The Law of the Few

Andrea Galeotti* Sanjeev Goyal†

This version: April 25, 2008
First Draft: August 2007.

Abstract

The *law of the few* refers to the following empirical phenomenon: in social groups a great proportion of individuals get most of their information from a very small subset of the group. This small set has many more connections than the average of the group. In some situations these few individuals acquire information as well as they convey it to others, while in other settings they act as pure connectors. Moreover, there are no significant observable differences between this select few and the rest of the social group. We are led to ask: can the law of the few be a result of strategic interaction among identical individuals?

We develop a model in which players personally acquire information as well as form connections with others to access the information which they have acquired. Our main finding is that the the 'law of the few' is a robust equilibrium outcome in this model.

JEL classification: C72, D00, D83, H41, O31.

Keywords: Opinion leaders, social influencers, connectors, mavens, network formation, public goods.

*Department of Economics, University of Essex. Email: agaleo@essex.ac.uk

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

We thank Bahskar Dutta, Francis Bloch, Yann Bramoulle, Myeonghwan Cho, Marco van der Leij and Brian Rogers for suggestions which have significantly improved the paper. We also thank seminar participants at Cambridge, Essex, Queen Mary University, Michigan, Microsoft (Cambridge) and the conference on "The Formation of Social Networks", Paris 2007, for useful comments.

1 Introduction

The classic early studies by Lazarsfeld, Berelson, and Gaudet (1948) and Katz and Lazersfeld (1955) investigate the impact of personal contacts, media advertising, newspaper advertising and alike on voting decisions and marketing decisions such as food brand choices, fashion changes and movies. They conceptualize interpersonal relations as potential networks of communication and define an opinion leader as a group member playing a key communication role–opinion leaders engage in information seeking and convey information to others. They find that the impact of personal contacts in shaping individuals' decisions dominates that of other sources and only 20% of their sample –800 women in Decatur, Illinois– are opinion leaders. Moreover, the observable economic and demographic characteristics of the opinion leaders appear to be much like those of other individuals.

Over the years, empirical research across a range of subjects–which include political science, sociology, management, marketing and computer science– has confirmed and generalized the scope of these findings.[1] For example, Cross *et al.* (2001) study the patterns of information flow between 40 managers working in Fortune 500 firms.[2] They find that a vast majority of the managers – about 85% – obtain critical information via personal communication with few others, and that this is significantly more than other sources of information. They also provide evidence for a very unequal distribution of connections in the social communication network within organizations. These findings extend to the pattern of information acquisition and exchange in virtual communities such as social networking websites and peer-to-peer networks.[3]

[1]In marketing, the influential study by Feick and Price (1987) introduces the concept of market maven: a person who is knowledgeable across a range of products. Their study focuses on food, household products, nonprescription drugs and beauty products and their sample is of roughly 1400 individuals. Market mavens are about 25% of the sample, they engage in market information seeking more intensively than others, they frequently convey information to others and their demographic characteristics appear to be much like those of other consumers. For more recent work on information collection and information transmission in the marketplace see Wiedman et al. (2001), Williams and Slama (1995) and Geisser and Edison (2005). The recent work in political science arrives at similar conclusions, see e.g., Beck *et al* (2002), and Huckfeldt, Johnson and Sprague (2004).

[2]Many other studies in the management literature arrive at similar conclusions. We refer to Cross and Parker (2004) for numerous case studies.

[3]Zhang, Ackerman and Adamic (2007) study the patterns of communication in the Java Forum: an on-line community of users who ask and respond to queries concerning the programme Java. They construct a

The Law of the Few subsumes these wide ranging empirical regularities in the pattern of information acquisition and informal social organization: in social groups there is a large majority of individuals who get a lot of critical information needed for their decisions from a very small subset of the group–the *influencers*. Social influencers have many more connections than the average of the group, in some instances they acquire personally and convey it to others most of the information–*mavens* and *opinion leaders*– while in other settings they act as pure communicators–*connectors*. There is some difference in observable economic and demographic characteristics between social influencers and the others, but this difference is slight.

In light of these findings, it is important to understand whether the law of the few can arise among identical individuals. To investigate this question, we develop a model with two key ingredients: individuals can choose to personally acquire information as well as form connections with others to access the information which they acquire. The incentives to acquire information and to form connections depend on the relative costs of doing so. Moreover, there is the issue of coordination: if no one else acquires any information then a player may have no choice but to acquire information himself.

We start with a basic model in which information can be either personally acquired or accessed from someone who has personally acquired it. In this model the main result is that, if the costs of forming a link are lower than the costs of directly acquiring information then, every (strict) equilibrium exhibits the law of the few (Propositions 3.2-3.3). In particular, we show that every strict equilibrium network will have the *core periphery architecture*, as depicted in Figure 1,[4] with all players in the core exerting positive effort while all peripheral

directed network using the information on queries and responses: there is a link from user A to user B if A asked a question to which B responded. They locate 13739 users in all and find that about 55% of the users only ask questions and provide no responses at all, 12% of the users both ask and answer queries, and about 13% of the users provide answers but do not have any queries. Adar and Huberman (2000) reports similar finding in the well known peer-to-peer decentralized network Gnutella.

[4]The notion of core periphery communication networks is common in social network analysis. Core periphery networks exhibit: (a) very unequal distribution of connections, (b) short distances between nodes and (c) high clustering. These are all empirically observed features of socially generated networks. See Cross et al. (2001) for examples of this pattern of social communication structure in small groups. See Jackson and Rogers (2007) for a summary of the key empirical regularities shared by large social networks such as e-mail networks, World Wide Web, phone-call networks. Formal definitions of networks are provided in section 2.

players choose zero effort and form links with each of the players in the core.

In actual practice, we receive information from friends and colleagues which they have themselves received from their friends. This idea leads us to extend the basic model to allow for indirect information transmission. Our key finding is that information spillovers gives rise to a new "type" of influential agent: *the connector* (Proposition 4.1). A connector is someone whose primary role is that of an 'informediary': he typically has many more links then the average person and he connects individuals who personally acquire information – the mavens – with other mavens or with individuals who personally acquire no information.[5] Proposition 4.2 shows that a 'small' amount of ex-ante heterogeneity in costs of acquiring information or forming links implies that all information is acquired by a single individual and there is a single connector who communicates this information to the rest of the population. This yields the law of the few in a very strong form.

In the results discussed so far, we assume that the costs of acquiring information are linear in effort and that efforts at information acquisition are a continuous variable. We also study the case of convex costs of acquiring information and forming links, and of discrete actions. Our analysis, summarized in Propositions 5.1-5.3, shows that specialization in information acquisition and social communication, reflected in the law of the few, are a robust feature of equilibrium in these settings as well.

The principal contribution of our paper is a simple model of strategic investments in information acquisition and link formation. Our results show that a combination of simple economic forces – the relative costs of acquiring information *versus* the costs of forming links – and strategic interaction together provide a simple explanation for the empirically observed specialization in information acquisition and social communication.

Our paper is also a contribution to the theory of networks. We develop a game in which individuals choose investments in information acquisition as well as decide with whom to form connections with a view to accessing the information acquired by these individuals.

[5]See, for example, Cross and Prusak (2002) for an extensive discussion of the importance and empirical relevance of connectors in informal communication networks.

Our model combines the approach to link formation introduced in Bala and Goyal (2000) with the approach to the study of local public goods developed in Bramoulle and Kranton (2007).[6] A model in which both information acquisition as well as networks are endogenous yields new insights and resolves some open problems in each these earlier literatures. In particular, endogenizing the network formation yields clear cut predictions on aggregate information acquisition and the distribution of efforts (in addition to the pattern of links of course). By contrast, with exogenous networks, Bramoulle and Kranton (2007) show that multiple equilibria typically exist and these equilibria exhibit very different levels of aggregate information acquisition as well as distributions. Similarly, endogenizing the effort level yields us the distinction between information collectors (the mavens) and the communicators (the connectors). This distinction is important in actual practice and has implications for the structure of social communication, but is ruled out by assumption in earlier models of pure link formation, where everyone has the same amount of exogenously provided information.

A recent paper by Cabrales, Calvo-Armengol and Zenou (2007) studies private investments and network formation. We have a model in which individuals decide on individual specific links while they assume that investments in links are *not* individual specific. In their paper individuals are ex-ante heterogenous, while the focus here is on showing how social differentiation can arise in the absence of any ex-ante heterogeneity. Thus the methods of analysis and the results in the two papers are quite different.[7]

We conclude by relating our results to the literature on public goods. Non-rival information acquisition is an instance of the private provision of public goods. In the context of global public goods, Bergstrom, Blume and Varian (1986) showed, roughly speaking, that if the players have the same preferences then the contributors will be those with higher en-

[6]The theory of network formation as well as the theory of games played on fixed networks are both active fields of study currently. For a survey of this research, see Goyal (2007). There is a small body of work which combines network formation and play in games, see e.g., Goyal and Vega-Redondo (2005), Jackson and Watts (2002), Bramoulle *et. al.* (2004) and Calvo-Armengol and Zenou (2004). This paper tackles a very different type of question as compared to these other papers, and so a discussion of the relation between the papers is not very useful.

[7]In this paper, our interest is in situations where information sharing is a social activity and it is plausible to suppose that there are no prices or bargaining on the terms at which information is exchanged; for a study of situations in which players can charge prices or bargain for their information, see Cabrales and Gottardi (2007) and Cho (2007).

dowments. It is possible to interpret information as a global public good in the sense that everyone can access it; but the key difference is that access is a matter of individual choice, it is costly and may be indirect (and takes the form of bilateral connections). The convex costs model studied in section 5.2 makes assumptions on preferences and costs of effort which are similar to those in Bergstrom, Blume and Varian (1986). The findings with regard to the existence of mavens and connectors are, to the best of our knowledge, novel in this literature.

The rest of the paper is organized as follows. Section 2 develops the basic model with local information flow and section 3 analyzes this model. Section 4 considers a model with indirect information transmission. Section 5 considers three extensions, one, to allow for one-way information flow, two, to allow for convex costs of acquiring information and linking and three, to allow for discrete action games. Section 6 concludes. The appendix contains all the proofs.

2 Model

Let $N = \{1, 2, .., n\}$ with $n \geq 3$ be the set of players and let i and j by typical members of this set. Each player i chooses an effort $x_i \in X$ and a set of links which is represented as a (row) vector $g_i = (g_{i1}, ..., g_{ii-1}, g_{ii+1}, ..., g_{in})$, where $g_{ij} \in \{0, 1\}$, for each $j \in N \setminus \{i\}$. We will suppose that $X \in [0, +\infty)$.[8] We say that player i has a link with player j if $g_{ij} = 1$. A link between player i and j allows both players to access the effort exerted by the other player.[9] The set of strategies of player i is denoted by $S_i = X \times \mathcal{G}_i$. Define $S = S_1 \times ... \times S_n$ as the set of strategies of all players. A strategy profile $s = (x, g) \in S$ specifies the effort of each player, $x = (x_1, x_2, ..., x_n)$, and the network of relations $g = (g_1, g_2, ..., g_n)$. The network of relations g is a directed graph; let $\mathcal{G}$ be the set of all possible directed graphs on n vertices. Define $N^d(i; g) = \{j \in N : g_{ij} = 1\}$ as the set of players with whom i has formed a link. Let $\eta_i(g) = |N^d(i; g)|$.

[8] We have also studied a model of discrete action choice in which $X = \{0, 1\}$; Section 5.3 presents our analysis of a model with discrete actions. We find that our main results on equilibrium efforts and networks are robust to a change in the action sets.

[9] In Section 5.1 we discuss a model where a link formed by i with j allows only player i to access the effort of player j.

The closure of g is a non-directed network denoted $\bar{g} = cl(g)$, where $\bar{g}_{ij} = \max\{g_{ij}, g_{ji}\}$ for each i and j in N. In words, the closure of a directed network simply means replacing every directed edge of g by a non-directed one. Define $N(i; \bar{g}) = \{j \in N : \bar{g}_{ij} = 1\}$ as the set of players directly connected to i.

The payoffs to player i under strategy profile $s = (x, g)$ are:

$$\Pi_i(s) = f\left(x_i + \sum_{j \in N(i;\bar{g})} x_j\right) - cx_i - \eta_i(g)k, \tag{1}$$

where $c > 0$ reflects the cost of effort and $k > 0$ is the cost of linking with one other person.[10] We will assume that $f(y)$ is twice continuously differentiable, increasing, and strictly concave in y. To focus on interesting cases we will assume that $f(0) = 0$, $f'(0) > c$ and $\lim_{y\to\infty} f'(y) = z < c$. Under these assumptions there exists a number $\hat{y} > 0$ such that $\hat{y} = \arg\max_{y \in X} f(y) - cy$.

Remark 2.1 *For concreteness, we may think of the action x as reflecting draws from a distribution, e.g., the price distribution for a product. If the different draws are independent across individuals and players are interested in the lowest price, then the value of an additional draw, which is the change in the average value of the lowest order statistic, is positive but declining in the number of draws.*

Remark 2.2 *In the above model, we interpret k as the time and resource cost of forming and maintaining a social tie. An alternative interpretation of k is that it is a gift which the player forming the link makes to the person receiving the link. In this case, k becomes a transfer and the payoffs to player i in a strategy profile $s = (g, s)$ are given by:*

$$\hat{\Pi}_i(s) = f\left(x_i + \sum_{j \in N(i;\bar{g})} x_j\right) - cx_i - \eta_i(g)k + \sum_{j \in N(i,g)} g_{ji}k. \tag{2}$$

We observe that in a simultaneous move game, the last term involving transfers is independent of the strategy of i. Given the linear-separable structure of payoffs, it then follows that for all $s_{-i} \in S_{-i}$, and $s_i, s'_i \in S_i$, $\hat{\Pi}_i(s_i, s_{-i}) \geq \hat{\Pi}_i(s'_i, s_{-i})$, if and only if

[10] Note that we assume pure local externalities: player i only accesses the effort of his immediate neighbors. In Section 4 we study indirect spillovers.

$\Pi_i(s_i, s_{-i}) \geq \Pi_i(s'_i, s_{-i})$.[11] *This means that our methods of analysis and our findings regarding equilibrium outcomes under payoffs given by (1) will carry over to an alternative model, where link formation costs are transfers made by link forming individuals to link receiving individuals.*

A Nash equilibrium is a strategy profile $s^* = (x^*, g^*)$ such that:

$$\Pi_i(s_i^*, s_{-i}^*) \geq \Pi_i(s_i, s_{-i}^*), \forall s_i \in S_i, \forall i \in N.$$

An equilibrium is said to be *strict* if the inequalities in the above definition are strict for every player.

We define social welfare to be the sum of individual payoffs. So that for any profile s social welfare is given by:

$$W(s) = \sum_{i \in N} \Pi_i(s). \tag{3}$$

A profile s^* is socially efficient if $W(s^*) \geq W(s)$, $\forall s \in S$.

We say that there is a path in $\bar{g}$ between i and j if either $\bar{g}_{ij} = 1$ or there exists players $j_1, ..., j_m$ distinct from each other and i and j such that $\{\bar{g}_{ij_1} = \bar{g}_{j_1 j_2} = ... = \bar{g}_{j_m j} = 1\}$; $i \overset{\bar{g}}{\longleftrightarrow} j$ indicates that there is a path between i and j in $\bar{g}$. Given a network $\bar{g}$, we define a component as a set $C(\bar{g}) \subset N$ such that $\forall i, j \in C(\bar{g})$ there exists a path between them and there does not exist a path between $\forall i \in C(\bar{g})$ and a player $j \in N \setminus C(\bar{g})$. A component $C(\bar{g})$ is non-singleton if $|C(\bar{g})| > 1$. A player i is isolated if $\bar{g}_{ij} = 0$, $\forall j \in N \setminus \{i\}$. Let $m(\bar{g})$ be the number of components of $\bar{g}$; we say that a network $\bar{g}$ is minimal if $m(\bar{g} - \bar{g}_{ij}) > m(\bar{g})$, for every link $\bar{g}_{ij} = 1$ in $\bar{g}$, where $\bar{g} - \bar{g}_{ij}$ is a network obtained starting from $\bar{g}$ and deleting a link $\bar{g}_{ij}$. A network $\bar{g}$ is minimally connected if it is composed of only one component and it is minimal.

A network g is a core periphery network if there are some players who are linked to everyone while the rest of the players only form links with these players. Formally in a *core periphery* network there are two groups of players, $\mathbf{N}_1(\bar{g})$ and $\mathbf{N}_2(\bar{g})$, with the feature that $N_i(\bar{g}) =$

[11] We follow convention, and suppose that, for any strategy profile s, $s_{-i} = (s_1, ...s_{i-1}, s_{i+1}, ..., s_n)$, refers to the profile of strategies of all players other than player i.

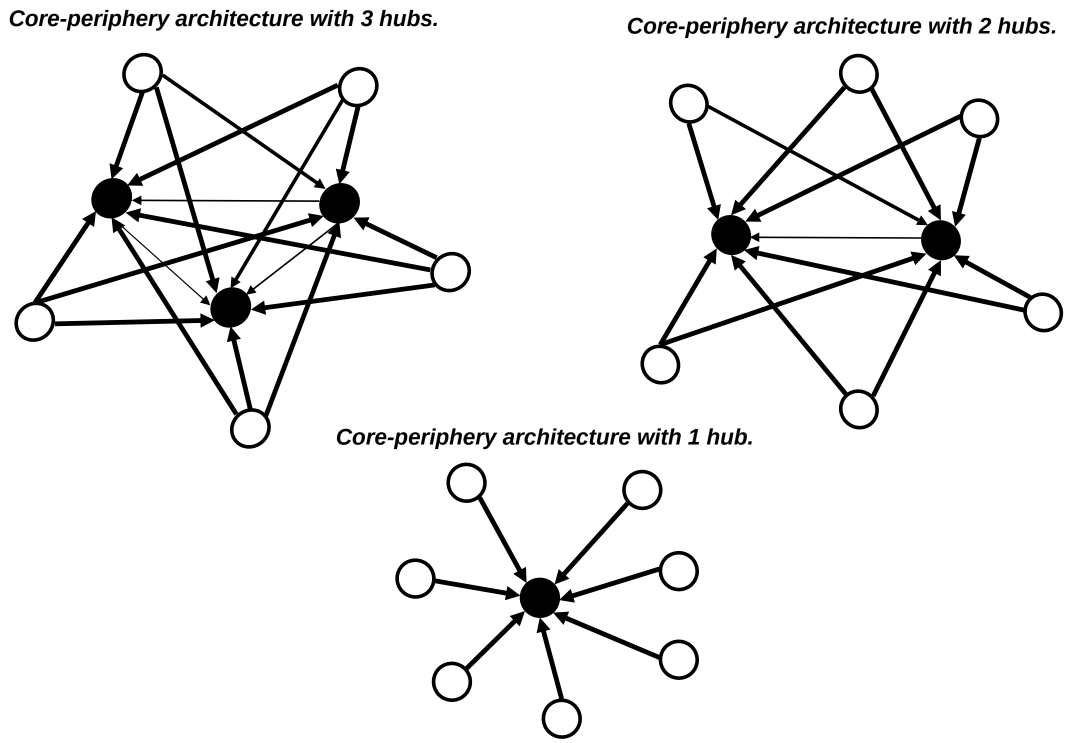

Figure 1: Core Periphery Networks, $n = 8$.

$\mathbf{N}_2(\bar{g})$ for $i \in \mathbf{N}_1(\bar{g})$, $N_j(\bar{g}) = N \backslash \{j\}$ for all $j \in \mathbf{N}_2(\bar{g})$, and, for all $i \in \mathbf{N}_2(\bar{g})$, $g_{ij} = 0$ for all $j \in \mathbf{N}_1(\bar{g})$. The periphery sponsored star is a special case of this architecture, in which $|\mathbf{N}_2(\bar{g})| = 1$ and $|\mathbf{N}_1(\bar{g})| = n - 1$. In a core periphery network, nodes which have $n - 1$ links are referred to as central nodes or as hubs, while the complementary set of nodes are referred to as peripheral nodes or as spokes. Figure 1 illustrates core periphery networks. In the figure there are $n = 8$ players; in each architecture the black nodes are the hubs (the set $\mathbf{N}_2$), the white nodes are the spokes (the set $\mathbf{N}_1$) and an edge starting at i with the arrowhead pointing at j indicates that i sponsors a link to j.

3 Analysis

The focus of our analysis will be on the distribution of effort and linking activity across players. Our main result is that *every* equilibrium exhibits the law of the few: a few players – the mavens – personally invest in gathering information while the rest of the players gather all their information by forming links with the mavens. This feature of equilibrium also tells us that in *every* equilibrium, the social network has a core periphery architecture. We also assess the social welfare properties of these outcomes and find that while the star network is socially attractive as it economizes on the number of links, individual incentives to personally collect information are generally too low, relative to what is socially desirable.

The first observation is that in every equilibrium, every player must access at least $\hat{y}$, and if a player collects some information himself then he must access exactly $\hat{y}$. Indeed, if a player accesses less than $\hat{y}$, then he gains by increasing effort, since marginal returns are larger than marginal costs. Similarly, if a player acquires some information himself and he accesses more than $\hat{y}$, then he can strictly increase payoffs by lowering effort. We next observe that if some player collects $\hat{y}$, and if $k < c\hat{y}$, then it is optimal for all other players to choose zero effort and simply form a link with this player. Lemma 3.1 summarizes these observations. For any equilibrium $s = (x, g)$, let us define $I(s) = \{i \in N | x_i > 0\}$ as the set of players who choose a positive effort, and let $y_i = \sum_{j \in N(i;\bar{g})} x_j$, i.e. the information that i accesses from his neighbors.[12]

Lemma 3.1 *In any equilibrium $s^* = (x^*, g^*)$, $x_i^* + y_i^* \geq \hat{y}$, for all $i \in N$. Moreover, if $x_i^* > 0$ then $x_i^* + y_i^* = \hat{y}$. If $k < c\hat{y}$, then in any equilibrium $s^* = (x^*, g^*)$, if $x_i^* = \hat{y}$, then $x_j^* = 0$, for all $j \neq i$.*

We build on Lemma 3.1 to provide the following (partial) characterization of Nash equilibrium.

Proposition 3.1 *Suppose payoffs are given by (1). If $k > c\hat{y}$ then there exists a unique equilibrium in which every player exerts effort $\hat{y}$ and no one forms any links. Suppose $k < c\hat{y}$ and let $s^* = (x^*, g^*)$ be an equilibrium. If $\sum_{i \in N} x_i^* = \hat{y}$ then g^* is core periphery network, hubs choose positive efforts and spokes choose 0 effort. If $\sum_{i \in N} x_i^* > \hat{y}$ then there are only two possible configurations.*

1. *Every player $i \in I(s^*)$ has $\Delta \in \{1, ..., n-2\}$ links with positive effort players and chooses effort $x_i^* = \frac{\hat{y}}{\Delta+1} = \frac{k}{c}$, while every player $j \notin I(s^*)$ has $\Delta + 1$ links with positive effort players and there are no other links.*

[12] A similar result is proved with regard to equilibrium efforts in a fixed network, by Bramoulle and Kranton (2007).

2. *Every player chooses positive efforts and there are two types of players. High effort players choose $\bar{x}^* = \frac{k}{c}$, while every low effort player has η links with high effort players, they are not neighbors of each other and choose effort $\underline{x}^* = \hat{y} - \eta\frac{k}{c}$, where $\frac{\hat{y}c}{k} - 1 < \eta < \frac{\hat{y}c}{k}$.*

The focus is on the case $k < c\hat{y}$. Nash equilibria exhibit a number of interesting features which are worth noting. If the aggregate effort is $\hat{y}$, then the network has a core periphery architecture (see Figure 2A). This is because if aggregate effort is equal to $\hat{y}$, then, given Lemma 3.1, it must be the case that positive effort players are linked among themselves, and also that zero effort players form a link with every positive effort player and no other links. So the positive effort players are the hubs, while zero effort players are the spokes and the network is a core periphery network.

However, if the aggregate effort is greater than $\hat{y}$, then the above line of reasoning is no longer valid. Indeed, quite the converse must be true: all positive effort players cannot be linked to each other, as that would contradict the first part of Lemma 3.1. Figure 2B illustrates a range of possible Nash equilibrium outcomes in which aggregate efforts exceed $\hat{y}$. While the equilibrium structure is partially characterized in this case, strategic behaviour does place fairly sharp restrictions on the distribution of efforts: only two levels of effort are possible and the lower level effort players only form links with a subset of higher effort players

The equilibrium outcomes in which aggregate effort exceeds $\hat{y}$ are fragile in the following sense: in such equilibria players can form additional links and lower efforts and retain the same payoffs, or, similarly, players can delete one link and increase efforts and retain the same payoffs. In other words, the strategies of the players are not a *strict* best response to the strategies of others. This fragility is not present in the equilibria where aggregate effort is equal to $\hat{y}$. The following result summarizes our analysis on this issue.

Proposition 3.2 *Suppose payoffs are given by (1) and suppose that $k < c\hat{y}$. In every strict equilibrium $s^* = (x^*, g^*)$:*

1. $\sum_{i \in N} x_i^* = \hat{y}$.

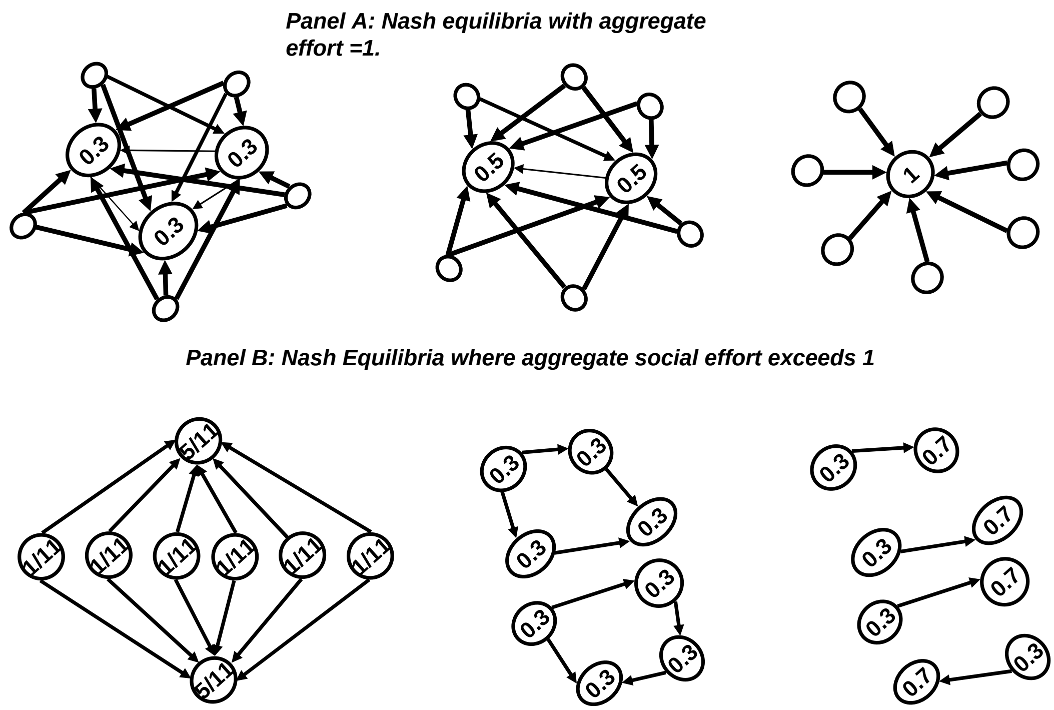

Figure 2: Examples of Nash Equilibria, $n = 8$, $\hat{y} = 1$.

2. *Every equilibrium network has a core periphery architecture, hub players exert positive effort and the spokes choose zero effort.*

3. *For given c and k, with $k < c\hat{y}$, the ratio $|I(s^*)|/n \to 0$ as $n \to \infty$.*

The key point is that in every strict equilibrium the aggregate effort is equal to $\hat{y}$ and this derives from the following basic economic intuition. If a positive effort player strictly prefers to retain a link with another player j, it must be the case that the costs of linking with j is strictly lower than the cost of the extra information that player i would collect himself, should he discard the link with j. But this suggests that the neighbors of player i should find it optimal to access j's information. Hence, in a strict equilibrium positive effort players are linked among themselves and, taken together with Lemma 3.1, this yields the desired conclusion that the sum of effort must equal $\hat{y}$. Part (2) of Proposition 3.2 then follows by Proposition 3.1.

The last step is to derive bounds on the number of positive effort players. For any c and k it must be the case that if i links with j then the cost of the link must be less than the cost of providing the effort accessed from j, in other words, $cx_j > k$. This gives us a lower bound of k/c on the effort of j. Since total effort in equilibrium is $\hat{y}$, it follows that the maximum value of $|I(s)|$ is bounded above by $(\hat{y}c)/k$. This number is independent of n, and so it follows that for any (strict) equilibrium, the ratio $|I(s)|/n$ can be made arbitrarily small by suitably raising n. Thus Proposition 3.2 obtains the law of the few as the outcome of strategic interaction among individuals who are ex-ante identical.[13]

While Proposition 3.2 is a strong result, it does not pin down the number of positive effort players for fixed n, and it also does not say anything about who will be in the core set of positive contributors. We now explore the role of individual heterogeneity in mitigating this residual uncertainty.

A recurring theme in the empirical literature is that people who gather information enjoy collecting it. A natural way to model this is to suppose that some players may have slightly lower costs of collecting information. We consider a situation where $c_i = c$ for all $i \neq 1$, while $c_1 = c - \epsilon > 0$, where $\epsilon > 0$ is a small number. Let $\hat{y}_1 = \arg\max_y f(y) - c_1 y$. Clearly, as long as $\epsilon > 0$, $\hat{y}_1 > \hat{y}$ and when ϵ goes to zero then $\hat{y}_1$ goes to $\hat{y}$. We focus on strict Nash equilibria.

Proposition 3.3 *Suppose payoffs are given by (1), $c_i = c$ for all $i \neq 1$ and $c_1 = c - \epsilon$, $\epsilon > 0$. If $k < f(\hat{y}_1) - f(\hat{y}) + c\hat{y}$ then in every strict equilibrium $s^* = (x^*, g^*)$:*

1. $\sum_{i \in N} x_i^* = \hat{y}_1$.

2. *The network is a periphery sponsored star and player 1 is the hub.*

3. $x_1^* > x_i^* = \tilde{x}$, *for all* $i \neq 1$ *and* $\tilde{x} \to 0$ *as* $\epsilon \to 0$.

The main point of this result is to show that a very small difference in the cost of collecting information is sufficient to considerably sharpen the findings of Proposition 3.2. In Proposition 3.2 an equilibrium network has the core periphery structure where one or more players collect information directly and the identity of these players is not determined. Proposition 3.3 shows that a slight cost advantage in acquiring information resolves the multiplicity

[13]We observe that the periphery sponsored star network where the hub exerts effort $\hat{y}$ and all other players exert effort 0 is a strict equilibrium for all $k < c\hat{y}$. Furthermore, if $k/\hat{y} \in (c/2, c)$, then this is the only strict equilibrium.

problem in both these dimensions: in every equilibrium (with a non-empty network) there is a single hub and the low cost player is the hub player. Moreover, for small ϵ the hub chooses an effort approximately equal to $\hat{y}$ while the spokes choose an effort close to 0.[14]

A few words on the arguments underlying this result are in order. Observe that for the low cost player the optimal information level $\hat{y}_1$ is higher than the optimal information level for other players, $\hat{y}$. It then follows that if player 1 chooses $\hat{y}_1$ then it is optimal for all other players to simply form a link with player 1 and choose zero effort. The proof shows that the communication structure in which the low cost player is the hub and all others form a link with him is the *only* one sustainable in equilibrium. So suppose that the effort of player 1 is $x_1 \in (0, \hat{y}_1)$. In any equilibrium, following Lemma 3.1, it must be the case that $x_1 + y_1 = \hat{y}_1$. However, if player 1 accesses information from some player i then, since $x_1 + y_1 = \hat{y}_1 > \hat{y} = x_i + y_i$, it must be the case that there is a player j whom player 1 accesses but who is not accessed by player i. The key step in the proof shows that any two players i and j who choose positive effort, who are linked with player 1 and not with each other, have *no other links with positive effort players.* This however implies that $x_i + x_1 = \hat{y} = x_j + x_1$ and so $x_i = x_j$. Moreover, player 1 does not link with player i: otherwise j would have a strict preference to do likewise, which would contradict the previous observation. In other words, $x_1 > x_i$, and players like i form the link with player 1. Variations on these arguments establish that all other players must choose the same level of effort and form a link with player 1.

Proposition 3.3 is also interesting in the context of the empirical work in this area. Recall, from the introduction, that a recurring finding of the empirical work is that mavens are different from others in observed characteristics but the differences do not appear to be large. Proposition 3.3 shows that a small cost advantage is sufficient to pin down who will be a maven – and personally collect information – and who will form connections with others to gather information.

[14] In a recent paper on a model of pure network formation with *exogenous* effort levels, Hojman and Szeidl (2006) obtain a result on how small differences between players can help select the identity of the hub players. Their result relies on stochastic stability type arguments. Our result shows that heterogeneity has powerful effects in the selection of hubs when we endogenize information collection. Moreover, we emphasize that, with endogenous efforts, the selection of hubs obtains within a one-shot static model.

To summarize, Propositions 3.2-3.3 say that in social contexts where individuals can invest in collecting information or invest in linking with others who collect information, strategic interaction leads to a very simple and sharp form of role differentiation: a small subset of individuals collect all the information, while the rest of the individuals simply form connections with this small set of mavens. This differentiation in turn generates a simple and elegant structure of social communication: information collectors constitute hubs and the rest of the individuals constitute spokes of a star type network.

3.1 Efficient Outcomes

Given their salience it is important to understand the welfare properties of these forms of social differentiation and communication. Proposition 3.1 tells us that in a Nash equilibrium for every player $i \in N$, $x_i + y_i = \hat{y}$. Thus in every equilibrium, aggregate gross returns are $nf(\hat{y})$. If $k < c\hat{y}$, given the linearity in costs of effort and linking, the most efficient equilibrium minimizes the total costs of effort and links; this immediately yields the periphery sponsored star architecture, where the hub provides all the effort.

Proposition 3.4 *Suppose payoffs are given by (1). If $k < \hat{y}c$, the efficient equilibrium is the periphery sponsored star network where the hub provides effort $\hat{y}$ and every spoke provides effort* 0.

However, it is clear that an equilibrium will not be socially efficient in general. To see this, note that in the star the hub player chooses effort $\hat{y}$, and at this point $f'(\hat{y}) = c$. But marginal social returns are given by $nf'(\hat{y})$, which is certainly larger than c, for $n \geq 2$. Hence, all equilibria are inefficient if $k < \hat{y}c$. This is an implication of the public good nature of individual effort. So long as equilibrium entails any links, it will also imply an under provision of effort relative to the social optimum. The following proposition characterizes efficient outcomes.

Proposition 3.5 *Suppose payoffs are given by (1). For every c, there exists a $\bar{k} > c\hat{y}$ such that if $k < \bar{k}$ then the socially optimal outcome is a star network in which the hub chooses effort $\tilde{y}$ such that $nf'(\tilde{y}) = c$, while all other players choose effort* 0*. If $k > \bar{k}$, then in the socially optimal outcome every player chooses effort $\hat{y}$ and no one forms links.*

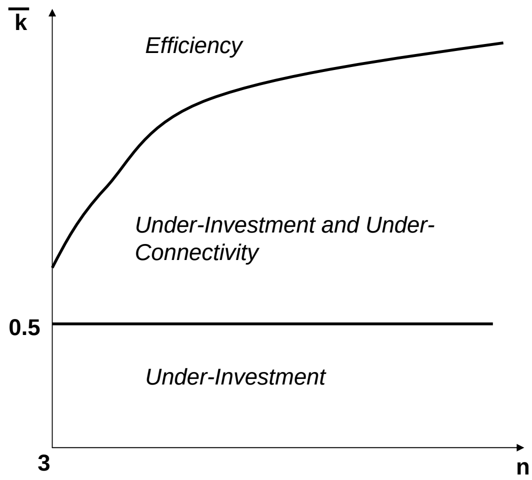

Figure 3: Trade-off Between Efficient Outcomes and Equilibrium Outcomes.

The first thing to note is that given any profile of efforts and linking, there is a corresponding star network in which the hub does all the information collection which is as good. This is a consequence of the linear costs of information collection and the positive costs of linking. The next step is to see that two components which are both stars cannot be optimal as there exists a profile in which all the spokes are connected to a single hub and this hub chooses corresponding higher effort. These arguments show that if the optimal social organization is a non-empty network then it must be a star. The value of $\bar{k}$ is obtained by equating the best social welfare attained in the empty network and the star network.

The next example illustrates the relation between equilibrium and socially efficient collection and communication of information.

Example 1: Suppose $c = 1/2$ and $f(y) = \ln(1 + y)$. In this case $\hat{y} = 1$, while $\tilde{y} = 2n - 1$. In Figure 3 we plot $\bar{k}$ as a function of the number of players. For a given n there are three regions. For low costs of linking, $k < 1/2$, the most efficient equilibrium is a star where the hub provides effort 1 and the spokes choose 0. As compared to socially optimal outcomes, in equilibrium there is under investment. For moderate costs of linking, $k \in (1/2, \bar{k})$, in equilibrium we have under investment and under connectivity relative to socially optimal outcomes (noting that $\bar{k} > 1/2$). In the remaining region, equilibrium outcomes coincide with socially optimal outcomes. ■

4 Indirect flow of information

In the basic model, a person can either acquire information personally or get it from another person who has directly acquired it herself. In actual practice, we often receive information from friends and colleagues which they have themselves received from their friends. The aim of this section is to examine the implications of this form of indirect information transmission for information acquisition and the social structure of communication.

The main insight we obtain is that information spillovers give rise to a new "type" of influential agent: *the connector*. A connector is someone whose primary role is that of acting as an intermediary between other people. He typically has many more links then the average person and he connects individuals who collect information – the mavens – with other mavens or with other individuals who collect no information at all.

We now extend the basic model to allow for indirect flow of information. Given two players i and j in g, the geodesic distance, $d(i, j; \bar{g})$, is defined as the length of the shortest path between i and j in $\bar{g}$. If no such path exists, the distance is set equal to infinity. Let $N^l(i; \bar{g}) = \{j \in N : d(i, j; \bar{g}) = l\}$, that is $N^l(i; \bar{g})$ is the set of players who are at finite distance l from i in $\bar{g}$. We measure the level of spillovers by a vector $a = \{a_1, a_2, ..., a_{n-1}\}$, where $a_1 \geq a_2, ..., \geq a_{n-1}$ and $a_l \in [0, 1]$ for all $l \in \{1, ..., n-1\}$. It is assumed that if $j \in N^l(i; \bar{g})$, then agent's j information accessed by i is $a_l x_j$. The payoffs to player i under strategy profile $s = (x, g)$ can now be written as follows,

$$\Pi_i(s) = f\left(x_i + \sum_{l=1}^{n-1} \sum_{j \in N^l(i;\bar{g})} a_l x_j\right) - cx_i - \eta_i(g)k. \tag{4}$$

We observe that the case in which $a_1 = 1$ and $a_2 = 0$ corresponds to the pure local spillovers model analysed in Section 3. To bring out the role of indirect information transmission in the simplest form we start by considering the polar case of no decay or delay in flow across links, i.e., $a_{n-1} = 1$.

Given a network $\bar{g}$, let $y_i(\bar{g})$ be the information which i derives from others, and define $y_{ij}(\bar{g}) = y_i(\bar{g}) - y_i(\bar{g} - \bar{g}_{ij})$, i.e., the effort that i accesses exclusively via j. The following proposition obtains some important properties of all equilibria with indirect information flow.

Proposition 4.1 *Suppose payoffs are given by* (4) *and suppose that* $a_{n-1} = 1$.

1. *If* $k < c\hat{y}$*, then* $s^* = (x^*, g^*)$ *is an equilibrium if and only if: (i)* $\sum_{i \in N} x_i^* = \hat{y}$*, (ii)* $\bar{g}^*$ *is minimally connected, (iii) If* $g_{ij}^* = 1$ *then* $k \leq c y_{ij}(\bar{g}^*)$.

2. *If* $k > c\hat{y}$*, there exists a unique equilibrium: every player exerts effort* $\hat{y}$ *and no one forms links.*

Frictionless information flow implies that, in any equilibrium, a network is minimal. From Lemma 3.1 we know that in equilibrium every individual must access at least $\hat{y}$ effort level. If the costs of linking k are smaller than the threshold level of effort $\hat{y}$, then standard considerations, relying on network externalities, imply that the network is connected. Finally note that the costs of a link that player i forms with player j must be lower than the value of information that player i accesses exclusively via the link with j. This implies that either player j collects enough information on his own, or that player j is a *connector*, and accesses others who have enough information.

We now explore the distribution of effort and the architecture of social communication in more detail. We have been unable to provide a complete characterization of network equilibrium architectures. So we proceed by discussing special classes of equilibria, and this helps us to highlight some novel aspects of the implications of strategic interaction in a setting with indirect information transmission.

Connector links with mavens, while others link with the connector: In this class of equilibria, there is an individual (say) 1 who does not acquire any information himself while he forms links with all players who are acquiring information, i.e., the mavens. All other players acquire no information on their own but form a link with player 1. Player 1 is the connector. Figure 4A provides some examples of such equilibria. Within the class of these equilibria property (iii) in part (2) of Proposition 4.1 has the following implication: an increase in the costs of forming links leads to a fall in the number of mavens. To see why this must be true

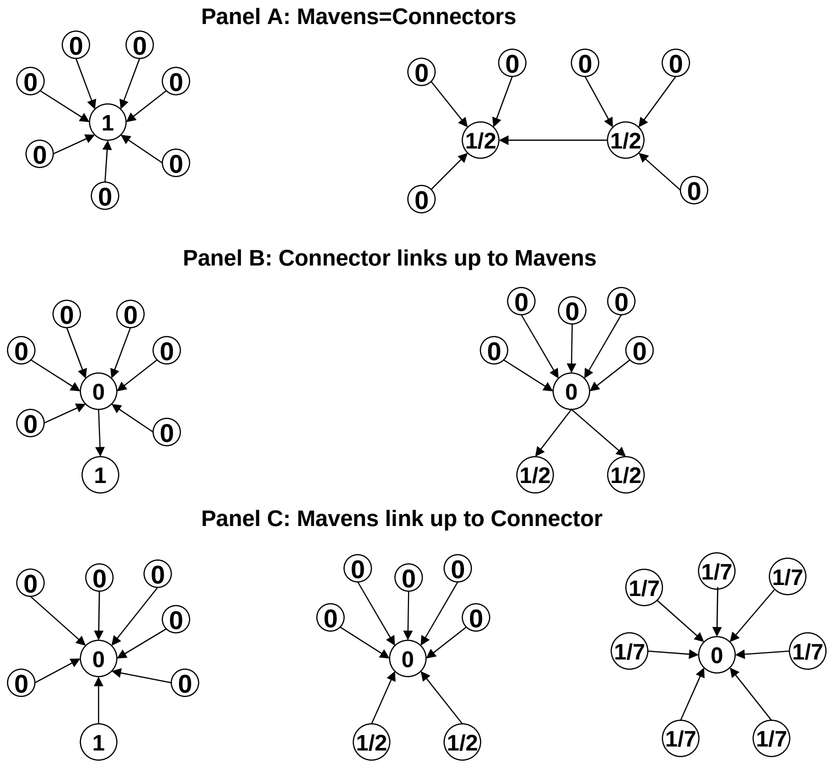

Figure 4: Mavens, Connectors and Others, $n = 8$.

note that an equilibrium with M mavens exists whenever the costs of linking k are smaller than the benefits of accessing a single maven $c\hat{y}/M$. Clearly the latter expression is falling in M. In particular, when $k \in (\hat{y}c/2, c\hat{y})$ we have an equilibrium with one maven and one connector.

Mavens and others link with a Connector: In this class of equilibria, all players – some of whom are mavens while others are not – form a link to a single player; this single player is the connector. Figure 4B illustrates this type of equilibrium. In this class of equilibria, if the costs of linking increase, it is profitable for a maven to link with the connector only if the benefits of linking increase as well. Since mavens are symmetric and the total effort is given by $\hat{y}$, this is only possible if each maven puts in less effort and there are more mavens in equilibrium. Thus an increase in costs of linking leads to an increase in the number of mavens.

Mavens = Connectors: In this class of equilibria connectors coincide with the players who personally collect information; Figure 4C illustrates these equilibria. In the star network, the hub chooses $\hat{y}$ and every other player accesses information directly from the hub. In the other network, the mavens are also connectors, and all other players choose zero effort and simply link with a connector, who gives them access to personally collected information as well as information collected by others mavens.

This range of equilibrium possibilities is consistent with the extensive evidence we have on the distribution of efforts and the patterns of social communication. In their classic study of personal influence, Katz and Lazersfeld (1955) emphasize that the social influencers typically have more social ties and also acquire more information (via radio, newspapers and television). We interpret this as a situation in which the connectors are also the mavens. In other instances, we observe that connectors acquire some information personally but their numerous contacts also provide new information which they then communicate to their neighbors and friends. Here the highly connected individual functions primarily as a connector. See Gladwell (1999) for an engaging discussion of such patterns of information acquisition and social communication. See also Cross et al. (2001) for a description of connectors in social communication networks within firms.

The different equilibria described in Figure 4 are also relevant because they arise as the only equilibrium outcomes under small *ex-ante* differences between players: heterogeneity in the costs of effort and the costs of linking pins down both the equilibrium network architectures as well as the identity of players occupying prominent positions in social communication structures.

To see this, suppose that player 1 has the lowest marginal costs of effort, i.e., $c_1 < c_i = c$ for all $i \in N \setminus \{1\}$. Let $k_{ii'}$ be the costs that player i has to "pay" for the link with i'. For simplicity, we assume that, for all $i' \neq j$, $k_{ii'} = k_{i'} = k$, while the cost of accessing player j is $k_{ij} = k_j < k$. We characterize equilibria in two situations. First, $j = 1$ so that player 1 has the lowest marginal costs of effort as well as linking. Second, $j \neq 1$, so that the player with low costs of effort and the player with low costs of linking are different. Recall that $\hat{y} = \arg\max_y f(y) - cy$, while $\hat{y}_1 = \arg\max_y f(y) - c_1 y$.

Proposition 4.2 *Suppose payoffs are given by* (4) *and suppose that* $a_{n-1} = 1$, $c_1 < c = c_i$, *for all* $i \neq 1$ *and* $k_j < k = k_i$ *for all* $i \neq j$.

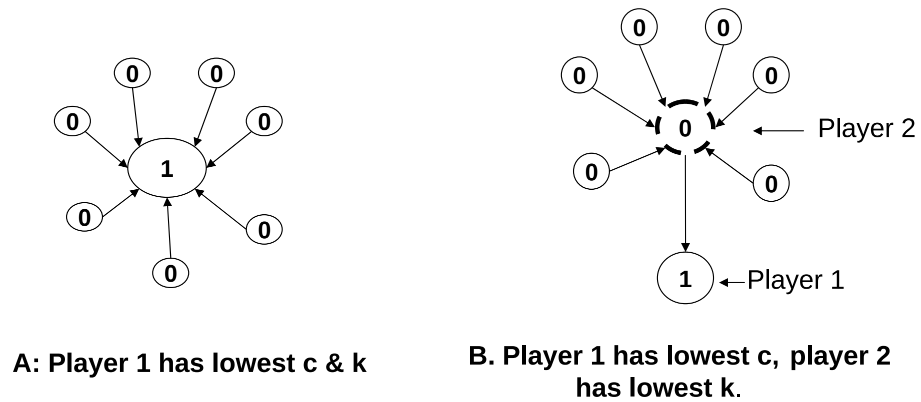

Figure 5: Indirect Flow of information: Heterogeneity

1. *Assume* $j = 1$. *If* $k_1 < f(\hat{y}_1) - f(\hat{y}) + c\hat{y}$ *then there is a unique equilibrium,* $s = (x^*, g^*)$*;* $x_1^* = \hat{y}_1$, $x_i^* = 0$ *for all* $i \neq 1$, g^* *is a periphery sponsored star and player* 1 *is the hub.*

2. *Assume* $j \neq 1$. *If* $k < f(\hat{y}_1) - f(\hat{y}) + c\hat{y}$ *then in every equilibrium* $s^* = (x^*, g^*)$, $x_1^* = \hat{y}_1$, $x_i^* = 0$ *for all* $i \neq 1$, g^* *is such that: there are* $l \in \{0, n-2\}$ *players who form a link with* j*, and all other players lie in the path between player* j *and player* 1*, with each link directed towards player* 1*. Moreover, there is a unique strict equilibrium with* $l = n - 2$.

This result illustrates how a strong form of the law of the few obtains even in the presence of indirect information transmission, if there is a small heterogeneity in costs of acquiring information and costs of forming links. Figure 5 depicts the equilibria. Figure 5A presents the unique equilibrium when player $j = 1$ has the lowest c as well as the lowest cost of being accessed by others. This equilibrium corresponds to the "mavens=connectors" equilibrium. Figure 5B shows some equilibria in the case in which the player with low costs of effort and the player with low costs of linking are different. In the only strict equilibrium of this setting the player with low costs of being accessed is the connector, he links with the maven, who is the player with the low costs of effort, while all other players link with the connector.

To summarize, in the polar case with pure local information transmission, strategic information acquisition and networking gives rise to core periphery network equilibria: few players collects information, while the others access information by networking with this select few. In this case, the mavens coincide with the connectors (Propositions 3.2-3.3). At the other polar case, when information flows seamlessly through links, strategic information acquisition and networking gives rise to the phenomenon of the *connector*. In some equilibria the connector is a pure go-between who collects no personal information himself, and acts as a pure intermediary. In other equilibria, the connector also plays the role of a maven (Proposition 4.1-4.2).

These findings lead us to ask whether these patterns of role differentiation and social communication networks are robust equilibrium properties of intermediate formulations of indirect information transmission. We have been unable to obtain a clear characterization of information acquisition and social communication under a general formulation of decay and so we restrict ourselves to making two general comments.

The first comment concerns the robustness of core-periphery network equilibria. We have observed that these patterns arise under pure local information transmission as well as with frictionless or "global" information transmission (mavens=connectors equilibria). We now observe that a *necessary* condition for them to arise is that $a_1 = 1$. Indeed, in any such configuration aggregate effort equals $\hat{y}$ and there is always a player who relies at least partly on information collected by others. If $a_1 < 1$, then for this player total information available is strictly less than $\hat{y}$, which contradicts Lemma 3.1. On the other hand, it is easy to see that as long as $a_1 = 1$ the periphery sponsored star where the hub exerts all the effort is always equilibrium given that $k < c\hat{y}$.

Our second comment is about periphery sponsored networks with hubs as pure connectors. We observe that if $a_1 < 1$ and $a_2 > a_3 \geq 0$, then a network in which each of the peripheral players chooses $x_i = \hat{y}/[1 + a_2(n-2)]$ is an equilibrium so long as n is sufficiently large. The higher is n, the more attractive it is to form a link with the hub, rather than the peripheral players. Overall, star type of communication networks and the presence of mavens and connectors are salient and robust equilibrium phenomena of social contexts where individuals can collect information personally and/or accessing information indirectly by networking with others.

5 Extensions

In this section, we discuss three assumptions which are maintained in the paper so far, one, a link formed by i with j allows both players to access the information of the other player, two, the costs of acquiring information and linking are linearly increasing and three, the actions sets are continuous.

5.1 One-way information transmission

In our basic formulation we assume that if i forms a link with j then i accesses the effort of j and *vice-versa*. An alternative model is one where a link $g_{ij} = 1$ allows only player i to access the information acquired by player j, i.e., one-way information flow. Let $N(i; g) = \{j \in N : g_{ij} = 1\}$; the payoffs to player i in a strategy profile $s = (x, g)$ are then given by

$$\Pi_i(s) = f\left(x_i + \sum_{j \in N(i;g)} x_j\right) - cx_i - \eta_i(g)k. \tag{5}$$

We observe that some of our main results carry over to the one-way information flow formulation. To illustrate this, consider the equilibria described in Proposition 3.2. Recall that at equilibrium $s = (x, g)$, aggregate effort equals $\hat{y}$, g is a core periphery network, hubs acquire information and spokes do not exert effort. Let g' be a network in which if $g_{ij} = 1$ then $g'_{ij} = 1$ and if $g_{ij} = 0$ and $x_i, x_j > 0$, then $g'_{ij} = 1$. It is easy to see that $s' = (x, g')$ is an equilibrium under the one-way flow formulation (for appropriately chosen k). In particular, in the one way flow formulation, the periphery sponsored star, with the hub collecting all the information is an equilibrium whenever $k < c\hat{y}$. An analogous argument can be used to extend the results of Proposition 3.3. Indeed, it is easy to see that a periphery sponsored star, where the player with the lowest cost of effort is the hub, he personally acquires information while all other players do not provide effort is an equilibrium under the assumption of one-way information flow as well.

5.2 Convex Costs

This section studies the case where the costs of acquiring information and linking are increasing and strictly convex. The idea is that the time and resources needed for information gathering or linking are being diverted from alternatives uses. Under reasonable conditions on preferences – such as diminishing marginal utility – it then follows that marginal costs of effort in the game under study are increasing in effort. This section makes two points: one, convexity in costs of links and information pushes towards outcomes in which all players collect some information and two, that in spite of this pressure, unequal information acquisition and star type networks of social communication remain robust feature of equilibrium.

Define $z_i = x_i + \eta_i k$, and let $C(z_i)$ satisfy the following properties: $C(0) = 0$, $C'(0) = C''(0) = 0$, $C'(z_i) > 0$, for $z_i > 0$, and $C''(z_i) > 0$, for $z_i > 0$. We focus on the case of frictionless information transmission, i.e., $a_{n-1} = 1$. Recall that $y_i \geq 0$ is the information accesses by player i from the others. The payoffs to a player from strategy s_i faced with strategy profile s_{-i} are then given by

$$\Pi_i(s_i, s_{-i}) = f(x_i + y_i) - C(x_i + \eta_i k). \tag{6}$$

What is the distribution of efforts and what is the architecture of the social network in equilibrium? The following result addresses this question. Let $\bar{x}$ be an effort level such that $f'(\bar{x}) = C'(\bar{x})$.

Proposition 5.1 *Suppose payoffs are given by (6) and suppose that $k \neq \bar{x}$ and $s^* = (x^*, g^*)$ is an equilibrium. Then $\bar{g}^*$ is minimally connected or empty. In every non-empty network equilibrium players sponsoring the same number links choose the same effort and the effort of players is declining in the number of links sponsored. If $k < \bar{x}$, the periphery sponsored star network in which the hub chooses x_h and the spokes choose x_s, with $x_h > x_s > 0$ is an equilibrium.*

Note that if the network is non-empty then it is minimally connected, which implies that there is at least one player who forms zero links, while there are clearly some players who form links (since the network is connected). This implies that the number of links is unequal in every

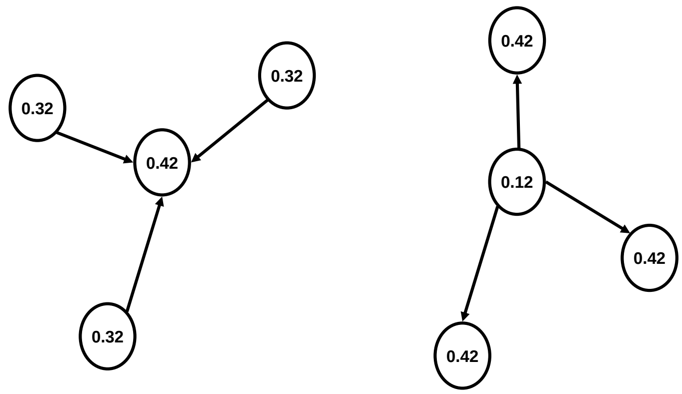

Figure 6: Two Equilibria under Convex Costs, $f(y) = \ln(1+y)$, $C(z) = z^2/2$, k=0.1.

equilibrium and so is personal information acquisition. Thus specialization in information acquisition remains an essential aspect of equilibrium behavior even under convex costs. In particular, the level of information a player personally acquires is declining in the number of links sponsored by that player. We have been unable to fully characterize equilibrium networks but the latter part of the result suggests that star type of network architecture is a robust form of social communication as well. Figure 6 illustrates some equilibrium networks. In both equilibria the network is a star. In the first equilibrium the hub does not form links and he acquires more information relatively to other players. In contrast, in the second equilibrium the hub sponsors all the links and therefore he is the player who invests less in acquiring information personally.

In the basic model with linear costs Propositions 3.2, 3.3, and 4.2 we show that aggregate effort is constant as we vary the number of players, and also as we vary the costs of forming links (so long as $c\hat{y} > k$). If costs are convex, then we would expect that more players would allow for a more even distribution of effort which would lead to lower costs and this would push up aggregate effort. Similarly, we expect that link formation costs will smoothly affect aggregate effort, as these costs now enter the first order conditions of individual optimization for all positive effort players. These considerations lead us to the following result.

Proposition 5.2 *Suppose that payoffs are given by (6) and that $s^* = (x^*, g^*)$ is a non-empty network equilibrium in which all players put positive effort. Then aggregate information acquisition is equal to $n\tilde{x} - (n-1)k$, where $\tilde{x}$ is the solution to $f'(n\tilde{x} - (n-1)k) = C'(\tilde{x})$.*

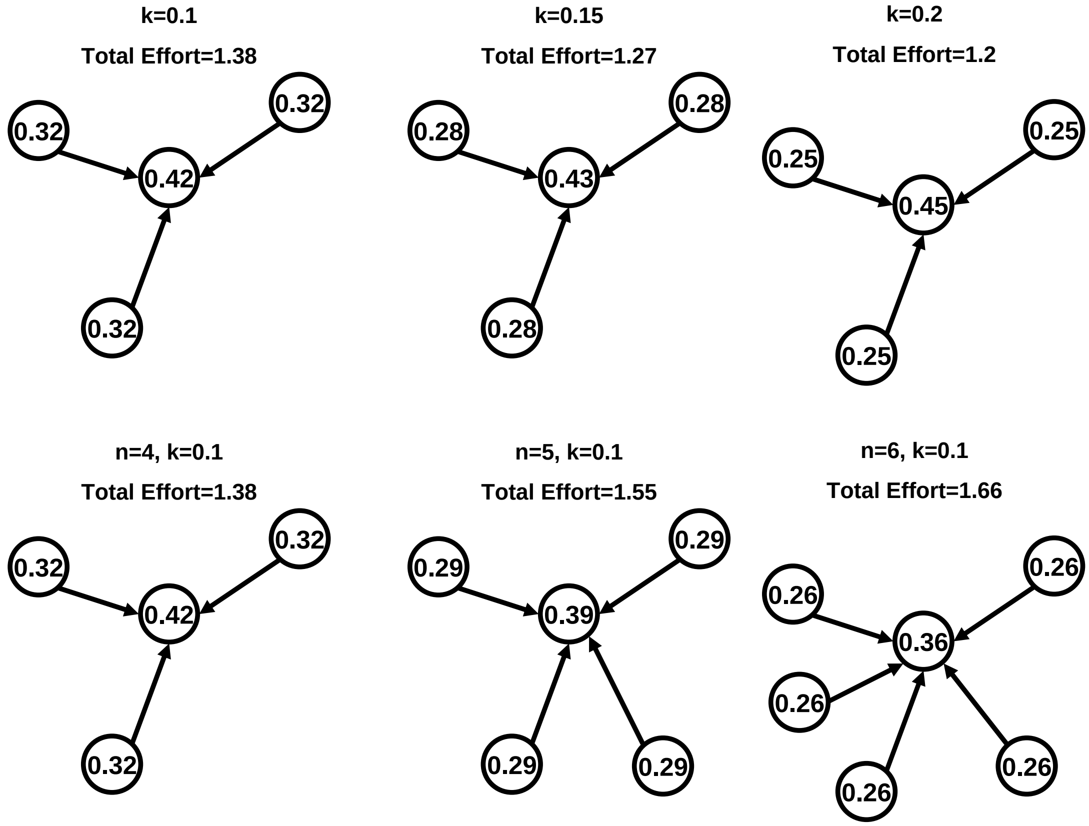

Figure 7: Aggregate Effort under Convex Costs, $f(y) = \ln(1 + y)$, $C(z) = z^2/2$.

This aggregate effort is decreasing in k and increasing in n.

The first part of the result characterizes equilibrium effort in an equilibrium where every player chooses positive effort. Here we exploit the fact that if a player i chooses positive effort then the first order condition of optimization $f'(x_i + y_i) = C'(x_i + k\eta_i(g))$ must hold. Defining $\tilde{x}$ as the effort of a zero link player the aggregate effort follows. The comparative statics follow directly from the aggregate effort equation. Figure 7 illustrates these findings. Figure 7A illustrates periphery sponsored star equilibria for $n = 4$ and different values of k: $k = 0.1$, $k = 0.15$ and $k = 0.2$. Figure 7B illustrates periphery sponsored stars equilibria for $k = 0.1$ and different values of n: $n = 4$, $n = 5$ and $n = 6$.

We now turn to the relation between aggregate effort and specialization in information acquisition. A recurring theme in the paper is the emergence of specialization in information acquisition. In the linear costs case, aggregate effort is constant across different levels of specialization. With convex costs it would appear that specialization would entail lower aggregate information acquisition. However, in our model the costs of links are also important

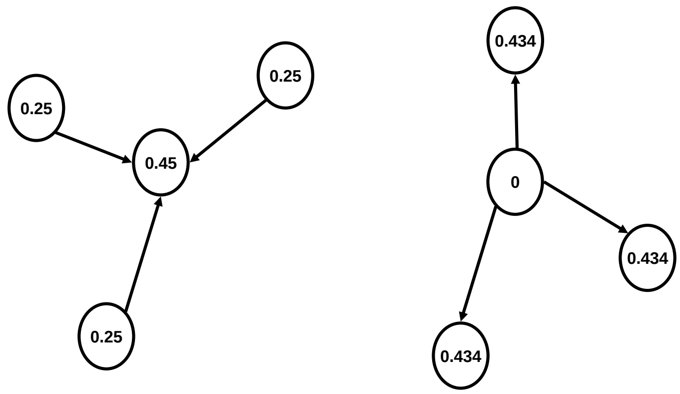

Figure 8: Example. $f(y) = \ln(1 + y)$, $C(z) = z^2/2$, $k = 0.2$, $n = 4$.

and the following example shows that aggregate effort can be higher in a specialized equilibrium as compared to an equilibrium in which all players put positive effort. For $k = 0.2$ the two configurations depicted in Figure 8 are equilibria. In the first equilibrium (the periphery sponsored star) all players put positive effort and aggregate effort is equal to 1.2. In the second equilibrium the hub sponsors all the links and he chooses 0 while the spokes choose positive effort and aggregate effort is 1.3.

5.3 The Best Shot Game

We now study a model in which players have a choice between two actions 0 and 1, where we interpret 1 as acquire information and 0 as not acquire information. The main point we wish to bring out is that every equilibrium exhibits the law of the few and that equilibria are efficient. We focus on the case where players only access the information personally acquired by their direct neighbors.

Formally, a player either acquires information at a cost c or he does not provide effort at all, i.e. $X = \{0, 1\}$. The returns to a player from acquiring information are $f(x_i + y_i) = 1$ if $x_i + y_i \geq 1$, otherwise $f(x_i + y_i) = 0$, where $y_i = \sum_{j \in N(i;\bar{g})} x_j$. We assume that $c < 1$. This specification resembles the best shot game which has been widely studied in economics. The best-shot game is a good metaphor for situations in which there are significant externalities between players' effort.[15]

[15]For a discussion of best-shot games within the context of public good games see, e.g., Hirshleifer (1983) and Harrison and Hirshleifer (1989).

The following proposition characterizes the equilibria in the best shot game.

Proposition 5.3 *Suppose* $X = \{0, 1\}$. *If* $k < c$ *then every equilibrium has a periphery sponsored star architecture, the hub chooses* 1 *and every spoke chooses* 0. *If* $k > c$ *then there exists a unique equilibrium: every player chooses* 1 *and no one forms any links.*

We finally note that in best shot game every equilibrium is efficient. This is in sharp contrast with the case in which effort is a continuous variable.

Proposition 5.4 *Suppose* $X = \{0, 1\}$. *If* $k < c$, *then the socially optimal outcome is a star network, the hub chooses* 1 *and every spoke chooses* 0. *If* $k > c$, *then in the socially optimal outcome every player chooses* 1 *and no one forms links.*

6 Concluding remarks

We have defined the law of the few as the empirical phenomenon of a small subset of individuals collecting and providing information while the rest of the group invests in connections with this select few. The main contribution of our paper is to develop a simple model of investments in information acquisition and links in which the law of the few – reflected in specialization in information acquisition and social communication – emerges as a robust equilibrium phenomenon.

There are two directions in which the analysis of this paper can be extended which appear to us to be especially promising. In actual practice individuals decide on information acquisition and links with others over time, and it is important to understand the dynamics. A second line of investigation concerns the benefits of own efforts and the efforts of others. We have focused on the case where they are substitutes; it would be interesting to study the case where efforts are complements.

7 Appendix

For a strategy profile $s = (x, g)$, define, with some abuse of notation, $y_i = \sum_{j \in N(i;\bar{g})} x_j$ as the total effort accessible to player i from the others. Recall that $\hat{y} = \arg\max_{y \in X} f(y) - cy$.

Proof of Lemma 3.1: We first prove statement 1 in the lemma. Suppose not and $x_i+y_i < \hat{y}$ for some i in equilibrium. Under the maintained assumptions $f'(x_i + y_i) > c$ and so player i can strictly increase his payoffs by increasing effort. Next suppose that $x_i > 0$ and $x_i+y_i > \hat{y}$. Under our assumptions on $f(.)$ and c, if $x_i + y_i > \hat{y}$ then $f'(x_i + y_i) < c$; but then i can strictly increase payoffs by lowering effort. This completes the proof.

Next we prove statement 2 in the lemma. Suppose that $s = (x, g)$ is an equilibrium in which $x_i = \hat{y}$ and there is $x_j > 0$, for some $j \neq i$. First, since $x_i > 0$, it follows from Lemma 3.1 that $x_i + y_i = \hat{y}$. This also implies that every player in the neighborhood of i must exert effort 0. Now consider j, with $x_j > 0$. This means that $\bar{g}_{ij} = 0$. It follows from Lemma 3.1 that $x_j + y_j = \hat{y}$. If $x_j = \hat{y}$ then this player must get payoff $f(\hat{y}) - c\hat{y}$. If he switched to a link with i and reduced effort to 0, his payoff is $f(\hat{y}) - k$. Since $k < c\hat{y}$, $x_j = \hat{y}$ is clearly not an optimal strategy for player j. So s is not an equilibrium. Next suppose that $x_j < \hat{y}$. From Lemma 3.1 we know that in equilibrium $x_j + y_j = \hat{y}$, and so there there is some player $l \neq i$ such that $\bar{g}_{jl} = 1$ and $x_l \in (0, \hat{y})$. It is clear that if $g_{jl} = 1$ then player j can strictly increase his payoffs by switching the link from l to i. Similarly, if $g_{lj} = 1$, then player l gains strictly by switching link from j to i. So s cannot be an equilibrium. A contradiction which completes the proof. ■

Proof of Proposition 3.1: The proof for the case $k > c\hat{y}$ is trivial and therefore omitted; we focus on $k < c\hat{y}$. First suppose that $\sum_{i \in N} x_i = \hat{y}$. In this case it follows from Lemma 3.1 that $I(s)$ must be a clique. Furthermore, $\bar{g}_{ij} = 0$ for all $i, j \notin I(s)$ and $g_{ij} = 0$ for all $i \in I(s)$, $j \notin I(s)$. Therefore, each player choosing 0 effort must sponsor a link with every positive effort player. This also shows that the network must be a core periphery network.

Hereafter, let $s = (x, g)$ be an equilibrium where $\sum_{i \in N} x_i > \hat{y}$. The proof for this case is developed in three steps. In the first step, we consider the case in which positive effort players choose the same effort. In the second step we consider situations in which positive effort players choose different level of efforts. The third step uses the observations derived in the previous two steps to conclude the proof.

Step 1. We prove that if all positive effort players choose the same level of effort then s

satisfies Part 1 of Proposition 3.1. Suppose $x_i = x$, $\forall i \in I(s)$. If $x = \hat{y}$, Lemma 3.1 implies that $|I(s)| = 1$ and therefore aggregate effort is $\hat{y}$, a contradiction. Assume $x \in (0, \hat{y})$; from Lemma 3.1 it follows that $x_i + y_i = \hat{y}$, $\forall i \in I(s)$. Since, by assumption, $x_i = x$, $\forall i \in I(s)$, it follows that every positive effort player accesses the same amount of effort from his neighbors, which immediately implies that every positive effort player has the same number of links with positive effort players; let Δ be this number. Note that for all $i \in I(s)$, $x_i + y_i = x + \Delta x = \hat{y}$, which implies that $x = \frac{\hat{y}}{\Delta+1}$. Since aggregate effort is strictly higher than $\hat{y}$ it follows that $\Delta + 1 < |I(s)|$. Also, from Lemma 3.1 we know that $x < \hat{y}$, which implies that $\Delta \geq 1$. Thus, there exist two positive effort players who are neighbors, implying that $k \leq cx$. Also, since, by assumption, $\sum_{i \in N} x_i > \hat{y}$, there exist two positive effort players who are not neighbors, implying that $k \geq cx$. Hence, $k = cx$. Finally, if $I(s) = N$, the proof follows. If not, select $j \notin I(s)$. Clearly, in equilibrium no player forms a link with j. So, in equilibrium j must sponsor $\Delta + 1$ links with positive effort players. This concludes the proof of part 1 of the proposition.

Step 2. Let g' be the subgraph of g defined on $I(s)$. Let $C(\bar{g}')$ be a component of $\bar{g}'$. By construction each player in $C(\bar{g}')$ chooses positive effort. Suppose that (A1) total sum of efforts in $C(\bar{g}')$ is strictly higher than $\hat{y}$ and (A2) there exists at least a pair of players in $C(\bar{g}')$ who choose a different level of effort. The following Lemma is key.

Lemma 7.1 *Suppose that (A1) and (A2) hold in $C(\bar{g}')$. Then there are two types of players in $C(\bar{g}')$: high effort players choose $\bar{x}$ and low effort players choose $\underline{x} < \bar{x}$. Moreover, every low effort player forms η links with high effort players, there are no links between low effort players, $k = c\bar{x}$, $\underline{x} = \hat{y} - \eta\bar{x}$ and $\frac{\hat{y}c}{k} - 1 < \eta < \frac{\hat{y}c}{k}$.*

Proof Lemma 7.1. Without loss of generality label players in $C(\bar{g}')$, so that $\hat{y} > x_1 \geq x_2 \geq ... \geq x_m$. (A2) implies that there exists $l \in C(\bar{g}')$, $l \neq m$, such that $x_j = x_l = \bar{x}$, for all $j \leq l$, and $\bar{x} > x_{l+1}$. We start by proving two claims.

Claim 1. For all $j > l$, $g_{ji} = 1$ for some $i \leq l$.

Proof of Claim 1. Suppose that there exists a $j > l$ such that $g_{ji} = 0$, $\forall i \leq l$. This implies that j does not sponsor links. If, on the contrary, player j sponsors links, then these links are directed to players $j' > l$, but then player j could strictly gain by switching a link from j' to some $i \leq l$. Note that, it must also be the case that j does not receive any links. Suppose j receives a link from a player j'. Then it must be the case that player l is also j's neighbor, otherwise j' strictly gains by switching the link from j to l. But this says that every player who sponsors a link to j is l's neighbor and since player j only receives links, this means that player j accesses from his neighbors at most as much effort as player l does. Since $x_j + y_j = x_l + y_l = \hat{y}$, this implies that $x_j \geq x_l$, contradicting our hypothesis that $x_j < \bar{x} = x_l$. Thus j does not receive links. But then $x_j + y_j = x_j < \hat{y}$, which contradicts Lemma 3.1. Hence, claim 1 follows. ■

Claim 2. There exists some $i, i' \leq l$ such that $\bar{g}_{ii'} = 0$.

Proof of Claim 2. Suppose $\{1, .., l\}$ is a clique. Since aggregate effort in component exceeds $\hat{y}$, it must be the case then that for every $i \leq l$, there is one player $j > l$ such that $\bar{g}_{ij} = 0$. Select such a player j. Clearly, $g_{jj'} = 0$ for all $j' > l$, otherwise j strictly gains by switching the link from j' to i. Analogously, if j receives a link from some $j' > l$, then i must also be a neighbor of j'. Therefore, since $\{1, .., l\}$ is a clique, it follows that every neighbor of j is also i's neighbor, and this contradicts the hypothesis that $x_j < x_l = \bar{x}$. So $\bar{g}_{jj'} = 0$, $\forall j' > l$. Finally note that this implies that i accesses a superset of the players accessed by j, i.e., $y_i \geq y_j$. We know that $x_i + y_i = \hat{y} = x_j + y_j$ and so $x_j \geq x_i$, which contradicts the hypothesis that $j > l$. Claim 2 follows. ■

We can now conclude the proof of Lemma 7.1. From claim 1, there exists a player $j > l$ who sponsors a link to a player $i \leq l$, so $k \leq c\bar{x}$. Similarly, claim 2 implies that there exists $i, i' \leq l$ such that $\bar{g}_{ii'} = 0$; this implies $k \geq c\bar{x}$. Hence, we have $k = c\bar{x}$.

Next, since $k = c\bar{x}$ and $x_j < \bar{x}$ for all $j > l$, it follows that $\bar{g}_{j'j} = 0$ for all $j > l$. Therefore, every player $j > l$ forms only links with players in $\{1, .., l\}$. We now show that $x_j = x_{j+1}$ for all $j > l$. Select $j > l$ and assume that $x_j > x_{j+1} > 0$. Then, $x_j + y_j = x_j + \eta_j(g)\bar{x}$ and $x_{j+1} + y_{j+1} = x_{j+1} + \eta_{j+1}(g)\bar{x}$. Lemma 3.1 implies that $x_j + y_j = x_{j+1} + y_{j+1} = \hat{y}$, which holds whenever $x_j - x_{j+1} = (\eta_{j+1}(g) - \eta_j(g))\bar{x}$. Since $x_j > x_{j+1}$, then $\eta_{j+1(g)} - \eta_j(g) \geq 1$, but then $(\eta_{j+1}(g) - \eta_j(g))\bar{x} \geq \bar{x} > x_j - x_{j+1}$, where the last inequality follows because, by assumption, $x_j < \bar{x}$. Thus, all players $j > l$ choose the same effort, say $\underline{x}$, and from Lemma 3.1 it follows that $\underline{x} + \eta_j(g)\bar{x} = \hat{y}$. Thus, every low effort player sponsors the same number of links with high effort players, say η, and $\underline{x} + \eta\bar{x} = \hat{y}$. This concludes the proof of Lemma 7.1. ■

Step 3: We now conclude the proof of Proposition 3.1. Recall that g' is the subgraph of g defined on $I(s)$. We need to consider two cases: one, $\bar{g}'$ is connected, and two $\bar{g}'$ is not connected.

$\bar{g}'$ **is connected:** first observe that that (A1) holds by assumption. If all positive effort players choose same action then step 1 applies and the proof follows. If (A2) holds then Lemma 7.1 applies. we next observe that in this case every player $i \in N$ must choose positive effort. To see this note that since $k = c\bar{x}$ every player $j \notin I(s)$ will only sponsor links to high effort players. Then, by symmetry, low effort players must obtain the same payoffs as players $j \notin I(s)$. It is easy to check that this is possible if and only if $\underline{x} = \bar{x}$, which contradicts (A2).

g' **is not connected;** Let C_1 and C_2 be two components in g'. We observe that $x_i < \hat{y}$, and from Lemma 3.1 it follows that the components must contain at least two players each. Here, note that for every $i, i' \in C_1$ and $j, j' \in C_2$ such that $g_{ii'} = g_{jj'} = 1$, $x_{i'} = x_{j'} = x \geq x_i, x_j$ and $k = cx$. Indeed, $x_{i'} = x_{j'} = x$ follows because, if $x_{i'} < x_{j'}$ then player i would strictly gain by switching a link from i' to j'; for analogous reasons it follows that $x_i, x_j \leq x$; Since i sponsors a link to i', $k \leq cx$, while i' does not sponsor a link to j', and so $k \geq cx$. Thus $k = cx$. Together, these observations imply that every player who receives a link in C_1 and every player who receives a link in C_2 chooses effort x. Thus, if in C_1 and C_2 every positive player receives at least one link, every player chooses the same effort and the proof follows from step 1.

Suppose next that there is some player in C_1 who does not receive a link *and* effort is not equal across players. If the aggregate effort in C_1 equals $\hat{y}$, then C_1 is a clique and therefore there is at most one player who only sponsors links and receives no links. Since C_1 is a clique and aggregate effort is $\hat{y}$, this player will choose $\underline{x} = \hat{y} - (|C_1| - 1)x$.

Finally consider the case where aggregate effort in C_1 exceeds $\hat{y}$, and efforts are not equal. Then Lemma 7.1 applies and there are two positive effort levels, x and x', with $x' < x$. We observe that as in the case of connected network above, it is possible to rule out j such that $x_j = 0$. Since C_1 was arbitrary, this completes the proof of Proposition 3.1. ■

Proof of Proposition 3.2: We first prove that in every strict Nash equilibrium $s = (x, g)$ the aggregate effort is equal to $\hat{y}$, i.e., $\sum_{i \in N} x_i = \hat{y}$. Let $s = (x, g)$ be a Nash equilibrium in which aggregate effort exceeds $\hat{y}$. First, suppose s satisfies part 1 of Proposition 3.1. We know that a positive effort player i chooses $x_i = k/c < \hat{y}$ and forms Δ links such that $x_i[\Delta + 1] = \hat{y}$, and that $\Delta + 1 < I(s)$. Moreover $cx = k$. Then it is immediate that this player is indifferent between a link and exerting additional effort k/c. This means equilibria in part 1 are not strict. Second, suppose s satisfies part 2 of Proposition 3.1. Again a player with positive effort is indifferent between forming a link and exerting extra effort himself, since $c\bar{x} = k$. Hence, equilibria in part 2 of the proposition are are not strict. Taken together with Proposition 3.1 this implies that in every strict equilibrium aggregate effort is equal to $\hat{y}$. The core-periphery architecture of equilibrium networks follows directly from Proposition 3.1.

We now consider the proportion $|I(s)|/n$ in every strict equilibrium s. Observe first that in a strict equilibrium aggregate effort $\sum_{i \in N} x_i = \hat{y}$. From Lemma 3.1 we know that $x_i + y_i \geq \hat{y}$, for all $i \in N$, in equilibrium. This means that every positive effort player is accessed by every player in equilibrium. This implies that there is at most one player i with no incoming links, i.e., $g_{ji} = 0$, for all $j \in N$. For all other players l, it must be the case there is at least one player j such that $g_{jl} = 1$; but this implies that that $x_l > k/c$. So the number of accessed players with incoming links, $I(s) - 1$, is bounded above by $(\hat{y}c)/k$. It follows that $I(s)/n \leq ([\hat{y}c/k] + 1)/n$, which can be made arbitrarily small by raising n. ■

Proof Proposition 3.3: It is immediate to see that if $x_1 = \hat{y}_1$ then the proposition follows. Next, if $x_1 = 0$ then we can use Proposition 3.2 to show that in a strict equilibrium aggregate effort equals $\hat{y}$. Note however that if $x_1 = 0$, player 1 must access at least $\hat{y}_1 > \hat{y}$ from his neighbors, a contradiction. We now take up the case of $x_1 \in (0, \hat{y}_1)$.

Claim 1. $\forall i, j \in I(s) \setminus \{1\}$, if $\bar{g}_{ij} = 1$ then i and j share the same neighbors, i.e., for every $l \in I(s) \setminus \{i, j\}$, $l \in N(i; \bar{g})$ if and only if $l \in N(j; \bar{g})$.

Proof Claim 1. Let $\bar{g}_{ij} = 1$, $i, j \in I(s) \setminus \{1\}$, and suppose, without loss of generality, that $x_i \leq x_j$. We first prove that for every $l \in I(s) \setminus \{i, j\}$, if $l \in N(i; \bar{g})$ then $l \in N(j; \bar{g})$. Suppose not and there exists a player $l \in I(s)$, with $l \in N(i; \bar{g})$ and $l \notin N(j; \bar{g})$. If $g_{li} = 1$, then, since $x_i \leq x_j$, l (weakly) gains by switching the link from i to j. Hence, let $g_{il} = 1$. Since $x_i > 0$, it follows from Lemma 3.1 that $x_i + y_i = \hat{y}$ and the payoffs to i in equilibrium s are $f(\hat{y}) - cx_i - \eta_i(g)k$. Suppose that i deletes the link with player l and chooses an effort $\tilde{x}_i = x_i + x_l$, then he obtains payoffs $f(\hat{y}) - cx_i - cx_l - (\eta_i(g) - 1)k$. Since s is a strict equilibrium this deviation strictly decreases i's payoffs, which requires that $k < cx_l$. Let $k < cx_l$ and consider the following two possibilities.

(1:) $x_j \geq x_l$. In this case, since $\bar{g}_{jl} = 0$, and since s is a strict equilibrium, player j must strictly loose if he forms and additional link with l and chooses effort $\tilde{x}_j = x_j - x_l$. That is, $f(\hat{y}) - cx_j - \eta_j(g)k > f(\hat{y}) - c(x_j - x_l) - (\eta_j(g) + 1)k$, which holds if and only if $k > cx_l$; but this contradicts that $k < cx_l$.

(2:) $x_j < x_l$. Here we have two sub-cases. (2a:) Suppose $g_{ij} = 1$; this implies that the costs for i to link with j are strictly lower than the costs of effort that i accesses from j, i.e., $k < cx_j$. Since $k < cx_j$, $\bar{g}_{lj} = 0$ and, by assumption, $x_l > x_j$, then l strictly gains if he links with j and chooses effort $\tilde{x}_l = x_l - x_j$. So s is not a strict equilibrium. (2b:) Suppose $g_{ji} = 1$. Since j does no access l but he sponsors a link to i, it follows that $x_i > x_l$. Since, by assumption, $x_j < x_l$, it follows that $x_i > x_l > x_j$, which contradicts that $x_i \leq x_j$. We have then shown that for every $l \in I(s) \setminus \{i, j\}$, if $l \in N(i; \bar{g})$ then $l \in N(j; \bar{g})$.

We now show that if $l \in I(s) \setminus \{i, j\}$ and $l \in N(j; \bar{g})$ then $l \in N(i; \bar{g})$. Suppose not; then player j accesses all positive effort players that i accesses plus some other positive effort players. But this would contradict that $x_i \leq x_j$. This concludes the proof of Claim 1. ■

Claim 2 Suppose $i, j \in N(1; \bar{g})$, $i, j \in I(s)$, and $\bar{g}_{ij} = 0$, then $\bar{g}_{li} = \bar{g}_{lj} = 0$ for all $l \neq 1$, $l \in I(s)$.

Proof of Claim 2. Suppose, without loss of generality, $x_i \leq x_j$. We first show that $\bar{g}_{li} = 0$ for all $l \neq 1$, $l \in I(s)$. Suppose, on the contrary, that $\bar{g}_{li} = 1$, for some $l \neq 1$, $l \in I(s)$. In view of claim 1, since $\bar{g}_{ij} = 0$, then $\bar{g}_{lj} = 0$; this fact and $x_i \leq x_j$ implies that $g_{il} = 1$, and since $\bar{g}_{ij} = 0$, then from the strictness of equilibrium, it follows that $x_l > x_j$. Since $x_l > x_j$, $g_{il} = 1$, and $x_j + y_j = \hat{y} = x_i + y_i$, it must be the case that there exists some $l' \in I(s)$, $l' \in N(j; \bar{g})$ and $l' \notin N(l; \bar{g})$. Claim 1 implies that $l' \notin N(i; \bar{g})$. Since $x_l > x_j$ and $l' \notin N(l; \bar{g})$, then $g_{jl'} = 1$. But $g_{jl'} = 1$ and $l \notin N(j; \bar{g})$ implies that $x_{l'} > x_l$; similarly, $g_{il} = 1$ and $l' \notin N(i; \bar{g})$ implies that $x_l > x_{l'}$, a contradiction. Thus, the only neighbor of i is player 1. It is easy to see that the same holds for player j. Indeed, if $l \in I(s)$ and $l \in N(j; \bar{g})$, then claim 1 implies that $l \notin N(i; \bar{g})$, but then player j accesses strictly higher effort than player i, which contradicts our initial hypothesis that $x_j \geq x_i$. Claim 2 follows. ∎

Final step in proof of Proposition 3.3: We are concerned with the case $x_1 \in (0, \hat{y}_1)$. Since $x_1 + y_1 = \hat{y}_1$, there exists some $i \in I(s)$ such that $i \in N(1; \bar{g})$. Observe that given such an i, there exists a $j \in N(1; \bar{g})$ such that $j \notin N(i; \bar{g})$ and $x_j > 0$. This is because otherwise $x_i + y_i \geq \hat{y}_1 > \hat{y}$, and this contradicts $x_i > 0$ and Lemma 3.1. From Claim 2 above we know that players i and j do not have any links with players in $I(s)$. This means that $x_i + x_1 = \hat{y} = x_j + x_1 = \hat{y}$, and so $x_i = x_j$. Given that we are in a strict equilibrium, it follows that $k < cx_i$. This implies that i and j choose the same effort and form a link with 1. We observe that this also means that $\hat{y} > x_1 > x_i = x_j > 0$.

We now show that the player 1 constitutes the hub of his component and that all other players behave as players i and j identified above. In a path of length of two or more starting at 1, there are four possible patterns: two players choosing 0, two players choosing positive effort, and two mixed cases. Clearly it is not possible to have two players choosing 0 as the costs of linking are strictly positive. Next consider a positive effort player followed by a zero effort player. Suppose player $l \notin I(s)$ and suppose there is a link with player m such that $\bar{g}_{m1} = 1$. Since $l \notin I(s)$, it must be the case that $g_{lm} = 1$ and so $m \in I(s)$. However, it is profitable for l to form this link only if $cx_i < k < cx_m$. Putting together this fact with the effort levels of i we get that $x_m > x_i$, and so $x_m + y_m \geq x_m + x_1 > x_i + x_1 = \hat{y}$. This contradicts Lemma 3.1. Thus we have ruled out case 2. The case of two positive efforts is ruled out by noting that in that case there is a sequence of players 1, l and l' such that $x_1 + x_l + x_{l'} \leq \hat{y}$, but this means that $x_l, x_{l'} < x_i$, and so a link $\bar{g}_{ll'} = 1$ is not profitable for either l or l'. The last case to consider has a sequence 1, l and l', with $x_l = 0$ and $x_{l'} > 0$. Clearly then $g_{ll'} = 1$ and so $cx_{l'} > k$. If $x_{l'} > x_1$, then it is strictly profitable for l' to lower effort and form a link with 1, while if $x_{l'} < x_1$ then it is strictly profitable for 1 to lower effort and form a link with l'. We have thus shown there cannot exist a path of length of two or more starting at player 1. So player 1 constitutes a hub. Now we can exploit the fact that there exists a player i such that $g_{i1} = 1$ and $x_i + x_1 = \hat{y}$ to conclude that there cannot exist any links between the neighbors of player 1. This proves that player 1 is a hub of his component and that all other players choose effort level x and form a link with player 1.

The above argument is done for a single component. The connectedness of non-empty strict equilibrium networks follows form standard arguments; the details are omitted. Finally, note that if $x_1 = \hat{y}_1$, then $x_i = 0$ for all $i \neq 1$ and therefore property (3) follows. Suppose $x_1 \in (0, \hat{y}_1)$, then we know that each player $i \neq 1$ chooses $x_i = x$. Since spokes sponsor only a link to the hub, in a strict equilibrium it must be the case that $x_1 > x$. Furthermore, for a player i to play x is optimal only if $x = \hat{y} - x_1$ and, similarly, for player 1 to play x_1 is optimal only if $x_1 + (n-1)x = \hat{y}_1$. It is not easy to see that as $\epsilon \to 0$, then $x \to 0$. ■

Proof of Proposition 3.4: Suppose $k < c\hat{y}$. First consider strict equilibria. From Proposition 3.2 it follows that the sum of total efforts is $\hat{y}$, that $x_i + y_i = \hat{y}$ for all $i \in N$ and that g has a core periphery architecture. Given the linearity of costs of effort as well as the costs of linking, it follows that the most efficient strict equilibrium is the star. Let s^* be such configuration, then $SW(s^*) = nf(\hat{y}) - c\hat{y} - (n-1)k$.

We now show that the social welfare of every non-strict Nash equilibrium is strictly lower that $SW(s^*)$. Suppose $s = (x, g)$ is a non-strict Nash equilibrium. From Proposition 3.1 we know that $x_i + y_i = \hat{y}$ for all $i \in N$ and that $\sum_{i \in N} x_i > \hat{y}$. If g is connected, then there are at least $n-1$ links and therefore the proof follows. Suppose g is not connected and suppose there are p components. Let $C_1(g)$ be a component of g. Since s is a non-strict equilibrium then $x_i + y_i = \hat{y}$ for all $i \in C_1(g)$ and $\sum_{i \in C_1(g)} x_i \geq \hat{y}$. Also, the number of links in $C_1(g)$ is at least $m \geq |C_1(g)| - 1$. So the sum of players' payoffs in $C_1(g)$ is $|C_1(g)|f(\hat{y}) - c\sum_{i \in C_1(g)} x_i - mk$. This is (weakly) lower than a profile in which $C_1(g)$ is a star, the hub chooses $\hat{y}$ and all the spokes choose 0. So, $SW(s) \leq nf(\hat{y}) - cp\hat{y} - (n-p)k < SW(s^*)$. This concludes the proof of Proposition 3.4 ■

Proof of Proposition 3.5: Suppose $s = (x, g)$ corresponds to an efficient profile. We first show that if g is not empty, then g is a star. Let g be a non-empty network and suppose that C is a component in g. Let $|C| \geq 3$ be the number of players in C. Suppose that y is the total effort exerted in component C. Then it follows that the total payoff of all players in component C is at most $|C|f(y) - cy - (|C| - 1)k$. Consider a star network with $|C|$ players in which the hub player alone exerts effort equal to y. It then follows that this configuration attains the maximum possible aggregate payoff given effort y. Moreover, note that aggregate payoff in any profile s, in which two or more players exert effort is strictly less than this, since it will entail the same total costs of effort but a strictly higher cost of linking or a strictly lower payoff to at least one of the players. So the star network with the hub exerting effort is the optimal profile for each component.

Next consider two or more components in an efficient profile s. It is easy to see that in a component of size m, efficiency dictates that effort y satisfy $mf'(y) = c$. If the components are of unequal size then efforts will be unequal and a simple switching of spoke players across components raises social welfare. So in any efficient profile with two or more components, the components must be of equal size. Let m be the size and let the effort y satisfy $mf'(y) = c$.

Suppose next that the network contains two components C_1 and C_2 of size m. Consider the network in which the spoke players in component 2 are all switched to component 1. This yields a network g' with components C'_1 and C'_2 with the former containing $2m - 1$ players while the latter contains 1 player. Then the payoff remains unchanged. However, the effort level y is no longer optimal in either of the components. So, for instance, effort can be lowered in component 2 and the aggregate payoff thereby strictly increased, under the assumptions on $f(\cdot)$. A similar argument also applies to networks with three or more components, and so we have proved that no profile with two or more components can be efficient. Thus, if g is not empty then g is a star and the effort of the central player is $\tilde{y} = \arg\max_{y \in X} nf(y) - cy$. The social welfare associated to such profile is: $SW = nf(\tilde{y}) - c\tilde{y} - (n-1)k$.

Finally, note that if s is socially efficient and g is not a star, then g must be empty and every player will choose $\hat{y}$. The social welfare is then $SW = n[f(\hat{y}) - c\hat{y}]$. The expression for $\bar{k}$ is obtained by equating the social welfare in these two configurations, i.e. $(n-1)\bar{k} = n[f(\tilde{y}) - f(\hat{y})] + c[(n-1)\hat{y} - \tilde{y}] + c\hat{y}$. To see that $\bar{k} > c\hat{y}$, note that if $\bar{k} \leq c\hat{y}$, then $n[f(\tilde{y}) - f(\hat{y})] + c[(n-1)\hat{y} - \tilde{y}] + c\hat{y} \leq (n-1)c\hat{y}$, which holds if and only if $nf(\tilde{y}) - c\tilde{y} \leq nf(\hat{y}) - c\hat{y}$. Given that $\tilde{y} = \arg\max_{y \in X} nf(y) - cy$, $\hat{y} = \arg\max_{y \in X} f(y) - cy$ and that $f(\cdot)$ is strictly concave, the above inequality cannot hold. This concludes the proof of Proposition 3.5. ■

Proof of Proposition 4.1 The proof of Part 2 of Proposition 4.1 is straightforward and it is omitted. Hereafter, we focus on Part 1. We first prove that if s satisfies properties (i.)-(iii.) in the proposition then s is a Nash equilibrium. Take a player i; since $\sum_{j \in N} x_j = \hat{y}$, $\bar{g}$ is minimally connected, and there is no decay, then $x_i + y_i = \hat{y}$, so player i does not want to change his own effort level and also he does not want to form an additional link. The payoffs to i at equilibrium s are $f(\hat{y}) - cx_i - \eta_i(g)k$. If $\eta_i(g) = 0$, then player i plays a best reply. Suppose $\eta_i(g) > 0$, then $g_{ij} = 1$ for some j. Note that player i is indifferent between keeping the link with j and switching the link from j to a player that i accessed via j. Also, property (iii.) says that $k \leq cy_{ij}(g)$ and therefore player i does not gain by deleting the link with j. Hence, s is a Nash equilibrium.

We now prove the converse. Let $s = (x, g)$ be an equilibrium. Frictionless information flow implies that every component of $\bar{g}$ must be minimal. Also, in every component the aggregate effort is $\hat{y}$. If not, then a positive effort player strictly gains by either increasing his own effort (if aggregate effort is lower than $\hat{y}$) or decreasing his own effort (if aggregate effort is higher than $\hat{y}$). Next, suppose $\bar{g}$ is not connected. Let C_1 be a component of $\bar{g}$. If $x_i = \hat{y}$ for some $i \in C_1$, then all i's neighbors choose effort 0 and sponsors a link to i, so i's payoffs are

$f(\hat{y}) - c\hat{y}$ and $k < c\hat{y}$. But then player i strictly gains if he chooses 0 and forms a link with a player $j \in C_2$. Thus, in C_1 there are at least two players choosing positive effort; since C_1 is minimal it must be the case that there is a link $g_{i'j'} = 1$ such that player i' accesses via the link with j' strictly less information than $\hat{y}$, say $z < \hat{y}$. It is then clear that if player i' deletes the link with j' and form a new link with a player in C_2, he will incur the same costs but he will access strictly higher information. Therefore player i' can strictly improve his payoffs, a contradiction. Thus, $\bar{g}$ is connected. Finally, it is readily seen that if $g_{ij} = 1$ and s is equilibrium, then $k \leq c y_{ij}(g)$. This concludes the proof. ■

Proof Proposition 4.2: The assumption on frictionless information transmission implies that in every equilibrium the network is minimal. Let $z_{C(g)}$ be the aggregate effort in component $C(g)$

Claim 1. Suppose $s = (x, g)$ is a non-empty equilibrium network and let $C(g)$ be a component. If $1 \in C(g)$ then $z_{C(g)} = \hat{y}_1$. If $1 \notin C(g)$ then $z_{C(g)} = \hat{y}$

Proof of Claim 1. First, note that in every component $C(g)$, $x_i > 0$ for some i. Second, let $1 \in C(g)$. Suppose $x_1 = 0$, then from Lemma 3.1 we know that $x_1 + y_1 \geq \hat{y}_1$; but this contradicts Lemma 3.1 because $\hat{y}_1 > \hat{y}$ and there exists a player $i \in C(g)$ with $x_i > 0$. So, $x_1 > 0$, and again from Lemma 3.1 it follows that $x_1 + y_1 = \hat{y}_1 = z_{C(g)}$. The second part of the claim can be proved using analogous arguments, the details of which are omitted. ■

Proof of Part 1. Suppose $j = 1$, $k_1 < f(\hat{y}_1) - f(\hat{y}) + c\hat{y}$ and $s = (x, g)$ is equilibrium. We first show that g is connected. Suppose g is not connected. First, assume that $1 \in C(g)$ and $C(g)$ is singleton. Claim 1 implies that $z_{C(g)} = \hat{y}_1$. Select a component $\tilde{C}(g)$, and let $i \in \tilde{C}(g)$, $i \neq 1$; claim 1 implies that $z_{\tilde{C}(g)} = \hat{y}$. If $\tilde{C}(g)$ is singleton, $x_i = \hat{y}$ and $\Pi_i(s) = f(\hat{y}) - c\hat{y}$. If player i forms a link to 1 and chooses zero effort, he gets $f(\hat{y}_1) - k_1 > \Pi_i(s)$, where the inequality follows from $k_1 < f(\hat{y}_1) - f(\hat{y}) + c\hat{y}$. This contradicts the hypothesis that s is equilibrium. Thus, $i, j \in \tilde{C}(g)$ and assume, without loss of generality, that $g_{ij} = 1$. The payoffs to i are $\Pi_i(s) = f(\hat{y}) - cx_i - k\eta_i(g)$, where $x_i \geq 0$ and $\eta_i(g) \geq 1$. If i deletes all his links and forms a link to 1, then he gets at least $f(\hat{y}_1) - cx_i - k_1 > f(\hat{y}) - cx_i - k \geq \Pi_i(s)$, where the first inequality follows from $k_1 < k$ and $\hat{y}_1 < \hat{y}$, while the second inequality follows because $\eta_i(g) \geq 1$. This contradicts the hypothesis that s is equilibrium.

The case where 1 belongs to a non-singleton component can be ruled out using analogous arguments, the details of which are omitted. Hence, g is connected.

Since the network g is minimally connected, Claim 1 implies that aggregate effort is $\hat{y}_1$. If $x_i > 0$, $i \neq 1$, then Lemma 3.1 implies that $x_i + y_i = \hat{y} < \hat{y}_1$, which is a contradiction. Thus, $x_i = 0$ for all $i \neq 1$ and therefore $x_1 = \hat{y}_1$ and player 1 does not form any links. This implies that if $g_{ij_1} = 1$ and $j_1 \neq 1$, then i must access 1 via j_1; for otherwise player i would not access any effort via j_1 and therefore i would strictly gain by setting $g_{ij_1} = 0$. But then the payoff of player i is $\Pi(s) = f(\hat{y}_1) - k\eta_i < f(\hat{y}_1) - k_1$, where the last expression is the payoff to i by deleting all his links and sponsoring a link to player 1. This contradicts the hypothesis that s is equilibrium. Hence, $g_{ij} = 0$ for all $i, j \neq 1$, which implies that g is a star, the hub is player 1 and each spoke forms one link with the hub. To conclude the proof of Part 1, note that the payoffs to player $i \neq 1$ are $\Pi_i(s) = f(\hat{y}_1) - k_1$ and, given the assumption on k_1, $\Pi_i(s) > f(\hat{y}) - c\hat{y}$, where the last expression is the best payoff that i can earn should he delete his link and become isolated.

Proof of part 2: Suppose $j \neq 1$, $k < f(\hat{y}_1) - f(\hat{y}) + c\hat{y}$ and $s = (x, g)$ is equilibrium. Without loss of generality, set $j = 2$. From standard arguments we can establish that g is connected and from the assumption of frictionless flow we conclude that g is minimally connected. Claim 1 implies that aggregate effort equals $\hat{y}_1$. Therefore, $x_i = 0$ for all $i \neq 1$ and $x_1 = \hat{y}_1$. There are three facts that follow. Fact 1: player 1 does not form any links. Fact 2: for all $i \neq 1$, if $g_{ii'} = 1$, then i accesses player 1 via i' (including the possibility that $i' = 1$); for otherwise, i strictly gains by deleting the link to i'. Fact 3: for all $i \neq 1$, with $g_{ii'} = 1$ and $i' \neq 2$, if i accesses l via i', then $l \neq 2$; for otherwise player i strictly gains by switching from i' to 2. These three facts immediately imply that player 1 is an end-agent. We then have two possibilities. One, $g_{21} = 1$; in this case, fact 2 and 3 imply that g is a star, player 2 is the hub, and $g_{i2} = 1$ for all $i \neq 1$. Two, $g_{21} = 0$. Consider the path between 2 and 1: $\bar{g}_{2j_1} = \bar{g}_{j_1j_2} = ... = \bar{g}_{j_d1} = 1$. Fact 2 implies that $g_{2j_1} = g_{j_1j_2} = ... = g_{j_d1} = 1$; fact 3 implies that players $2, j_1, ..., j_d$ do not sponsor additional links. Let $i \neq \{2, j_1, ..., j_d, 1\}$, fact 3 implies that player i only links with player 2, i.e., $g_{i2} = 1$. This concludes the characterization of equilibria.

Finally, it is easy to verify that the only non-star network involves a path such as $g_{2j_1} = g_{j_1 j_2} = ... = g_{j_d 1} = 1$; and player 2 is indifferent between a link with j_1 and 1; so such paths are not sustainable in a strict equilibrium. ■

Proof of Proposition 5.1 Suppose $s = (x, g)$ is an equilibrium. We first show that g is either minimally connected or empty. Since information transmission is frictionless, g is minimal. Suppose that g is non-empty; let $C(g)$ contain i and j and let us focus on the case where $l \notin C(g)$ is a singleton component. Also suppose, without loss of generality, that $g_{ij} = 1$.

By definition of $\bar{x}$, it follows that $\Pi_l(s) = f(\bar{x}) - C(\bar{x})$. Player $i \in C(g)$ can always ensure himself $\Pi_l(s)$ by deleting all his links and choosing effort $\bar{x}$; thus, $\Pi_i(s) \geq \Pi_l(s)$. Next, note that l can ensure herself $\Pi_i(s)$ by choosing effort x_i and forming a link $g_{li} = 1$; thus, $\Pi_l(s) \geq \Pi_i(s)$. Putting together these inequalities we get $\Pi_l(s) = \Pi_i(s)$. We now consider two cases.

Case A: There is only one player $j \in C(g)$ such that $x_j > 0$. Clearly player j will not form any links in $C(g)$, and also $y_j = 0$, so we have $x_j = \bar{x}$. Observe that if $k > \bar{x}$ then it is not optimal for player i to have a link with j, so $k < \bar{x}$. It now follows that player l gains strictly by forming a link with j and setting own effort to 0. So, s is not an equilibrium, a contradiction.

Case B: There are two or more positive effort players belonging to $C(g)$. Suppose that one of these players (say) j' also forms $g_{j'm} = 1$. The payoff to player l from linking with m and choosing effort 0 is $f(\sum_{a \in C(g)} x_a) - C(k)$, which is strictly larger than $f(\sum_{a \in C(g)} x_a) - C(x_{j'} + \eta_{j'} k) = \Pi_l(s)$ (where the last equality follows from an argument analogous to the one which shows that $\Pi_i(s) = \Pi_l(s)$). Hence, s is not an equilibrium, a contradiction. Finally, consider the case when there is no player $j' \in C(g)$, with $x_{j'} > 0$ and $g_{j'm} = 1$ for some $m \in C(g)$.

It follows then that for every positive effort player j', $f'(\sum_{a\in C(g)} x_a) = C'(x_{j'})$. Since two or more players exert effort in $C(g)$ and none of them form any links there is some player i who lies in the path between them, with $x_i = 0$, and i forms one or more links. Now it is easy to see that if i forms only one link then payoff $\Pi_i(s) = f(y) - C(k) > f(\bar{x} + d) - C(\bar{x}) > f(\bar{x}) - C(\bar{x}) = \Pi_l(s)$, where $d > 0$ since the effort of player m is accessed by i after he deletes its link. This contradicts the equilibrium hypothesis. Similarly, if i forms two or more links then l can do strictly better than i (and himself in equilibrium) by forming a single link and setting $x_l = 0$, which again contradicts the equilibrium hypothesis.

The case where l belongs to a non-singleton component can be ruled out using analogous arguments, the details of which are omitted. Hence, the equilibrium network is either minimally connected or empty.

Second, let $s = (x, g)$ be a non-empty network equilibrium. We start by proving that players sponsoring the same number of links choose the same effort and that effort is declining in the number of links a player sponsors. Notice that g is minimally connected, and so every player i accesses the same aggregate information, say y. From the optimality condition for a player it follows that if $x_i, x_j > 0$ then $f'(y) = C'(x_i + \eta_i(g)k) = C'(x_j + \eta_j(g)k)$; in other words $x_i + \eta_i(g)k = x_j + \eta_j(g)k$. So, if $\eta_i(g) = \eta_j(g)$ then $x_i = x_j$ and if $\eta_i(g) > \eta_j(g)$ then $x_i < x_j$. Similarly, if $x_i = 0$ then $f'(y) \leq C'(\eta_i(g)k)$ and therefore all players j with $\eta_j(g) > \eta_i(g)$ will also choose $x_j = 0$.

Finally, we prove the last statement of the proposition. Suppose $k < \bar{x}$. We know that $x_h > x_s$ and it is clear that at least one player should provide effort. Thus, $x_h > x_s \geq 0$. Let $y = (n-1)x_s + x_h$. Notice that for the hub the strategy of choosing x_h is optimal if and only if $f'(y) = C'(x_h)$. We now show that $x_s > 0$. Suppose, on the contrary, that $x_s = 0$; for the spoke player the strategy of forming one link and choosing $x_s = 0$ is optimal only if $f'(y) \leq C'(k)$. But if $x_s = 0$, then $x_h = \bar{x}$ and $y = \bar{x}$; since $\bar{x} > k$, $f'(\bar{x}) = C'(\bar{x}) > C'(k)$, a contradiction.

So, suppose $x_s > 0$; for the spoke player the strategy of forming one link and choosing $x_s > 0$ is optimal if and ony if $f'(y) = C'(x_s + k)$ and $f(y) - C(k + x_s) \geq f(\bar{x}) - C(\bar{x})$. We note that these inequalities are satisfied. To see this we first observe that $y > \bar{x}$. Indeed, if $y \leq \bar{x}$ then $f'(y) = C'(x_s + k) \geq f'(\bar{x}) = C'(\bar{x})$, so $x_s + k = x_h \geq \bar{x}$, which implies that $y = x_h + (n-1)x_s \geq \bar{x} + (n-1)x_s > \bar{x}$, a contradiction. So, $y > \bar{x}$ which implies that $x_h = k + x_s < \bar{x}$ and therefore the payoff of the spoke player is $f(y) - C(k + x_s) > f(\bar{x}) - C(k + x_s) > f(\bar{x}) - C(\bar{x})$, where the last expression is the best payoff that a spoke player can earn should he delete his link and become isolate. ■

Proof of Proposition 5.2: Let $s = (x, g)$ be a non-empty network equilibrium in which $x_i > 0$ for all i. Let $n(l)$ be the number of players who sponsor l links, $l = 0, ..., n-1$; also let $x(l)$ be the effort of a player who sponsors l links. Note that in a minimally connected network there are $n-1$ links and n nodes and therefore $n(0) > 0$. Since every player puts in positive effort, we have that $x(0) = x(l) + kl$. Thus, aggregate effort is $\sum_{l=0}^{n-1} n(l)x(l) = \sum_{l=0}^{n-1} n(l)(x_0 - kl) = nx_0 - k(n-1)$. Since $x(0)$ is part of equilibrium, it is the solution to $f'(nx(0) - k(n-1)) = C'(x(0))$.

For given k, let $s = (x, g)$ be a non-empty network equilibrium in which $x_i > 0$ for all $i \in N$; similarly, for given $k' > k$, let $s' = (x', g')$ be a non-empty network equilibrium in which $x'_i > 0$ for all $i \in N$. Let y and y' be the aggregate effort under s and s', respectively. We claim that $y > y'$. Suppose $y \leq y'$, then $f'(y) = C'(x(0)) \geq f'(y') = C'(x'(0))$ (where recall that $x(0)$ is the effort of a player who does not sponsor any links.); so, $x(0) \geq x'(0)$. But then, since $k' > k$, $y = nx(0) - (n-1)k \geq nx'(0) - (n-1)k > nx'(0) - (n-1)k' = y'$, a contradiction.

Next, fix k; for a given n, let $s = (x, g)$ be a non-empty network equilibrium with $x_i > 0$ for all $i \in N$; similarly, for given $n' > n$, let $s' = (x', g')$ be a non-empty network equilibrium with $x'_i > 0$ for all $i \in N$. We claim that $y' > y$. Suppose that $y' \leq y$, then $f'(y) = C'(x(0)) \leq f'(y') = C'(x'(0))$; so, $x(0) \leq x'(0)$. But then, $y = nx(0) - (n-1)k \leq nx'(0) - (n-1)k < n'x'(0) - (n'-1)k = y'$, where the last inequality follows because $nx'(0) - (n-1)k$ is increasing in n whenever $x'(0) > k$, which holds because $x'(l) = x'(0) - kl > 0$ for all $l \geq 1$. ■

Proof of Proposition 5.3: Suppose $k < c$ and let $s = (x, g)$ be an equilibrium. We claim that there exists an $i \in N$ such that $x_i = 1$ and that $x_j = 0$, $\forall j \neq i$. First, since $k < c$, there must be at least a player who chooses effort 1. Second, suppose both i and j choose effort 1. Then, it must be the case that $x_{i'} = 0$, $\forall i' \in N(i;\bar{g})$; for if a neighbor of i chooses effort 1, player i strictly gains by choosing effort 0. Since $x_{i'} = 0$, $\forall i' \in N(i;\bar{g})$, then $g_{il} = 0$ for all l. Hence, player i's payoffs in equilibrium s are $1 - c$. If player i chooses 0 and forms a link with j then he obtains $1 - k$. Since $k < c$, $1 - k > 1 - c$ and therefore s cannot be an equilibrium. Next, let $x_i = 1$ and $x_j = 0$, $\forall j \neq i$. Trivially, $g_{j'j} = 0$, $\forall j' \in N$, $j \neq i$, and, since $k < c$, every player $j \neq i$ has a link with i. This completes the proof for the case $k < c$. The proof for the case $k > c$ is trivial and therefore omitted. ■

Proof of Proposition 5.4: Suppose $k < c$ and suppose that $s = (x, g)$ is efficient. It is easy to see that the only links in g are between pair of players (i, j) with $x_i \neq x_j$. Also, if player i chooses 0 then player i has only one link with a player choosing 1. Indeed, if player i had two distinct links with two players choosing 1, then welfare can be made strictly higher by deleting one of the link. Hence, the total number of links are $(n - m)$, where m is the number of players choosing 1, and each player gets returns of 1. Then the social welfare is $n - mc - (n - m)k$. If $k > c$, this expression decreases with m and therefore $m = 1$, which implies the result. Suppose now that $k > c$. The above arguments show that if there are $m < n$ players choosing 1, and s is efficient then the social welfare is $n - mc - (n - m)k$, but then welfare can be increased by setting $m = n$, which implies the result. This concludes the proof. ■

References

[1] Adar E. and B. A. Huberman (2000), Free Riding on Gnutella, *FirstMonday*, October.

[2] Bala V. and S. Goyal (2000), A non-cooperative model of network formation, *Econometrica*, 68, 1181-1229.

[3] Beck, P., R. Dalton, S. Green, and R. Huckfeldt (2002), The Social Calculus of Voting: Interpersonal, Media and Organizational Influences on Presidential Choices, *American Political Science Review*, 96, 1, 57-73.

[4] Bergstrom, T. L. Blume and H. Varian (1986), On the private provision of public goods, *Journal of Public Economics*, 29, 25-49.

[5] Bramoulle, Y, D. Lopez, S. Goyal and F. Vega-Redondo (2004), Social interaction in anti-coordination games, *International Journal of Game Theory,* 33, 1-20.

[6] Bramoullé, Y. and R. Kranton (2007), Strategic Experimentation in Networks, *Journal of Economic Theory*, forthcoming.

[7] Cabrales, A., Calvó-Armengol, A. and Y. Zenou (2007), Effort and Synergies in Network Formation. Mimeo.

[8] Cabrales, A., and P. Gottardi (2007), Markets for Information: Of Inefficient Farewalls and Efficient Monopolies. Mimeo.

[9] Calvó-Armengol, A. and Y. Zenou (2004), Social Networks and Crime Decisions: The Role of Social Structure in Facilitating Delinquent Behavior. *International Economic Review*, 45(3), 939-958.

[10] Cho, M. (2007), Endogenous formation of networks for local public goods, *mimeo*, Penn State University.

[11] Cross, R., A. Parker, L. Prusak and S.P. Borgatti (2001), Knowing What We Know: supporting knowledge creation and sharing in social networks, *Organizational Dynamics*, 30, 2, 110-120.

[12] Cross, R. and L. Prusak (2002), The People Who Make Organizations Go-or Stop, *Harvard Business Review*, 104-112.

[13] Cross, R. and A. Parker (2004), The Hidden Power of Social Networks: Understanding How Work Really Gets Done In Organizations. Harvard Business School Press.

[14] Feick, L. F. and L. L. Price (1987), The Market Maven: A Diffuser of Marketplace Information, *Journal of Marketing*, 51, 1, 83-97.

[15] Galdwell M. (2000), *The Tipping Point: How Little Things Can Make a Big Difference.* Little Brown.

[16] Geissler G.L., and S.W. Edison (2005), Market Mavens' Attitudes Towards General Technology: Implications for Marketing Communications, *Journal of Marketing Communications*, 11, 2, 73-94.

[17] Goyal, S. (2007), *Connections: an introduction to the economics of networks.* Princeton University Press. October 2007.

[18] Goyal, S. and F. Vega-Redondo (2005), Network formation and social coordination, *Games and Economic Behavior*, 50, 178-207.

[19] Jackson, M. and B. W. Rogers (2007), Meeting Strangers and Friends of Friends: How Random Are Social Networks, *American Economic Review*, 97, 3, 890-915.

[20] Jackson, M. and A. Watts (2002), On the Formation of Interaction Networks in Social Coordination Games, *Games and Economic Behavior*, 41, 2, 265-291.

[21] Katz, E., and P. Lazersfeld (1955), Personal Influence. New York: The Free Press.

[22] Harrison G.W. and J. Hirshleifer (1989), An Experimental Evaluation of Weakest Link/Best Shot Models of Public Goods, *Journal of Political Economy*, 97, 1, 201-225.

[23] Hirshleifer, Jack (1983), From Weakest-Link to Best-Shot: The Voluntary Provision of Public Goods, *Public Choice,*, 41, 371-386.

[24] Hojman, D. and A. Szeidl (2006), Core and Periphery in networks, *Journal of Economic Theory*. Forthcoming.

[25] Huckfeldt, R., P. Johnson, and J. Sprague (2004), *Political Disagreement: The survival of diverse opinions within Communication networks.* New York: Cambridge University Press.

[26] Lazarsfeld, P., B. Berelson, and H. Gaudet (1948), The People's Choice. New York: Columbia University Press.

[27] Wiedmann K.P., G. Walsh and V.W. Mitchell (2001), The Mannmaven: an agent for diffusing market information, *Journal of Marketing Communications*, 7, 185-212.

[28] Williams, T.G. and M.E. Slama (1995), Market mavens' purchase decision evaluative criteria: implications for brand and store promotion efforts, *Journal of Consumer Marketing*, 12, 3, 4-21.

[29] Zhang J., M. Ackerman and L. Adamic (2007), Expertise Networks in Online Communities: Structures and Algorithms, *WWW2007*, Banff, Canada.