American Economic Review 100 (September 2010) 1468–1492
http://www.aeaweb.org/articles.php?doi=10.1257/aer.100.4.1468

The Law of the Few

By Andrea Galeotti and Sanjeev Goyal*

Empirical work shows that a large majority of individuals get most of their information from a very small subset of the group, viz., the influencers*; moreover, there exist only minor differences between the observable characteristics of the influencers and the others. We refer to these empirical findings as the Law of the Few. This paper develops a model where players personally acquire information and form connections with others to access their information. Every (robust) equilibrium of this model exhibits the law of the few.* (*JEL* D83, D85, Z13)

Individuals often have to make a choice between alternatives whose advantages are imperfectly known. In order to make an informed decision, they personally acquire information and gather information through social contacts. Personal acquisition of information is costly; similarly, creating and maintaining personal contacts takes time and resources. Rational individuals therefore compare the relative costs of these different sources of information. This paper explores the implications of such individual choices for social communication and the aggregate information available in a society.

Our point of departure is a series of empirical studies about information acquisition and communication in social groups. The classical early work of Paul F. Lazarsfeld, Bernard Berelson, and Hazel Gaudet (1948) and Elihu Katz and Lazersfeld (1955) investigated the impact of personal contacts and mass media on voting and consumer choice with regard to product brands, films, and fashion changes. They found that personal contacts play a dominant role in disseminating information which in turn shapes individuals' decisions. In particular, they identified 20 percent of their sample of 4,000 individuals as the primary source of information for the rest. Similarly, Lawrence F. Feick and Linda L. Price (1987) found that 25 percent of their sample of 1,400 individuals acquired a great deal of information about food, household goods, nonprescription drugs, and beauty products and that they were widely accessed by the rest.[1]

Research on virtual social communities reveals a similar pattern of communication. Jun Zhang, Mark S. Ackerman, and Lada Adamic (2007) study the Java Forum, an online community of users who ask and respond to queries concerning Java. They identify 14,000 users and

* Galeotti: Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK (e-mail: agaleo@essex.ac.uk); Goyal: Faculty of Economics and Christ's College, University of Cambridge, Sidgewick Avenue, Cambridge CB3 9DD, UK (e-mail: sg472@econ.cam.ac.uk). We thank the editor and three anonymous referees for useful suggestions. We are grateful to Bhaskar Dutta, Francis Bloch, Yann Bramoulle, Myeonghwan Cho, Matthew Jackson, Willemien Kets, Marco van der Leij, Brian Rogers and a number of seminar audiences for suggestions which have significantly improved the paper. We thank the ESRC for financial support via Grant RES-000-22-1901.

[1] Recent work in political science arrives at similar conclusions; see, e.g., Paul Beck et al. (2002) and Robert Huckfeldt, Paul E. Johnson, and John Sprague (2004). For recent work on information acquisition about products, see Gary L. Geissler and Steve W. Edison (2005). For evidence on patterns and the importance of informal communication in firms, see Robert Cross and Andrew Parker (2004).

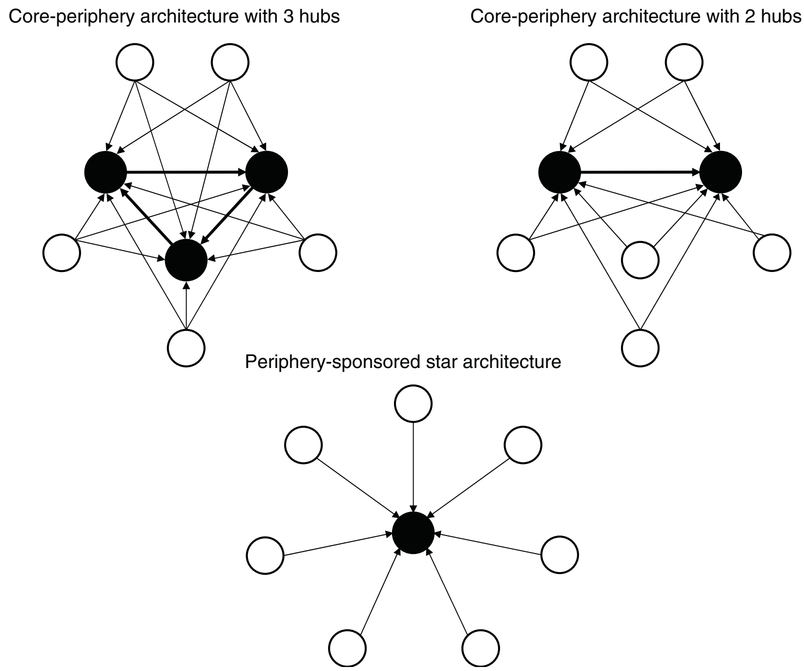

FIGURE 1. CORE-PERIPHERY NETWORKS

find that 55 percent of these users only ask questions, 12 percent both ask and answer queries, and about 13 percent only provide answers.[2]

The Law of the Few subsumes these empirical findings: in social groups, a majority of individuals get most of their information from a very small subset of the group, viz., *the influencers*. Moreover, research suggests that there are minor differences between the observable economic and demographic characteristics of the influencers and the others. We are thus led to ask: can the law of the few be understood as a consequence of strategic interaction among identical individuals?

We propose a game in which individuals choose to personally acquire information and to form connections with others to access the information these contacts acquire. Our main finding is that every (strict) equilibrium of the game exhibits the law of the few (propositions 1–3). The network has a *core-periphery architecture*; the players in the core acquire information personally, while the peripheral players acquire no information personally but form links and get all their information from the core players. Figure 1 illustrates equilibrium configurations.

We informally outline the ideas underlying this result. In our model, returns from information are increasing and concave, while the costs of personally acquiring information are linear. This implies that on his own an individual would acquire a certain amount of information, which we denote by 1. The second element in our model is the substitutability of information acquired by different individuals. This implies that if A acquires information on his own and receives information from player B, then in the aggregate he must have access to 1 unit of information (else he could strictly increase his payoff by modifying personal information acquisition). The third

[2] Eytan Adar and Bernardo A. Huberman (2000) report similar findings with regard to provision of files in the peer-to-peer network Gnutella.

and key element is that links are costly and rationally chosen by individuals. The implication of this is that if A finds it optimal to maintain a link with B, then so must every other player. Hence, the group of individuals who acquire information must be completely linked, and the aggregate information acquired in the society must equal exactly 1. Moreover, since linking is costly, A will link with B only if B acquires a certain minimum amount of information. Since total information acquired is 1, it follows that there is an upper bound to the number of people who will acquire information, and so the proportion of information acquirers in a large group is very small. Finally, we observe that since the aggregate information acquired in the group is 1, everyone who does not personally acquire information must be linked to all those who acquire information, yielding the core-periphery network.

The result mentioned above is derived in a setting where individuals are ex ante identical. A recurring theme in the empirical work is that influencers have similar demographic characteristics as the others. But this work also finds that they have distinctive attitudes which include higher attention to general market information and enjoyment in acquiring information (see, e.g., Feick and Price 1987). This motivates a study of the consequences of small heterogeneity in individual characteristics. Our main finding is that a slight cost advantage in acquiring information (or a greater preference for information) leads to a unique equilibrium in which the low cost (or high information need) individual player is the single hub: he acquires all the information, while everyone else simply connects with him (Proposition 3). Small heterogeneities thus help select individuals who play dramatically different roles in social organization.

In actual practice, we receive information from friends and colleagues which they have themselves received from others. We then extend the basic model to allow for indirect information transmission. Our main insight here is that indirect information transmission gives rise to a new type of influencer: *the connector* (Proposition 5). A connector is someone whose primary role is that of an "informediary": he does not acquire (much) information himself but connects individuals who personally acquire most of the information.

We now relate our paper to the theory of network formation as well as the theory of games played on fixed networks.[3] Our model builds on the approach to link formation introduced in Goyal (1993) and Venkatesh Bala and Goyal (2000) and the model of local public goods developed in Yann Bramoullé and Rachel Kranton (2007). We show that two economic ideas—(i) the substitutability between information acquired personally and information acquired by others, and (ii) the possibility of forming costly links with others who acquire information—explain the specialization in information acquisition and social communication reflected in the law of the few. Moreover, in line with the findings of empirical work, we show that a small difference—with regard to the costs of acquiring information, the need for information, or with regard to personal sociability—is sufficient to perfectly select influencers in a social group. We observe that existing models cannot explain the law of the few: the pure link formation model cannot account for the specialization in information acquisition, while the public goods model cannot account for the specific patterns of social communication.

Our results also resolve important theoretical questions in these two strands of the literature. First, in the public goods model with exogenous networks, Bramoulle and Kranton (2007) show that multiple equilibria typically exist, and these equilibria exhibit very different individual and aggregate information acquisition. In contrast, introducing link formation yields clear cut predictions on individual as well as aggregate information acquisition.

Second, in a model of network formation with pure local information sharing, the equilibrium network is either the complete or the empty network, depending on whether the cost of linking is

[3] For surveys of this literature, see Goyal (2007) and Matthew O. Jackson (2008).

lower or higher than the value of individual information. By contrast, we find the unique equilibrium is a core-periphery network with the hubs acquiring all the information. Thus endogenous information acquisition provides an alternative theoretical foundation for core-periphery (and star) networks. Furthermore, our model yields an interesting cost of link effect: an increase in this cost leads to a fall in the number of players who acquire information, an increase in the amount of information acquired by each player who acquires information, and a decrease in the total number of links. By contrast, changes in costs of linking have no effect on personal information acquisition in a model of pure link formation, since everyone has exogenously specified information (see, e.g., Bala and Goyal 2000; Galeotti 2006; Galeotti, Goyal, and Jurjen Kamphorst 2006; and Daniel Hojman and Adam Siezdl 2008).[4]

Nonrival information acquisition is an instance of the private provision of public goods. For global public goods, Theodore Bergstrom, Lawrence Blume, and Hal Varian (1986) showed that the contributors will be those with higher endowments. The key difference in our model is that access to public good is a matter of individual choice; it is costly and takes the form of bilateral connections. The findings with regard to the existence of information hubs and connectors and the core-periphery network structure are, to the best of our knowledge, novel in this literature.

The rest of the paper is organized as follows. Section I develops the basic model with local information flow, and Section II analyzes this model. Section III considers a model with indirect information transmission. Section IV discusses two important aspects of our model, linear costs of acquiring information and forming links and the link formation protocol, respectively. Section V concludes.

I. Model

Let $N = \{1, 2, \ldots, n\}$ with $n \geq 3$ be the set of players and let i and j be typical members of this set. Each player i chooses a level of personal information acquisition $x_i \in X$ and a set of links with others to access their information, which is represented as a (row) vector $\mathbf{g}_i = (g_{i1}, \ldots, g_{ii-1}, g_{ii+1}, \ldots, g_{in})$, where $g_{ij} \in \{0, 1\}$, for each $j \in N \backslash \{i\}$. We will suppose that $X = [0, +\infty)$ and that $\mathbf{g}_i \in G_i = \{0, 1\}^{n-1}$.[5] 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 information personally acquired by the other player.[6] The set of strategies of player i is denoted by $S_i = X \times G_i$. Define $S = S_1 \times \cdots \times S_n$ as the set of strategies of all players. A strategy profile $\mathbf{s} = (\mathbf{x}, \mathbf{g}) \in S$ specifies the personal information acquired by each player, $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, and the network of relations $\mathbf{g} = (\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_n)$.

The network of relations $\mathbf{g}$ is a directed graph; let G be the set of all possible directed graphs on n vertices. Define $N_i(\mathbf{g}) = \{j \in N : g_{ij} = 1\}$ as the set of players with whom i has formed a link. Let $\eta_i(\mathbf{g}) = |N_i(\mathbf{g})|$. The closure of $\mathbf{g}$ is an undirected network denoted by $\overline{\mathbf{g}} = cl(\mathbf{g})$, where $\overline{g}_{ij} = \max\{g_{ij}, g_{ji}\}$ for each i and j in N. In words, the closure of a directed network involves replacing every directed edge of $\mathbf{g}$ by an undirected one. Define $N_i(\overline{\mathbf{g}}) = \{j \in N : \overline{g}_{ij} = 1\}$ as the set of players directly connected to i. The undirected link between two players reflects bilateral information exchange between them.

[4] There is a small body of work which combines network formation and play in games (see, e.g., Goyal and Fernando Vega-Redondo 2005; Jackson and Alison Watts 2002; Bramoullé et al. 2004; and Antoni Calvó-Armengol and Yves Zenou 2004). The game studied and the questions addressed in the present paper are quite different from this literature, and so a detailed discussion of the relation with these papers is omitted.

[5] We have completely solved the model with discrete information choice variable. We find that the main results on equilibrium information acquisition and networks are robust to a change in action sets. The details of these derivations are available in a Web Appendix for the paper.

[6] The Web Appendix presents and analyzes a model in which link formation and information flow are both one way.

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

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

where $c > 0$ reflects the cost of information and $k > 0$ is the cost of linking with one other person. We assume pure local externalities: player i accesses only the information personally acquired by his immediate neighbors. Section III studies a model with indirect information transmission.

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 $\lim_{y \to \infty} f'(y) = m < c$. Under these assumptions there exists a number $\hat{y} > 0$ such that $\hat{y} = \arg\max_{y \in X} f(y) - cy$, i.e., $\hat{y}$ solves $f'(\hat{y}) = c$.

We now discuss the key elements of our model.

First, consider the returns from information. We may think of the action x as 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. Another possible interpretation is in terms of individuals choosing an action whose payoffs are unknown. Every individual has access to a costly sample of observations—which may reflect personal experience with a product or a technology. A link with another player then allows access to his personal experience. Under reasonable conditions, the returns from accessing more samples of information—own and others—are increasing but concave.[7]

Second, consider the protocol of link formation. We assume a person can form a binary link (it takes value one or zero) and that this link is formed once a cost is incurred. One possible interpretation of unilateral link formation and two-way exchange of information is that one player pays for a telephone call and the call involves an exchange of information. Alternatively, we may interpret the cost incurred in the formation of a link as a "gift" or a social favor that 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 $\mathbf{s} = (\mathbf{x}, \mathbf{g})$ are given by:

$$\hat{\Pi}_i(\mathbf{s}) = f\left(x_i + \sum_{j \in N_i(\bar{\mathbf{g}})} x_j\right) - cx_i - \eta_i(\mathbf{g})k + \sum_{j \in N_i(\mathbf{g})} g_{ji}k.$$

The last term involving transfers is independent of the strategy of i. It then follows that for all $\mathbf{s}_{-i} \in S_{-i}$, and $\mathbf{s}_i, \mathbf{s}_i' \in S_i$, $\hat{\Pi}_i(\mathbf{s}_i, \mathbf{s}_{-i}) \geq \hat{\Pi}_i(\mathbf{s}_i', \mathbf{s}_{-i})$, if and only if $\Pi_i(\mathbf{s}_i, \mathbf{s}_{-i}) \geq \Pi_i(\mathbf{s}_i', \mathbf{s}_{-i})$. Therefore our methods of analysis and our equilibrium findings with payoffs (1) carry over to the alternative model, where link formation costs are transfers from one individual to another.

Third, we assume that a player has no interest in misleading others and that there are no incentives for refusing to share information. This assumption is reasonable in a number of contexts such as informal social communication between individual consumers.

[7] For a detailed discussion of these examples on information sharing, see Bramoullé and Kranton (2007) and Hojman and Szeidl (2008).

Finally, we assume that there is no bargaining and no prices for information. We are aware that this is a strong assumption. As a first step toward incorporating prices, in Section IV, we study the case where players can ask for transfers in exchange for sharing information.[8]

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

$$\Pi_i(\mathbf{s}_i^*, \mathbf{s}_{-i}^*) \geq \Pi_i(\mathbf{s}_i, \mathbf{s}_{-i}^*), \ \forall \mathbf{s}_i \in S_i, \forall i \in N. \tag{2}$$

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. For any profile $\mathbf{s}$ social welfare is given by:

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

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

We say that there is a path in $\overline{\mathbf{g}}$ between i and j if either $\overline{g}_{ij} = 1$ or there exist players $j_1, \ldots, j_m$ distinct from each other and i and j such that $\{\overline{g}_{ij_1} = \overline{g}_{j_1 j_2} = \ldots = \overline{g}_{j_m j} = 1\}$. A network $\overline{\mathbf{g}}$ is connected if there exists a path between every pair of players; we say that a network $\overline{\mathbf{g}}$ is minimally connected if it is connected and there exists only one path between every pair of players. In a *core-periphery* network there are two groups of players, $\hat{N}_1(\mathbf{g})$ and $\hat{N}_2(\mathbf{g})$, with the feature that $N_i(\overline{\mathbf{g}}) = \hat{N}_2(\overline{\mathbf{g}})$ for all $i \in \hat{N}_1(\overline{\mathbf{g}})$, and $N_j(\overline{\mathbf{g}}) = N \backslash \{j\}$ for all $j \in \hat{N}_2(\mathbf{g})$. 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. A core-periphery network with a single hub is referred to as a periphery-sponsored star. Figure 1 illustrates core-periphery networks. There are $n = 8$ players; in each architecture the black nodes are the hubs (the set $\hat{N}_2(\overline{\mathbf{g}})$), the white nodes are the spokes (the set $\hat{N}_1(\overline{\mathbf{g}})$) and an edge starting at i with the arrowhead pointing at j indicates that i sponsors a link to j. We say that $\overline{\mathbf{g}}$ is a regular network of degree v if each player has v connections in $\overline{\mathbf{g}}$. A complete network is a regular network with $v = n - 1$.

II. Analysis

The main result of this section is that *every* (strict) equilibrium exhibits the law of the few: a small subset of individuals personally acquire information, while the rest of the population of individuals form connections with this small set of information acquirers. This differentiation in turn generates an elegant architecture of social communication: players who personally acquire information constitute hubs, while the rest of the players are spokes in a core-periphery social network.

We start by noting that in an equilibrium every player must access at least $\hat{y}$ information (where $\hat{y}$ is the optimal information acquired by an isolated player). Moreover, the perfect substitutability of own and neighbor's information and the linearity in the costs of acquiring information imply that if a player personally acquires information then the sum of the information he acquires and the information acquired by his neighbors must equal $\hat{y}$. We next observe that if some player acquires $\hat{y}$, and if $k < c\hat{y}$, then it is optimal for all other players to acquire no information personally and to form a link with this player. Lemma 1 summarizes these

[8] For other recent studies of models in which players can charge prices or bargain for their information, see Antonio Cabrales and Piero Gottardi (2007) and Myeonghwan Cho (2007).

observations. For a strategy profile $\mathbf{s} = (\mathbf{x}, \mathbf{g})$, let us define $I(\mathbf{s}) = \{i \in N | x_i > 0\}$ as the set of players who acquire information personally, and let $y_i(\bar{\mathbf{g}}) = \sum_{j \in N_i(\bar{\mathbf{g}})} x_j$ be the information that i accesses from his neighbors.

LEMMA 1: *In any equilibrium* $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$, $x_i^* + y_i^*(\bar{\mathbf{g}}) \geq \hat{y}$, *for all* $i \in N$, *and if* $x_i^* > 0$ *then* $x_i^* + y_i^*(\bar{\mathbf{g}}) = \hat{y}$. *Moreover, if* $k < c\hat{y}$ *and* $x_i^* = \hat{y}$ *then* $x_j^* = 0$, *for all* $j \neq i$.

The first statement in the lemma is similar to a result obtained in Bramoulle and Kranton (2007) in the context of local public good provision in a fixed network; we note that the second statement arises out of the endogenous linking. The proof is given in the Appendix.

This lemma tells us a great deal about the information accessed by individuals but relatively little about the distribution of personal information acquisition, the aggregate information acquired in a social group, and the structure of social communication. Our next result addresses these concerns.

PROPOSITION 1: *Suppose payoffs are given by* (1). *If* $k > c\hat{y}$ *then there exists a unique equilibrium in which every player acquires information* $\hat{y}$ *and no one forms any links. Suppose* $k < c\hat{y}$ *and let* $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$ *be an equilibrium.*

(*i*) *If* $\sum_{i \in N} x_i^* = \hat{y}$ *then* $\mathbf{g}^*$ *is a core-periphery network, hubs acquire information personally, and spokes acquire no information personally.*

(*ii*) *If* $\sum_{i \in N} x_i^* > \hat{y}$ *then:*

a) *Every player* $i \in I(\mathbf{s}^*)$ *chooses* $x_i^* = (\hat{y}/(\Delta + 1)) = k/c$ *and has* $\Delta \in \{1, \ldots, n-2\}$ *links within* $I(\mathbf{s}^*)$, *while every player* $j \notin I(\mathbf{s}^*)$ *forms* $\Delta + 1$ *links with players in* $I(\mathbf{s}^*)$.

b) *High information level players choose* $\overline{x}^* = k/c$, *low information level players have* η *links with high information level players, they are not neighbors of each other and choose information* $\underline{x}^* = \hat{y} - \eta\,(k/c)$, *where* $(\hat{y}c/k) - 1 < \eta < \hat{y}c/k$.

If the cost of a link exceeds the cost of acquiring the threshold information, $k > c\hat{y}$, then no one forms any links. Otherwise every equilibrium is characterized by linking activity, i.e., the network is nonempty. Figure 2 illustrates equilibrium outcomes for $n = 8$, $\hat{y} = 1$ and $k < c$. There are two types of equilibria: one, where aggregate information is equal to $\hat{y}$ (Figure 2A) and two, where it exceeds $\hat{y}$ (Figure 2B). When aggregate information equals $\hat{y}$, equilibrium networks are connected, they have the core-periphery structure, and players in the core are the only ones who acquire any information personally. Moreover, as the relative cost of linking k/c grows the number of hubs decreases, each hub player acquires more information, and the total number of links decreases. When $k/c \in (\hat{y}/2, \hat{y})$ there is only one hub, and the social communication structure takes the form of a periphery sponsored star.

When aggregate information exceeds $\hat{y}$, equilibrium networks may not be connected, but players acquire at most two levels of information (as in Figure 2B). Moreover, since aggregate information acquired exceeds $\hat{y}$ we know, from Lemma 1, that there must exist players who are accessed by some but not by other players. This means that the cost of linking is *exactly* equal to the cost of information acquired by such a player. But then players are indifferent between forming a link with such a player and acquiring information personally. In other words, the strategies of the players are not a *strict* best response to the strategies of others. The following result brings out the general implications of this observation.

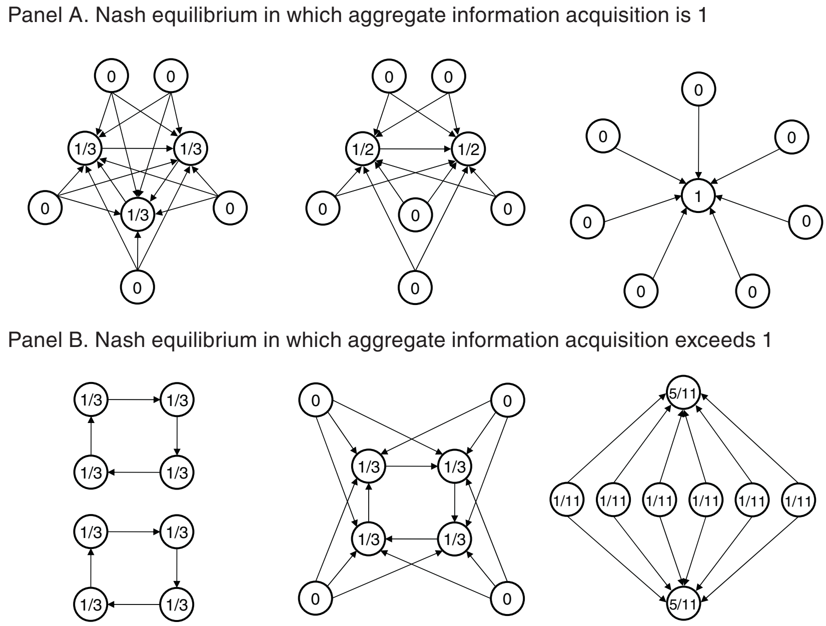

FIGURE 2. EXAMPLE OF NASH EQUILIBRIUM, $n = 8$, $\hat{y} = 1$

PROPOSITION 2: *Suppose payoffs are given by* (1) *and* $k < c\hat{y}$. *In every strict equilibrium* $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$: (*i*) $\sum_{i \in N} x_i^* = \hat{y}$, (*ii*) *the network has a core-periphery architecture, hubs acquire information personally, and spokes acquire no information personally, and* (*iii*) *for given c and k, the ratio* $|I(\mathbf{s}^*)|/n \to 0$ *as* $n \to \infty$.

PROOF:

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

We now consider the proportion $|I(\mathbf{s})|/n$ in every strict equilibrium $\mathbf{s}$. Recall that in a strict equilibrium $\sum_{i \in N} x_i = \hat{y}$ and that $x_i + y_i(\overline{\mathbf{g}}) \geq \hat{y}$, for all $i \in N$. This means that every player who acquires information personally is accessed by every player in equilibrium. This implies that there is at most one player $i \in I(s)$ with no incoming links, i.e., $g_{ji} = 0$, for all $j \in N$. For all other players $l \in I(\mathbf{s})$, it must be the case there is at least one player $j \in I(s)$ such that $g_{jl} = 1$; but this implies that $x_l > k/c$. So the number of accessed players who acquire information personally and have incoming links, $I(\mathbf{s}) - 1$, is bounded above by $(\hat{y}c)/k$. It follows that $I(\mathbf{s})/n \leq ([\hat{y}c/k] + 1)/n$, which can be made arbitrarily small by raising n. This completes the proof.

If a positive information player strictly prefers to retain a link with another player j, it must be the case that the costs of linking with j are strictly lower than the cost of the information that player j acquires. But then all other players find it optimal to access j's information, and so all positive information players are linked among themselves. Lemma 1 tells us that any player who acquires information personally must access exactly $\hat{y}$ information. Hence the desired conclusion: aggregate information acquisition must equal $\hat{y}$. Part (ii) of Proposition 2 then follows from Proposition 1.

The last part of the result derives bounds on the number of hub players. If player i links with j then the cost of the link must be less than the cost of personally acquiring the information accessed from j, i.e., $cx_j > k$, and so player j must be acquiring at least k/c information. Since aggregate information is $\hat{y}$, it follows that there is an upper bound on the number of players who acquire information, $I(\mathbf{s}^*)$, and this number is independent of n. It then follows that the ratio $|I(\mathbf{s}^*)|/n$ can be made arbitrarily small by suitably raising n. Proposition 2 shows that these properties arise in *every* equilibrium of our game.[9] Thus, the law of the few obtains as a consequence of strategic interaction among ex ante identical and rational individuals.

Proposition 2 tells us that the number of players personally acquiring information is small relative to the number of players in large societies. But for fixed n, it does not determine the number of hubs, nor does it tell us who the hub players are. A recurring theme in the empirical literature is that even though hubs seem to have similar demographic characteristics as the others, they have distinctive attitudes that include higher attention to general market information and greater enjoyment in collecting information. A natural way to model this difference is to suppose that some players have slightly lower costs of acquiring 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 $\hat{y}_1 \to \hat{y}$ as $\epsilon \to 0$. We focus on strict Nash equilibria.

PROPOSITION 3: *Suppose payoffs are given by* (1), $c_i = c$ *for all* $i \neq 1$ *and* $c_1 = c - \epsilon$, *where* $\epsilon > 0$. *If* $k < f(\hat{y}_1) - f(\hat{y}) + c\hat{y}$ *then in a strict equilibrium* $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$: (*i*) $\sum_{i \in N} x_i^* = \hat{y}_1$, (*ii*) *the network is a periphery-sponsored star and player* 1 *is the hub, and* (*iii*) *either* $x_1^* = \hat{y}_1$ *and spokes choose* $x_i^* = 0$, *OR* $x_1^* = [(n-1)\hat{y} - \hat{y}_1]/[n-2]$ *and* $x_i^* = [\hat{y}_1 - \hat{y}]/[n-2]$, *for all* $i \neq 1$.

The proof is given in the Appendix. Proposition 3 shows that a very small difference in the cost of acquiring information is sufficient to separate the player who will acquire information and act as a hub from those who will acquire little or no information personally and will only form connections.[10]

We now discuss the ideas underlying this result. First, observe that for the low cost player the optimal information level is greater than the optimal information level for other players, i.e., $\hat{y}_1 > \hat{y}$. From the arguments developed in Proposition 2 we know that aggregate information acquired by all players other than player 1 will be at most $\hat{y}$. This implies that in equilibrium, player 1 must acquire information personally, $x_1 > 0$. If $x_1 = \hat{y}_1$, the best reply of every other player is to acquire no information and to form a link with player 1. In case $x_1 < \hat{y}_1$ we know, from Lemma 1, that $x_1 + y_1(g) = \hat{y}_1$ and so there is a player $i \neq 1$ with $x_i > 0$ and $x_i + y_i(\bar{\mathbf{g}}) = \hat{y}$.

[9] Indeed, part (iii) of Proposition 2 can be strengthened to read: for every strict equilibrium, $\lim_{n\to\infty} I(\mathbf{s}^*)/n^{\alpha} \to 0$, for all $\alpha > 0$. We thank a referee for this observation.

[10] We have focused on slight differences in costs of acquiring information; analogous arguments show that if one player derives greater marginal benefits from acquiring information, as compared to others, then he will constitute the hub of the social network and acquire more information than the others.

If some player wants to link with i then so must everyone else. But then player i accesses all information $\hat{y}_1$; since $\hat{y}_1 > \hat{y}$, this contradicts Lemma 1. Thus *no* player must have a link with player $i \neq 1$ in equilibrium. Hence, i must form a link with player 1, and, from the optimality of linking, so must every other player. Finally, since every player is choosing positive effort, the equilibrium values of x_1 and x_i can be derived from the two equations $x_1 + (n-1)x_i = \hat{y}_1$ and $x_1 + x_i = \hat{y}$.[11]

A. *Efficient Outcomes*

Given their salience it is important to understand the welfare properties of specialization in information acquisition and social communication. Proposition 1 tells us that in a Nash equilibrium for every player $i \in N$, $x_i + y_i(\bar{\mathbf{g}}) = \hat{y}$. Thus, in every equilibrium, aggregate gross returns are $nf(\hat{y})$. If $k < c\hat{y}$, given the linearity in costs of information and linking, the efficient equilibrium minimizes the total costs of information and links. This immediately implies that the efficient equilibrium is a periphery-sponsored star network in which the hub acquires information $\hat{y}$ and every spoke chooses 0.

However, individual information acquisition is a local public good, and this implies that so long as equilibrium entails any links, there will be underprovision of information acquisition relative to the social optimum. To see this, note that in the star the hub player chooses $\hat{y}$, and at this point $f'(\hat{y}) = c$. But marginal social returns are given by $nf'(\hat{y})$, which are larger than c, for $n \geq 2$. The following proposition characterizes efficient outcomes.

PROPOSITION 4: *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* $\tilde{y}$ (*where* $nf'(\tilde{y}) = c$), *while all other players choose* 0. *If* $k > \bar{k}$, *then in the socially optimal outcome every player chooses* $\hat{y}$ *and no one forms links.*

PROOF:

Suppose $\mathbf{s} = (\mathbf{x}, \mathbf{g})$ corresponds to an efficient profile. We first show that if $\mathbf{g}$ is not empty, then $\mathbf{g}$ is a star. Let $\mathbf{g}$ be a nonempty network and suppose that C is a component in $\mathbf{g}$. Let $|C| \geq 3$ be the number of players in C. Suppose that y is the total information acquired 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 chooses y. It then follows that this configuration attains the maximum possible aggregate payoff given effort y. Moreover, note that aggregate payoff in any profile $\mathbf{s}$ in which two or more players acquire positive information is strictly less than this, since it will entail the same total costs of information acquisition and 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 acquiring all the information personally is the optimal profile for each component.

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

[11] In a recent paper on network formation with *exogenous* information levels, Hojman and Szeidl (2008) obtain a result on how small differences between players can help select the identity of the hub players. Their result is derived in a dynamic model and relies on stochastic stability arguments. By contrast, our result shows that with endogenous information acquisition, a slight amount of individual heterogeneity is sufficient for the selection of hubs in a *one-shot static* model.

more components, the components must be of equal size. Let m be the size, and let the effort y satisfy $mf'(y) = c$. Suppose now 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 $\mathbf{g}'$ 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 information level y is no longer optimal in either of the components. So, for instance, information 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 $\mathbf{g}$ is not empty then $\mathbf{g}$ is a star, and the information 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 $\mathbf{s}$ is socially efficient and $\mathbf{g}$ is not a star, then $\mathbf{g}$ must be empty and every player will choose information $\hat{y}$. 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 4.

The key point to note is that given any profile of information acquisition and linking, there is a corresponding star network in which the hub does all the information acquisition, which is strictly better. This is a consequence of the linear costs of information acquisition and the positive costs of linking. So, if the optimal social organization is a nonempty network, then it must be a star where the hub acquires all the information. The value of $\bar{k}$ is obtained by equating the social welfare attained in the empty network and such a star network. The following example illustrates the relation between equilibrium and socially efficient outcomes.

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 efficient equilibrium is a star where the hub acquires 1 and the spokes choose 0. As compared to socially optimal outcomes, in equilibrium there is underinvestment in information. For moderate costs of linking, $k \in (1/2, \bar{k})$, in equilibrium we have underinvestment and underconnectivity relative to socially optimal outcomes (noting that $\bar{k} > 1/2$). In the remaining region, equilibrium outcomes coincide with socially optimal outcomes.

III. 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 himself. In this context, Propositions 2 and 3 show that equilibrium leads to core-periphery social communication structures where players in the core acquire all information. In actual practice, we often receive information from friends and colleagues which they have themselves received from other friends. The aim of this section is to examine the implications of this form of indirect information transmission. We show that information spillovers give rise to a new type of hub player/social influencer: *the connector*. A connector does not acquire information personally but acts as an intermediary between other people who acquire information.

Given two players i and j in $\mathbf{g}$, the geodesic distance, $d(i, j; \bar{\mathbf{g}})$, is defined as the length of the shortest path between i and j in $\bar{\mathbf{g}}$. If no such path exists, the distance is set equal to infinity. Let $N_i^l(\bar{\mathbf{g}}) = \{j \in N : d(i, j; \bar{\mathbf{g}}) = l\}$, that is $N_i^l(\bar{\mathbf{g}})$ is the set of players who are at distance l from i in $\bar{\mathbf{g}}$.

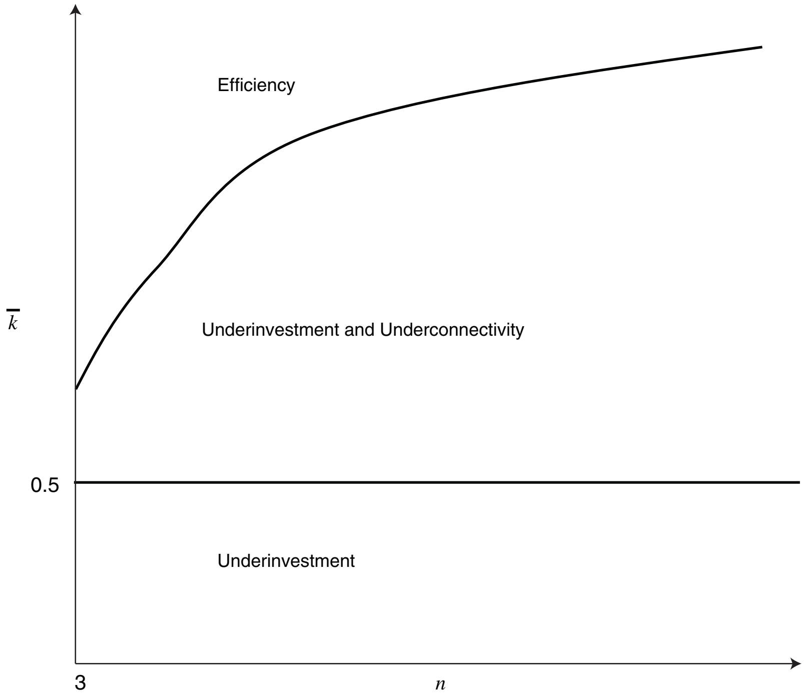

FIGURE 3. EQUILIBRIUM AND EFFICIENT OUTCOMES

We measure the level of spillovers by a vector $\mathbf{a} = \{a_1, a_2, \ldots, a_{n-1}\}$, where $a_1 \geq a_2, \ldots, \geq a_{n-1}$ and $a_l \in [0, 1]$ for all $l \in \{1, \ldots, n-1\}$. It is assumed that if $j \in N_i^l(\overline{\mathbf{g}})$, then the value of agent j's information to i is given by $a_l x_j$. Observe that the case $a_1 = 1$ and $a_2 = 0$ corresponds to the pure local spillovers model analyzed in Section II. To bring out the role of indirect information transmission in the simplest form we start by considering the polar case of no decay in flow across links, i.e., $a_{n-1} = 1$.[12]

The payoffs to player i under strategy profile $\mathbf{s} = (\mathbf{x}, \mathbf{g})$ can be written as:

$$\Pi_i(\mathbf{s}) = f\left(x_i + \sum_{l=1}^{n-1} \sum_{j \in N_i^l(\overline{\mathbf{g}})} x_j\right) - cx_i - \eta_i(\mathbf{g})k. \tag{4}$$

In network $\mathbf{g}$, define $y_{ij}(\overline{g}) = y_i(\overline{\mathbf{g}}) - y_i(\overline{\mathbf{g}} - \overline{g}_{ij})$, i.e., the information which i accesses exclusively via j. Our next result derives the properties of equilibrium with indirect information transmission.

PROPOSITION 5: *Suppose payoffs are given by* (4). *If* $k > c\hat{y}$, *there exists a unique equilibrium: every player personally acquires information* $\hat{y}$, *and no one forms links. If* $k < c\hat{y}$, *then* $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$ *is an equilibrium if and only if:* (*i*) $\sum_{i \in N} x_i^* = \hat{y}$, (*ii*) $\mathbf{g}^*$ *is minimally connected, and* (*iii*) $k \leq c y_{ij}(\overline{\mathbf{g}}^*)$ *for all* $g_{ij}^* = 1$, $i, j \in N$.

[12] See Hojman and Szeidl (2008) for an elegant model of interpersonal communication which leads to declining value of information with respect to distance in a social network.

PROOF:

The proof for $k > c\hat{y}$ of Proposition 5 is straightforward, and it is omitted. Hereafter, we focus on $k < c\hat{y}$. We first prove that if $\mathbf{s}$ satisfies properties (*i*)–(*iii*) in the proposition then $\mathbf{s}$ is a Nash equilibrium. Take a player i; since $\sum_{j \in N} x_j = \hat{y}$, $\overline{\mathbf{g}}$ is minimally connected, and there is no decay, then $x_i + y_i(\overline{\mathbf{g}}) = \hat{y}$, so player i does not want to change his own information level, and also he does not want to form an additional link. The payoffs to i at equilibrium $\mathbf{s}$ are $f(\hat{y}) - cx_i - \eta_i(\mathbf{g})k$. If $\eta_i(\mathbf{g}) = 0$, then player i plays a best reply. Suppose $\eta_i(\mathbf{g}) > 0$, 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 $k \leq cy_{ij}(\overline{\mathbf{g}})$, and therefore player i does not gain by deleting the link with j. Hence, $\mathbf{s}$ is a Nash equilibrium.

We now prove the converse. Let $\mathbf{s} = (\mathbf{x}, \mathbf{g})$ be an equilibrium. Frictionless information flow implies that every component of $\overline{\mathbf{g}}$ must be minimal. Also, frictionless information flow together with Lemma 1 imply that in every component the aggregate information is $\hat{y}$. Next, suppose $\overline{\mathbf{g}}$ is not connected. Let C_1 be a component of $\overline{\mathbf{g}}$. If $x_i = \hat{y}$ for some $i \in C_1$, then all i's neighbors choose information 0 and sponsor 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 level of personal information acquisition; since C_1 is minimal it must be the case that there is a link $g_{i'j'} = 1$ for some $i', j' \in C_1$, 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 forms 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, $\overline{\mathbf{g}}$ is connected. Finally, it is readily seen that if $g_{ij} = 1$ and $\mathbf{s}$ is equilibrium, then $k \leq cy_{ij}(\overline{\mathbf{g}})$. This concludes the proof.

Frictionless information flow implies that equilibrium networks are minimal.[13] From Lemma 1 we know that in equilibrium every individual must access at least $\hat{y}$ information. If the costs of linking k are smaller than the costs of acquiring the threshold level of information $\hat{y}$, then standard considerations 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, i.e., $k \leq cy_{ij}(\overline{\mathbf{g}}^*)$. This implies that either player j acquires enough information on his own, or that player j is a *connector* and accesses others who have enough information.

We explore the distribution of information acquisition and the architecture of social communication via an examination of different classes of equilibria.

Hubs Acquire Information.—Here the hubs personally acquire information. Figure 4A illustrates these equilibria. At a superficial level these equilibria are similar to the core-periphery equilibria of the basic model. However, there is a key difference: in the present context, the information hubs acquire as well as aggregate information. This can be seen in the equilibrium on the right in Figure 4A: each hub personally acquires information but also passes on information from the other hub. This transmission in turn explains the minimality of the equilibrium network.

Hub as Active Connector.—The hub is an active connector in the sense that he acquires no information himself but forms links with all players who are acquiring information. All players who acquire no information in turn link with the hub. Figure 4B provides examples of such equilibria. An equilibrium with $|I(\mathbf{s}^*)|$ peripheral players acquiring information exists whenever the costs of linking are smaller than the benefits of accessing a single information acquirer, i.e.,

[13] This builds on a result of Bala and Goyal (2000), which establishes minimality of equilibrium networks in a model of pure network formation and frictionless information flow.

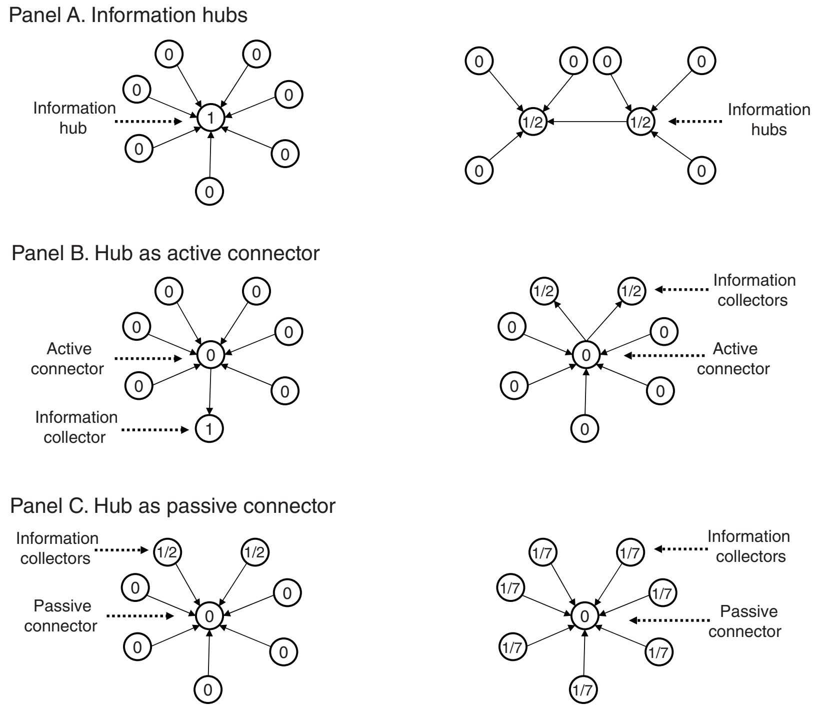

FIGURE 4. HUBS, CONNECTORS, AND OTHERS, $n = 8$

$k < c\hat{y}/|I(\mathbf{s}^*)|$. As k increases the number of information acquirers falls, and each positive information agent must acquire more information. This is in line with one of the results in the basic model: the number of players acquiring information declines with the cost of linking.

Hub as Passive Connector.—All players link with a single player who himself acquires no information but serves as a gateway to those who do. Figure 4C illustrates this type of equilibrium. As costs of linking increase, initially, the number of players acquiring information increases, and each of them acquires less information. That is, an increase in costs of linking necessitates a greater number of active information sources. If every peripheral player acquires information, an increased cost of linking implies that each periphery player acquires less information, which is possible only if the hub acquires some information. In other words, an equilibrium where the hub acts as a passive connector is no longer sustainable for large costs of linking.

We conclude this section with three remarks. *First*, the equilibria illustrated in Figure 4 exhibit patterns which are consistent with the empirical work on social communication. In their study of personal influence, Katz and Lazersfeld (1955) emphasize that 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 influencers acquire information. In other instances, hubs acquire some—possibly a small amount of—information personally, but their numerous contacts provide new information which they then communicate to their neighbors and friends. Here the highly connected individual functions primarily as a connector. See Malcolm Gladwell (2000) for an engaging discussion of such patterns of information acquisition and social communication and Cross and Parker (2004) for a description of social connectors within firms.

Second, we note that the equilibria presented in Figure 4 are the only outcomes in the presence of small heterogeneities across players. In particular, if there is a player with sightly lower costs of acquiring information and slightly lower costs of being accessed by others, then he becomes the information hub, and everyone else forms a link with him. If, instead, the player with the lowest cost of information acquisition and the most sociable player are different, then in the unique (strict) equilibrium the most sociable player is a pure connector: he links with the low cost information player, who acquires all the information, while all other players form a link with him. These claims are formally stated and proved in the Online Appendix.

Our *final* remark is about equilibrium in a general model of decay. We note that lack of decay with respect to information acquired by immediate neighbors is a *necessary* condition for equilibrium where hubs acquire all information and spokes acquire no information. Indeed, in every such equilibrium the aggregate information acquisition is $\hat{y}$, and there is always a player who relies at least partly on information acquired by others. If $a_1 < 1$, then information available for this player is strictly less than $\hat{y}$, which contradicts Lemma 1. Next observe that lack of decay with respect to immediate personal contacts is also *sufficient*: an equilibrium in which the network is a periphery-sponsored star and the hub acquires all the information exists so long as $a_1 = 1$ and $k < c\hat{y}$.[14] We next turn to the nature of influencers under gradual decay. Suppose $a_1 < 1$ and $a_2 > a_3 \geq 0$. A periphery-sponsored star network in which each of the peripheral players chooses $x = \hat{y}/[1 + a_2(n-2)]$ while the central hub player acquires *no* information personally is an equilibrium so long as n is sufficiently large. Hence, in a general model of decay, we expect hubs to play the role of pure connectors.

IV. Discussion

In this section we discuss two aspects of the model which play a prominent role in our analysis: (*i*) the linear costs of acquiring information and forming links, and (*ii*) the link formation protocol.

A. *Convex Costs*

The linear costs of acquiring information and forming links has the following implication: for any player acquiring information the total information accessed is $\hat{y}$, and this level is independent of the amount of information acquired by the neighbors. However, if the costs of personally acquiring information are increasing and convex, this is no longer true. This section explores equilibrium outcomes under convex costs.

Define $z_i = x_i + \eta_i(\mathbf{g})k$, 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 $y_i(\overline{\mathbf{g}}) \geq 0$ is the information accessed by player i from the others. The payoffs to player i facing a strategy profile $\mathbf{s} = (\mathbf{x}, \mathbf{g})$ are given by:

$$\Pi_i(\mathbf{s}) = f(x_i + y_i(\overline{\mathbf{g}})) - C(x_i + \eta_i(\mathbf{g})k). \tag{5}$$

We first note that, in equilibrium, a network is either empty or minimally connected.[15] Moreover, the lack of decay in information transmission implies that in a minimally connected network

[14] We thank an anonymous referee for drawing our attention to this fact.

[15] This follows from standard arguments which rely on network externalities and the lack of decay in information transmission through the network; see, e.g., Bala and Goyal (2000).

every player accesses the information acquired by all players. Let y be the aggregate information acquired in equilibrium. For any player i who acquires information personally, the following first order condition must hold:

$$f'(y) = C'(x_i + \eta_i(\mathbf{g})k). \tag{6}$$

It then follows that players sponsoring an equal number of links must acquire an equal amount of information, and this personal information acquisition is declining in the number of links sponsored. In any minimally connected network there are $n - 1$ links, and so at least one player forms no links. Thus specialization in information acquisition remains an essential aspect of equilibrium behavior even under convex costs.

We now turn to aggregate information acquisition in equilibrium. Recall that in the basic model with linear costs, propositions 2, 3, and 5 prove that aggregate information is invariant with regard to the number of players and the costs of forming links (so long as $k < c\hat{y}$). A greater number of players allows for smaller per capita acquisition of information; under convex costs this leads to lower marginal costs. So as we raise the number of players, aggregate information should increase. Similarly, we expect that link formation costs will affect aggregate information, as these costs now enter the first order conditions of individual optimization (for all positive information players).

We now develop the implications of these ideas formally. We restrict attention to nonempty networks and suppose all players acquire information. Define $\tilde{x}$ as the information acquired by the zero link player. From the optimality of equilibrium actions and equation (6) it follows that for the zero link player $f'(y) = C'(\tilde{x})$ (where y is aggregate information acquired), while for a player i with η_i links we have that $f'(y) = C'(x_i + \eta_i k)$. So we infer that for any player i, $x_i + \eta_i k = \tilde{x}$. Summing across all players we obtain an expression for aggregate information acquisition $y = n\tilde{x} - (n - 1)k$.

PROPOSITION 6: *Suppose that payoffs are given by* (5) *and that* $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$ *is a nonempty network equilibrium in which all players acquire information. Aggregate information is decreasing in the costs of linking and increasing in the number of players.*

The proof of this result is given in the Web Appendix to the paper. Figure 5 illustrates these findings. Figure 5A presents periphery-sponsored star equilibria for $n = 4$ and different values of k. Figure 5B illustrates periphery-sponsored star equilibria for $k = 0.1$ and different values of n.

Figure 5 also helps us explore further the role of differentiation and the architecture of social communication under convex costs. To fix ideas, suppose that every player acquires information and the network is a periphery-sponsored star. From the above arguments we know that in such an equilibrium the hub acquires information $\tilde{x}$, each peripheral player acquires information x_p, where $x_p + k = \tilde{x}$, and $f'(n\tilde{x} - (n - 1)k) = C'(\tilde{x})$. Assume that $\lim_{x \to \infty} f'(x) \to 0$; since $C'(0) = 0$ and is strictly increasing thereafter, it now follows that $x_p \to 0$ and that $\tilde{x} \to k$, as n gets large. The example in Figure 5B satisfies these hypotheses. Thus we have shown that, even with convex costs, sharp role differentiation with a core-periphery communication network emerges in large societies.

B. *The Link Formation Protocol*

We have so far assumed that a player can unilaterally form a link with another player. This is a convenient and simple way to model link formation, and the research on network formation

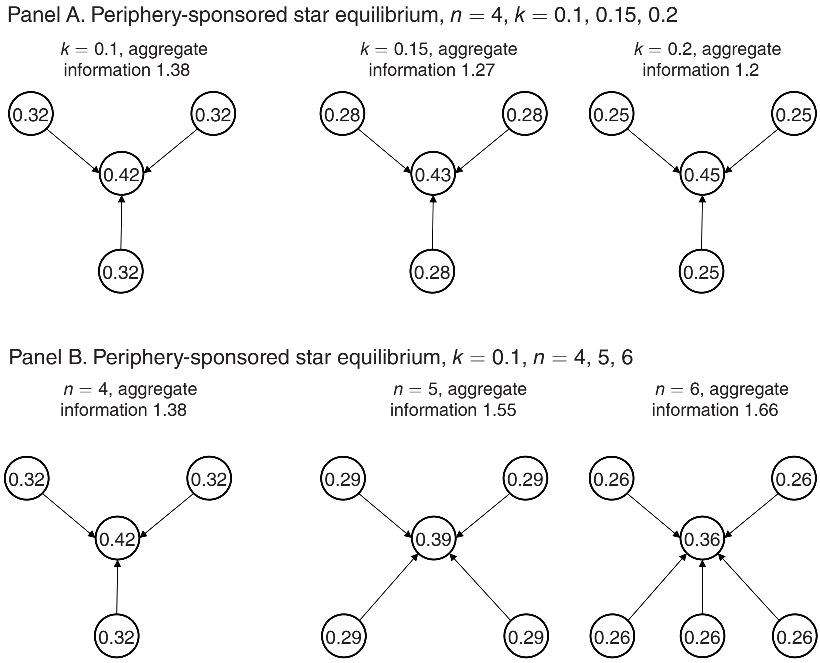

FIGURE 5. AGGREGATE INFORMATION UNDER CONVEX COSTS, $f(y) = \ln(1 + y)$, $C(z) = z^2/2$

over the last decade has shown that it offers a tractable framework to address a range of interesting questions. However, from a descriptive point of view, in social contexts it is more natural to suppose that social communication requires the active participation of the two players involved in a link. This in turn suggests that both players in a link must bear some costs. This section examines role differentiation and core-periphery social communication in a model with two-sided network formation.

Suppose a link is formed only when both players agree and the costs of the link are shared equally between linked players. Formally, $\overline{g}_{ij} = 1$ if and only if $g_{ij}g_{ji} = 1$, and for each link $\overline{g}_{ij} = 1$ both players i and j pay $k/2$. The payoffs to player i facing a strategy profile $\mathbf{s} = (\mathbf{x}, \mathbf{g})$ are given by:

$$\Pi_i(\mathbf{s}) = f\left(x_i + \sum_{j \in Ni(\overline{g})} x_j\right) - \; cx_i - \frac{k}{2}\sum_{j \in N} \overline{g}_{ij}. \tag{7}$$

Following the convention in the network literature, we allow for players to delete links unilaterally (see Jackson and Asher Wolinsky 1996). In addition, we allow for coordinated changes in information acquisition by players contemplating a new link. These considerations are summarized in the following solution concept.

DEFINITION 1. *A strategy profile* $\mathbf{s} = (\mathbf{x}, \mathbf{g})$ *is a pairwise equilibrium if* (*i*) $\mathbf{s}$ *is a Nash Equilibrium, and* (*ii*) *for all* $\overline{g}_{ij} = 0$, *if* $\Pi_i(\overline{\mathbf{g}} + \overline{g}_{ij}, x_i', x_j', \mathbf{x}_{-ij}) > \Pi_i(\mathbf{s})$ *then* $\Pi_j(\overline{\mathbf{g}} + \overline{g}_{ij}, x_i', x_j', \mathbf{x}_{-ij}) < \Pi_j(\mathbf{s})$, $\forall x_i', x_j' \in X$.

We now provide a partial characterization of pairwise equilibrium.

PROPOSITION 7: *Suppose payoffs are given by* (7) *and* $k < c\hat{y}$. *In a pairwise equilibrium* $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$, $x_i^* \geq k/(2c)$ *for all* $i \in N$. *Moreover,* $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$ *with a regular network* $v \in \{1, \ldots, n-2\}$, *and each player choosing effort* $x^* = \hat{y}/[v+1]$ *is pairwise equilibrium if and only if* $k \in [c\hat{y}/(v+1), 2c\hat{y}/(v+1)]$. *The complete network with* $x_i^* = x^* = \hat{y}/n$ *is a pairwise equilibrium if and only if* $k \leq 2c\hat{y}/n$.

The proof of this result is given in the online Appendix of the paper. If link formation requires mutual acceptance *and* if costs of linking are equally shared between two players, then every individual must acquire some information personally. Indeed, given costs of linking k and costs of acquiring information c, there is a lower bound to information acquired by each player, and this is independent of the number of players. Moreover, regular social communication networks in which everyone acquires the same amount of information is an equilibrium. These equilibria are in sharp contrast to our main results, propositions 1–3 and 5, in the model with one-sided link formation. We now examine the sources of this big difference in results in the two models.

Since social communication is costly, a player who acquires information will agree to form a link with someone else who has no (or much less) information only if he is offered some compensation. More generally, an informed person may well ask for some compensation for his efforts at acquiring the information in the first place. This compensation may take the form of social favors or direct transfers (monetary or in kind). These ideas motivate a model of two-sided link formation with transfers.

Suppose each player chooses information acquisition and also proposes a set of transfers to every other player. The transfers may be positive (contributing to the costs of communication) or they may be negative (asking compensation for information acquired). Transfers proposed by player i are denoted by $\boldsymbol{\tau}_i = \{\tau_{ij}\}_{j \in N}$ where $\tau_{ij} \in \mathcal{R}$. We assume that $\overline{g}_{ij} = 1$ if and only if $\tau_{ij} + \tau_{ji} \geq k$. So, a strategy profile $\mathbf{s} = (\mathbf{x}, \boldsymbol{\tau})$ specifies efforts $\mathbf{x}$ and transfers $\boldsymbol{\tau}$. A strategy profile $\mathbf{s} = (\mathbf{x}, \boldsymbol{\tau})$ supports a core-periphery network $\overline{\mathbf{g}}(\boldsymbol{\tau})$ if there are two groups of players $N_1(\overline{\mathbf{g}})$ and $N_2(\overline{\mathbf{g}})$ such that (*i*) $\tau_{ij} + \tau_{ji} < k$ for all $i, j \in N_1(\mathbf{g})$, (*ii*) for all $i, j \in N_2(\overline{\mathbf{g}})$, $\tau_{ij} + \tau_{ji} \geq k$, and (*iii*) for all $i \in N_2(\overline{\mathbf{g}})$ and $j \in N_1(\overline{\mathbf{g}})$, $\tau_{ij} + \tau_{ji} \geq k$.

The payoffs to i given a profile $\mathbf{s} = (\mathbf{x}, \boldsymbol{\tau})$ are

$$\Pi_i(\mathbf{s}) = f\left(x_i + \sum_{j \in Ni(\overline{g}(\boldsymbol{\tau}))} x_j\right) - cx_i - \sum_{j \in N} \overline{g}_{ij}(\boldsymbol{\tau})\tau_{ij}. \tag{8}$$

Building on the work of Francis Bloch and Jackson (2007) we propose the following solution concept for our game with transfers.

DEFINITION 2. *A strategy profile* $\mathbf{s} = (\mathbf{x}, \boldsymbol{\tau})$ *is pairwise equilibrium if and only if* (*i*) $\mathbf{s}$ *is a Nash equilibrium and* (*ii*) *for all* $\tau_{ij} + \tau_{ji} < k$, *if* $\Pi_i(\mathbf{s}_{-ij}, \tau'_{ij}, \tau'_{ji}, x'_i, x'_j) > \Pi_i(\mathbf{s})$ *then* $\Pi_j(\mathbf{s}_{-ij}, \tau'_{ij}, \tau'_{ji}, x'_i, x'_j) < \Pi_j(\mathbf{s})$, $\forall x'_i, x'_j \in X$, *and* $\forall \tau'_{ij}, \tau'_{ji}$.

The following proposition shows that if individuals ask for compensation to communicate the information they acquire, the core-periphery communication network in which the hub acquires information is an equilibrium. We assume $\epsilon > 0$.

PROPOSITION 8: *Consider the model with transfers. Suppose that payoffs are given by* (8) *and assume that* $k < c\hat{y}$. *The profile* $\mathbf{s}^* = (\mathbf{x}^*, \boldsymbol{\tau}^*)$ *such that:* (*i*) $\mathbf{g}^*(\boldsymbol{\tau}^*)$ *is a core-periphery network,* *every hub acquires information* $x^* = \hat{y}/|I(\mathbf{s}^*)|$, *and spokes acquire no information personally,* (*ii*) *for every pair of hubs* (i, j), $\tau^*_{ij} = \tau^*_{ji} = k/2$, *and* (*iii*) *for every hub* i *and spoke player* j, $\tau^*_{ij} = -\epsilon c x^*$ *and* $\tau^*_{ji} = k + \epsilon c x^*_i$ *is a pairwise equilibrium so long as* $k + \epsilon c\hat{y}/I(\mathbf{s}^*) < c\hat{y}/I(\mathbf{s}^*)$.

The proof of this result is straightforward and omitted. While core-periphery structures with information acquiring hubs are sustainable in equilibrium, regular networks with homogenous information acquisition are also equilibrium outcomes. For example, the following profile constitutes a pairwise equilibrium in the game with transfers if $k/c \in [\hat{y}/(v+1), 2\hat{y}/(v+1)]$: *For each* i, $x^*_i = c\hat{y}/[v+1]$ *and there is a set of players* $\{i_1, i_2, \ldots, i_v\}$, *such that* $\tau^*_{ij} = k/2$ *and* $\tau_{ji} = k/2$ *for exactly* $j \in \{i_1, i_2, \ldots, i_v\}$ *players.* These different possibilities naturally raise a question about relative robustness of equilibria. This is an interesting question which we leave for future research.[16]

V. Conclusions

The determination of information people have when making decisions is a central problem in the social sciences. Our paper makes two contributions to the study of this problem. One, we develop a model in which players choose investments in personal information acquisition as well as in forming links with others to access their information. Two, we show that the *law of the few*—the phenomenon where a large majority of individuals get most of the information needed for their decisions from a very small subset of the group—is a robust equilibrium phenomenon in such a model.

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 these dynamics. A second line of investigation concerns the relation between personally acquired information and information acquired by others. We have focused on the case where they are substitutes; it would be interesting to study the case where they are complements.

Appendix

Given a network $\overline{\mathbf{g}}$, we define a component as a set $C(\overline{\mathbf{g}})$ of players such that $\forall i, j \in C(\overline{\mathbf{g}})$, there exists a path between them, and there does not exist a path between any $i \in C(\overline{\mathbf{g}})$ and a player $j \in N \setminus C(\overline{\mathbf{g}})$. A component $C(\overline{\mathbf{g}})$ is non-singleton if $|C(\overline{\mathbf{g}})| > 1$. A player i is isolated if $\overline{g}_{ij} = 0$, $\forall j \in N \setminus \{i\}$.

PROOF OF LEMMA 1:

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

We now prove statement 2 in the lemma. Suppose that $\mathbf{s} = (\mathbf{x}, \mathbf{g})$ 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 the first part of

[16] We note that the incentives to acquire information are not affected by transfers as transfers are not conditional on information acquisition. This also limits the ability of transfers to facilitate efficient information acquisition.

Lemma 1 that $x_i + y_i(\overline{\mathbf{g}}) = \hat{y}$. This also implies that every player in the neighborhood of i must acquire no information personally. Now consider j, with $x_j > 0$. This means that $\overline{g}_{ij} = 0$. It follows from Lemma 1 that $x_j + y_i(\overline{\mathbf{g}}) = \hat{y}$. 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 personal information acquisition 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 $x_j < \hat{y}$. In equilibrium $x_j + y_j(\overline{\mathbf{g}}) = \hat{y}$, and so there is some player $l \neq i$ such that $\overline{g}_{jl} = 1$ 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 $\mathbf{s}$ cannot be an equilibrium. A contradiction which completes the proof.

PROOF OF PROPOSITION 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 1 that $I(\mathbf{s})$ must be a clique. Furthermore, $\overline{g}_{ij} = 0$ for all $i, j \notin I(\mathbf{s})$, and $g_{ij} = 0$ for all $i \in I(\mathbf{s}), j \notin I(\mathbf{s})$. Therefore, each player choosing 0 must sponsor a link with every positive information player. This shows that the network must be a core-periphery network.

Hereafter, let $\mathbf{s} = (\mathbf{x}, \mathbf{g})$ 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 information players choose the same level of information. In the second step we consider situations in which positive information players choose different levels of information. The third step uses the observations derived in the previous two steps to conclude the proof.

Step 1: We prove that if all players who acquire information personally choose the same level, then $\mathbf{s}$ satisfies $ii(a)$ of Proposition 1. Suppose $x_i = x, \forall i \in I(\mathbf{s})$. If $x = \hat{y}$, Lemma 1 implies that $|I(\mathbf{s})| = 1$ and therefore aggregate information is $\hat{y}$, a contradiction. Assume $x \in (0, \hat{y})$; from Lemma 1 it follows that $x_i + y_i(\overline{\mathbf{g}}) = \hat{y}, \forall i \in I(\mathbf{s})$. Since, by assumption, $x_i = x, \forall i \in I(\mathbf{s})$, it follows that every player who accesses information personally also gets an equal amount of information from his neighbors, which immediately implies that every positive information player has the same number of links with positive information players; let Δ be this number. Note that for all $i \in I(\mathbf{s})$, $x_i + y_i(\mathbf{g}) = x + \Delta x = \hat{y}$, which implies that $x = \hat{y}/(\Delta + 1)$. Since aggregate information is strictly higher than $\hat{y}$ it follows that $\Delta + 1 < |I(\mathbf{s})|$. Also, from Lemma 1 we know that $x < \hat{y}$, which implies that $\Delta \geq 1$. Thus, there exist two positive information 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 information players who are not neighbors, implying that $k \geq cx$. Hence, $k = cx$. Finally, if $I(\mathbf{s}) = N$, the result follows. If not, select $j \notin I(\mathbf{s})$. Clearly, in equilibrium no player forms a link with j. So, in equilibrium j must sponsor $\Delta + 1$ links with positive information players. This concludes the proof of $ii(a)$ of the proposition.

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

LEMMA 2: *Suppose that (A1) and (A2) hold in $C(\overline{\mathbf{g}}')$. Then there are two types of players in $C(\overline{\mathbf{g}}')$: high information players choose $\overline{x}$ and low information players choose $\underline{x} < \overline{x}$. Moreover, every low information player forms η links with high information players, there are no links between low information players, $k = c\overline{x}$, $\underline{x} = \hat{y} - \eta\overline{x}$, and $(\hat{y}c/k) - 1 < \eta < \hat{y}c/k$.*

PROOF OF LEMMA 2:

Without loss of generality label players in $C(\overline{\mathbf{g}}')$, so that $\hat{y} > x_1 \geq x_2 \geq \ldots \geq x_m$. (A2) implies that there exists $l \in C(\overline{\mathbf{g}}')$, $l \neq m$, such that $x_j = x_l = \overline{x}$, for all $j \leq l$, and $\overline{x} > x_{l+1}$. 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 1 is also a neighbor of j', 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 information as player l does. Since $x_j + y_j(\overline{\mathbf{g}}) = x_l + y_l(\overline{\mathbf{g}}) = \hat{y}$, this implies that $x_j \geq x_l$, contradicting our hypothesis that $x_j < \overline{x} = x_l$. Thus j does not receive links. But then $x_j + y_j(\overline{\mathbf{g}}) = x_j < \hat{y}$, which contradicts Lemma 1. Hence, claim 1 follows.

CLAIM 2: *There exists some i, $i' \leq l$ such that $\overline{g}_{ii'} = 0$.*

PROOF OF CLAIM 2:

Suppose $\{1, \ldots, l\}$ is a clique. Since aggregate information in component $C(\overline{\mathbf{g}})$ exceeds $\hat{y}$, it must be the case then that for every $i \leq l$, there is one player $j > l$ such that $\overline{g}_{ij} = 0$. 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 $\{1, \ldots, l\}$ is a clique, it follows that every neighbor of j is also i's neighbor, and this contradicts the hypothesis that $x_j < x_l = x$. So $\overline{g}_{jj'} = 0$, $\forall j' > l$. Finally note that this implies that i accesses a superset of the players accessed by j, i.e., $y_i(\overline{\mathbf{g}}) \geq y_j(\overline{\mathbf{g}})$. We know that $x_i + y_i(\overline{\mathbf{g}}) = \hat{y} = x_j + y_j(\overline{\mathbf{g}})$, 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 2. From claim 1, there exists a player $j > l$ who sponsors a link to a player $i \leq l$, so $k \leq c\overline{x}$. Similarly, claim 2 implies that there exists i, $i' \leq l$ such that $\overline{g}_{jj'} = 0$; this implies $k \geq c\overline{x}$. Hence, we have $k = c\overline{x}$. Next, since $k = c\overline{x}$ and $x_j < \overline{x}$ for all $j > l$, it follows that $\overline{g}_{j'j} = 0$ for all $j, j' > l$. Therefore, every player $j > l$ forms only links with players in $\{1, \ldots, l\}$. 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, $x_j + y_j(\overline{\mathbf{g}}) = x_j + \eta_j(\mathbf{g})\overline{x}$, and $x_{j+1} + y_{j+1}(\overline{g}) = x_{j+1} + \eta_{j+1}(\mathbf{g})\overline{x}$. Lemma 1 implies that $x_j + y_j(\overline{\mathbf{g}}) = x_{j+1} + y_{j+1}(\overline{\mathbf{g}}) = \hat{y}$, which holds whenever $x_j - x_{j+1} = (\eta_{j+1}(\mathbf{g}) - \eta_j(\mathbf{g}))\overline{x}$. Since $x_j > x_{j+1}$, then $\eta_{j+1}(\mathbf{g}) - \eta_j(\mathbf{g}) \geq 1$, but then $(\eta_{j+1}(\mathbf{g}) - \eta_j(\mathbf{g}))\overline{x} \geq \overline{x} > x_j - x_{j+1}$, where the last inequality follows because, by assumption, $x_j < \overline{x}$. Thus, all players $j > l$ choose the same information, say $\underline{x}$, and from Lemma 1 it follows that $\underline{x} + \eta_j(\mathbf{g})\overline{x} = \hat{y}$. Thus, every low information player sponsors the same number of links with high information players, say η, and $\underline{x} + \eta\overline{x} = \hat{y}$. This concludes the proof of Lemma 2.

Step 3: We now conclude the proof of Proposition 1. Recall that $\mathbf{g}'$ is the subgraph of $\mathbf{g}$ defined on $I(\mathbf{s})$. We need to consider two cases: (*i*) $\overline{\mathbf{g}}'$ is connected, and (*ii*) $\overline{\mathbf{g}}'$ is not connected.

$\overline{\mathbf{g}}'$ is connected: first observe that (A1) holds by assumption. If all positive information players choose same action then step 1 applies, and the proof follows. If (A2) holds then Lemma

2 applies. We next observe that in this case every player $i \in N$ must choose positive information. To see this note that since $k = c\bar{x}$ every player $j \notin I(\mathbf{s})$ will only sponsor links to high information players. Then, by symmetry, low information players must obtain the same payoffs as players $j \notin I(\mathbf{s})$. It is easy to check that this is possible if and only if $\underline{x} = \bar{x}$, which contradicts (A2).

$\bar{\mathbf{g}}'$ is not connected: Let C_1 and C_2 be two components in $\bar{\mathbf{g}}'$. We observe that $x_i < \hat{y}$, and from Lemma 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 information x. Thus, if in C_1 and C_2 every positive player receives at least one link, every player chooses the same information, and the proof follows from Step 1.

Suppose next that there is some player in C_1 who does not receive a link, *and* information acquisition is not equal across players. If the aggregate information 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 information is $\hat{y}$, this player will choose $\underline{x} = \hat{y} - (|C_1| - 1)x$.

Finally, consider the case where aggregate information in C_1 exceeds $\hat{y}$, and personal information acquisition is not equal. Then Lemma 2 applies and there are two positive information acquisition 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 1.

PROOF OF PROPOSITION 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 2 to show that in a strict equilibrium aggregate information 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 3: *$\forall i, j \in I(\mathbf{s}) \setminus \{1\}$, if $\bar{g}_{ij} = 1$ then i and j share the same neighbors; i.e., for every $l \in I(\mathbf{s}) \setminus \{i, j\}$, $l \in N_i(\bar{\mathbf{g}})$ if and only if $l \in N_j(\bar{\mathbf{g}})$.*

PROOF OF CLAIM 3:

Let $\bar{g}_{ij} = 1$, $i, j \in I(\mathbf{s}) \setminus \{1\}$, and suppose, without loss of generality, that $x_i \leq x_j$. We first prove that for every $l \in I(\mathbf{s}) \setminus \{i, j\}$, if $l \in N_i(\bar{\mathbf{g}})$ then $l \in N_j(\bar{\mathbf{g}})$. Suppose not and there exists a player $l \in I(\mathbf{s})$, with $l \in N_i(\bar{\mathbf{g}})$ and $l \notin N_j(\bar{\mathbf{g}})$. 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 1 that $x_i + y_i(\bar{\mathbf{g}}) = \hat{y}$ and the payoffs to i in equilibrium $\mathbf{s}$ are $f(\hat{y}) - cx_i - \eta_i(\mathbf{g})k$. Suppose that i deletes the link with player l and chooses an information level $\tilde{x}_i = x_i + x_l$, then he obtains payoffs $f(\hat{y}) - cx_i - cx_l - (\eta_i(\mathbf{g}) - 1)k$. Since $\mathbf{s}$ 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 $\mathbf{s}$ is a strict equilibrium, player j must strictly lose if he forms an additional link with l and chooses information level $\tilde{x}_j = x_j - x_l$. That is, $f(\hat{y}) - cx_j - \eta_j(\mathbf{g})k > f(\hat{y}) - c(x_j - x_l) - (\eta_j(\mathbf{g}) + 1)k$, 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 subcases. (2a) Suppose $g_{ij} = 1$; this implies that the costs for i to link with j are strictly lower than the costs of information that i accesses from j, i.e., $k < cx_j$. Since $k < cx_j$, $\overline{g}_{lj} = 0$, and, by assumption, $x_l > x_j$, then l strictly gains if he links with j and chooses information level $\tilde{x}_l = x_l - x_j$. So $\mathbf{s}$ is not a strict equilibrium. (2b) Suppose $g_{ji} = 1$. Since j does not access l but he sponsors a link to i, it follows that $x_i > x_l$. Next note that, 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 $l \in I(\mathbf{s}) \setminus \{i, j\}$, if $l \in N_i(\overline{\mathbf{g}})$ then $l \in N_j(\overline{\mathbf{g}})$.

We now show that if $l \in I(\mathbf{s}) \setminus \{i, j\}$ and $l \in N_j(\overline{\mathbf{g}})$ then $l \in N_i(\overline{\mathbf{g}})$. Suppose not; then player j accesses all positive information players that i accesses plus some other positive information players. But this would contradict that $x_i \leq x_j$, since $y_i \geq y_j$. This concludes the proof of Claim 3.

CLAIM 4: *Suppose* $i, j \in N_1(\overline{\mathbf{g}})$, $i, j \in I(\mathbf{s})$, *and* $\overline{g}_{ij} = 0$, *then* $\overline{g}_{li} = \overline{g}_{lj} = 0$ *for all* $l \neq 1$, $l \in I(\mathbf{s})$.

PROOF OF CLAIM 4:

Suppose, without loss of generality, $x_i \leq x_j$. We first show that $\overline{g}_{li} = 0$ for all $l \neq 1$, $l \in I(\mathbf{s})$. Suppose, on the contrary, that $\overline{g}_{li} = 1$, for some $l \neq 1$, $l \in I(\mathbf{s})$. In view of claim 3, since $\overline{g}_{ij} = 0$, then $\overline{g}_{lj} = 0$; this fact and $x_i \leq x_j$ implies that $g_{il} = 1$, and since $\overline{g}_{ij} = 0$, then from the strictness of equilibrium, it follows that $x_l > x_j$. Since $x_l > x_j$, $g_{il} = 1$, and $x_j + y_j = \hat{y} = x_i + y_i(\overline{\mathbf{g}})$, it must be the case that there exists some $l' \in I(\mathbf{s})$, $l' \in N_j(\overline{\mathbf{g}})$, and $l' \notin N_i(\overline{\mathbf{g}})$. Claim 3 implies that $l' \notin N_l(\overline{\mathbf{g}})$. Since $x_l > x_j$ and $l' \notin N_l(\overline{\mathbf{g}})$, then $g_{jl'} = 1$. But $g_{jl'} = 1$ and $l \notin N_j(\overline{\mathbf{g}})$ implies that $x_{l'} > x_l$; similarly, $g_{il} = 1$ and $l' \notin N_i(\overline{\mathbf{g}})$ 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 $l \in I(\mathbf{s})$ and $l \in N_j(\overline{\mathbf{g}})$, then claim 3 implies that $l \notin N_i(\overline{\mathbf{g}})$, but then player j accesses strictly higher information than player i, which contradicts our initial hypothesis that $x_j \geq x_i$. Claim 4 follows.

Final Step in Proof of Proposition 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 $i \in I(\mathbf{s})$ such that $i \in N_1(\overline{\mathbf{g}})$. Observe that given such an i, there exists a $j \in N_1(\overline{\mathbf{g}})$ such that $j \notin N_i(\overline{\mathbf{g}})$ and $x_j > 0$. This is because otherwise $x_i + y_i(\overline{\mathbf{g}}) \geq \hat{y}_1 > \hat{y}$, and this contradicts $x_i > 0$ and Lemma 1. From claim 4 above we know that players i and j do not have any links with players in $I(\mathbf{s})$. 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 information 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 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 information, 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 information player followed by a zero information player. Suppose player $l \notin I(\mathbf{s})$, and suppose there is a link with player m such that $\overline{g}_{m1} = 1$. Since $l \notin I(\mathbf{s})$, it must be the case that $g_{lm} = 1$ and so $m \in I(\mathbf{s})$. However, it is profitable for l to form this link only if $k < cx_m$. Moreover, claim 3 implies that $\overline{g}_{mi} = 0$; if $x_i \geq x_m$, then, since $k < cx_m$, player i strictly gains by forming a link with m. So $x_i < x_m$, and since $\overline{g}_{mi} = 0$ it follows that $k > cx_i$. Putting together these facts 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 1. Thus we have ruled out case 2. The case of two positive levels of information acquisition 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 $\overline{g}_{ll'} = 1$ 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 information acquisition and form a link with 1, while if $x_{l'} < x_1$ then it is strictly profitable for 1 to lower information acquisition 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 information level x and form a link with player 1.

The above argument is done for a single component. The connectedness of nonempty strict equilibrium networks follows from 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 (iii) 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 now easy to see that as $\epsilon \to 0$, then $x \to 0$.

REFERENCES

Adar, Eytan, and Bernardo A. Huberman. 2000. "Free Riding on Gnutella." *FirstMonday*, 5(10).

Bala, Venkatesh, and Sanjeev Goyal. 2000. "A Noncooperative Model of Network Formation." *Econometrica*, 68(5): 1181–229.

Beck, Paul Allen, Russel J. Dalton, Steven Green, and Robert Huckfeldt. 2002. "The Social Calculus of Voting: Interpersonal, Media, and Organizational Influences on Presidential Choices." *American Political Science Review*, 96(1): 57–73.

Bergstrom, Theodore, Lawrence Blume, and Hal Varian. 1986. "On the Private Provision of Public Goods." *Journal of Public Economics*, 29(1): 25–49.

Bloch, Francis, and Matthew O. Jackson. 2007. "The Formation of Networks with Transfers among Players." *Journal of Economic Theory*, 133(1): 83–110.

Bramoullé, Yann, and Rachel Kranton. 2007. "Public Goods in Networks." *Journal of Economic Theory*, 135(1): 478–94.

Bramoullé, Yann, Dunia Lopez-Pintado, Sanjeev Goyal, and Fernando Vega-Redondo. 2004. "Network Formation and Anti-Coordination Games." *International Journal of Game Theory*, 33(1): 1–19.

Cabrales, Antonio, and Piero Gottardi. 2007. "Markets for Information: Of Inefficient Farewalls and Efficient Monopolies." http://163.117.2.172/acabrales/research/IM.pdf.

Calvó-Armengol, Antoni, and Yves Zenou. 2004. "Social Networks and Crime Decisions: The Role of Social Structure in Facilitating Delinquent Behavior." *International Economic Review*, 45(3): 939–58.

Cho, Myeonghwan. Forthcoming. "Endogenous Formation of Networks for Local Public Goods." *International Journal of Game Theory*.

Cross, Robert, and Andrew Parker. 2004. *The Hidden Power of Social Networks: Understanding How Work Really Gets Done In Organizations*. Boston: Harvard Business School Press.

Feick, Lawrence F., and Linda L. Price. 1987. "The Market Maven: A Diffuser of Marketplace Information." *Journal of Marketing*, 51(1): 83–97.

Galeotti, Andrea. 2006. "One-Way Flow Networks: The Role of Heterogeneity." *Economic Theory*, 29(1): 163–79.

Galeotti, Andrea, Sanjeev Goyal, and Jurjen Kamphorst. 2006. "Network Formation with Heterogeneous Players." *Games and Economic Behavior*, 54(2): 353–72.

Geissler, Gary L., and Steve W. Edison. 2005. "Market Mavens' Attitudes Towards General Technology: Implications for Marketing Communications." *Journal of Marketing Communications*, 11(2): 73–94.

Gladwell, Malcolm. 2000. *The Tipping Point: How Little Things Can Make a Big Difference*. Boston: Little Brown.

Goyal, Sanjeev. 1993. "Sustainable Communication Networks." Tinbergen Institute Discussion Paper TI 93–250.

Goyal, Sanjeev. 2007. *Connections: An Introduction to the Economics of Networks*. Princeton, NJ: Princeton University Press.

Goyal, Sanjeev, and Fernando Vega-Redondo. 2005. "Network Formation and Social Coordination." *Games and Economic Behavior*, 50(2): 178–207.

Hojman, Daniel A., and Adam Szeidl. 2008. "Core and Periphery in Networks." *Journal of Economic Theory*, 139(1): 295–309.

Huckfeldt, Robert, Paul E. Johnson, and John Sprague. 2004. *Political Disagreement: The Survival of Diverse Opinions within Communication Networks*. Cambridge, UK: Cambridge University Press.

Jackson, Matthew O. 2008. *Social and Economic Networks*. Princeton, NJ: Princeton University Press.

Jackson, Matthew O., and Alison Watts. 2002. "On the Formation of Interaction Networks in Social Coordination Games." *Games and Economic Behavior*, 41(2): 265–91.

Jackson, Matthew O., and Asher Wolinsky. 1996. "A Strategic Model of Social and Economic Networks." *Journal of Economic Theory*, 71(1): 44–74.

Katz, Elihu, and Paul F. Lazarsfeld. 1955. *Personal Influence: The Part Played by People in the Flow of Mass Communications*. New York: The Free Press.

Lazarsfeld, Paul F., Bernard Berelson, and Hazel Gaudet. 1948. *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*. New York: Columbia University Press.

Zhang, Jun, Mark S. Ackerman, and Lada Adamic. 2007. "Expertise Networks in Online Communities: Structure and Algorithms" Paper presented at WWW2007, Banff, Canada.

This article has been cited by:

1. LUCA DALL'ASTA, PAOLO PIN, ABOLFAZL RAMEZANPOUR. 2011. Optimal Equilibria of the Best Shot Game. *Journal of Public Economic Theory* **13**:6, 885-901. [CrossRef]