Network multipliers and the optimality of indirect communication

Andrea Galeotti* Sanjeev Goyal†

September 28, 2007

Abstract

We study the problem of a firm $\mathcal{M}$ which wishes to inform a community of individuals about its product. Information travels within the community because of the social interactions between individuals. Our interest is in understanding how the firm can incorporate the network of social interactions in the design of its communication strategy.

We study a model of undirected networks and start by showing that social interactions appear in the payoff of the firm in the form of a network multiplier. We establish that the network multiplier is an increasing function of both the mean and the variance in the distribution of connections of the network. This implies in particular that denser and more dispersed degree distributions are better for the firm. We then show that the degree distribution of the neighbor first order dominates the degree distribution of a node at large and so it is always better for a firm to use *indirect* communication, i.e., *viz.* picking the neighbor of a node rather than a node itself as the target of communication. Finally, we show that the advantages of indirect communication are increasing with dispersion in the degree distribution.

Keywords: Social interactions, network multiplier, targeting neighbors.

*Department of Economics, University of Essex. E-mail: `agaleo@essex.ac.uk`

†Faculty of Economics, University of Cambridge. E-mail: `sg472@econ.cam.ac.uk`

This note arose out of comments made by Fernando Vega-Redondo and we are grateful to him. We thank the ESRC for financial support via Grant RES-000-22-1901.

1 Introduction

We study the problem of a firm $\mathcal{M}$ which wants to inform a community of individuals about its product. Information travels within the community because of the social interactions between individuals. Our interest is in understanding how the firm can incorporate the social interactions in the design of its communication strategy. The social interactions are modeled as a network and our question can therefore be rephrased as saying: how does the structure of the network affect the optimal strategy of the firm?

We show that social interactions are captured in a network multiplier which is a simple function of the mean and the variance of the distribution of connections of the network. The changes in social interactions then find expression in changes in the network multiplier. The optimal communication strategy of the firm equates the marginal costs of communication to the network multiplier. This allows us to develop a number of results on the relation between social interactions and optimal influence strategies of the firm.

Our first result is that the network multiplier is an increasing function of both the mean and the variance in the distribution of connections of the network. This implies in particular that denser and more dispersed degree distributions lead to higher action from the firm and also raise the profits of the firm. We then show that the degree distribution of the neighbor first order dominates the degree distribution of a node at large. Hence, it is always better for a firm to pick the neighbor of a node than a node itself as the target of communication. Moreover, the advantages of choosing a neighbor are increasing with the dispersion in the degree distribution.

The role of optimal strategies in the face of social interactions has been studied in a number of disciplines over the years. For a discussion of this work and its relation to the economic approach, see our earlier paper Galeotti and Goyal (2007). This note builds on the ideas developed in Goyal and Galeotti (2007), which studies the optimal diffusion strategy of a firm facing a set of consumers with social interactions. In that paper we focus on *directed* links of information transmission: in such a setting person i may get information from person j without the converse flow of information. In this note our interest is in *undirected* social interactions: this reflects a situation in which interaction is mutual and a link implies two-way flow of information. Two-way flow of information is descriptively appealing in some contexts and also has distinct technical implications. In particular, in a context with random directed links the distribution on in-degrees of the neighbor is the same as the distribution of in-degrees of a node at large. By contrast, in an undirected links network the fact that j is the neighbor of i (rather than not) means that it is more likely that j has more (undirected) links than a node taken at random. Consequently, the degree distribution of a neighbor is different from the unconditional degree distribution of nodes.

This note explores the implications of this distinction for the network multiplier and the optimal communication of the firm. In particular, our result that more dispersed degree distributions are better for communication arises entirely due to this distinction. Similarly, our result of the optimality of indirect communication and that its increasing profitability under mean preserving spreads in the degree distribution is founded on this distinction as well.

The importance of this distinction has been hinted at in the earlier literature; for instance, Krackhardt (1996) explores the attraction of targeting a node as against targeting the neighbor of a node and develops some numerical examples which show why the latter is better. Our analysis provides a general analytical result that the degree distribution of the neighbor first order dominates the degree distribution of a node at large and therefore shows that, for a general class of random networks, it is *always* better to target the neighbor of a node rather than a node picked at random.

Section 2 lays out the basic model and section 3 presents the results.

2 Model

There is a countable infinity of individuals and we define the set of individuals as $\mathcal{N} = \{1, 2, ..., \infty\}$. Each individual i is located at a node i of an *undirected* infinite network g. The notation $g_{i,j} = 1$ signifies that a link exists between i and j, while $g_{i,j} = 0$ means that no link exists between i and j. The set of direct neighbors of individual i is $\mathcal{N}_i(g) = \{j \in \mathcal{N} : g_{i,j} = 1\}$, and $k_i(g) = |\mathcal{N}_i(g)|$ is the degree of i in g. $\mathcal{N}_i^r(g)$ denotes the r-neighborhood of individual i in g; that is all the individuals that can be reached from i by paths of length no more than r.

The strategy of $\mathcal{M}$ is to directly informed $T \in \mathcal{N}$ individuals in the network. This comes at a cost $C(T)$; we assume that $C(\cdot)$ is increasing, convex, $C(0) = 0$ and $\lim_{T\to\infty} C(T) = \infty$. Suppose $\mathcal{M}$ informs individual i, then i will report the information to his friends, i.e., $\forall j \in \mathcal{N}_i$, and, in turn, these individuals will transmit further the information to their neighbors, i.e., $\forall j \in \mathcal{N}_i^2 \cap \mathcal{N}_i$, and so on, till the information becomes obsolete. We assume that the information becomes obsolete after it travels r steps and we term r as the radius of information.

For a given r, the choice of $\mathcal{M}$ will depend on her knowledge about the architecture of the network. We consider a situation in which $\mathcal{M}$ does not know the architecture of the network, but she only holds some beliefs about it. Specifically, the probability structure is formalized as follows. The degree of an arbitrary individual i in the network, k_i, is a random variable defined on $\mathcal{I} = \{1, 2, ...\bar{k}\}$, where $\bar{k} < \infty$.[1] The

[1] The infinite network taken together with a finite $\bar{k}$ implies that two randomly picked nodes will have a common neighbor with probability 0; we will use this observation in what follows, and assume that the intersection of the neighborhoods of randomly picked nodes is empty.

distribution of k_i is $P : \mathcal{I} \rightarrow [0,1]$. We shall refer to P as the degree distribution of the network, $z_P = \sum_{k\in\mathcal{I}} kP(k)$ is the average degree of the network, and $\sigma_P = \sum_{k\in\mathcal{I}} P(k)(k - z_P)^2$ denotes the variance of P.

We assume that the degrees of any two neighboring nodes are stochastically independent. This would be the case if, for example, the network has been generated via the classical Erdos and Renyi (1960) type of link formation process.[2] The probability distribution of a randomly selected neighbor is the conditional degree distribution, which is given by:

$$\tilde{P}(k) = \frac{kP(k)}{z_P}.$$

We denote by $\tilde{z}_P$ the average degree under $\tilde{P}$; this is the average degree of a neighbor. It is easy to verify that:

$$\tilde{z}_P = \sum_{k\in\mathcal{I}} \tilde{P}(k)k = \frac{1}{z_P}\sum_{k\in\mathcal{I}} P(k)k^2 = \frac{\sigma_P + z_P^2}{z_P}. \tag{1}$$

We can now define the expected payoffs to $\mathcal{M}$. Given the radius of information transmission, r, and the degree distribution P, if the monopolist directly informed T individuals, his expected payoffs are:

$$\Pi(T|r,P) = T\left[1 + \sum_{s=1}^{r}\sum P(k)k\left[\sum \tilde{P}(l)(l-1)\right]^{s-1}\right] - C(T).$$

These payoffs can be simplified and rewritten as follows:

$$\Pi(T|r,P) = T\left[1 + z_P\sum_{s=1}^{r}(\tilde{z}_P - 1)^{s-1}\right] - C(T).$$

Note that since $k \geq 1$ then $z_P \geq 1$ and $\tilde{z}_P \geq 1$. The optimal strategy of $\mathcal{M}$ is

$$T^*(r,P) = \arg\max_{T\in\mathcal{N}} \Pi(T|r,P),$$

and we denote by $\Pi(T^*(r,P)|r,P)$, the expected payoffs obtained under the optimal strategy.

[2]In the modern literature of complex networks alternative mechanisms have been proposed which generate networks in which neighboring nodes have independent degrees; see Vega-Redondo (2007) for a survey

3 Analysis

We define the network multiplier as the (expected) number of individuals that eventually obtain the information from an individual directly informed by $\mathcal{M}$: Formally, this is written as:

$$\Psi(r,P) = E[\mathcal{N}^r(g)] = \sum_{s=1}^{r} \sum P(k)k \left[\sum \tilde{P}(l)(l-1)\right]^{s-1} = z_P \sum_{s=1}^{r} (\tilde{z}_P - 1)^{s-1} \quad (2)$$

We can now rewrite the expected payoffs in terms of the multiplier as:

$$\Pi(T|r,P) = T\left[1 + \Psi(r,P)\right] - C(T).$$

It now follows that $1 + \Psi(r,P)$ is the marginal return to $\mathcal{M}$ from directly informing an additional agent in the network. Since $\Psi(r,P)$ is positive and bounded for finite r, given the assumptions on $C(\cdot)$, it follows that $T^*(r,P)$ is finite. Taking T to be a continuous variable, it follows that $T^*(r,P)$ solves:

$$1 + \Psi(r,P) = C'(T^*(r,P)). \quad (3)$$

To illustrate how the degree distribution shapes the network multiplier let us consider a few simple examples.

Example 1: First suppose that information only travels one step, i.e., $r = 1$. Then, the network multiplier is:

$$\Psi(1,P) = z_P.$$

In other words, the higher is the average degree of the network the higher is the network multiplier. More generally, this means that if the degree distribution P' first order dominates the degree distribution P, then $\Psi(1,P') \geq \Psi(1,P)$. An increase in the average degree raises the marginal returns and from equation (3) we know this means that the optimal number of individuals directly informed by $\mathcal{M}$ also increases. We also observe that since expected payoffs are larger under a first order shift in degree distribution even if the strategy is kept constant, the same must hold if the firm optimally adjusts its strategy. So, both optimal actions and profits increase with a first order shift in the degree distribution. We conclude by noting that changes in the variance of the degree distribution have no bearing on the network multiplier if $r = 1$. ■

To illustrate the role of the variance of the degree distribution, we briefly discuss an example where information travels two links.

Example 2: Suppose now that $r = 2$; in this case the network multiplier is

$$\Psi(2, P) = \sum P(k)k \sum_{s=1}^{r=2} \left[\sum \tilde{P}(l)(l-1) \right]^{s-1} \tag{4}$$

Now we see that both the degree distribution as well as the conditional degree distribution come into play. Observe that a first order shift in both these distributions raises the network multiplier unambiguously and thus raises the marginal returns; from equation (3) this implies that the optimal strategy increases and it is immediate that the profits increase as well.[3] Let us next consider changes in the dispersion of the degree distribution. The simplest way to see the effect of dispersion is to look at a degree distribution P' which is a mean preserving spread of P. Combining equation (1) with (4) we immediately see that $\Psi(2, P') \geq \Psi(2, P)$. Thus a mean preserving spread in the degree distribution yields a higher network multiplier and this in turn means, following from our earlier arguments, an increase in the optimal action as well as profits of $\mathcal{M}$.

■

These examples allow us to state the following general result on the effects of radius of information transmission and the degree distribution on the network multiplier.

Proposition 3.1 *The network multiplier $\Psi(r, P)$ is:*

1. *increasing in the radius of information transmission, r;*
2. *increasing under a first order shift in the degree distribution P and the conditional degree distribution $\tilde{P}$;*
3. *increasing under a mean preserving spread of the degree distribution P.*

The proof is immediate from expression (2) and is omitted. Equipped with this result on the network multiplier, we can now establish a general result on optimal strategies and payoffs.

Proposition 3.2 *The optimal strategy $T^*(r, P)$ and the payoffs $\Pi(T^*(r, P)|r, P)$ of the firm $\mathcal{M}$ are increasing with the radius of information transmission r, increasing under a first order shift of the degree distribution P and the conditional degree distribution $\tilde{P}$, and also increasing under a mean preserving spread of the degree distribution P.*

The proof follows from Proposition 3.1 and equation (3) and is omitted.

[3] It is worth noting that a first order shift in P does not imply a first order shift in $\tilde{P}$. The following example taken from Galeotti et. al. (2006) illustrates this point. Consider two degree distributions P and P', where P' assigns one half probability to degrees 2 and 10 each, while distribution P assigns one half probability to degrees 8 or 10 each. Clearly P FOSD P'. As mentioned above, when neighboring degrees are independent, the probability of having a link with a node is (at least roughly, depending on the process) proportional to the degree of that node, so that for all k, $P(k'|k) = k'P(k')/\sum P(l)l$. Let $\tilde{P}(k')$ be the neighbor's degree distribution. Under $\tilde{P}'$, the probability that a neighbor has degree 10 is 5/6, while under $\tilde{P}$, the same probability is 5/9. Thus, $\tilde{P}$ does not FOSD $\tilde{P}'$.

3.1 The optimality of indirect communication

So far we have studied a strategy of direct communication: $\mathcal{M}$ picks T individuals in the network and informs them. Suppose instead that $\mathcal{M}$ picks T people, but instead it informs a neighbor of each of these T people. We shall refer to this strategy of sending information to the neighbors of nodes as *indirect communication.*

The network multiplier with indirect communication is given by

$$\tilde{\Psi}(r,P) = \sum_{s=1}^{r} \sum \tilde{P}(k)k \left[\sum \tilde{P}(l)(l-1)\right]^{s-1} = \tilde{z}_P \sum_{s=1}^{r} (\tilde{z}_P - 1)^{s-1}$$

Therefore the expected payoffs to $\mathcal{M}$ from an indirect communication strategy of informing T individuals is:[4]

$$\tilde{\Pi}(T|P,r) = T[1 + \tilde{\Psi}(r,P)] - C(T)$$

The distinction between payoffs under direct and indirect communication rests on the difference between z_P and $\tilde{z}_P$. Indeed, let this difference be $\Delta(P) = \tilde{\Pi}(T|P,r) - \Pi(T|P,r)$, then

$$\Delta = T(\tilde{z}_P - z_P) \sum_{s=1}^{r} (\tilde{z}_P - 1)^{s-1}. \qquad (5)$$

We now establish a simple general result on the relation between $\tilde{P}$ and P which allows us to rank the two averages unambiguously.

Proposition 3.3 *The conditional degree distribution $\tilde{P}$ first order stochastically dominates the unconditional degree distribution P.*

Proof: Recall that $\tilde{P}(k) = kP(k)/z_P$, where z_P is the average degree under P. Define $\mathcal{P}(x) = \sum_{k=1}^{k=x} P(k)$. Note that by definition $\mathcal{P}(\bar{k}) = \tilde{\mathcal{P}}(\bar{k}) = 1$. We wish to show that for all $x \leq \bar{k}$, $\mathcal{P}(x) \geq \tilde{\mathcal{P}}(x)$. From the definition of $\tilde{P}$ it follows that $\tilde{\mathcal{P}}(x) \leq \mathcal{P}(x)$, for all $x \leq z_P$. Suppose that $\tilde{\mathcal{P}}(y) > \mathcal{P}(y)$ for some $z_P < y < \bar{k}$ and let $A = \tilde{\mathcal{P}}(y) - \mathcal{P}(y) > 0$. Now observe that for any $y' \geq y$,

$$\begin{aligned}
\tilde{\mathcal{P}}(y') - \mathcal{P}(y') &= \sum_{k=1}^{y'} P(k)\left[\frac{k}{z_P} - 1\right] \\
&= \sum_{k=1}^{y} P(k)\left[\frac{k}{z_P} - 1\right] + \sum_{k=y+1}^{y'} P(k)\left[\frac{k}{z_P} - 1\right] \\
&\geq A
\end{aligned}$$

[4] Here we are again exploiting the infinite nodes and finite maximum degree assumptions to suppose that neighborhoods of neighbors have an empty intersection.

This however contradicts $\mathcal{P}(\bar{k}) = \tilde{\mathcal{P}}(\bar{k}) = 1$. ■

This first order dominance result immediately implies that $\tilde{z}_P \geq z_P$, and we can now state the following result as a corollary.

Proposition 3.4 *Given any degree distribution P, indirect communication yields higher profits as compared to direct communication.*

We now turn to the effects of changing networks on the relative attractiveness of indirect communication. Observe that if the probability of any degree is 1 under P, then the conditional and the unconditional degree distributions coincide. So the profit differential arises when there is some variation in degrees. The following result builds on this intuition and shows that the difference in profits Δ is increasing with greater dispersion in the degree distribution.

Proposition 3.5 *The difference in profits between direct and indirect communication is increasing under a mean preserving spread in the degree distribution P.*

Proof: First observe that variance is increasing under a mean preserving spread of the degree distribution. We now use equations (1) and (5) to show that

$$\Delta = T\left(\frac{\sigma_P + z_P^2}{z_P} - z_P\right)\sum_{s=1}^{r}(\tilde{z}_P - 1)^{s-1}.$$

This in turn allows us to conclude that the difference in profits between direct and indirect profits is increasing under a mean preserving spread of the degree distribution P. ■

References

[1] Galeotti, A., Goyal, S., Jackson, M., Vega-Redondo, F., Yariv, L. (2006), Network Games. Mimeo, Essex and Caltech.

[2] Galeotti, A. and S. Goyal (2007), A theory of strategic diffusion. Mimeo, Essex and Cambridge.

[3] Krackhardt, D. (1996), Structural Leverage in Marketing. In Dawn Iacobucci (Ed.) Networks in Marketing, Sage, Thousand Oacks CA. pp. 50-59.

[4] Vega-Redondo, F. (2007), *Complex social networks*, Econometric Society Monographs. Cambridge University Press, Cambridge.