University of Essex

Department of Economics

Discussion Paper Series

No. 636 August 2007

The Law of the Few

Andrea Galeotti and Sanjeev Goyal

The Law of the Few

Andrea Galeotti*
Sanjeev Goyal†

August 19, 2007
First Draft.

Abstract

The *law of the few* refers to the following empirical phenomenon: in social groups a very small subset of individuals invests in collecting information while the rest of the group invests in forming connections with this select few. In many instances, there are no observable differences in characteristics between those who invest in information and those who invest in forming connections. This paper shows that the law of few naturally emerges in environments with *identical* rational agents.
We develop a strategic game in which players have the opportunity to invest in collecting information as well as in investing in bilateral connections with others. We find that *every* strict equilibrium of this game exhibits the 'law of the few'. We also show that this pattern of social differentiations is efficient in some cases.

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

Keywords: Opinion leaders, Strategic Network Formation, Network Games.

*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 the participants at the seminars in Cambridge, Essex and at the conference on "The Formation of Social Networks", Paris 2007, for their valuable comments.

1 Introduction

In their classic study, Katz and Lazersfeld (1955) found that in making purchase decisions across a range of products, most individuals relied on the information they received from a small group of other individuals. They called these few individuals, *opinion leaders*. We refer to the phenomenon of a small subset of individuals collecting information while the rest of the group invests in connections with this select few as the *Law of the Few*. Over the years, a large body of research in different subjects – which include political science, sociology, and marketing – has confirmed the generality of the findings of Katz and Lazersfeld (1955).[1] Most of this literature, after confirming the law of the few, has examined the individual characteristics of opinion leaders and in many instances they found somewhat surprisingly that there were no significant observable differences between the opinion leaders and the rest.[2] Of course, in these studies one cannot exclude that such patterns of social differentiations are caused by the presence of unobservable heterogeneity. Nevertheless, we think it is worth exploring whether the law of the few can be obtained in environments with *identical* rational agents.

There are three key ingredients in social information gathering: individuals can choose how much to invest in collecting information and how much to invest in forming connections, there are costs to each of these activities, and there is no difference across individuals with regard to these costs and the corresponding benefits of information. The incentives to acquire information and to form connections will naturally depend on the relative costs of doing so. Moreover, there is also an issue of coordination: if no one else acquires any information then a player may have no choice but to acquire information himself. The main result of the paper is that these economic pressures together yield a clear cut prediction: *if the costs of forming links are lower than the costs of directly acquiring information then every strict equilibrium*

[1]The classic paper in marketing is Feick and Price (1987); for more recent work see, e.g., Geisser and Edison (2005), Wiedman, Walsh and Mitchel (2001), and Williams and Slama (1995). In political science the classic work is Lazarsfeld *et al.* (1955); more recent papers are Beck *et al* (2002), Huckfeldt and Sprague (1995).

[2]For example, Feick and Price (1987) showed that market mavens – individuals who are well defined across several product categories – are prominent and that their characteristics do not differ significantly from other individuals who are not well informed. Geisser and Edison (2005), Wiedman, Walsh and Mitchel (2001), and Williams and Slama (1995) arrive at similar conclusions.

exhibits the law of the few: a small fraction of players acquire information directly while the rest of the players invest in forming connections and only acquire information indirectly.

We now briefly sketch the main arguments underlying this finding. In our model, the returns from information are increasing and concave in total information received by a player, the costs of acquiring information are linear in amount of information while the costs of forming connections are linear in the number of connections. Under reasonable restrictions on the marginal returns we get the property that on his own an individual will choose an interior information level, say 1. This leads to our *first* observation: in any equilibrium the total information available to an individual must be at least 1. Moreover, if a player exerts any effort at all then the total information he gets must indeed be equal to 1. If the total information was less than 1, then he gains by increasing effort, since marginal returns are larger than marginal cost. Similarly, if own effort is positive but total information is greater than 1, then the player can strictly increase payoffs by lowering effort.

The *second* observation constitutes the key to the result: in every strict equilibrium, the sum total of information directly acquired in a society is equal to 1. There are two steps in the proof of this equilibrium property. The first step shows that if any player chooses 1, then it is optimal for everyone to choose 0 and simply link to this player. So let us consider an equilibrium in which no player chooses 1. The second step shows that if two players choose positive effort and they are neighbors then they must have common neighbors in a strict equilibrium.[3]

The basic idea underlying this proof is the following. Suppose i and j are neighbors, they both choose positive effort and l is a neighbor of i but not of player j. Suppose also that player l chooses higher effort than player j. To fix ideas, figure 1 illustrates this configuration. In the figure an arrow starting from i and pointing at j signifies that i sponsors the link with j. Since i links with j then the costs of the link must be smaller than the costs of the information that i accesses from j. Similarly, since player l is not a neighbor of j, it must be the case that the costs of a link with j are higher than the costs of information that j has

[3]In a network, j and i are said to be neighbors if there is a direct link between them.

acquired on his own. Together, this tells us that the costs of j's information equal the costs of a link. This makes player i indifferent between keeping the link with j and substituting the link with own additional effort. Hence, this configuration cannot be sustained in a strict equilibrium. Similar considerations are obtained in other configurations between these three players. Thus, this second step implies that all linked players choosing positive effort share the same neighbors and if aggregate effort were higher than 1 this would contradict our first observation. Hence, in any strict equilibrium the aggregate information in the society must be exactly 1.

The *third* observation is that every equilibrium network has the inter-linked stars architecture. Figure 2 illustrates this architecture. Roughly speaking, an *inter-linked stars architecture* contains a set of hub players (the black nodes in figure 2) who are linked to everyone while every other player (the white nodes in figure 2) forms a link with each of the hubs.[4] To see how this comes about suppose that the number of players choosing to directly acquire information is less than n. We know that in any strict equilibrium the total effort is 1, and that every player has at least 1 unit of information. It then follows that every positive effort player will form a link with every other positive effort player, while the zero effort players will form a link with all the positive effort players. Thus any equilibrium network will have the inter-linked stars architecture.

The *final* observation combines the above steps to derive the law of the few property for every equilibrium. For any cost of acquiring information and any cost of forming link it must be the case that if player i forms a connection with j then the cost of linking must be less than the cost of providing the effort accessed from j. This, however, gives us a lower bound on the size of effort of j. Since aggregate effort is 1, the maximum number of players who choose positive effort is bounded above and is independent of the number of players. This means that the fraction of players who choose positive effort can be made arbitrarily small by suitably raising the number of players, and the Law of the Few obtains.

We also study the social efficiency of different patterns of efforts and linking. We find

[4] For a formal definition of this architecture, see section 2.

that when the costs of forming links are low, then the star network in which the central player acquires all the information is socially efficient.[5] The intuition for this result is that if many individuals make investments in acquiring information then social efficiency also entails the formation of several links among these players. Since links are costly and the costs of information gathering is linear in effort, concentrating all efforts with one player economizes on costs and yields the efficient outcome. In contrast, if costs of forming links are high, the socially efficient outcome is characterized by each player acquiring information and no player forms links. Comparing socially optimal outcomes with equilibrium outcomes we can conclude that for low costs of linking in equilibrium players under-invest in effort, for moderate costs of linking in equilibrium there is under-investment and under-connectivity, while if costs of linking are sufficiently high then equilibria are socially efficient.

Finally, we also study a discrete version of this model, which is called best shot game. Players can only choose either to provide a unit of effort at a cost c or not providing effort at all. A player gets 1 if he accesses at least a unit of effort, otherwise the returns are 0.[6] Here we confirm that if the costs of linking are lower than the costs of providing effort in every equilibrium the law of the few obtains. In contrast with the continuous model, in the discrete model every equilibrium is socially efficient.

The main contribution of our paper is to develop a simple model of strategic investments in information collection and link formation which can address the empirical finding of the Law of the Few. Our analysis shows that in settings with identical rational agents, a combination of simple economic factors – the relative costs of acquiring information versus the costs of forming links – and strategic interaction together provide a simple explanation for this law.

From a theoretical point of view, our paper is a contribution to the recent theory of networks. We develop a simple game in which individuals choose investments in information acquisition as well as decide with whom to form connections with a view to accessing the

[5]The star architecture is a network in which there is one player, the hub, who is linked with all other players, the spokes, and there are no other links. Figure 3 contains an example of this architecture.

[6]The best-shot game is a good metaphor for situations in which there are significant externalities between players' effort. For a discussion of best-shot games within the contexts of public good games see, e.g., Hirshleifer (1983) and Harrison and Hirshleifer (1989).

information acquired by these individuals. Our model combines the approach to link formation introduced in Bala ad Goyal (2000) with the approach to the study of local public goods developed in Bramoulle and Kranton (2007).[7] As the above discussion illustrates, this combination yields a tractable framework and sharp predictions. A recent paper by Cabrales, Calvo-Armengol and Zenou (2007) also presents a model of private investments and network formation. There are two key differences between our paper and their papers. We have a model in which individuals decide on individual specific links while in their papers investments in links are *not* individual specific. This implies that the strategy set of players and the methods of analysis are completely different. The second difference is that in their models individuals are ex-ante different, while in our paper the focus is on understanding how significant differentiation and the law of the few can arise in settings with identical individuals.[8]

The rest of the paper is organized as follows. Section 2 develops the model, while section 3 contains the main results. Section 4 considers two extensions. The first extension studies the effect of richer patters of spill overs. The second extensions studies a best shot game. Section 5 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 $X \in [0, +\infty)$. 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. 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 $g = (g_1, g_2, ..., g_n)$.

[7] The literature on network formation and the literature on games played on fixed networks are both extensive and rich. For a survey of this work see Goyal (2007).

[8] In this paper, our interest is in situations where information sharing is a social activity; for a study of situations in which players can charge prices for their information, see Cabrales and Gottardi (2007).

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 $\eta_i(g) = |N^d(i;g)|$.

The closure of g is a non-directed network denoted $\bar{g} = cl(g)$, where $\bar{g}_{ij} = \max\{g_{ij}, g_{j,i}\}$ 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

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

where $c > 0$ reflects the cost of effort and $k > 0$ is the cost of linking with one other person. 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 $\hat{y} = \arg\max_{y \in X} f(y) - cy$.

For any strategy profile s, let $s_{-i} = (s_1, ...s_{i-1}, s_{i+1}, ..., s_n)$, be the strategies of all players other player i. 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:

$$W(s) = \sum_{i \in N} \Pi_i(s) \tag{2}$$

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 $\{\bar{g}_{ij_1} = \bar{g}_{j_1,j_2} = ... = \bar{g}_{i_n,j_m} = 1\}$. 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 $\bar{g}_{ij} = 0$, $\forall j \in N$. 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 an inter-linked stars 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 an *inter-linked stars* network there are two groups of players, $\mathbf{N}_1(\bar{g})$ and $\mathbf{N}_2(\bar{g})$, with the feature that $N_i(\bar{g}) = \mathbf{N}_2(\bar{g})$ for $i \in \mathbf{N}_1(\bar{g})$ and $N_j(\bar{g}) = N \backslash \{i\}$, for all $j \in \mathbf{N}_2(\bar{g})$. The star a special case of this architecture, in which $|\mathbf{N}_2(\bar{g})| = 1$ and $|\mathbf{N}_1(\bar{g})| = n - 1$. In an inter-linked star 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 3 illustrates inter-linked stars networks. In the figure there are $n = 8$ players; in each architecture the black nodes are the hubs (the set $\mathbf{N}_1$), while the white nodes are the spokes (the set $\mathbf{N}_2$).

3 Analysis

The focus of our analysis will be on the distribution of effort and linking activity across players in strict equilibria. Our main result has three parts. First, we show that if costs of linking are smaller than the costs of effort that a player would provide on his own, say $\hat{y}$, then in equilibrium the aggregate social effort will be equal to $\hat{y}$, irrespective of the number of players. Second, we show that the inter-linked stars network is the unique equilibrium architecture. In this network every player who exerts a positive effort is a hub, while the no effort players are the spokes. Third, the set of hub players is very small relative the total number of players, i.e., the law of the few obtains.

Given any equilibrium $s = (x, g)$, define $I(s) = \{i \in N | x_i > 0\}$ as the set of players who choose a positive effort.

Theorem 3.1 *Suppose payoffs are given by (1). If $k < c\hat{y}$, then in every strict equilibrium $s^* = (x^*, g^*)$, $\sum_{i \in N} x_i^* = \hat{y}$. Every strict equilibrium has the interlinked stars architecture and hub players exert positive efforts while the spokes choose zero effort. Finally, for given c and k, with $k < c\hat{y}$, in any strict equilibrium s^* the ratio $|I(s^*)|/n$ can be made arbitrarily close to 0 by raising n. If $k > c\hat{y}$ then there exists a unique equilibrium: every player exerts effort $\hat{y}$ and no one forms any links.*

We now briefly sketch the main arguments underlying the proof of this result. The focus will be on the case where $k < c\hat{y}$. Define $y_i = x_i + \sum_{j \in N_i(\bar{g})} x_j$. The first step in the proof exploits the assumption that $f(0) = 0$, $f'(0) > c$ and $\lim_{y \to \infty} f'(y) < c$, to show that in every equilibrium $y_i \geq \hat{y}$ and if $x_i > 0$ then $y_i = \hat{y}$. Note that if $y_i < \hat{y}$, then a player gains by increasing effort, since marginal returns are larger than marginal cost. Similarly, if $x_i > 0$ and $y_i > \hat{y}$, then player i can strictly increase payoffs by lowering effort.

The key step in the proof shows that in any strict equilibrium $\sum_{i \in N} x_i = \hat{y}$. This relies in the following equilibrium properties: in any strict equilibrium every linked pair of positive effort players must share the same neighbors. To see the intuition underlying this suppose that i and j exert some positive efforts and that $g_{ij} = 1$. Suppose also that $x_i \leq x_j$ and that player i has a neighbor, say l, who is not linked with j. Figure 1 illustrates a possible configuration between these three players.

First note that since $x_i \leq x_j$ then $g_{li} = 0$, otherwise player j would weakly gain by switching the link from player i to player j. Hence, $g_{il} = 1$; this immediately implies that the costs of a link sponsored to l are sufficiently low, namely $k < cx_l$. The fact that $k < cx_l$ also implies that the effort of player j must be strictly less than the effort of player l, i.e., $x_j < x_l$. Indeed, if j were providing higher effort than l, then, since $k < cx_l$, player j would strictly gain by forming an additional link with player l and reducing his own effort to $x_j - x_l$.

However, given that $x_j < x_l$, then l may have an incentive to form a link with j and reducing his own effort to $x_l - x_j$. This is not profitable only if the effort provided by j is sufficiently low, namely $k > cx_j$. But note that in this case j must have sponsored a link to i and since j is not linked to l then i must exert strictly higher effort than l, $x_i > x_l$. Since the effort of player j is lower than the effort of l, we conclude that the effort of i is strictly higher than the effort of j, which is in contradiction with the initial hypothesis. This proves that every neighbor of i must also be a neighbor of j. The reverse then follows by noting that if j were accessing the effort of some player not in the neighbor of i, then player j would access from his own neighbor strictly more effort than what player i would access from his own neighbor, but then player j should provide less effort than player i, which is in contradiction with our initial hypothesis.

It is then clear that since every linked pair of players $i, j \in I(s)$ share the same neighbors, the players in $I(s)$ constitute a clique and this implies that the total effort must equal $\hat{y}$.

The third step in the proof shows that if $\sum_{i \in N} x_i = \hat{y}$ then an equilibrium network has the inter-linked stars architecture. To see why this is true, let us focus on the case $|I(s)| < n$. Since in any strict equilibrium $\sum_{i \in N} x_i = \hat{y}$, it follows that for every pair of players $i, j \in I(s)$, $\bar{g}_{ij} = 1$. Since $k > 0$, for every $i \in N$, and for every $l \notin I(s)$, $g_{il} = 0$. Thus, for every $l \notin I(s)$, since $x_l = 0$ it must be the case that and $g_{li} = 1$ for all $i \in I(s)$. This establishes the required architecture of strict equilibrium networks.

The last step in the proof derives the law of the few property for every equilibrium. For any c and k it must be the case that if i links with j then the cost of link must be less than the cost of providing the effort accessed from j, in other words $cx_j > k$. This, however, 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 equilibrium, the ratio $|I(s)|/n$ can be made arbitrarily small by suitably raising n.

Theorem 3.1 obtains the law of the few, for large n. What can we say about number of players who exert effort for fixed n? The following result addresses this issue.

Proposition 3.1 *Suppose payoffs are given by (1). If $k < \hat{y}c$, then there exists an equilibrium in which the network is a star, the hub player exerts effort $\hat{y}$, all the other players exert effort* 0 *and each forms a link with the hub player. Moreover, if $\frac{k}{\hat{y}} \in (c/2, c)$ then this is the unique strict equilibrium outcome.*

If all players choose zero effort and link with player i, then it is clearly a best response for player i to choose $\hat{y}$. For a spoke player the payoff is $f(\hat{y}) - k$. Exerting effort while maintaining the link is not profitable under the assumptions on $f(.)$. Deleting the link is not profitable since $k < \hat{y}c$. Next note from Theorem 3.1 that every strict equilibrium is an inter-linked star with every player who exerts effort being a hub. However, if $k > \hat{y}c/2$ then a link is only profitable if a player chooses effort $x_i > \hat{y}/2$. Since sum of efforts is equal to $\hat{y}$, in equilibrium at most one player can exert effort.

We now turn to the social welfare of equilibrium networks. We first observe that if $k < \hat{y}c$ then in any strict equilibrium the sum of total effort is $\hat{y}$, each player accesses exactly $\hat{y}$ units of effort, but the number of links vary. Given the linearity of costs of effort as well as the costs of linking it then follows that these equilibria can be ranked by the number of links they contain. In particular, since the star minimizes the number of links, it is the most efficient equilibrium. The second observation is that in every equilibrium, every player must access $\hat{y}$ of effort. This observation follows from Proposition 5.1 which is presented in the appendix. In that proposition we provide a partial characterization of all equilibria in this game. Since every player must access effort $\hat{y}$ in every equilibrium, it follows that in every equilibrium the aggregate gross returns are $nf(\hat{y})$. The most efficient equilibrium will clearly minimizes the total costs of effort and total costs of forming links, which immediately leads to the star architecture where the hub provides all the effort. These observations are summarized in the following proposition.

Proposition 3.2 *Suppose payoffs are given by (1). If $k < \hat{y}c$, then the strict equilibria can be ranked by the number of links they contain. Furthermore, the efficient equilibrium is a star where the hub provides effort $\hat{y}$, every spoke provides effort* 0 *and forms a link with the hub.*

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 3.3 *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.*

The value of $\bar{k}$ is obtained by equating the social welfare attained by the two configurations presented in Proposition 3.3 and it is formally derived in the appendix. To illustrate more in details the trade-off between equilibrium and efficiency, consider the following example. Suppose $c = 1/2$ and $f(y) = \ln(1 + y)$. In this case $\hat{y} = 1$, while $\tilde{y} = 2n - 1$. In figure 4 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. In the remaining region equilibrium outcomes coincide with socially optimal outcomes.

4 Extensions

In this section we consider two extensions of the model presented in Section 2. The first extension studies a situation in which the effort provided by a player may spill over along links in the network. The second extension considers a model in which effort is a discrete choice.

4.1 General Decay Model

The model in section 2 assumes that the efforts of players only spill over on direct neighbors. A more general model is one where efforts spill over along links and the intensity of these synergies depends on the distance in the network between players. We now extend the model presented in section 2 to allow for a richer patterns of effort's externalities.[9]

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 not such path exists, the distance is set 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 $\bar{g}$. For a strategy $s = (x, g)$ the total amount of efforts that player i has is given by $y_i = x_i + \sum_{l=1}^{n-1} \sum_{j \in N^l(i;\bar{g})} a_l x_j$, where $a_1, a_2, ..., a_{n-1}$ are weights measuring the intensity of spill overs at different distances. We shall assume that $a_l \in [0, 1]$ for all $l = 1, ..., n-1$ and that $a_1 \geq a_2 \geq ... \geq a_{n-1}$; this last assumption signifies that spill overs are decreasing in the geodesic distance.

The payoffs to player i under strategy profile $s = (x, g)$ can be rewritten as follows,

$$\Pi_i(s) = f(y_i) - cx_i - \eta_i(g)k, \tag{3}$$

and the assumptions on c, k and $f(\cdot)$ are the same as in section 2. Note that the model in section 2 is obtained by setting $a_1 = 1$ and $a_l = 0$ for all $l > 1$.

The following proposition provides some preliminary results on the effect of allowing spill overs along links.

Proposition 4.1 *Suppose payoffs are given by* (3)*.*

I Suppose $a_l = 1$ for all $l = 1...., n-1$. If $k < c\hat{y}$ then $s = (x, g)$ is an equilibrium if and only if (*a*) *aggregate effort equals $\hat{y}$,* (*b*) *$\bar{g}$ is minimally connected and* (*c*) *if $g_{ij} = 1$ then $k \geq cy_j$. Moreover, the star where the hub chooses $\hat{y}$, every spoke chooses* 0 *and forms a link with the hub is always an equilibrium.*

[9] We model spill overs following Hojman and Szeidl (2006).

II *Suppose* $a_1 = 1 > a_2$. *If* $k < c\hat{y}$, *then there exists a strict equilibrium in which the network is a star, the hub chooses* $\hat{y}$, *every spoke chooses* 0 *and forms a link with the hub.*

Part I of Proposition 4.1 covers the extreme situation in which efforts perfectly spill over along links. In this case it is clear that in every equilibrium the aggregate effort must equal the effort that a player would provide on his own and that the network must be connected and minimal. Equilibrium condition (c) says that not every minimally connected network may be part of an equilibrium. Indeed, it must be the case that if the costs of a link, say from i to j, must be lower than the effort that i accesses via j. This implies that either player j provides enough effort on his own, or that player j allows i to access the effort provided by other players. This suggests that when the costs of a link are sufficiently high (sufficiently closed to $c\hat{y}$), then even if in equilibrium there may be many players who provide efforts, only few players provide most of the total effort collected in the group. The following example illustrates this idea.

Consider a star architecture with 4 spokes and suppose that the hub chooses 0 and each spokes choose $\hat{y}/4$. First, suppose also that the hub forms a link with each spoke. Clearly, if $k > \hat{y}/4$ this configuration cannot be an equilibrium. Second, consider now that every spoke forms a link with the hub. In this case, even if the hub does not provide effort, he allows each hub to access $3\hat{y}/4$ units of effort. Clearly, if $k \leq 3c\hat{y}/4$ this configuration cannot be part of an equilibrium. In contrast, note that if the hub provides all the effort then this is always an equilibrium.

Finally, Part II of Proposition 4.1 covers the case in which the effort accessed from direct neighbors is as valuable as own effort, while the effort accessed from non-neighbors players is less valuable. In this case the star architecture where only the hub invests in effort is always a strict equilibrium. It is also possible to check that inter-linked stars are equilibria for appropriately chosen levels of costs of linking and effort.

4.2 The Best Shot Game

In this section we study a model similar to the model presented in section 2, but where a player can either acquire information at a cost c or he does not provide effort at all, i.e. $X = \{0, 1\}$. We assume that the returns to a player from acquiring information are $f(y_i) = 1$ if $y_i \geq 1$, otherwise $f(y_i) = 0$, where recall that $y_i = x_i + \sum_{j \in N(i;\bar{g})} x_j$. We assume that $c < 1$. This specification resembles the best shot game which has been widely studied in economics.[10] The following proposition characterizes the equilibria in the best shot game.

Proposition 4.2 *Suppose $X = \{0, 1\}$. If $k < c$ then every equilibrium has a star architecture, the hub chooses 1, every spoke chooses 0 and forms a link with the hub. If $k > c$ then there exists a unique equilibrium: every player chooses 1 and no one forms any links.*

The proof of this proposition relies on the observation that if $k < c$ then only one player can provide effort. Suppose, on the contrary, that players i and j provide effort. If they are neighbors, then player i would strictly gain by choosing effort 0. This implies that each player belonging to i's neighbor does not provide effort. But then player i would strictly gain by choosing effort 0 and linking up with player j.

There are two remarks we would like to emphasize. First, Proposition 4.2 shows that even if players can choose a discrete amount of effort the law of the few obtains. We note that this is true in a more general model where the returns to a player are: $f(y) = 1$ if $y \geq z$, otherwise 0, $z \geq 1$. When $z > 1$ this specification is reminiscent of the weakest link model studied within the contexts of public good games by, among others, Harrison and Hirshleifer (1989). Note that if $z > 1$, then the efforts of players are complements if $y < z$, while strict substitutes if $y \geq z$. In this model, for low costs of linking, in every equilibrium the sum of total efforts would be z and every equilibrium has the inter-linked stars architecture, hub players choose 1 and every spoke chooses 0.[11]

[10] The best-shot game is a good metaphor for situations in which there are significant externalities between players' effort. For a discussion of best-shot games within the contexts of public good games see, e.g., Hirshleifer (1983) and Harrison and Hirshleifer (1989).

[11] A full characterization of this "weakest-link" public good model is available from the authors upon requests.

The second remark is that in this model every equilibrium is efficient. This is in sharp contrast with the case in which effort is a continuous variable.

Proposition 4.3 *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.*

5 Concluding Remarks

We have defined the law of the few as the empirical phenomenon of a small subset of individuals collecting 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 strategic investments in information collection and link formation in which the law of the few emerges as an equilibrium outcome with identical rational players. We also studies the efficiency properties of these patterns of social differentiation.

From a theoretical point of view, our paper combines the approach of link formation introduced in Bala and Goyal (2000) with the approach to the study of network games with strategic substitutes developed in Bramoulle and Kranton (2007). On the one hand, the main drawback of the existing literature on strategic network formation is that the benefits that players obtained when belonging to a certain network are primitives of the model. That is, the architecture of the network influences with whom a player would like to link up, but it does not influence other decision variables, such as provision of effort, collecting information and alike, which naturally also determine the value of the network. On the other hand, the existing literature on network games assumes that the network of relations is given and focuses on how the location of a player in the network affects his behavior. In many instances, both dimensions are endogenous: individuals form connections with others depending on their behavior and the behavior of individuals depends on the social network. This paper shows that the combination of these two approaches yields a tractable framework and sharp predictions.

Before concluding we would like to make a remark on the implications of our results for the design of prevention policies interventions. Many social programmes attempt to create awareness among individuals about different risk behavior that can lead, for example, to sexually transmitted disease. Our analysis suggests that the data collection of interpersonal communication networks of a community is key to design effective prevention policies interventions.[12] As a very simple illustration suppose that a government realizes that a particular community lacks information on how to prevent the transmission of a particular diseas. Suppose that the policy of the government is to contact and inform 1 individual with the hope that the information will spread among other community members. Without knowledge of the network, the government will choose the individual randomly and for a large community almost surely that individual will not be an opinion leader. In this case every dollar that the government spends to inform the individual will only spill over to a small subsets of the community. On the other hand, by collecting information about the communication network, for example by asking a subset of the community members to report "with whom they talk to" about a particular matter, the government can identify an opinion leader, the individual who receives most nominations. Each dollar spent on this opinion leader will then spill over to all community members.

[12]For example, in a recent report of the World Bank "The Africa Multi-Country AIDS Programm 2000-2006" there are many examples of effective prevention social programmes based on interpersonal communication networks. See also Valente *et al.* (2003) and Kelly *et al.* (1991) for a discussion about the empirical importance of prevention policies intervention which incorporate information on social networks.

Appendix A.

This appendix provides proofs of the results in section 3. We also provides Proposition 5.1 which provides a partial characterization of Nash equilibria. We star with the proof of Theorem 3.1. This proof consists of a number of steps and it is useful to present it as a sequence of lemmas. The first step in the proof obtains a general property of every equilibrium configuration. For a strategy profile $s = (x, g)$, define, with some abuse of notation, $y_i = x_i + \sum_{j \in N(i;\bar{g})} x_j$ as the total effort accessible to player i. Recall that $\hat{y} = \arg\max_{y \in X} f(y) - cy$.

Lemma 5.1 *In any equilibrium $s = (x, g)$, $y_i \geq \hat{y}$, for all $i \in N$. Moreover, if $x_i > 0$ then $y_i = \hat{y}$.*

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

If a player chooses $x_i = \hat{y}$ and $k < c\hat{y}$ then this leads to a specially simple equilibrium profile. The following lemma clarifies this point.

Lemma 5.2 *Suppose $k < c\hat{y}$. In any equilibrium $s = (x, g)$, if $x_i = \hat{y}$, then $x_j = 0$, for all $j \neq i$.*

Proof: 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 5.1 that $y_i = \hat{y}$. This also implies that every player in the neighbor of i must exert effort 0. Now consider j, with $x_j > 0$. This means that $\bar{g}_{ij} = 0$. It follows from Lemma 5.1 that $y_j = \hat{y}$. 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 5.1 we know that in equilibrium $y_j = \hat{y}$, 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 payoff 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. ■

Lemma 5.2 implies that, for given $k < c\hat{y}$, if some player chooses $\hat{y}$, then in any equilibrium aggregate effort is $\hat{y}$. We now turn to equilibria in which no player chooses $\hat{y}$. We show that in any strict equilibrium aggregate effort is also equal to $\hat{y}$. This is the key step in the proof of Theorem 3.1.

Lemma 5.3 *Suppose $k < c\hat{y}$. In every strict equilibrium $s = (x, g)$, $\sum_{i \in N} x_i = \hat{y}$.*

Proof: In view of Lemma 5.2 we can focus on the case where no player chooses $\hat{y}$. Recall that $I(s)$ is the set of players who choose positive effort in equilibrium s. We now show that if two players belonging to $I(s)$, say i and j, are linked, then every player in the neighbor of i who exerts positive effort also belongs to the neighboor of j, and *vice versa*.

Claim 1. Suppose s is a strict equilibrium. Let $i, j \in I(s)$ and $\bar{g}_{ij} = 1$. Then, 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 $\bar{g}_{i,j} = 1$, $i, j \in I(s)$, 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 5.1 that $y_i = \hat{y}$ 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 choose 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.

(I:) $x_j \geq x_l$. 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 choose efforts $\tilde{x}_j = x_j - x_l$. 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 $k < cx_l$.

(II:) $x_j < x_l$. Here we have two sub-cases. (IIa:) 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. (IIb:) 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. ■

It is now easy to complete the proof of Lemma 5.3. Consider the subgraph of g defined on players belonging to $I(s)$. If this subgraph is connected, then Claim 1 implies that it is a clique. In this case Lamma 5.3 immediately follows from Lemma 5.1. Next, suppose this subgraph is not connected and let C_1 and C_2 be two components. Claim 1 implies that each component is a clique and, from Lemma 5.1, the total effort in each component is $\hat{y}$. Let $i, i' \in C_1$ and $j, j' \in C_2$ with $g_{i,i'} = 1$ and $g_{j,j'} = 1$. Suppose, without loss of generality, that $x'_i \leq x'_j$; note that player i (weakly) gains by switching link from i' to j', a contradiction with the hypothesis that s is a strict equilibrium. This concludes the proof of Lemma 5.3. ■

We are now ready to complete the proof of Theorem 3.1.

Proof of Theorem 3.1: We first consider the case $k < c\hat{y}$. From Lemma 5.3 we know that aggregate effort in any strict equilibrium is equal to $\hat{y}$. We now take up the issue of architecture. Suppose that $s = (x, g)$ is a strict equilibrium and it is not an inter-linked stars network. Clearly then there is no player i such that $x_i = \hat{y}$; for if there were such a player then from Lemma 5.2 the equilibrium network would be a star. Since $\sum_{i \in N} x_i = \hat{y}$, it follows that for all $i, j \in I(s)$, $\bar{g}_{ij} = 1$. If $|I(s)| = n$, the claim follows. Suppose $|I(s)| < n$; since $k > 0$, for every $i \in N$, and for every $l \notin I(s)$, $g_{il} = 0$. Thus, for every $l \notin I(s)$, $g_{l,i} = 1$

for all $i \in I(s)$. This proves that every strict equilibrium network has the inter-linked stars architecture.

We now consider the proportion $I(s)/n$. Fix c and k. Consider first a strict equilibrium in which there is some player $j \notin I(s)$. In such an equilibrium j forms a link with every $i \in I(s)$. For a player j to link to i, it must be true that $cx_i > k$. This means that $x_i > k/c$ for every $i \in I(s)$ and so the maximum number of players who can contribute is given by $(\hat{y}c)/k$. Clearly, for given c and k, the ratio $(\hat{y}c)/nk$ can be made arbitrarily small by suitably increasing n. Now consider an equilibrium in which $|I(s)| = n$. Note that for any $i \in I(s)$, if there is some $j \in N$ such that $g_{ji} = 1$ then $cx_i > k$. So the number of players who will have incoming links is bounded above by $(\hat{y}c)/nk$, as before. The rest of the players will have no in-coming links but since $|I(s)| = n$, and $\sum_{i \in N} x_i = \hat{y}$, it follows that $y_i < \hat{y}$, for all $i \in N$. This contradicts Lemma 5.1 and so $|I(s)| = n$ is not possible in a strict equilibrium, for large n.

We finally consider the case $k > c\hat{y}$. In any equilibrium $s = (x, g)$, $x_i \leq \hat{y}$, for all $i \in N$. But this means that if $k > c\hat{y}(c)$ then no player will form a link in equilibrium. Under our assumptions on $f(\cdot)$, it now follows that in equilibrium $x_i = \hat{y}$, for every $i \in N$. This completes the proof of Theorem 3.1. ■

Proof of Proposition 3.1: First we show that a star network in which the hub exerts effort $\hat{y}$ and all other players exert 0 effort but each forms a single link with the hub is an equilibrium. Suppose that $x_i = \hat{y}$ for some $i \in N$. The payoff to player i is $f(\hat{y}) - c > 0$. All players are linked with him, so forming links is clearly not profitable. A lowering of effort lowers payoff since $f(.)$ is strictly concave and $f'(\hat{y}) = c$. Consider a player $j \neq i$. His payoff is $f(\hat{y}) - k$. A possible deviation is to retain the link and increase effort, but this is not profitable since $f'(\hat{y}) = c$ and $f(\cdot)$ is strictly concave. Linking with a player $l \neq i$ is clearly not profitable since this player chooses effort 0. The only other alternative is to delete the link with player i and increase effort. The optimal effort level with zero links is $x_j = \hat{y}$, but then the payoff is $f(\hat{y}) - c\hat{y}$. Since $k < c\hat{y}$, this is less than the payoff $f(\hat{y}) - k$, which player j obtains in the stipulated profile.

Consider next the case where $k > (\hat{y}c)/2$; suppose that $s = (x, g)$ is a strict equilibrium. We know from Lemma 5.3 that $\sum_{i \in I(s)} x_i = \hat{y}$. Suppose $|I(s)| \geq 2$. Then from Theorem 3.1 we know that $\bar{g}_{ij} = 1$, for every pair $i, j \in I(s)$. But $g_{ij} = 1$ implies that $cx_j > k$, and under the hypothesis $k > (\hat{y}c)/2$ this means that $x_j > \hat{y}(c)/2$. There are two possible situations: one, $I(s) = n$ and two, $|I(s)| < n$. In the former case, there are $n(n-1)/2$ links and $n(n-1)/2 \times \hat{y}(c)/2 > \hat{y}$, for all $n \geq 3$; from Lemma 5.3 this contradicts the hypothesis that s is a strict equilibrium. Finally, if $|I(s)| < n$, then every player in $I(s)$ is linked to by every player outside $I(s)$. However then $\sum_{i \in I(s)} x_i > (|I(s)|\hat{y})/2 \geq \hat{y}$ so long as $|I(s)| \geq 2$; from Lemma 5.3 this contradicts the hypothesis that s is a strict equilibrium. Thus the only possible equilibrium involves $|I(s)| = 1$. The result now follows. ■

Proof of Proposition 3.2:

Suppose $k < c\hat{y}$. First consider strict equilibria. From Theorem 3.1 it follows that the sum of total efforts is $\hat{y}$, that $y_i = \hat{y}$ for all $i \in N$ and that g has an inter-linked stars 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 nonstrict Nash equilibrium is strictly lower that $SW(s^*)$. Suppose $s = (x, g)$ is a nonstrict Nash equilibrium. From Proposition 5.1 we know that $y_i = \hat{y}$ 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 nonstrict equilibrium then $y_i = \hat{y}$ 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 3.2 ■

Proof of Proposition 3.3: 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 not empty network and suppose that C is a component in g. Let $|C| \geq 3$ be the number of players in C. Suppose that x is the total effort exerted in component C. Then it follows that the total payoff of all players in component C is at most $|C|f(x) - cx - (|C| - 1)k$. Consider a star network with $|C|$ players in which the hub player alone exerts effort equal to x. It then follows that this configuration attains the maximum possible aggregate payoff given effort x. 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 x satisfy $mf'(x) = c$. 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 x satisfy $mf'(x) = c$.

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 x 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 of $\bar{k}$ 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 3.3. ■

The following proposition provides a partial characterization of Nash equilibria.

Proposition 5.1 *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 an inter-linked stars, hubs choose positive efforts and every spoke chooses effort* 0*. If $\sum_{i \in N} x_i > \hat{y}$ there are two possibilities:*

I *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 other player has $\Delta + 1$ links with positive effort players and there are not other links.*

II *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 $\frac{\hat{y}c}{k} - 1 < \eta < \frac{\hat{y}c}{k}$.*

Proof of Proposition 5.1:

First suppose that $\sum_{i \in N} x_i = \hat{y}$. In this case it is clear that $I(s)$ must be a clique. Furthermore, $\bar{g}_{i,j} = 0$ for all $j \notin I(s)$. Therefore, each player choosing 0 effort must sponsor a link with every positive effort players.

Hereafter, let $s = (x, g)$ be an equilibrium where $\sum_{i \in N} x_i > \hat{y}$. The proof now consists of two steps. In the first step, we characterize equilibria 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.

Step 1. We prove that if all positive effort players choose the same level of effort then s satisfies Part I of Proposition 5.1. Suppose $x_i = x$, $\forall i \in I(s)$. If $x = \hat{y}$, Lemma 5.2 implies that $|I(s)| = 1$ and therefore aggregate effort is $\hat{y}$, a contradiction. Assume $x \in (0, \hat{y})$; from Lemma 5.1 it follows that $y_i = \hat{y}$, $\forall i \in I(s)$. 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 $i \in I(s)$, $y_i = x + \Delta x = \hat{y}$, which implies that $x = \frac{\hat{y}}{\Delta+1}$. Since aggregate effort is strictly higher than $\hat{y}$ it follows that $\Delta < |I(s)| - 1$. Also, from Lemma 5.2 we know that $x < \hat{y}$, which implies that $\Delta \geq 1$. Thus, there exists two positive effort players who are neighbors. Since s is equilibrium, then $k \leq cx$. Also, since, by assumption, $\sum_{i \in N} x_i > \hat{y}$, there exists two positive effort players who are not neighbors. Since s is equilibrium, then $k \geq cx$. 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 I of Proposition 5.1.

Step 2. Let g' be the subgraph of g defined on $I(s)$. Let $C(\bar{g}')$ be a component $\bar{g}'$. 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 5.4 *Suppose that (A1) and (A2) holds 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 not links between low effort players and $k = c\bar{x}$, $\underline{x} = \hat{y} - \eta\bar{x}$ and $\frac{\hat{y}c}{k} - 1 < \eta < \frac{\hat{y}c}{k}$.*

Proof of Lemma 5.4 Without loss of generality label players in $C(\bar{g}')$, so that $x_1 \geq x_2 \geq \ldots \geq x_m$. (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 $i \leq l$.

To see this, 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 a 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 a 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. This is in contradiction with our hypothesis that $x_j < \bar{x} = x_l$. Hence, claim 1 follows.

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

Suppose not; then $\{1, .., l\}$ is a clique. This implies that for all $i \leq l$, there exists at least a player $j > l$ such that $\bar{g}_{ij} = 0$. If not, i would access everyone and from A1 it follows that $y_i > \hat{y}$, which contradicts Lemma 5.1. Next, 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 also i must be a neighbor of j'. Therefore, since $\{1, .., l\}$ is a clique, it follows that every neighbor of j is also a i's neighbor, and this contradicts the assumption that $x_j < \bar{x}$. Hence, claim 2 follows.

We can now conclude the proof of Lemma 5.4. First note that an implication of claim 1 and claim 2 is that $k = c\bar{x}$. Indeed, from claim 1 we know that there exists a player $j > l$ who sponsors a link to a player $i \leq l$. Since s is equilibrium, this implies that $k \leq c\bar{x}$. Similarly, claim 2 implies that there exists $i, i' \leq l$ such that $\bar{g}_{ii'} = 0$; since s is an equilibrium this implies that $k \geq c\bar{x}$. Hence, $k = c\bar{x}$.

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 $x_j > x_{j+1}$. Then, $y_j = x_j + \eta_j(g)\bar{x}$ and $y_{j+1} = x_{j+1} + \eta_{j+1}(g)\bar{x}$. Lemma 5.1 implies that $y_j = y_{j+1} = \hat{y}$, which holds whenever $x_j - x_{j+1} = (\eta_{j+1} - \eta_j)\bar{x}$. Since $x_j > x_{j+1}$, then $\eta_{j+1} - \eta_j \geq 1$, but then $(\eta_{j+1} - \eta_j)\bar{x} \geq \bar{x} > x_j - x_{j+1}$, where the last inequality follows because, by assumption, $x_j < \bar{x}$. Thus, all players $j > l$ chooses the same effort, say $\underline{x}$, and from Lemma 5.1 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 5.4. ■

We now conclude the proof. Recall that g' is the subgraph of g defined on positive efforts players. We need to consider two cases: one, $\bar{g}'$ is connected, and two $\bar{g}'$ is not connected. One, if $\bar{g}'$ is connected, then (A1) holds by assumption. If (A2) does not hold then step 1 applies and the proof follows. If (A2) holds then Lemma 5.4 applies. We then need to show that every player 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 of 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).

Two, suppose g' is not connected and let C_1 and C_2 be two arbitrary components. Here, note that for every $i, i' \in C_1$ and $j, j' \in C_2$ such that $g_{i,i'} = g_{j,j'} = 1$, then $x'_i = x'_j = x \geq x_i, x_j$ 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 reasonings it follows that $x_i, x_j \leq x$; $k = cx$ follows because i sponsors a link to i', thus $k \leq cx$, and i' does not sponsor a link to j', thus $k \geq cx$. 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 player receives at least a link, positive efforts players choose the same effort and the proof follows from step 1. Suppose in C_1 some player does not receive a link and effort is not homogeneous across players. If the aggregate effort in C_1 equals $\hat{y}$, then C_1 is a clique and therefore at most one player can only sponsor links. Since C_1 is a clique and aggregate effort is $\hat{y}$, this player will choose $\underline{x} = \hat{y} - (|C_1| - 1)x$. Alternatively, if in C_1 the aggregate effort is higher than 1, then lemma 5.4 applies. Similar considerations hold for any other components. The proof of Proposition 5.1 now follows from the combination of these observations. ■

Appendix B.

This appendix provides proffs of the results in section 4.

Proof of Proposition 4.1: We start with Part I. Suppose s satisfies the condition in the proposition. Take a player i; since $\sum_{j \in N} x_j = \hat{y}$, $\bar{g}$ is minimally connected, and $a_l = 1$ for all finite l, then $y_i = \hat{y}$, 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 $f(\hat{y}) - cx_i - \eta_i(g)$.

If $\eta_i(g) = 0$, then player i plays a best reply. Suppose $\eta_i(g) > 0$, then $g_{i,j} = 1$ 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. So, the only possible deviation to check is that player i deletes the link with j; since $k \leq cy_j$ player i does not gain by doing so. Hence, s is an equilibrium.

We now prove the reverse. Let $s = (x, g)$ be an equilibrium. Since $a_l = 1$ for all finite l, then, in equilibrium, every component of $\bar{g}$ must be minimal. Also, the aggregate effort in each component must be $\hat{y}$. 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 $\bar{g}$; note that it cannot be the case that a player $i \in C_1$ chooses $x_i = \hat{y}$. Suppose, on the contrary, that $x_i = \hat{y}$, then all i's neighbors choose effort 0 and sponsors a link to i, so i's payoffs are $f(\hat{y}) - c\hat{y}$, but, since $k < c\hat{y}$, 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; moreover, since C_1 is minimal it must be the case that at least a player who chooses positive effort also sponsors a link. So let $g_{jj'} = 1$, $j, j' \in C_1$ and $x_j \in (0, \hat{y})$. Then, since $\sum_{j \in N} x_j = \hat{y}$, $y'_j < \hat{y}$ and therefore player j strictly gains by switching the link from j' to a player j'' belonging to a different component. Thus, $\bar{g}$ is connected. Finally, it is readily seen that if $g_{ij} = 1$ and s is equilibrium, then $k \leq cy_j$. It is not easy to conclude the proof of Part I of Proposition 4.1. It is easy to verify Part II of Proposition 4.1. This concludes the proof of Proposition 4.1 ■

Proof of Proposition 4.2: 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 4.3: 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] Bala V. and S. Goyal (2000), A non-cooperative model of network formation, *Econometrica*, **68** , 1181-1229.

[2] 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.

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

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

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

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

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

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

[9] 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.

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

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

[12] Kelly, J.A., Lawrence, J.S., Diaz, Y. Stevenson, L.Y., Hauth, A.C., T.L. Brasfield, et al. (1991), HIV risk behavior reduction following intervention with key opinion leaders of population: An experimental analysis. *American Journal of Public Health*, 81, 168-171.

[13] 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.

[14] Hirshleifer, Jack (1983), From Weakest-Link to Best-Shot: The Voluntary Provision of Public Goods, *Public Choice,*, **41(3)**, 371-386.

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

[16] Huckfeldt, R., P. Johnson, and J. Sprague (2004), Political Disagreemeting: The survival of Diverse Opinions within Communication Networks. New York: Cambridge University Press.

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

[18] 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.

[19] 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.

[20] Valente, T., Hoffman, B., Ritt-Olson, A., Lichtman, K., and Johnson, C. (2003), Effects of a social network method for group assignment strategies in peer-led tobacco prevention programs in schools, *American Journal of Public Health*, 93, 11, 1837-1843.

Figure 1. $x_i,x_j,x_l>0,\ x_l>x_j$

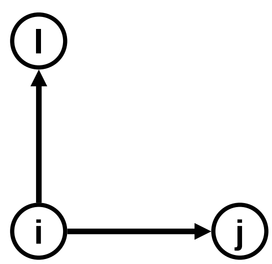


Figure 2. *Inter-linked stars architecture with three hubs, n=8.*

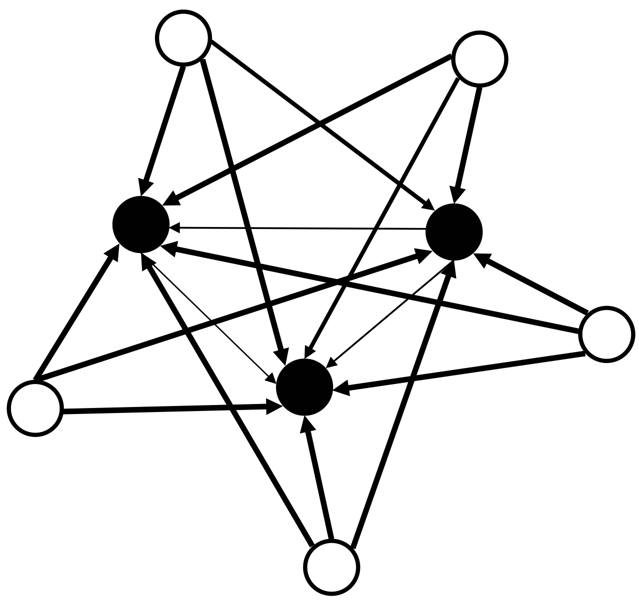
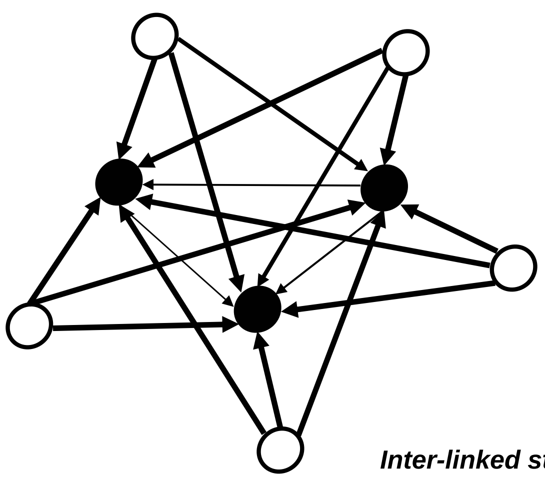
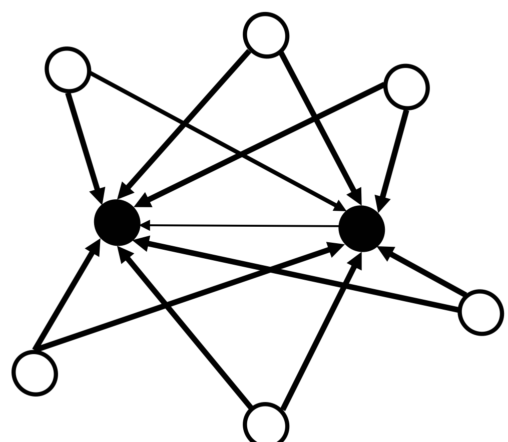


Figure 3. *Inter-linked stars architectures,*
n=8

Inter-linked stars architectures with 3 hubs.



Inter-linked stars architectures with 2 hubs.



Inter-linked stars architectures with 1 hub.

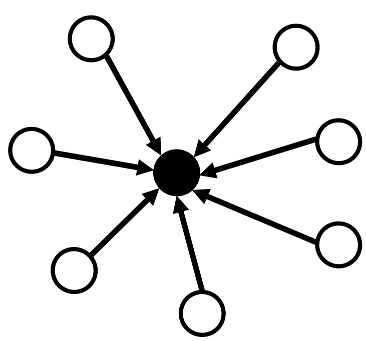


Figure 4. *Trade-off: Efficient Outcomes and Equilibrium Outcomes*