NET Institute*

www.NETinst.org

Targeted Information Release in Social Networks

Junjie Zhou
Shanghai University of Finance and Economics

Ying-Ju Chen
University of California at Berkeley

Targeted information release in social networks[*]

Junjie Zhou[†]   Ying-Ju Chen[‡]

September 2013

Abstract

As a common practice, various firms initially make information and access to their products/services scarce within a social network; identifying influential players that facilitate information dissemination emerges as a pivotal step for their success. In this paper, we tackle this problem using a stylized model that features payoff externalities and local network effects, and the network designer is allowed to release information to only a subset of players (leaders); these targeted players make their contributions first and the rest followers move subsequently after observing the leaders' decisions. In the presence of incomplete information, the signaling incentive drives the optimal selection of leaders and can lead to a first-order materialistic effect on the equilibrium outcomes. We propose a novel index for the key leader selection (i.e., a single player to provide information to) that can be substantially different from the key player index in Ballester et al. (2006) and the key leader index with complete information proposed in Zhou and Chen (2013). We also show that in undirected graphs, the optimal leader group identified in Zhou and Chen (2013) is exactly the optimal follower group when signaling is present. The pecking order in complete graphs suggests that the leader

[†]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.

should be selected by the ascending order of intrinsic valuations. We also examine the out-tree hierarchical structure that describes a typical economic organization. The key leader turns out to be the one that stays in the middle, and it is not necessarily exactly the central player in the network.

1 Introduction

In the past decade, we have witnessed the explosive growth of social networks, as exemplified by the success of Facebook, Linkedin, Twitter, and various online game producers. One common feature in these social networks is the presence of strong *network effects*, i.e., a user's (player's) "consumption utility" depends not only on how much time and effort she invests in the network, but also on the behaviors of other users. These interactions suggest that any marketing strategy cannot be determined in isolation, because its effect on one individual can be propagated to the close friends/ neighbors, the neighbors of neighbors, and ultimately spread out through the entire network. Second, as these companies frequently launch novel games and services that go beyond the users' original imagination, users may have *limited knowledge* about the level of satisfaction while making their purchasing or subscription decisions (Jackson (2008)). This inherent uncertainty constitutes a critical barrier that prevents users from quickly adopting new services or technology. To overcome this issue, managing the *information dissemination* becomes a critical business strategy for these social network designers when marketing a product, promoting an idea, inducing innovation, or disciplining the social behavior (Acemoglu et al. (2011), Acemoglu et al. (2013), and Bala and Goyal (1998)).

While intuitively it is sensible to make the information as transparent as possible, in practice various firms intentionally *restrict* the users that have access to the quality of new services/ games, even though mass information spreading is not technically infeasible. This targeted advertising approach turns out to be extremely popular amongst the practitioners. For example, Ty's initial launch of bean-stuffed toys is made available only to an exclusive set of consumers. When the European music site Spotify was first launched in the United States, free versions were accessible by invitation only (Dye (2000)). Another example is Google+, a social network site that can only be joined through invitations by friends who have already been in the circle. At the early stage of Facebook, only individuals with email addresses from a restricted set of schools are eligible to join. Consequently, *identifying influential users* (also known as "opinion leaders," "trend setters," "fashion influentials," and "mavens") in the networks becomes a pivotal step and is nowadays an active marketing area of research. There are some e-commerce sites and startup companies that focus exclusively on identifying influentials (such as Klout.com), and the advance of big data analytics makes it

possible to conduct large-scale identifications for various kinds of social networks. See Section 2 for a detailed literature review of academic papers. The targeted advertising approach is particularly attractive when spending a limited advertising resource can generate a huge momentum of buzz through the network interactions (Campbell et al. (2012)).

This paper attempts to provide some guidelines for targeting influentials (seeding) that facilitates the optimal information dissemination in such a social network context. In pursuit of this goal, we consider a network game amongst multiple self-interested players. The game exhibits payoff externalities (i.e., a player's contribution has direct impacts on other players' payoffs) and strategic complementarity (i.e., an increase of one player's contribution leads to a weakly increasing marginal benefit of other players' contributions). There is an underlying state ("product quality") that influences the players' valuations uniformly, and it is ex ante unknown to all the players. The network designer is allowed to choose a subset of players (leaders) to move first, and then the rest follow after observing the leaders' contributions. In compliance with the targeted advertising practice and research, the network designer is able to make the leaders informed about the product quality via experiencing the trial versions or having free samples beforehand. Since the leaders are informed about the state, the followers shall speculate the information content they have received by observing their contributions. A priori, the information transmissions in this network game are not so straightforward because of payoff externalities: leaders care directly about the followers' contributions, and therefore it is not a dominant strategy to simply disclose information. In the absence of direct communication channel among players, the contributions serve as credible *signaling* instruments in this sequential-move network game.

We characterize the unique linear perfect Bayesian Nash equilibrium in which each leader's contribution is linear in the underlying state. The inversibility of linear functions implies that the equilibrium is fully separating; thus, in equilibrium every player ultimately knows the true state. We then build upon these equilibrium characterizations to examine the selection problem of leader group. This is in the same spirit of Zhou and Chen (2013), in which we consider a complete-information version of this problem. In that paper, we show that it is isomorphic to a weighted maximum-cut problem, and sequential-moving gives rise to a second-order enhancement of aggregate contribution. In the presence of incomplete information, we prove that this problem is fundamentally *different* from the maximum-cut problem due to the signaling incentives. Furthermore, the appropriate selection of leader

4

group can lead to a *first-order* materialistic effect on the equilibrium outcomes. Thus, in this social network context a little incomplete information takes us a long way.

The above result subsequently provides some non-trivial recommendation of the network designer's information management strategies. We propose a novel index for the key leader selection (i.e., a single player to provide information to). We show that this index can be substantially different from the key player index in Ballester et al. (2006) and the key leader index with complete information proposed in Zhou and Chen (2013). In undirected graphs, the optimal leader group identified in Zhou and Chen (2013) is exactly the optimal follower group when signaling is present. In particular, if the graphs are complete (i.e., there exists an unweighted link between each pair of players), the pecking order suggests that the leader should be selected by the *ascending* order of intrinsic valuations. This is exactly the opposite criterion used in the case with complete information (Zhou and Chen (2013)). We also examine the out-tree hierarchical structure in which the externality is generated only one way from a player to another. This pyramidal network is representative in the organization structure context, and two common approaches have been widely studied and proposed: top-down (passing information from the root of the tree) and bottom-up (starting from the end leaves). However, our analysis reveals that despite their popularity, neither of these two is optimal. The key leader turns out to be the one that stays in the *middle*, and it is not necessarily exactly the central player in the network.

The remainder of this paper is organized as follows. In Section 2, we review some relevant literature. Section 3 gives a simple two-player example to illustrate the main ideas of the paper. Section 4 introduces the model setup. Section 5 characterizes the equilibrium outcomes and discusses the key-leader problem in this context. Section 6 provides further examples of network structures for illustration. We draw some concluding remarks in Section 7. All the technical proofs are relegated to the appendix.

2 Literature review

There has been vast literature on network externality that elaborates on how a player's utility depends directly on other players' decisions. The classical papers primarily focus on the aggregate level of network externality (i.e., the aggregate number of players in the game is

a sufficient statistics of this effect); see, e.g., Rohlfs (1974), Katz and Shapiro (1985), Farrell and Saloner (1986), and Economides (1996) for an extensive survey of this literature. The influential paper by Ballester et al. (2006) incorporates the local network effects (i.e., players may care more about some players' actions than others). In the simultaneous-move network game, they show that the measure "weighted Katz-Bonacich Centrality" can be used to describe the Nash equilibrium outcomes; this nicely bridges the network economics literature and the sociology literature. Candogan et al. (2012) incorporate the pricing decisions into the framework of Ballester et al. (2006) and examine both the perfect discrimination scenario and the case when the seller is restricted to use a limited number of prices. See also Ballester and Calvó-Armengol (2010), Bramoullé and Kranton (2007), and Corbo et al. (2006) for further discussions and Jackson (2008) for a comprehensive survey.

As aforementioned, we build upon our earlier work Zhou and Chen (2013). In that paper, we introduce the possibility of sequential-moving using the framework of Ballester et al. (2006). In the two-stage game of Zhou and Chen (2013), we examine the problem of group selection: which subset of players should be included in the first-movers (leaders). We show that the problem is isomorphic to the classical weighted maximum-cut problem, and we proceed to characterize the optimal group selections for several special network structures. If the network designer is allowed to determine the entire hierarchy (sequence) of moves, Zhou and Chen (2013) show that the optimal hierarchy turns out to be a chain structure. The new feature in the current paper is the information asymmetry amongst players. Thus, the current paper applies more suitably to the situations in which products or services involve substantial uncertainty, thereby creating room for information dissemination. In such a scenario, the leaders' actions serve as effective signaling instruments that convey their private information to the followers. We show that this signaling incentive leads to first-order contribution improvement; in contrast, the benefit of sequentiality in Zhou and Chen (2013) is of second order, irrespective of whether it is a restricted two-stage game or a general hierarchy design problem. This subsequently leads us to prove that the group selection problem is substantially different from the weighted maximum-cut problem, and some design principles are radically different (e.g., the swap between the leader and follower groups, and the implications of the pecking order based on intrinsic valuations).

Our paper also adds to the social learning literature. Unlike Banerjee (1992) and Bikhchandani et al. (1992), we endogenize the flow of information transmissions via group

selection. There are some papers that examine the learning aspect in dynamic networks (e.g., Acemoglu et al. (2011) and Acemoglu et al. (2013)). However, due to the payoff externalities, in our paper truthful information disclosure is not a priori obvious since players care about other players' actions and therefore how their own actions affect other players' beliefs. Consequently, signaling comes in to play at its full force. The identification of influential players is reminiscent of the idea of targeted advertising, and it has recently been re-examined in various sociology and marketing contexts (such as Coulter et al. (2003), Gladwell (2006), Van den Bulte and Joshi (2007), and Weimann (1994)). Our work addresses directly the critiques by Watts and Dodds (2007) that this research stream should be grounded by economic micro-foundation, as we provide a theoretical framework that explicitly accounts for individual decision making and information dissemination. A very recent paper by Campbell et al. (2012) examines the buzz management from the perspective of social status signaling, where information is exchanged via random individual meetings. In our context, the underlying uncertainty concerns the quality of new products/ services rather than the individual players' social types and the signaling instrument is the contribution level rather than communication. Thus, the economic drivers and accordingly the strategy recommendations are fundamentally different. The identification of influentials for seeding is also examined in the computer science field (such as Chen et al. (2009) and Kempe et al. (2005)); nevertheless, the signaling effect we study here has no counterparts in that research stream.

3   A two-player toy example

To illustrate the main ideas of this paper, we hereby present a toy example with two players. Each player i's payoff is given in the following form:

$$u_i = \alpha_i \theta x_i - \frac{1}{2} x_i^2 + \delta x_i x_{3-i}, i \in \{1, 2\}. \tag{1}$$

In (1), $\alpha_i > 0$ measures the *intrinsic marginal utility (valuation)* for player i. The parameter θ captures the *product quality*. The quadratic term $-\frac{1}{2} x_i^2$ is adopted to capture the diminishing marginal return of the player's own contribution. The last term captures the network effect among the players. The parameter $\delta > 0$ controls the strength of this effect, and it is common across the two players. Each player is entitled to determine the level of contribution (effort) x_i. In the following we investigate several different scenarios of information and game

structures.

We start with the simplest case in which the product quality, θ, is publicly known and the two players make their contributions simultaneously. In this simultaneous-move game with complete information, we can write down the players' best responses and solve the game easily. The Nash equilibrium is described below:

$$\mathbf{x}^N = (x_1^N, x_2^N)' = \left(\frac{\alpha_1 + \delta\alpha_2}{1 - \delta^2}\theta, \ \frac{\alpha_2 + \delta\alpha_1}{1 - \delta^2}\theta \right)',$$

where the superscript N denotes the Nash equilibrium.

Sequential-move game. Now consider the case in which player 1 moves first and assume again that θ is publicly known. As aforementioned, this "targeting" can be facilitated by offering free samples, demonstrating an innovation, or showcasing the idea/ behavior to player 1. In this case, player 2 observes x_1 and plays her best response to $x_2 = \alpha_2\theta + \delta x_1$ (from the first-order condition). Anticipating player 2's response in the subsequent stage, player 1 determines her optimal level of contribution. In this sequential-move game with complete information, the outcomes of the unique subgame perfect Nash equilibrium are:

$$x_1^* = \frac{\alpha_1 + \delta\alpha_2}{1 - 2\delta^2}\theta, \text{ and } x_2^* = \alpha_2\theta + \delta x_1^* = \frac{(1 - \delta^2)\alpha_2 + \delta\alpha_1}{1 - 2\delta^2}\theta.$$

Comparing the contribution levels in these two games, we find that $x_1^* > x_1^N$ and $x_2^* > x_2^N$. Thus, sequential-moving boosts both players' contribution levels.

Moreover, the aggregate effort under sequential move is $\frac{(1+\delta)\alpha_1 + (1+\delta-\delta^2)\alpha_2}{1 - 2\delta^2}\theta$. If we are free to choose any player as the first mover, the player with higher α should be the one, since the coefficient $1 + \delta$ is greater than $1 + \delta - \delta^2$. This is shown in any complete graph in Zhou and Chen (2013), and we include it here for comparison purposes.

Incomplete information: simultaneous-move game. Next, we introduce incomplete information regarding the product quality. In such a scenario, we assume that θ is a random draw from $\Theta = [0, \bar{\theta}] \subset R^+$ with some distribution. Furthermore, we impose *information asymmetry* between players: player 1 knows the realization of θ but player 2 does not. In a simultaneous-move game, from player 2's perspective she faces different types of player 1. Thus, a Bayesian game is played and θ refers to player 1's type. Given player 2's contribution x_2, a type-θ player 1's best response is $\alpha_1\theta + \delta x_2$. On the other hand, facing different types of player 1, player 2's goal is to find the contribution level that maximizes

her expected payoff. These collectively give rise to the following equilibrium outcomes:

$$\bar{x}_1(\theta) = \alpha_1\theta + \delta\frac{\alpha_2 + \delta\alpha_1}{1 - \delta^2}\mathbb{E}[\theta] = \frac{\alpha_1\left((1 - \delta^2)\theta + \delta^2\mathbb{E}[\theta]\right) + \delta\alpha_2}{1 - \delta^2}, \text{ and } \bar{x}_2 = \frac{\alpha_2 + \delta\alpha_1}{1 - \delta^2}\mathbb{E}[\theta].$$

The above analysis utilizes the linearity of the best responses. Consequently, player 2, who is uninformed about the true state, simply takes the expectation over all possible states.

Incomplete information: sequential-move game. Now suppose that player 1 moves first. In this sequential-move game, information asymmetry complicates the players' decision making. Since player 1 knows the realization of θ, player 2 will infer this information from x_1. Accordingly, player 2's contribution will reflect this inference and this conversely affects player 1's decision. In this sense, player 1's contribution x_1 serves as a *signaling* instrument to convey her information. Later in the paper(Theorem 1) we show that there exists a unique separating perfect Bayesian equilibrium which is linear in θ. In that equilibrium, the strategy of player 1 is:

$$\tilde{x}_i(\theta) = \frac{\alpha_1 + 2\delta\alpha_2}{1 - 2\delta^2}\theta,$$

which suggests that player 1 contributes more when the product quality is higher. This credibly conveys the quality information to player 2, because player 2 can simply invert this linear function and update her belief after observing x_1:

$$\hat{\theta} = \frac{1 - 2\delta^2}{\alpha_1 + 2\delta\alpha_2} \cdot x_1.$$

Thus, her best response is to choose effort $\hat{\theta}\alpha_2 + \delta x_1$.

On the equilibrium path, the players' contribution levels are:

$$\hat{x}_1 = \frac{\alpha_1 + 2\delta\alpha_2}{1 - 2\delta^2}\theta, \text{ and } \hat{x}_2 = \theta\alpha_2 + \delta\hat{x}_1 = \frac{\alpha_2 + \delta\alpha_1}{1 - 2\delta^2}\theta,$$

and the aggregate contribution is $\frac{(1+\delta)\alpha_1 + (1+2\delta)\alpha_2}{1 - 2\delta^2}\theta$. We can then compare these contributions with the above scenarios:

$$\hat{x}_1 > x_1^* > x_1^N \quad \text{and} \quad \hat{x}_2 > x_2^* > x_2^N,$$

(here we assume δ is small enough, specially $\delta < \sqrt{1/2}$). Therefore, we conclude that the signaling effect in this sequential-move game can further enhance both players' incentives to contribute. The informed player as the leader contributes more because she wants to
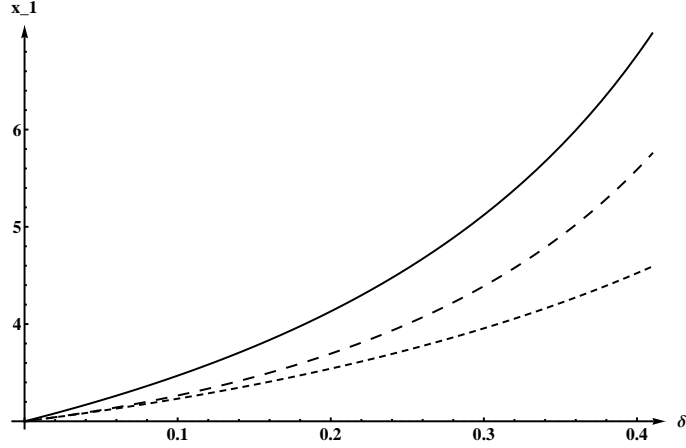
Figure 1: Equilibrium efforts of player 1 under different scenarios
$\alpha_1 = 3, \alpha_2 = 2, \theta = 1$; short dash line: x_1^N; long dash line: x_1^*; solid line: $\hat{x}_1$.

credibly convey this information, and because of this the uninformed player as the follower also contributes more.

Magnitude of effort increments. To visualize the impacts of sequential-moving and information asymmetry, we plot player 1's efforts in Figure 1 for different values of δ. Note that the difference between x_i^N and x_1^* accounts for the effect of sequential-moving, because they are both obtained under complete information. On the other hand, the difference between x_1^* and $\hat{x}_1$ accounts for the effect of incomplete information, whereas the players move sequentially. From Figure 1, we observe that the incremental effort from information asymmetry is *more significant* than that from sequential-moving, and this dominance is particularly apparent when δ is small. This is because the increments are of different orders (one is of order δ, and the other is of order δ^2); later we analytically verify this in Corollary 2.

Leader selection. In addition, in the sequential-move game with incomplete information, the aggregate contribution in equilibrium is $\frac{(1+\delta)\alpha_1+(1+2\delta)\alpha_2}{1-2\delta^2}\theta$. Note that the coefficient of $\alpha_1 (1+\delta)$ is less than that of $\alpha_2 (1+2\delta)$. If we are free to target any player to disclose information to (and make her the leader), the player with a *lower* intrinsic valuation should be the leader. This is because the player with a lower intrinsic valuation is more credible if her contribution turns out to be high, i.e., the signaling effect gets substantially amplified. Thus, our analysis also provides a simple guideline for the selection of "key leader." Inciden-

10

tally, this result is in strict contrast with that under complete information. When all the players perfectly know the true state, Zhou and Chen (2013) show that the key leader in this two-player example is the one with the higher intrinsic valuation. This illustrates the radical difference between the information channel and the pure strategic complementarity. We later return to this point and articulate the underlying rationale for this discrepancy.

How general are these results? To address this question in the sequel we present a general setup with more than two players. We will also investigate other related research questions with the help of the general setup.

4   Model

In our general model, we consider a network game with n players and their payoffs are specified as follows:

$$u_i(x) = \theta \alpha_i x_i - \frac{1}{2} x_i^2 + \delta \sum_{j \neq i} g_{ij} x_i x_j,$$

where $g_{ij} \geq 0$ and $g_{ii} = 0\ \forall i, j$. The cross term $g_{ij} x_i x_j$ 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). We also require that $g_{ii} = 0$, i.e., there is no self-loop; this indicates that the cross effect appears only amongst different players' contributions. In most of the paper, it is the adjacent matrix of an undirected graph. As in our toy example, α_i measures the intrinsic marginal utility for player i, and θ captures the intrinsic product quality.

Notation. 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 $T_{ii}^D = t_{ii}, i = 1, \cdots N$, and off-diagonal entries $T_{ij}^D = 0, \forall i \neq j$. Unless indicated otherwise, vector $\mathbf{x} = (x_1, \cdots, x_N)'$ is a column vector. For any subset A of $\mathcal{N}$, $x_{\mathbf{A}}$ (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 $\mathbf{x}, \mathbf{y}$.

11

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 $\lim \sup_{\delta \to 0} |\frac{f_1(\delta)}{f_2(\delta)}| < \infty$, 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 with complete information. Now we introduce the two benchmarks with complete information (i.e., θ is commonly known). The simultaneous-move game has been studied by Ballester et al. (2006). They show that the *Nash equilibrium* is characterized by

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

where $\alpha = (\alpha_1, \cdots, \alpha_n)'$. Let $\mathbf{M} := [\mathbf{I} - \delta G]^{-1}$ and

$$m_{ij} = \sum_{k=0}^{+\infty} \delta^k g_{ij}^{[k]} = \delta_{ij} + \delta g_{ij} + \delta^2 g_{ij}^{[2]} + \cdots,$$

where $g_{ij}^{[k]}$ is the ij entry of G^k. Using m_{ij}, we can rewrite (2) as:

$$x_i^N = b_i(G, \delta, \alpha\theta) = \sum_{j=1}^{N} m_{ij}\alpha_j\theta,$$

where $\mathbf{b}(G, \delta, \alpha\theta) = [\mathbf{I} - \delta G]^{-1}\alpha\theta$ is called the *weighted Katz-Bonacich Centrality* of parameter δ and weight vector $\alpha\theta$.

Sequential-move game with complete information. Zhou and Chen (2013) introduce the sequential-move feature to the above network game. In the two-stage game, we divide the players into two groups, leader group L and follower group F. For convenience, let us rewrite the matrix G as a block matrix: $G = \begin{pmatrix} G_{LL} & G_{LF} \\ G_{FL} & G_{FF} \end{pmatrix}$. Let vectors $\mathbf{x}_L$ and $\mathbf{x}_F$ denote the contributions chosen by the nodes in L and those in F, respectively. The following proposition is adopted from Zhou and Chen (2013).

Proposition 1. *For sufficiently small δ,[1] the unique subgame perfect Nash equilibrium of*

---

[1]For example, the result holds when $\delta < \frac{1}{2\mu_1(G)}$ if G is symmetric, where $\mu_1(G)$ is the largest eigenvalue of G. 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.

the two-stage game is given by:[2]

$$\begin{pmatrix} \mathbf{x}_L \\ \mathbf{x}_F \end{pmatrix} = \mathbf{S} \begin{pmatrix} \alpha_L \theta \\ \alpha_F \theta \end{pmatrix} \tag{3}$$

with

$$\mathbf{S} = \begin{pmatrix} \left[1 - \delta(T + T^D)\right]^{-1} & \delta \left[1 - \delta(T + T^D)\right]^{-1} G_{LF} U \\ \delta U G_{FL} \left[1 - \delta(T + T^D)\right]^{-1} & U + \delta^2 U G_{FL} \left[1 - \delta(T + T^D)\right]^{-1} G_{LF} U \end{pmatrix}, \tag{4}$$

where

$$T = G_{LL} + \delta G_{LF} U G_{FL}, \text{ and } U = [I - \delta G_{FF}]^{-1}.$$

Incomplete information. Having discussed the above two benchmarks, we now in-corporate incomplete information into the setup. Ex ante, θ is drawn from the distribution $F(\cdot) : [0, \bar{\theta}] \to [0, 1]$, and for simplicity we assume that F has full support on $\Theta = [0, \bar{\theta}]$. We assume that a group of players (leaders) move first, and then the rest follow after observing the leaders' contributions. We further assume that the network designer is able to make the leaders informed about the product quality θ. This can be implemented as the network designer invites a subset of customers/ players to experience the trial versions or to give free samples to.

In this case, a group L of leaders with $|L| = M$ are all fully informed about the state. Accordingly, we use F to denote the set of the originally uninformed players, with $|F| = N - M$. In such a scenario, we use $\{x_l, l \in L\}$ or $x_{\mathbf{L}}$ to determine these leaders' effort decisions in the first stage. In the next stage, the remaining players observe the quantities $x_{\mathbf{L}}$ chosen by all the leaders, update their beliefs about the state, and subsequently choose $\{x_f, f \in F\}$ or $x_{\mathbf{F}}$ simultaneously.

5 Analysis

In this section, we carry out the equilibrium analysis.

---

[2]It is verifiable that $\mathbf{S} \succeq \mathbf{M}$, where the matrix $\mathbf{M}$ captures the sensitivities of intrinsic valuations in a simultaneous-move game. This component-wise dominance therefore implies that the equilibrium contribu-tions in this two-stage sequential-move game are higher.

5.1 Equilibrium characterization

Recall that in our setting, the leaders know the true state. Thus, their effort decisions shall be made contingent on the state realization.

Signaling. We will focus on a particular form of equilibrium in which the leaders choose their efforts according to the following:

$$x_l^*(\theta) = \kappa_l \theta, \ l \in L,$$

where k_l is a constant yet to be determined. This particular form indicates a one-to-one correspondence between the value of θ and the effort $x_l^*(\theta)$. Consequently, upon observing the leader's effort choice, each player perfectly infers the realization of θ. This suggests that the equilibrium is *fully separating.*

Note that a follower may observe the effort choices by multiple informed players. Furthermore, even though on the equilibrium path the leaders may disclose the same information, the sequential structure of the game admits all possible deviations in the first stage. Thus, we shall specify how a follower updates her belief in each instance, not only on the equilibrium path but also off-equilibrium. We will adopt the *pessimistic belief* in the following sense.[3] After observing $\{x_l, l \in L\}$, a follower's belief about the state is $\hat{\theta}(x_{\mathbf{L}}) = \min_{l \in L}(\frac{x_l}{\kappa_l})$, which is the same among all followers, as they observe the same effort vector $x_{\mathbf{L}}$.

Let $U = [1 - \delta G_{FF}]^{-1}$ and $T = G_{LL} + \delta G_{LF} U G_{FL}$. The results are summarized in the following theorem.

Theorem 1. *With pessimistic beliefs, there exists a fully separating perfect Bayesian equilibrium (PBE) in which each leader i in group L chooses*

$$\tilde{x}_i(\theta) = \kappa_i \theta, \ for \ all \ \ i \in L,$$

while the coefficients $\{\kappa_i\}$'s are given by (in matrix form):

$$\kappa_{\mathbf{L}} = [1 - \delta(T + T^D)]^{-1}(\alpha_{\mathbf{L}} + 2\delta G_{LF} U \alpha_{\mathbf{F}}).$$

The followers' strategies are:

$$x_{\mathbf{F}}(x_{\mathbf{L}}) = U(\hat{\theta}(x_{\mathbf{L}})\alpha_{\mathbf{F}} + \delta G_{FL} x_{\mathbf{L}}), \ with \ belief \ \ \hat{\theta}(x_{\mathbf{L}}) = \min_{l \in L}(\frac{x_l}{\kappa_l}).$$

---

[3]If there is only a single leader, i.e., $|L| = 1$, this assumption degenerates.

In a single-leader case ($|L| = 1$), Theorem 1 characterizes the unique separating PBE. With multiple leaders, however, the equilibrium is not unique as there are other possible belief systems that give rise to different equilibrium outcomes. In the sequel, we conduct various sorts of comparative statics based on Theorem 1. For example, we can compare the equilibrium effort characterized in Theorem 1 and that with complete information in the two-stage game (see Proposition 1). To this end, we present the equilibrium efforts in Theorem 1 using the following matrix form:

$$\tilde{x} = \begin{pmatrix} \tilde{x}_{\mathbf{L}} \\ \tilde{x}_{\mathbf{F}} \end{pmatrix} = \hat{\mathbf{S}} \begin{pmatrix} \theta \alpha_{\mathbf{L}} \\ \theta \alpha_{\mathbf{F}} \end{pmatrix},$$

with

$$\hat{\mathbf{S}} = \begin{pmatrix} \left[1 - \delta(T + T^D)\right]^{-1} & 2\delta \left[1 - \delta(T + T^D)\right]^{-1} G_{LF} U \\ \delta U G_{FL} \left[1 - \delta(T + T^D)\right]^{-1} & U + 2\delta^2 U G_{FL} \left[1 - \delta(T + T^D)\right]^{-1} G_{LF} U \end{pmatrix}.$$

Corollary 2. *The equilibrium effort in the two-stage signaling is higher than the equilibrium efforts in the two-stage game without signaling, which is higher than the equilibrium efforts in the simultaneously-move game. Moreover, the effort increase compared with the simultaneous-move game has the following expression:*

$$\delta\theta \begin{pmatrix} \mathbf{0} & G_{LF} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \alpha_{\mathbf{L}} \\ \alpha_{\mathbf{F}} \end{pmatrix} + \delta^2\theta \begin{pmatrix} (G_{LF}G_{FL})^D & G_{LL}G_{LF} + G_{LF}G_{FF} \\ \mathbf{0} & G_{FL}G_{LF} \end{pmatrix} \begin{pmatrix} \alpha_{\mathbf{L}} \\ \alpha_{\mathbf{F}} \end{pmatrix} + \mathcal{O}(\delta^3). \quad (5)$$

It is worth mentioning that the group selection decision (L) affects the players' payoffs in the order of δ. In contrast, when θ is publicly known, Zhou and Chen (2013) show that sequential-moving only gives rise to a second-order enhancement of effort exertion. Thus, in this social network context a little incomplete information takes us a long way, since it has a *first-order* materialistic effect on the equilibrium outcomes. This result therefore provides some non-trivial recommendation of the network designer's information management strategies.

5.2   Key leader problem

Of particular interest is a special case of the above group selection problem. If we intend to target one player to provide information, what is the criterion for this *key leader*? In essence, this problem shares a similar spirit with the key player problem in Ballester et al. (2006). In

a simultaneous-move game, Ballester et al. (2006) consider the possibility of removing one player from the criminal network and therefore are interested in identifying the player that impacts the network most. To make notation simple, we assume G is the adjacent matrix of an undirected graph in this subsection.

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. According to Theorem 1, her equilibrium strategy is

$$\tilde{x}_i(\theta) = \frac{\alpha_i + 2\delta \left\langle \beta_i, \ (\mathbf{I} - \delta G_{-i})^{-1} \alpha_{-i} \right\rangle}{1 - 2\delta^2 \left\langle \beta_i, \ (\mathbf{I} - \delta G_{-i})^{-1} \beta_i \right\rangle} \theta, \tag{6}$$

where matrix G is rewritten as follows:

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

Using block matrix inversion for $\mathbf{I} - \delta G$ yields

$$\mathbf{M} = [\mathbf{I} - \delta G]^{-1} = \begin{pmatrix} \frac{1}{1 - \delta^2 \left\langle \beta_i, \ (\mathbf{I} - \delta G_{-i})^{-1} \beta_i \right\rangle} & \frac{\delta \beta_i' [\mathbf{I} - \delta G_{-i}]^{-1}}{1 - \delta^2 \left\langle \beta_i, \ (\mathbf{I} - \delta G_{-i})^{-1} \beta_i \right\rangle} \\ \frac{\delta [\mathbf{I} - \delta G_{-i}]^{-1} \beta_i}{1 - \delta^2 \left\langle \beta_i, \ (\mathbf{I} - \delta G_{-i})^{-1} \beta_i \right\rangle} & [\mathbf{I} - \delta G_{-i}]^{-1} + \frac{[\mathbf{I} - \delta G_{-i}]^{-1} \beta_i \beta_i' [\mathbf{I} - \delta G_{-i}]^{-1}}{1 - \delta^2 \left\langle \beta_i, \ (\mathbf{I} - \delta G_{-i})^{-1} \beta_i \right\rangle} \end{pmatrix}.$$

Therefore, player i's equilibrium contribution in the simultaneous-move game is

$$x_i^N = \frac{\alpha_i + \delta \left\langle \beta_i, \ (\mathbf{I} - \delta G_{-i})^{-1} \alpha_{-i} \right\rangle}{1 - \delta^2 \left\langle \beta_i, \ (\mathbf{I} - \delta G_{-i})^{-1} \beta_i \right\rangle} \theta = \theta \sum_{j=1}^N m_{ij} \alpha_j. \tag{7}$$

Comparing the coefficients of α_j in (7), we obtain that

$$m_{ii} = \frac{1}{1 - \delta^2 \left\langle \beta_i, \ (\mathbf{I} - \delta G_{-i})^{-1} \beta_i \right\rangle}, \tag{8}$$

$$\frac{m_{ij}}{m_{ii}} = j\text{-th entry of } (\mathbf{I} - \delta G_{-i})^{-1} \delta \beta_i, \quad j \neq i.$$

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*

$$S_i := \frac{b_i(G, \delta, 1)(b_i(G, \delta, \alpha) - \alpha_i)}{2 - m_{ii}} \tag{9}$$

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

Recall that in the toy example (Section 3), we have already seen that if $\alpha_1 > \alpha_2$, the player with the highest b_i, c_i, L_i is player 1. However, player 2 has the highest S_i and therefore should be the key leader with incomplete information. Proposition 2 generalizes this result to any complete graph : the rationale for choosing the player with the lowest α_i continues to be valid in any complete graph with N nodes.

To see this, we first rewrite b_i using (2):

$$b_i(G, \delta, \alpha) = \alpha_i + \delta \sum_j g_{ij} b_j(G, \delta, \alpha).$$

Substituting it into the definition of S_i, (9) yields

$$S_i := \frac{b_i(G, \delta, 1)(b_i(G, \delta, \alpha) - \alpha_i)}{2 - m_{ii}} = \frac{b_i(G, \delta, 1)(\delta \sum_j g_{ij} b_j(G, \delta, \alpha))}{2 - m_{ii}}.$$

For a complete graph, $g_{ii} = 0$ and $g_{ij} = 1, \forall i \neq j$, and $b_i(G, \delta, 1) = b_j(G, \delta, 1)$ for any i, j by symmetry. Therefore,

$$\sum g_{ij} b_j(G, \delta, \alpha) = \sum_{j \neq i} b_j(G, \delta, \alpha) = \left(\sum_k b_k(G, \delta, \alpha) \right) - b_i(G, \delta, \alpha).$$

Note that the first term is common across all players (it does not depend on the index i), and $b_i(G, \delta, \alpha) > b_j(G, \delta, \alpha)$ if and only if $\alpha_i > \alpha_j$.

The heterogeneity regarding the players' intrinsic valuations constitutes the main driver that distinguishes between the cases with complete and incomplete information. If we eliminate this heterogeneity (i.e., all $\{\alpha_i\}$'s are the same and normalized to 1), and G is a *regular* undirected graph, i.e., every node has the same number of links, then S_i simplifies to $b_i(b_i - 1)/(2 - m_{ii})$. In this case, we shall pick the one with the highest m_{ii} because b_i are the same for regular graphs with homogeneous players.[4] This prediction coincides with the key leader index L_i proposed in Zhou and Chen (2013) with complete information.

5.3   Impact of incomplete information

We can also articulate the impact of incomplete information by comparing our results with those in Zhou and Chen (2013). With incomplete information, the signaling effect leads to

---

[4]If every node in G has degree d, then $G\mathbf{1} = d\mathbf{1}$, so $[1 - \delta G]^{-1} \mathbf{1} = \frac{1}{1 - \delta d} \mathbf{1}$.

the following group selection criterion:

$$\max_L \mathbf{1}' G_{LF} \alpha_{\mathbf{F}}. \tag{10}$$

In contrast, when θ is publicly known, Zhou and Chen (2013) show that the optimal leader set is determined by $\max \mathbf{1}'(G_{LF}G_{FL})^D \alpha_{\mathbf{L}}$. As a special case, suppose that the graph is undirected, i.e., $g_{ij} = g_{ji} \in \{0,1\}, \forall i \neq j$. In such a scenario, we find that

$$\max_L \mathbf{1}'(G_{LF}G_{FL})^D \alpha_{\mathbf{L}} = \max_L \mathbf{1}' G_{FL} \alpha_{\mathbf{L}}. \tag{11}$$

We can see that the optimization program (11) is exactly the same as program (10) after we *swap* the roles of F and L. Here, the optimal leader group identified in Zhou and Chen (2013) is exactly the optimal follower group when signaling is present. We formally state this result below.

Proposition 3. *Suppose that G is the adjacent matrix of an undirected graph, i.e., $g_{ij} = g_{ji} \in \{0,1\}, \forall i \neq j$, and $g_{kk} = 0, \forall k$, and the parameter δ is small enough. The optimal follower group with incomplete information is exactly the optimal leader group in the case with complete information.*

As a corollary, the two combinatorial optimization programs are equally difficult to solve, as they have the same complexity. Going beyond undirected graphs (i.e., G is not symmetric), incomplete information leads to radically different group selection strategies from the case with complete information. The following example provides a crispy contrast between the two.

Another two-player example. In this example, there are two players with the following payoff functions:

$$u_1(x_1, x_2) = \theta \alpha_1 x_1 - \frac{1}{2} x_1^2 + \delta x_1 x_2, \text{ and } u_2(x_1, x_2) = \theta \alpha_2 x_2 - \frac{1}{2} x_2^2,$$

or equivalently

$$G = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

With complete information, in the simultaneous-move game player 2's best reply is to play $x_2^* = \theta \alpha_2$. Thus, player 1 will choose $x_1^* = \theta \alpha_1 + \delta x_2^* = \theta(\alpha_1 + \delta \alpha_2)$. Notice that this is also the equilibrium outcomes in the sequential-move game (with either player 1 or player 2 being the leader). This is because $g_{12} g_{21} = 1 \cdot 0 = 0$ here.

Now let us introduce incomplete information. When player 1 is the leader, her equilibrium strategy is

$$x_1(\theta) = (\alpha_1 + 2\delta\alpha_2)\theta.$$

Accordingly, player 2's belief is $\hat{\theta}(x_1) = \frac{x_1}{\alpha_1 + 2\delta\alpha_2}$, and hence player 2 chooses $x^*(x_1) = \hat{\theta}(x_1)\alpha_2 = \frac{x_1\alpha_2}{\alpha_1 + 2\delta\alpha_2}$. Therefore, on the equilibrium path, their efforts are $((\alpha_1 + 2\delta\alpha_2)\theta, \alpha_2\theta)$; these dominate $((\alpha_1 + \delta\alpha_2)\theta, \alpha_2\theta)$. On the other hand, when player 2 is the leader, this signaling effect disappears (because $g_{21} = 0$). Therefore, players' equilibrium efforts are the same as those in the case with complete information about θ.

The above discussions reveal that in this case, the sequence of moves does not matter in the case of complete information. However, with incomplete information player 1 should be the leader. Note that in this example, player 1's payoff is positively affected by her neighbor's effort. This is precisely where her signaling incentive comes from. We summarize these results in Table 1.

players	m_{ii}	b_i	c_i	L_i	S_i
1	1	$\alpha_1 + \delta\alpha_2$	$\alpha_1 + \delta\alpha_2$	0	$\delta\alpha_2$
2	1	α_2	$(1+\delta)\alpha_2$	0	0

Table 1: Comparison of different measures: $m_{12} = \delta$, $m_{21} = 0$, $m_{11} = m_{22} = 1$.

As indicated in Table 1, player 1 has a higher S_i (because $S_2 = 0$). Nevertheless, if $\alpha_2 > \alpha_1/(1-\delta)$, player 2 has higher values of b_i and c_i than player 1 does; however, player 1 has the highest S_i. If $\alpha_1 < \alpha_2 < \alpha_1/(1-\delta)$, then player 2 has a higher c_i but lower b_i, S_i; if $\alpha_2 < \alpha_1$, then player 1 has higher values of b_i, c_i, S_i. These indicate the substantial differences amongst the various indices proposed in the existing literature. In the sequential-move game with incomplete information, player 1 should be the leader ($S_1 < 0 = S_2$). With complete information, however, both players are key leaders ($L_1 = L_2 = 0$). In the simultaneous-move game, the selection of key player depends on the relative sizes of $\alpha_1, \alpha_2, \delta$. For example, when $\alpha_1 = 0.8, \alpha_2 = 1, \delta = 0.1, b_1 = 0.9 < b_2 = 1$; thus, player 1 is the key player. With other parameters, the choice may be different.

Revisiting the two-player toy example. Finally, let us return to the toy example. The corresponding network matrix G is $G = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, and we summarize the results in Table

19

2. If $\alpha_1 > \alpha_2$, then player 1 has higher values of b_i, c_i, and L_i but a lower S_i.

players	m_{ii}	b_i	c_i	L_i	S_i
1	$\frac{1}{1-\delta^2}$	$\frac{(\alpha_1+\delta\alpha_2)}{1-\delta^2}$	$\frac{(\alpha_1+\delta\alpha_2)}{1-\delta}$	$\frac{\delta^2(\alpha_1+\delta\alpha_2)}{(1-\delta)(1-2\delta^2)}$	$\frac{\delta(\alpha_2+\delta\alpha_1)}{(1-\delta)(1-2\delta^2)}$
2	$\frac{1}{1-\delta^2}$	$\frac{(\alpha_2+\delta\alpha_1)}{1-\delta^2}$	$\frac{(\alpha_2+\delta\alpha_1)}{1-\delta}$	$\frac{\delta^2(\alpha_2+\delta\alpha_1)}{(1-\delta)(1-2\delta^2)}$	$\frac{\delta(\alpha_1+\delta\alpha_2)}{(1-\delta)(1-2\delta^2)}$

Table 2: Comparison of different measures: $m_{11} = m_{22} = \frac{1}{1-\delta^2}$, $m_{12} = m_{21} = \frac{\delta}{1-\delta^2}$ ($\delta \in (0, \sqrt{1/2})$).

6 Some specific network structures

In this section, we consider several special cases to illustrate the results of group selection problems.

6.1 Hierarchical structure or out-tree

In the organization structure context, sometimes it is sensible to assume that the externality is generated only one way from a player to another. Putting it in our network setup, this can be conveniently modeled as a a hierarchical structure,[5] which is represented by a pyramidal network, A single individual, called the root/principal, is at the top, and each other individual is assigned a unique direct superior. Each hierarchical structure defines a matrix G: $g_{ij} = 1$ *if and only if i is the direct superior of j*. A hierarchical structure, in graph theory literature, is also defined as *out-tree*: an oriented tree with only one vertex of in-degree zero, see Bang-Jensen and Gutin (2008) for classical notation and concepts.

The two-player example given in Section 5.3 is a special case of the directed trees: player 2's effort generates some externality to player 1's payoff, but the converse is not true. In the economics of organization literature, a player's decision is impacted by her subordinates, thereby giving rise to an out-tree structure.

For any given hierarchical structure, we can now provide the exact indices for different

---

[5]See Demange (2004) for the example of hierarchical structures.

purposes as follows:

$$c_i = \frac{(\sum_j m_{ji})b_i(G, \delta, \alpha)}{m_{ii}},$$ (12)

$$S_i = \frac{(\sum_j m_{ji})(b_i(G, \delta, \alpha) - \alpha_i)}{2 - m_{ii}},$$

$$L_i = \frac{(m_{ii} - 1)}{(2 - m_{ii})}\frac{(\sum_j m_{ji})}{m_{ii}}b_i(G, \delta, \alpha).$$

In other words, we just substitute $b_i(G, \delta, 1)$ by $\sum_j m_{ji}$ in the original index. Note that if a network is undirected, then G is symmetric and $\mathbf{M}$ is symmetric and $m_{ij} = m_{ji}$. This returns to the original definitions of those indices. To visualize the generalization from the example in Section 5.3, consider the following directed chain:

$$1 \longrightarrow 2 \longrightarrow 3 \longrightarrow 4 \longrightarrow 5$$

The adjacent matrix G is as follows:

$$G = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \text{ and } \mathbf{M} = [\mathbf{I} - \delta G]^{-1} = \begin{bmatrix} 1 & \delta & \delta^2 & \delta^3 & \delta^4 \\ 0 & 1 & \delta & \delta^2 & \delta^3 \\ 0 & 0 & 1 & \delta & \delta^2 \\ 0 & 0 & 0 & 1 & \delta \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Notice that $m_{ii} = 1, \forall i$ as there is no loop in any directed tree, and m_{ij} is the number of paths from i to j with discounting. In the simultaneous-move game, the Nash equilibrium is

$$\mathbf{x}^N = \mathbf{b}(G, \delta, \theta\alpha) = \begin{bmatrix} 1 & \delta & \delta^2 & \delta^3 & \delta^4 \\ 0 & 1 & \delta & \delta^2 & \delta^3 \\ 0 & 0 & 1 & \delta & \delta^2 \\ 0 & 0 & 0 & 1 & \delta \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \theta\alpha_1 \\ \theta\alpha_2 \\ \theta\alpha_3 \\ \theta\alpha_4 \\ \theta\alpha_5 \end{bmatrix}.$$

Different measures are calculated in Table 3.

In the following discussions, we assume $\delta \in (0, 1)$ to restrict the impact of social interactions. From Table 3, we observe that if $\alpha_i = 1$, then player 1 has the highest b_i, player 2 has the second highest b_i, ..., and player 5 has the lowest b_i. Since the externality is only one-way ($m_{ii} = 1$), all of them have zero-value L_i. Interestingly, the key leader problem in

node	b_i	c_i	L_i	S_i
1	$(\alpha_1 + \delta\alpha_2 + \delta^2\alpha_3 + \delta^3\alpha_4 + \delta^4\alpha_5)$	$(\alpha_1 + \delta\alpha_2 + \delta^2\alpha_3 + \delta^3\alpha_4 + \delta^4\alpha_5)$	0	$(\delta\alpha_2 + \delta^2\alpha_3 + \delta^3\alpha_4 + \delta^4\alpha_5)$
2	$(\alpha_2 + \delta\alpha_3 + \delta^2\alpha_4 + \delta^3\alpha_5)$	$(\alpha_2 + \delta\alpha_3 + \delta^2\alpha_4 + \delta^3\alpha_5)(1 + \delta)$	0	$(\delta\alpha_3 + \delta^2\alpha_4 + \delta^3\alpha_5)(1 + \delta)$
3	$(\alpha_3 + \delta\alpha_4 + \delta^2\alpha_5)$	$(\alpha_3 + \delta\alpha_4 + \delta^2\alpha_5)(1 + \delta + \delta^2)$	0	$(\delta\alpha_4 + \delta^2\alpha_5)(1 + \delta + \delta^2)$
4	$(\alpha_4 + \delta\alpha_5)$	$(\alpha_4 + \delta\alpha_5)(1 + \delta + \delta^2 + \delta^3)$	0	$(\delta\alpha_5)(1 + \delta + \delta^2 + \delta^3)$
5	(α_5)	$(\alpha_5)(1 + \delta + \delta^2 + \delta^3 + \delta^4)$	0	0

Table 3: Comparison of different measures for directed chain C_5: Bonacich centrality measure b_i; Inter-centrality measure c_i (Ballester et al. (2006)); Leading index L_i (Zhou and Chen (2013)); Signaling index S_i;

our context turns out to be the one that stays in the *middle*, because she is more likely to have the highest S_i; moreover, it is not necessarily exactly the central player in the network. In this example, it could be either player 2 or player 3 depending on the value of δ, i.e., depending on the relative importance between the intrinsic valuations and social interactions. This result also has an intriguing implication on the information management problem in the organization structure. In the literature on economics of organizations, two common approaches have been widely studied and proposed: top-down (passing information from the root of the tree) and bottom-up (starting from the end leaves). However, our analysis reveals that neither of these two is optimal, despite their popularity.

To articulate the underlying rationale, let us first revisit the key player index c_i. In this example, it should be exactly in the middle of the chain (i.e., player 3) because

$$(1 + \delta + \delta^2)^2 > (1 + \delta)(1 + \delta + \delta^2 + \delta^3) > (1 + \delta + \delta^2 + \delta^3 + \delta^4).$$

From (12), the index c_i has the following interpretation:

$$c_i = \frac{(\sum_j m_{ji})b_i(G, \delta, \alpha)}{m_{ii}} = (\sum_j m_{ji}))(\sum_k m_{ik})).$$

$m_{ii} = 1, \forall i$ as there is no loop in any out-tree. Note that $\sum_j m_{ji}$ counts the number of paths from any player to i, and $\sum_k m_{ik}$ counts the number of paths that start from i to any player. Therefore, the product of them simply counts all the paths that go through player i with discounting. Remember that in the example C_5, player 1 has the largest number of paths that start from herself (the highest $\sum_k m_{ik}$), and the smallest number of paths that end with herself (the lowest $\sum_j m_{ji}$). Yet, the player 3 has the highest c_i, as it lies in the center of the figure, because the key-player selection has to balance between $\sum_k m_{ik}$ and $\sum_j m_{ji}$.

22

The same logic applies to the key-leader selection as well, but with some minor twist. In a directed tree, there is a unique path from i to j, if such a path exists. Therefore, $\sum_k m_{ik}$ counts the number of followers discounted by the distance, while $\sum_j m_{ji}$ counts the number of predecessors of i discounted by the distance. We can then express the index S_i as follows:

$$S_i = \frac{(\sum_j m_{ji})(b_i(G, \delta, \alpha) - \alpha_i)}{2 - m_{ii}} = (\sum_j m_{ji})(\sum_k m_{ik} - 1).$$

The term $\sum_k m_{ik} - 1$ counts the number of paths from player i that have a length of at least 2 with discounting (excluding the trivial path from player i to herself). Therefore, the product is the discounted number of paths that go through player i but do not end with her.

In closing this subsection, we present another out-tree with three tiers of branches in Figure 2. The root player (player R) has q direct followers ("children"), and each child M_i has p direct followers; all the grandchildren $\{F_j\}$'s have no followers in this three-tier tree. Table 4 summarizes the measures in this example. To be concrete, let us take $p = 4, q = 3$, and $\delta = 0.1$. In this case, $b_R = 1.42 > b_M = 1.4 > b_F = 1$, and $c_M = 1.54 > c_R = 1.42 > c_F = 1.11$ and $S_M = 0.44 > S_R = 0.42 > S_F = 0$. Therefore, player R has the highest b_i, while each player in the second tier M_i has the highest c_i and S_i.
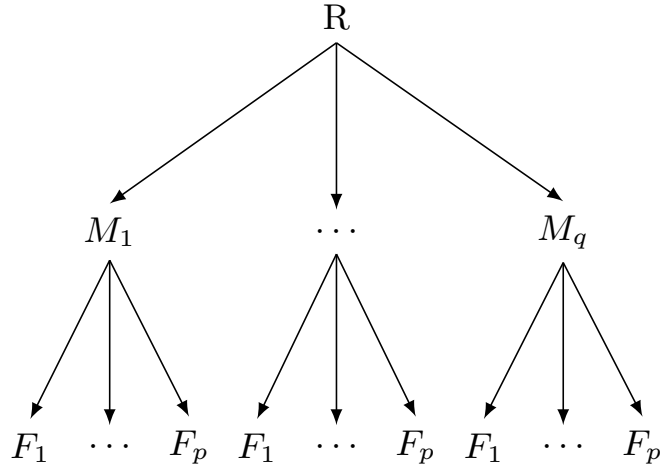

Figure 2: A branched tree.

node	b_i	c_i	L_i	S_i
R	$(1 + q\delta + pq\delta^2)$	$(1 + q\delta + pq\delta^2)$	0	$(q\delta + pq\delta^2)$
M	$(1 + p\delta)$	$(1 + p\delta)(1 + \delta)$	0	$p\delta(1 + \delta)$
F	1	$(1 + \delta + \delta^2)$	0	0

Table 4: Comparison of different measures for a branched tree: $\alpha_i = 1$.

6.2 Complete graphs

The second example we investigate is the family of complete graphs. In this case, let J_{mn} be the matrix of 1s with size m by n, and I_k be the identity matrix with size k by k. Then the network matrix of a complete graph can be expressed as $G = J_{NN} - I_N$, and (5) is reduced to

$$\Delta \mathbf{x} = \delta\theta \begin{pmatrix} \mathbf{0} & J_{LF} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \alpha_{\mathbf{L}} \\ \alpha_{\mathbf{F}} \end{pmatrix} + \delta^2\theta \begin{pmatrix} FI_L & (N-2)J_{LF} \\ \mathbf{0} & LJ_{FF} \end{pmatrix} \begin{pmatrix} \alpha_{\mathbf{L}} \\ \alpha_{\mathbf{F}} \end{pmatrix} + \mathcal{O}(\delta^3).$$

Hence the aggregate difference is:

$$\Delta x = \delta L \theta \alpha_F + \delta^2 \theta \left(F\alpha_L + L(N-2)\alpha_F + LF\alpha_F \right) + \mathcal{O}(\delta^3). \tag{13}$$

As a special case, when $\alpha_i = 1, \forall i$, (13) reduces to $\delta\theta L(N-L) + \mathcal{O}(\delta^2)$. Although the optimal group size when players are homogeneous on a complete graph is still $\lfloor \frac{N}{2} \rfloor$, but the optimal choice of players with heterogeneous $\{\alpha_i\}$'s is different for any fixed group size L. To see this, we note that from (13) the dominant term is $\delta L \theta \alpha_F$. Therefore, the players with higher $\{\alpha_i\}$'s should be placed in the follower group F. In contrast, with complete information they are selected as leaders (Zhou and Chen (2013)).

6.3 Complete bipartite graphs

Next, we examine complete bipartite graphs. This is appropriate for situations in which there is a natural separation between players (e.g., buyers and sellers, boys and girls, and employers and workers). Figure 3 is a bipartite graph with $M = 3$ nodes on one side, and $N = 5$ nodes on the other side. In general, the network matrix G can be written as $G = \begin{pmatrix} \mathbf{0} & J_{MN} \\ J_{NM} & \mathbf{0} \end{pmatrix}$. Let $\alpha_i, i \in M$ and $\beta_j, j \in N$ denote the intrinsic valuations of players in the leader and follower groups, respectively. Define $\bar{\alpha} = \frac{\sum_{i \in M} \alpha_i}{M}, \bar{\beta} = \frac{\sum_{j \in N} \beta_h}{N}$ as the average
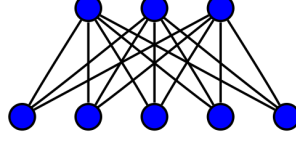
Figure 3: A complete bipartite graph $K_{5,3}$.

intrinsic valuations. Using (5) in Corollary 2, the increment of effort contribution is

$$\Delta \mathbf{x} = \delta\theta \begin{pmatrix} \mathbf{0} & J_{MN} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \alpha_{\mathbf{M}} \\ \beta_{\mathbf{N}} \end{pmatrix} + \delta^2\theta \begin{pmatrix} NI_M & \mathbf{0} \\ \mathbf{0} & MJ_{NN} \end{pmatrix} \begin{pmatrix} \alpha_{\mathbf{M}} \\ \beta_{\mathbf{N}} \end{pmatrix} + \mathcal{O}(\delta^3),$$

and the aggregate increment is

$$\Delta x = \delta\theta M\beta_N + \delta^2\theta \left(N\alpha_M + MN\beta_N \right) + \mathcal{O}(\delta^3) = \delta\theta MN\bar{\beta} + \delta^2\theta MN(\bar{\alpha} + N\bar{\beta}) + \mathcal{O}(\delta^3).$$

When δ is relatively small, the dominant term is $\delta\theta MN\bar{\beta}$.

Similarly, if players of group N move first, the aggregate difference is

$$\Delta' x = \delta\theta MN\bar{\alpha} + \delta^2\theta MN(\bar{\beta} + M\bar{\alpha}) + \mathcal{O}(\delta^3).$$

Corollary 3. *In a complete bipartite graph, when δ is small the group with a lower average intrinsic valuation should be the leader group.*

Again, this corollary provides a recommendation that is exactly the opposite of that in Zhou and Chen (2013) when we eliminate incomplete information. Moreover, when $\bar{\alpha} = \bar{\beta}$, the first term, which is dominant, is the same: $\delta\theta MN\bar{\beta} = \delta\theta MN\bar{\alpha}$. The second term $\delta^2\theta MN(\bar{\alpha} + N\bar{\beta})$ in Δx is greater than the term $\delta^2\theta MN(\bar{\beta} + M\bar{\alpha})$ in $\Delta' x$ if and only if $N > M$. This suggests that the leader group should be of a smaller size.

Corollary 4. *If δ is small and the average characteristic numbers are the same across two sides, i.e., $\bar{\alpha} = \bar{\beta}$, the group with a smaller size should be the leader group.*

The above corollary provides a theoretical ground for why focus groups and fashion influentials are typically small. This result holds both with and without incomplete information, thereby suggesting some form of robustness. In particular, for a star (hub-spoke) network, the center or the star player, or the hub, should be the unique leader.

7 Conclusions

Motivated by the targeted information release practice, this paper studies how network designers identify influential players that serve as the seeds for information dissemination. We build upon our earlier work Zhou and Chen (2013) in which the players in a network game with complete information move sequentially. We consider a two-stage version of the same setup and introduce information asymmetry amongst players. In such a scenario, the leaders' actions serve as effective signaling instruments that convey their private information to the followers. We show that this signaling incentive leads to first-order contribution improvement; this stands in strict contrast with the second-order improvement in Zhou and Chen (2013), irrespective of whether it is a restricted two-stage game or a general hierarchy design problem. This subsequently leads us to prove that the group selection problem is substantially different from the weighted maximum-cut problem.

We propose a novel index for the key leader selection and show that it can be substantially different from the indices in Ballester et al. (2006) and Zhou and Chen (2013). We also show that in undirected graphs, there is a swap between leader and follower groups when incomplete information is introduced. Furthermore, for complete graphs, the pecking order suggests that the leader should be selected by the *ascending* order of intrinsic valuations. This is exactly the opposite criterion used in the case with complete information (Zhou and Chen (2013)). Thus, while Zhou and Chen (2013) study a context in which the industry is mature, the current paper shows that a fast changing industry, for which informational issues are critical, may demand substantially different design principles. We also examine the out-tree hierarchical structure that is representative in the organization structure context. We prove that two common approaches – top-down and bottom-up – are both generically suboptimal despite their popularity, as the key leader turns out to stay around the middle but not exactly the center of the network. Overall, the analysis provides a theoretical ground for some marketing practitioners' recommendations of influencer choices or opinion leaders, and complements the existing literature of network economics by incorporating both signaling and strategic complementarity in a sequential-move context.

A Appendix. Proofs

Proof of Theorem 1. In the second-stage, after observing $x_{\mathbf{L}}$ the followers in F form a common belief $\hat{\theta}$. Accordingly, their effort choices are:

$$x_{\mathbf{F}}(x_{\mathbf{L}}) = [1 - \delta G_{FF}]^{-1}(\hat{\theta}\alpha_{\mathbf{F}} + \delta G_{FL}x_{\mathbf{L}}) = U(\hat{\theta}\alpha_{\mathbf{F}} + \delta G_{FL}x_{\mathbf{L}}).$$

The individual effort by player j can be expressed as $x_j(x_{\mathbf{L}}) = \sum_{k\in F} U_{jk}(\hat{\theta}\alpha_k + \delta \sum_{l\in L} g_{kl}x_l)$, $\forall j \in F$. We can then go backwards and characterize the equilibrium outcomes. Going backwards, in the first stage, the players i's payoff is given by

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

Recalling $T = G_{LL} + \delta G_{LF}UG_{FL}$, we can rewrite the above equation as follows:

$$u_i = \left(\theta\alpha_i + \delta\hat{\theta}\sum_{j\in F}\sum_{k\in F} g_{ij}U_{jk}\alpha_k \right) x_i - \frac{1}{2}x_i^2 + \delta \sum_{j\in L} T_{ij}x_i x_j, \ \forall i \in L. \tag{14}$$

Assuming linear strategies $x_i(\theta) = \kappa_i\theta, i \in L$ and the pessimistic belief, the equilibrium condition for each player i in L is the following:

$$\kappa_i\theta \in \arg\max_{x_i\in R} \left\{ \begin{array}{c} \left(\theta\alpha_i + \delta\min(\theta, \frac{x_i}{\kappa_i})\sum_{j\in F}\sum_{k\in F} g_{ij}U_{jk}\alpha_k \right) x_i - \frac{1}{2}x_i^2 \\ +\delta\sum_{j\in L,j\neq i} T_{ij}x_i\kappa_j\theta + \delta T_{ii}x_i^2 \end{array} \right\}. \tag{15}$$

Notice that here we have plugged in $x_j = \kappa_j\theta$ for $j \in L\backslash\{i\}$; thus, $\hat{\theta} = \min(\theta, \frac{x_i}{\kappa_i})$. Observe that $\hat{\theta}$ is not differentiable at $x_i = \kappa_i\theta$. Therefore, the first-order condition for (15) comprises two parts:

$$\left(\theta\alpha_i + \delta\theta\sum_{j\in F}\sum_{k\in F} g_{ij}U_{jk}\alpha_k \right) - \kappa_i\theta + \delta\sum_{j\in L,j\neq i} T_{ij}\kappa_j\theta + 2\delta T_{ii}\kappa_i\theta \ \leq \ 0.$$

$$\begin{array}{c} \left(\theta\alpha_i + \delta\theta\sum_{j\in F}\sum_{k\in F} g_{ij}U_{jk}\alpha_k \right) + \delta\frac{1}{\kappa_i}\sum_{j\in F}\sum_{k\in F} g_{ij}U_{jk}\alpha_k\kappa_i\theta \\ -\kappa_i\theta + \delta\sum_{j\in L,j\neq i} T_{ij}\kappa_j\theta + 2\delta T_{ii}\kappa_i\theta \end{array} \ \geq \ 0.$$

Rewriting these inequalities using matrices, we obtain that:

$$\alpha_{\mathbf{L}} + \delta G_{LF} U \alpha_{\mathbf{F}} - \kappa_{\mathbf{L}} + \delta(T + T^D)\kappa_{\mathbf{L}} \leq \mathbf{0} \leq \alpha_{\mathbf{L}} + 2\delta G_{LF} U \alpha_{\mathbf{F}} - \kappa_{\mathbf{L}} + \delta(T + T^D)\kappa_{\mathbf{L}}. \quad (16)$$

The above equation leads to multiple equilibria. We will choose the largest solution to (16), which gives the result in Theorem 1. Actually, the smallest one corresponds to the case with complete information characterized in Zhou and Chen (2013). In other words, if we make the first inequality in (16) bind, then $\kappa_{\mathbf{L}} = [1 - \delta(T + T^D)]^{-1}(\alpha_{\mathbf{L}} + \delta G_{LF} U \alpha_{\mathbf{F}})$. This gives exactly the equilibrium effort in the unique subgame NE of the two stage game in Proposition 1. $\square$

Proof of Corollary 2. In the sequential-move game with complete information, Proposition 1 shows that the equilibrium efforts are characterized by the matrix $\mathbf{S} \begin{pmatrix} \theta\alpha_{\mathbf{L}} \\ \theta\alpha_{\mathbf{F}} \end{pmatrix}$, and the NE with complete information is $\mathbf{M} \begin{pmatrix} \theta\alpha_{\mathbf{L}} \\ \theta\alpha_{\mathbf{F}} \end{pmatrix}$, where

$$\mathbf{M} = [1 - \delta G]^{-1} = \begin{pmatrix} [1 - \delta(T + \mathbf{0}T^D)]^{-1} & \mathbf{1}\delta[1 - \delta(T + \mathbf{0}T^D)]^{-1} G_{LF} U \\ \delta U G_{FL}[1 - \delta(T + \mathbf{0}T^D)]^{-1} & U + \mathbf{1}\delta^2 U G_{FL}[1 - \delta(T + \mathbf{0}T^D)]^{-1} G_{LF} U \end{pmatrix}.$$

Zhou and Chen (2013) have calculated the difference between $\mathbf{S}$ and $\mathbf{M}$ as

$$\mathbf{S} - \mathbf{M} = \delta^2 \begin{pmatrix} (G_{LF}G_{FL})^D & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} + \mathcal{O}(\delta^3).$$

Note that $\hat{\mathbf{S}} - \mathbf{M} = (\hat{\mathbf{S}} - \mathbf{S}) + (\mathbf{S} - \mathbf{M})$, so it's suffice to show that the difference between $\hat{\mathbf{S}}$ and $\mathbf{S}$ has the following expression.

$$\begin{aligned}
(\hat{\mathbf{S}} - \mathbf{S}) &= \begin{pmatrix} \mathbf{0} & \delta[1 - \delta(T + T^D)]^{-1} G_{LF} U \\ \mathbf{0} & \delta^2 U G_{FL}[1 - \delta(T + T^D)]^{-1} G_{LF} U \end{pmatrix} \\
&= \delta \begin{pmatrix} \mathbf{0} & G_{LF} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} + \delta^2 \begin{pmatrix} \mathbf{0} & G_{LL}G_{LF} + G_{LF}G_{FF} \\ \mathbf{0} & G_{FL}G_{LF} \end{pmatrix} + \mathcal{O}(\delta^3).
\end{aligned}$$

To this end, note that $U = [I - \delta G_{FF}]^{-1} = I + \delta G_{FF} + \mathcal{O}(\delta^2)$. Also $T = G_{LL} + \delta G_{LF} U G_{FL} = G_{LL} + \mathcal{O}(\delta)$, and $T^D = (G_{LL} + \mathcal{O}(\delta))^D = \mathcal{O}(\delta)$ because $G_{LL}^D = 0$ (the diagonal entries of G

are zeros). Combing these equations yields:

$$
\begin{aligned}
\delta \left[I - \delta(T + T^D) \right]^{-1} G_{LF} U &= \delta \left[I + \delta(T + T^D) + \mathcal{O}(\delta^2) \right] G_{LF} U \\
&= \delta \left[I + \delta G_{LL} + \mathcal{O}(\delta^2) \right] G_{LF} (I + \delta G_{FF} + \mathcal{O}(\delta^2)) + \mathcal{O}(\delta^3) \\
&= \delta G_{LF} + \delta^2 (G_{LL} G_{LF} + G_{LF} G_{FF}) + \mathcal{O}(\delta^3).
\end{aligned}
$$

Similarly,

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

□

Proof of Proposition 2. Comparing (6) and (7) gives us the relation between player i's equilibrium contribution under the two scenarios:

$$
\begin{aligned}
x_i(\theta) &= \frac{2(\alpha_i + \delta \left\langle \beta_i, \ (\mathbf{I} - \delta G_{-i})^{-1} \alpha_{-i} \right\rangle) - \alpha_i}{1 - 2\delta^2 \left\langle \beta_i, \ (\mathbf{I} - \delta G_{-i})^{-1} \beta_i \right\rangle} \theta = \frac{\frac{2b_i(G,\delta,\alpha)}{m_{ii}} - \alpha_i}{1 - 2(1 - \frac{1}{m_{ii}})} \theta \qquad (17) \\
&= \frac{2b_i(G, \delta, \alpha) - \alpha_i m_{ii}}{2 - m_{ii}} \theta
\end{aligned}
$$

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 (8).

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.,

$$
x^*_{-i}(x_i) = \mathbf{b}(G_{-i}, \delta, \alpha_{-i}\theta + \delta x_i \beta_i) = (\mathbf{I} - \delta G_{-i})^{-1} (\alpha_{-i}\theta + \delta x_i \beta_i). \qquad (18)
$$

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

$$
\Delta x^*_{-i}(x_i) = (\mathbf{I} - \delta G_{-i})^{-1} \delta \Delta x_i \beta_i \qquad (19)
$$

According to (8), 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

$$
(1 + \sum_{j \neq i} \frac{m_{ij}}{m_{ii}}) \Delta x_i = \frac{\sum_{k=1}^{N} m_{ik}}{m_{ii}} \Delta x_i = \frac{b_i(G, \delta, 1)}{m_{ii}} \Delta x_i. \qquad (20)
$$

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

29

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

$$\frac{b_i(G, \delta, 1)}{m_{ii}} \left(\frac{2b_i(G, \delta, \alpha) - \alpha_i m_{ii}}{2 - m_{ii}}\theta - b_i(G, \delta, \alpha)\theta \right) = \frac{b_i(G, \delta, 1)(b_i(G, \delta, \alpha) - \alpha_i)}{2 - m_{ii}}\theta. \quad (21)$$

This then leads to the S-index specified in the proposition. $\square$

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.

Bala, V. and S. Goyal (1998). Learning from neighbours. *The Review of Economic Studies 65*(3), 595–621.

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.

Banerjee, A. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics 107*(3), 797–817.

Bang-Jensen, J. and G. Z. Gutin (2008). *Digraphs: theory, algorithms and applications.* Springer.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *The Journal of Political Economy 100*(5), 992–1026.

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

Campbell, A., D. Mayzlin, and J. Shin (2012). A model of buzz and advertising. Summer Institute in Competitive Strategy (SICS) 2012, University of California at Berkeley.

Candogan, O., K. Bimpikis, and A. Ozdagla (2012). Optimal pricing in networks with externalities. Forthcoming in Operations Research.

Chen, W., Y. Wang, and S. Yang (2009). Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 199–208. ACM.

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.

Coulter, R. A., L. L. Price, and L. Feick (2003). Rethinking the origins of involvement and brand commitment: insights from postsocialist central europe. *Journal of Consumer Research 30*(2), 151–169.

Demange, G. (2004). On group stability in hierarchies and networks. *Journal of Political Economy 112*(4), 754–778.

Dye, R. (2000). The buzz on buzz. *Harvard Business Review 78*(6), 139–146.

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

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

Gladwell, M. (2006). *The tipping point: How little things can make a big difference.* Little, Brown.

Jackson, M. (2008). *Social and economic networks.* New Jersey: Princeton University Press.

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

Kempe, D., J. Kleinberg, and É. Tardos (2005). Influential nodes in a diffusion model for social networks. In *Automata, languages and programming*, pp. 1127–1138. Springer.

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

Van den Bulte, C. and Y. V. Joshi (2007). New product diffusion with influentials and imitators. *Marketing Science 26*(3), 400–421.

Watts, D. and P. S. Dodds (2007). Influentials, networks, and public opinion formation. *Journal of Consumer Research 34*(4), 441–458.

Weimann, G. (1994). *The influentials: People who influence people.* SUNY Press.

Zhou, J. and Y.-J. Chen (2013). The benefit of sequentiality in social networks. Working paper, University of California at Berkeley.