p-1)}{2}, p\right]$ and any positive integer k. Moreover, $\mathbf{1}' \cdot \boldsymbol{E}_{1, t-\frac{p(p-1)}{2}+1} \cdot \boldsymbol{QC}^k \cdot \mathbf{1} \geq \mathbf{1}' \cdot \boldsymbol{E}_{1, t-\frac{p(p-1)}{2}+1} \cdot \bar{\boldsymbol{G}}^k \cdot \mathbf{1}$.

Note that $\boldsymbol{E}_{i,j} \cdot \boldsymbol{G}^k \cdot \mathbf{1}$ is a n-dimensional vector with i-th and j-th elements being total number of k-length paths starting from nodes i and j in network $\boldsymbol{G}$ respectively. Thus, the first statement in step 2 is to show that, for any agent in $\left[2, t - \frac{p(p-1)}{2}\right] \cup \left[t + 2 - \frac{p(p-1)}{2}, p\right]$, the total number of k-length paths starting from this agent in network $\boldsymbol{QC}$ is larger than that in $\bar{\boldsymbol{G}}$. The second statement is to show that the sum of k-length paths from agents 1 and $t - \frac{p(p-1)}{2} + 1$ in network $\boldsymbol{QC}$ is larger than that in $\bar{\boldsymbol{G}}$. We adapt mathematical induction to show step 1. Denote the j-th element of vectors $\boldsymbol{QC}^l \cdot 1$ and $\bar{\boldsymbol{G}}^l \cdot 1$ as $\mathring{a}_j^l$ and $\bar{a}_j^l$ respectively.

- $k=1$, $\boldsymbol{E}_{p+1,i}\cdot\boldsymbol{QC}^1\cdot\mathbf{1}$ is a vector with i-th entry $\mathring{a}_i^1=p-1$ and $p+1$-th entry $\mathring{a}_{p+1}^1=t+1-\frac{p(p-1)}{2}$ and other elements $\mathring{a}_j^1=0$ for all $j\neq i\in\left[2,t-\frac{p(p-1)}{2}\right]\cup\left[t+2-\frac{p(p-1)}{2},p\right]$ and $j\neq p+1$; Moreover, $\bar{a}_i^1=p-1$ and $\bar{a}_{p+1}^1=t-\frac{p(p-1)}{2}$. Apparently, $\mathring{a}_i^1=\bar{a}_i^1$ and $\mathring{a}_{p+1}^1>\bar{a}_{p+1}^1$. Moreover, $\mathbf{1}'\cdot\boldsymbol{E}_{1,t-\frac{p(p-1)}{2}+1}\cdot\boldsymbol{QC}^1\cdot\mathbf{1}=\mathbf{1}'\cdot\boldsymbol{E}_{1,t-\frac{p(p-1)}{2}+1}\cdot\bar{\boldsymbol{G}}^1\cdot\mathbf{1}$ since the sum of 1's and $t+1-\frac{p(p-1)}{2}$'s degree are identical in $\boldsymbol{QC}$ and $\bar{\boldsymbol{G}}$.

- Suppose for all $k\leq l$ the statements holds. That is, $\bar{a}_i^l\leq\mathring{a}_i^l$ for all $i\in\left[2,t-\frac{p(p-1)}{2}\right]\cup\left[t+2-\frac{p(p-1)}{2},p+1\right]$ and $\bar{a}_1^l+\bar{a}_{t-\frac{p(p-1)}{2}+1}^l\leq\mathring{a}_1^l+\mathring{a}_{t-\frac{p(p-1)}{2}+1}^l$. For $k=l+1$, we have

$$
\begin{aligned}
&\boldsymbol{E}_{p+1,i}\cdot\bar{\boldsymbol{G}}^{l+1}\cdot\mathbf{1}=\boldsymbol{E}_{p+1,i}\cdot\big(\boldsymbol{G}+\boldsymbol{E}_{p+2,1}\big)\cdot\bar{\boldsymbol{G}}^l\cdot\mathbf{1}=\boldsymbol{E}_{p+1,i}\cdot\boldsymbol{G}\cdot\bar{\boldsymbol{G}}^l\cdot\mathbf{1}\\
= &\begin{bmatrix}\mathbf{0}\\ \sum_{j\in N}g_{p+1,j}\bar{a}_j^l\\ \mathbf{0}\\ \sum_{j\in N}g_{ij}\bar{a}_j^l\\ \mathbf{0}\end{bmatrix}\leq\begin{bmatrix}\mathbf{0}\\ \sum_{j\in N}g_{p+1,j}\bar{a}_j^l+\bar{a}_{t-\frac{p(p-1)}{2}+1}^l\\ \mathbf{0}\\ \sum_{j\in N}g_{ij}\bar{a}_j^l\\ \mathbf{0}\end{bmatrix}\\
&\overset{p+1\text{ connects with first }t-\frac{p(p-1)}{2}\text{ nodes}}{\underset{i\text{ belongs to the complete subnetwork}}{=}}\begin{bmatrix}\mathbf{0}\\ \sum_{j\in\left[2,t-\frac{p(p-1)}{2}\right]}g_{p+1,j}\bar{a}_j^l+\bar{a}_1^l+\bar{a}_{t-\frac{p(p-1)}{2}+1}^l\\ \mathbf{0}\\ \sum_{j\in\left[2,t-\frac{p(p-1)}{2}\right]\cup\left[t+2-\frac{p(p-1)}{2},p+1\right]}g_{ij}\bar{a}_j^l+\bar{a}_1^l+\bar{a}_{t-\frac{p(p-1)}{2}+1}^l\\ \mathbf{0}\end{bmatrix}\\
&\underset{\text{by induction}}{\leq}\begin{bmatrix}\mathbf{0}\\ \sum_{j\in\left[2,t-\frac{p(p-1)}{2}\right]}g_{p+1,j}^l\mathring{a}_j^l+\mathring{a}_1^l+\mathring{a}_{t-\frac{p(p-1)}{2}+1}^l\\ \mathbf{0}\\ \sum_{j\in\left[2,t-\frac{p(p-1)}{2}\right]\cup\left[t+2-\frac{p(p-1)}{2},p+1\right]}g_{ij}^l\mathring{a}_j^l+\mathring{a}_1^l+\mathring{a}_{t-\frac{p(p-1)}{2}+1}^l\\ \mathbf{0}\end{bmatrix}\\
= &\ \boldsymbol{E}_{p+1,i}\cdot\boldsymbol{QC}^l\cdot\mathbf{1}.
\end{aligned}
$$

On the other hand, by the decomposition of l-length paths

$$\begin{aligned}
\mathring{a}_1^l + \mathring{a}_{t-\frac{p(p-1)}{2}+1}^l &= \sum_{j\in[2,p+1]} \mathring{a}_j^{l-1} + \sum_{j\in[1,p+1]\setminus\left\{t-\frac{p(p-1)}{2}+1\right\}} \mathring{a}_j^{l-1}; \\
\bar{a}_1^l + \bar{a}_{t-\frac{p(p-1)}{2}+1}^l &= \sum_{j\in[2,p+2]} \bar{a}_j^{l-1} + \sum_{j\in[1,p]\setminus\left\{t-\frac{p(p-1)}{2}+1\right\}} \bar{a}_j^{l-1} \\
&\le \sum_{j\in[2,p+1]} \bar{a}_j^{l-1} + \bar{a}_{p+1}^{l-1} + \sum_{j\in[1,p]\setminus\left\{t-\frac{p(p-1)}{2}+1\right\}} \bar{a}_j^{l-1}.
\end{aligned}$$

By the induction condition when $k = l-1$, we have $\mathring{a}_1^l + \mathring{a}_{t-\frac{p(p-1)}{2}+1}^l \ge \bar{a}_1^l + \bar{a}_{t-\frac{p(p-1)}{2}+1}^l$. $\square$

Step 2: $\mathbf{1}' \cdot \boldsymbol{QC}^k \cdot \mathbf{1} \ge \mathbf{1}' \cdot \bar{\boldsymbol{G}}^k \cdot \mathbf{1}$ for all positive integer k.

- When $k = 1$ we have $\mathbf{1}' \cdot \boldsymbol{QC} \cdot \mathbf{1} = \mathbf{1}' \cdot \bar{\boldsymbol{G}} \cdot \mathbf{1}$ since the sum of degree of the two networks are identical.
- When $k = 2$, apparently $\mathbf{1}' \cdot \boldsymbol{QC}^2 \cdot \mathbf{1} \ge \mathbf{1}' \cdot \bar{\boldsymbol{G}}^2 \cdot \mathbf{1}$ since the sum of degree square of $\boldsymbol{QC}$ is larger than that of $\bar{\boldsymbol{G}}$.
- Now we focus on the case in which $k \ge 3$.

$$\begin{aligned}
\mathbf{1}' \cdot \boldsymbol{QC}^k \cdot \mathbf{1} &= \underbrace{\sum_{i\in\left[1,t-\frac{p(p-1)}{2}+1\right]} \sum_{j\in[1,p+1]\setminus\{i\}} \mathring{a}_j^{k-1}}_{\text{total number of } k\text{-length paths of players connects with } p+1} \\
&\quad + \underbrace{\left(\frac{p(p+1)}{2} - t - 1\right) \cdot \sum_{j\in[1,p-1]} \mathring{a}_j^{k-1}}_{\text{total number of } k\text{-length paths of players in the complete subgraph that does not connect with } p+1} \\
&\quad + \underbrace{\sum_{j\in\left[1,t-\frac{p(p-1)}{2}+1\right]} \mathring{a}_j^{k-1}}_{\text{total number of } k\text{-length paths of } p+1} \\
&= p\left(\mathring{a}_1^{k-1} + \mathring{a}_{t-\frac{p(p-1)}{2}+1}^{k-1}\right) + \left(t - \frac{p(p-1)}{2} + 1\right)\mathring{a}_{p+1}^{k-1} \\
&\quad + \sum_{i\in\left[2,t-\frac{p(p-1)}{2}\right]\cup\left[t+2-\frac{p(p-1)}{2},p+1\right]} \sum_{j\in[1,p+1]\setminus\{i\}} \mathring{a}_j^{k-1}
\end{aligned}$$

$$
\begin{aligned}
\mathbf{1}' \cdot \bar{\boldsymbol{G}}^k \cdot \mathbf{1} \;=\;& \underbrace{\sum_{j \in [2, p+2]} \bar{a}_j^{k-1}}_{\text{total number of } k\text{-length paths of } 1} + \underbrace{\bar{a}_1^{k-1}}_{\text{total number of } k\text{-length paths of } p+2} \\
&+ \underbrace{\sum_{i \in \left[2, t - \frac{p(p-1)}{2}\right]} \sum_{j \in [1, p+1] \setminus \{i\}} \bar{a}_j^{k-1}}_{\text{total number of } k\text{-length paths of players connects with } p+1 \text{ except } 1} \\
&+ \underbrace{\sum_{i \in \left[t - \frac{p(p-1)}{2} + 1, p\right]} \sum_{j \in [1, p-1] \setminus \{i\}} \bar{a}_j^{k-1}}_{\text{total number of } k\text{-length paths of players in the complete subgraph that does not connect with } p+1} \\
&+ \underbrace{\sum_{j \in \left[1, t - \frac{p(p-1)}{2}\right]} \bar{a}_j^{k-1}}_{\text{total number of } k\text{-length paths of } p+1} . \\
=\;& (p-1)\left(\bar{a}_1^{k-1} + \bar{a}_{t - \frac{p(p-1)}{2} + 1}^{k-1}\right) + 2\bar{a}_1^{k-1} + \bar{a}_{p+2}^{k-1} \\
&+ \left(t - \frac{p(p-1)}{2}\right) \mathring{a}_{p+1}^{k-1} + \sum_{i \in \left[2, t - \frac{p(p-1)}{2}\right] \cup \left[t + 2 - \frac{p(p-1)}{2}, p+1\right]} \sum_{j \in [1, p+1] \setminus \{i\}} \bar{a}_j^{k-1}
\end{aligned}
$$

By step 1 we have $\bar{a}_i^k \le \mathring{a}_i^k$ for all $i \in \left[2, t - \frac{p(p-1)}{2}\right] \cup \left[t + 2 - \frac{p(p-1)}{2}, p\right]$, and $\bar{a}_1^k + \bar{a}_{t - \frac{p(p-1)}{2} + 1}^k \le \mathring{a}_1^k + \mathring{a}_{t - \frac{p(p-1)}{2} + 1}^k$. Therefore, to prove $\mathbf{1}' \cdot \boldsymbol{QC}^k \cdot \mathbf{1} \ge \mathbf{1}' \cdot \bar{\boldsymbol{G}}^k \cdot \mathbf{1}$ we only need to show

$$
\mathring{a}_1^{k-1} + \mathring{a}_{t - \frac{p(p-1)}{2} + 1}^{k-1} + \mathring{a}_{p+1}^{k-1} \ge 2\bar{a}_1^{k-1} + \bar{a}_{p+2}^{k-1}.
$$

Moreover, since

$$
\begin{aligned}
\mathring{a}_1^{k-1} + \mathring{a}_{t - \frac{p(p-1)}{2} + 1}^{k-1} + \mathring{a}_{p+1}^{k-1} &= \sum_{j \in [2, p+1]} \mathring{a}_j^{k-2} + \sum_{j \in [1, p+1] \setminus \left\{t - \frac{p(p-1)}{2} + 1\right\}} \mathring{a}_j^{k-2} + \sum_{j \in \left[1, t - \frac{p(p-1)}{2} + 1\right]} \mathring{a}_j^{k-2} \\
2\bar{a}_1^{k-1} + \bar{a}_{p+2}^{k-1} &= 2 \sum_{j \in [2, p+2]} \bar{a}_j^{k-2} + \bar{a}_1^{k-2},
\end{aligned}
$$

we could therefore show that

$$
\begin{aligned}
& \mathring{a}_1^{k-1} + \mathring{a}_{t - \frac{p(p-1)}{2} + 1}^{k-1} + \mathring{a}_{p+1}^{k-1} - 2\bar{a}_1^{k-1} + \bar{a}_{p+2}^{k-1} \\
\ge\;& \mathring{a}_1^{k-2} + \mathring{a}_{t - \frac{p(p-1)}{2} + 1}^{k-2} + \sum_{j \in \left[2, t - \frac{p(p-1)}{2}\right]} \mathring{a}_j^{k-2} - 2\bar{a}_{p+2}^{k-2} \\
\ge\;& 0. \;\square
\end{aligned}
$$

Proof of Corollary 1: The proof is clearly stated in the main text. $\square$

Proof of Proposition 2: Bernardo M. Ábrego (2009)'s Theorem 2 and corollary 2 state that the sum of squares of degrees in quasi-star network is strictly larger than that in other networks when the total number of links $t \in [4, \frac{n^2-3n}{4})$, while that in quasi-complete network is the strict largest one when $t \in (\frac{n^2+n}{4}, \frac{n^2-n}{2}]$. Therefore, when $t \in [4, \frac{n^2-3n}{4})$, there exists $\varepsilon_1 > 0$ such that

$$\mathbf{1}'\boldsymbol{QS}(t)\,\boldsymbol{QS}(t)\,\mathbf{1} \geq \mathbf{1}'\boldsymbol{GG}\mathbf{1} + \varepsilon_1,$$

for any $\boldsymbol{G}$ with t links.

If $t \in (\frac{n^2+n}{4}, \frac{n^2-n}{2}]$, then there exists $\varepsilon_2 > 0$ such that

$$\mathbf{1}'\boldsymbol{QC}(t)\,\boldsymbol{QC}(t)\,\mathbf{1} \geq \mathbf{1}'\boldsymbol{GG}\mathbf{1} + \varepsilon_2,$$

for any $\boldsymbol{G}$ with t links.

Note that, the aggregate welfare activities of $\Gamma(\boldsymbol{G})$ is given by

$$\frac{1}{2} \cdot \mathbf{1}'(\boldsymbol{I} - \phi\boldsymbol{G})^{-1}(\boldsymbol{I} - \phi\boldsymbol{G})^{-1}\,\mathbf{1} = \frac{1}{2}\sum_{k=0}(k+1)\,\phi^k \mathbf{1}'\boldsymbol{G}^k\mathbf{1}$$

Moreover, $\sum_{k=3} \phi^k \mathbf{1}'\boldsymbol{G}^k\mathbf{1} \leq \sum_{k=3} \phi^k \mathbf{1}'\boldsymbol{C}^k\mathbf{1}$ where $\boldsymbol{C}$ is complete network formed by n players.

By mathematical induction, it is easy to show that, for any $k \in Z^+$, $\left[\boldsymbol{C}^k\right]_{ii} = \frac{(n-1)^k-(n-1)(-1)^{k-1}}{n}$ and $\left[\boldsymbol{C}^k\right]_{ij} = \frac{(n-1)^k+(-1)^{k-1}}{n}$ $\forall i$ and $j \neq i$. Thus, $\mathbf{1}'\boldsymbol{C}^k\mathbf{1} = n \cdot (n-1)^k$.

For any network $\boldsymbol{G}$, we have

$$\mathbf{1}'(\boldsymbol{I} - \phi\boldsymbol{G})^{-1}(\boldsymbol{I} - \phi\boldsymbol{G})^{-1}\,\mathbf{1} < \mathbf{1}'\,\boldsymbol{I}\mathbf{1} + 2\phi\mathbf{1}'\,\boldsymbol{G}\mathbf{1} + 3\phi^2\mathbf{1}'\,\boldsymbol{GG}\mathbf{1} + \sum_{k=3}(k+1)\,\phi^k\mathbf{1}'\boldsymbol{C}^k\mathbf{1}$$

Let $S = \sum_{k=3}(k+1)\,\phi^k\mathbf{1}'\boldsymbol{C}^k\,\mathbf{1} = n\sum_{k=3}(k+1)\,\phi^k(n-1)^k$, and thus $\phi(n-1)\,S = n\sum_{k=3}(k+1)\,\phi^{k+1}(n-1)^{k+1}$. Then we have,

$$\begin{aligned} S - \phi(n-1)\,S &= 3n\phi^3(n-1)^3 + n\sum_{k=3}\phi^k(n-1)^k \\ &= n\phi^3(n-1)^3\left(3 + \frac{1}{1-\phi(n-1)}\right). \end{aligned}$$

Finally, we have $S = \sum_{k=3}(k+1)\,\phi^k\mathbf{1}'\boldsymbol{C}^k\mathbf{1} = \frac{n\phi^3(n-1)^3(4-3\phi(n-1))}{(1-\phi(n-1))^2}$.

Therefore,

$$\begin{aligned}\mathbf{1}'(\boldsymbol{I}-\phi\boldsymbol{G})^{-1}(\boldsymbol{I}-\phi\boldsymbol{G})^{-1}\mathbf{1} &< \mathbf{1}'\boldsymbol{I}\mathbf{1}+2\phi\mathbf{1}'\boldsymbol{G}\mathbf{1}+3\phi^2\mathbf{1}'\boldsymbol{G}\boldsymbol{G}\mathbf{1}+\frac{n\phi^3(n-1)^3(4-3\phi(n-1))}{(1-\phi(n-1))^2}\\ &< \mathbf{1}'\boldsymbol{I}\mathbf{1}+2\phi\mathbf{1}'\boldsymbol{G}\mathbf{1}+3\phi^2\mathbf{1}'\boldsymbol{G}\boldsymbol{G}\mathbf{1}+\frac{4n\phi^2(n-1)^2}{(1-\phi(n-1))^2}.\end{aligned}$$

The last inequality comes from the fact that $0<\phi\le\frac{1}{n-1}$.

From $\frac{4n\phi_1^2(n-1)^2}{(1-\phi_1(n-1))^2}\le\varepsilon_1$, we can get $\phi_1\le\frac{\sqrt{\frac{\varepsilon_1}{4n}}}{1+\sqrt{\frac{\varepsilon_1}{4n}}}$. Then, when $t\in[4,\frac{n^2-3n}{4})$, for any $\phi<\phi_1$ and network $\boldsymbol{G}$ with link t,

$$\begin{aligned}\mathbf{1}'(\boldsymbol{I}-\phi\boldsymbol{QS}(t))^{-1}(\boldsymbol{I}-\phi\boldsymbol{QS}(t))^{-1}\mathbf{1} &\ge \mathbf{1}'\boldsymbol{I}\mathbf{1}+2\phi\mathbf{1}'\boldsymbol{QS}(t)\mathbf{1}+3\phi^2\mathbf{1}'\boldsymbol{G}\boldsymbol{G}\mathbf{1}+\varepsilon_1\\ &> \mathbf{1}'(\boldsymbol{I}-\phi\boldsymbol{G})^{-1}(\boldsymbol{I}-\phi\boldsymbol{G})^{-1}\mathbf{1}.\end{aligned}$$

Similarly, let $\phi_2\le\frac{\sqrt{\frac{\varepsilon_2}{4n}}}{1+\sqrt{\frac{\varepsilon_2}{4n}}}$, we have $1'(\boldsymbol{I}-\phi\boldsymbol{QC}(t))^{-1}(\boldsymbol{I}-\phi\boldsymbol{QC}(t))^{-1}1\ge 1'(\boldsymbol{I}-\delta\boldsymbol{G})^{-1}(\boldsymbol{I}-\delta\boldsymbol{G})^{-1}1$ for all network $\boldsymbol{G}$ with link t when $t\in(\frac{n^2+n}{4},\frac{n^2-n}{2}]$. $\square$

Proof of Lemma 5: The proof is similar with that of Lemma 1. The agents heterogeneity just amplifies the inequalities in the proof. $\square$

Proof of Proposition 3: Any strategy $p=(p_t)_{t=1}^T$ of Problem (7) induces a strategy of Problem (4) $s=(\boldsymbol{G}(t))_{t=1}^T$. Therefore, the "if" part directly holds.

Meanwhile, for any optimal strategy $s^*=(\boldsymbol{G}^*(t))_{t=1}^T$ of Problem (4), we can construct a strategy $p=(p_t)_{t=1}^T$ of Problem (7) inducing a same sequence of networks $(\boldsymbol{G}^*(t))_{t=1}^T$. Specifically, let $p_t=i$, where $\boldsymbol{G}^*(t+1)=\boldsymbol{G}^*(t)+\boldsymbol{E}_{ij}$ and $N_i(\boldsymbol{G}(t))\subseteq N_j(\boldsymbol{G}(t))$. That is, if $\boldsymbol{G}^*(t+1)$ is obtained by connecting agents i,j and i's degree is no larger than j, then $p_t=i$. We argue that, agent i will connects agent j to maximize his own utility. According to König et al. (2014)'s Proposition 1 (ii), agent i will connects the agent with highest (possible) degree. It is agent j since otherwise $\boldsymbol{G}^*(t)+\boldsymbol{E}_{ij}$ can not be an NSG. The "only if" part has been verified. $\square$

References

Allouch, N. (2017). On the private provision of public goods on networks. *Journal of Economic Theory 157*, 527–552.

Atack, J., F. Bateman, M. Haines, and R. A. Margo (2010). Did railroads induce or follow economic growth? urbanization and population growth in the american midwest, 1850-1860. *Social Science History 34*(2), 171–197.

Baetz, O. (2015). Social activity and network formation. *Theoretical Economics 10*, 315–340.

Bala, V. and S. Goyal (2000). A noncooperative model of network formation. *Econometrica 68*(5), 1181–1229.

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

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013). The diffusion of microfinance. *Science 341*(6144).

Banerjee, A., E. Duflo, and N. Qian (2020). On the road: Access to transportation infrastructure and economic growth in china. *Journal of Development Economics 145*, 102442.

Belhaj, M., S. Bervoets, and F. Deroïan (2016). Efficient networks in games with local complementarities. *Theoretical Economics 11*(1), 357–380.

Bernardo M. Ábrego, Silvia Fernández-Merchant, M. G. N. W. W. (2009). Sum of squares of degrees in a graph. *Journal of Inequalities in Pure and Applied Mathematics 10*(3).

Billand, P., C. Bravard, J. Durieu, and S. Sarangi (2015). Efficient networks for a class of games with global spillovers. *Journal of Mathematical Economics 61*, 203–210.

Bloch, F. and N. Quérou (2013). Pricing in social networks. *Games and Economic Behavior 80*, 243 – 261.

Bramoullé, Y., A. Galeotti, and B. Rogers (2016). *The Oxford handbook of the economics of networks.* Oxford University Press.

Bramoullé, Y. and R. Kranton (2007). Public goods in networks. *Journal of Economic Theory 135*(1), 478 – 494.

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009). Peer effects and social networks in education. *The Review of Economic Studies 76*(4), 1239–1267.

David, G. and M. Dina (2004). Using online conversations to study word-of-mouth communication. *Marketing Science 23*(4), 545–560.

Duranton, G., P. M. Morrow, and M. A. Turner (2014). Roads and Trade: Evidence from the US. *The Review of Economic Studies 81*(2), 681–724.

Dutta, B., S. Ghosal, and D. Ray (2005). Farsighted network formation. *Journal of Economic Theory 122*(2), 143–164.

Dutta, B. and S. Mutuswami (1997). Stable networks. *Journal of Economic Theory 76*(2), 322–344.

Elliott, M. L., S. Goyal, and A. Teytelboym (2019). Networks and economic policy. *Oxford Review of Economic Policy 35*(4), 565–585.

Galeotti, A. and S. Goyal (2010). The law of the few. *American Economic Review 100*(4), 1468–1492.

Helsley, R. W. and Y. Zenou (2014). Social networks and interactions in cities. *Journal of Economic Theory 150*, 426 – 466.

Hiller, T. (2017). Peer effects in endogenous networks. *Games and Economic Behavior 105*, 349 – 367.

Jackson, M. O., B. W. Rogers, and Y. Zenou (2017, March). The economic consequences of social-network structure. *Journal of Economic Literature 55*(1), 49–95.

Jackson, M. O. and A. Watts (2002). On the formation of interaction networks in social coordination games. *Games and Economic Behavior 41*(2), 265–291.

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

König, M. D., C. J. Tessone, and Y. Zenou (2014). Nestedness in networks: A theoretical model and some applications. *Theoretical Economics 9*(3), 695–752.

König, M. D., X. Liu, and Y. Zenou (2019). R&D Networks: Theory, Empirics, and Policy Implications. *The Review of Economics and Statistics 101*(3), 476–491.

Li, X. (2020). Optimal complementarity networks. *Working paper*.

Song, Y. and M. van der Schaar (2020, June). Dynamic network formation with foresighted agents. *International Journal of Game Theory 49*(2), 345–384.

van Leeuwen, B., T. Offerman, and A. Schram (2019, 02). Competition for Status Creates Superstars: an Experiment on Public Good Provision and Network Formation. *Journal of the European Economic Association 18*(2), 666–707.

Watts, A. (2001). A dynamic model of network formation. *Games and Economic Behavior 34*(2), 331–341.