The benefit of sequentiality in social networks*

Junjie Zhou† Ying-Ju Chen‡

February 16, 2013

Abstract

This paper examines the benefit of sequentiality in the social networks. We adopt the elegant theoretical framework proposed by Ballester et al. (2006) wherein a fixed set of players non-cooperatively determine their contributions. This setting features payoff externalities and strategic complementarity amongst players. We first analyze the two-stage game in which players in the leader group make contributions prior to the follower group. Compared with the simultaneous-move benchmark, the equilibrium contribution by any individual player in any two-stage sequential-move game is unambiguously higher. We establish the isomorphism between the socially optimal selection of the leader and follower groups and the classical weighted maximum-cut problem. We give an exact index to characterize the key leader problem, and show that the key leader can be substantially different from the key player who impacts the networks most in the simultaneous-move game. We also provide some design principles for unweighted complete graphs and bipartite graphs.

We then examine the structure of optimal mechanism and allow for arbitrary sequence of players' moves. We show that starting from any fixed sequence, the aggregate contribution always goes up while making simultaneous-moving players move sequentially. This suggests a robust rule of thumbs – any local modification towards the sequential-move game is beneficial. Pushing this idea to the extreme, the optimal sequence turns out to be a chain structure, i.e., players should move one by one. Our results continue to hold when either players exhibit strategic substitutes instead or the network designer's goal is to maximize the players' aggregate payoff rather than the aggregate contribution.

Keywords: social network, dynamic games, key leader, sequencing, game theory

JEL classifications: D21, D29, D82

---

*We have benefited from the discussions with Santiago Balseiro, Kostas Bimpikis, Ozan Candogan, Yongmin Chen, Bar Ifrach, Ananth Iyer, Ilan Lobel, Joan de Mart Beltran, Cheng-Zhong Qin, Adam Szeidl, Guofu Tan, Yves Zenou, Yang Zhang, and seminar participants at UC Berkeley, SUFE, 2012 China International Conference on Game Theory and Applications–Qingdao, the 9th Workshop on Industrial Organization and Management Strategy– SJTU, and INFORMS. All remaining errors are our own.

†School of International Business Administration, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai, 200433, China; e-mail: zhoujj03001@gmail.com.

‡University of California at Berkeley, 4121 Etcheverry Hall, Berkeley, CA 94720; e-mail: chen@ieor.berkeley.edu.

1 Introduction

Many modern organizations are highly sophisticated. Within an organization, members (individual workers, departments, or divisions) are assigned distinct roles, endowed with heterogeneous abilities, and have the discretion of their own decision making. Since the ultimate organizational profitability and sustainability depend on their collective efforts, there is interdependence amongst these members' individual payoffs; subsequently, it gives rise to strong incentives for them to interact with each other. Coordinating amongst multiple divisions within the organization requires extensive communication, and this is regarded as one of the main concerns in managing modern organizations (Calvó-Armengol and Beltran (2009)). See also Milgrom and Roberts (1992) for more elaboration on the importance of coordination for the organizations. Since members' interactions take place in very refined levels (a member may intensively exchange ideas with close colleagues but may never speak to others), the structure of communication is best modeled as a nexus of network. In this networked structure, members (players) are represented as distinct nodes, and the possibilities of communication or interactions between a pair of members are represented by a link that connects them. This allows us to explicitly incorporate the local effects. In addition, the heterogeneity (regarding members' abilities) can be easily described as the nodes' characteristics (see Calvó-Armengol and Beltran (2009) for details).

This paper investigates how an organization designer can promote the communication amongst their members (*players* hereafter) by choosing the *sequence* of players that make their decisions accordingly. One could interpret this sequential-move feature as a hierarchy within the organization; thus, players that are assigned to move earlier are promoted to higher ranks and their decisions become prominent and readily observable by others. We show in this parsimonious setup that *sequentiality alone* can be substantially beneficial for the aggregate benefit of the organization. Accordingly, we provide some simple, yet non-trivial, *design principles* for the "configuration" of optimal sequence. To the best of our knowledge, no prior work has ever explored this particular angle. Furthermore, at a higher level, our model setup may be interpreted as different sorts of social networks (such as crime organizations, political connections, and labor markets). Thus, our theoretical analysis may also provide justifications or suggestions to other social networks beyond the organization structure applications.

To achieve this goal, we adopt the elegant framework proposed by Ballester et al. (2006) with a social network and a set of players, which originally was introduced to describe crime organizations but then was found to have broad applications in network economics. In this setting, players are heterogeneous in terms of their intrinsic valuations (which can be alternatively interpreted as their abilities), and each player may be connected to only a subset of other players as described by the network structure. Each player's contribution generates some *positive externality* to others. The peer effects have widely been documented in various empirical and experimental studies (see, e.g., Bandiera et al. (2004), Falk and Ichino (2006), and the book by Jackson (2008) for an extensive survey). Moreover, there is *strategic complementarity* among the players (i.e., a more aggressive decision by a player reinforces other players' decisions). The model setup incorporates the *local* network effects and includes the micro-level descriptions of underlying networks. We use the adjacency matrix to describe the physical network structure embedded among the players.

We first analyze a two-stage game in which a group of players (*leaders*) move first, and their contributions are observed by others; the remaining players (*followers*) then determine their contributions simultaneously. Building upon the equilibrium characterizations, we examine the optimal selection of the leader and follower groups, with the objective being the aggregate contribution amongst players. This allows us to examine how the organization designer should promote members given the organization's network structure and their heterogeneous intrinsic valuations (or abilities). We show that the problem is mathematically equivalent to the classical weighted *maximum-cut* (MAX-CUT) problem, i.e., to find a set partition such that the number of cross links, weighted by the players' intrinsic valuations, between the two groups is maximized. While this problem is in general NP-hard, numerous heuristics and approximation algorithms have been proposed in the literature of computer science and operations research. Furthermore, when players possess homogeneous intrinsic valuations, there are some polynomial-time solvable cases such as the planar graphs (Hadlock (1975)), graphs without K_5 minors (Barahona (1983)), and weakly bipartite graphs with non-negative weights (Grötschel and Pulleyblank (1981)).

A special case of the above group selection problem may be of particular interest: if we are allowed to pick up only one player to move first, who should be the person of interest? This *key leader* problem parallels the key player problem in Ballester et al. (2006), who argue that the outsider may be able to remove one player from the criminal network and therefore intends to identify the key player that impacts the network most. We give an exact index to characterize this key leader problem without using any approximation. We show that the key leader can be substantially different from the key player identified in Ballester et al. (2006), and it need not be the player with the highest intrinsic valuation. We then proceed to investigate the two-stage network design problem for some specific network structures. In the case of unweighted complete graphs, we establish an intuitive *pecking order* for any given size of leader group: it is always beneficial to assign players with higher intrinsic valuations to the leader group. We then provide a concrete method to determine the optimal size of leader group. In particular, with homogeneous intrinsic valuations (and hence homogeneous players), it is optimal to split them into two halves. This provides a theoretical ground for a simple rule of thumbs – the 50-50 rule. In complete bipartite graphs, it is always optimal to split the players based on the two groups naturally defined by the bipartite structure. Furthermore, when the two groups have the same size, we shall nominate the group with a higher *average* intrinsic valuation as the leaders. On the other hand, when the average intrinsic valuations are the same between two groups, the one with a *smaller* size should move first; thus, early adopters should be relatively rare. As an example, in a star (hub-spokes) network, the hub, a natural influencer or trend-setter, should make the decision before all the spokes. We also provide the general formula for other mixed scenarios.

Next, we relax the two-stage restriction and examine the structure of optimal mechanism. We show that starting from any fixed sequence, the aggregate contribution always goes up while inserting any sequentiality, i.e., making any set of simultaneous-moving players choose their contributions sequentially. Note that this result does not hinge on any structural assumptions of the underlying network. This result suggests that any "local" modification towards the sequential-move game is beneficial. It also serves as a robust design principle if we are bound by some physical restrictions for freely choosing any sequence. Pushing this idea further, the optimal sequence turns out to be a *chain* structure, i.e., players should move one by one, as it maximally capitalizes the positive feedback effects. With homogeneous intrinsic valuations, we further show that any chain structure yields the same maximum aggregate contribution, irrespective of the configuration of the chain.

Finally, with strategic substitutes, most of our results continue to hold in this alternative setting. If instead the network designer aims at maximizing the players' aggregate payoff, the chain structure remains optimal with strategic complements, but this result no longer holds with strategic substitutes. The group selection criterion differs from the maximum-cut problem, but some design principles continue to apply (such as the pecking order based on their intrinsic valuations and the $50-50$ rule). These design principles hopefully help the network/organization management problem in various scenarios.

The remainder of this paper is organized as follows. In Section 2, we review some relevant literature. Section 3 introduces the model setup and the benchmark simultaneous-move game. Section 4 studies the two-stage games wherein some players decide their contribution levels before others. Section 5 discusses the key leader problem, and Section 6 provides further examples of network structures for illustration. Section 7 examines general sequences of events and establishes the optimality of chain structure. Section 8 extends our analysis to some alternative scenarios. We draw some concluding remarks in Section 9. All the technical proofs are relegated to the appendix.

2 Literature review

Our paper is related to the vast literature on network externality. The central premise of this research stream is that the utility generated from possessing a product gets higher as more other players/consumers use it. The classical papers take the macro-economic perspective and primarily focus on the aggregate level of network externality; see, e.g., Rohlfs (1974) for the self-enhancing and self-fulfilling characteristics of telephone and fax machine industries, Katz and Shapiro (1985) for the complementary goods, Farrell and Saloner (1986) on technology adoption, and Economides (1996) for an extensive survey of this literature. In contrast, we acknowledge the local network effects and explicitly model the physical network structure amongst players. This is in line with the network economics literature whereby researchers are motivated by ample empirical evidence and start introducing the local network effects into their theoretical constructs. Applications can be found in labor markets, developing countries, risk sharing, diffusion and social structure, and social learning; see Jackson (2008) for a comprehensive survey.

As aforementioned, our paper is closely related to the influential paper by Ballester et al. (2006), who study the simultaneous-move network game. They show that the Nash equilibrium extensively uses the measure "weighted Katz-Bonacich Centrality" and therefore establishes the connection between the network economics literature and the sociology literature. We extend their analysis by allowing for sequential moves. We characterize the optimal group selection problem in the two-stage game and show that the optimal hierarchy turns out to be a chain structure. See also Ballester and Calvó-Armengol (2010), Bramoullé and Kranton (2007), and Corbo et al. (2006) for further discussions. A recent contribution by Candogan et al. (2012) incorporates the pricing decisions into the framework of Ballester et al. (2006). They characterize the optimal price discrimination as a function of the underlying social interactions. They also investigate the alternative scenarios with uniform or two-price schemes when the seller's price discrimination power is limited. All the above papers consider simultaneous-move games amongst a network of the players, whereas we introduce the sequential-move feature into this network game. There are some papers that examine the learning aspect in dynamic networks (e.g., Acemoglu et al. (2011) and Acemoglu et al. (2013)); however, the players' strategies and the economic forces in these papers are fundamentally different from ours.

Our study is closely related to the literature on strategic complementarities. Stemming from the classical contribution by Bulow et al. (1985), it has been observed that strategic complementarities can arise in various forms. To the best of our knowledge, no prior work incorporates both the local network effects and examines the strategic consequences of sequential-move games. Our work also adds to the vast literature on the timing of decision making among players. This includes the classical comparison between Cournot competition and Stackelberg leader-follower game, and the discussions on the first-mover and second-mover advantages; see, e.g., Amir and Stepanova (2006), Hoppe (2000), Hoppe and Lehmann-Grube (2005), and Kerin et al. (1992). Our study shows that in the presence of strategic complementarity, players in a network game can benefit from moving sequentially; furthermore, it leads to a Pareto improvement amongst all players irrespective of the underlying network structure.

The organization structure design problem has been a central topic in economics. Stemming from the seminal work by Radner (1962), this research stream examines how a team of players should be organized while facing costly communication and information processing. This includes Alonso et al. (2008), Aoki (1986), Crémer (1980), and Marschak and Radner (1972). We abstract away the communication costs and therefore complement this research stream by studying the pure effect of strategic complementarity. The network structure is explicitly described in Calvó-Armengol and Beltran (2009), but they emphasize the information gathering aspect and consider a simultaneous-move game. In a two-player setup, Huck and Biel (2012) show that sequential-moving may be beneficial if conformity (i.e., behaving similarly) is inherently important for the players. The conformity also generates strategic complementarity between the two players. Thus, their paper is in spirit similar to our two-node example, and we proceed to characterize the optimal design of sequential-move game for the multiple-player network game.

3 The model

We adopt the elegant framework proposed by Ballester et al. (2006) with a social network and a set of players $\mathcal{N} = \{1, 2, \cdots, N\}$. Each player, indexed by i, is represented as a *node* of the network and is entitled to determine the level of contribution x_i. Since we aim at deriving general principles to various kinds of social networks, we will keep the model descriptions generic.

Payoff structure. We use the following payoff structure to capture these two features:

$$\pi_i(x_1, x_2, \cdots, x_n) = \alpha_i x_i - \frac{1}{2} x_i^2 + \delta \sum_{j=1}^{N} g_{ij} x_i x_j, \tag{1}$$

where $\{x_j\}$'s correspond to the contributions by these players. In (1), $\alpha_i > 0$ measures the *intrinsic marginal utility* for player i; as aforementioned, it could also be interpreted as the player's ability in other contexts. The first two terms collectively suggest the diminishing marginal return of the player's own contribution. The last term captures the network effect among the players. Parameter $\delta > 0$ controls the strength of this effect, and it is common across all the players.

The cross term $g_{ij}x_ix_j$ indicates the interaction between the pair of players i, j, and we assume that $g_{ij} \geq 0$ to capture the *strategic complementarity.* The matrix $G = (g_{ij})$ summarizes the cross effects between players. If two players are frequently involved in the same community or group, their cross effect is strong (g_{ij} is large). For some examples, G is the adjacency matrix of a directed graph. Naturally, the cross effect appears only amongst different players' contributions; thus, $g_{ii} = 0$, i.e., there is no self-loop. We do not require G to be symmetric, and some of our analysis goes through even if some components of G are negative. Nevertheless, in the majority of this paper we will refrain from making implications of the negative components, as our primary goal is to study the strategic complementarity in social networks. In Section 8, we discuss why some of our results continue to hold when strategic substitution occurs.

Network game and notation. The above descriptions give rise to a concrete environment of social network regarding the players' actions and payoffs. To specify a network game, we shall introduce the timing and information structure. We will examine various scenarios to illustrate the impacts of sequentiality. Before we proceed, we introduce some notation that will be intensively used throughout the paper.

For a matrix T, the transpose is denoted as T'. The zero matrix (of suitable dimensions) is denoted as $\mathbf{0}$. If T is a square matrix, then T^D is a matrix with diagonal entries $T^D_{ii} = t_{ii}, i = 1, \cdots N$, and off diagonal entries $T^D_{ij} = 0$, $\forall i \neq j$. Unless indicated otherwise, vector $\mathbf{x} = (x_1, \cdots, x_N)'$ is a column vector. For any subset A of $\mathcal{N}$, $\mathbf{x}_A$ (in bold) denotes the vector of $(x_i)_{i\in A}$; that is, it is a sub-vector wherein the sequence of selected components follows their original sequence in vector $\mathbf{x}$. The (non-bold) term $x_A = \sum_{i\in A} x_i$ is the sum of these selected components. Let $\langle \mathbf{x}, \mathbf{y}\rangle$ denote the inner product of two column vectors $\mathbf{x}, \mathbf{y}$.

We say that two matrices A, B satisfy $A \succeq B$ if and only if $A_{ij} \geq B_{ij}, \forall i, j$. In other words, this dominance relationship applies to the component-wise comparisons. For any pair of functions f_1 and f_2, we call $f_1(\delta) = \mathcal{O}(f_2(\delta))$ as $\delta \to 0$, if $\limsup_{\delta\to 0} |\frac{f_1(\delta)}{f_2(\delta)}| < \infty$, and $f_1(\delta) = o(f_2(\delta))$, as $\delta \to 0$, if $\lim_{\delta\to 0} |\frac{f_1(\delta)}{f_2(\delta)}| = 0$. In this paper, the function f_2 is a power function of δ (i.e., δ^k for an integer $k = 1, 2, \cdots$).

Simultaneous-move game. Now we introduce the benchmark scenario wherein all players determine their contribution levels simultaneously. This game has been studied by Ballester et al. (2006), and we include it for completeness.

First, we note that the game is supermodular and the payoffs are quadratic and concave in $\{x_i\}$'s. Therefore, the best response function for a player i is linear and increasing in other players' contributions:

$$BR_i(\mathbf{x}_{-i}) = \alpha_i + \delta \sum_{j\neq i} g_{ij}x_j.$$

Therefore, the *Nash equilibrium* in this simultaneous-move game is just the solution to:

$$x^N_i = BR_i(\mathbf{x}^N_{-i}) = \alpha_i + \delta \sum_{j\neq i} g_{ij}x^N_j.$$

Rewriting the above (using matrix notation), we obtain:

$$\mathbf{x}^N = \alpha + \delta G \cdot \mathbf{x}^N \Leftrightarrow \mathbf{x}^N = [\mathbf{I} - \delta G]^{-1}\alpha. \tag{2}$$

where $\alpha = (\alpha_1, \cdots, \alpha_n)'$.

Taylor expansions. The matrix is well defined and invertible if δ is small enough.[1] In this case, we have the following Taylor series:

$$[\mathbf{I} - \delta G]^{-1}\alpha = (\mathbf{I} + \delta G + \delta^2 G^2 + \cdots)\alpha = \alpha + \delta G\alpha + \delta^2 G^2 \alpha + \cdots .$$

For sufficiently small δ, the matrix $\mathbf{M} := [\mathbf{I} - \delta G]^{-1}$ is well defined and nonnegative, with

$$m_{ij} = \sum_{k=0}^{+\infty} \delta^k g_{ij}^{[k]} = ((\mathbf{I} - \delta G)^{-1})_{ij} = \delta_{ij} + \delta g_{ij} + \delta^2 g_{ij}^{[2]} + \cdots ,$$

where $g_{ij}^{[k]}$ is the ij entry of G^k. The term m_{ij} counts the number of paths in G that start at node i and end at node j, and paths of length k are weighted by δ^k. This matrix measures the impact of player i's contribution on player j's contribution through direct and indirect influences. The direct influence arises from the payoff externality between players i and j, and the indirect influences follow from the impact of player i's contribution upon other players that ultimately connect to player j.

In compliance with the network economics literature, we shall implicitly assume that δ is sufficiently small such that the equilibrium is well-defined; otherwise, some players intend to make infinite contributions due to very strong positive feedback effects. We will not be explicit in describing the exact bound for δ in the majority of our analysis, since this does not help understanding the economic intuition of the problem. This applies to both the simultaneous-move game and sequential-move games (to be formally defined momentarily). In some worked examples, however, we will briefly indicate the bounds for completeness.

Equilibrium outcomes. The best response functions immediately lead to the unique Nash equilibrium:

$$x_i^N = b_i(G, \delta, \alpha),$$

where the vector

$$\mathbf{b}(G, \delta,\ \alpha) = [\mathbf{I} - \delta G]^{-1}\alpha$$

is called the *weighted Katz-Bonacich Centrality* of parameter δ and weight vector α. This measure stands out among various ones proposed by sociology researchers, because it naturally ties in the Nash equilibrium in the simultaneous-move games (Ballester et al. (2006)).

In the analysis below, we need a variant of this equilibrium outcome when the diagonal entries of matrix G are nonzero. This is given in the next lemma.

Lemma 1. *For sufficiently small δ, the unique Nash equilibrium outcome* $\mathbf{x}$ *of the simultaneous-move game with payoff functions*

$$u_i(\mathbf{x}) = \beta_i x_i - \frac{1}{2}x_i^2 + \delta x_i (\sum_{j=1}^{N} t_{ij} x_j).$$

[1] The exact upper bound is $1/\mu_1(G)$ if G is symmetric, where $\mu_1(G)$ is the largest eigenvalue of G.

is given by

$$\mathbf{x} = \left[1 - \delta(T + T^D)\right]^{-1} \beta, \tag{3}$$

where $T = (t_{ij})_{N\times N}$*, and* T^D *is a matrix with diagonal entries* $T^D_{ii} = t_{ii}, i = 1, \cdots N$*, and off diagonal entries* $T^D_{ij} = 0$*,* $\forall i \neq j$*. Moreover,* $\left[1 - \delta(T + T^D)\right]^{-1}$ *is symmetric if* T *is symmetric.*

Having discussed the simultaneous-move game, we next consider the two-stage sequential-move games and compare them with this benchmark. We will start with the general two-stage problems, where a single leader case is just a special case. We will then build upon the results and observations in the two-stage games to characterize the optimal mechanism with a general hierarchy.

4 Two-stage games

In this section, we analyze the two-stage case.

Leaders versus followers. In any two-stage game, a group of players move first, and their contributions are observed by others; following this, the remaining players then determine their contributions simultaneously. Thus, we can partition the players into two groups: the *leader* group A and the *follower* group B. In this simplest sequential-move game, the leaders correspond to the influencers, opinion leaders, or trend-setters.

The equilibrium concept is subgame perfect Nash equilibrium (SPNE) because now the game involves multiple rounds of interactions (Fudenberg and Tirole (1991)). For convenience, let us rewrite the matrix G as a block matrix:

$$G = \begin{pmatrix} G_{AA} & G_{AB} \\ G_{BA} & G_{BB} \end{pmatrix}.$$

Second stage. The equilibrium can be solved using backward induction. Let vector $\mathbf{x}_A$ denote the contributions chosen by the nodes in A. In the second stage, since the contributions of the leader group have been determined, the subgame is again a network game with nodes B and adjacency matrix G_{BB}, and parameters $\alpha'_B = \alpha_B + \delta G_{BA}\mathbf{x}_A$. By Lemma 1, the equilibrium in the subgame after observing $\mathbf{x}_A$ is given by

$$\mathbf{x}_B(\mathbf{x}_A) = U\alpha'_B = U(\alpha_B + \delta G_{BA}\mathbf{x}_A), \text{ where } U = [I - \delta G_{BB}]^{-1}. \tag{4}$$

Moreover, the matrix U is symmetric if G_{BB} is, and $G^D_{BB} = \mathbf{0}$ because $g_{ii} = 0$, $\forall i$. The inclusion of $\mathbf{x}_A$ shows that the followers' contributions are crafted by the leaders' prominent decisions.

First stage. Let us now go backwards to the first stage. Anticipating the followers' actions, the nodes in A will play a game with payoff functions given by:

$$u_i = \alpha_i x_i - \frac{1}{2}x_i^2 + \delta x_i \left(\sum_{j\in A} g_{ij}x_j + \sum_{j\in B} g_{ij}x_j(\mathbf{x}_A)\right), \quad \forall i \in A,$$

where $x_j(\mathbf{x}_A) = \sum_{k\in B} U_{jk}(\alpha_k + \delta \sum_{l\in A} g_{kl}x_l)$ by (4). Plugging in $x_j(\mathbf{x}_A)$, we obtain that

$$u_i = \left(\alpha_i + \delta \sum_{j\in B}\sum_{k\in B} g_{ij}U_{jk}\alpha_k\right) x_i - \frac{1}{2}x_i^2 + \delta x_i \left(\sum_{j\in A} g_{ij}x_j + \delta \sum_{j\in B}\sum_{k\in B}\sum_{l\in A} g_{ij}U_{jk}g_{kl}x_l\right).$$

In other words, it is a (modified) network game on the set A with parameters

$$\beta = \alpha_A + \delta G_{AB}\, U\, \alpha_B, \text{ and } T = G_{AA} + \delta G_{AB}\, U\, G_{BA}.$$

Equilibrium characterization. We can then apply Lemma 1 to the above modified game. We summarize the characterization below.

Theorem 1. *For sufficiently small δ,*[2] *the unique subgame perfect Nash equilibrium of the two-stage game is given by:*

$$\begin{pmatrix} \mathbf{x}_A \\ \mathbf{x}_B \end{pmatrix} = S \begin{pmatrix} \alpha_A \\ \alpha_B \end{pmatrix} \tag{5}$$

with

$$S = \begin{pmatrix} \left[1-\delta(T+T^D)\right]^{-1} & \delta\left[1-\delta(T+T^D)\right]^{-1} G_{AB}U \\ \delta U G_{BA}\left[1-\delta(T+T^D)\right]^{-1} & U + \delta^2 U G_{BA}\left[1-\delta(T+T^D)\right]^{-1} G_{AB}U \end{pmatrix}, \tag{6}$$

where

$$T = G_{AA} + \delta G_{AB}\, U\, G_{BA}, \text{ and } U = [I - \delta G_{BB}]^{-1}.$$

Moreover, if G is symmetric, then S is symmetric as well.

Theorem 1 leads to some simple interpretations of the equilibrium outcomes. First, the equilibrium contributions are linearly increasing in each component of α, where the sensitivities are given by the matrix S. This suggests that an increase of each player's intrinsic valuation has positive and straightforward impacts on all the players' contributions, and the magnitudes are succinctly summarized by the matrix S. Second, while followers (in set B) play their best responses to other players' contributions, leaders anticipate the followers' subsequent reactions and incorporate these effects into their decision making. This is a signature feature of Stackelberg game, or more generally any sequential-move game.

Third, it is verifiable that $S \succeq \mathbf{M}$, where the matrix $\mathbf{M}$ captures the sensitivities of intrinsic valuations in a simultaneous-move game. To see this, it suffices to compare these two matrix component-wise. Using Block matrix inversion formula, we can write $\mathbf{M}$ as follows:

$$\mathbf{M} = [1-\delta G]^{-1} = \begin{pmatrix} [1-\delta(T+0)]^{-1} & \delta[1-\delta(T+0)]^{-1} G_{AB}U \\ \delta U G_{BA}[1-\delta(T+0)]^{-1} & U + \delta^2 U G_{BA}[1-\delta(T+0)]^{-1} G_{AB}U \end{pmatrix}, \tag{7}$$

[2] For example, the result holds when $\delta < \frac{1}{2\mu_1(G)}$ if G is symmetric. The exact upper bound of the parameter δ depends on both the network G and the leader group A, and its expression is complicated. Here we just give an upper bound which does not depend on A.

where the matrices U and T are defined in Theorem 1. In other words, If we replace the matrix T^D by the zero matrix in (6), we will get the matrix $[1-\delta G]^{-1}$.

Comparing the entries of (6) and (7), we observe the following. To show that $S \succeq \mathbf{M} = [1-\delta G]^{-1}$, it suffices to show that

$$\left[1-\delta(T+T^D)\right]^{-1} \succeq [1-\delta T]^{-1}$$

This follows immediately from

$$\left[1-\delta(T+T^D)\right]^{-1} - [1-\delta T]^{-1} = [1-\delta T]^{-1}\,\delta T^D \left[1-\delta(T+T^D)\right]^{-1} \succeq \mathbf{0},$$

for sufficiently small δ. Here we use the fact $A^{-1} - B^{-1} = B^{-1}(B-A)A^{-1}$. This component-wise dominance therefore implies that the equilibrium contributions in this two-stage sequential-move game are higher. This is formally stated in the following proposition.

Proposition 1. *For sufficiently small δ, the equilibrium contribution profile in any two-stage game is component-wise higher than that in the simultaneous-move game.*

Proposition 1 suggests that the sequential-move game effectively utilizes the *positive feedbacks* and partially restores the social efficiency. In particular, each player contributes more in the two-stage game, irrespective of the underlying network structure. This is reminiscent of the findings in Beckmann (2005), who shows in the tax competition context that sequentiality is beneficial in the presence of strategic complementarities.

Before we proceed, we note that the results are extendable to the (modified) network game wherein the diagonal entries of G are not all zeros. We summarize our findings as the following corollary, and we will use this fact in the proof of the Proposition 6 later on.

Corollary 1. *If the diagonal entries of G are not all zeros, the results in Theorem 1 still hold, except that the matrix U is replaced by $\tilde{U} = [1-\delta(G_{BB}+G_{BB}^D)]^{-1}$. In other words, we have:*

$$\tilde{S} = \begin{pmatrix} \left[1-\delta(\tilde{T}+\tilde{T}^D)\right]^{-1} & \delta\left[1-\delta(\tilde{T}+\tilde{T}^D)\right]^{-1} G_{AB}\tilde{U} \\ \delta\tilde{U}G_{BA}\left[1-\delta(\tilde{T}+\tilde{T}^D)\right]^{-1} & \tilde{U}+\delta^2\tilde{U}G_{BA}\left[1-\delta(\tilde{T}+\tilde{T}^D)\right]^{-1} G_{AB}\tilde{U} \end{pmatrix},$$

where

$$\tilde{T} = G_{AA} + \delta G_{AB}\tilde{U}G_{BA}, \text{ and } \tilde{U} = [1-\delta(G_{BB}+G_{BB}^D)]^{-1}.$$

Moreover, the statement of Proposition 1 remains valid.

Contribution differential. To get more quantitative results about the difference between the two-stage game and the simultaneous-move game, we express the differences of contributions by comparing the entries of (6) and (7), we have:

$$\begin{pmatrix} \mathbf{x}_A \\ \mathbf{x}_B \end{pmatrix} - \begin{pmatrix} \mathbf{x}_A^N \\ \mathbf{x}_B^N \end{pmatrix} = \begin{pmatrix} \Delta & \delta\Delta G_{AB}U \\ \delta U G_{BA}\Delta & \delta^2 U G_{BA}\Delta G_{AB}U \end{pmatrix} \begin{pmatrix} \alpha_A \\ \alpha_B \end{pmatrix} \tag{8}$$

where

$$\Delta = \left[1 - \delta(T + T^D)\right]^{-1} - \left[1 - \delta T\right]^{-1}.$$

Moreover, for small δ, we can get a more concise expression using Taylor expansion.

Theorem 2. *For sufficiently small δ, we can characterize the increment of equilibrium contribution profile between the two-stage and simultaneous-move games using Taylor expansion up to the order of δ^3 :*

$$\begin{pmatrix} \mathbf{x}_A \\ \mathbf{x}_B \end{pmatrix} - \begin{pmatrix} \mathbf{x}_A^N \\ \mathbf{x}_B^N \end{pmatrix} = \delta^2 \begin{pmatrix} (G_{AB}G_{BA})^D & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \alpha_A \\ \alpha_B \end{pmatrix} + \mathcal{O}(\delta^3). \tag{9}$$

Theorem 2 provides a succinct expression of the incremental benefit of the sequential-move game. The first nontrivial term in the difference (9) is of order δ^2, which means that these two contribution vectors have the same linear term. The square term $\delta^2 \begin{pmatrix} (G_{AB}G_{BA})^D \alpha_A \\ \mathbf{0} \end{pmatrix}$ only depends on α_A but not on α_B. The literature on network economics is primarily interested in the situations wherein δ is reasonably small (for stability consideration) but not completely negligible (so that the peer effects remain active). In such a scenario, we can concentrate on the δ^2 term and ignore the higher order term.

Group selection. The primary advantage of Theorem 2 is to facilitate the network design problem after characterizing the equilibrium outcomes. In our two-stage setup, suppose that we are free to choose any subset of players into the leader group A. If our goal, as a central planner, is to maximize the aggregate contribution, the problem can be written as follows:

$$\max_{A \subset N} \; L(A) = \mathbf{1}'_A (G_{AB}G_{BA})^D \alpha_A = \sum_{i \in A} \sum_{j \in B} g_{ij} g_{ji} \alpha_i, \tag{10}$$

where $\mathbf{1}_A$ is a vector of 1s with length $|A|$. The aggregate contribution is used as the criterion because it represents the collective efforts by all the members in the organization, and empirically this may be measured (through some proxies) by outside observers. In Section 8, we discuss the alternative criterion – the aggregate payoff – and reexamine the same research question.

This program (10) is a combinatorial optimization problem. As an example, suppose that G is the adjacency matrix of an undirected graph; i.e., $g_{ij} \in \{0, 1\}$, and $g_{ij} = g_{ji}$. In this case, let N^i denote the set of neighbors of i. The objective function is then:

$$L(A) = \sum_{i \in A} \sum_{j \in B} g_{ij} g_{ji} \alpha_i = \sum_{i \in A} \sum_{j \in B} g_{ij} g_{ij} \alpha_i = \sum_{i \in A} \sum_{j \in B} g_{ij} \alpha_i = \sum_{i \in A} \#|N^i \cap B| \alpha_i,$$

where $\#|N^i \cap B|$ counts the number of i's neighbor nodes that are not in the set A ($B = N \backslash A$), and $L(A)$ is the sum of these nodes weighted by the numbers α_i over all the nodes in the set A. Here we use the fact that $g_{ij}^2 = g_{ij}$ as $g_{ij} \in \{0, 1\}$.

When the nodes are homogeneous ($\alpha_i = \alpha, \forall i \in N$), the problem is then to find a set A such that the number of *cuts* between A and its complement $N \backslash A$ is maximized. This problem is labeled as "maximum-cut" (MAX-CUT) and is well studied in the literature of computer science and operations research. It is easy to describe but in general an NP-hard problem. Polynomial-time solvable cases include the planar graphs (Hadlock (1975)), graphs without K_5 minors (Barahona (1983)), and weakly bipartite graphs with non-negative weights (Grötschel and Pulleyblank (1981)). Numerous heuristics and approximation algorithms have been proposed. For example, Goemans and Williamson (1995) provide a 0.878-approximation randomized algorithm (i.e., the worst-case efficiency loss is around 12%). Most of these algorithms fall into one of the two categories: best-in and worst-out. The former starts with an empty graph and in each step determines which node to add to the existing subgraph, whereas the latter starts with an initial graph and identifies the node that performs the worst to kick out. Notably, since $L(A)$ corresponds not only to the weighted number of cross links but also the second-order term of the aggregate contribution, the approximation ratio (worst-case bound) of any polynomial-time approximation scheme for MAX-CUT applies directly to the aggregate contribution.

Since the algorithmic treatments of the MAX-CUT problems are abundant in the literature, in this paper we suppress the discussions and simply present some analytically solvable cases. First, we observe that an obvious upper bound of $L(A)$ is $\#|N^i|\alpha$, the number of all the links in N multiplied by α. The solutions are easy for some special graphs. For example, if $\alpha_i = 1, \forall i$, the upper bound of the program (10) is the total number of edges on the graph G. For a bipartite graph $K_{m,n}$, all the links are between two groups and there is no edges between any pair of nodes within each group. Given the structure, the optimal solution to the program (10) is $A = M$ or $A = N$. Another example is the circle of $2n$ nodes, O_{2n}. It is a subgraph of the bipartite graph $K_{n,n}$. Therefore, the solution is given by $A = \{1, 3, 5, \cdots, 2n-1\}$ (which is unique up to isomorphism). For a circle with odd nodes O_{2n+1}, $2n+1$ remains the upper bound, but this cannot be achieved. The maximum cut has only $2n$ edges. As the third example, consider a complete graph K_n in which every node is connected to all other nodes. The optimal size of A is $\arg\max_{k=1,2,\cdots,n} k(n-k)$, which is $n/2$ if n is even, or $(n \pm 1)/2$ if n is odd.

The above discussions suggest a clear principle of finding the optimal partition for the two-stage game for arbitrary graphs. In the next two sections, we work out some specific examples and seek additional principles for this network design problem.

5 Key leader problem

A special case of the above group selection problem may be of particular interest: if we are allowed to pick up only one player to move first, who should be the person of interest? This *key leader* problem parallels the key player problem in Ballester et al. (2006), who argue that the outsider may be able to remove one player from the criminal network and therefore intends to identify the key player that impacts the network most. Thanks to equations (6) and (7), we can give an exact index, called L-index, to characterize this key leader problem without using Taylor expansions. To simplify the notation, we assume G is symmetric in this section.

To identify the key leader, we consider a sequential-move game in which player i moves in the first stage and the rest move simultaneously in the second stage. The equilibrium contribution of player i, using (6), is

$$x_i^L = \frac{\alpha_i + \delta \left\langle \beta_i,\ (\mathbf{I} - \delta G_{-i})^{-1} \alpha_{-i} \right\rangle}{1 - 2\delta^2 \left\langle \beta_i,\ (\mathbf{I} - \delta G_{-i})^{-1} \beta_i \right\rangle}. \tag{11}$$

where matrix G is rewritten as follows:

$$G = \begin{pmatrix} 0 & \beta_i' \\ \beta_i & G_{-i} \end{pmatrix}.$$

Meanwhile, using (7), the equilibrium contribution of player i in the simultaneous-move game is

$$x_i^N = \frac{\alpha_i + \delta \left\langle \beta_i,\ (\mathbf{I} - \delta G_{-i})^{-1} \alpha_{-i} \right\rangle}{1 - \delta^2 \left\langle \beta_i,\ (\mathbf{I} - \delta G_{-i})^{-1} \beta_i \right\rangle} = b_i(G, \delta, \alpha) = \sum_{j=1}^{N} m_{ij} \alpha_j. \tag{12}$$

Comparing the coefficients of α_j in (12), we obtain that

$$\begin{aligned} m_{ii} &= \frac{1}{1 - \delta^2 \left\langle \beta_i,\ (\mathbf{I} - \delta G_{-i})^{-1} \beta_i \right\rangle}, \\ \frac{m_{ij}}{m_{ii}} &= j\text{-th entry of } (\mathbf{I} - \delta G_{-i})^{-1} \delta \beta_i, \quad j \neq i. \end{aligned} \tag{13}$$

We can then derive the equilibrium contributions of other players. Afterwards, we compare across scenarios with different leaders to determine the key leader. The results are summarized in the following proposition.

Proposition 2. *Define*

$$L_i := \frac{(m_{ii} - 1)}{(2 - m_{ii})} \frac{b_i(G, \delta, 1)}{m_{ii}} b_i(G, \delta, \alpha) \tag{14}$$

as the leading index *of player i. The solution to the key leader problem, i^*, has the highest leading index,i.e, $L_{i^*} \geq L_j, \forall j \in N$.*

The above approach also gives a new derivation of the inter-centrality measure defined in Ballester et al. (2006) who study the overall impact if a player is removed from the network. Removing a player i from the network is equivalent to changing the player i's contribution from $x_i^N = b_i(G, \delta, \alpha)$ to 0, or $\Delta x_i = -b_i(G, \delta, \alpha)$. The following corollary re-establishes (Ballester et al. 2006: Theorem 3) using our derivations.

Corollary 2. *Define $c_i := \frac{b_i(G,\delta,1)}{m_{ii}} b_i(G, \delta, \alpha)$ as the inter-centrality measure. The solution to the key player problem, j^*, has the highest inter-centrality measure, i.e. $c_{j^*} \geq c_j, \forall j \in N$.*

In Ballester et al. (2006), the inter-centrality measure is defined as $c_i = \frac{b_i^2(G,\delta,\alpha)}{m_{ii}}$, which is consistent with our new definition, as they assume that $\{\alpha_i\}$'s are homogeneous (and are normalized to 1). Our new definition works even if $\{\alpha_i\}$'s are heterogeneous. Also, when $\{\alpha_i\}$'s are heterogeneous, one cannot naively use $\frac{b_i^2(G,\delta,\alpha)}{m_{ii}}$ as the inter-centrality measure; the appropriate extension of their result is stated in the above corollary.

To see the difference of inter-centrality measure and L-index, let us assume that $\alpha_i = 1, \forall i$. Consider two nodes i, j in a network G and suppose that $b_i(G,\delta,1) = b_j(G,\delta,1)$[3] and $1 \le m_{jj} < m_{ii} < 2$. It is easy to see that $c_i < c_j$, but $L_i > L_j$.[4] In other words, for regular graph, the key player is the node with lowest m_{ii}, while the key leader is the node with highest m_{ii}. This can be seen from the following example. In Figure 1, we revisit a regular graph studied in Calvo-Armengol and Jackson (2004), where there are three kinds of players: 1, 2, 3. The calculations in Table 5 show that player 3 is the key leader, whereas player 1 is the key player.

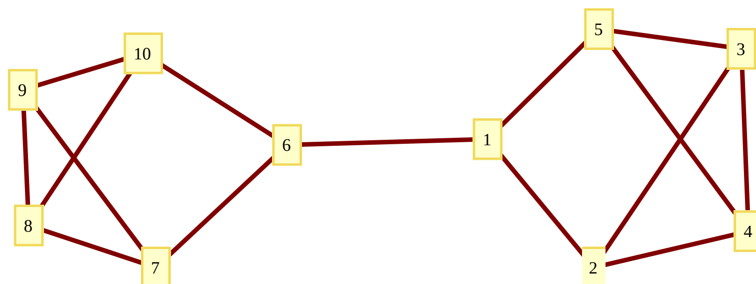

Figure 1: A regular graph, from an example in Calvo-Armengol and Jackson (2004)

players	m_{ii}	b_i	c_i	L_i
1	1.168	2.5	**5.347**	1.086
2	1.198	2.5	5.216	1.288
3	**1.216**	2.5	5.139	**1.417**

Table 1: Comparison of different measures with $\delta = 0.2$

6 Some specific network structures

To better articulate the nature of equilibrium outcomes, in this section we study some special graphs.

[3]For example, if the graph G is regular with degree k, $b_i(G,\delta,1) = \frac{1}{1-d\delta}$ for every node.

[4]Because $\frac{\partial(b_i^2/m_{ii})}{\partial m_{ii}} = -b_i^2/m_{ii}^2 < 0$ and $\frac{\partial\left(\frac{(m_{ii}-1)}{(2-m_{ii})}\frac{b_i^2}{m_{ii}}\right)}{\partial m_{ii}} = \frac{b_i^2(1+(m_{ii}-1)^2)}{((2-m_{ii})m_{ii})^2} > 0$.

6.1 Complete graph

Let us first start with unweighted complete graphs, i.e., every node is connected to the rest of the graph. Since it is unweighted, without loss of generality we make $g_{ij} = 1$, $\forall i \neq j$. This does not imply that all the nodes are identical, however, because players may be endowed with heterogeneous valuations ($\{\alpha_i\}$'s). Figure 2 presents an example of the complete graph with $N = 4$ nodes. For general N, the adjacency matrix can be written as $G = \mathbf{J} - \mathbf{I}$, where $\mathbf{I}$ is the identity matrix and $\mathbf{J}$ is the matrix of 1s.

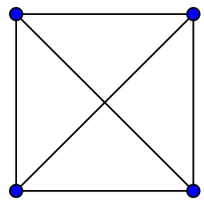

Figure 2: a complete graph with 4 nodes

Using Theorem 1, we can obtain the equilibrium outcomes in the sequential-move game and compare them with those in the simultaneous-move game. The results are summarized in the following proposition.

Proposition 3. *In an N-node unweighted complete graph, the aggregate equilibrium contribution in the two-stage game with L leaders is given by:*[5]

$$\frac{\sum_{j\in N\setminus L}\alpha_j}{1-(N-L-1)\delta} + \frac{1+\delta}{1-(N-L-1)\delta}\frac{(1-(N-L-1)\delta)\sum_{i\in L}\alpha_i + \delta L\sum_{j\in N\setminus L}\alpha_j}{1-\delta(N-2)-\delta^2(2N-L-1)}. \tag{15}$$

Proposition 3 gives a simple expression of the aggregate contribution, which is our primary objective in this paper. It is possible to characterize the equilibrium contributions at the individual levels, although they depend heavily on the detailed network structure and therefore the exact expressions are very complicated. Incidentally, we can also apply Theorem 1 to derive the above results; nonetheless, it requires inverting the matrices.

We now build upon Proposition 3 and investigate the optimal partition of leader-follower groups. In the two-node case, this problem is trivial – one player is the leader, whereas the other is the follower. Proposition 3 implies that the player with a higher intrinsic valuation should move first. To see this, we let $N = 1, F = 1$; (15) then reduces to $\frac{(1+\delta)\alpha_l+(1+\delta-\delta^2)\alpha_f}{1-2\delta^2}$. Since the weight

[5]Here we need the condition that δ is small enough such that

$$1-\delta(N-1)>0,\ 1-\delta(N-2)-\delta^2(2N-L-1)>0, \forall L=1,2,\cdots,N.$$

This is equivalent to $1-(N-2)\delta-(2N-2)\delta^2>0$, or $\delta < \frac{2}{N-2+\sqrt{N^2+4N-4}}$. A sufficient condition is $\delta < \frac{1}{N}$.

on α_l is greater, the player with a higher α should move first. With more than two nodes, the set partitioning becomes involved. To this end, we decompose our analysis into two steps.

Optimal selection of leaders. First, we assume that the size of leader groups is fixed exogenously. In this case, the remaining question is which players to select into the leader group. Proposition 3 leads to a clear-cut prediction (that is similar to the two-node case) regarding the selection of leader group.

Corollary 3. *Consider the two-stage game with L leaders for an N-node unweighted complete graph. To maximize the aggregate contribution, the players with the L highest $\{\alpha_i\}$'s should be chosen to move in the first stage.*

We have argued this feature for the two-node case. The same intuition carries over to arbitrary complete nodes with any fixed number of leaders. Specifically, the leaders' contributions induce positive feedbacks from all the followers. Each leader anticipates these positive feedbacks and therefore her contribution is amplified by N times, the size of follower group that she has direct influences upon. The pecking order suggested by Corollary 3 follows as a higher leader's intrinsic valuation leads to a more significant amplification.

Optimal size of leader group. Corollary 3 allows us to simplify the problem of finding the optimal partition. We now proceed to examine the optimal number of leaders. To concentrate on the timing rather than on the individual heterogeneity, we assume that the players are homogeneous. In this case, $\alpha_i = \alpha_j$, and without loss of generality $\alpha_i = 1$, $\forall i \in N$. With this simplification, we can express the aggregate contribution, $\phi(L)$, as follows:

$$\begin{aligned}\phi(L) &= \frac{\sum_{j\in N\setminus L}\alpha_j}{1-(N-L-1)\delta} + \frac{1+\delta}{1-(N-L-1)\delta} \times \frac{(1-(N-L-1)\delta)\sum_{i\in L}\alpha_i + \delta L \sum_{j\in N\setminus L}\alpha_j}{1-\delta(N-2)-\delta^2(2N-L-1)} \\ &= \frac{N-L}{1-(N-L-1)\delta} + \frac{1+\delta}{1-(N-L-1)\delta} \times \frac{(1+\delta)L}{1-\delta(N-2)-\delta^2(2N-L-1)}.\end{aligned}$$

For sufficiently small δ, we again use Taylor expansions to get the leading terms of $\phi(L)$.

Proposition 4. *The first three terms of $\phi(L)$ as a function of δ:*

$$\phi(L) = N + \delta N(N-1) + \delta^2 \left\{N(N-1)^2 + L(N-L)\right\} + \mathcal{O}(\delta^3).$$

Note that the linear term does not depend on L at all. As a comparison, in the simultaneous-move game, $L = 0$:

$$\phi(0) = \frac{N}{1-\delta(N-1)} = N + \delta N(N-1) + \delta^2 N(N-1)^2 + \mathcal{O}(\delta^3),$$

which corresponds to the aggregate contribution. Hence,

$$\phi(L) - \phi(0) = L(N-L)\delta^2 + \mathcal{O}(\delta^3).$$

This has an intuitive interpretation using Theorem 2. Since the original graph is complete, each node in L is connected to all the $N-L$ nodes in the second stage. Therefore, the total number between these two groups are $L(N-L)$. Obviously, the optimal size is $\lfloor \frac{N}{2} \rfloor$ if δ is small. This provides a theoretical ground for a simple rule of thumbs – the 50-50 rule.

Finally, let us comment on the heterogeneous case. If $\{\alpha_i\}$'s are different, the problem of finding the optimal size of leader group can be done in N steps. In each step, we calculate the sum of the L highest $\{\alpha_i\}$'s multiplied by $N-L$. Afterwards, we compare the maximum of these numbers. This is certainly doable but the insights are less transparent.

6.2 Complete bipartite graph

In this subsection, we study another important family of networks – the bipartite graphs. For a complete bipartite graph $K_{M,N}$, the adjacency matrix is

$$G^{(M,N)} = \begin{bmatrix} \mathbf{0} & J_{M\times N} \\ J_{N\times M} & \mathbf{0} \end{bmatrix},$$

where J_{mn} is an m by n matrix of 1s. By the definition of bipartite graph, there is no direct link between any pair of players within the same group. The graph is complete when every node in group M is connected to every node in group N. See Figure 3 for an example of a complete bipartite graph with $M=3$ and $N=2$. In other words, there is an important feature that naturally distinguishes them into two groups. In the social network literature, the canonical example is the match between men and women; the star (hub-spokes) network is a special case of bipartite graph with a single node in one group.

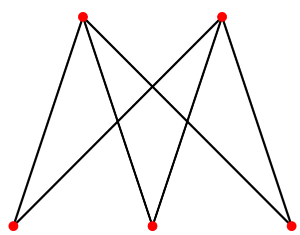

Figure 3: A complete bipartite graph $K_{3,2}$

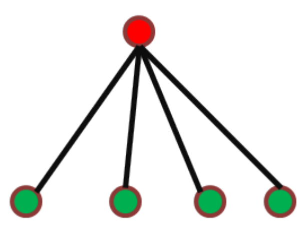

Figure 4: A star with 5 nodes

We again apply Theorem 1 to this specific graph to obtain the equilibrium outcomes in the sequential-move game. Let $\bar{\alpha} = \frac{\sum_{i \in M} \alpha_i}{M}$ and $\bar{\beta} = \frac{\sum_{j \in N} \beta_j}{N}$ denote the average marginal utilities in these two groups. We obtain the following result.

Proposition 5. *Consider the two-stage game of a complete bipartite graph. The aggregate contribution when group M moves first is higher than the aggregate contribution when group N moves first if and only if*

$$\frac{(1 + N\delta)(\bar{\alpha} + N\delta\bar{\beta})}{(1 - \delta^2(MN + N))} > \frac{(1 + M\delta)(\bar{\beta} + M\delta\bar{\alpha})}{(1 - \delta^2(MN + M))}. \tag{16}$$

As a corollary, we have:

Corollary 4. *In the two-stage game of a complete bipartite graph, if the two groups have the same size, i.e., $M = N$, the group with higher average α should move first.*

Corollary 4 re-establishes the insight we obtain from complete graphs. From the social efficiency perspective, we shall allow players who are highly intrinsically motivated to lead others. As another special case, we suppose that the average valuations are the same between the two groups: $\bar{\alpha} = \bar{\beta}$. In particular, this holds true when groups are homogeneous, i.e., $\alpha_i = \beta_j, \forall i, j$.

Corollary 5. *In the two-stage game of a complete bipartite graph, if $\bar{\alpha} = \bar{\beta}$, the group with a smaller size should move first.*

Given that the star network is a special case of the bipartite graph, Corollary 5 shows that the center hub should move first. This somehow provides a justification for why fashion influencers are typically composed of a relatively *small* group of people, and a priori they are substantially different from others. Incidentally, (16) is equivalent to:

$$\bar{\alpha} - \bar{\beta} + (N - M)\delta(\bar{\alpha} + \bar{\beta}) + \delta^2 \left((N^2 + N)\bar{\beta} - (M^2 + M)\bar{\alpha}\right) > 0.$$

This condition is more likely to hold if $\bar{\alpha}$ is high or the group size N is large. These results are consistent with Corollaries 4 and 5. Also, if δ is relatively small, the dominant term is $\bar{\alpha} - \bar{\beta}$. In other words, if the strength of the complementarity is reasonably bounded, the group with a higher average marginal utility should move first.

7 General sequence

In this section, we characterize the structure of optimal mechanism. Building upon the results from the two-stage settings, we observe that making the players move sequentially yields a higher aggregate contribution. In this section, we prove that this is a general principle. Pushing this idea to the extreme, the optimal sequence turns out to be a *chain* wherein players move one by one. To set up the general sequence problem, we first define some technical terms.

Definition 1. *A sequence* $\mathcal{S} = (P_1, P_2, \cdots, P_k)$ *is just a partitioning of* $\mathcal{N}$ *such that* $P_i \cap P_j = \emptyset, \forall i \neq j$ *and* $\cup_{1\leq i\leq k} P_i = \mathcal{N}$*. The number* $|\mathcal{S}| = k$ *corresponds to the number of steps of this sequence.*

For convenience, let $\emptyset$ denote the sequence for the simultaneous-move game. For each $\mathcal{S}$, we can define an *extensive-form* game with complete information as follows. First, players in P_1 move in the first period simultaneously, players in P_2 move in the second period simultaneously, $\cdots$, and players in P_k move in the k-th period simultaneously. The actions, once taken, are observable to all the remaining players who move later.

Fixed sequence. We start with the problem with a fixed sequence of moves. This intermediate step not only is necessary for establishing the optimality of chain structure but also distills the key drivers of sequential moves. We use backward induction to characterize the equilibrium outcomes and state our main results in the next theorem.

Theorem 3. *For sufficiently small* δ*, for any sequence of moves* $\mathcal{S}$*, the equilibrium outcome* $\mathbf{x}(\mathcal{S}, G, \delta, \alpha)$ *is a linear function of* α*:*

$$\mathbf{x}(\mathcal{S}, G, \delta, \alpha) = Z(\mathcal{S}, G, \delta)\alpha,$$

where $Z(\mathcal{S}, G, \delta)$ *is a matrix independent of* α*. The aggregate equilibrium contribution is*

$$\sum x_i(\mathcal{S}, G, \delta, \alpha) = \mathbf{1}' Z(\mathcal{S}, G, \delta)\alpha,$$

where $\mathbf{1}$ *is a vector of 1s. Moreover, if* G *is symmetric, then* $Z(\mathcal{S}, G, \delta)$ *is also symmetric.*

As a special case, consider the sequence $\mathcal{S} = (N)$, i.e., all players move together in the "first stage." In this case, $Z(\emptyset, G, \delta) = [1 - \delta G]^{-1}$ and it coincides with the result in the simultaneous-move game. The sequence $\mathcal{S} = (A, N\backslash A)$ corresponds to the two-stage game wherein group A is the leader group, and the remaining players move in the second stage. In this case, $Z(\mathcal{S}, G, \delta)$ is derived by (6) in Theorem 1.

The proof of Theorem 3 also gives us an *algorithm* to compute the matrix $Z(\mathcal{S})$ for any sequence $\mathcal{S} = (P_1, P_2, \cdots, P_k)$. Recall that in analyzing the two-step game, $U = [I - \delta G_{BB}]^{-1}$ plays an important role in equilibrium characterizations. In the general sequential-move game, a corresponding matrix (labeled as U too for consistency) can be established via the induction steps. Given the matrix U, we can apply the formula in (31) in the appendix to find $Z(\mathcal{S})$. This can be done in $O(N^3)$ steps and thus does not lead to computational burden. Nevertheless, to find the matrix U we need to apply Theorem 3 to the subsequence $\mathcal{S}' = (P_2, \cdots, P_k)$. In the sequel, when G and δ are obvious in the context, we write $Z(\mathcal{S})$ and $\mathbf{x}(\mathcal{S}, a)$ for short.

Optimality of chain structure. Having characterized the equilibrium outcome for a fixed sequence, we now proceed to compare equilibria for different sequences. To this end, we shall formally define how to refine a given sequence.

Definition 2. *Sequence* $\tilde{\mathcal{S}}$ *is a refinement of* $\mathcal{S} = (P_1, P_2, \cdots, P_k)$ *if there exists* r *such that* $\tilde{\mathcal{S}} = (P_1, \cdots, P_{r-1}, Q_1, Q_2, P_{r+1}, \cdots P_k)$ *where* $Q_1 \cup Q_2 = P_r$ *and* $Q_1 \cap Q_2 = \emptyset$*.*

Obviously, $(A, N\backslash A)$ is a refinement of the sequence $\emptyset$. Previously, we have shown that the matrix for the two-stage game dominates the matrix in the simultaneous-move game. The next proposition proves that this pattern holds for any refinement.

Proposition 6. *For any fixed G, suppose that $\mathcal{S}'$ is a refinement of $\mathcal{S}$. Then $Z(\tilde{\mathcal{S}}, G, \delta) \succeq Z(\mathcal{S}, G, \delta)$.*

Proposition 6 suggests that sequentiality improves the social efficiency in a strong sense, as every refinement necessarily leads to a higher aggregate contribution. Note that this result does not hinge on any structural assumptions of the underlying network; it only requires the strategic complementarity, i.e., $G \geq 0$, embedded in this social network context. This result is not obvious as it stands, as changing the sequence affects the incentives of the nodes in the beginning of the sequence as well. Proposition 6 also leads to a handy guideline. If we are bound by some physical restrictions and cannot freely choose any sequence, we can always improve upon the current situation by injecting more sequentiality "*locally*" to the process. This unambiguously and strictly improves the aggregate contribution and social welfare.

Pushing the above idea to the extreme, we can characterize the optimal sequence. Define a *chain* as a sequence with step N, i.e., players determine their contributions one by one. As a corollary of Theorem 6, we obtain the following.

Corollary 6. *For any fixed G, a chain maximizes the aggregate contribution among the players.*

Corollary 6 establishes the optimality of chain structure. The intuition is as follows. Recall that in the two-stage games we show that the aggregate contribution is tightly connected to the number of cross links between the leader and follower groups. This constitutes a proxy of the magnitude of positive feedback effects in the two-stage games. When we are entitled to determine the sequence arbitrarily, naturally we shall intensify the feedback effects to the extent possible. It turns out that the best way to exploit the indirect influences is to maximize the numbers of links across groups from different layers, and this is attained by the chain structure.

Also, we can show a counterpart of Theorem 2 for general sequence $\mathcal{S}$, which might be useful in its own right.

Theorem 4. *For any sequence $\mathcal{S} = (P_1, P_2, \cdots, P_k)$, we have the Taylor expansion for*

$$Z(\mathcal{S}, G, \delta) - Z(\emptyset, G, \delta) = \delta^2 \Lambda(\mathcal{S}) + \mathcal{O}(\delta^3)$$

where

$$\Lambda(\mathcal{S}) = \begin{pmatrix} \left(\sum_{j=2}^{k} G_{P_1 P_j} G_{P_j P_1}\right)^D & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \left(\sum_{j=3}^{k} G_{P_2 P_j} G_{P_j P_2}\right)^D & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \left(G_{P_{k-1} P_k} G_{P_k P_{k-1}}\right)^D & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix},$$

and $Z(\emptyset, G, \delta) = [1 - \delta G]^{-1}$.

The difference of the matrix $D(\mathcal{S}) = Z(\mathcal{S}, G, \delta) - Z(\emptyset, G, \delta)$, up to the error term of $\mathcal{O}(\delta^3)$, is a diagonal matrix, with entry $d_{ii} = \delta^2 \sum_{j \in \cup_{s=r+1}^{k} P_s} g_{ij} g_{ji}$ if $i \in P^r$. For the perspective of the nodes in P_1, $\left(\sum_{j=2}^{k} G_{P_1 P_j} G_{P_j P_1}\right)^D$ only depends on the set $\cup_{r=2}^{k} P_r = N \backslash P_1$, but not on the exact decomposition of the subsequence $\mathcal{S}' = (P_2, \cdots, P_k)$. Similarly, the equilibrium contributions of nodes in P_2 depend only upon $\cup_{r=3}^{k} P_r$ but not the exact decomposition.

We can define the function on the sequence $\mathcal{S} = (P_1, \cdots, P_k)$ as follows:

$$L(\mathcal{S}) = \sum_{r=1}^{k} \sum_{i \in P^r} \left(\sum_{j \in \cup_{s=r+1}^{k} P_s} g_{ij} g_{ji} \right) \alpha_i.$$

If G is the adjacency matrix for an undirected graph, then the matrix L counts the path from each node to the nodes who move after her, weighted by the α_i.

Corollary 7. *If $\mathcal{S}'$ is a refinement of $\mathcal{S}$, then $L(\mathcal{S}') \geq L(\mathcal{S})$.*

This result can also be shown directly by the definition of L, and it echoes Proposition 6 in that any sort of sequentiality strictly improves the social welfare. Finally, we consider a special case with homogeneous intrinsic valuations.

Corollary 8. *Suppose that G is the adjacency matrix for an undirected graph without self loop ($g_{ii} = 0$), and $\alpha_i = 1, \forall i$. If the sequence of $\mathcal{S}$ is a chain, then $L(\mathcal{S}) = \frac{1}{2} \sum_{ij} g_{ij}$, which counts the number of links in G. In particular, this number does not depend on the configuration of the chain.*

Notably, Corollary 8 implies that the sequencing given a chain is irrelevant for the aggregate contribution, as long as players have homogeneous intrinsic valuations. Thus, the implementation of the optimal hierarchy is surprisingly simple. The underlying reason for this irrelevance is the following. When the players have homogeneous intrinsic valuations, the aggregation contribution depends mainly on the aggregation of positive feedbacks within players. Therefore, when a chain structure is adopted, the aggregate contribution is reflected by the summation of all the feedback effects, irrespective of the order by which these effects are aggregated.

An example. To close this section, we use the following example to illustrate the intuition behind our main theorem. The same idea applies to any general graph. Consider the complete graph with 6 nodes (see Figure 5). In this case, the adjacent matrix is symmetric, and all the off-diagonal entries are equal to 1, i.e, $g_{ij} = 1, \forall i \neq j$. Recall that the aggregate contribution is directly pinned down by the best response matrix. For the sequence $\mathcal{S}_1 = \{\{1, 2\}, \{3, 4, 5, 6\}\}$, $L(\mathcal{S}_1) = 4\alpha_1 + 4\alpha_2$ (see Table 2 for details). This measures the positive feedbacks and is a direct index of aggregate contribution.

Now we consider an alternative sequence $\mathcal{S}_2 = \{\{1, 2\}, \{3, 4\}, \{5, 6\}\}$, which is apparently a refinement of $\mathcal{S}_1$ as we split the group $\{3, 4, 5, 6\}$. Under this sequence, $L(\mathcal{S}_2) = 4\alpha_1 + 4\alpha_2 + 2\alpha_3 + 2\alpha_4$ (see Table 3). After further refinements, in the end we reach the chain structure, sequence $\mathcal{S}_3 =$

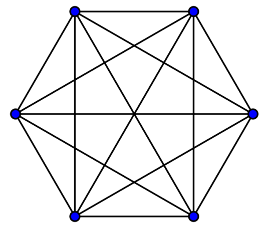

Figure 5: K_6: a complete graph with 6 nodes

$\{\{1\},\{2\},\{3\},\{4\},\{5\},\{6\}\}$. The corresponding matrix is $L(\mathcal{S}_3) = 5\alpha_1 + 4\alpha_2 + 3\alpha_3 + 2\alpha_4 + \alpha_5$ (see Table 4). It is obvious to see that

$$\begin{aligned} & L(\mathcal{S}_3) = 5\alpha_1 + 4\alpha_2 + 3\alpha_3 + 2\alpha_4 + \alpha_5 \\ > \; & L(\mathcal{S}_2) = 4\alpha_1 + 4\alpha_2 + 2\alpha_3 + 2\alpha_4 \\ > \; & L(\mathcal{S}_1) = 4\alpha_1 + 4\alpha_2. \end{aligned}$$

0	1	**1**	**1**	**1**	**1**
1	0	**1**	**1**	**1**	**1**
1	1	0	1	1	1
1	1	1	0	1	1
1	1	1	1	0	1
1	1	1	1	1	0

Table 2: Sequence $\mathcal{S}_1 = \{\{1,2\},\{3,4,5,6\}\}$, $L(\mathcal{S}_1) = 4\alpha_1 + 4\alpha_2$

0	1	**1**	**1**	**1**	**1**
1	0	**1**	**1**	**1**	**1**
1	1	0	1	**1**	**1**
1	1	1	0	**1**	**1**
1	1	1	1	0	1
1	1	1	1	1	0

Table 3: Sequence $\mathcal{S}_2 = \{\{1,2\},\{3,4\},\{5,6\}\}$, $L(\mathcal{S}_2) = 4\alpha_1 + 4\alpha_2 + 2\alpha_3 + 2\alpha_4$

This is consistent with Corollary 7. Thus, the aggregate contribution gets higher when we make any set of simultaneous-moving players choose their contributions sequentially. As illustrated in the tables, the additional blocks account for this additional feedback effect from the sequentiality.

0	**1**	**1**	**1**	**1**	**1**
1	0	**1**	**1**	**1**	**1**
1	1	0	**1**	**1**	**1**
1	1	1	0	**1**	**1**
1	1	1	1	0	**1**
1	1	1	1	1	0

Table 4: Sequence $\mathcal{S}_3 = \{\{1\},\{2\},\{3\},\{4\},\{5\},\{6\}\}$, $L(\mathcal{S}_3) = 5\alpha_1 + 4\alpha_2 + 3\alpha_3 + 2\alpha_4 + \alpha_5$

8 Extensions

In this section, we consider some variants of our model characteristics.

8.1 Strategic substitution

While we primarily focus on the strategic complementarity, we note that Ballester et al. (2006) also include the case with strategic substitutes. To incorporate this effect, we should relax the assumption that $g_{ij} \geq 0$ for all i, j. We shall also allow for some combinations of positive and negative components in matrix G. Since $\{g_{ij}\}$'s capture the local relationships amongst players, a reasonable assumption is the reciprocity (and its inverse): if $g_{ij} \geq 0$, then $g_{ji} \geq 0$; on the other hand, if $g_{ij} < 0$, then $g_{ji} < 0$. This ensures that if player j's action imposes a negative externality on player i's utility, then so does player i's action on player j's utility.

With this assumption, most of our results continue to hold in this alternative setting. To see this, first consider the two-stage game. Recall from (9) that the contribution differential depends on $\begin{pmatrix} (G_{AB}G_{BA})^D & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \alpha_A \\ \alpha_B \end{pmatrix}$. This suggests that g_{ij} and g_{ji} always appear together in the product form $g_{ij}g_{ji}$. Regardless of whether they represent strategic substitutes or complements, the product term is necessarily positive. The group selection problem is $\max_{A \subset N} \sum_{i \in A} \sum_{j \in B} g_{ij} g_{ji} \alpha_i$. Therefore, this maximum cut problem is mathematically equivalent to that for an alternative graph for which each component g_{ij} is replaced by $|g_{ij}|$. Our analysis then applies to this alternative graph (with strategic complementarity). The examples with complete and bipartite graphs can be analyzed in exactly the same manner.

When we are able to choose the general sequence, the aggregate contribution can be expressed as: $L(\mathcal{S}) = \sum_{r=1}^{k} \sum_{i \in P^r} \left(\sum_{j \in \cup_{s=r+1}^{k} P_s} g_{ij} g_{ji} \right) \alpha_i$. We observe again the product terms $g_{ij}g_{ji}$. Therefore, even if the original graph features a mixture of strategic substitution and complementarity, any refinement of sequence strictly increases the aggregate contribution. As a result, the optimal sequence turns out to be a chain as predicted in Section 7. When the players have homogeneous intrinsic valuations but differ in their network-related payoffs ($\{g_{ij}\}$'s), the aggregate contribution given a chain does not depend on its configuration, and thus the implementation remains simple. All the above arguments do not require any a priori symmetry in $\{g_{ij}\}$'s that goes beyond the reciprocity assumption.

8.2 Aggregate payoff

Insofar we use the aggregate contribution as the proxy of network activities. An alternative measure in this network game is the aggregate payoff. The aggregate payoff, given the contribution vector $\mathbf{x}$, is:

$$W(\mathbf{x}) = \sum_{i \in N} \left(\alpha_i x_i - \frac{1}{2} x_i^2 + \delta \sum_{j \in N} g_{ij} x_i x_j \right) = \langle \mathbf{x}, \alpha + \delta G \mathbf{x} \rangle - \frac{1}{2} \langle \mathbf{x}, \mathbf{x} \rangle.$$

In the simultaneous-move benchmark, (Ballester et al. 2006: page 1412) have shown that

$$W(\mathbf{x}^N) = \langle \mathbf{x}^N, \alpha + \delta G \mathbf{x}^N \rangle - \frac{1}{2} \langle \mathbf{x}^N, \mathbf{x}^N \rangle = \frac{1}{2} \sum_{i \in N} (\mathbf{x}_i^N)^2.$$

Now consider the general sequence. Recall the contribution differential in (9). As a consequence, we have the following welfare comparison result.

Lemma 2. *For small δ, the welfare difference is given by*

$$\Delta W(\mathcal{S}) = W(\mathbf{x}(\mathcal{S})) - W(\mathbf{x}(\emptyset)) = \delta^3 \alpha' G \Lambda(\mathcal{S}) \alpha + \mathcal{O}(\delta^4).$$

From Lemma 2, the leading term of ΔW is $\delta^3 \alpha' G \Lambda(\mathcal{S}) \alpha$. The order of welfare improvement is δ^3, which is one order lower than that of aggregate contribution improvement. Moreover, we observe that the finer the sequence $\mathcal{S}$, the greater the leading term of aggregate payoff. The above expression immediately leads to the following result:

Proposition 7. *If $\mathcal{S}'$ is a refinement of $\mathcal{S}$, then $\Lambda(\mathcal{S}') \succeq \Lambda(\mathcal{S})$. Therefore, the aggregate payoff under $\mathcal{S}'$ is higher than that under $\mathcal{S}$. Consequently, the sequence that maximizes the aggregate payoff is a chain.*

The above proposition suggests that using the aggregate payoff as the proxy of network design yields the same prediction as the aggregate contribution. Note that unlike the aggregate contribution, this result does *not* apply when there are strategic substitutes. In fact, if all $\{g_{ij}\}$'s are negative, i.e., the network game involves purely strategic substitutes, the chain structure leads to the worst outcome of aggregate payoff. This is because with negative $\{g_{ij}\}$'s, in the simultaneous-move game each player neglects the negative externality she imposes on others. Thus, each player tends to contribute too much compared with the first-best (coordinated) level. As discussed in Section 8.1, any increased sequentiality boosts each player's contribution; thus, the equilibrium contributions are further away from the first-best levels. This is in line with the observation of Von Stengel (2010), who considers a duopoly game with negative externalities and demonstrates that sequential-move is harmful to firms' profitability.

In addition, since Lemma 2 applies to any general sequence, we can also articulate the group selection problem in the two-stage game:

$$\max_{L \subset N} \delta^3 \alpha' G \Lambda \alpha, \text{ where } \Lambda \text{ is induced by L.}$$

This is *not* equivalent to a max-cut problem. Thus, in general aggregate contribution and aggregate payoff lead to different selections of leader groups. However, some of our findings continue to hold.

For example, suppose that the network is a complete graph. In this case, let L denote the number of leaders, and $F = N - L$ is the number of followers. Note that

$$\alpha' G = (\sum_{k \neq 1} \alpha_k, \cdots, \sum_{k \neq n} \alpha_k)',$$

and

$$\Lambda \alpha = (N - L) \cdot (\alpha_1, \cdots, \alpha_L, 0, \cdots, 0).$$

Hence, the inner product of these vectors is

$$(N - L) \sum_{i \in L} \alpha_i \sum_{k \neq i} \alpha_k = (N - L) \sum_{i \in L} \alpha_i (A - \alpha_i).$$

where $A = \sum_{i=1}^{N} \alpha_i$.

If we fix L, the number of leaders, and intend to pick L players such that the above expression is maximized. It turns out that the same pecking order applies: we shall always choose the players with the highest $\{\alpha_i\}$'s. This follows from the same pairwise interchange argument.[6] On the other hand, when players have homogeneous intrinsic valuations (all the $\{\alpha_i\}$'s are the same), then the objective becomes $\max_L (N - L)(N - 1)L$ and the maximum is attained at $L^* = \frac{N}{2}$. This leads to the $50 - 50$ rule again.

9 Conclusions

In this paper we investigate how an organization designer can promote the communication amongst their membes. Our model setup features payoff externalities and strategic complementarity amongst players. We show in this parsimonious setup that sequentiality alone can be substantially beneficial to the aggregate benefit. Specifically, we first analyze the two-stage game whereby we categorize them as the leader and follower groups. Compared with the simultaneous-move benchmark, the

[6] Suppose that there are two players i, j such that $\alpha_i \leq \alpha_j$. Consider a selection in which player i is selected into the leader group but player j is not. If we make the pairwise interchange between players i and j but keep all other players in their original groups, the contribution differential is

$$\alpha_j (\sum_{k \neq j} \alpha_k) - \alpha_i (\sum_{k \neq i} \alpha_k) = (\alpha_j - \alpha_i) \sum_{k \neq i,j} \alpha_k \geq 0.$$

This unambiguously improves the aggregate payoff (if $\{g_{ij}\}$'s are non-negative).

equilibrium contribution by any individual player in any two-stage sequential-move game is unambiguously higher. We establish the isomorphism between the optimal selection of the leader and follower groups and the classical weighted maximum-cut (MAX-CUT) problem. We give an exact index to characterize this key leader problem without using any approximation, and show that the key leader can be substantially different from the key player identified in Ballester et al. (2006).

We then apply our results to some leading examples of specific network structures. For unweighted complete graphs, we establish an intuitive pecking order for any given size of leader group, and with homogeneous intrinsic valuations it is optimal to split the players into two halves. For the complete bipartite graphs, players shall be split based on the two groups naturally defined by the bipartite structure. Furthermore, when the two groups have the same size, we shall nominate the group with a higher average intrinsic valuation as the leaders. On the other hand, when the average intrinsic valuations are the same between two groups, the one with a smaller size should move first. We then relax the two-stage restriction and examine the structure of optimal mechanism. We show that any form of sequentiality strictly improves the aggregate contribution, and consequently the optimal sequence turns out to be a chain structure, i.e., players should move one by one. Finally, with strategic substitutes, most of our results continue to hold in this alternative setting. If instead the network designer aims at maximizing the players' aggregate payoff, the chain structure remains optimal with strategic complements, but this result no longer holds with strategic substitutes. The group selection criterion differs from the maximum-cut problem, but some design principles continue to apply (such as the pecking order based on their intrinsic valuations and the $50-50$ rule). Our flexible framework may be interpreted as different sorts of social networks; therefore, the above design principles may be applicable to a broad class of contexts.

A Appendix. Proofs

Before we proceed, we first introduce some definitions and then prove a technical lemma. We say that two matrices A, B satisfy $A \succeq B$ if and only if $A_{ij} \geq B_{ij}, \forall i, j$. Let $\mathbf{0}$ denote the matrix of 0s. Then, $A \succeq B$ iff $A - B \succeq \mathbf{0}$. It is easy to show the following. (1) If $A \succeq \mathbf{0}, B \succeq \mathbf{0}$, then $A + B \succeq \mathbf{0}$. (2) If $A \succeq \mathbf{0}, B \succeq \mathbf{0}$, then $AB \succeq \mathbf{0}$. (3) If $A \succeq B \succeq \mathbf{0}, C \succeq D \succeq \mathbf{0}$, then $AC \succeq BD$.

Lemma 3. *The n-by-n matrix $a\mathbf{I} + b\mathbf{J}$ is invertible if and only if $a \neq 0$ and $a + bn \neq 0$. Moreover, the inverse matrix of $a\mathbf{I} + b\mathbf{J}$ is $\frac{1}{a}[\mathbf{I} - \frac{b}{a+bn}\mathbf{J}]$ if $a \neq 0$ and $a + bn \neq 0$.*

Proof: The eigenvalues of $\mathbf{J}$ are $(0, \cdots, 0, n)$. Hence, the eigenvalues $a\mathbf{I} + b\mathbf{J}$ are $(a, \cdots, a, a + bn)$. This then implies that it is invertible if and only if $a \neq 0, a + bn \neq 0$. The second part can be verified directly by using the fact $\mathbf{J}^2 = n\mathbf{J}$. □

Now we proceed to establish the main results in the paper.

Proof of Lemma 1. The equilibrium conditions are:

$$BR_i(\mathbf{x}_{-i}) = \beta_i + \delta(\sum_{j=1}^{N} t_{ij}x_j) + \delta x_i t_{ii} = x_i.$$

Equivalently, it means that $\left[1-\delta(T+T^D)\right]\mathbf{x}=\beta$, and therefore $\mathbf{x}=\left[1-\delta(T+T^D)\right]^{-1}\beta$. Also, if T is symmetric, $\left[1-\delta(T+T^D)\right]$ is symmetric; consequently, $\left[1-\delta(T+T^D)\right]^{-1}$ is symmetric as well. $\square$

Proof of Theorem 1. By Lemma 1, the equilibrium contribution profile $\mathbf{x}_A$ is given by:

$$\begin{aligned}\mathbf{x}_A &= \left[1-\delta(T+T^D)\right]^{-1}\beta = \left[1-\delta(T+T^D)\right]^{-1}(\alpha_A+\delta G_{AB}U\alpha_B)\\ &= \left[1-\delta\left((G_{AA}+\delta G_{AB}UG_{BA})+(G_{AA}+\delta G_{AB}UG_{BA})^D\right)\right]^{-1}(\alpha_A+\delta G_{AB}U\alpha_B).\end{aligned}$$

Also, (4) gives rise to the equilibrium contribution profile $\mathbf{x}_B$:

$$\begin{aligned}\mathbf{x}_B(\mathbf{x}_A) &= U(\alpha_B+\delta G_{BA}\left[1-\delta(T+T^D)\right]^{-1}(\alpha_A+\delta G_{AB}U\alpha_B))\\ &= \delta UG_{BA}\left[1-\delta(T+T^D)\right]^{-1}\alpha_A+\left(U+\delta^2 UG_{BA}\left[1-\delta(T+T^D)\right]^{-1}G_{AB}U\right)\alpha_B.\end{aligned}$$

For the second part, if G is symmetric, then $G'_{AA}=G_{AA}, G'_{BB}=G_{BB}$, $G'_{AB}=G_{BA}, G'_{BA}=G_{AB}$. Therefore, $U=[1-\delta G_{BB}]^{-1}$ is also symmetric, and subsequently $T=G_{AA}+\delta G_{AB}UG_{BA}$ is symmetric as well. It is then readily observable that S is symmetric. $\square$

Proof of Proposition 1. It suffices to show that the matrix S given in (6) dominates the matrix $[1-\delta G]^{-1}$, i.e., $S\succeq[1-\delta G]^{-1}$. Then, it is obvious that:

$$\begin{pmatrix}\mathbf{x}_A\\ \mathbf{x}_B\end{pmatrix}=S\begin{pmatrix}\alpha_A\\ \alpha_B\end{pmatrix}\geq\mathbf{x}^N=[1-\delta G]^{-1}\begin{pmatrix}\alpha_A\\ \alpha_B\end{pmatrix}.$$

To establish the result, we need to find a convenient expression for the $[1-\delta G]^{-1}$ using the block matrix:

$$[1-\delta G]^{-1}=\begin{pmatrix}[1-\delta(T+0)]^{-1} & \delta\,[1-\delta(T+0)]^{-1}G_{AB}U\\ \delta UG_{BA}\,[1-\delta(T+0)]^{-1} & U+\delta^2 UG_{BA}\,[1-\delta(T+0)]^{-1}G_{AB}U\end{pmatrix}, \tag{17}$$

where the matrices U and T are defined in Theorem 1. In other words, If we replace the matrix T^D by the zero matrix in (6), we will get the matrix $[1-\delta G]^{-1}$.

To this end, it is equivalent to show that the solution to this linear equation:

$$[1-\delta G]\begin{pmatrix}\mathbf{y}_A\\ \mathbf{y}_B\end{pmatrix}=\begin{pmatrix}1-\delta G_{AA} & -\delta G_{AB}\\ -\delta G_{BA} & 1-\delta G_{BB}\end{pmatrix}\begin{pmatrix}\mathbf{y}_A\\ \mathbf{y}_B\end{pmatrix}=\begin{pmatrix}\alpha_A\\ \alpha_B\end{pmatrix}$$

is given by

$$\begin{pmatrix}\mathbf{y}_A\\ \mathbf{y}_B\end{pmatrix}=\begin{pmatrix}[1-\delta(T+0)]^{-1} & \delta\,[1-\delta(T+0)]^{-1}G_{AB}U\\ \delta UG_{BA}\,[1-\delta(T+0)]^{-1} & U+\delta^2 UG_{BA}\,[1-\delta(T+0)]^{-1}G_{AB}U\end{pmatrix}\begin{pmatrix}\alpha_A\\ \alpha_B\end{pmatrix}. \tag{18}$$

We can express $\mathbf{y}_B$ in terms of $\mathbf{x}_A$, i.e., $-\delta G_{BA}\mathbf{y}_A+(1-\delta G_{BB})\mathbf{y}_B=\alpha_B$ or

$$\mathbf{y}_B=(1-\delta G_{BB})^{-1}(\alpha_B+\delta G_{BA}\mathbf{y}_A)=U(\alpha_B+\delta G_{BA}\mathbf{y}_A).$$

Plugging this into $(1-\delta G_{AA})\mathbf{y}_A - \delta G_{AB}\mathbf{y}_B = \alpha_A$, we have

$$(1-\delta G_{AA})\mathbf{y}_A - \delta G_{AB}U(\alpha_B + \delta G_{BA}\mathbf{y}_A) = \alpha_A.$$

Therefore,

$$\mathbf{y}_A = \left[1-\delta\left((G_{AA} + \delta G_{AB}UG_{BA})\right)\right]^{-1}(\alpha_A + \delta G_{AB}U\alpha_B) = [1-\delta T]^{-1}(\alpha_A + \delta G_{AB}U\alpha_B).$$

Accordingly,

$$\mathbf{y}_B = U(\alpha_B + \delta G_{BA}\mathbf{x}_A) = \delta UG_{BA}[1-\delta T]^{-1}\alpha_A + \left(U + \delta^2 UG_{BA}[1-\delta T]^{-1}G_{AB}U\right)\alpha_B,$$

which is exactly (18).

Comparing the entries of (6) and (17), we observe that $S \succeq [1-\delta G]^{-1}$ is implied by

$$\left[1-\delta(T+T^D)\right]^{-1} \succeq [1-\delta T]^{-1}.$$

This follows immediately from

$$\left[1-\delta(T+T^D)\right]^{-1} - [1-\delta T]^{-1} = [1-\delta T]^{-1}\delta T^D\left[1-\delta(T+T^D)\right]^{-1} \succeq \mathbf{0},$$

for sufficiently small δ. Here we use the fact $A^{-1} - B^{-1} = B^{-1}(B-A)A^{-1}$. Thus, the result holds true. $\square$

Proof of Corollary 1. The first part is straightforward. To compare the sequential-move and simultaneous-move games, we shall examine the modified simultaneous-move game. The equilibrium matrix for the simultaneous-move game is given by $[1-\delta(G+G^D)]^{-1}$, and (17) now becomes:

$$[1-\delta(G+G^D]^{-1} = \begin{pmatrix} \left[1-\delta\hat{T}\right]^{-1} & \delta\left[1-\delta\hat{T}\right]^{-1}G_{AB}\tilde{U} \\ \delta\tilde{U}G_{BA}\left[1-\delta\hat{T}\right]^{-1} & \tilde{U} + \delta^2\tilde{U}G_{BA}\left[1-\delta\hat{T}\right]^{-1}G_{AB}\tilde{U} \end{pmatrix},$$

while here $\hat{T} = G_{AA} + G_{AA}^D + \delta G_{AB}UG_{BA}$, and $\tilde{U} = [1-\delta(G_{BB} + G_{BB}^D)]^{-1}$.

Note that

$$T + T^D = G_{AA} + G_{AA}^D + \delta G_{AB}UG_{BA} + \delta(G_{AB}UG_{BA})^D \succeq \hat{T} = G_{AA} + G_{AA}^D + \delta G_{AB}UG_{BA}.$$

Thus, $\left[1-\delta(T+T^D)\right]^{-1} \succeq \left[1-\delta\hat{T}\right]^{-1}$ for sufficiently small δ. By the similar argument, we also have $\tilde{S} \succeq [1-\delta(G+G^D]^{-1}$ even if G contains nonzero diagonal entries. Moreover, the superiority of sequential-move game over the simultaneous-move game still holds. In other words, Proposition 1 remains valid even if G has nonzero diagonal entries. $\square$

Proof of Theorem 2. First, we define:

$$\Delta = \left[1-\delta(T+T^D)\right]^{-1} - [1-\delta T]^{-1}.$$

Comparing the entries of (6) and (17), we have:

$$\begin{pmatrix} \mathbf{x}_A \\ \mathbf{x}_B \end{pmatrix} - \begin{pmatrix} \mathbf{x}_A^N \\ \mathbf{x}_B^N \end{pmatrix} = \begin{pmatrix} \Delta & \delta\Delta G_{AB}U \\ \delta U G_{BA}\Delta & \delta^2 U G_{BA}\Delta G_{AB}U \end{pmatrix} \begin{pmatrix} \alpha_A \\ \alpha_B \end{pmatrix}.$$

We observe that

$$\Delta = \left[1-\delta(T+T^D)\right]^{-1} - [1-\delta T]^{-1} = [1-\delta T]^{-1}\,\delta T^D \left[1-\delta(T+T^D)\right]^{-1}.$$

Note that

$$\delta T^D = \delta(G_{AA}+\delta^2 G_{AB}UG_{BA})^D = \delta G_{AA}^D + \delta^2 (G_{AB}UG_{BA})^D = \delta^2 (G_{AB}UG_{BA})^D,$$

where we use the fact that $G_{AA}^D = \mathbf{0}$ ($g_{ii}=0, \forall i$). Therefore, Δ is at least of order δ^2, and consequently $\delta\Delta G_{AB}U = \mathcal{O}(\delta^3)$. Similarly, we can show that $\delta U G_{BA}\Delta = \mathcal{O}(\delta^3)$, and $\delta^2 U G_{BA}\Delta G_{AB}U = \mathcal{O}(\delta^4)$.

Therefore, the only δ^2 term is contained in Δ :

$$\Delta = \delta^2 [1-\delta T]^{-1} (G_{AB}UG_{BA})^D \left[1-\delta(T+T^D)\right]^{-1}.$$

Here, $U = [1-\delta G_{BB}]^{-1} = I + \mathcal{O}(\delta)$,

$$(G_{AB}UG_{BA})^D = (G_{AB}(1+\mathcal{O}(\delta))G_{BA})^D = (G_{AB}G_{BA})^D + \mathcal{O}(\delta).$$

Also note that $[1-\delta T]^{-1} = 1+\mathcal{O}(\delta)$, and $\left[1-\delta(T+T^D)\right]^{-1} = 1+\mathcal{O}(\delta)$. Thus,

$$\begin{aligned} \Delta &= \delta^2 [1-\delta T]^{-1} (G_{AB}UG_{BA})^D \left[1-\delta(T+T^D)\right]^{-1} \\ &= \delta^2 (1+\mathcal{O}(\delta)) \left[(G_{AB}G_{BA})^D + \mathcal{O}(\delta)\right] (1+\mathcal{O}(\delta)) \\ &= \delta^2 (G_{AB}G_{BA})^D + \mathcal{O}(\delta^3). \end{aligned}$$

Combing all the results, we get:

$$\begin{pmatrix} \Delta & \delta\Delta G_{AB}U \\ \delta U G_{BA}\Delta & \delta^2 U G_{BA}\Delta G_{AB}U \end{pmatrix} = \delta^2 \begin{pmatrix} (G_{AB}G_{BA})^D & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} + \mathcal{O}(\delta^3),$$

which completes the proof. □

Proof of Proposition 2. Comparing (11) and (12) gives us the relation between player i's equilibrium contribution under the two scenarios:

$$\begin{aligned} x_i^L &= \frac{1-\delta^2 \left\langle \beta_i,\ (\mathbf{I}-\delta G_{-i})^{-1}\beta_i \right\rangle}{1-2\delta^2 \left\langle \beta_i,\ (\mathbf{I}-\delta G_{-i})^{-1}\beta_i \right\rangle} x_i^N \\ &= \frac{1-(1-\frac{1}{m_{ii}})}{1-2(1-\frac{1}{m_{ii}})} x_i^N = \frac{1}{2-m_{ii}} x_i^N = \frac{1}{2-m_{ii}} b_i(G,\delta,\alpha). \end{aligned} \tag{19}$$

Here we have used the fact that $\delta^2 \left\langle \beta_i,\ (\mathbf{I}-\delta G_{-i})^{-1}\beta_i \right\rangle = 1-\frac{1}{m_{ii}}$ by (13).

The next step is to study the impact of i's contribution on other players. Note that the rest of the group ($N\backslash\{i\}$) play their best-responses in both scenarios, i.e.,

$$x^*_{-i}(x_i) = \mathbf{b}(G_{-i}, \delta, \alpha_{-i} + \delta\beta_i x_i) = (\mathbf{I} - \delta G_{-i})^{-1} (\alpha_{-i} + \delta\beta_i x_i). \tag{20}$$

If the player i's contribution changes by Δx_i, the incremental contributions of players in $N\backslash\{i\}$ are given by:

$$\Delta x^*_{-i}(x_i) = (\mathbf{I} - \delta G_{-i})^{-1} \delta\beta_i \Delta x_i \tag{21}$$

According to (13), we obtain the following expression: $\Delta x^*_j(x_i) = \frac{m_{ij}}{m_{ii}}\Delta x_i$, $\forall j \neq i$. Therefore, the change of aggregate contribution due to Δx_i is

$$(1 + \sum_{j\neq i} \frac{m_{ij}}{m_{ii}})\Delta x_i = \frac{\sum_{k=1}^N m_{ik}}{m_{ii}}\Delta x_i = \frac{b_i(G, \delta, 1)}{m_{ii}}\Delta x_i. \tag{22}$$

Here we use the fact that $\sum_{k=1}^N m_{ik} = b_i(G, \delta, 1)$.

Notice that player i's contribution increases from $x_i^N = b_i(G, \delta, \alpha)$ to $x_i^L = \frac{1}{2-m_{ii}}x_i^N = \frac{1}{2-m_{ii}}b_i(G, \delta, \alpha)$. Thus, if i is selected as the leader, the change of aggregate contributions is

$$\frac{b_i(G, \delta, 1)}{m_{ii}}\left(\frac{1}{2 - m_{ii}}b_i(G, \delta, \alpha) - b_i(G, \delta, \alpha)\right) = \frac{(m_{ii} - 1)}{(2 - m_{ii})}\frac{b_i(G, \delta, 1)}{m_{ii}}b_i(G, \delta, \alpha). \tag{23}$$

This then leads to the L-index specified in the proposition. □

Proof of Corollary 2. Using (22), the change of aggregate contribution is

$$\frac{b_i(G, \delta, 1)}{m_{ii}}\Delta x_i = -\frac{b_i(G, \delta, 1)}{m_{ii}}b_i(G, \delta, \alpha).$$

□

Proof of Proposition 3. We again start with the benchmark case. In this simultaneous-move game, the equilibrium contributions are given by

$$\mathbf{x}^N = [I - \delta G]^{-1}\alpha = [(1 + \delta)\mathbf{I} - \delta\mathbf{J}]^{-1}\alpha = \frac{1}{1 + \delta}[\mathbf{I} - \frac{\delta}{1 - (N - 1)\delta}\mathbf{J}]\alpha.$$

The last equality follows from Lemma 3. For each player i, the equilibrium contribution profile is given by:

$$x_i^N = \frac{\alpha_i + \frac{\delta}{1-(N-1)\delta}\sum_{j=1}^N \alpha_j}{1 + \delta},\ \forall i = 1, 2, \cdots, N,$$

and the aggregate contribution is

$$\sum_{i=j}^N x_j = \frac{\sum_{j=1}^N \alpha_j}{1 - (N - 1)\delta}. \tag{24}$$

We can interpret the factor $(N - 1)\delta$ as the amplification factor. An implicit assumption is that δ is sufficiently small such that $(N - 1)\delta < 1$.

Next, we examine the two-stage game. Let L denote the set of first-movers (leaders), and $N\backslash L$ represents the rest of players (followers). First, let us study the subgame in the second stage. This can be done similarly as the subgraph is still complete but with only $N-L$ nodes (this is the primary advantage of working with a complete graph). In addition, we only care about the sum of the contributions of the players in $N\backslash L$. Given the contributions $x_i, i \in L$, the subgame is a network game, where parameters $\{\alpha_j, j \in N\backslash L\}$'s are modified as $\alpha_j \to \alpha_j + \delta(\sum_{i\in L} x_i)$. Hence, the aggregate equilibrium contribution is (from (24)):

$$\sum_{j\in N\backslash L} x_j(x_1, x_2, \cdots, x_L) = \frac{\sum_{j\in N\backslash L}\alpha_j + \delta(N-L)(\sum_{i\in L} x_i)}{1-(N-L-1)\delta}. \tag{25}$$

Anticipating the feedback in the second stage, each player i in L plays another network game with the following payoff:

$$\begin{aligned} u_i &= \alpha_i - \frac{1}{2}x_i^2 + \delta x_i \left(\sum_{j\in L, j\neq i} x_j + \sum_{j\in N\backslash L} x_j(x_1, x_2, \cdots, x_L) \right) \\ &= \alpha_i - \frac{1}{2}x_i^2 + \delta x_i \left(\sum_{j\in L, j\neq i} x_j + \frac{\sum_{j\in N\backslash L}\alpha_j + \delta(N-L)(\sum_{i\in L} x_i)}{1-(N-L-1)\delta} \right). \end{aligned}$$

The equilibrium is obtained by taking the first-order conditions:

$$\alpha_i - x_i + \delta \left(\sum_{j\in L, j\neq i} x_j + \frac{\sum_{j\in N\backslash L}\alpha_j + \delta(N-L)(\sum_{i\in L} x_i)}{1-(N-L-1)\delta} \right) + \delta x_i \frac{\delta(N-L)}{1-(N-L-1)\delta} = 0,\ i \in L.$$

Summing over $i \in L$ yields:

$$\begin{aligned} &\textstyle\sum_{i\in L}\alpha_i - \sum_{i\in L} x_i + \frac{\delta^2(N-L)}{1-(N-L-1)\delta}\sum_{i\in L} x_i \\ &+\delta\left((L-1)\textstyle\sum_{i\in L} x_i + L\frac{\sum_{j\in N\backslash L}\alpha_j + \delta(N-L)(\sum_{i\in L} x_i)}{1-(N-L-1)\delta} \right) \end{aligned} = 0.$$

Therefore, we can find the aggregate contribution in L:

$$\sum_{i\in L} x_i = \frac{\sum_{i\in L}\alpha_i + \frac{\delta L}{1-\delta(N-L-1)}\sum_{j\in N\backslash L}\alpha_j}{1-\delta(L-1) - \frac{\delta^2 L(N-L)}{1-(N-L-1)\delta} - \frac{\delta^2(N-L)}{1-(N-L-1)\delta}}. \tag{26}$$

Plugging this into (25), we can find the aggregate contribution in this two-stage game.

From (25) and (26), we have:

$$\sum_{i\in L} x_i + \sum_{j\in N\backslash L} x_j = \frac{\sum_{j\in N\backslash L}\alpha_j}{1-(N-L-1)\delta} + \frac{(1+\delta)}{1-(N-L-1)\delta}\frac{(1-(N-L-1)\delta)\sum_{i\in L}\alpha_i + \delta L\sum_{j\in N\backslash L}\alpha_j}{1-\delta(N-2)-\delta^2(2N-L-1)}.$$

□

Proof of Corollary 3. It suffices to show that in (15), the coefficient of the term $\sum_{i\in L}\alpha_i$ is greater than that of $\sum_{j\in N\setminus L}\alpha_j$, as the sum of these terms is fixed. To this end, it suffices to show that:

$$\begin{aligned}
& \frac{-1}{1-(N-L-1)\delta}+\frac{1+\delta}{1-(N-L-1)\delta}\times\frac{(1-(N-L-1)\delta)-\delta L}{1-\delta(N-2)-\delta^2(2N-L-1)}>0 \\
\iff\ & (1+\delta)(1-(N-1)\delta)>1-\delta(N-2)-\delta^2(2N-L-1) \\
\iff\ & 1-\delta(N-2)-(N-1)\delta^2>1-\delta(N-2)-\delta^2(2N-L-1) \\
\iff\ & 2N-L-1>N-1 \\
\iff\ & N>L,
\end{aligned}$$

where the last inequality obviously holds. □

Proof of Proposition 4. This follows from tedious but straightforward calculations. □

Proof of Proposition 5. Let us start with the simultaneous-move benchmark. To characterize the equilibrium, we only need to find the inverse matrix of $1-\delta G$. We observe that the payoff functions for nodes in each group are:

$$\begin{cases} u_i &= \alpha_i x_i - \frac{1}{2}x_i^2 + \delta x_i(\sum_{j\in N} y_j), i\in M, \\ u_j &= \beta_j y_j - \frac{1}{2}y_j^2 + \delta y_j(\sum_{i\in M} x_i), j\in N, \end{cases}$$

and the best responses are:

$$\begin{cases} x_i &= \alpha_i + \delta(\sum_{j\in N} y_j), i\in M, \\ y_j &= \beta_j + \delta(\sum_{i\in M} x_i), j\in N. \end{cases} \tag{27}$$

Thus, summing over them yields:

$$\begin{cases} \sum_{i\in M} x_i &= \sum_{i\in M}\alpha_i + M\delta(\sum_{j\in N} y_j), \\ \sum_{j\in N} y_j &= \sum_{j\in N}\beta_j + N\delta(\sum_{i\in M} x_i). \end{cases} \tag{28}$$

Equivalently,

$$\begin{pmatrix}\sum_{i\in M} x_i \\ \sum_{j\in N} y_j\end{pmatrix} = \frac{1}{1-MN\delta^2}\begin{pmatrix}\sum_{i\in M}\alpha_i + M\delta\sum_{j\in N}\beta_j \\ \sum_{j\in N}\beta_j + N\delta\sum_{i\in M}\alpha_i\end{pmatrix}.$$

Theorem 2 suggests that to maximize the aggregate contribution, we shall partition the players based on the bipartite structure (because this maximizes the number of cross links). Therefore, there are only two candidates: either group M moves first (and becomes the leader group), or group M moves in the second stage. We shall start with the former scenario, as the latter is simply the mirror image.

In this case, in the second stage the best responses of y_j are given in (27) as the subgraph in N is empty. Consequently, the aggregate contribution in group N is:

$$\sum_{j\in N} y_j(\mathbf{x_M}) = \sum_{j\in N}\beta_j + N\delta(\sum_{i\in M} x_i(\mathbf{x_M})). \tag{29}$$

Given the feedback from the followers, the payoff of a leader i in M is

$$u_i(\mathbf{x_M}) = \alpha_i x_i - \frac{1}{2}x_i^2 + \delta x_i \left(\sum_{j \in N} y_j(\mathbf{x_M}) \right) = \alpha_i x_i - \frac{1}{2}x_i^2 + \delta x_i \left(\sum_{j \in N} \beta_j + N\delta(\sum_{i \in M} x_i) \right).$$

The first-order conditions are

$$\alpha_i - x_i + \delta \left(\sum_{j \in N} \beta_j + N\delta(\sum_{i \in M} x_i) \right) + N\delta^2 x_i = 0, \ \forall i \in M.$$

This system of linear equations can be solved via simple Gaussian eliminations. Similarly, we can derive the equilibrium contributions when group N moves first.

For our purpose, we only need to compare the aggregate contribution. To this end, taking summation over $i \in M$ yields:

$$\sum_{i \in M} \alpha_i - (1 - \delta^2 N) \sum_{i \in M} x_i + M\delta \left(\sum_{j \in N} \beta_j + N\delta(\sum_{i \in M} x_i) \right) = 0 \tag{30}$$

$$\iff \sum_{i \in M} x_i = \frac{\sum_{i \in M} \alpha_i + M\delta \sum_{j \in N} \beta_j}{1 - \delta^2(MN + N)}.$$

Hence, the aggregate contribution in N is just

$$\sum_{j \in N} y_j = \sum_{j \in N} \beta_j + N\delta(\sum_{i \in M} x_i) = \sum_{j \in N} \beta_j + N\delta \left(\frac{\sum_{i \in M} \alpha_i + M\delta \sum_{j \in N} \beta_j}{1 - \delta^2(MN + N)} \right).$$

Let Δ_X denote the *increment* of aggregate contribution from group M, and similarly for Δ_Y from group N. As the nodes in N are playing best responses, by (29) we have $\Delta_Y = N\delta\Delta_X$. Now we can compare the increment of aggregate contribution:

$$\begin{aligned} \Delta Y + \Delta X &= (1 + N\delta)\Delta X = (1 + N\delta) \left(\frac{\sum_{i \in M} \alpha_i + M\delta \sum_{j \in N} \beta_j}{1 - \delta^2(MN + N)} - \frac{\sum_{i \in M} \alpha_i + M\delta \sum_{j \in N} \beta_j}{1 - \delta^2 MN} \right) \\ &= \frac{N\delta^2(1 + N\delta)(\sum_{i \in M} \alpha_i + M\delta \sum_{j \in N} \beta_j)}{(1 - \delta^2(MN + N))(1 - \delta^2 MN)}. \end{aligned}$$

Recalling that $\bar{\alpha} = \frac{\sum_{i \in M} \alpha_i}{M}$ and $\bar{\beta} = \frac{\sum_{j \in N} \beta_j}{N}$, we obtain that

$$\Delta = \Delta Y + \Delta X = \frac{MN\delta^2}{1 - MN\delta^2} \times \frac{(1 + N\delta)(\bar{\alpha} + N\delta\bar{\beta})}{(1 - \delta^2(MN + N))}.$$

Similarly, we can derive the increment of aggregate contribution when group N moves first:

$$\Delta' = \frac{MN\delta^2}{1 - MN\delta^2} \times \frac{(1 + M\delta)(\bar{\beta} + M\delta\bar{\alpha})}{(1 - \delta^2(MN + M))}.$$

The comparison immediately leads to the proposition. Note that to make these numbers meaningful (for stability considerations), we assume that $\delta < 1/\sqrt{MN + \max(M,N)}$. □

Proof of Corollary 4. If $M = N$, then (16) is equivalent to

$$\bar{\alpha} + N\delta\bar{\beta} > \bar{\beta} + M\delta\bar{\alpha} = \bar{\beta} + N\delta\bar{\alpha} \Leftrightarrow (1 - N\delta)(\bar{\alpha} - \bar{\beta}) > 0.$$

The result follows because $1 - N\delta > 0$. □

Proof of Corollary 5. To this end, it suffices to show that:

$$N > M \Longleftrightarrow \frac{(1+N\delta)^2}{1 - \delta^2 MN - \delta^2 N} > \frac{(1+M\delta)^2}{1 - \delta^2 MN - \delta^2 M}.$$

This obviously holds because $(1+N\delta)^2 > (1+M\delta)^2$ and $1 - \delta^2 MN - \delta^2 N < 1 - \delta^2 MN - \delta^2 M$. □

Proof of Theorem 3. We prove this theorem by induction on the step of the sequence. Suppose that $k = 1$. In this case, it is simply the simultaneous-move game, and we know that $x(N, G, \delta, \alpha) = [I - \delta G]^{-1}\alpha$. Thus, the matrix T is equal to $[I - \delta G]^{-1}$, which is symmetric if G is.

Now suppose that the results hold for any sequence with steps smaller than k. Consider a sequence $\mathcal{S} = (P_1, P_2, \cdots, P_k)$ with step k. Let $A = P_1$ denote the nodes in the first step, let $B = N \backslash A = \cup_{2 \le i \le k} P_i$, and $\mathcal{S}' = (P_2, \cdots, P_k)$ as the subsequence. For each $\mathbf{x}_A$, the subgame is again a network game with sequence $\mathcal{S}'$, nodes B with adjacency matrix G_{BB}, and parameter $\alpha'_B = \alpha_B + \delta G_{BA} x_A$. Note that the step of sequence $\mathcal{S}'$ is $k - 1$. Therefore, by induction the equilibrium in the subgame is given by $\mathbf{x}_B = U\alpha'_B = U(\alpha_B + \delta G_{BA}\mathbf{x}_A)$ for a matrix U. Moreover, the matrix U is symmetry if G_{BB} is by induction.

Given the feedback in the subgame, the nodes in A choose x_i to maximize:

$$u_i = \alpha_i x_i - \frac{1}{2}x_i^2 + \delta x_i \left(\sum_{j \in A} g_{ij} x_j + \sum_{j \in B} g_{ij} x_j(\mathbf{x}_A) \right), \quad \forall i \in A.$$

The rest of the analysis is similar to the case with two stages in the proof of Theorem 1, except that the matrix U is given by the induction step.

In all, we get:

$$Z(\mathcal{S}) = \begin{pmatrix} \left[1 - \delta(T + T^D)\right]^{-1} & \delta\left[1 - \delta(T + T^D)\right]^{-1} G_{AB} U \\ \delta U G_{BA}\left[1 - \delta(T + T^D)\right]^{-1} & U + \delta^2 U G_{BA}\left[1 - \delta(T + T^D)\right]^{-1} G_{AB} U \end{pmatrix} \tag{31}$$

where U is given by the induction step, and $T = G_{AA} + \delta G_{AB} U G_{BA}$. □

Proof of Proposition 6. The proof is composed of several steps. In Step 1, we show that if the response function in the subgame is stronger, the equilibrium in the first stage is also stronger. Step 2 establishes the equivalence between the general sequence and a specific two-stage game regarding the equilibrium outcomes. Finally, Step 3 shows that the equilibrium contribution profile under two-stage (A_1, A_2) is higher than the single-stage game for group A.

Step 1.

Without loss of generality, we assume that $r = 1$, i.e., we just split the first set P_1. This can be observed in the following lemma. Define two matrices

$$Z = \begin{pmatrix} \left[1-\delta(T+T^D)\right]^{-1} & \delta\left[1-\delta(T+T^D)\right]^{-1} G_{AB}U \\ \delta U G_{BA}\left[1-\delta(T+T^D)\right]^{-1} & U+\delta^2 U G_{BA}\left[1-\delta(T+T^D)\right]^{-1} G_{AB}U \end{pmatrix}$$

and

$$\tilde{Z} = \begin{pmatrix} \left[1-\delta(\tilde{T}+\tilde{T}^D)\right]^{-1} & \delta\left[1-\delta(\tilde{T}+\tilde{T}^D)\right]^{-1} G_{AB}\tilde{U} \\ \delta \tilde{U} G_{BA}\left[1-\delta(\tilde{T}+\tilde{T}^D)\right]^{-1} & \tilde{U}+\delta^2 \tilde{U} G_{BA}\left[1-\delta(\tilde{T}+\tilde{T}^D)\right]^{-1} G_{AB}\tilde{U} \end{pmatrix},$$

where $T = G_{AA} + \delta G_{AB} U G_{BA}$, $\tilde{T} = G_{AA} + \delta G_{AB}\tilde{U} G_{BA}$.

We claim that if $\tilde{U} \succeq U$, then $\tilde{Z} \succeq Z$. To prove this, observe that if $\tilde{U} - U \succeq \mathbf{0}$, $\tilde{T} - T = \delta G_{AB}(\tilde{U} - U)G_{BA} \succeq \mathbf{0}$. Therefore, $\tilde{T} + \tilde{T}^D - (T + T^D) \succeq \mathbf{0}$, and hence

$$\begin{aligned} &\left[1-\delta(\tilde{T}+\tilde{T}^D)\right]^{-1} - \left[1-\delta(T+T^D)\right]^{-1} \\ =& \left[1-\delta(T+T^D)\right]^{-1}(\tilde{T}+\tilde{T}^D-(T+T^D))\left[1-\delta(\tilde{T}+\tilde{T}^D)\right]^{-1} \succeq \mathbf{0}. \end{aligned}$$

In other words,

$$\left[1-\delta(\tilde{T}+\tilde{T}^D)\right]^{-1} \succeq \left[1-\delta(T+T^D)\right]^{-1}.$$

Combined with $\tilde{U} \succeq U$, we can show that

$$\delta\left[1-\delta(\tilde{T}+\tilde{T}^D)\right]^{-1} G_{AB}\tilde{U} \succeq \delta\left[1-\delta(T+T^D)\right]^{-1} G_{AB}U.$$

The comparisons for the rest terms are similar and hence are omitted.

The above argument shows that if the response function in the subgame is stronger, the equilibrium in the first stage is also stronger. This positive feedback then makes the whole response function stronger. Therefore, by the induction argument, we only need to check the result when we just split the first set P_1 into two sets, i.e., the case with $r = 1$.

Step 2.

If $r = 1$, without loss of generality we can assume that $k = 2$. This is intuitive since from the perspective of the nodes in P_1, they only care about the response functions of nodes in $P_2 \cup P_2 \cup \cdots \cup P_k$ as a whole. Consequently, we can simply reduce their response functions to the equilibrium response matrix U in the subsequence $(P_2 \cup P_2 \cup \cdots \cup P_k)$. The existence of U is given in the proof of Theorem 3.

Now suppose that $r = 1$ and $k = 2$. In this case, let $\mathcal{S} = (A, B)$ and $\tilde{\mathcal{S}} = (A_1, A_2, B)$. For each $\mathbf{x}_A$, the equilibrium in the subgame is given by $\mathbf{x}_B = U\alpha'_B = U(\alpha_B + \delta G_{BA}\mathbf{x}_A)$ for matrix U. Taking the feedback into account, the nodes in A are playing a network game with payoffs:

$$u_i = \alpha_i x_i - \frac{1}{2}x_i^2 + \delta x_i \left(\sum_{j\in A} g_{ij}x_j + \sum_{j\in B} g_{ij}x_j(\mathbf{x}_A)\right), \quad \forall i \in A,$$

where $x_j(\mathbf{x}_A) = \sum_{k\in B} U_{jk}(\alpha_k + \delta \sum_{l\in A} g_{kl}x_l)$ by (4). Plugging in and switching dummy variables if necessary, we obtain that

$$\begin{aligned} u_i &= \alpha_i x_i - \frac{1}{2}x_i^2 + \delta x_i \left(\sum_{j\in A} g_{ij}x_j + \sum_{j\in B} g_{ij} \left(\sum_{k\in B} U_{jk}(\alpha_k + \delta \sum_{l\in A} g_{kl}x_l) \right) \right) \\ &= \left(\alpha_i + \delta \sum_{j\in B}\sum_{k\in B} g_{ij}U_{jk}\alpha_k \right) x_i - \frac{1}{2}x_i^2 \\ &\quad + \delta x_i \left(\sum_{j\in A} g_{ij}x_j + \delta \sum_{j\in B}\sum_{k\in B}\sum_{l\in A} g_{ij}U_{jk}g_{kl}x_l \right). \end{aligned}$$

In other words, the nodes in A play a new network game with parameters:

$$\tilde{\alpha}_A = \alpha_A + \delta G_{AB}U\alpha_B, \text{ and } \tilde{G}_{AA} = G_{AA} + \delta G_{AB}UG_{BA}.$$

$\mathcal{S} = (A, B)$ corresponds to the simultaneous-move game in the new network game on A, and the sequence $\tilde{\mathcal{S}} = (A_1, A_2, B)$ maps to the two-stage (A_1, A_2) game in the new network game on A. Hence, we can reduce the proof to the case with $k = 1$.

Step 3.

The next step is to show that the equilibrium contribution profile under two-stage (A_1, A_2) is higher than the single-stage game for group A. If this is true, then the equilibrium of the nodes in B is also higher as their response functions $\mathbf{x}_B = U\alpha'_B = U(\alpha_B + \delta G_{BA}\mathbf{x}_A)$ are increasing in $\mathbf{x}_A$. All in all, we can reduce everything into the case with $k = 1$.

If $k = 1$, the basic argument is given in Proposition 1. However, after reducing everything into the case $k = 1$ on the nodes in A, the "effective" influence matrix

$$\tilde{G}_{AA} = G_{AA} + \delta G_{AB}UG_{BA}$$

may have nonzero diagonal entries. Therefore, we have to use the analogous part to establish the result. Combing these steps, we then complete the proof. □

Proof of Corollary 6. This follows immediately from Proposition 6. □

Proof of Theorem 4. We again prove it by induction on the step k. For $k = 1$, the result is trivial. For $k = 2$, it is shown in Theorem 2.

For genera k, define sequence $\mathcal{S} = (P_1, P_2, \cdots, P_k)$. Furthermore, $A = P_1$ denotes the nodes in the first step, and $B = N\backslash A$. Thus,

$$Z(\mathcal{S}) = \begin{pmatrix} \left[1 - \delta(T + T^D)\right]^{-1} & \delta\left[1 - \delta(T + T^D)\right]^{-1} G_{AB}U \\ \delta UG_{BA}\left[1 - \delta(T + T^D)\right]^{-1} & U + \delta^2 UG_{BA}\left[1 - \delta(T + T^D)\right]^{-1} G_{AB}U \end{pmatrix},$$

by (31), where U is given by the induction step, and $T = G_{AA} + \delta G_{AB}UG_{BA}$. Also, by (17), we have

$$[1 - \delta G]^{-1} = \begin{pmatrix} \left[1 - \delta\hat{T}\right]^{-1} & \delta\left[1 - \delta\hat{T}\right]^{-1} G_{AB}\hat{U} \\ \delta\hat{U}G_{BA}\left[1 - \delta\hat{T}\right]^{-1} & \hat{U} + \delta^2\hat{U}G_{BA}\left[1 - \delta\hat{T}\right]^{-1} G_{AB}\hat{U} \end{pmatrix},$$

with $\hat{U} = [1 - \delta G_{BB}]^{-1}$, and $\hat{T} = G_{AA} + \delta G_{AB}\hat{U}G_{BA}$. By induction, we obtain that

$$U - \hat{U} = \delta^2 \begin{pmatrix} \left(\sum_{j=3}^{k} G_{P_2P_j}G_{P_jP_2}\right)^D & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \left(G_{P_{k-1}P_k}G_{P_kP_{k-1}}\right)^D & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} + \mathcal{O}(\delta^3).$$

To compare these two matrices, we first show that:

$$\begin{aligned} (T + T^D) - \hat{T} &= G_{AA} + \delta G_{AB}UG_{BA} + (G_{AA} + \delta G_{AB}UG_{BA})^D - (G_{AA} + \delta G_{AB}\hat{U}G_{BA}) \\ &= \delta G_{AB}(U - \hat{U})G_{BA} + \delta(G_{AB}UG_{BA})^D \\ &= \delta(G_{AB}G_{BA})^D + \mathcal{O}(\delta^2). \end{aligned}$$

Because $G_{AA}^D = \mathbf{0}$ and $(U - \hat{U}) = \mathcal{O}(\delta^2)$, $U = I + \mathcal{O}(\delta)$. Therefore:

$$\begin{aligned} \left[1 - \delta(T + T^D)\right]^{-1} - \left[1 - \delta\hat{T}\right]^{-1} &= \left[1 - \delta\hat{T}\right]^{-1} \delta\left((T + T^D) - \hat{T}\right)\left[1 - \delta(T + T^D)\right]^{-1} \\ &= \delta^2(G_{AB}G_{BA})^D + \mathcal{O}(\delta^3). \end{aligned}$$

Similarly, we can show that:

$$\begin{aligned} \delta\left[1 - \delta(T + T^D)\right]^{-1} G_{AB}U - \delta\left[1 - \delta\hat{T}\right]^{-1} G_{AB}\hat{U} &= \mathcal{O}(\delta^3), \\ \delta UG_{BA}\left[1 - \delta(T + T^D)\right]^{-1} - \delta\hat{U}G_{BA}\left[1 - \delta\hat{T}\right]^{-1} &= \mathcal{O}(\delta^3), \end{aligned}$$

and

$$\begin{aligned} & U + \delta^2 UG_{BA}\left[1 - \delta(T + T^D)\right]^{-1} G_{AB}U - \left(\hat{U} + \delta^2\hat{U}G_{BA}\left[1 - \delta\hat{T}\right]^{-1} G_{AB}\hat{U}\right) \\ = \; & U - \hat{U} + \mathcal{O}(\delta^3) \\ = \; & \delta^2 \begin{pmatrix} \left(\sum_{j=3}^{k} G_{P_2P_j}G_{P_jP_2}\right)^D & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \left(G_{P_{k-1}P_k}G_{P_kP_{k-1}}\right)^D & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} + \mathcal{O}(\delta^3). \end{aligned}$$

Plugging in these terms, we establish the results. □

Proof of Corollary 7. By Proposition 6, $Z(\mathcal{S}') \succeq Z(\mathcal{S})$. Therefore,

$$\frac{Z(\mathcal{S}') - Z(\emptyset)}{\delta^2} \succeq \frac{Z(\mathcal{S}') - Z(\emptyset)}{\delta^2}.$$

Taking the limits as $\delta \to 0$, the result follows directly from Theorem 4. □

Proof of Corollary 8. Under these conditions, if $\mathcal{S}$ is a chain, we can assume $\mathcal{S} = (1, \cdots, N)$. Then the term $L(\mathcal{S})$ simply counts the number of 1s on the upper diagonal part of G. In other words, $L(\mathcal{S}) = \sum_{j>i} g_{ij} = \frac{1}{2}\sum_{ij} g_{ij}$, where we use the fact that $g_{ii} = 0$, $\forall i$. □

Proof of Lemma 2. First, the gradient of $W(\mathbf{x})$ is $\nabla W(\mathbf{x}) = \alpha - \mathbf{x} + \delta(G + G')\mathbf{x}$, which equals $\delta G'\mathbf{x}^N$ at $\mathbf{x}(\emptyset) = \mathbf{x}^N = [1 - \delta G]^{-1}\alpha$. Note that $\mathbf{x}(\mathcal{S}) - \mathbf{x}(\emptyset) = \delta^2\Lambda\alpha + \mathcal{O}(\delta^3)$ from (9). Therefore,

$$\begin{aligned}\Delta W &= W(\mathbf{x}(\mathcal{S})) - W(\mathbf{x}(\emptyset)) \\ &= \langle \delta G'\mathbf{x}^N, \mathbf{x}(\mathcal{S}) - \mathbf{x}(\emptyset)\rangle + \mathcal{O}(\delta^4) \\ &= \langle \delta G'\mathbf{x}^N, \delta^2\Lambda\alpha + \mathcal{O}(\delta^3)\rangle + \mathcal{O}(\delta^4) \\ &= \delta^3\alpha' G\Lambda\alpha + \mathcal{O}(\delta^4).\end{aligned}$$

Here we use the fact that $\mathbf{x}^N = \alpha + \mathcal{O}(\delta)$. □

Proof of Proposition 7. This is a direct consequence of Lemma 2. □

References

Acemoglu, D., K. Bimpikis, and A. Ozdaglar (2013). Dynamics of information exchange in endogenous social networks. forthcoming in Theoretical Economics.

Acemoglu, D., M. A. Dahleh, I. Lobel, and A. Ozdaglar (2011). Bayesian learning in social networks. *The Review of Economic Studies 78*(4), 1201–1236.

Alonso, R., W. Dessein, and N. Matouschek (2008). When does coordination require centralization? *The American Economic Review 98*(1), 145–179.

Amir, R. and A. Stepanova (2006). Second-mover advantage and price leadership in bertrand duopoly. *Games and Economic Behavior 55*(1), 1–20.

Aoki, M. (1986). Horizontal vs. vertical information structure of the firm. *The American Economic Review 76*(5), 971–983.

Ballester, C. and A. Calvó-Armengol (2010). Interactions with hidden complementarities. *Regional Science and Urban Economics 40*(6), 397–406.

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

Bandiera, O., I. Barankay, and I. Rasul (2004). Relative and absolute incentives: Evidence on worker productivity. Working paper, London School of Economics.

Barahona, F. (1983). The max-cut problem on graphs not contractible to K_5. *Operations Research Letters 2*(3), 107–111.

Beckmann, K. (2005). Tax competition and strategic complementarity. Working paper, Andrassy University.

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

Bulow, J., J. Geanakoplos, and P. Klemperer (1985). Multimarket oligopoly: Strategic substitutes and complements. *Journal of Political Economy 93*(3), 488–511.

Calvó-Armengol, A. and J. d. M. Beltran (2009). Information gathering in organizations: equilibrium, welfare, and optimal network structure. *Journal of the European Economic Association 7*(1), 116–161.

Calvo-Armengol, A. and M. O. Jackson (2004). The effects of social networks on employment and inequality. *The American Economic Review 94*(3), 426–454.

Candogan, O., K. Bimpikis, and A. Ozdaglar (2012). Optimal pricing in networks with externalities. *Operations Research 60*(4), 883–905.

Corbo, J., A. Calvó-Armengol, and D. Parkes (2006). A study of Nash equilibrium in contribution games for peer-to-peer networks. *Operating Systems Review 40*(3), 61–66.

Crémer, J. (1980). A partial theory of the optimal organization of a bureaucracy. *Bell Journal of Economics 11*(2), 683–693.

Economides, N. (1996). The economics of networks. *International Journal of Industrial Organization 14*(6), 673–699.

Falk, A. and A. Ichino (2006). Clean evidence on peer effects. *Journal of Labor Economics 24*(1), 39–57.

Farrell, J. and G. Saloner (1986). Installed base and compatibility: Innovation, product preannouncement, and predation. *American Economic Review 76*, 940–955.

Fudenberg, D. and J. Tirole (1991). *Game Theory*. The MIT Press, Massachusetts.

Goemans, M. and D. Williamson (1995). Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM 42*(6), 1115–1145.

Grötschel, M. and W. Pulleyblank (1981). Weakly bipartite graphs and the max-cut problem. *Operations Research Letters 1*(1), 23–27.

Hadlock, F. (1975). Finding a maximum cut of a planar graph in polynomial time. *SIAM Journal on Computing 4*, 221–225.

Hoppe, H. (2000). Second-mover advantages in the strategic adoption of new technology under uncertainty. *International Journal of Industrial Organization 18*(2), 315–338.

Hoppe, H. and U. Lehmann-Grube (2005). Innovation timing games: a general framework with applications. *Journal of Economic Theory 121*(1), 30–50.

Huck, S. and P. Biel (2012). Endogenous leadership in teams. Working paper, University of College London.

Jackson, M. (2008). *Social and Economic Networks*. Princeton University Press, Princeton, NJ.

Katz, M. and C. Shapiro (1985). Network externalities, competition, and compatibility. *The American Economic Review 75*(3), 424–440.

Kerin, R., P. Varadarajan, and R. Peterson (1992). First-mover advantage: A synthesis, conceptual framework, and research propositions. *The Journal of Marketing 56*, 33–52.

Marschak, J. and R. Radner (1972). *Economic theory of teams*. Yale University Press (New Haven).

Milgrom, P. R. and J. Roberts (1992). *Economics, organization and management*. Prentice-Hall International.

Radner, R. (1962). Team decision problems. *The Annals of Mathematical Statistics 33*(3), 857–881.

Rohlfs, J. (1974). A theory of interdependent demand for a communications service. *The Bell Journal of Economics and Management Science 5*(1), 16–37.

Von Stengel, B. (2010). Follower payoffs in symmetric duopoly games. *Games and Economic Behavior 69*(2), 512–516.