

Acquiring Information Through Peers*

Bernard Herskovic

João Ramos

UCLA

NYU

January 2016

Abstract

We study information acquisition from peers when agents' actions balance adaptation and coordination motives. Agents acquire information personally and may obtain additional information by connecting to other agents. Although equally informative regarding adaptation, the source's relative position in the information structure is relevant to form expectations about actions of other players. In our setting, information sources are not perfectly substitutable, and the information of an "opinion maker"—an agent whose information is more public—is more informative of how others act. We show that, when players choose their connections, (i) it is always preferable to connect to opinion makers, and (ii) opinion makers have less incentives to form links. These two results characterize the endogenous shape of the network: Any strict equilibrium of the network formation game generates a hierarchical information structure. Furthermore, if the marginal cost of acquiring information is increasing, the information structure is "core-periphery". We take advantage of the simplicity of the equilibrium information structure to provide two applications. First, we use data on earnings-per-share forecasts to provide an example of how much of the aggregate volatility of forecast can the information structure account for. Second, we focus on the origins of leadership: how individual characteristics influence the role of the agent in the information structure.

Keywords: Network Formation, Information Acquisition, Coordination.

JEL Codes: D83, D85.

*This project first circulated as "Network Formation and Information Acquisition." We are extremely grateful to David Pearce, for his invaluable support and input to this project. We are also indebted to Alberto Bisin, Debraj Ray, and Laura Veldkamp for their numerous comments and suggestions. We would also like to thank Jess Benhabib, Jarda Borovicka, Katka Borovickova, Brianna Chang, Sylvain Chassang, Wouter Dessein, Jack Fanning, Raquel Fernandez, Douglas Gale, Boyan Jovanovic, Qingmin Liu, Mihai Manea, Laurent Mathevet, Tymofiy Mylovanov, Tano Santos, Venky Venkateswaran, Chris Woolnough, Sevgi Yuksel and participants of seminars at NYU, LAMES, Midwest Economic Theory, NSF-Information Transmission in Networks. Updates: <https://sites.google.com/a/nyu.edu/joaoramos/home>. Corresponding author: ramos@nyu.edu

1 Introduction

Social networks are embedded in our social and economic lives. Individual decisions concerning career choices; which policies to support or for whom to vote; and whether to buy a new product or to adopt a new technology are examples of decisions influenced and informed by peers.¹ A pervasive force in the situations above is the need for individuals to adapt to the environment as well as to coordinate their actions. As an agent acquires information from her peers, she also acquires information regarding what her peers know, and thus regarding how they act. Information obtained through peers does double-duty, being instrumental in adaptation to the environment and in coordination with other agents. When agents acquire information from their peers, they form an endogenous network that implies correlations between what agents know, and what they know that others know, and so on. In this paper we consider information acquisition through peers and its effects on individual incentives to acquire information. In an environment in which agents care about adaptation as well as coordination, we show that any strict equilibrium information structure is a hierarchical network.

There are several real world examples that highlight the importance of coordination and adaptation motives. For instance, a political party activist would like to support the best policy, while balancing the need for the Party to display unity. A financial analyst would like to forecast earnings-per-share close to the true value and, due to career concerns, close to market consensus.² A consumer would like to adopt the best looking clothes, while balancing the advantages of being in trend. By following the fashion blog or Instagram account of a friend, a consumer is informed not only of fashion trends, but also of what her friend (and other followers of her friend's account) know about it and how they may choose to dress. Furthermore, as she and others acquire information from her friend they also change that friend's incentives to further acquire information, as the information he already knew becomes more informative of what is on trend.

In our model, agents balance the need to adapt to an unknown state of the world with the need for coordination among themselves. Before choosing an action, an agent receives a signal of the state of the world; and can also, at a cost, tap into the signals received by other agents. Even though agents are ex-ante identical, which agent is the source of information is relevant to how useful the information is. While equally informative about the state of the world, signals from different agents inform differently regarding what others know. A more public signal—a signal from a source observed by more players—is more influential to the actions of other players, thus being

¹De Weerd and Dercon (2006) on insurance; Foster and Rosenzweig (1995), Conley and Udry (2010), Bandiera and Rasul (2006), and Munshi (2004) on technology, product and crop adoption; Calvó-Armengol (2004) on job-market and drop-out; and Calvó-Armengol and Zenou (2004) and Glaeser et al. (1996) on criminality.

²The complementarity here is due to career concerns, as documented in Hong and Kubik (2003). In Section 4 we focus on sell-side analysts and the impacts of different information structures on the volatility of the average forecasts.

more useful for an agent to coordinate with others. The influence of a player’s signal is determined in equilibrium and depends who do other agents acquire information from. In equilibrium, the need to coordinate actions translates into coordination of sources, as agents focus on the same players’ signals. While some players, namely opinion-makers, are influential with highly public signals, others have no influence as no one taps into their signal.

The equilibrium information structure captures coordination on information acquisition decisions and some agents endogenously emerge as opinion makers. Our main result is that any equilibrium information structure is a *directed hierarchical network*. Furthermore, if the marginal cost associated with tapping into other players’ signals is also increasing with respect to the number of signals tapped, the information structure is *core-periphery*. We elaborate these concepts in the next two paragraphs.

A hierarchical directed network is characterized by a ladder of informational importance. Players of each tier tap into the signal of all players in higher tiers. A top tier individual is thus very influential, as her signal is tapped into by members of all other tiers. A second tier individual’s signal is observed by all members of tiers below her, and so on. Three standard features of social hierarchies are replicated by the equilibrium networks of our model.³ First, even though agents are identical, in equilibrium agents endogenously separate themselves into tiers—players in the same tier exert the same influence over the society and symmetrically acquire the same information—with players in different tiers having very distinct influences over the society. Second, the information hierarchy is self-enforcing: as more players tap into a top-tier agent’s signal, the signal’s influence increases, and so does the incentive to tap into it. Third, a player at the top of the informational hierarchy has a higher equilibrium payoff than a player at the bottom. We interpret an individual’s tier as her informational status, a player with higher status reaping higher rewards.

While a hierarchical directed network allows the existence of multiple social levels, a core-periphery network partitions the set of players in only two: the core and the periphery. All members of the periphery are connected to all members of the core. All members of the core are connected to each other; however, members of the periphery are not connected to each other. Members of the core are influential while members of the periphery are not. Furthermore, the signals of core players endogenously act as public information and thus are common knowledge to all players. Finally, core-periphery networks have been widely documented empirically, from information spreading inside a firm to fashion and voting decisions.

Our model is a two-stage game, in which there is an unknown state of the world and each agent receives (for free) an equally informative independent signal.⁴ Each agent’s signal is a source of

³See Magee and Galinsky (2008) for a definition and many examples of social hierarchy.

⁴Section 5 explores a generalization of the model in which agents may differ in their signal accuracies, cost of information acquisition, and preferences.

information that can be tapped via costly social connections by other agents. In the first stage players choose which other players' signals to tap into, forming links and defining an information network. We make two assumptions on how information propagates along the network. First, information flow is one-way: player i tapping into agent j 's signal acquires j 's information, but j does not acquire i 's information, unless she forms another linkage herself. Second, the information acquired by a player through the network is limited to her immediate connections. A player does not acquire information from other agents' connections.

In the second stage each player uses the information obtained through the network to choose an action. A player wants to be as close as possible to both the unknown state of the world and to the average action of the economy. The introduction of this beauty contest element to the payoff generates complementarity and has two immediate effects on the incentives in the network formation game. First, information is not perfectly substitutable. All players' signals have the same precision, but some agents' signals are more informative about the average action, depending on their positions in the network. Second, there is strategic complementarity between the agent's connection decisions.⁵ When a player taps into an opinion maker's signal, she reinforces the opinion maker's role in the network. Her action is influenced by the opinion maker's information which increases its impact on the average action.

To characterize the network structure common to all strict Nash equilibria, we show two key monotonicity properties that drive the information acquisition process. The first states that a player would rather tap into an opinion maker's signal, since it is observed by more players. An individual's endogenous status as an opinion maker makes her signal more public, and thus more informative about the average action of the economy. The second monotonicity property states that an opinion maker has less incentive to tap into signals. The intuition is that an opinion maker's own signal is more informative about the average action, disincentivizing the acquisition of additional information.

The central result of this paper, that each strict Nash equilibrium of this game produces a directed hierarchical network, is a consequence of the two monotonicity properties above. We show that any strict equilibrium satisfy both properties in our network formation game. However, this equilibrium structure is more general than our information acquisition framework. We define a more general class of games, the directed link formation game, in which agents simultaneously choose to form directed links (to tap into another player's signal) and an action strategy. We show that any strict Nash equilibrium of a directed link formation game that satisfies both properties is a Hierarchical Directed Network. Furthermore, we define a third property, which states that opinion makers tap into each others' signal. We show that, in our model, if the marginal cost of acquiring information is increasing, then any strict equilibrium satisfies this third property. Finally, we show

⁵See Bulow, Geanakoplos, and Klemperer (1985) for a definition.

that any strict Nash equilibrium of a directed link formation game that satisfies the three properties is a core-periphery network.

We focus on two implications of a hierarchical information structure for coordination on information acquisition. First, the average action depends heavily on the information of the most influential players. Action coordination is the endogenous result of agents' optimization, but creates volatility in the average action. In a core-periphery network, players' actions become endogenously correlated because all players focus on the signals of those in the core. Information propagates differently and a signal from a core player is more influential on the average action. To verify the quantitative implications of hierarchical information structures for average action volatility, we provide a numerical example using the Institutional Brokers' Estimate System (IBES) data to calibrate the information structure of the model. Depending on how much an analyst weights conformity relative to adaptation, a core-periphery information structure increases 11% to 14% the earnings-per-share forecast volatility. The high impact of the information structure on economic performance highlighted in this example is in line with experimental research in the field.⁶

Second, agents' coordination of information acquisition implies that some agents' information influence all other agents, literally coordinating their actions. This implies a position of leadership for opinion makers. Given that agents are ex-ante identical, our baseline model has no predictive power of who assumes a leadership role. To address this issue, we extend the model to account for different sources of heterogeneity among the agents: (i) communication ability, (ii) listening ability, (iii) quality of private information, and (iv) resoluteness. In line with the literature⁷ we find that a player who is either better at communicating her signal or a player who is more resolute will take leadership positions.

The rest of the paper is organized as follows: First, we provide examples of information acquisition through peers. Next, we discuss our contribution to the literature. In Section 2, we present and solve the model. Section 3 has our main result: we characterize hierarchical and core-periphery networks and show that any equilibrium information structure will have these architectures. Finally, we discuss implications of such architectures and provide some comparative statics. Section 4 discusses how core-periphery networks generate volatility of the average action, while Section 5 extends the baseline model to account for agent heterogeneity. The last section concludes and proposes extensions that can be handled by the model, as well as directions for future research.

⁶For instance, Centola (2010), Centola and Baronchelli (2015), and Suri and Watts (2011). In particular, Centola (2010) shows that the adoption rate of a social behavior doubles if agents communicate in a more clustered network than if all communication happens through a random network. A more clustered network captures the intuition that the friends of my friends are also my friends. For the adoption of a social good it implies more reinforcing signals, as one friend adopting it influences my other friends to do the same and so on. A similar intuition is behind the self-enforcing nature of our information hierarchy.

⁷For instance, see Bolton et al. (2013), Hermalin (2012) and Kaplan et al. (2012).

Acquiring Information Through Peers First, consider *movie critics* writing a Best-of-the-Year list. First, when a movie is released, each critic watches the movie and writes an independent review of it. Each review is an unbiased signal of the true quality of the film. In December, when asked to write about the top movies of the year, critics would like to coordinate and identify the same movies. Besides her own impression of the movies, each critic has now a plethora of reviews available—one by each other critic—to base her report on. Each critic can tap into the signal of any other critic, at a certain cost; and information flows in only one way: the player whose review is being read receives no information regarding the review of the reader. Finally, note how information cannot be rebroadcasted, and thus the content of one review will be known only by those who pay the cost to read it. The equilibrium information structure specifies which reviews each movie critic reads.

Second, consider *sell-side analysts* forecasting a firm's earnings-per-share. Hong and Kubik (2003) documents that career concerns by the analysts imply a tendency to conform to the consensus in their evaluations. Furthermore, Brown et al. (2015) documents that analysts rely, for instance, on (i) private calls to a company's C.F.O.s, (ii) company or plant visits, as sources of information for their forecasts, (iii) road shows and investor day events, and (iv) Q & A following earnings conference calls. Each financial analyst could in principle acquire each one of these signals, calling the C.F.O., taking part in the Q & A portion of the earnings conference call, visiting the company, and going to investor day events; but each signal would come at a cost. There is no reason for these signals to be equally costly for all analysts. For instance, if analyst i_1 went to college with the C.F.O. in question, she will have a much lower cost in developing a relationship that includes frequent private calls. Or if analyst i_2 is based in the same city as the company, she will have a negligible cost of visiting the company. Our model would consider four analysts,⁸ each observing one signal for free: i_1 observes signal (i), while i_2 observes signal (ii), and so on. Once more, information flows in only one direction. If i_1 decides to pay the associated cost and visit the company, this informs investor i_2 nothing regarding what i_1 discussed with the C.F.O. in a private call. Note also that there is no rebroadcasting of information, as Brown et al. (2015) documents there is almost no evidence of direct communication between analysts. An equilibrium information structure specifies which signal is acquired by which analyst.

Third, consider *political partisans* deciding how much to support a specific policy. The importance of coordinating their actions is consequence of partisans' interest in displaying Party unity. Partisans acquire information regarding different policy qualities by watching the T.V. news. However, even though they all watch *CNN Newsroom*, each partisan finds a different time show more convenient. For instance, there are those who prefer to watch the morning show, as they have breakfast; others prefer watching the news while having lunch; and some dine in front of the night

⁸Alternatively, there could be four types of analysts, each type with $\frac{1}{4}$ analysts.

news. Partisans enjoy their routines and value their time, but any signal is available to any partisan at a cost. A partisan of the first group could change her routine and also watch the night news, but that would imply—for instance—spending less time with her family at night. And a partisan of the third group could change her routine and also watch the morning news, but that would imply waking up earlier than usual. An equilibrium information structure would specify which shows each partisan decides to watch.

Observe that the sense in which there is information acquisition from peers in the first example is very different from the other two. Movie critics read the reviews written by one another, thus literally acquiring the information known and produced by another critic. However, the sense in which analysts acquire information from their peers is much more subtle. Analysts (and political partisans) acquire different signals because other players are acquiring them as well: the value of a signal depends on how informative it is regarding the action of other players. In both interpretations, an agent wants to tap into a signal to obtain information regarding what her peers know, independently of who produced the information. Furthermore, as a review of a movie critic is more public—influencing other critics—the author is more informed about how will others act. The reasoning is similar for analysts: If all analysts chose to call the C.F.O. of the company, the analyst that went to school with the C.F.O. would be more informed about how others will act, even though she did not acquire any new signal. Thus, as an analyst acquires a signal that another analyst has, she is effectively affecting the incentives of such a peer to further acquire information.

Related Literature This paper contributes to the literature on information acquisition, first studied in Morris and Shin (2002) focusing on the family of Gaussian-quadratic economies. Beauty-contest models have been applied to a variety of settings, from political leadership to investment games.⁹ For instance, Angeletos and Pavan (2004, 2007) focused on the efficient use of information. This literature compares the equilibrium and the efficient influence of public and private signals, concluding that complementarities in preference lead to excessive influence of the public signal in equilibrium. This feature is part of our results, but the scope of our paper is quite different. Angeletos and Pavan (2007) focuses on the impact of public information dissemination by a central organization in equilibrium outcomes, while we focus on how the excessive influence of public information endogenously affects the publicity of other signals.

The role of individual information acquisition with strategic complementarity has been explored recently in Myatt and Wallace (2012, 2015), Colombo et al. (2014), and Pavan (2014). Closer to our paper, Hellwig and Veldkamp (2009) shows that complementarity in actions generates complementarity in information acquisition; in all of the multiple equilibria, agents coordinate

⁹See for instance Angeletos and Pavan (2004) regarding investment games; Dewan and Myatt (2008) and Dewan and Myatt (2012) regarding political leadership; Allen et al. (2006) regarding financial markets; and Hellwig and Myatt and Wallace (2014) regarding monopolistic and Cournot competition.

on which signals to acquire. The coordination on the acquisition of information exacerbates the excessive influence of public information. Myatt and Wallace (2012) shows the uniqueness of an information acquisition equilibrium if agents can choose how carefully to observe a signal, instead of choosing whether to observe it or not, linking information acquisition to rational inattention. Pavan (2014) extends such formalization to a setting of bounded recall: agents cannot recall all the signals that have influenced their posterior, even though they know the posterior itself. Our paper contributes to this literature by including learning from peers, since in our model the signals come from other agents, and not from outside of the economy. We introduce a new mechanism present in information acquisition decisions. As the signal of a player becomes more public, that player has less incentive to obtain more information, since her own signal is more informative regarding the average action.

In a different literature strand, our paper contributes to the theory of network formation and information acquisition, in particular to the theory of one-sided link formation initiated by Bala and Goyal (2000). Closer to our work, Galeotti and Goyal (2010) provide a model of information acquisition through social networks that aims to explain the “law of the few”, an empirical observation that most individuals acquire their information from a small subset of people. Their results are driven by two characteristics of their model: (i) information sources are perfectly substitutable, and (ii) it is possible to form costly links with players that invest to acquire information. The authors characterize core-periphery networks as an equilibrium outcome. Our model differs from theirs in two key dimensions. First, we consider a unilateral flow of information. Second, in our model the complementarity in payoff space makes information sources non-substitutable. While the Galeotti and Goyal result is a direct consequence of the concavity in payoffs with respect to information and linearity of costs of acquiring it, our result comes from such complementarity.

Our paper is related to a set of studies combining network formation and strategic interaction, for instance Calvó-Armengol and Zenou (2004), Goyal and Vega-Redondo (2005), Hiller (2013), and Baetz (2015). Close to our results, Königa et al. (2014) constructs a dynamic network formation model, aiming to explain nestedness: the neighborhood of a node is contained in the neighborhood of a node with higher degree. The concept of nestedness in network connections can be interpreted as a hierarchy of nodes—a more connected node has more importance since it overlaps all connections of a less connected node. In their paper, the hierarchy is a result of both strategic link formation and exogenous link destruction: links connecting central players have a lower probability of being broken. In contrast, our result is a consequence only of the strategic information acquisition of players in the intent of knowing what others know. Focusing on the emergence of social hierarchies, Baetz (2015) constructs a network formation model that has hierarchical equilibrium. The paper assumes that a player’s value function is concave on the level of activity of her peers and that her best response is increasing in her peers’ activity. The hierarchy

is a direct result of these two assumptions. By contrast, our model provides a micro structure for the hierarchy result, showing how it relies on the complementarity as well as on how information spreads through the network.

In contrast to the literature above, this paper uses a network formation model to discuss equilibrium information structures, and its impact on economic variables. How different information structures affect the incentives in a game has been studied in Bergemann and Morris (2013a), Bergemann and Morris (2013b), Azrieli and Lehrer (2008), and Lehrer et al. (2010). Closer to our work, Dessein et al. (2015) studies the optimal information structure in a similar game. The paper studies the team problem of allocation of attention to many signals when the team needs to coordinate in many different tasks as well as to locally adapt to each task. Each member of the team has pertinent information regarding a particular task. The optimal allocation of attention resembles our core-periphery result: all attention is allocated to a select number of leaders. In a modeling environment full of externalities, it is surprising that equilibrium and optimal organization analysis results in similar intuition: agents coordinate and focus on the information from a subset of the agents in the economy. Our results highlight that organizational focus occurs both in equilibrium and in the optimal allocation of attention, even with ex-ante identical agents.

Finally, our paper contributes to the literature on granular networks. In our model, idiosyncratic shocks (i.e. the individual signals) don't aggregate in equilibrium, and the signal of certain agents matter for the average action. Specifically, some agents will have more influence on the equilibrium outcome than others. This is related to recent works on the origins of aggregate fluctuations (Gabaix (2011); Acemoglu et al. (2012)). Our paper presents this result in a game-theoretical framework. Typically, in this literature, the sparsity of the network is capable of generating aggregate fluctuations from idiosyncratic shocks. That is true in our paper as well, but in our paper the network sparsity is endogenous, as agents choose to form their connections. Our paper extends Jovanovic (1987) paper on complementarities and aggregate fluctuations to an endogenous network formation model.

2 Model

We consider an economy populated by a finite set of players, $N = \{1, 2, \dots, n\}$. Each agent i simultaneously chooses an action $a_i \in \mathbb{R}$ in an attempt to balance two forces: the need for adaptation and the need for coordination. We model these forces by considering an agent who wants to choose an action that, at the same time, is close to the true state of the world and close to the average action in the economy. However, neither the average action nor the true state is known by any agent in the economy. Before choosing actions, agents acquire information about these two variables.

Information Structure There exists a common unknown state of the world, θ , concerning which all agents share a common prior,

$$\theta \sim \mathcal{N}(0, 1).$$

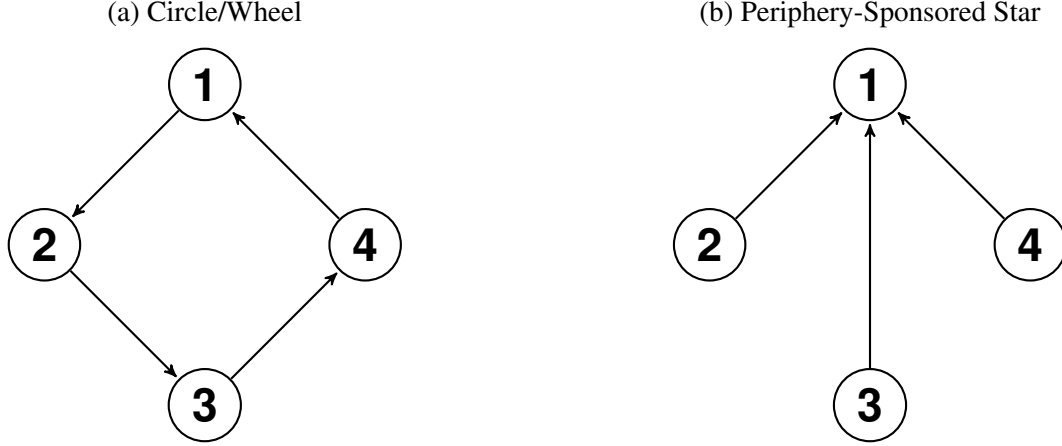
Each agent i observes at no cost a private signal, e_i , about θ , with distribution:

$$e_i = \theta + \varepsilon_i, \text{ where } \varepsilon_i \stackrel{iid}{\sim} \mathcal{N}(0, \sigma).$$

Besides observing her own private signal, e_i , and the common prior, $e_0 = 0$, each agent i can tap, at a cost, into any other agent's private signal of θ . For instance, a movie critic can read the review of another movie critic, or an analyst that is based in the same city as the company, can call the company's C.F.O.

The agents' interconnections implied by the decisions of tapping into each others signals constitute a directed network, called the economy's *information structure*. An information structure $G = \{g_{i,j}\}_{i,j}$ is a list of ordered pairs of agents, such that if agent i taps into the signal of agent j , then the pair is in the list: $g_{i,j} = 1$. Otherwise, $g_{i,j} = 0$. The links carry no intensity attachments and are non reciprocal, in the sense that an agent tapping into another player's signal does not imply the symmetric relation. The benefit of connecting to an agent is that if i connects to j , agent i is able to observe j 's signal, even though j does not observe i 's signal. In addition, there is no retransmission of information—by connecting to j , agent i does not observe any signal that j has tapped into, unless i taps into those signals as well. Making connections is costly and the cost is solely carried by the linking party.

Figure 1 presents examples of information structures and their respective information sets. In Panel (a), each agent is tapping into one other player's signal, and thus obtains three signals: her own, the common prior, and the signal being tapped into. In Panel (b), agents coordinate on whom to focus their attention. All players are tapping into the signal of the first player, while she does not tap into anyone's signal. The first agent's information set is composed only of her own signal and the common prior, while another agent's information set is composed of his signal, the common prior and the first player's signal.



Agent 1's information set is $\mathbb{I}_1 = \{e_0, e_1, e_2\}$, while for agent 4, it is $\mathbb{I}_4 = \{e_0, e_1, e_4\}$.

Agent 1's information set is $\mathbb{I}_1 = \{e_0, e_1\}$, while for agent 4, it is $\mathbb{I}_4 = \{e_0, e_1, e_4\}$.

Figure 1: Two different information structures (networks) and the differences implied in agents' information sets.

Action Optimality After obtaining a collection of signals, an agent chooses an action to maximize her payoff. Each agent's signal is equally informative *ex-ante*, since signals are equally precise about the true state of the world. However, tapping into an agent's signal informs about more than the true state of the world: it also informs about what other players know and how they may act. Information from different agents is endogenously different. The position of an agent in the network defines how much information her signal carries about the average action.

Once agent i learns from the signals she has tapped into, agent i chooses an action, a_i , to maximize her expected payoff given other players' actions, a_{-i} ,

$$\Pi(a_i, a_{-i}) = -(a_i - a_i^*)^2 - c(K_i),$$

where $a_i^* = (1 - r)\theta + r\bar{a}_{-i}$ and $\bar{a}_{-i} = \frac{1}{n-1} \sum_{j \neq i} a_j$.

We capture the adaptation and coordination motives in an agent's actions with the bliss action, namely a_i^* . The bliss action is a linear combination of the true state of the world and the average action of all other players in the economy, excluding player i .¹⁰ If an agent knew both the true state and the average action, she would choose a linear combination of both variables. The parameter r captures conformity: how much an agent cares about the average action. An agent's payoff is a function of the distance between her action and the bliss action. The assumptions of normality of

¹⁰Alternatively, we could have defined $a_i^* = (1 - \tilde{r})\theta + \tilde{r}\bar{a}$, where $\bar{a} = \frac{1}{n} \sum_j a_j$. This leads to the same optimal action, by setting $\tilde{r} = \frac{rn}{r+n-1}$.

the signal structure and quadratic payoffs are standard in the organizational economics literature¹¹.

Consistent with our motivation, we only consider the case of action complementarities. An agent would like her action to be close to the true state of the world as well as to the average action.

Assumption 1. *Actions are strategic complements: $r \in (0, 1)$.*

Relevant for a player's payoff is the cost incurred in acquiring information, represented by $c(K_i)$, where K_i is the number of signals agent i is tapping into, $K_i = \sum_{j=1}^n g_{i,j}$. By making the cost simply a function of the number of connections an agent has, we assume that the cost is identical for all agents and that the cost is anonymous—it is equally costly to connect to any agent. To capture the fact that tapping into another player's signal is costly, we assume the cost function to be increasing. We refrain from any other assumption on the cost function at this moment. Later, in Section 3.3 we discuss how increasing marginal cost affects the information structure in the economy.

Assumption 2. *Tapping into an additional player's signal is always costly: $C(K_i + 1) - C(K_i) > 0$.*

Timing of the Game In the first period, agents simultaneously choose connections (which signals to tap into), forming a directed network. A strategy for player i is a vector $g_i = (g_{i,1}, g_{i,2}, \dots, g_{i,n})$, where $g_{i,j} \in \{0, 1\}$ for each $j \in N$, $j \neq i$.¹² A connection, $g_{i,j} = 1$, is to be interpreted as agent i tapping into agent j 's signal.¹³ Private signals are realized, and agents observe their own private signal as well as the signals of other agents they chose to connect to. In the second period, agents simultaneously choose an action a_i , and payoffs are realized.

2.1 Complete Information

Consider a complete information version of the model. In this setting, it is common knowledge that all agents perfectly know the true state of the world which is a limiting case of the model, when $\sigma = 0$.

Proposition 1. *In a game of complete information there exists a unique equilibrium in which all players choose $a_i = \theta$. No connections are formed and each player receives the same payoff, 0.*

Proof. To verify that it is an equilibrium is trivial. To show the first statement, we proceed by contradiction: suppose $a_i = \theta$ is not the unique equilibrium. Without loss of generality, $\theta < \bar{a}$.

¹¹See Garicano and Prat for a survey and Calvò-Armengol et al. (2015) for an interesting interpretation.

¹²Throughout the paper we restrict our attention to pure strategies.

¹³Given that a player always observes the common prior and her own signal, and for simplicity of notation, we assume $g_{i,0} = g_{i,i} = 1$.

There is at least one player choosing $a_i \geq \bar{a}$. Contradiction. For the second statement, connections are costly and generate no benefit. \square

2.2 Information Acquisition

We solve the model by backward induction. In the final stage, each agent chooses an action, given her information set. In the first stage, agents acquire information about the state of the world by tapping into the signals of other agents in the economy, forming an information structure. The information set of agent i is composed by the common prior, her own signal, and the signals she has chosen to tap into in the previous stage, and is described by

$$\mathbb{I}_i = \{e_j \text{ with } j = 0, 1, 2, \dots, i, \dots, n, \text{ such that } g_{ij} = 1\} = \mathcal{X}_i e,$$

where the matrix \mathcal{X}_i selects the signals agent i observes and e is the vector of signals in the economy, one for each agent and the common prior, e_0 .

Optimal Action The agent chooses an action in order to minimize the distance of her action to the bliss action. An agent does not know either the true state of the world or the average action, but uses the information available to her to choose an optimal action that solves:

$$\text{Max}_{a_i} -\mathbb{E} \left[(a_i - a_i^*)^2 + c(K_i) \right].$$

The first order condition leads to the following optimal action:

$$a_i = E[a_i^* | \mathbb{I}_i] = (1 - r)E[\theta | \mathbb{I}_i] + rE[\bar{a}_{-i} | \mathbb{I}_i].$$

An agent's optimal action is a linear combination of the best predictor of the true state and the best predictor of the average action. The weight given to each is determined by the parameter r , which captures how much the agent conforms to the average action. The optimal action of an agent is a combination of the signals she observes—the signals of others she has tapped into, her own signal, and the common prior.

Definition 1. An *action-strategy* is a choice of action, conditional on the information received by the agent. Formally, it is a function that maps all possible realizations of the agent's information set, given the information structure, to the real line, $a_i : \mathbb{R}^{K_i+2} \rightarrow \mathbb{R}$.

A *linear action-strategy* is an action-strategy that is a linear combination of the signals the agent observes.

A linear action-strategy equilibrium is a Nash equilibrium in which all agents play a linear action-strategy.

The following proposition shows the existence and uniqueness of a linear action-strategy equilibrium, for any information structure. The restriction to linear equilibrium is a standard assumption in the literature using the normal-quadratic approach.¹⁴

Proposition 2. *For any information structure in the first stage of the game, there exists a unique linear action-strategy equilibrium.*

The proof of the proposition is presented in the Appendix. The proof consists of guessing and verifying a linear equilibrium, and showing it is unique. The restriction to linear action-strategy equilibrium implies that an action is a linear combination of the signals in the economy—with the obvious restriction that a signal not observed must be assigned to a coefficient of zero. Also, the average action of the economy is a linear combination of the signals. Furthermore, if player i 's action and the average action are linear combinations of signals, it must be that the average action of all players, except player i , is also a linear combination. Finally, note that the average action and the average action excluding player i 's action are related, implying a relationship between the linear coefficients.

Corollary 1. *The following properties regarding linearity of the actions are true:*

- *The individual action of any agent i , a_i , the average action, \bar{a} , and the average action not including an agent i , \bar{a}_{-i} , are linear combinations of the signals in the economy and can be written as:*

$$\begin{aligned} - a_i &= \sum_{j=0}^n g_{i,j} \lambda_{i,j} e_j \\ - \bar{a} &= \frac{1}{n} \sum_{i=1}^n a_i = \sum_{j=0}^n \beta_j e_j = \beta' e \\ - \bar{a}_{-i} &= \frac{1}{n-1} \sum_{s=1, s \neq i}^n a_s = \sum_{j=0}^n \beta_{-i,j} e_j = \beta'_{-i} e. \end{aligned}$$

where $g_{i,j}$ is interpreted as whether agent i pays attention to signal j , and $\lambda_{i,j}$ as how much agent i 's action is influenced by j 's signal. Similarly, β_j and $\beta_{-i,j}$ are interpreted as the influence that the signal of agent j has over the average action, and over the average not including player i .

- *Since $\bar{a} = \frac{n-1}{n} \bar{a}_{-i} + \frac{1}{n} a_i$, we have that $\beta_j = \frac{n-1}{n} \beta_{-i,j} + \frac{1}{n} \lambda_{i,j}$.*
- *The linear coefficients are interpreted as relative influence, given that they sum to one, $\sum_{j=0}^n \lambda_{i,j} = 1$, $\sum_{j=0}^n \beta_j = 1$ and $\sum_{j=0}^n \beta_{-i,j} = 1$.*

¹⁴For instance, see Angeletos and Pavan (2007), Calvò-Armengol et al. (2015) and Dewan and Myatt (2008).

The linear coefficients for the average action, β_j , and for the average action excluding a player, $\beta_{-i,j}$, are network centrality measures that indicate an agent's signal influence. The coefficient formulas are explicitly stated in the Appendix, and the corollary is shown at the proof of Proposition 2. Focusing on the influence over the average action, we have:

$$\beta_j = \frac{1 - \tilde{r}}{n} \sum_{i=1}^n \frac{g_{ij}}{\sigma^2 + K_i + 1} + \frac{\tilde{r}}{n} \left[\beta_j (\bar{K}_j + 1) + \sum_{i=1}^n \frac{g_{ij}}{\sigma^2 + K_i + 1} - \sum_{i=1}^n \sum_{s=0}^n \frac{\beta_s g_{is} g_{ij}}{\sigma^2 + K_i + 1} \right],$$

where $\tilde{r} = \frac{rn}{rn+(n-1)(1-r)}$, and $\bar{K}_i = \sum_{j=1}^n g_{i,j} = 1$.

The influence of agent j 's signal on the average action is a measure of how central is her position on the network. It is a function of how many players are looking at her, \bar{K}_j , but also of how many and whose signals are the players that are looking at her tapping into. And how many and whose signals are those players tapping into, and so on. Influence is expressed as a recursive formula, depending on the influence of each and every other player. A player is very influential if many other players are tapping into her signal, and only into her signal. For instance, if all other players have only her signal to base their action decisions on, her signal has significant impact on the average action.

Network Formation When an agent decides which connections to form, she has not observed any signal. The first step is to compute the agent ex-ante expected payoff, as a function of other players' connections and their resulting optimal actions (that is, the expectation over all signals in the economy, except the prior, taking as given only the connection strategies of the others).

$$\text{Max}_{g_i} E[\Pi(g_i, g_{-i})] = -E[(E[a_i^*|I_i] - a_i^*|I_i)^2|G_{-i}] - c(K_i) \quad (1)$$

Proposition 3. *The ex-ante expected payoff of agent i choosing a set of signals to tap into, can be written as a function of her own information choices, g_i , as well as the influence that players have over the average action not including agent i 's action:*

$$E[\Pi(g_i, g_{-i})] = -\frac{\sigma^2}{K_i + \sigma^2} \left(1 - r \sum_{j=0}^n g_{i,j} \beta_{-i,j} \right)^2 - \sigma^2 r^2 \sum_{j=0}^n (1 - g_{i,j}) \beta_{-i,j}^2 - c(K_i)$$

The proof of the proposition is presented in the Appendix. It consists of a long sequence of algebraic manipulation, where the last step is a direct result of the Sherman-Morrison theorem regarding the inversion of matrix additions. It is worth noting that $\beta_{-i,j}$ is a function of the linear action-strategies used by all other agents. Given that the information structure is not observed at

any stage, $\beta_{-i,j}$ is not a function of player i 's action or connection strategies. It depends on which connections the other players expect player i to form, and in equilibrium these expectations must be correct.

The payoff specification above makes explicit the trade-off involved in making a connection. By tapping into another agent's signal, player i increases her payoff in two ways. First, by adding a link, she increases the information she knows regarding the average action, bringing the first term closer to zero. At the same time she reduces the information she does not know regarding the average action, bringing the second term closer to zero as well. Finally, her payoff decreases by the increase in the cost associated with tapping into another agent's signal.

Another feature of the payoff suggested by the formulation above is concavity of the benefits of information acquisition. Suppose all players had the same influence on the average action not including agent i , $\beta_{-i,j} = b \quad \forall j$. Note that the first connections add more to the payoff than subsequent ones.

Finally, if agent i were to connect to a set of players, she would choose those players that are more influential regarding the average action excluding her own action. The following proposition captures this monotonicity result.

Lemma 1. *For any agent i , and set of connections g_i , if agent i is willing to tap into one other player signal, she will choose the most influential agent with respect to the average action not including her own. That is, agent i will choose to tap into the signal that has the highest $\beta_{-i,j}$ among the signals she is not observing.*

Proof. Observe that player's i payoff derivative with respect to $\beta_{-i,j}$ is strictly positive. □

3 Equilibrium Information Structures

In this section, we characterize properties that any equilibrium information structure satisfies. The payoff specification from Proposition 3 allows us to understand two features of an agent's best response function. First, an agent would rather connect to a more influential player, and second, tapping into other players' signals has diminishing marginal returns. In subsection 3 we define two network properties, and show how they correspond to the two main forces behind information acquisition decisions in our model. Subsection 3.2 presents our main result. We show that these two properties alone are sufficient to characterize all information structures in equilibrium as hierarchical networks. We also present interpretations of this particular information structure. Subsection 3.3 discusses sufficient conditions on the link formation cost function to have only core-periphery networks in equilibrium, a particular case of a hierarchical network. Furthermore, we discuss the

prevalence of core-periphery networks and the interpretations of it. Finally, subsection 3.4 presents numerical examples as well as comparative statics.

3.1 Link Formation Incentives

The fact that agents are ex-ante identical does not imply that the source of information is irrelevant. In equilibrium the source affects how much that information is considered. While informing equally about the true state of the world, each signal informs differently regarding the average action, depending on its influence on other players' actions.

Given the connections other players are choosing, and the resulting influence on the average action not including player i 's action, player i chooses to connect to the most influential players (with highest β_{-i} , according to Lemma 1). When making a decision, player i is concerned with the most influential signals regarding the average action not including her own action; hence different agents may rank other players' influence differently. That is, two players, i and s , may rank signals' influence differently, since they rank based on β_{-i} and β_{-s} respectively. However, we show that in equilibrium all agents share the same ranking over signals' influence. We show that an agent's influence is linked to her in-degree centrality, a network centrality measure, the number of players tapping into a signal. We interpret in-degree centrality as the *publicity* of a given signal: a signal that is more observed is more public. Property 1 summarizes this feature.

Property 1. *Influence is Publicity*

If an agent is tapping into another agent's signal, she must be also tapping into the signal of any other, more in-degree central, agent.

$$\forall i \ g_{i,l} = 1 \implies g_{i,m} = 1, \forall m : \bar{K}_m \geq \bar{K}_l,$$

where $\bar{K}_j = \sum_i g_{i,j}$.

The intuition for Property 1 is that, when tapping into a signal, an agent obtains information regarding both the average action and the true state of the world. Even though all signals are equally informative about the state of the world, the same is not true for the average action: a signal from a more observed source is more informative about the average action. All players would rather connect to a more observed player, since by tapping into a more observed player's signal, player i acquires the same amount of information about the state of the economy and more information about the average action. There exists a monotonicity regarding whom to connect to: one always wants to observe the players that are more observed, because their signals are more public. If agent i is connected to agent l , and there is another agent, m , whose signal is more tapped than l , then

i must also be connected to m in equilibrium. These more observed players are opinion makers, given their influence on the average action.

The fact that a player's influence, β_j , can be mapped into in-degree centrality is not a feature of the influence, but an equilibrium result. To illustrate that property 1 may not hold for some information structures, we provide the following example.

Example 1. Out of equilibrium, influence is not in-degree centrality

Consider the following network: four players are tapping into the signals of a set of players, C_2 , while three other players are tapping into the signal of a single player, C_1 . It is easy to see that if C_2 is a singleton, it is more in-degree central and more influential than C_1 . However, as we increase the number of players in the set C_2 , as we see in the figure, we diminish the influence a particular node in C_2 has. It diminishes in such fashion that if there are enough players in C_2 , a node in C_2 becomes less influential than C_1 , even though the in-degree centrality remains the same.

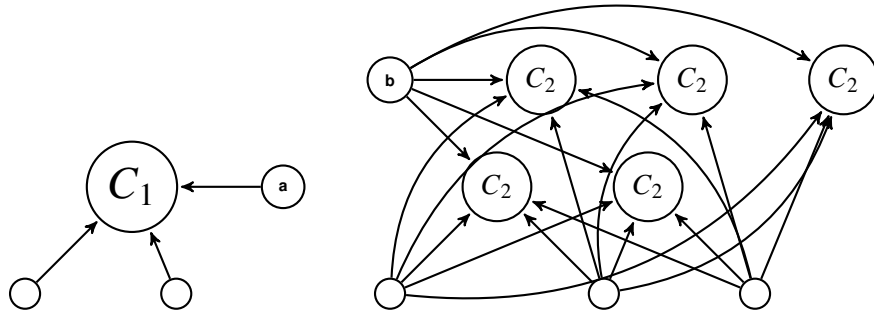


Figure 2: A node C_2 is more in-degree central, but is less influential than C_1 .

The equilibrium requirement allows us to rule out networks like the above by comparing the influence of C_1 and of an agent in C_2 over the average action without a and without b . Observe that:

$$\beta_{C_1} = \frac{n-1}{n}\beta_{-a,C_1} + \frac{1}{n}\lambda_{a,C_1} \quad (2)$$

$$\beta_{C_1} = \frac{n-1}{n}\beta_{-b,C_1} + \frac{1}{n}\lambda_{b,C_1} = \frac{n-1}{n}\beta_{-b,C_1} \quad (3)$$

$$\beta_{-b,C_1} > \beta_{-a,C_1} \quad (4)$$

By a similar reasoning, for a player in C_2 , we have that $\beta_{-b,C_2} < \beta_{-a,C_2}$. By agent a optimality, we have that $\beta_{-a,C_1} > \beta_{-a,C_2}$, which implies $\beta_{-b,C_1} > \beta_{-b,C_2}$, a contradiction to agent b optimality.

The fact that agents are ex-ante identical does not imply that signals received for free are equally informative. A more tapped-into signal is more informative about the average action.

Combined with concavity of the benefits of information, we have a monotonicity result in information acquisition. If a player signal is being observed by more players, her own signal is more informative of the average action. As a result, she has less incentives than a less observed player has to tap into other players' signals. Property 2 summarizes this feature.

Property 2. *Know your place in the network*

If a player's signal is more tapped into than another player's, she cannot be observing more signals.

$$\bar{K}_f \geq \bar{K}_h \implies K_f \leq K_h \quad \forall f, h,$$

where $\bar{K}_j = \sum_i g_{i,j}$ and $K_i = \sum_j g_{i,j}$.

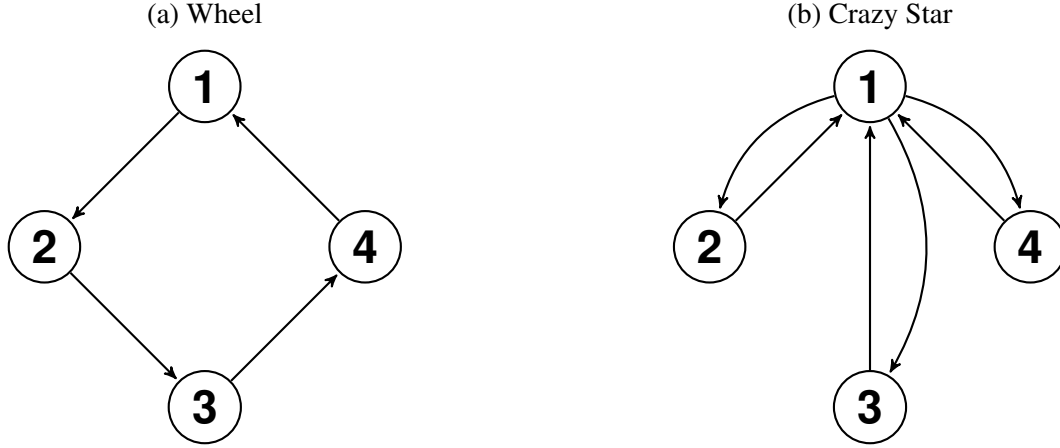
Information acquisition is more valuable to players who have less information. A player whose signal is tapped into by many others has, in fact, more information than a player whose signal is not tapped into. In equilibrium, players are aware of their place in the network—how many players are tapping into their signal—and act accordingly.

3.2 Characterizing Equilibria

First, we establish that the properties above hold true in equilibrium.

Proposition 4. *Given Assumptions 1 and 2, any linear action-strategy strict Nash equilibrium of the game described above satisfies Properties 1 and 2.*

The proof is presented in the Appendix. The proposition restricts which information structures can occur in equilibrium. For instance, neither the Wheel nor the Crazy Star depicted below occurs in equilibrium. The wheel does not satisfy Property 1, while the Crazy Star does not satisfy Property 2.



The Wheel information structure is not an equilibrium because it violates Property 1.

The Crazy Star information structure is not an equilibrium because it violates Property 2.

Figure 3: Property 1 and Property 2 restrict the set information structures that can occur in equilibrium

The first property that an information network satisfies in equilibrium suggests that agents are organized in layers of influence, which translate into layers of in-degree centrality. More influential agents have positions at the top layer of the hierarchy. The second property shows that a more influential agent has less incentive to acquire information. The hierarchical ordering of agents according to their influence implies that agents who are at the top of the hierarchy acquire less information than those at the base. The following definition captures these intuitions.

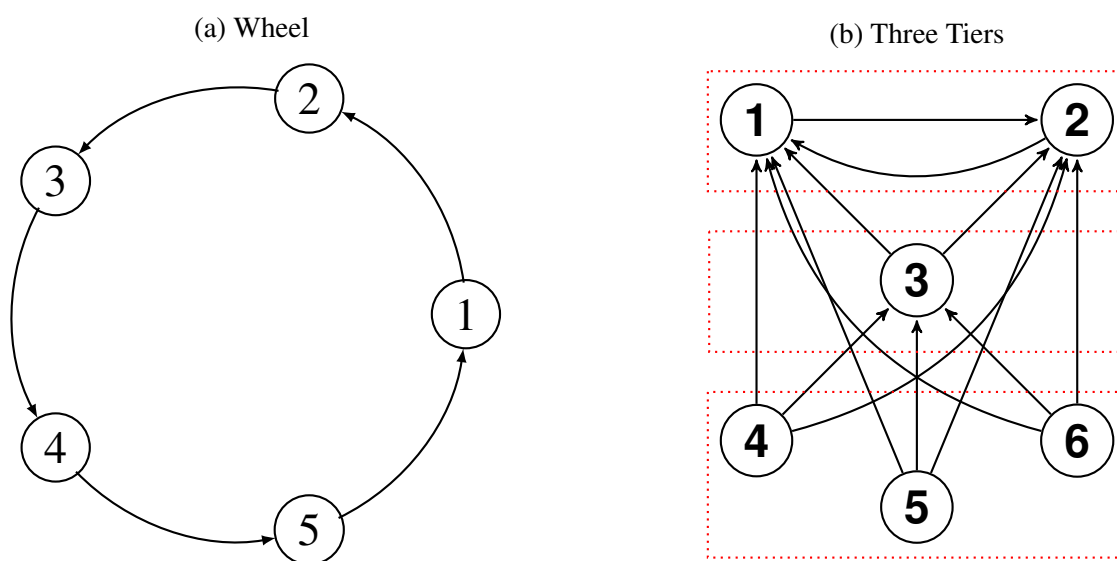
Definition 2. A network is a hierarchical directed network if, and only if, there exists a partition on the set of agents, $A_{s \in \{1, 2, \dots, N\}}$, such that:

- i) $i \in A_1$ if, and only if, $\forall j \notin A_1, g_{ji} = 1$.
- ii) $i \in A_s$ if, and only if, $\forall j \notin \cup_{k=1}^s A_k, g_{ji} = 1$.
- iii) if there exist $i \in A_s$ & $j \notin \cup_{k=1}^{s-1} A_k$, such that $g_{i,j} = 1$, then $g_{l,m} = 1 \forall l, m \in A_s$.
- iv) If there exist $i \in A_s$ & $j \in \cup_{k=1}^{s-1} A_k$, such that $g_{i,j} = 1$, then $j \in A_{s-1}$ and for any $l \in A_s, g_{l,j} = 1$. Furthermore, $g_{l,m} = 0$ for any other $m \in \cup_{k=1}^{s-1} A_k$.

A network is a hierarchical directed network if agents can be partitioned in tiers, such that the signal of any agent of a certain tier is tapped into by all agents in tiers below hers. Furthermore, within a certain tier, either all agents are tapping into the signals of all other agents in their same tier or no one is. Finally, if a player of a certain tier is tapping into the signal of a player in a tier

below hers, all other members of her tier are as well and that is the only member of a tier below that any of them observes.

To help understand the definition we present a couple of examples in Figure 4. Consider the wheel network in Panel (a). It is not hierarchical, since if all players are in the same level the level is neither full nor empty and if the players are in different levels, not all players from a level below are looking to all in upper levels. On the other hand, consider a three tier network, depicted in Panel (b). Such network satisfies the hierarchical definition. Consider the partition that places players 1 and 2 in one group, player 3 in a second group, and the other players in a third group. Note that all groups are either empty or full, with the first group being full and the third empty. Also, note that all players in the second and third groups do observe all players in the first group. Finally, all players in the third group observe the player in the second group.



The Wheel is not a hierarchical directed network.

The Three Tiers network is a hierarchical directed network.

Figure 4: Examples of networks: one not hierarchical and one hierarchical.

Properties 1 and 2 restrict the possible networks that can occur in equilibrium. However, this restriction is not particular to the model presented. Formally, Properties 1 and 2 restrict the set of equilibria of a more general class of games than the one discussed here so far. That is, if a strict equilibrium of a different game of network formation satisfies both properties, then the resulting network is restricted in the same way. Indeed, we can extend the original simultaneous link-announcement game presented in Jackson and Wolinsky (1996) and first defined by Myerson (1977) to include directed network formation and a richer strategy space, that includes actions as well as the link formation process.

Definition 3. A *directed link formation game*, $\Gamma = \langle N, \mathbf{G}, A, u \rangle$, consists of a finite set of players, $N = \{1, 2, \dots, n\}$; link-choice sets $G_i = \{0, 1\}^{N \setminus \{i\}}$, where if $G \in G_1 \times \dots \times G_n$ is the link-choice profile played, then i is connected to j if $g_{i,j} = 1$; action-choice sets A_i ; and payoff functions $u_i : G \times A \rightarrow \mathbb{R}$.

For instance, in our game, an element of G_i is a list of players' signals that player i chooses to tap into, while an element of A_i is a map from all the possible signal realizations player i observes to a real valued action.

Theorem 1. Consider a directed link formation game. If **Property 1** and **Property 2** hold in a strict Nash equilibrium of this game, then the resulting network is a Hierarchical Directed Network.

For the particular game analyzed in this paper, the theorem implies that all strict Nash equilibria present a Hierarchical Directed Network information structure. We phrase the conditions for the theorem in terms of the network properties instead of parameter restrictions, since Property 1 and Property 2 hold true in our model and the theorem is more general than the model presented. We interpret Property 1 and Property 2 as payoff assumptions in a more general framework, with Theorem 1 as a game-theoretical result. We present the proof below. We inductively construct the sets of definition 2, A_1, \dots, A_n . To characterize the set A_1 , we show that all players not in such a tier must be observing any player in it. We do so by an inductive argument.

Proof. Consider an equilibrium of the network formation game. Define the following set: $i \in B_1 \iff \bar{K}_i \geq \bar{K}_k \forall k$. The proof will proceed inductively, and we first show that the set B_1 above is indeed the set A_1 of the definition of hierarchical directed network.

First of all, notice that if $G = \emptyset$ the proof is done, since the empty network is a hierarchical network.

Thus let us focus on the interesting case: G is not empty, and thus B_1 is not empty with k_1 being an element of B_1 .

The first step is to show that if $k_2 \notin B_1$, then it must be the case that $g_{k_2, k_1} = 1$. That is, if k_2 is not among the set of most connected players it must be observing k_1 . The structure of the argument will be based on contradiction.

As usual, assume not. $\exists k_2 \notin B_1 : g_{k_2, k_1} = 0$

If $g_{k_2, j} = 1$ for some j we are done, since we can use Property 1 and arrive at a contradiction already. Player k_2 should be observing player k_1 . Observe that this holds true even if $k_2 \in B_1$.

But what if $g_{k_2, j} = 0 \forall j$? Then, we know that:

1. $g_{k_2, s} = 0 \forall s : \bar{K}_s \geq \bar{K}_{k_2}$
2. Let B_N be the group of the least observed players. $k_2 \notin B_N$, otherwise by Property 2 we would have an empty network.
3. $k_2 \notin B_1 \implies \bar{K}_{k_2} < \bar{K}_{k_1}$, thus $\exists k_3 : g_{k_3, k_2} = 0$ and $g_{k_3, k_1} = 1$
4. By Property 2, it must be that $\bar{K}_{k_3} < \bar{K}_{k_2}$
5. By the same arguments as before, it must be that $\exists k_4 : g_{k_4, k_3} = 0$ and $g_{k_4, k_2} = 1$
6. If we proceed by induction, we will finally conclude that $\exists k_N : g_{k_N, k_{N-1}} = 0$ and $g_{k_N, k_{N-2}} = 1$.
 1. Observe that since this is the last step in the induction, it must be that $k_N \in B_N$.
7. However, if $g_{k_N, k_{N-1}} = 0$, then $\bar{K}_{k_{N-1}} = 0$. And thus, $k_N - 1 \in B_N$. Contradiction.

Second, we need to show that if one element of the set B_1 is being observed by any other element of B_1 , it must be observed by all elements in B_1 and also it must observe all players in B_1 .

1. If $k_1 \in B_1$ is being observed by one other player in B_1 it must be the case that all players in B_1 are being observed by one other member in B_1 . The reason for this is that they all receive the same number of looks from players outside B_1 and they also have the same number of total players looking at them.
2. If all players are looking at someone, from Property 1, we already have that all players are looking at everyone else in B_1 .
3. What if only a subset of players from B_1 are looking at someone in B_1 ? Call this subset $S_{looking}$ and let the complement be $S_{blind} = B_1 \setminus S_{looking}$. By Property 1, every member of $S_{looking}$ is looking at everyone else in B_1 .
4. Observe then that the number of looks an element of $S_{looking}$ receives is strictly smaller than the number of looks an element of S_{blind} receives. The reason for that is that a player in S_{blind} receives from other players in B_1 $\bar{K}_{S_{looking}} + 1$ looks, where $\bar{K}_{S_{looking}}$ is the cardinality of the set $S_{looking}$ and the extra one is from himself. While a player in $S_{looking}$ receives $\bar{K}_{S_{looking}}$ looks, contradiction.

In order to finish constructing the set B_1 , we need to show that if a player in the first tier is looking down at a player not in the first tier, then: (a) it must be looking at all players in the first tier, (b) all other players in the first tier are looking at that particular player, (c) it is looking at a player in the second tier, and (d) it is not looking at any other player in any level below the first.

The proofs of items (a) and (c) are trivial applications of Property 1, while for item (b) the same proof (that a level is either empty or full) presented above holds.

To show item (d), first suppose not: that is $g_{i,j} = 1 \& g_{i,l} = 1 \ j, l \in A_2$. The second tier cannot be full, otherwise players j and l would receive as many looks as player i , in the first tier, and they would not be members of the second tier, but of the first. The next step is to count the number of looks by each agent, let us start with agent j : $K_j = \bar{N} \equiv$ cardinality of $[\cup_1^n A_k]$, while $K_i = \bar{N} - 1 + 2$. Contradiction to Lemma 2.

Thus the set B_1 constructed satisfies the definition to be the set A_1 .

The next step is to construct the second level, B_2 and to show it actually is A_2 . Begin by ignoring any looks cast by the first level and simply follow the same steps detailed above. By induction we can proceed to construct B_2, B_3, \dots

□

The proposition restricts the set of information structures that can occur in equilibrium to have a particular shape. Furthermore, it simplifies the sufficient statistics to compute equilibrium payoff. Instead of the whole network, the pair of vectors containing the information of how many signals is each player observing and how many agents are observing each signal is sufficient.

Corollary 2. *Equilibrium payoffs depend only on the number of signals players tap into, and on how many players are tapping into each signal, the pair of vectors $(K, \bar{K})_i \in \{1, 2, \dots, n\}$.*

The proof is the algorithm to construct G from the pair of vectors (K, \bar{K}) . The only involved point concerns the fact that a whole tier may be looking at one player of a tier below.

Proof. The algorithm is done by induction.

- 1 Rank players by \bar{K} . The first group will be those with highest \bar{K} . Those are the members of A_1 . Let $\#A_1$ be the number of players in A_1 .
- 2 Let K_1 be value of K for the members of the set A_1 . If $K_1 \leq \#A_1 - 1$, then A_2 will be the set of players with higher \bar{K} , excluding A_1 .
- 3 If $K_1 = \#A_1$ then subtract $\#A_1$ from \bar{K} for the player with higher \bar{K} , excluding A_1 . Using this modified vector, A_2 will be the set of players with higher \bar{K} , excluding A_1 .
- 4 repeat the steps.

□

Before we proceed, we show that the unusual phrasing of Theorem 1 is not vacuous. There exist payoff functions that satisfy Property 1 and Property 2 outside of the model, which underscores that Theorem 1 applications go beyond the scope of the information model presented.

Example 2. A Different Payoff Specification

Consider the following payoff function. A player receives a payoff that depends on how many players tap into her signal, on how many players’ signals she is tapping into, and also on the relative quality of those signals—captured by the ratio of how many players tap into those signals compared to how many tap into her own signal. Finally, observe that such payoff specification satisfies both Property 1 and Property 2.

$$\Pi(G) = \sum_{j \neq i} g_{i,j} \frac{\bar{K}_j + 1}{\bar{K}_i + 1} + n^2 \bar{K}_i - cK_i$$

We parametrize the model with $c = 1.53$ and $n = 40$, and present a three tier equilibrium shown in the graph below. Only one player is in the first tier, while 15 are in the second and 24 are in the third. A player in the second tier observes only the player in the first, while players in the third tier tap into the signal of players in the first and second tiers.

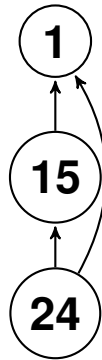


Figure 5: With this different payoff specification, that satisfies Property 1 and 2, the three tier network represented is an equilibrium. It is a particular hierarchical directed network.

Information Hierarchies The central result so far is the endogenous emergence of a hierarchical information structure. The hierarchical information structure we have defined resembles what has been commonly defined as a social hierarchy: an implicit or explicit rank of individuals with respect to a certain social dimension.¹⁵ In our model, the number of people tapping into an agent’s private signal serves as the social dimension used to rank those agents. Furthermore, this implies a ranking regarding the influence that each agent’s signal has over the average action, with more influential agents being at the top of the social hierarchy.

In equilibrium, agents endogenously separate themselves into groups (what we have called tiers), characterized by the fact that all players belonging to a certain group not only exert the same influence over the average action (and have their private signals tapped by the same individuals),

¹⁵For a formal definition, see Magee and Galinsky (2008).

but also choose symmetric information acquisition strategies. It is not true that agents of the same tier choose the same action, however they all choose the same signals to tap into. The hierarchy of influence is self-enforcing in the model. As more players choose to tap into a particular signal, the influence of that signal for the average action increases, which increase the payoff of tapping into it. Social hierarchies have been documented to be self-enforcing, for examples see the seminal study of Rosenthal and Jacobson (1968), focused on student’s learning and teacher’s expectations of student’s quality, and the review contained in Magee and Galinsky (2008). Finally, it is worth noting that the payoff of a player at the top of the hierarchy must be higher than the payoff of a player at the bottom of it. A top tier player can always copy the information acquisition of any player that is at any level below her (since all those players are tapping into her signal), and if she chose not to do so her payoff must be higher. This is true even though in any equilibrium players in the bottom of the structure have better information than players at the top. The information set of a player at the top is weakly contained in the information set of a player at the top, and thus players at the bottom choose actions closer to their bliss action.

The definition of hierarchical network suggests a nestedness of influence. If an agent is influenced by a player of a certain tier, she is influenced by all players of tiers above. The idea of nestedness is a well-documented aspect of real-world networks, for instance see Königa et al. (2014).

Another key element of the endogenous emergence of a hierarchical information structure lies in the allocation of attention in the society. In equilibrium, some agents’ signals are very public and widely influential (top tier) while others’ signals are private and have very little impact. This asymmetry occurs even though agents are ex-ante identical. The ex-post asymmetry implies payoff differences among agents, with agents in the top of the social ladder accruing higher payoffs. This result is in stark opposition to Garicano (2000), in which ex-ante identical agents differentiate themselves through costly information acquisition and occupy different positions in a social hierarchy, but all receive the same payoff.

3.3 Core-Periphery Information Structure

We have shown that any equilibrium information structure has a hierarchical directed structure. In this subsection, we refine the set of possible equilibria even further by restricting the cost of forming links to have weakly increasing marginal costs.

Assumption 3. *The cost of obtaining information, $C(K_i)$, is a strictly increasing function with*

weakly increasing differences:

$$C(K_i + 1) > C(K_i)$$

$$C(K_i + 2) - C(K_i + 1) \geq C(K_i + 1) - C(K_i)$$

The increasing marginal cost assumption implies that it is cheaper to tap into an extra link if you have tapped into fewer links. Note that the assumption requires only weakly increasing marginal costs, and in particular allows for constant marginal cost.

As an agent forms links and acquires information, she expands her information set. If a player i is tapping into player j 's signal in a Hierarchical Directed Network, this implies that player i 's information set is at least as comprehensive as player j 's information set. For instance, consider the Three-Tier hierarchical network presented in Figure 4. The information set of a player in the bottom tier, for instance $I_5 = \{e_0, e_1, e_{2,3}, e_5\}$, is composed of the common prior, her own private signal, and the signals of players 1, 2, and 3. Meanwhile the information set of a player in the top tier, for instance $I_1 = \{e_0, e_1, e_2\}$, is composed of the common prior, and the signals of players 1 and 2. Thus the information set of a player in the bottom tier is a strict superset of the information set of a player in the top tier. Furthermore, it would still be a superset even if the player in the bottom tier stopped tapping into the middle tier. As the benefit of acquiring information is concave, we have a restriction on how different the information sets of different agents can be.

Property 3. *All of those that influence me, influence each other*

If a player taps into two distinct players' signals, then both of these players tap into each other's signal:

$$g_{i,j} = 1 \text{ and } g_{i,k} = 1 \implies g_{k,j} = 1 \text{ and } g_{j,k} = 1,$$

for any distinct players, i, j , and k .

Information acquisition is more valuable to players who have less information. Furthermore, as the cost of acquiring additional information is increasing, this implies that players that have acquired more information are less willing to further acquire information. First, we establish that the property above holds true in equilibrium.

Proposition 5. *Given Assumptions 1, 2, and 3, any linear action-strategy strict Nash equilibrium of the game described above satisfies Properties 1, 2, and 3.*

The proof is presented in the Appendix. We show that, in any hierarchical information structure that does not satisfy Property 3, a player in a lower tier is acquiring strictly more information than a player in the upper tier—her information set contains the information set of the player in the upper tier. Furthermore, she would still acquire more information even if she had dropped a link. Due to

increasing marginal cost, if it is not worthwhile for the upper tier player to connect to such player, it cannot be for the more informed player in the lower tier.

The proposition restricts which hierarchical directed information structures can occur in equilibrium. For instance, given Assumption 3, the Three Tiers hierarchical network presented in Figure 4 will not occur in equilibrium. Player 4 is looking to both player 3 and player 2, but player 2 is not looking to player 3. Indeed, any hierarchical directed network that has three or more tiers of influence will not occur in equilibrium for this same reason.

We show that, if the cost function is strictly increasing with weakly increasing marginal costs, then any equilibrium information structure is core-periphery. A core-periphery network is a particular hierarchical directed network, in which the set of players is partitioned in at most two groups—the core and the periphery. All players tap into the private signal of all members of the core, making their signals effectively public, while members of the periphery do not tap into each others' signals.

Definition 4. *A network is a directed **core-periphery** if the set of agents can be partitioned in two: the core $A_1 = \{i_1, i_2, \dots, i_n\}$ and the periphery $A_2 = \{j_1, j_2, \dots, j_n\}$. For any i and j , $g_{j,i} = 1$, $g_{i_s, i_r} = 1$, and $g_{j_s, j_r} = 0$.*

To help understand the definition above, we discuss some examples from Figure 6. Panel (a) considers the traditional example of a core-periphery network: a hierarchical network with only two levels, in which players in the lowest level and players in the top level all tap into the signals of all players of the top level. Panel (b) considers a more subtle type of core-periphery network: a hierarchical network with only two levels, in which players in the lowest level and players in the top level all tap into the signals of all players of the top level, however players in the top level also tap into the signal of one particular player of the periphery. Finally, Panel (c) considers a hierarchical network with only two levels, in which players in the lowest level tap into the signals of all players of the top level, but players at the top level do not tap into each others' signals. It is not a core-periphery, since the signal of a core player is not tapped by other members of the core. Although highly public, core players' signals are not public information. Also, note that the empty network is one possible core-periphery network, in which the core is empty, while the complete network is also a core-periphery network, in which all players are in the core.

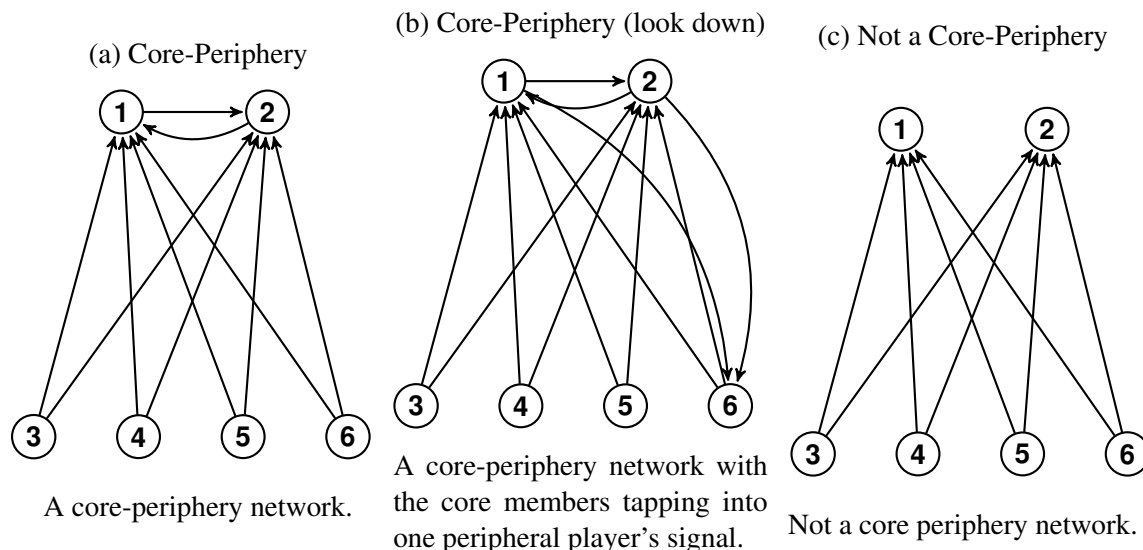


Figure 6: Different hierarchical networks. Only the first two are core-periphery networks.

Before we proceed it is useful to present some new graphical notation. The simplicity of a core-periphery structure allows us to represent any information structure as two tiers of information, and we associate with each tier the number of players in that particular tier. Furthermore, we associate the top tier with a black node, to denote that all agents in that tier are tapping into the signal of one another (full tier). Similarly, we associate the bottom tier with a white node, since no agent in that tier is tapping into a same-tier-agent's signal (empty tier). Finally, the notation used to indicate the case in which all members of a tier decide to look at one member of a lower tier is a dotted arrow from a tier to a lower one. The two core-periphery networks represented in Panels (a) and (b) of Figure 6 are represented in this notation in Figure 7.

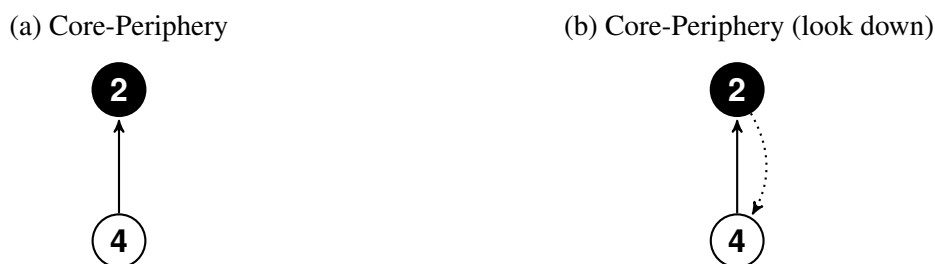


Figure 7: Two possible core-periphery networks for a 6 player configuration. Both networks have two players in the core, and the second network has the core players tapping into the signal of one peripheral player (looking down).

Theorem 2. *Consider a directed link formation game. If Properties 1, 2, and 3 hold in a strict Nash equilibrium, then such equilibrium present a Core-Periphery information structure.*

Proof. Consider an equilibrium of the network formation game. We know, by Theorem 1, that it is a hierarchical directed network. First, we show that it must be comprised of only two tiers. Second, we show that all players in the top tier must be connected to each other.

First, suppose there are more than two tiers. Thus, there exists a player in the bottom tier looking to all players in the top tier and all players in a medium tier. There also exists a player in a medium tier not being observed by the players in the top tier (otherwise there would not exist a medium tier). This violates property 3.

Second, suppose there is more than one player on the top tier, and at least one player on the bottom tier. From the definition of hierarchical networks, we know that either those players look at each other or at to no one. By contradiction, suppose the latter. This violates property 3, since the player on the bottom tier is looking at both, but they are not looking at each other. \square

For the particular game analyzed in this paper, this implies the following: if the marginal cost of acquiring information is increasing, then all strict Nash equilibrium of the game present a core-periphery information structure. Furthermore, within core-periphery networks, there are only two types of equilibrium information structures possible. In the first (call it type *A*), all players tap into the signals of all core players. The private signal of any core player is effectively public, being observed by all players. The signal of a periphery player is effectively private, with each signal from the periphery being observed only by its respective player. Given that all core players have the same information, they all choose the same action and accrue the same payoff. The expected payoff of a core player must be higher than that of a periphery player, since a core player could tap into the signal of one peripheral player to obtain the same information and the same payoff. Any peripheral player is always more informed than a player at the core. In the second type of equilibrium network (call it type *B*), all players tap into the signals of all core players, who in turn also tap into the signal of one particular member of the periphery, agent *s*. The private signal of any core player is again effectively public, but it is not true that the signal of any peripheral player is strictly private. The signal of the particular player *s* is observed by herself as well as by all core players. Again, it holds that all core players have the same information, thus they all choose the same action and accrue the same payoff. However, it is not true that the expected payoff of a core player must be higher than of any periphery player, since player *s* has the same information and accrues the same payoff as a core player.

Before we proceed, we highlight that the unusual phrasing of Theorem 2 is not a vacuous extension. There exist directed link formation games that satisfy Properties 1, 2, and 3 beyond the present model. For instance, consider a directed citation model, in which player *i* receives payoff for being cited as well for citing the most cited papers.¹⁶

¹⁶Consider a directed link formation game such that agent *i* payoff, as a function of the connections formed by all players, is given by: $u_i(g_i, g_{-i}) = \bar{K}_i + \sum_{j \neq i} g_{i,j} \bar{K}_j - cK_i$. Any strict equilibrium of the game with this payoff specification

Example 3. A Three-Tier Information Structure

In order to highlight the necessity of Assumption 3 for the result in Theorem 2, we momentarily abandon such an assumption and show the existence of an equilibrium that does not have a core-periphery structure: one of the equilibria below has three tiers. We parametrize the model with 5 agents that care equally about adapting to the state of the world and conforming to the average action, $r = 0.5$. We consider the private signal to be as informative as the common prior, $\sigma = 1$, and the cost of connecting to be the non-convex function:

$$c(K_i) = \begin{cases} 0 & \text{if } K_i = 0 \\ 0.1214 & \text{if } K_i = 1 \\ 0.1652 & \text{if } K_i = 2 \\ \infty & \text{if } K_i \geq 3. \end{cases}$$

There are five different equilibrium information structures. From left to right in the figure, we have the empty network, the star, the three tier, and the core periphery with two agents in the core, both with the core looking and not looking, respectively, to an agent in the periphery.

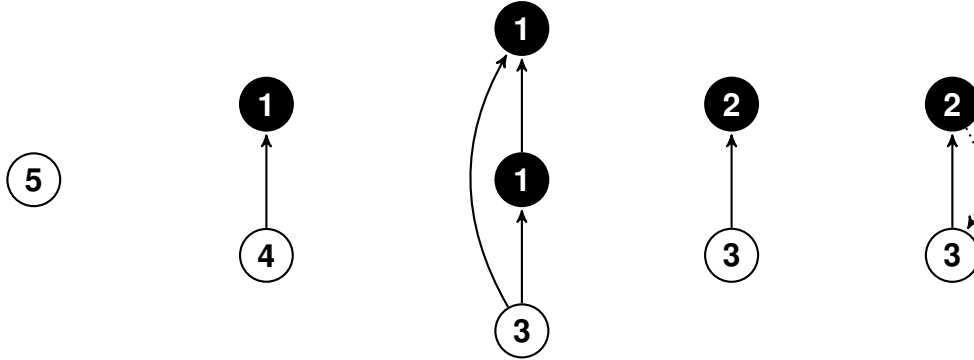


Figure 8: Five equilibrium architectures, among which one with three tiers.

Consider the equilibrium with three tiers. Note that the player in the top of the hierarchy has only her signal and the common prior on which to base her decisions, while a player in the third tier has his own signal, the common prior, and the signals of the players in the first and second tier. Thus even if he stopped paying attention to the player in the second tier, he would have more information than the player in the first tier. The non-convexity in the cost of acquiring information makes a more informed player at the bottom of the hierarchy want to look at two players (one player in the first tier and one in the second tier), while the less informed player at the top of

satisfies Property 1 since a player wants to cite the most cited paper. It also satisfies Property 2 and Property 3, since the incentives that j has to cite a paper are independent of how much j 's paper is cited.

the hierarchy wants to look at none. This happens because tapping into one player's signal is fairly expensive when compared to the price of tapping into two. The relative cheapness of further acquiring signals is what Assumption 3 prevents.

Core-Periphery In equilibrium, agents coordinate on their sources of information. Even though agents are identical ex-ante, the equilibrium information structure paints a very different picture: some players have great influence over the society, with everyone paying attention to them, while others are not paid any attention and have no influence. All agents coordinate and focus on a subset of signals available as a direct result of the need to coordinate actions. Because agents want to choose similar actions, each agent wants to know what others know to better predict how they act. It is important to note that the focus result is a consequence of complementarity, and not of discrete linkage choices. In the Appendix we solve a model with continuous choices of linkage, in which a player can acquire a fraction of the information of another player, and we show that equally spreading attention among all other players is not an equilibrium.

Tapping into another player's signal plays a dual role in our model, it enables learning about the world and, at the same time, inferring the information of others. As agents try to know what other players know, they are effectively trying to forecast the forecast of others, and higher order beliefs could affect their actions. The fact that in any equilibrium players coordinate, focusing on the same set of players, makes higher order beliefs irrelevant. Even though there are two types of equilibria, both share a common feature: the signals of all core players are public and known by all players, effectively common knowledge. We can interpret any equilibria as a “rational ritual”,¹⁷ a social practice creating common knowledge. The creation of common knowledge among agents help them coordinate, and thus to achieve higher payoffs.

Finally, core-periphery networks have been documented in a wide array of information acquisition situations—from product brands and fashion changes to voting decisions—Galeotti and Goyal (2010) presents stylized facts and a review of the literature. For instance, in a study regarding the differences between formal and informal organization structures of a firm and how information dispersion inside it impacts productivity, Cross and Parker (2004) divides the employees of a firm in “central people” and “peripheral people”. Regarding the spread of social behavior, Gladwell (2006) reserves a special importance for agents that connect all others, thus dispersing the information and allowing the social behavior to spread. Kumar et al. (2010) documents the dynamic evolution of *Flicker* and *Yahoo! 360*, social networks with more than 5 million users. They document the existence of multiple connected components, each organized roughly as a core-periphery structure. In a paper regarding online diffusion processes, Goel et al. (2012) paints a very startling picture of how information and behavior spreads online. Instead of a diffusion process similar to

¹⁷See Chwe (2013), for definition and a plethora of examples.

the spread of infectious disease—characterized by long sequences of links of contagion, triggering large cascades—online diffusion is characterized by almost all adoption happening within one degree of a seed node. By tracking the diffusion within a variety of internet based apps and websites, the authors document that 94% to 99% of adoptions happen within one degree of a seed. This implies that re-transmission of information, although tempting as a story, is not a central point to the diffusion of information. Bakshy et al. (2011) finds similar numbers when analyzing contagion on twitter.

3.4 Numerical Examples

Although we restrict the set of networks that emerge in equilibrium to hierarchical structures (Theorem 1) and, when the cost of acquiring information is convex, core-periphery (Theorem 2), multiplicity of equilibria is pervasive in the model for two reasons. The first one regards the arrangement of agents. Given one particular information structure, there are multiple ways of arranging the agents in each position. All of those will be an equilibrium since agents are ex-ante identical. In Section 5, we tackle this source of multiplicity by assuming agents are ex-ante heterogeneous. However, there exists another source of multiplicity because more than one information structure may arise in equilibrium. This is a direct result of agents’ conformity concerns. For instance, if all players have very high conformity standards, then they like to choose an action close to the average action—almost ignoring the true state of the world. The importance of coordinating on which information to obtain increases. If all players are looking at one player this will be an equilibrium, but if all were looking at two players it would also be an equilibrium, and so on.

In this subsection, we present all equilibrium networks for a set of parameters and do comparative statics by varying all parameters of the model. We focus on linear cost functions, $C(K_i) = cK_i$, thus any equilibrium information structure is a core-periphery. A particular model is characterized by four parameters: the number of players, n , the level of conformity, r , the precision of each agent’s individual signal, σ , and the cost of forming links, c .

The restriction of the equilibrium set provided by Theorem 2 simplifies the equilibrium computation in two ways. First, only networks that are core-periphery are candidates as equilibrium information structures. Without any restriction on the set of possible equilibria, one would have to check whether each of $2^{(n-1)n}$ different networks was an equilibrium. The restriction imposed by Theorem 2 restricts this number to $2n$ networks, being $n + 1$ core-periphery networks and $n - 1$ core-periphery networks with core members looking down. One information structure differs from the other only regarding how many players are part of the core.

Second, to check whether a particular information structure is an equilibrium is simplified given the symmetry of the problem, the monotonicity of the benefit of tapping into a link, and the

core-periphery structure. Without any restriction on the problem, one would have to check $2^{n-1} - 1$ possible deviations for each agent, a total of $n(2^{n-1} - 1)$ equilibrium conditions. Given that the benefit of tapping into a signal is monotonic in the number of players tapping into that signal, for each agent we need to check at most two deviations: an agent would like to form a link with the most influential agent she is not observing or she would like to break a link with the least influential she is observing. Furthermore, the symmetry of the problem tells us that we only need to check for the incentives of two players: one representative core player and one representative periphery player.¹⁸ However, we can further reduce this number to only two deviations per information structure. For core-periphery structures in which a core player does not tap into the signal of any peripheral player, we have that the peripheral player has more information than a core player. Thus we only have to check for the incentive that a core player has for forming a link with a peripheral player and for the incentives that a peripheral player has for dropping the link formed to a core player. If the incentives for these two possible deviations make them unprofitable, then the information structure is an equilibrium. In core-periphery information structures with core players tapping into the signals of a peripheral player, the core players have the same amount of information as any peripheral player. Thus, we only have to check for the incentive that a core player has for dropping a link with the special peripheral player, s , and for the incentives that a peripheral player has for forming a link with the special peripheral player, s .

Parametrization First we characterize all equilibrium networks for one particular parametric specification. We consider a model with 5 agents and equal precision of the private signal and the common prior: the variance of the private signal is equal to 1. Each agent chooses an action that is an average of the true state of the world and the average action, since the conformity is $r = 0.5$. Finally, the cost of forming links is set to be $c(K_i) = 0.03K_i$.

Figure 9 represents the four different information structures that are equilibrium. From left to right the number of players in the core increases from 2 players to 5. In the first equilibrium the three members of the periphery observe the two members of the core, who also tap into each others' signals. In the last equilibrium, all agents tap into the signals of all other agents in the economy. This figure displays an example of the multiplicity of equilibrium structures, while ignoring that for each structure there are multiple agent arrangements. For instance, a core-periphery network with two members in the core and three in the periphery can be arranged in different ways (with players 1 and 2 or with players 3 and 4 in the core, etc). These are different networks and will have different payoff implications for different agents. We abstract from these considerations and focus

¹⁸In a core-periphery equilibrium with players looking down, we would in principle have to check for the incentives of three representative players: the core, the periphery and the s player. However, note that the s player has the same incentive to not look at a peripheral player as the central player has. And the s player has more incentives to not drop the link with another central player, than a central player has.

only on the multiplicity of structures.

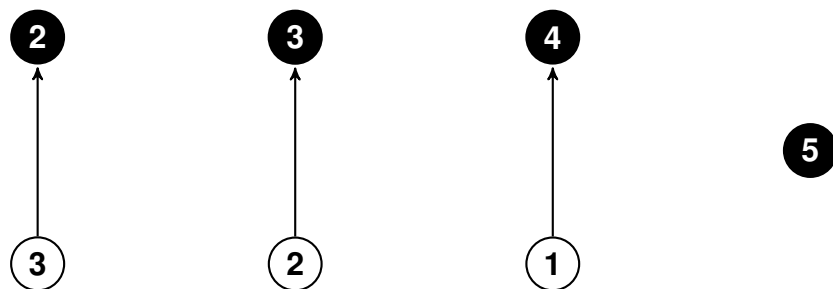


Figure 9: Four equilibrium information structures

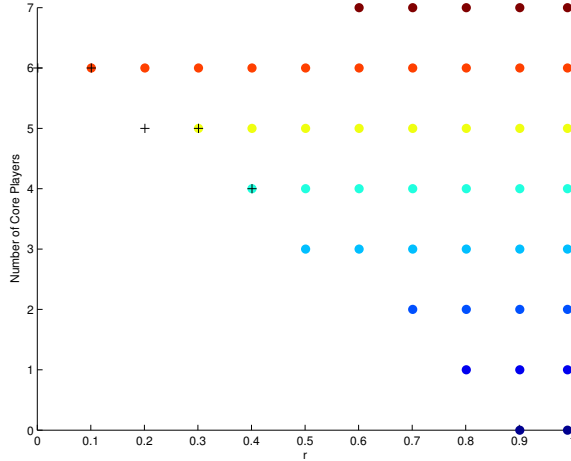
3.4.1 Comparative Statics

Conformity We focus on how changes in the level of conformity affect the set of equilibria. We allow the weight of the average action on the bliss action to vary, but keep the other parameters fixed: there are 10 players in the economy and the precision of each private signal is the same as the precision of the common prior, 1. To tap into another player's signal has an additional cost of 0.016, independent of how many signals a player has tapped into.

Figure 10 shows the set of equilibria as conformity, r , varies from zero to one. Each point in the graph represents a different equilibrium with the number of core players plotted against the y -axis. Since there are two types of equilibrium with the same number of core players, we represent equilibrium of type A as a filled dot and of type B as a plus sign.

If players care little about matching the average action (low r), the set of equilibria is quite small. Agents balance how much information they want to acquire in order to predict the true state and its cost. At the parameters above, this balance is achieved when each player is tapping into 6 other signals, forming a core-periphery network with 6 agents in the core and 4 in the periphery and all agents in the core tapping into the signal of 1 agent in the periphery. It is always true that for low enough levels of conformity, there exists an equilibrium of type *B*.

As the level of conformity increases, agents are more concerned with matching what other agents are doing and the set of equilibria increases. To know what other players know becomes more valuable and as a result information structures that involve more (or less) than 6 players at the core become equilibrium.



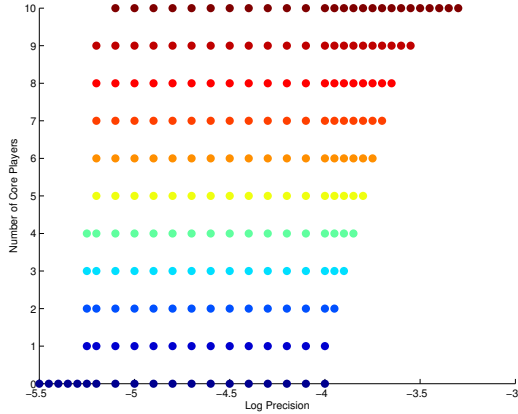
This picture reports the set of equilibria by varying r from zero to one (x-axis). Each point in the graph represents a different equilibrium, with the number of core players plotted against the y-axis. Equilibrium of type A is represented with a filled dot, while of type B with a plus sign.

Figure 10: As conformity increases, the set of equilibria expands.

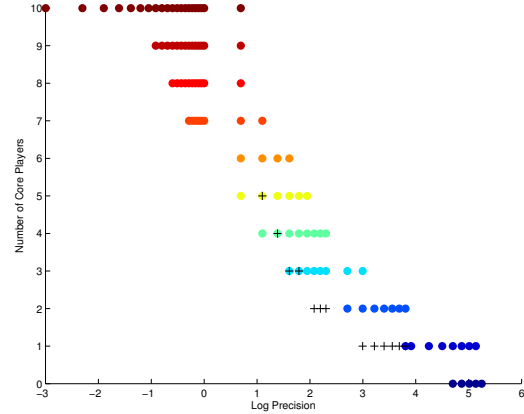
Precision We now focus on how changes in the precision of private signals affect the set of equilibria. Figure 11 plots the set of equilibria, for different values of precision of the private signal. As the precision varies the number of players is fixed at 10, and the cost of forming a link is 0.005. Agents care equally about the average action and the true state, and conformity is set at 0.5.

We can see that if the precision is close to zero, any private signal is very uninformative and thus not worth its cost. Even if all other agents were focusing on one agent and tapping one signal, it would not be worthwhile for an agent to do the same, and the empty network is the unique equilibrium. Panel (a) of Figure 11 shows that as precision increases a little, it is worthy it to look at a player only if all others do as well, and the set of possible equilibria expands, including larger and larger cores. As precision continues to increase (Panel a), to tap into another agent's signal becomes profitable, even if no one else is doing the same. Thus the set of equilibria shrinks monotonically, with the empty network being the first information structure not to be an equilibrium anymore, followed by the star, and so on. If the precision of the private signal is close to e^{-3} , the full network is the unique equilibrium.

As precision continues to increase, diminishing marginal returns to information implies that the number of agents in the core diminishes. And if the precision is high enough, an agent is already sufficiently well informed by her own private signal that connecting to another agent is not profitable enough. Panel b shows that the set of equilibria shrinks again as signals become sufficiently precise.



(a) If the precision is small, as precision increases the number of agents in the core increases.

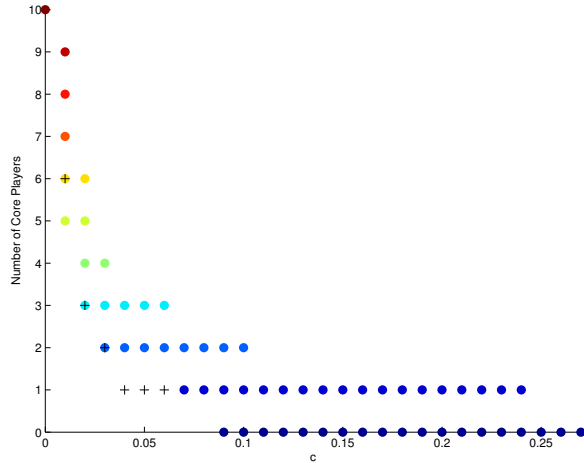


(b) If the precision is larger, as precision increases the number of agents in the core diminishes.

This picture reports the set of equilibria by varying σ (x-axis). Panel (a) reports the set of equilibria of low values of σ , while Panel (b) uses higher values for σ . Each point in the graph represents a different equilibrium with the number of core players plotted against the y-axis. Since there are two types of equilibrium with the same number of core players, we represent equilibrium of type A as a filled dot and of type B as a plus sign.

Figure 11: If the precision is small, the empty network is the unique equilibrium. As precision increases, the set of equilibria monotonically increases, incorporating larger core sizes. As it further increases, the set of possible equilibria monotonically shrinks and the full information structure is the only equilibrium. Increasing the precision even more, the number of agents in the core diminishes.

Cost of Acquiring Information As we decrease the cost of acquiring information, keeping the number of agents fixed at 10, the precision of the private signal fixed at 1 and the conformity at 0.5, the number of players in the core increases (see Figure 12). We can interpret a decrease in the cost of tapping into a signal as an improvement in communication technology. An improvement in communication technology implies a more horizontal society, in which players coordinate on a larger number of other players' signals. If acquiring information is more costly, players will organize themselves in a more vertical information structure, with a very narrow focus. Recent technological changes affect the cost of tapping into a signal, changing the information structure. For instance, internet has made reading the reviews of other movie critics easier, since now all information can be found on-line; while recording devices have made the access to different news television shows easier and more convenient.

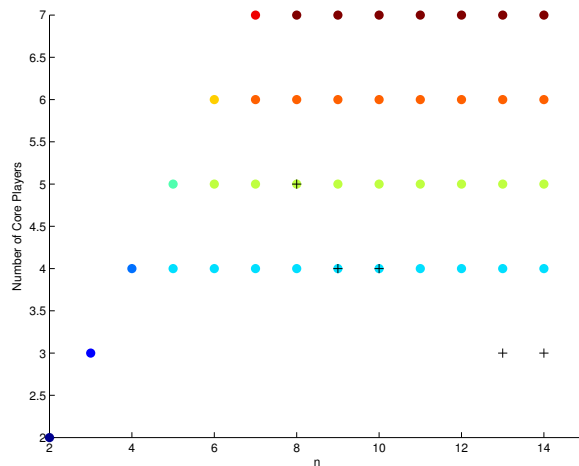


This picture reports the set of equilibria by varying c (x-axis). Each point in the graph represents a different equilibrium with the number of core players plotted against the y-axis. Since there are two types of equilibrium with the same number of core players, we represent equilibrium of type A as a filled dot and of type B as a plus sign.

Figure 12: As the cost of tapping into an agent signal increases, the equilibrium number of signals tapped into by a player diminishes monotonically.

Number of Agents Finally, we focus on how the number of agents in the economy affects the equilibrium set. As we vary the number of agents, we consider the precision of the private signal to be fixed at 1; the conformity of an agent at 0.5; and the cost of forming a connection at 0.015.

Figure 13 shows that as the number of agents in the economy increases, the number of agents in the core increases, and the set of equilibria expands.



This picture reports the set of equilibria by varying n (x-axis). Each point in the graph represents a different equilibrium with the number of core players plotted against the y-axis. Since there are two types of equilibrium with the same number of core players, we represent equilibrium of type A as a filled dot and of type B as a plus sign.

Figure 13: As the number of agents in the economy increases, the set of possible equilibria expands and the number of agents in the core increases.

4 Aggregate Volatility

One implication of hierarchical network structures is that information aggregates inefficiently. The signals of players in the core of the network, namely opinion makers, affect actions and payoffs of all other players more than the signal of players in the outer layers of the hierarchy. Even though there are several connections in the network and more information available in the economy, agents rely on the core members' signals when choosing actions. All the players are learning from the same source, and the coordination in attention translates into coordination in actions among agents.

Action coordination is an endogenous outcome of agents' optimization, and results in more volatility of the average action if compared to other network structures. If a player receives a realized signal that is incorrect, then the misinformation propagates through the network. More importantly, the information propagates differently in the economy depending on the source. In a hierarchical network, the average action heavily depends on the signals from core players which generates more average action volatility relative to an economy in which the average action equally depends on players' signals.

Therefore, our model endogenously generates excess average action volatility in a environment of network formation to acquire information. For example, consider a game with ten players, and let's compare how different informational structures affect the spreading of an informational shock into the average action. First, consider a core-periphery network with one agent in the core and the other nine players in the periphery. Second, consider a small-world network in which players are connected to a neighbor forming an incomplete wheel: player one looks at player two, player two looks at player three, and so on. The last agent (player ten) does not look at anyone, so that both networks have exactly the same number of links. Additionally, let's assume $r = 0.5$ and $\sigma = 1$.

The average action is a linear combination of the common prior about the state of the economy and all signals available in the economy:

$$\bar{a} = \sum_{i=0}^n \beta_i e_i,$$

where e_i is the signal from player i , β_i is the equilibrium influence of signal e_i and $e_0 = 0$ is the common prior. In the first network, the average action heavily depends on the prior about the state of the economy and on the signal of the core players, with both betas equal to 0.41. In this case the unconditional standard deviation of the average action is 0.72. In the second network, the average action depends more on the common prior with a beta of roughly 0.5, but all players' signals have very low impact on the average action, close to 0.05. Moreover, the second network features an average action volatility of only 0.52. Although both networks have exactly the same number of connections and thus the same magnitude of informational flow, the hierarchical network (first

network) has nearly 40% more volatility in the average action. This is what we call a line network. Therefore, the hierarchical network structures generate volatility in the average action and players are more likely to be further away from the true state of the economy.

Although this is a relatively simple example with arbitrary parameters, it stresses how the network structure has significant implications for volatility. To further illustrate this finding, we present a calibration of the model using Institutional Brokers' Estimate System (IBES) data, which provide several analysts' estimates of different variables of publicly listed firms. We focus on estimates of one-year-ahead earnings per share from 1981 to 2014.¹⁹

To match the IBES data into our model, we assume that the action taken by the agents (a_i) is the analyst i 's estimate of the earnings per share, and the realized value of earnings per share is the state of the economy (θ). In our model, analysts acquire information from each other—forming a network of informational flow. As discussed in the Introduction, we interpret analyst information acquisition decisions within our framework: an analyst acquire signals in order to forecast earnings-per-share of a given company. Each analyst receives one signal for free (for instance, calling the C.F.O. of the company), and acquires other signals at a cost. Furthermore, as forecast analysts care about having their estimates close to both the correct value and the other analysts' estimates (Hong and Kubik (2003); Horton et al. (2015); Hong et al. (2000); Nolte et al. (2014)), our payoff specification is a good representation of analysts' objective function.

We assume the link formation cost function is linear with marginal cost c . This cost function implies that every equilibrium has core-periphery network structure (see Theorem 2).²⁰ Although the information network itself is not observed, we assume that senior analysts are in the core of the network, and we identify core analysts using the frequency of forecast revision. According to recent studies on analyst's performance and incentives, more senior and experienced analysts revise their forecast less frequently than junior analysts. Hong et al. (2000) documents that the frequency of forecast revision decreases with experience, while Bradley et al. (2013) documents that a forecasters influence increases with experience.²¹ For each company and each year, we define as core analysts those who update their estimates the least. In our sample, firms have on average 26.4 analysts from which 5.6 are core analysts.²²

¹⁹We remove the top and bottom one percent to remove outliers, and we only use companies with at least 20 analysts, which reduces the number of analyzed firms to 655 different companies.

²⁰Other papers have suggested similar information structures, focusing on herding behavior. For instance, see Cooper et al. (2001) and Stickel (1990) for different approaches.

²¹Clement and Tse (2003) documents that the frequency of revision decreases the influence of the revision.

²²To further validate our choice of core analysts, we test whether the estimate from peripheral analysts depend on core players' forecast. Our theoretical model suggests that the peripheral analyst forecast positively depends on core analysts's forecast. For each company, we regress average forecast of peripheral analysts on the average forecast of core analysts, controlling for the realized earnings per share. In 94% of the regressions, the coefficient is positive, and 73% of all regressions have t-statistic greater than two. Alternatively, for each company, we can run a pooled regression of peripheral analyst forecast on the average forecast of core analysts, controlling for the realized earnings

The average action in the model is a linear combination of the signals in the economy, while each signal is equal to the true state of the world, plus the error term:

$$\bar{a} = \sum_{j=0}^n \beta_j e_j = \sum_{j=0}^n \beta_j \theta + \sum_{j=0}^n \beta_j \epsilon_j = \theta + \sum_{j=0}^n \beta_j \epsilon_j,$$

where the last equality follows from the fact that $\sum_{j=0}^n \beta_j = 1$. Also, the error terms ϵ are independently distributed following a normal distribution centered around zero.

Because the error terms are independent of the true state of the world, we write the variance of the average action as a sum of variances. Furthermore, we interpret the variance of the prior as the variance after all information contained in ex-ante public announcements have been accounted for. Assuming that the public and private signals are equally informative—with variance equal to σ^2 —we can write the variance of the average action as:

$$V(\bar{a}) = \sigma^2 + \sum_{j=0}^n \beta_j^2 \sigma^2 = \sigma^2 \left[1 + \sum_{j=0}^n \beta_j^2 \right] = \sigma^2 \left[1 + HHI(\beta_j) \right],$$

where $HHI(\beta_j)$ is the Herfindahl-Hirschman index, measuring influence concentration.

The equilibrium variance of the average action is directly related to how concentrated the influence over the average action is. If influence is very concentrated—all agents rely on the information of a few—the index will be higher, and in the extreme case of all influence on one signal it will be equal to 1.

If all agents know the true state, the average action would always be equal to it, as seen in Proposition 1. This implies that the fundamental variance of the average action is given by:

$$V^F(\bar{a}) = \sigma^2.$$

While the fundamental variance assumes that all agents know the true state of the world, the equilibrium variance considers that agents receive the endogenously chosen imprecise signals of it. This introduces two distinct forces, information and correlation of information, both increasing the equilibrium average action variance. To disentangle these two effects, we compute the informational variance of the average action by assuming that agents have only a private signal to rely on—no common prior or acquisition of information. We compensate the precision of the private signal such that each individual is equally informed regarding the true state of the world as they were under the core-periphery equilibrium structure.

We use the information structure calibrated from the data, with 26 agents and 6 on the core, per share. In this case, 93% of the companies have a positive coefficient and 83% have t-statistic greater than two.

Table 1: Average action variance as a function of information structures

The first column show the complementarity parameter used. The second column shows the fundamental variance of the average action, $V^F(\bar{a})$. The third column shows the variance of the average action once imperfect information is accounted for, $V^I(\bar{a})$. The fourth column shows the calibrated information structure average action variance, $V^E(\bar{a})$. The last column reports the percentage increase in variance from the second to the fourth column. We use 26 players, with 6 core members. Also, we normalize the fundamental variance to be 100.

r	$V(\bar{a})^F$	$V(\bar{a})^I$	$V(\bar{a})^E$	<i>Increase</i>
0.01	100.00	100.50	111.74	11.74
0.10	100.00	100.50	111.94	11.94
0.25	100.00	100.50	112.28	12.28
0.50	100.00	100.50	112.89	12.89
0.75	100.00	100.50	113.56	13.56
0.90	100.00	100.50	113.99	13.99
0.99	100.00	100.50	114.26	14.26

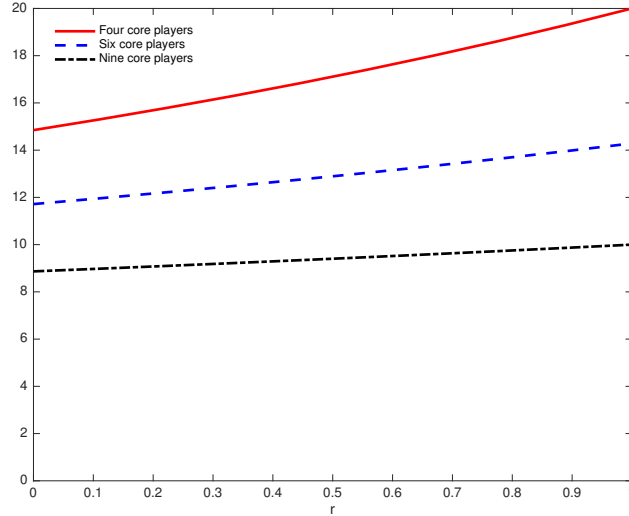
to compute the variance of the average action implied by the equilibrium influence. Table 1 reports the average action variance for different complementarity levels (r), under the equilibrium core-periphery network, $V^E(\bar{a})$. We also report the informational variance $V^I(\bar{a})$ and the fundamental variance, $V^F(\bar{a})$, which is normalized to 100. Finally, the last column reports the increase in variance due to the core-periphery information structure.

For almost no complementarity ($r = 0.01$), the core-periphery structure increases average action variance by almost 12%. As agents care more about the average action (higher r), then the equilibrium network structure increases the average action variance even more. For example if $r = 0.75$ variance increases almost 14%. Furthermore, imperfect information increases the average action variance by a negligible 0.5%. Therefore, a significant fraction of the average action variance is explained by the informational structure endogenously generated as the outcome of the network formation process. Furthermore, the informational structure affects the average action volatility through the endogenous correlation of information, not through the imperfectness of it.

The implication of this example is straightforward: Concentration of influence increases the average action variance, making average market predictions more likely to be wrong. Furthermore, such increase is not due to the lack of information. While this result relies on the core-periphery equilibrium structure, we show that it does not depend on the particular information structure calibrated from the data. Figure 14 plots the increase in average action variance due to different core-periphery information structures when compared to the fundamental variance. It plots the increase in variance as a function of the complementarity parameter, r . For a network with 26 players, and 4 core players, average action variance increases between 15% and 20%. While for a network with the same number of players and 9 core players, it increases between 9% and 10%.

Figure 14: Robustness: Average action variance in different core-periphery networks

The solid line shows the increase in the variance of the average action, compared to the fundamental variance, when there are four players in the core. The increase in variance is plotted as a function of agent’s conformity level, r . The dashed line plots the increase in variance for a network with six players in the core. Finally, the dotted-dashed line shows the increase in variance for a network with nine players in the core. For all networks, we consider 26 players.



5 Origins of Leadership

Our central result in the previous sections is that players coordinate on which players’ signals to watch, focusing on a subset of the signals available. These central players are opinion makers, influencing the actions of other players. A natural interpretation is to have core players as *leaders* and periphery players as *followers*. Although leadership has been defined in many ways,²³ a common element is that the leaders exert a special influence over the others.²⁴ Thus one of the primary functions of a leader is to serve as a coordinator of other players.²⁵

In Dewan and Myatt (2008), followers would like to coordinate their actions, but cannot because each follower is unsure about what the others believe. The leader helps to coordinate their actions, by providing information. Our model expands this idea, by making the agent’s role of leader or follower endogenous. In our model, the leader serves as a focus point, however not through her actions or example, as in Hermalin (1998), and Komai et al. (2007). It is only by providing information to other players that the leader helps to coordinate their actions. In Section 4, the core player’s signals become effective public information. It is common knowledge in equilibrium that everyone in the economy knows such signals, and thus all players are influenced by

²³For instance, see Hermalin (2012), and Glynn and DeJordy (2010) for a variety of definitions.

²⁴Gardner and Laskin (2011), and Bass and Stogdill (1990) are good examples.

²⁵See Hermalin (2012) for a discussion and literature review.

them. Thus, in equilibrium a leader coordinates players' higher order beliefs. A leader may be interpreted as an endogenous coordinating "rational ritual", as proposed by Chwe (2013).

The model discussed so far shows the endogenous emergence of leadership, given that agents are ex-ante identical. However it does not provide insights regarding which agent will perform each role. All agents have the same chance of being leaders, given that any permutation of agents and positions in the information structure is an equilibrium. In this section we abandon the assumption of ex-ante homogeneity in order to focus on which individual characteristics determine leaders.

Following Dewan and Myatt (2008) and Kaplan et al. (2012), we analyze four individual characteristics, listed below, and compare how conducive to leadership each trait is. In order to introduce heterogeneity in the model, we let agents differ in one trait at a time. Furthermore, we focus on a simplified version of the model, in which agents are restricted to tap into at most one signal. Given that agents are not symmetric anymore, the concept of leadership becomes more complicated. It is possible that some players tap into a signal while others choose not to do so.

Definition 5. *We say an economic environment produces a **Leader** if there exists at least one equilibrium in which all players focus their attention on one player, by all tapping into her signal.*

Communication ability How good a communicator is each player? As discussed in Dewan and Myatt (2008), different agents have different abilities to communicate, and this may be determinant to their quality as leaders. For instance, a political partisan may be better at communicating her view, a financial analyst may write clearer recommendations, and a movie critic may write a movie review that is funny and interesting in itself, leading to a smaller cost to be borne by the reader.

In our model this translates to the cost of tapping into signals being a function not only of how many links a player is forming but also of to whom are the links being formed. Given that each agent forms at most one connection, we formalize such heterogeneity as each agent having an individual specific vector of costs of acquiring information, restricted to the fact that such cost depends only on the source of information. That is, an individual i has a vector of costs of tapping into each of the other players' signals $[c_{i,1}, \dots, c_{i,n}]$ such that the cost of player i tapping into player j 's signal is $c_{i,j} = c_j$. That the cost depends only on the source of the information implies that c_j is also the cost of player s tapping into player j 's signal. Agents can have different costs of having their signals being tapped into, with better communicator implying a smaller cost to be born by the listener.

Listening ability How good a listener/reader is each player? For instance, a particular political partisan may have different costs of watching other TV news shows and thus acquiring different information. In our model this translates to the cost of tapping into a signal being different for each agent. The cost of player i tapping into player j 's signal is $c_{i,j} = c_i$, which is the same as the cost

of player i tapping into player s 's signal. However, if a worse listener were to tap into any of those signals, she would incur a higher cost.

Precision How well informed is each player? Different agents may receive signals of different qualities. For instance, a financial analyst may have better information regarding a certain stock or company, or a particular movie critic may be more apt in determining a movie's quality. This is related to what Kaplan et al. (2012) and Gabaix and Landier (2008) defined as individual talent. In our model this can be simply expressed as the individual variance of the private signal being agent specific, σ_i . A better informed agent has a more precise private signal.

Resoluteness How much does an agent conform to other players in the economy? All agents in our model balance adaptation and coordination motives, however different agents may balance those same motives differently. A particular financial analyst may be more resolute, regarding conformity as less important, depending on her career concerns. The relationship between resoluteness and leadership—focusing on the importance of the CEO—has been studied at length, see Bolton et al. (2013) for a review. Resoluteness translates, in our model, to different agents having different conformity parameters, r_i . We interpret an agent with lower conformity parameter to be a more resolute one.

5.1 Analysis and Interpretation

Consider an economy populated by n players, each characterized by an individual level of conformity r_i , an individual precision of her private signal σ_i , and an individual cost vector of acquiring information $[c_{i,1}, \dots, c_{i,n}]$. An economy is parametrized by a number of players, n , and three distributions: (i) distribution \mathbb{R} , of individual levels of conformity, (ii) distribution \mathbb{S} , of individual precision, and (iii) distribution \mathbb{C} , of cost vectors. Let us proceed by considering agents to be heterogeneous in only one dimension at a time.

Proposition 6. *Leadership as a result of individual traits:*

1. *Consider an economy in which all n individuals have the same level of conformity, the same precision of their private signal, the same listening ability but differ in their communication ability. For each parametrization such that the environment produces a leader, there exists an equilibrium in which the good communicator is the leader. Furthermore, there exist parametrizations such that leadership occurs only if the good communicator is the leader.*
2. *Consider an economy in which all n individuals have the same level of conformity, precision, communication ability, but differ in their listening ability. There exist parametrizations*

such that the environment produces a leader, but the good listener is not the leader in any equilibrium.

3. *Consider an economy in which all n individuals have the same level of precision, communication ability, listening ability, but differ in their level of conformity. For each parametrizations such that the environment produces a leader, there exists an equilibrium in which the more resolute player is the leader. Furthermore, there exist parametrizations such that leadership occurs only if the more resolute player is the leader.*
4. *Consider an economy in which all n individuals have the same level of conformity, communication ability, listening ability, but differ in the precision of their private signal. There exist parametrizations such that the environment produces a leader, but the better informed agent is not the leader in any equilibrium.*

The proposition above is proven in the Appendix. If we interpret each possible parametrization as a possible situation in the economy, the first part of the proposition states that all situations that are conducive to the emergence of a leader, are also conducive to the emergence of the best communicator as the leader. It is possible that for a certain situation, everybody choosing to tap into the signal of the worst communicator is an equilibrium. However, for this same situation it will also be an equilibrium for all players to tap into the signal of the good communicator. Furthermore, there are situations in which the only equilibrium with a leader is the one where all players focus on the good communicator. This result is aligned with empirical findings, and is intuitive in the model.²⁶ The benefit of connecting to a player is additively separable from the cost of forming such connection. Thus the benefit of tapping into any signal (given that all other players are doing it) is always the same, and if it is worth paying a cost for such benefit it is also worth paying the smallest cost, thus tapping into the good communicator's signal.

The second part of the proposition states that there are situations in which leadership occurs, but everybody focusing on the good listener agent is not an equilibrium. The intuition behind this result can be understood in a three player example. The crucial point lies in that the benefit of tapping into a signal is independent of the cost of doing so. Suppose one agent has a very high cost of acquiring signals (bad listener), while both other agents have smaller costs (good listeners). Note that the benefit of tapping into a signal (if another player is doing the same) is independent of which signal is being tapped, and suppose that such benefit is slightly larger than the cost of tapping into a signal for the good listener. This implies that both good listeners focusing on the bad listener signal is an equilibrium. However, given that the cost for the bad listener to tap into a signal is significantly higher than the benefit, he will never do so. Thus there will not be an

²⁶See Dewan and Myatt (2008), for a discussion.

equilibrium in which a good listener emerges as a leader. This result contradicts the conventional wisdom emphasizing empathy and the “team player” qualities of leaders, however it finds support in the findings of Kaplan et al. (2012) and Bolton et al. (2013), that “CEO’s who stick to their guns tend to be better leaders than good listeners.”

The third part of the proposition states that in any situation conducive to the emergence of leadership, there exists an equilibrium where the more resolute player is the leader. Furthermore, there are situations such that in the unique equilibrium with leadership, the resolute player is the leader. The intuition behind this result is that a more resolute player takes her private information more into consideration than a less resolute one; thus in order to predict how the more resolute player will act, knowing her private information is important. This implies that the benefit of tapping into the more resolute player (if all other players are also doing so) is higher than of tapping into a less resolute one.

The importance of resoluteness and overconfidence²⁷ for leadership is widely studied, for instance Goel and Thakor (2008), Gervais and Goldstein (2007), Bolton et al. (2013), and Kaplan et al. (2012). While in our model the leader is endogenously determined, Bolton et al. (2013) constructs a model in which an exogenously assigned leader does a better job in leading the more resolute she is. Kaplan et al. (2012) empirically confirms the importance of resoluteness for the quality of leadership. In our model, a more resolute player does provide a better quality of leadership, given that a more resolute player provides a more steady leadership, with the leader’s actions being more predictable, and closer to the true state,

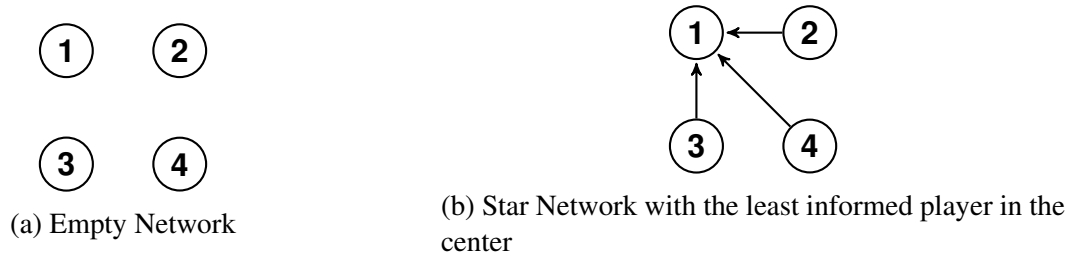
The final part of the proposition above states that there exist situations that are conducive to leadership, but the best informed player is not the leader in any equilibrium. Interpreting the quality of the private signal as the agent’s talent, the proposition argues that there are situations in which all of the equilibria involving leadership exclude the most talented player from being the leader. The intuition behind this result can be understood in the four player example presented below. It relies on the correlation of signals being a function of the precision of those signals. Even though this result may seem counter intuitive, it finds support in other papers that focus on the role of the leader as a coordinator, for instance Andreoni (2006) and Huck and Rey-Biel (2006), or see Hermalin (2012) for a review. Finally, Kaplan et al. (2012) empirically show that a more talented CEO leadership is better for the firm. Although our model shows that the most talented agent may not be the leader, it is true that if the leader is the more talented player agents’ payoffs are higher.

Example 4. Leadership and Talent

First: Consider a baseline model with four agents, variance of the private signal $\sigma^2 = 0.01$,

²⁷Overconfidence refers to a player who overweights some information compared to other players. In our model, a more resolute player overweights the private signals she has observed, reducing the influence of the prior when compared to a less resolute player.

conformity $r = 0.9$, and the cost of forming one link $c(1) = 0.0002$. In this example, the unique equilibrium is the empty network, and no player taps into any other signal (see Panel a). Now, let us introduce heterogeneity by making the first player's private signal worse than the others. Let us increase the variance of her private signal by 50%, to $\sigma_1^2 = 0.015$, while all other parameters remain constant. Two information structures are equilibria, the empty network and the star network with the worse informed agent in the center (Panels a and b).



The key intuition here is that the private signals of players 2, 3 and 4 are highly correlated due to their high precision (note that the variance is quite small). This implies that none of these players is willing to tap into the signal of one of the others (not even if player 1 were doing so), since there is little gain of information. However, the signal of player 1 is less correlated and thus, even though it is less informative about the state of the world, the more informed agents find it optimal to tap into it (given that all others are doing so). They coordinate on the signal of the first player because it is less precise, which makes the action of the first player less predictable.

6 Extensions and Final Remarks

In this paper we characterize all equilibrium information structures of an information acquisition game characterized by action complementarities and by the lack of retransmission of information. We show that all strict Nash equilibria present a hierarchical directed information structure, which is characterized by the existence of different tiers of informational importance. An individual that belongs to the top tier is very influential, as her signal is tapped into by members of all other tiers. A second tier individual's signal is observed by all members of tiers below her, and so on. A hierarchical directed network allows the existence of multiple social levels, ranked by their influence over economic outcomes. This theoretical result does not rely on the cost of forming connections. In fact, our main result is more general than the information acquisition model presented in this paper, as it only depends on Property 1 and 2. Property 1 states that an agent ranks other players' signal by how public they are, that is, according to their in-degree centrality in the network.

Property 2 states that a more central agent benefits less from tapping into another player's signal. Any network formation game satisfying Properties 1 and 2 has a hierarchical directed network as equilibrium outcome.

An example of a hierarchical network is a core-periphery. In this case, players are partitioned in two groups: the core and the periphery. All members of the periphery are connected to all members of the core. All members of the core are connected to each other and members of the periphery are not connected to each other. Members of the core are highly influential, with their private signals being effectively public information, while members of the periphery are not. *If the cost of tapping into an additional signal is weakly increasing in the number of signals tapped into, any equilibrium information structure is a core-periphery network.*

The network formation game discussed in this paper is insightful to describe different real-world situations, which can range from a farmer choosing a particular crop, or a political partisan choosing which policies to support, to sell-side investment analysts. These situations rely on strategic complementarities of players' actions—a farmer gains from choosing a crop similar to his neighbors', while a partisan wants the Party to display unity—as well as on endogenous information transmission. The equilibrium information structure affects economic variables, for instance average action volatility. In the model, as players coordinate on only a few signals those signals become highly influential for the average action—making the average action almost as volatile as only those few signals. In particular, we focus on sell-side analysts and earnings-per-share predictions. We use data from I/B/E/S to calibrate the equilibrium information structure of the model, and we argue that a substantial part of the volatility of the average action can be explained by the core-periphery information structure. When comparing a core-periphery with other networks in which each player is equally informed, we conclude that 12% to 14% of the average action's variance can be explained by the core-periphery network.

In the equilibrium information structure, ex-ante identical agents have different influence levels over the society. As an agent's signal is more tapped into, other agents pay more attention to her information and she becomes more influential. This suggests a leader/follower structure in the society. Since the baseline model has no predictive power regarding who the leaders will be, we extend the model to include ex-ante heterogeneity in four dimensions: (i) communication ability, (ii) listening ability, (iii) resoluteness, and (iv) talent. We conclude that communication ability and resoluteness are good predictors of which player will be the leader, while a player's talent is not.

Beyond those insights, our model can be extended in a variety of directions. In what follows we briefly describe some extensions to our model.

No names What would happen if players could not identify some of the others? A central point of our result is that players coordinate on which signals to tap, however players may not know who

is who to coordinate on. For instance, movie critics may not know the name of other movie critics, and thus don't know on which movie reviews to focus. If players are not able to recognize anyone, the only possible information structure is the random network, and the only choice a player can make regarding information acquisition is how many signals to tap into. A more interesting case arises when the absence of names is only partial: there are some players who all other players know the name of, but those are the only names they all recognize. For instance, movie critics may know the names of Mahola Dargis of the *New York Times* and Roger Ebert of *Chicago Sun Times*, but not of many others. All equilibria will include coordination on this two signals, however if players want to tap into more signals than that, they must randomize among the other signals. This implies a new type of information structure in equilibrium: a network that is a mixture of a core-periphery and a random network.

Private information Maybe a movie critic could go and watch the movie again, giving it a second chance, and obtaining a new independent observation of its quality. An agent could acquire another private signal, not yet observed by any other player in the game. Notice that this does not affect other players' connection incentives, and let us assume that the cost of acquiring such a signal is the same of tapping into another player's signal. The first property implies that no agent chooses to obtain an additional private signal, since it would inform her the same regarding the true state of the world as tapping into another agent's signal, while informing her strictly less regarding the average action. A player would acquire a new private signal only if she was already informed by all other signals in the economy.

Discrete choice model In this paper we assume agents can choose over a continuum of possible actions. However in many complementary settings agents choose to partake or not in one possible action: citizens choose to join or not anti-government protests; agents choose whether or not to buy a new good or adopt a new technology; friends decide whether or not to go to a certain party. Under this new light, we can reinterpret a_i^* —our model's bliss action—as an individually optimal cost cutoff. To adopt a technology, agents have to pay an unknown cost. Agent's choice involves a cost cutoff such that they will adopt the technology only if the individually drawn cost is lower than the cutoff. The individual cost is a random draw from a distribution that depends on the state of the world and on whether other players are joining the technology or not. Given the complementarity, an individual chooses an optimal cost cutoff that is a linear combination of the true state of the world (the quality of the technology) and of the average of the cutoffs chosen by other players (an indicator of other players' willingness to adopt the technology). If the quality of the technology is higher, the individual threshold is higher; and if other players are more willing to adopt the technology, the threshold will be higher. As before, an agent observes neither the true quality of

the technology nor the other players' individual cutoffs. However, she receives a signal regarding the quality of the technology and she can, at a cost, tap into the signals of other players.

Dynamic information acquisition A key simplifying point of our model is that information acquisition decisions happen only once, simultaneously for all agents. We relax this assumption slightly by assuming that agents form connections one link at a time. Consider a discrete time model, in which the game ends after any period with probability $1 - p$. If the game ends, movie critics write their best-of-the-year reviews based on their information set, with the same objective as earlier. At each period, a critic is allowed to, at a certain cost, acquire the review of at most one other critic. Acquiring the review should be understood as printing it and storing it in the critic's briefcase, so that it will be available when the game ends and the critic can go to her isolated cabin to write her best-of-the-year review in peace. Critics do not observe at each period how many times a certain review has been printed nor do they observe how many reviews have been printed. There is no communication between critics. If the probability of the game ending after this period is high enough (a very small p), it must be the case that in any equilibrium all critics will coordinate on the signal of the same movie critic. This implies that if no one was tapping into any signal in the beginning of the period, the period ends with a core-periphery network. Furthermore, even if the game survives the first period, the probability of the game reaching the third period is so small that again we will have a core-periphery at the end of the second period. If the game lasts for enough periods, only an equilibrium from the simultaneous model can be an equilibrium for the dynamic game.

Although we cannot extend our model to contemplate sequential acquisition of information or multiple rounds of information acquisition choices, we believe these are interesting directions to extend the research in the future. Specifically for the sell-side analysts, information acquisition is sequential, with a substantial part of it coming from individual calls to the C.F.O. of the company and from the Q & A subsequent to firm presentations.²⁸ This implies a need to account for sequential arrival of information to the agent, and for how agents revise their strategies following it.

Information re-transmission In many applications, an agent can re-transmit a signal she has tapped into. For instance, in *Twitter* a member can retweet, re-posting someone else's Tweet. An obvious direction to extend this research further is to consider a model that includes costly information re-transmission decisions, in which players may acquire information to rebroadcast it.²⁹ This would drastically change Property 1, since linking to a player that was an aggregator of

²⁸See Brown et al. (2015) for a comprehensive survey of information sources for sell-side analysts focused on earnings per share estimations, based on more than 360 interviews.

²⁹See, for instance, Cha et al. (2010) for a study regarding the relevance of retweet for influence.

all the information would be very desirable, independently of how many players were linking to her.

Consider a model in which by linking to a player I observe not only her signal, but also all the signals she has observed, and all the signals the players she is observing have observed and so on... That is, a model in which information is perfectly and costlessly retransmitted. Information flows in only one direction, and the cost of a link is still borne unilaterally. The equilibrium, if the cost is convex, is a wheel or a line network. This result follows from the following argument. First, the resulting network is always connected; if there were 2 disjoint components a player in the less populous one would rather switch and look to the most informative guy in the more populous one. Second, no player ever receives two links. Third, given weakly-convex costs of forming links no player would ever establish two links. Fourth, the network will be minimally connected, thus a wheel or a line network.

Continuous information acquisition A final direction to extend our research concerns the information acquisition discreteness. In our model, we assume that a player can tap or not tap into another player's signal; there is no gradual or partial acquisition of another player's information. In the Appendix we provide a model where a player chooses how much attention to allocate to each player's signal. We consider a fixed endowment of attention that the agent freely allocates among the signals in the economy—other than her private one. We show that all players equally dividing their attention among all signals is not an equilibrium. That is, our central result that agents coordinate and focus on a sub-group of the agents is not a direct result of the discreteness involved in acquiring information.

References

- Acemoglu, D., Carvalho, V.M., Ozdaglar, A., Tahbaz-Salehi, A., 2012. The network origins of aggregate fluctuations. *Econometrica* 80, 1977–2016.
- Allen, F., Morris, S., Shin, H.S., 2006. Beauty contests and iterated expectations in asset markets. *Review of Financial Studies* 19, 719–752.
- Andreoni, J., 2006. Leadership giving in charitable fund-raising. *Journal of Public Economic Theory* 8, 1–22.
- Angeletos, G.M., Pavan, A., 2004. Transparency of information and coordination in economies with investment complementarities. *The American Economic Review* 94, 91.
- Angeletos, G.M., Pavan, A., 2007. Efficient use of information and social value of information. *Econometrica* 75, 1103–1142.
- Azrieli, Y., Lehrer, E., 2008. The value of a stochastic information structure. *Games and Economic Behavior* 63, 679–693.
- Baetz, O., 2015. Social activity and network formation. *Theoretical Economics* 10, 315–340.
- Bakshy, E., Hofman, J.M., Mason, W.A., Watts, D.J., 2011. Everyone’s an influencer: quantifying influence on twitter, in: *Proceedings of the fourth ACM international conference on Web search and data mining*, ACM. pp. 65–74.
- Bala, V., Goyal, S., 2000. A noncooperative model of network formation. *Econometrica* 68, 1181–1229.
- Bandiera, O., Rasul, I., 2006. Social networks and technology adoption in northern mozambique. *The Economic Journal* 116, 869–902.
- Bass, B.M., Stogdill, R.M., 1990. *Bass & Stogdill’s handbook of leadership: Theory, research, and managerial applications*. Simon and Schuster.
- Bergemann, D., Morris, S., 2013a. The comparison of information structures in games: Bayes correlated equilibrium and individual sufficiency. *Economic Theory Center Working Paper* .
- Bergemann, D., Morris, S., 2013b. Robust predictions in games with incomplete information. *Econometrica* 81, 1251–1308.
- Bolton, P., Brunnermeier, M.K., Veldkamp, L., 2013. Leadership, coordination, and corporate culture. *The Review of Economic Studies* 80, 512–537.

- Bradley, D., Gokkaya, S., Liu, X., 2013. Does knowledge matter? Industry experience and analyst performance. Technical Report. Working Paper, University of South Florida.
- Brown, L.D., Call, A.C., Clement, M.B., Sharp, N.Y., 2015. Inside the black box of sell-side financial analysts. *Journal of Accounting Research* 53, 1–47.
- Calvó-Armengol, A., 2004. Job contact networks. *Journal of economic Theory* 115, 191–206.
- Calvó-Armengol, A., de Martí, J., Prat, A., 2015. Communication and influence. *Theoretical Economics* 10, 649–690.
- Calvó-Armengol, A., Zenou, Y., 2004. Social networks and crime decisions: The role of social structure in facilitating delinquent behavior. *International Economic Review* 45, 939–958.
- Centola, D., 2010. The spread of behavior in an online social network experiment. *Science* 329, 1194–1197.
- Centola, D., Baronchelli, A., 2015. The spontaneous emergence of conventions: An experimental study of cultural evolution. *Proceedings of the National Academy of Sciences* 112, 1989–1994.
- Cha, M., Haddadi, H., Benevenuto, F., Gummadi, P.K., 2010. Measuring user influence in twitter: The million follower fallacy. *ICWSM* 10, 30.
- Chwe, M.S.Y., 2013. *Rational ritual: Culture, coordination, and common knowledge*. Princeton University Press.
- Clement, M.B., Tse, S.Y., 2003. Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters? *The Accounting Review* 78, 227–249.
- Colombo, L., Femminis, G., Pavan, A., 2014. Information acquisition and welfare. *The Review of Economic Studies* .
- Conley, T.G., Udry, C.R., 2010. Learning about a new technology: Pineapple in Ghana. *The American Economic Review* , 35–69.
- Cooper, R.A., Day, T.E., Lewis, C.M., 2001. Following the leader: a study of individual analysts earnings forecasts. *Journal of Financial Economics* 61, 383–416.
- Cross, R.L., Parker, A., 2004. *The hidden power of social networks: Understanding how work really gets done in organizations*. Harvard Business Press.
- De Weerd, J., Dercon, S., 2006. Risk-sharing networks and insurance against illness. *Journal of Development Economics* 81, 337–356.

- Dessein, W., Galeotti, A., Santos, T., 2015. Rational inattention and organizational focus .
- Dewan, T., Myatt, D.P., 2008. The qualities of leadership: Direction, communication, and obfuscation. *American Political Science Review* 102, 351–368.
- Dewan, T., Myatt, D.P., 2012. On the rhetorical strategies of leaders: Speaking clearly, standing back, and stepping down. *Journal of Theoretical Politics* , 0951629811429006.
- Foster, A.D., Rosenzweig, M.R., 1995. Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy* , 1176–1209.
- Gabaix, X., 2011. The granular origins of aggregate fluctuations. *Econometrica* 79, 733–772.
- Gabaix, X., Landier, A., 2008. Why has ceo pay increased so much? *The Quarterly Journal of Economics* 123, 49–100.
- Galeotti, A., Goyal, S., 2010. The law of the few. *The American Economic Review* 100, 1468–1492.
- Gardner, H., Laskin, E., 2011. *Leading minds: An anatomy of leadership*. Basic Books.
- Garicano, L., 2000. Hierarchies and the organization of knowledge in production. *Journal of Political Economy* 108, 874–904.
- Garicano, L., Prat, A., . Organizational economics with cognitive costs. *Advances in Economics and Econometrics: Economic Theory*. Volume I. 49, 342.
- Gervais, S., Goldstein, I., 2007. The positive effects of biased self-perceptions in firms. *Review of Finance* 11, 453–496.
- Gladwell, M., 2006. *The tipping point: How little things can make a big difference*. Little, Brown.
- Glaeser, E.L., Sacerdote, B., Scheinkman, J.A., 1996. Crime and social interactions. *The Quarterly Journal of Economics* 111, 507–548.
- Glynn, M.A., DeJordy, R., 2010. Leadership through an organization behavior lens, in: *Handbook of Leadership Theory and Practice: An HBS Centennial Colloquium on Advancing Leadership*, Harvard Business Review Press. p. 119.
- Goel, A.M., Thakor, A.V., 2008. Overconfidence, ceo selection, and corporate governance. *The Journal of Finance* 63, 2737–2784.

- Goel, S., Watts, D.J., Goldstein, D.G., 2012. The structure of online diffusion networks, in: Proceedings of the 13th ACM conference on electronic commerce, ACM. pp. 623–638.
- Golub, G.H., Van Loan, C.F., 2012. Matrix computations. volume 3. JHU Press.
- Goyal, S., Vega-Redondo, F., 2005. Network formation and social coordination. *Games and Economic Behavior* 50, 178–207.
- Hellwig, C., . Heterogeneous information and the welfare effects of public information disclosures .
- Hellwig, C., Veldkamp, L., 2009. Knowing what others know: Coordination motives in information acquisition. *The Review of Economic Studies* 76, 223–251.
- Hermalin, B.E., 1998. Toward an economic theory of leadership: Leading by example. *American Economic Review* , 1188–1206.
- Hermalin, B.E., 2012. Leadership and corporate culture. *Handbook of organizational economics* , 432–78.
- Hiller, T., 2013. Peer effects in endogenous networks. Available at SSRN 2331810 .
- Hong, H., Kubik, J.D., 2003. Analyzing the analysts: Career concerns and biased earnings forecasts. *Journal of Finance* , 313–351.
- Hong, H., Kubik, J.D., Solomon, A., 2000. Security analysts' career concerns and herding of earnings forecasts. *The Rand journal of economics* , 121–144.
- Horton, J., Serafeim, G., Wu, S., 2015. Career concerns of banking analysts .
- Huck, S., Rey-Biel, P., 2006. Endogenous leadership in teams. *Journal of Institutional and Theoretical Economics JITE* 162, 253–261.
- Jackson, M.O., Wolinsky, A., 1996. A strategic model of social and economic networks. *Journal of economic theory* 71, 44–74.
- Jovanovic, B., 1987. Micro shocks and aggregate risk. *The Quarterly Journal of Economics* 102, 395–409.
- Kaplan, S.N., Klebanov, M.M., Sorensen, M., 2012. Which ceo characteristics and abilities matter? *The Journal of Finance* 67, 973–1007.

- Komai, M., Stegeman, M., Hermalin, B.E., 2007. Leadership and information. *The American economic review* , 944–947.
- Königa, M.D., Tesson, C.J., Zenou, Y., 2014. Nestedness in networks: A theoretical model and some applications. *Theoretical Economics* .
- Kumar, R., Novak, J., Tomkins, A., 2010. Structure and evolution of online social networks, in: *Link mining: models, algorithms, and applications*. Springer, pp. 337–357.
- Lehrer, E., Rosenberg, D., Shmaya, E., 2010. Signaling and mediation in games with common interests. *Games and Economic Behavior* 68, 670–682.
- Magee, J.C., Galinsky, A.D., 2008. Social hierarchy: The self-reinforcing nature of power and status. *The Academy of Management Annals* 2, 351–398.
- Morris, S., Shin, H.S., 2002. Social value of public information. *The American Economic Review* 92, 1521–1534.
- Munshi, K., 2004. Social learning in a heterogeneous population: technology diffusion in the Indian green revolution. *Journal of Development Economics* 73, 185–213.
- Myatt, D.P., Wallace, C., 2012. Endogenous information acquisition in coordination games. *The Review of Economic Studies* 79, 340–374.
- Myatt, D.P., Wallace, C., 2014. Cournot competition and the social value of information. *Journal of Economic Theory* .
- Myatt, D.P., Wallace, C., 2015. Information use and acquisition in price-setting oligopolies .
- Myerson, R.B., 1977. Graphs and cooperation in games. *Mathematics of operations research* 2, 225–229.
- Nolte, I., Nolte, S., Vasios, M., 2014. Sell-side analysts career concerns during banking stresses. *Journal of Banking & Finance* 49, 424–441.
- Pavan, A., 2014. Attention, coordination, and bounded recall. Technical Report.
- Rosenthal, R., Jacobson, L., 1968. Pygmalion in the classroom. *The Urban Review* 3, 16–20.
- Stickel, S.E., 1990. Predicting individual analyst earnings forecasts. *Journal of Accounting Research* , 409–417.
- Suri, S., Watts, D.J., 2011. Cooperation and contagion in web-based, networked public goods experiments. *ACM SIGecom Exchanges* 10, 3–8.

Appendix

A Proof of Proposition 2

Let $\bar{a} \equiv \frac{1}{n} \sum_{j=1}^n a_j$ be the average action, and let $\bar{a}_{-i} \equiv \frac{1}{n-1} \sum_{j \neq i} a_j = \frac{n}{n-1} \bar{a} - \frac{1}{n-1} a_i$ be the average action without agent i . We will verify the following guess:

$$\bar{a} = \sum_{j=0}^n \beta_j e_j$$

From the first order condition, agent i 's optimal action satisfies:

$$\begin{aligned} a_i &= (1 - r_i) \mathbb{E}[\theta | \mathbb{I}_i] + r_i E[\bar{a}_{-i} | \mathbb{I}_i] \\ &= (1 - r_i) \mathbb{E}[\theta | \mathbb{I}_i] + r_i E\left[\frac{n}{n-1} \bar{a} - \frac{1}{n-1} a_i | \mathbb{I}_i\right] \\ &= (1 - \tilde{r}_i) \mathbb{E}[\theta | \mathbb{I}_i] + \tilde{r}_i E[\bar{a} | \mathbb{I}_i] \end{aligned}$$

where $\tilde{r}_i = \frac{r_i n}{r_i + n - 1}$. Using bayes updating, the expected value of the state of the world given agent i 's informational set is given by $\mathbb{E}[\theta | \mathbb{I}_i] = \sum_{j=0}^n \tilde{g}_{ij} e_j \equiv \bar{e}_i$, where $\tilde{g}_{ij} = \frac{g_{ij} \sigma_j^{-2}}{\sum_{s=0}^n g_{is} \sigma_s^{-2}}$, $e_0 = 0$ is the prior's mean and $\sigma_0 = 1$ is the prior standard deviation. The expected value of the average action given i 's information is given by

$$E[\bar{a} | \mathbb{I}_i] = \sum_{j=0}^n \beta_j E[e_j | \mathbb{I}_i] = \sum_{j=0}^n \beta_j g_{ij} e_j + \sum_{j=0}^n \beta_j (1 - g_{ij}) \bar{e}_i$$

Thus, player i 's action is simplified to

$$a_i = (1 - \tilde{r}_i) \bar{e}_i + \tilde{r}_i \sum_{j=0}^n \beta_j g_{ij} e_j + \tilde{r}_i \sum_{j=0}^n \beta_j (1 - g_{ij}) \bar{e}_i$$

In order to verify our initial guess, we take the average over i ,

$$\begin{aligned} \bar{a} &= \frac{1}{n} \sum_{i=1}^n a_i \\ n\bar{a} &= \sum_{i=1}^n \left[(1 - \tilde{r}_i) \bar{e}_i + \tilde{r}_i \sum_{j=0}^n \beta_j g_{ij} e_j + \tilde{r}_i \sum_{j=0}^n \beta_j (1 - g_{ij}) \bar{e}_i \right] \\ &= \sum_{i=1}^n \bar{e}_i - \sum_{i=1}^n \tilde{r}_i \bar{e}_i + \sum_{j=0}^n \beta_j \sum_{i=1}^n \tilde{r}_i g_{ij} e_j + \sum_{j=0}^n \beta_j \sum_{i=1}^n \tilde{r}_i \bar{e}_i - \sum_{j=0}^n \beta_j \sum_{i=1}^n \tilde{r}_i g_{ij} \bar{e}_i \end{aligned}$$

Using matrix notation let

$$\tilde{r} = \begin{bmatrix} \tilde{r}_1 \\ \vdots \\ \tilde{r}_n \end{bmatrix}_{n \times 1}, \quad \bar{e} = \begin{bmatrix} \bar{e}_1 \\ \vdots \\ \bar{e}_n \end{bmatrix}_{n \times 1}, \quad e = \begin{bmatrix} e_0 \\ e_1 \\ \vdots \\ e_n \end{bmatrix}_{n+1 \times 1}, \quad G = \begin{bmatrix} G_{10} & G_{11} & \dots & G_{1n} \\ \vdots & & \ddots & \\ G_{n0} & G_{n1} & \dots & G_{nn} \end{bmatrix}_{n \times n+1}$$

Hence, the sum of all action becomes

$$n\bar{a} = \mathbf{1}'\bar{e} - \tilde{r}'\bar{e} + \beta' \text{diag}(G'\tilde{r})e + \beta' \mathbf{1}'\tilde{r}'\bar{e} - \beta' G' \text{diag}(\tilde{r})\bar{e}$$

where $\mathbf{1}$ is a column vector of ones, and $\text{diag}(\cdot)$ creates a diagonal matrix. Let

$$\tilde{G} = \begin{bmatrix} \tilde{g}_{10} & \tilde{g}_{11} & \dots & \tilde{g}_{1n} \\ \vdots & & \ddots & \\ \tilde{g}_{n0} & \tilde{g}_{n1} & \dots & \tilde{g}_{nn} \end{bmatrix}_{n \times n+1}$$

and we have that $\bar{e} = \tilde{G}e$. The sum of actions becomes

$$\begin{aligned} n\bar{a} &= \mathbf{1}'\tilde{G}e - \tilde{r}'\tilde{G}e + \beta' \text{diag}(G'\tilde{r})e + \beta' \mathbf{1}'\tilde{r}'\tilde{G}e - \beta' G' \text{diag}(\tilde{r})\tilde{G}e \\ &= (\mathbf{1} - \tilde{r})' \tilde{G}e + \beta' \left(\text{diag}(G'\tilde{r}) + \mathbf{1}\tilde{r}'\tilde{G} - G' \text{diag}(\tilde{r})\tilde{G} \right) e \\ \bar{a} &= \frac{1}{n} \left\{ (\mathbf{1} - \tilde{r})' \tilde{G} + \beta' \left(\text{diag}(G'\tilde{r}) + \mathbf{1}\tilde{r}'\tilde{G} - G' \text{diag}(\tilde{r})\tilde{G} \right) \right\} e \end{aligned}$$

Next, we use method of undetermined coefficient to solve for the vector of loadings β based on our initial guess $\bar{a} = \beta'e$.

$$\begin{aligned} \beta' &= \frac{1}{n} \left\{ (\mathbf{1} - \tilde{r})' \tilde{G} + \beta' \left(\text{diag}(G'\tilde{r}) + \mathbf{1}\tilde{r}'\tilde{G} - G' \text{diag}(\tilde{r})\tilde{G} \right) \right\} \\ \beta' &= \frac{1}{n} (\mathbf{1} - \tilde{r})' \tilde{G} \left[\mathbf{I} - \frac{1}{n} \left(\text{diag}(G'\tilde{r}) + \mathbf{1}\tilde{r}'\tilde{G} - G' \text{diag}(\tilde{r})\tilde{G} \right) \right]^{-1} \end{aligned}$$

We can verify that the average action loadings sum to 1. Starting from the equation above and

post-multiplying by a vector of ones.

$$\begin{aligned}
n\beta' \mathbf{1} &= (\mathbf{1} - \tilde{r})' \tilde{G} \mathbf{1} + \beta' (\text{diag}(G' \tilde{r}) \mathbf{1} + \mathbf{1} \tilde{r}' \tilde{G} \mathbf{1} - G' \text{diag}(\tilde{r}) \tilde{G} \mathbf{1}) \\
n\beta' \mathbf{1} &= (\mathbf{1} - \tilde{r})' \mathbf{1} + \beta' (G' \tilde{r} + \mathbf{1} \tilde{r}' \mathbf{1} - G' \text{diag}(\tilde{r}) \mathbf{1}) \\
n\beta' \mathbf{1} &= n - \tilde{r}' \mathbf{1} + \beta' G' \tilde{r} + \beta' \mathbf{1} \tilde{r}' \mathbf{1} - \beta' G' \tilde{r} \\
\beta' \mathbf{1} (n - \tilde{r}' \mathbf{1}) &= n - \tilde{r}' \mathbf{1} \\
\beta' \mathbf{1} &= 1
\end{aligned}$$

The action of each agent in vector notation is given by

$$\begin{aligned}
a &= \bar{e} - \text{diag}(\tilde{r}) \bar{e} + G \text{diag}(\beta) \text{diag}(\tilde{r}) e + \text{diag}(\tilde{r}) \bar{e} - \text{diag}(\beta' G') \text{diag}(\tilde{r}) \bar{e} \\
&= \left[\tilde{G} - \text{diag}(\tilde{r}) \tilde{G} + G \text{diag}(\beta) \text{diag}(\tilde{r}) + \text{diag}(\tilde{r}) \tilde{G} - \text{diag}(\beta' G') \text{diag}(\tilde{r}) \tilde{G} \right] e \\
&= \Lambda e
\end{aligned}$$

where Λ is a $n \times n + 1$ matrix of loadings

$$\Lambda = \tilde{G} - \text{diag}(\tilde{r}) \tilde{G} + G \text{diag}(\beta) \text{diag}(\tilde{r}) + \text{diag}(\tilde{r}) \tilde{G} - \text{diag}(\beta' G') \text{diag}(\tilde{r}) \tilde{G}$$

β s in sum notation:

$$\begin{aligned}
n\beta' &= (\mathbf{1} - \tilde{r})' \tilde{G} + \beta' \text{diag}(G' \tilde{r}) + \beta' \mathbf{1} \tilde{r}' \tilde{G} - \beta' G' \text{diag}(\tilde{r}) \tilde{G} \\
n\beta_j &= \sum_{i=1}^n (1 - \tilde{r}_i) \tilde{g}_{ij} + \beta_j \sum_{i=1}^n \tilde{r}_i g_{ij} + \sum_{i=1}^n \tilde{r}_i \tilde{g}_{ij} - \sum_{i=1}^n \sum_{s=0}^n \beta_s g_{is} \tilde{r}_i \tilde{g}_{ij}
\end{aligned}$$

Given that $\sum_{j=0}^n \beta_j = 1$, we have:

$$n\beta_j = \sum_{i=1}^n \tilde{g}_{ij} + \beta_j \sum_{i=1}^n \tilde{r}_i g_{ij} - \sum_{i=1}^n \sum_{s=0}^n \beta_s g_{is} \tilde{r}_i \tilde{g}_{ij}$$

If $r_i = r$ for every $i = 1, \dots, n$ and $\sigma_j = \sigma$ for every $j = 1, \dots, n$, then

$$\beta_j = \frac{1}{n} \sum_{i=1}^n \frac{g_{ij}}{\sigma^2 + K_i - 1} + \frac{\tilde{r}}{n} \left[\beta_j \bar{K}_j - \sum_{i=1}^n \sum_{s=0}^n \frac{\beta_s g_{is} g_{ij}}{\sigma^2 + K_i - 1} \right]$$

where $\bar{K}_j = \sum_{i=1}^n g_{ij}$ and $K_i = \sum_{j=0}^n g_{ij}$. If $\sigma = 1$, then we have

$$\beta_j = \frac{1}{n} \sum_{i=1}^n \frac{g_{ij}}{K_i} + \frac{\tilde{r}}{n} \left[\beta_j \bar{K}_j - \sum_{i=1}^n \sum_{s=0}^n \frac{\beta_s g_{is} g_{ij}}{K_i} \right]$$

B Payoff Function Derivation

In this subsection we derive the payoff function under the optimal action

The utility function of player i is given by

$$U_i = -(a_i - a_i^*)^2$$

where $a_i^* = (1 - r_i)\theta + r_i\bar{a}_{-i}$ is player i 's bliss action and $\bar{a}_{-i} = \frac{1}{n-1} \sum_{j=1, j \neq i}^n a_j = \sum_{j=1}^n \beta_{-i,j} e_j$ is the average action without i 's own action.³⁰ Notice that player i takes all $\beta_{-i,j}$'s as given.

In order to keep the payoff calculations tractable, let's write the signal structure of game in matrix notation as follows

$$\underbrace{\begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix}}_{\equiv e} = \underbrace{\begin{bmatrix} 1 & \sigma_1 & 0 & \cdots & 0 \\ 1 & 0 & \sigma_2 & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 1 & 0 & \cdots & 0 & \sigma_n \end{bmatrix}}_{\equiv \Gamma} \underbrace{\begin{bmatrix} \theta \\ \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}}_{\equiv \omega}$$

or simply

$$e = \Gamma\omega$$

Notice that the vector ω is a vector independent standard normal random variables.

However, player i only observes the signal e_j of players he is connected to, in addition to his own signal e_i . Thus, let the K_i by n matrix X_i be the matrix that selects the signals observed by player i :

$$\mathbb{I}_i = \{e_j\}_{j:g_{ij}=1} = X_i e = X_i \Gamma \omega$$

Notice that the optimal action of player i is set $a_i = \mathbb{E}[a_i^* | \mathbb{I}_i]$, and the expected payoff for a given network G is given by

$$\mathbb{E}[U_i | G] = -\mathbb{E}[(a_i - a_i^*)^2 | G] = -\mathbb{E}[\mathbb{E}[(a_i - a_i^*)^2 | \mathbb{I}_i] | G]$$

where the last equality hold based on law of iterated expectations. Using the optimal action choice, the expected value conditional on player i 's informational set is a conditional variance term and the payoff function is further simplified to

$$\mathbb{E}[U_i | G] = -\mathbb{E}[\text{Var}(a_i^* | \mathbb{I}_i) | G]$$

³⁰Remember that $e_0 = 0$, so we could have defined $\bar{a}_{-i} = \sum_{j=0}^n \beta_{-i,j} e_j$ instead.

Next, let's write the bliss action in matrix notation:

$$\begin{aligned}
a_i^* &= (1 - r_i)\theta + r_i\bar{a}_{-i} \\
&= (1 - r_i)\theta + r_i \sum_{j=1}^n \beta_{-i,j} e_j \\
&= (1 - r_i)\theta + r_i \sum_{j=1}^n \beta_{-i,j} \theta + r_i \sum_{j=1}^n \beta_{-i,j} \sigma_j \varepsilon_j \\
&= (1 - r_i \beta_{-i,0})\theta + r_i \sum_{j=1}^n \beta_{-i,j} \sigma_j \varepsilon_j \\
&= \underbrace{[1 - r_i \beta_{-i,0}, r_i \sigma_1 \beta_{-i,1}, r_i \sigma_2 \beta_{-i,2}, \dots, r_i \sigma_n \beta_{-i,n}]}_{\equiv F_i'} \omega \\
a_i^* &= F_i' \omega
\end{aligned}$$

where $\beta_{-i,0} = 1 - \sum_{j=1}^n \beta_{-i,j}$. The vector F_i does not depend on player i observed signal, it only depends on the network itself, thus player i 's expected payoff becomes

$$\mathbb{E}[U_i|G] = -\mathbb{E}[\text{Var}(a_i^*|\mathbb{I}_i)|G] = -F_i' \text{Var}(\omega|\mathbb{I}_i) F_i$$

and we can use bayes updating rule to compute the variance covariance term:³¹

$$\text{Var}(\omega|\mathbb{I}_i) = \text{Var}(\omega) - \text{Cov}(\omega, X_i \Gamma \omega)' \text{Var}(X_i \Gamma \omega)^{-1} \text{Cov}(\omega, X_i \Gamma \omega)$$

where

$$\begin{aligned}
\text{Var}(\omega) &= \mathbf{I} \\
\text{Var}(X_i \Gamma \omega) &= X_i \Gamma \Gamma' X_i' \\
\text{Cov}(\omega, X_i \Gamma \omega) &= X_i \Gamma
\end{aligned}$$

In order to successfully invert the variance-covariance matrix $\text{Var}(X_i \Gamma \omega)$, let's redefine Γ as

$$\Gamma = [\mathbf{1} \quad \Phi]$$

where $\mathbf{1}$ is a column vector of ones and

$$\Phi = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_n \end{bmatrix}_{n \times n}$$

³¹ Additionally, we also know the expected value: $\mathbb{E}[\omega|\mathbb{I}_i] = \text{Cov}(\omega, X_i \Gamma \omega)' \text{Var}(X_i \Gamma \omega)^{-1} X_i \Gamma \omega$

Using the above notation, we can simplify $\text{Var}(X_i\Gamma\omega)$ as follows

$$\text{Var}(X_i\Gamma\omega) = X_i\Gamma\Gamma'X_i' = X_i\Phi\Phi'X_i' + \mathbf{1}\mathbf{1}'$$

Notice that $X_i\Gamma\Gamma'X_i'$ is a diagonal matrix variance σ_j^2 's of signals that player i observes. This simplification is useful because $X_i\Gamma\Gamma'X_i'$ is easy to invert and we can apply Sherman-Morrison theorem:³²

$$\text{Var}(X_i\Gamma\omega)^{-1} = (X_i\Phi\Phi'X_i')^{-1} - \frac{1}{\phi_i}(X_i\Phi\Phi'X_i')^{-1}\mathbf{1}\mathbf{1}'(X_i\Phi\Phi'X_i')^{-1}$$

where $\phi_i = 1 + \mathbf{1}'(X_i\Phi\Phi'X_i')^{-1}\mathbf{1} = 1 + \sum_{j=1}^n g_{ij}\sigma_j^{-2}$. We can use the simplified inverse of the variance to compute the following:

$$\begin{aligned} \text{Cov}(\omega, X_i\Gamma\omega)' \text{Var}(X_i\Gamma\omega)^{-1} \text{Cov}(\omega, X_i\Gamma\omega) &= \begin{bmatrix} \mathbf{1}' \\ \Phi'X_i' \end{bmatrix} \text{Var}(X_i\Gamma\omega)^{-1} \begin{bmatrix} \mathbf{1} & X_i\Phi \end{bmatrix} \\ &= \mathcal{A} = \begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{bmatrix} \end{aligned}$$

where

$$\begin{aligned} \mathcal{A}_{11} &= \mathbf{1}'(X_i\Phi\Phi'X_i')^{-1}\mathbf{1} - \frac{1}{\phi_i}\mathbf{1}'(X_i\Phi\Phi'X_i')^{-1}\mathbf{1}\mathbf{1}'(X_i\Phi\Phi'X_i')^{-1}\mathbf{1} \\ \mathcal{A}_{12} &= \mathbf{1}'(X_i\Phi\Phi'X_i')^{-1}X_i\Phi - \frac{1}{\phi_i}\mathbf{1}'(X_i\Phi\Phi'X_i')^{-1}\mathbf{1}\mathbf{1}'(X_i\Phi\Phi'X_i')^{-1}X_i\Phi \\ \mathcal{A}_{21} &= \Phi'X_i'(X_i\Phi\Phi'X_i')^{-1}\mathbf{1} - \frac{1}{\phi_i}\Phi'X_i'(X_i\Phi\Phi'X_i')^{-1}\mathbf{1}\mathbf{1}'(X_i\Phi\Phi'X_i')^{-1}\mathbf{1} \\ \mathcal{A}_{22} &= \Phi'X_i'(X_i\Phi\Phi'X_i')^{-1}X_i\Phi - \frac{1}{\phi_i}\Phi'X_i'(X_i\Phi\Phi'X_i')^{-1}\mathbf{1}\mathbf{1}'(X_i\Phi\Phi'X_i')^{-1}X_i\Phi \end{aligned}$$

We can use $\mathbf{1}'(X_i\Phi\Phi'X_i')^{-1}\mathbf{1} = 1 - \phi_i$ to get

$$\mathcal{A} = \begin{bmatrix} \frac{\phi_i-1}{\phi_i} & \frac{1}{\phi_i}\mathbf{1}'(X_i\Phi\Phi'X_i')^{-1}X_i\Phi \\ \frac{1}{\phi_i}\Phi'X_i'(X_i\Phi\Phi'X_i')^{-1}\mathbf{1} & \mathcal{A}_{22} \end{bmatrix}$$

³² For any non-singular matrix A , column vectors u and v , and a scalar α , Sherman-Morrison theorem states that

$$(A + \beta uv')^{-1} = A^{-1} - \frac{\alpha}{\phi} A^{-1} uv' A^{-1}$$

where $\phi = 1 + \alpha v' A^{-1} u$. We apply this result by setting $A = X_i\Phi\Phi'X_i'$, $\alpha = 1$, $u = v = \mathbf{1}$. See Golub and Van Loan (2012) for more details.

Notice that

$$\Phi'X_i'(X_i\Phi\Phi'X_i')^{-1}\mathbf{1} = \begin{bmatrix} \frac{g_{i1}}{\sigma_1} \\ \frac{g_{i2}}{\sigma_2} \\ \vdots \\ \frac{g_{in}}{\sigma_n} \end{bmatrix}_{n \times 1}$$

$$\Phi'X_i'(X_i\Phi\Phi'X_i')^{-1}X_i\Phi = \begin{bmatrix} g_{i1} & 0 & \cdots & 0 \\ 0 & g_{i2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & g_{in} \end{bmatrix}_{n \times n}$$

$$\Phi'X_i'(X_i\Phi\Phi'X_i')^{-1}\mathbf{1}\mathbf{1}'(X_i\Phi\Phi'X_i')^{-1}X_i\Phi = \begin{bmatrix} \frac{g_{i1}g_{i1}}{\sigma_1\sigma_1} & \frac{g_{i1}g_{i2}}{\sigma_1\sigma_2} & \cdots & \frac{g_{i1}g_{in}}{\sigma_1\sigma_n} \\ & \vdots & & \\ \frac{g_{in}g_{i1}}{\sigma_n\sigma_1} & \frac{g_{in}g_{i2}}{\sigma_n\sigma_2} & \cdots & \frac{g_{in}g_{in}}{\sigma_n\sigma_n} \end{bmatrix}_{n \times n}$$

As a result, the matrix \mathcal{A} is further simplified to

$$\mathcal{A} = \frac{1}{\phi_i} \begin{bmatrix} \phi_i - 1, & \frac{g_{i1}}{\sigma_1}, & \frac{g_{i2}}{\phi_i\sigma_2}, & \cdots & \frac{g_{in}}{\sigma_n} \\ \frac{g_{i1}}{\sigma_1}, & g_{i1}\phi_i - \frac{g_{i1}g_{i1}}{\sigma_1\sigma_1}, & -\frac{g_{i1}g_{i2}}{\sigma_1\sigma_2}, & \cdots & -\frac{g_{i1}g_{in}}{\sigma_1\sigma_n} \\ \frac{g_{i2}}{\sigma_2}, & -\frac{g_{i2}g_{i1}}{\sigma_2\sigma_1}, & \phi_i g_{i2} - \frac{g_{i2}g_{i2}}{\sigma_2\sigma_2}, & \cdots & -\frac{g_{i2}g_{in}}{\sigma_2\sigma_n} \\ \vdots & \vdots & & \ddots & \vdots \\ \frac{g_{in}}{\sigma_n}, & -\frac{g_{in}g_{i1}}{\sigma_n\sigma_1}, & -\frac{g_{in}g_{i2}}{\sigma_n\sigma_2}, & \cdots & g_{in} - \frac{g_{in}g_{in}}{\sigma_n\sigma_n} \end{bmatrix}$$

and player i 's expected payoff becomes

$$\begin{aligned}
\mathbb{E}[U_i|G] &= -F'_i(\mathbf{I} - \mathcal{A})F_i \\
&= -\frac{1}{\phi_i}F'_i \begin{bmatrix} 1, & -\frac{g_{i1}}{\sigma_1}, & -\frac{g_{i2}}{\phi_i\sigma_2}, & \dots & -\frac{g_{in}}{\sigma_n} \\ -\frac{g_{i1}}{\sigma_1}, & (1 - g_{i1})\phi_i + \frac{g_{i1}}{\sigma_1\sigma_1}, & \frac{g_{i1}g_{i2}}{\sigma_1\sigma_2}, & \dots & \frac{g_{i1}g_{in}}{\sigma_1\sigma_n} \\ -\frac{g_{i2}}{\sigma_2}, & \frac{g_{i2}g_{i1}}{\sigma_2\sigma_1}, & (1 - g_{i2})\phi_i + \frac{g_{i2}}{\sigma_2\sigma_2}, & \dots & \frac{g_{i2}g_{in}}{\sigma_2\sigma_n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ -\frac{g_{in}}{\sigma_n}, & \frac{g_{in}g_{i1}}{\sigma_n\sigma_1}, & -\frac{g_{in}g_{i2}}{\sigma_n\sigma_2}, & \dots & (1 - g_{in})\phi_i + \frac{g_{in}}{\sigma_n\sigma_n} \end{bmatrix} F_i \\
&= -\frac{1}{\phi_i} \begin{bmatrix} 1 - r_i \sum_{j=0} g_{ij}\beta_{-i,j} \\ -g_{i1}\frac{1}{\sigma_1} + g_{i1}r_i\frac{1}{\sigma_1} \sum_{j=0}^n g_{ij}\beta_{-i,j} + (1 - g_{i1})\phi_i r_i \beta_{-i,1}\sigma_1 \\ \vdots \\ -g_{in}\frac{1}{\sigma_n} + g_{in}r_i\frac{1}{\sigma_n} \sum_{j=0}^n g_{ij}\beta_{-i,j} + (1 - g_{in})\phi_i r_i \beta_{-i,n}\sigma_n \end{bmatrix} F_i \\
&= -\frac{1}{\phi_i} \left\{ (1 - r_i\beta_{-i,0})(1 - r_i \sum_{j=0} g_{ij}\beta_{-i,j}) \right. \\
&\quad \left. - r_i \sum_{s=0}^n \beta_{-i,s} g_{is} (1 - r_i \sum_{j=0} g_{ij}\beta_{-i,j}) + \phi_i r_i^2 \sum_{j=1}^n (1 - g_{ij}\beta_{-i,j}^2 \sigma_j^2) \right\} \\
&= -\frac{1}{\phi_i} (1 - r_i \sum_{j=0} g_{ij}\beta_{-i,j})^2 - r_i^2 \sum_{j=0}^n (1 - g_{ij})\beta_{-i,j}^2 \sigma_j^2
\end{aligned}$$

where $\phi_i = 1 + \sum_{j=1}^n g_{ij}\sigma_j^{-2}$.

If $r_i = r$ for every $i = 1, \dots, n$ and $\sigma_j = \sigma$ for every $j = 1, \dots, n$, then $\phi_i = \frac{\sigma^2 + K_i - 1}{\sigma^2}$ and the expected payoff becomes

$$-\frac{\sigma^2}{\sigma^2 + K_i - 1} \left(1 - r \sum_{j=0} g_{ij}\beta_{-i,j} \right)^2 - r^2 \sigma^2 \sum_{j=0}^n (1 - g_{ij})\beta_{-i,j}^2$$

where $K_i = \sum_{j=0}^n g_{ij}$. If $\sigma = 1$ then the expected payoff is given by

$$-\frac{1}{K_i} \left(1 - r \sum_{j=0} g_{ij}\beta_{-i,j} \right)^2 - r^2 \sum_{j=0}^n (1 - g_{ij})\beta_{-i,j}^2$$

C Proof of Proposition 4

In this section we show Proposition 4, that any equilibrium satisfies both properties.

C.1 Property 1

We first start by showing Property 1. By using the Lemma 1, proven in the main text, we have that an agent best response to other agents choices of connections is to tap into the signal of the most influential agent. the next step is to show that the agent, j , whose signal is tapped into the most is also the agent j with higher $\beta_{-i,j}$ is, for all i .

The first step is to establish a relationship between $\beta_{-i,j}$ and β_j . More specifically $\beta_f > \beta_h$ implies that $\beta_{-i,f} > \beta_{-i,h}$ for every i that satisfies one of the following situations: (i) $g_{if} = g_{ih} = 0$; (ii) $g_{if} = g_{ih} = 1$; (iii) $g_{if} = 0$ and $g_{ih} = 1$; or (iv) $g_{if} = 1$ and $g_{ih} = 0$. We start with the fourth case. If $g_{if} = 1$ and $g_{ih} = 0$, it must be that $\beta_{-i,f} > \beta_{-i,h}$ by lemma 1.

For the other 3 cases, the loading $\beta_{-i,j}$ can be written as

$$\beta_{-i,j} = \frac{n}{n-1}\beta_j - \frac{1}{n-1}\lambda_{ij},$$

where

$$\begin{aligned} \lambda_{ij} &= (1 - \tilde{r})\frac{g_{ij}}{K_i + 1} + \tilde{r}g_{ij}\beta_j + \tilde{r}\frac{g_{ij}}{K_i + 1} - \tilde{r}\sum_{s=0}^n \beta_s g_{is} \frac{g_{ij}}{K_i + 1} \\ &= \frac{g_{ij}}{K_i + 1} + \tilde{r}\left[g_{ij}\beta_j - \sum_{s=0}^n \beta_s g_{is} \frac{g_{ij}}{K_i + 1}\right] \\ &= g_{ij}\left[\frac{1}{K_i + 1} + \tilde{r}\left[\beta_j - \frac{\sum_{s=0}^n \beta_s g_{is}}{K_i + 1}\right]\right] \end{aligned}$$

is the attention agent i pays to the j^{th} signal. A detailed derivation of the individual weights is provided in the Appendix, A.

Substituting in the $\beta_{-i,j}$ expression,

$$\begin{aligned} \beta_{-i,j} &= \frac{n}{n-1}\beta_j - \frac{1}{n-1}\gamma_{ij} \\ &= \frac{n}{n-1}\beta_j - \frac{g_{ij}}{n-1}\left[\frac{1}{K_i + 1} + \tilde{r}\left[\beta_j - \frac{\sum_{s=0}^n \beta_s g_{is}}{K_i + 1}\right]\right] \\ &= \frac{n}{n-1}\beta_j - \frac{g_{ij}}{n-1}\left[\frac{1}{K_i + 1} + \tilde{r}\left[\beta_j - \frac{\sum_{s=0}^n \beta_s g_{is}}{K_i + 1}\right]\right] \end{aligned}$$

If $g_{if} = g_{ih} = 0$, then $\beta_{-i,f} - \beta_{-i,h} = \frac{n}{n-1}(\beta_f - \beta_h) > 0$.

If $g_{if} = g_{ih} = 1$, then $\beta_{-i,f} - \beta_{-i,h} = \left(\frac{n}{n-1} - \frac{\tilde{r}}{n-1}\right)(\beta_f - \beta_h) > 0$.

If $g_{if} = 0$ and $g_{ih} = 1$, then $\beta_{-i,f} - \beta_{-i,h} = \frac{n}{n-1}(\beta_f - \beta_h) + \frac{1-\tilde{r}\sum_{s=0}^n \beta_s g_{is}}{(n-1)(K_i+1)} + \frac{\tilde{r}}{N-1}\beta_h > 0$.

Thus, we have that $\beta_f \geq \beta_h \implies \beta_{-i,f} \geq \beta_{-i,h} \forall i$.

All what's left to show of is that an agent whose signal is more tapped into has a higher influence on the average action, that is $\bar{K}_f \geq \bar{K}_h \implies \beta_f \geq \beta_h$.

Lemma 2. *Agent f 's private signal has a weakly higher impact in the average action than agent h 's has if, and only if, agent f is weakly more looked at than agent h .*

$$\beta_f \geq \beta_h \iff \bar{K}_f \geq \bar{K}_h$$

Proof. First of all, note that lemma 1 guarantees the only if part. As usual in economics, assume not. Assume $\beta_f < \beta_h$ & $\bar{K}_f \geq \bar{K}_h$.

From the fact that all players use the beta ranking to decide who to look, we have that $\beta_f < \beta_h$ implies $\beta_{-i,f} < \beta_{-i,h} \forall i$, and thus $g_{i,h} \geq g_{i,f} \forall i$. Since $\bar{K}_f \geq \bar{K}_h$, it must be that $\bar{K}_f = \bar{K}_h$, and also that $g_{i,h} = g_{i,f} \forall i$.

Let's now consider the formula for β_j .

$$\begin{aligned} \beta_j &= \frac{1-\tilde{r}}{n} \sum_{i=1}^n \frac{g_{ij}}{\sigma^2 + K_i - 1} + \frac{\tilde{r}}{n} \left[\beta_j \bar{K}_j + \sum_{i=1}^n \frac{g_{ij}}{\sigma^2 + K_i - 1} - \sum_{i=1}^n \sum_{s=0}^n \frac{\beta_s g_{is} g_{ij}}{\sigma^2 + K_i - 1} \right] \\ \beta_f - \beta_h &= \frac{(1-\tilde{r})}{n} \sum_{i=1}^n \frac{g_{if} - g_{ih}}{\sigma^2 + K_i - 1} + \frac{\tilde{r}}{n} \left[\beta_f \bar{K}_f - \beta_h \bar{K}_h + \sum_{i=1}^n \frac{g_{if} - g_{ih}}{\sigma^2 + K_i - 1} - \sum_{i=1}^n \sum_{s=0}^n \frac{\beta_s g_{is} (g_{if} - g_{ih})}{\sigma^2 + K_i - 1} \right] \\ &= \frac{\tilde{r}}{n} [(\beta_f - \beta_h) \bar{K}_f] \end{aligned}$$

Thus $\beta_f - \beta_h = 0$, a contradiction. □

C.2 Property 2

To show that Property 2 holds in equilibrium is a little more evolved. We will again show it by contradiction, so first assume that in equilibrium a player m is at the same time strictly tapping into more signals and has her signal being tapped more than another player l , that is:

$$\bar{K}_m > \bar{K}_l \text{ and } K_m > K_l$$

The contradiction will be constructed in the following way: If it is worth for player m to pay to tap into more signals, then player l 's deviation to look at those is also profitable.

Player m is tapping into d extra signals First, note that there is a set of signals such that player m is tapping into and player l is not. By Property 1, all players are ranked according to a common list and thus all signals that player l pays to tap into player m also taps into.

Let there be $d \geq 1$ signals—abusing notation, we call the corresponding set of signals, D_m and D_l —such that this is the largest number of signals such that if m were to disconnect to D_m she would be obtaining the same number of signals as l is. Accordingly, if l were to form connections to D_l she would be obtaining the same number of signals as m is. By convention, we define D_m and D_l to be the set of signals that give the best possible set for players m and l that satisfy the above.

The fact that player m receives signal m for free disciplines the sets D_m and D_l . Indeed in many situations, they are not equal, as agent m cannot deviate and stop looking at her own signal. Let's also define the set S_M , to be the set of signals that agent m observes and l does not observe, while the set S_L is the set of signals that agent l observes and m does not.

Example 5. Consider the following example, in which player m is the 5th most attractive signal to be tapped into while player l is the 9th. In each of the four situations above, we contemplate a different configuration between the two players. In the first, they both tap into each other signals, while on the last neither does so. In the second, even though m taps into l 's signal, l does not correspond, and finally, the third situation presents the inverse.

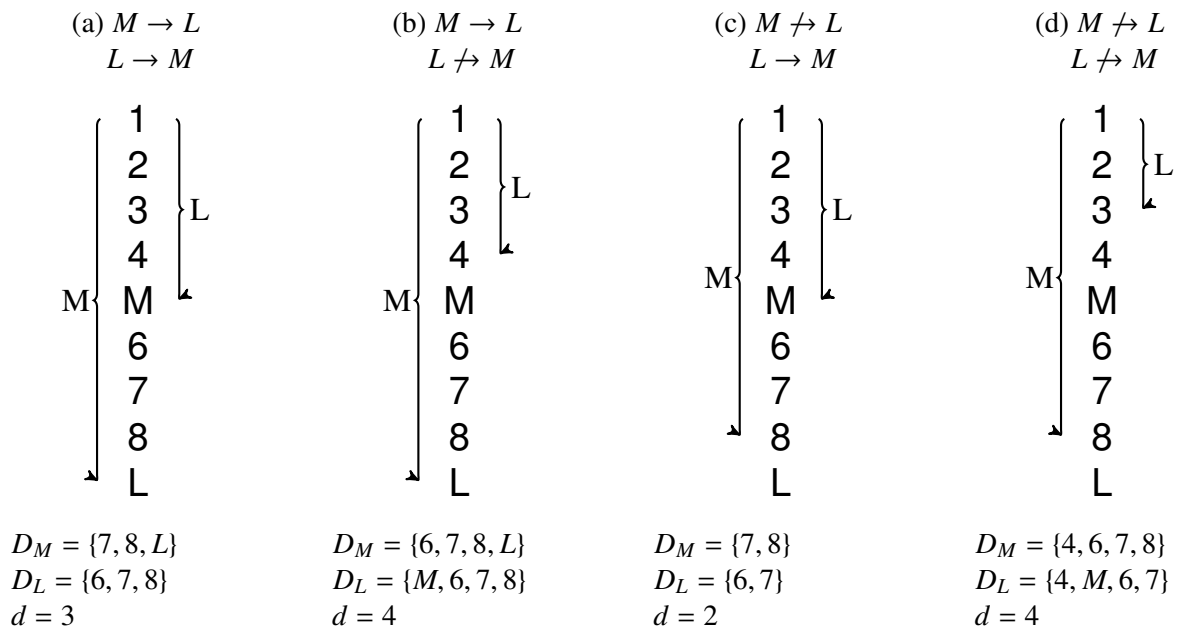


Figure 16: Four different configurations concerning players m and l choices of connections.

The four examples pictured above show different possibilities for the sets D_M and D_L , depending on the connections formed. It is also interesting to understand what is the information

set of players in each situation. In the first situation, the players Information sets are $I_M = \{e_0, e_1, e_2, e_3, e_4, e_M, e_6, e_7, e_8, e_L\}$ and $I_L = \{e_0, e_1, e_2, e_3, e_4, e_M, e_L\}$, and thus $S_M = \{e_6, e_7, e_8\}$ and $S_L = \{\emptyset\}$. In the second situation, they are $I_M = \{e_0, e_1, e_2, e_3, e_4, e_M, e_6, e_7, e_8, e_L\}$ and $I_L = \{e_0, e_1, e_2, e_3, e_4, e_L\}$ (and thus $S_M = \{e_6, e_7, e_8\}$ and $S_L = \{\emptyset\}$), while on the third they are $I_M = \{e_0, e_1, e_2, e_3, e_4, e_M, e_6, e_7, e_8\}$ and $I_L = \{e_0, e_1, e_2, e_3, e_4, e_M, e_L\}$, with $S_M = \{e_6, e_7, e_8\}$ and $S_L = \{e_L\}$. Finally, in the fourth configuration, $I_M = \{e_0, e_1, e_2, e_3, e_4, e_M, e_6, e_7, e_8\}$ and $I_L = \{e_0, e_1, e_2, e_3, e_L\}$, and thus $S_M = \{e_4, e_M, e_6, e_7, e_8\}$ and $S_L = \{\emptyset\}$.

In the figure above we can compare the sets D_m and D_l . For instance, we can note that the signals listed in D_l weakly dominate the ones listed in D_m . This is a direct result of the fact that player m receives the more attractive signal m for free and cannot stop tapping into it, while player l receives signal l .

The idea of the proof is to show that if player m is not willing to deviate and stop observing signals in D_m , then it is optimal for agent l to deviate and start looking at players in D_l . However, we do not proceed directly to it. Below we show that if player m is not willing to deviate and stop looking at signals D_l , then it would be optimal for agent l to deviate and start looking at players in D_l . However, by Lemma 1, we know that the deviation of not looking at signals in D_l is weakly dominated by the original deviation, thus proving the original claim. From now on, we call the set D_l as D , for deviation.

Before we proceed with the analysis, let us show the two lemmas below.

Lemma 3. *The summed influence of the signals in the set D is higher over the average action excluding agent l than over the average action excluding agent m .*

$$\sum_{j \in D} \beta_{-l,j} > \sum_{j \in D} \beta_{-m,j}$$

Proof. This is a direct result of the fact that l is not tapping into any signal in D while m is tapping into the signals of D . This implies that

$$\beta_j = \frac{n-1}{n} \beta_{-m,j} + \frac{1}{n} \lambda_{m,j} = \frac{n-1}{n} \beta_{-l,j} + \frac{1}{n} \lambda_{l,j} = \frac{n-1}{n} \beta_{-l,j}$$

$$\beta_{-m,j} = \beta_{-l,j} - \frac{\lambda_{m,j}}{n} < \beta_{-l,j}$$

□

Lemma 4. *The summed influence of the signals outside of the set D that agent m is paying attention*

through her connections is higher than the one that l is.

$$\sum_{j \notin D} g_{m,j} \beta_{-m,j} > \sum_{j \notin D} g_{l,j} \beta_{-l,j}$$

Proof.

Observe that l is not tapping into any signal in D , thus we have:

$$\sum_{j \notin D} g_{l,j} \beta_{-l,j} = \sum_{j=0}^n g_{l,j} \beta_{-l,j} = 1 - \sum_{j=0}^n (1 - g_{l,j}) \beta_{-l,j}.$$

Player m is tapping into all signals in the set D , and thus

$$\sum_{j \notin D} g_{m,j} \beta_{-m,j} = \sum_{j=0}^n g_{m,j} \beta_{-m,j} - \sum_{j \in D} \beta_{-m,j} = 1 - \sum_{j=0}^n (1 - g_{m,j}) \beta_{-m,j} - \sum_{j \in D} \beta_{-m,j}.$$

Subtracting the first from the second,

$$\sum_{j \notin D} g_{m,j} \beta_{-m,j} - \sum_{j \notin D} g_{l,j} \beta_{-l,j} = \sum_{j=0}^n (1 - g_{l,j}) \beta_{-l,j} - \sum_{j=0}^n (1 - g_{m,j}) \beta_{-m,j} - \sum_{j \in D} \beta_{-m,j}$$

We can partition all signals in the economy into four groups. The set of signals they are both observing S_B ; the set of signals neither is observing S_N ; the set of signals m is observing and l is not, S_M ; and the set of signals l is observing and m is not, S_L .

$$\sum_{j \notin D} g_{m,j} \beta_{-m,j} - \sum_{j \notin D} g_{l,j} \beta_{-l,j} = \sum_{j \in S_N} (\beta_{-l,j} - \beta_{-m,j}) + \sum_{j \in S_M} \beta_{-l,j} - \sum_{j \in S_L} \beta_{-m,j} - \sum_{j \in D} \beta_{-m,j}$$

There exists a relationship between the influence of a signal has over the average action excluding a player or another player. $\beta_j = \frac{n-1}{n} \beta_{-m,j} + \frac{1}{n} \lambda_{m,j}$. For a signal j in S_N , both $\lambda_{m,j}$ and $\lambda_{l,j}$ are zero. For a signal in S_M , $\lambda_{l,j} = 0$ and for a signal in S_L , $\lambda_{m,j} = 0$. Thus,

$$\begin{aligned} \sum_{j \notin D} g_{m,j} \beta_{-m,j} - \sum_{j \notin D} g_{l,j} \beta_{-l,j} &= \frac{n}{n-1} \left[\sum_{j \in S_M} \beta_j - \sum_{j \in S_L} \beta_j \right] - \sum_{j \in D} \beta_{-m,j} \\ &= \frac{n}{n-1} \left[\sum_{j \in S_M} \beta_j - \sum_{j \in S_L} \beta_j - \sum_{j \in D} \beta_j + \sum_{j \in D} \frac{\lambda_{m,j}}{n} \right] \end{aligned}$$

The final argument is to note that the set of signals player m observes and l does not, S_M , weakly dominate the union of the set of D with the set of signals that l observes while m does not S_L . It is

worth noting that the proof does not rely on player m being more observed than player l , only on the fact that m is observing more. \square

Let $\Delta\Pi_m$ be the difference of ex-ante expected payoff of player m between breaking those d extra links in D or maintaining them. Notice that if player m unilaterally deviates and breaks those links, no other player changes her action. Thus the influence of signals to other players action will not change, and $\beta_{-m,j}$ must be constant. Similarly, let $\Delta\Pi_l$ be the difference of ex-ante expected payoff of player l between tapping into those d signals or not forming those links. Given that by assumption we are at strict equilibrium, both expected payoff differences should be strictly smaller than zero.

Regarding notation, we keep $g_{i,j}$ to be the original linking strategy, of the proposed equilibrium.

$$\Delta\Pi_m = \underbrace{-\frac{\sigma^2 \left(1 - r \sum_{j \notin D} g_{mj} \beta_{-m,j}\right)^2}{\sigma^2 + K_m - d - 1}}_{\text{Without tapping}} + \underbrace{\frac{\sigma^2 \left(1 - r \sum_{j=0}^n g_{mj} \beta_{-m,j}\right)^2}{\sigma^2 + K_m - 1}}_{\text{Tapping into D}} \\ - \underbrace{r^2 \sigma^2 \left(\sum_{j=0}^n (1 - g_{mj}) \beta_{-m,j}^2 + \sum_{j \in D} \beta_{-m,j}^2 \right)}_{\text{Without tapping}} + \underbrace{r^2 \sigma^2 \sum_{j=0}^n (1 - g_{mj}) \beta_{-m,j}^2}_{\text{Tapping into D}} + \underbrace{\Delta C}_{\text{Cost difference of not looking to signals}}$$

Breaking up $\left(1 - r \sum_{j=0}^n g_{mj} \beta_{-m,j}\right)^2$ into $\left(1 - r \sum_{j \notin D} g_{mj} \beta_{-m,j}\right)^2 - 2 \left(1 - r \sum_{j \notin D} g_{mj} \beta_{-m,j}\right) \sum_{j \in D} \beta_{-m,j} + r^2 \left(\sum_{j \in D} \beta_{-m,j}\right)^2$, we have:

$$\Delta\Pi_m = - \left(1 - r \sum_{j \notin D} g_{mj} \beta_{-m,j}\right)^2 \left(\frac{\sigma^2}{\sigma^2 + K_m - d - 1} - \frac{\sigma^2}{\sigma^2 + K_m - 1} \right) - 2 \frac{r \sigma^2}{\sigma^2 + K_m - 1} \left(1 - r \sum_{j \notin D} g_{mj} \beta_{-m,j}\right) \sum_{j \in D} \beta_{-m,j} \\ + \frac{r^2 \sigma^2}{\sigma^2 + K_m - 1} \left(\sum_{j \in D} \beta_{-m,j}\right)^2 - r^2 \sigma^2 \sum_{j \in D} \beta_{-m,j}^2 + \Delta C$$

If we analyze the ex-ante expected payoff difference for player l , we have:

$$\Delta\Pi_l = \underbrace{-\frac{\sigma^2 \left(1 - r \sum_{j=0}^n g_{lj} \beta_{-l,j} - r \sum_{j \in D} \beta_{-l,j}\right)^2}{\sigma^2 + K_l + d - 1}}_{\text{Tapping into D}} + \underbrace{\frac{\sigma^2 \left(1 - r \sum_{j=0}^n g_{lj} \beta_{-l,j}\right)^2}{\sigma^2 + K_l - 1}}_{\text{Without tapping}} \\ - \underbrace{r^2 \sigma^2 \left(\sum_{j=0}^n (1 - g_{lj}) \beta_{-l,j}^2 - \sum_{j \in D} \beta_{-l,j}^2 \right)}_{\text{Tapping into D}} + \underbrace{r^2 \sigma^2 \sum_{j=0}^n (1 - g_{lj}) \beta_{-l,j}^2}_{\text{Without tapping}} - \underbrace{\Delta C}_{\text{Cost difference of looking to signals}}$$

Through the same steps as with m

$$\begin{aligned} \Delta\Pi_l &= \left(1 - r \sum_{j \notin D} g_{lj} \beta_{-l,j}\right)^2 \left(\frac{\sigma^2}{\sigma^2 + K_l - 1} - \frac{\sigma^2}{\sigma^2 + K_l + d - 1}\right) + 2 \frac{r\sigma^2}{\sigma^2 + K_l + d - 1} \left(1 - r \sum_{j \notin D} g_{lj} \beta_{-l,j}\right) \sum_{j \in D} \beta_{-l,j} \\ &\quad - \frac{r^2 \sigma^2}{\sigma^2 + K_l + d - 1} \left(\sum_{j \in D} \beta_{-l,j}\right)^2 + r^2 \sigma^2 \sum_{j \in D} \beta_{-l,j}^2 + \Delta C \end{aligned}$$

Before we proceed, note that the cost difference is the same in both cases, and also that $K_l + d = K_m$. Thus, as we sum the two differences, we have:

$$\begin{aligned} \Delta\Pi_l + \Delta\Pi_m &= \left(\frac{\sigma^2}{\sigma^2 + K_l - 1} - \frac{\sigma^2}{\sigma^2 + K_l + d - 1}\right) \left[\left(1 - r \sum_{j \notin D} g_{lj} \beta_{-l,j}\right)^2 - \left(1 - r \sum_{j \notin D} g_{mj} \beta_{-m,j}\right)^2 \right] \\ &\quad + 2 \frac{r\sigma^2}{\sigma^2 + K_l + d - 1} \left[\left(1 - r \sum_{j \notin D} g_{lj} \beta_{-l,j}\right) \sum_{j \in D} \beta_{-l,j} - \left(1 - r \sum_{j \notin D} g_{mj} \beta_{-m,j}\right) \sum_{j \in D} \beta_{-m,j} \right] \\ &\quad - \frac{r^2 \sigma^2}{\sigma^2 + K_l + d - 1} \left[\left(\sum_{j \in D} \beta_{-l,j}\right)^2 - \left(\sum_{j \in D} \beta_{-m,j}\right)^2 \right] + r^2 \sigma^2 \left[\sum_{j \in D} \beta_{-l,j}^2 - \sum_{j \in D} \beta_{-m,j}^2 \right] \end{aligned}$$

Reordering terms

$$\begin{aligned} \Delta\Pi_l + \Delta\Pi_m &= \left(\frac{\sigma^2}{\sigma^2 + K_l - 1} - \frac{\sigma^2}{\sigma^2 + K_l + d - 1}\right) \left[\left(1 - r \sum_{j \notin D} g_{lj} \beta_{-l,j}\right)^2 - \left(1 - r \sum_{j \notin D} g_{mj} \beta_{-m,j}\right)^2 \right] \\ &\quad + 2 \frac{r\sigma^2}{\sigma^2 + K_l + d - 1} \left[\left(1 - r \sum_{j \notin D} g_{lj} \beta_{-l,j}\right) \sum_{j \in D} \beta_{-l,j} - \left(1 - r \sum_{j \notin D} g_{mj} \beta_{-m,j}\right) \sum_{j \in D} \beta_{-m,j} \right] \\ &\quad + r^2 \sigma^2 \left[\left(\sum_{j \in D} \beta_{-l,j}^2 - \sum_{j \in D} \beta_{-m,j}^2\right) - \frac{\left(\sum_{j \in D} \beta_{-l,j}\right)^2 - \left(\sum_{j \in D} \beta_{-m,j}\right)^2}{\sigma^2 + K_l + d - 1} \right] \end{aligned}$$

By using lemma 3 and lemma 4, we have that the first two lines are strictly positive. We now focus to show that the third line is as well, characterizing the contradiction. That is equivalent to sign the following,

$$\begin{aligned}
& \frac{(\sum_{j \in D} \beta_{-l,j}^2 - \sum_{j \in D} \beta_{-m,j}^2)}{d} - \frac{d}{\sigma^2 + K_l + d - 1} \frac{(\sum_{j \in D} \beta_{-l,j})^2 - (\sum_{j \in D} \beta_{-m,j})^2}{d^2} \\
& \geq \frac{\sum_{j \in D} (\beta_{-l,j}^2 - \beta_{-m,j}^2)}{d} - \frac{(\sum_{j \in D} \beta_{-l,j})^2 - (\sum_{j \in D} \beta_{-m,j})^2}{d^2} \\
& = \frac{\sum_{j \in D} (\beta_{-l,j} - \beta_{-m,j})(\beta_{-l,j} + \beta_{-m,j})}{d} - \frac{(\sum_{j \in D} \beta_{-l,j} - \sum_{j \in D} \beta_{-m,j})(\sum_{j \in D} \beta_{-l,j} + \sum_{j \in D} \beta_{-m,j})}{d^2} \\
& = \frac{\sum_{j \in D} (\beta_{-l,j} - \beta_{-m,j})(\beta_{-l,j} + \beta_{-m,j})}{d} - \frac{\sum_{j \in D} (\beta_{-l,j} - \beta_{-m,j})}{d} \frac{\sum_{j \in D} (\beta_{-l,j} + \beta_{-m,j})}{d}
\end{aligned}$$

If we interpret $\beta_{-l,j} - \beta_{-m,j}$ as a variable and $\beta_{-l,j} + \beta_{-m,j}$ as another variable, we have that the above expression is simply the average of the product of two variables subtracted by the product of two averages. That is the sample covariance of these two variables when the sample is the set of signals in D .

$$= \text{Cov}(\beta_{-l,j} - \beta_{-m,j}, \beta_{-l,j} + \beta_{-m,j})$$

We can now use the traditional operations with covariance to obtain

$$\begin{aligned}
& = \text{Cov}(\beta_{-l,j}, \beta_{-l,j}) - \text{Cov}(\beta_{-m,j}, \beta_{-m,j}) \\
& = \text{Var}(\beta_{-l,j}) - \text{Var}(\beta_{-m,j})
\end{aligned}$$

From the fact that l does not observe any signal in D , we have that $\beta_{-l,j} = \beta_{-m,j} + \frac{1}{n-1} \lambda_{m,j}$.

$$\begin{aligned}
& = \text{Var}(\beta_{-m,j} + \frac{1}{n-1} \lambda_{m,j}) - \text{Var}(\beta_{-m,j}) \\
& = \text{Var}(\beta_{-m,j}) + \text{Var}(\frac{1}{n-1} \lambda_{m,j}) + 2\text{Cov}(\beta_{-m,j}, \frac{1}{n-1} \lambda_{m,j}) - \text{Var}(\beta_{-m,j}) \\
& = \frac{1}{(n-1)^2} \text{Var}(\lambda_{m,j}) + \frac{2}{n-1} \text{Cov}(\beta_{-m,j}, \lambda_{m,j})
\end{aligned}$$

While the first term is obviously positive, we still have to show that the second one is as well. It amounts to show that a player considers more, and thus gives it a higher weight, a signal that comes from a more influential source.

Lemma 5. *Consider a subset D of the agent i 's information set. The covariance between the influence a signal has over the average action not including agent i 's action and how influential*

that particular signal is to agent i 's action is positive.

$$\text{Cov}_{j \in D}(\beta_{-i,j}, \lambda_{i,j}) > 0$$

Proof. We start by computing the best action, as a function of signals observed. From Appendix 2, concerning the payoff function, we know that $a_i^* = F_i' \omega$, and that

$$a_i = \mathbb{E}[\omega | \mathbb{I}_i] = \text{Cov}(\omega, X_i \Gamma \omega)' \text{Var}(X_i \Gamma \omega)^{-1} X_i \underbrace{\Gamma \omega}_{[e_1, e_2, \dots, e_n]}.$$

Let's start by computing it:

$$\begin{aligned} \text{Cov}(\omega, X_i \Gamma \omega)' \text{Var}(X_i \Gamma \omega)^{-1} X_i &= \begin{bmatrix} B_{11 \times n} \\ B_{2n \times n} \end{bmatrix}_{(n+1) \times n} \\ B_1 &= \mathbf{1}'(X_i \Phi \Phi' X_i')^{-1} X_i - \frac{1}{\phi_i} \mathbf{1}'(X_i \Phi \Phi' X_i')^{-1} \mathbf{1} \mathbf{1}'(X_i \Phi \Phi' X_i')^{-1} X_i \\ B_2 &= \Phi' X_i'(X_i \Phi \Phi' X_i')^{-1} X_i - \frac{1}{\phi_i} \Phi' X_i'(X_i \Phi \Phi' X_i')^{-1} \mathbf{1} \mathbf{1}'(X_i \Phi \Phi' X_i')^{-1} X_i \end{aligned}$$

We can use the fact that agents are ex-ante identical, $\sigma_i = \sigma$ and $r_i = r$ for all agents, and thus

$$\begin{aligned} X_i \Phi \Phi' X_i' &= \sigma^2 \mathbf{I}_{K_i - 1} \\ \mathbf{1}'(X_i \Phi \Phi' X_i')^{-1} X_i &= \sigma^{-2} [g_{i,1}, g_{i,2}, \dots, g_{i,n}]_{1 \times n} \\ \mathbf{1}'(X_i \Phi \Phi' X_i')^{-1} \mathbf{1} &= \sigma^{-2} (K_i - 1)_{1 \times 1} \end{aligned}$$

Which gives us a simplified expression for B_1

$$B_1 = \frac{1}{\sigma^2 + K_i - 1} [g_{i,1}, g_{i,2}, \dots, g_{i,n}]_{1 \times n}$$

Working similarly with elements of B_2

$$\begin{aligned} \Phi' X_i'(X_i \Phi \Phi' X_i')^{-1} X_i &= \sigma^{-1} X_i' X_i = \sigma^{-1} \text{diag}([g_{i,1}, g_{i,2}, \dots, g_{i,n}])_{n \times n} \\ \Phi' X_i'(X_i \Phi \Phi' X_i')^{-1} \mathbf{1} \mathbf{1}'(X_i \Phi \Phi' X_i')^{-1} X_i &= \sigma^{-3} \begin{bmatrix} g_{i,1} \\ g_{i,2} \\ \dots \\ g_{i,n} \end{bmatrix}_{n \times 1} [g_{i,1}, g_{i,2}, \dots, g_{i,n}]_{1 \times n} \end{aligned}$$

For any signal e_j , $j \in \{1, 2, \dots, n\}$, we can compute the linear coefficient of that particular signal

over the action of agent i . We have that,

$$\lambda_{i,j} = \frac{1 - r\beta_{-i,0}}{\sigma^2 + K_i - 1} g_{i,j} + r\beta_{-i,j} g_{i,j} - \frac{r}{\sigma^2 + K_i - 1} g_{i,j} \sum_{s=1}^n \beta_{-i,s} g_{i,s}$$

$$\lambda_{i,j} = g_{i,j} \left[\frac{1 - r\beta_{-i,0}}{\sigma^2 + K_i - 1} + r\beta_{-i,j} - \frac{r}{\sigma^2 + K_i - 1} \sum_{s=1}^n \beta_{-i,s} g_{i,s} \right]$$

Finally, we can focus on the covariance between the influence, $\beta_{-i,j}$, and the linear coefficient, $\lambda_{i,j}$.

$$\begin{aligned} \text{Cov}_{j \in D}(\beta_{-i,j}, \lambda_{i,j}) &= \mathbb{E}_{j \in D} [\beta_{-i,j} \lambda_{i,j}] - \mathbb{E}_{j \in D} [\beta_{-i,j}] \mathbb{E}_{j \in D} [\lambda_{i,j}] \\ \text{Cov}_{j \in D}(\beta_{-i,j}, \lambda_{i,j}) &= r \mathbb{E}_{j \in D} [\beta_{-i,j}^2] - r \mathbb{E}_{j \in D} [\beta_{-i,j}]^2 = r \text{Var}_{j \in D}(\beta_{-i,j}) > 0 \end{aligned}$$

□

This gives us that $\Delta \Pi_m + \Delta \Pi_l \geq 0$, even though by assumption both elements were smaller than zero, characterizing our contradiction.

D Core-Periphery

A core-periphery network is organized in two tiers, with all members of the core observing each other, while members of the periphery do not observe each other. If an information architecture is a hierarchical directed network, but it is not a core periphery network, it must be that a player in the bottom tier (call such player c) observes a player b and a player a , while player a does not observe player b , nor c . This characterizes situations in which there are more than two tiers, situations in which the top tier does not observe each other, as well as situations in which members of the bottom tier observe each other. Player c is tapping into at least as many looks as player a is plus two signals. Thus, if player c stopped observing player b 's signal, her information set would still be a strict super set of player a 's.

We will compare two deviations, for player c to break the link with all players she is observing and player a is not, call such set D , and for player a to form links with such set. Note that set D is not empty, and has $d > 1$ elements. Note also that, even after player a forms links with all players in D , player a 's information set will still be a strict subset of player c 's. The reason for that is that c is observing a , but the converse is not true. First we present the payoff difference for player c , comparing the payoff of deviating with the payoff of maintaining the link.

$$\begin{aligned} \Delta\Pi_c &= \underbrace{-\frac{\sigma^2(1-r\sum_{j\notin D}g_{c,j}\beta_{-c,j})^2}{\sigma^2+K_c-d-1}}_{\text{Without tapping}} + \underbrace{\frac{\sigma^2(1-r\sum_{j=0}^n g_{c,j}\beta_{-c,j})^2}{\sigma^2+K_c-1}}_{\text{Tapping into D}} \\ &\quad - \underbrace{r^2\sigma^2\left(\sum_{j=0}^n(1-g_{c,j})\beta_{-c,j}^2+r\sum_{j\in D}\beta_{-c,j}^2\right)}_{\text{Without tapping}} + \underbrace{r^2\sigma^2\sum_{j=0}^n(1-g_{c,j})\beta_{-c,j}^2}_{\text{Tapping into D}} + \underbrace{C(K_c-1)-C(K_c-1-d)}_{\text{Cost difference of not looking to signals}} \\ \Delta\Pi_c &= -\left(1-r\sum_{j\notin D}g_{c,j}\beta_{-c,j}\right)^2\left(\frac{\sigma^2}{\sigma^2+K_c-d-1}-\frac{\sigma^2}{\sigma^2+K_c-1}\right)-2\frac{r\sigma^2}{\sigma^2+K_c-1}\left(1-r\sum_{j\notin D}g_{c,j}\beta_{-c,j}\right)\sum_{j\in D}\beta_{-c,j} \\ &\quad +\frac{r^2\sigma^2}{\sigma^2+K_c-1}\left(\sum_{j\in D}\beta_{-c,j}\right)^2-r^2\sigma^2\sum_{j\in D}\beta_{-c,j}^2+C(K_c-1)-C(K_c-1-d) \\ \Delta\Pi_c &= -\left(1-r\sum_{j\notin D}g_{c,j}\beta_{-c,j}\right)^2\left(\frac{\sigma^2}{\sigma^2+K_c-d-1}-\frac{\sigma^2}{\sigma^2+K_c-1}\right)-2\frac{r\sigma^2}{\sigma^2+K_c-1}\left(1-r\sum_{j\notin D}g_{c,j}\beta_{-c,j}\right)\sum_{j\in D}\beta_{-c,j} \\ &\quad -r^2\sigma^2\left[\sum_{j\in D}\beta_{-c,j}^2-\frac{1}{\sigma^2+K_c-1}\left(\sum_{j\in D}\beta_{-c,j}\right)^2\right]+C(K_c-1)-C(K_c-1-d) \end{aligned}$$

We know that ΔC_c is a positive number while the payoff difference is negative, thus it must be that the rest of the terms give us a negative enough difference such that the whole payoff

difference is negative.

$$C(K_c - 1) - C(K_c - 1 - d) < \left(1 - r \sum_{j \notin D} g_{c,j} \beta_{-c,j}\right)^2 \left(\frac{\sigma^2}{\sigma^2 + K_c - d - 1} - \frac{\sigma^2}{\sigma^2 + K_c - 1}\right) \\ + 2 \frac{r\sigma^2}{\sigma^2 + K_c - 1} \left(1 - r \sum_{j \notin D} g_{c,j} \beta_{-c,j}\right) \sum_{j \in D} \beta_{-c,j} + r^2 \sigma^2 \left[\sum_{j \in D} \beta_{-c,j}^2 - \frac{1}{\sigma^2 + K_c - 1} \left(\sum_{j \in D} \beta_{-c,j}\right)^2\right]$$

Following similar steps for player a , we have:

$$\Delta \Pi_a = \underbrace{-\frac{\sigma^2 \left(1 - r \sum_{j=0}^n g_{a,j} \beta_{-a,j}\right)^2}{\sigma^2 + K_a - 1}}_{\text{Without tapping}} + \underbrace{\frac{\sigma^2 \left(1 - r \sum_{j=0}^n g_{a,j} \beta_{-a,j} - r \sum_{j \in D} \beta_{-a,j}\right)^2}{\sigma^2 + K_a + d - 1}}_{\text{Tapping into b}} \\ - \underbrace{r^2 \sigma^2 \left(\sum_{j=0}^n (1 - g_{a,j}) \beta_{-a,j}^2\right)}_{\text{Without tapping}} + \underbrace{r^2 \sigma^2 \left(\sum_{j=0}^n (1 - g_{a,j}) \beta_{-a,j}^2 - \sum_{j \in D} \beta_{-a,j}^2\right)}_{\text{Tapping into b}} + \underbrace{C(K_a + d - 1) - C(K_a - 1)}_{\text{Cost difference of looking to b}} \\ \Delta \Pi_a = \left(1 - r \sum_{j \notin D} g_{a,j} \beta_{-a,j}\right)^2 \left(\frac{\sigma^2}{\sigma^2 + K_a + d - 1} - \frac{\sigma^2}{\sigma^2 + K_a - 1}\right) - 2 \frac{r\sigma^2}{\sigma^2 + K_a - 1} \left(1 - r \sum_{j \notin D} g_{a,j} \beta_{-a,j}\right) \sum_{j \in D} \beta_{-a,j} \\ + \frac{r^2 \sigma^2}{\sigma^2 + K_a + d - 1} \left(\sum_{j \in D} \beta_{-a,j}\right)^2 - r^2 \sigma^2 \sum_{j \in D} \beta_{-a,j}^2 + C(K_a + d - 1) - C(K_a - 1) \\ \Delta \Pi_a = \left(1 - r \sum_{j \notin D} g_{a,j} \beta_{-a,j}\right)^2 \left(\frac{\sigma^2}{\sigma^2 + K_a + d - 1} - \frac{\sigma^2}{\sigma^2 + K_a - 1}\right) - 2 \frac{r\sigma^2}{\sigma^2 + K_a - 1} \left(1 - r \sum_{j \notin D} g_{a,j} \beta_{-a,j}\right) \sum_{j \in D} \beta_{-a,j} \\ - r^2 \sigma^2 \left[\sum_{j \in D} \beta_{-a,j}^2 - \frac{1}{\sigma^2 + K_a + d - 1} \left(\sum_{j \in D} \beta_{-a,j}\right)^2\right] + C(K_a + d - 1) - C(K_a - 1)$$

Before we proceed, observe that

$$\frac{1}{d} \sum_{j \in D} \beta_{-a,j}^2 - \frac{1}{d} \frac{1}{\sigma^2 + K_a + d - 1} \left(\sum_{j \in D} \beta_{-a,j}\right)^2 > \frac{1}{d} \sum_{j \in D} \beta_{-a,j}^2 - \frac{1}{d^2} \left(\sum_{j \in D} \beta_{-a,j}\right)^2 \\ = \frac{1}{d} \sum_{j \in D} \beta_{-a,j}^2 - \left(\frac{\sum_{j \in D} \beta_{-a,j}}{d}\right)^2 = \text{Var}_{j \in D} \beta_{-a,j} \geq 0$$

We know that ΔC_a is a positive number and the difference in payoff between continuing to not form those D connections and forming them is also positive. However, all of the other elements are negative numbers. Thus it must be that the difference in linking cost is enough to counterbalance the benefits of a connection.

$$\begin{aligned}
C(K_a + d - 1) - C(K_a - 1) &> \left(1 - r \sum_{j \notin D} g_{a,j} \beta_{-a,j}\right)^2 \left(\frac{\sigma^2}{\sigma^2 + K_a - 1} - \frac{\sigma^2}{\sigma^2 + K_a + d - 1}\right) \\
&\quad + 2 \frac{r\sigma^2}{\sigma^2 + K_a - 1} \left(1 - r \sum_{j \notin D} g_{a,j} \beta_{-a,j}\right) \sum_{j \in D} \beta_{-a,j} + r^2 \sigma^2 \left[\sum_{j \in D} \beta_{-a,j}^2 - \frac{1}{\sigma^2 + K_a + d - 1} \left(\sum_{j \in D} \beta_{-a,j}\right)^2\right]
\end{aligned}$$

Given that $K_a + d < K_c$, by weak convexity of the cost curve, we have that

$$C(K_a + d - 1) - C(K_a - 1) < C(K_c - 1) - C(K_c - d - 1)$$

This inequality disciplines how the expressions we have just obtained relate.

$$\begin{aligned}
&\left(1 - r \sum_{j \notin D} g_{c,j} \beta_{-c,j}\right)^2 \left(\frac{\sigma^2}{\sigma^2 + K_c - d - 1} - \frac{\sigma^2}{\sigma^2 + K_c - 1}\right) \\
&\quad + 2 \frac{r\sigma^2}{\sigma^2 + K_c - 1} \left(1 - r \sum_{j \notin D} g_{c,j} \beta_{-c,j}\right) \sum_{j \in D} \beta_{-c,j} + r^2 \sigma^2 \left[\sum_{j \in D} \beta_{-c,j}^2 - \frac{1}{\sigma^2 + K_c - 1} \left(\sum_{j \in D} \beta_{-c,j}\right)^2\right] \\
&> \\
&C(K_c - 1) - C(K_c - d - 1) \\
&> \\
&C(K_a + d - 1) - C(K_a - 1) \\
&> \\
&\left(1 - r \sum_{j \notin D} g_{a,j} \beta_{-a,j}\right)^2 \left(\frac{\sigma^2}{\sigma^2 + K_a - 1} - \frac{\sigma^2}{\sigma^2 + K_a + d - 1}\right) \\
&\quad + 2 \frac{r\sigma^2}{\sigma^2 + K_a - 1} \left(1 - r \sum_{j \notin D} g_{a,j} \beta_{-a,j}\right) \sum_{j \in D} \beta_{-a,j} + r^2 \sigma^2 \left[\sum_{j \in D} \beta_{-a,j}^2 - \frac{1}{\sigma^2 + K_a + d - 1} \left(\sum_{j \in D} \beta_{-a,j}\right)^2\right]
\end{aligned}$$

This gives us the following inequality.

$$\begin{aligned}
& \left(1 - r \sum_{j \notin D} g_{c,j} \beta_{-c,j}\right)^2 \left(\frac{\sigma^2}{\sigma^2 + K_c - d - 1} - \frac{\sigma^2}{\sigma^2 + K_c - 1}\right) \\
& + 2 \frac{r\sigma^2}{\sigma^2 + K_c - 1} \left(1 - r \sum_{j \notin D} g_{c,j} \beta_{-c,j}\right) \sum_{j \in D} \beta_{-c,j} + r^2 \sigma^2 \left[\sum_{j \in D} \beta_{-c,j}^2 - \frac{1}{\sigma^2 + K_c - 1} \left(\sum_{j \in D} \beta_{-c,j}\right)^2 \right] \\
& > \\
& \left(1 - r \sum_{j \notin D} g_{a,j} \beta_{-a,j}\right)^2 \left(\frac{\sigma^2}{\sigma^2 + K_a - 1} - \frac{\sigma^2}{\sigma^2 + K_a + d - 1}\right) \\
& + 2 \frac{r\sigma^2}{\sigma^2 + K_a - 1} \left(1 - r \sum_{j \notin D} g_{a,j} \beta_{-a,j}\right) \sum_{j \in D} \beta_{-a,j} + r^2 \sigma^2 \left[\sum_{j \in D} \beta_{-a,j}^2 - \frac{1}{\sigma^2 + K_a + d - 1} \left(\sum_{j \in D} \beta_{-a,j}\right)^2 \right]
\end{aligned}$$

We will show now that the above inequality never holds, characterizing the contradiction. Observe that the following inequalities hold: (i) $K_c - d > K_a$, (ii) $\beta_{-a,j} = \beta_{-c,j} + \frac{1}{n-1} \lambda_{c,j} > \beta_{-c,j}$, $\forall j \in D$, (iii) $\sum_{j \notin D} g_{c,j} \beta_{-c,j} \geq \sum_{j \notin D} g_{a,j} \beta_{-a,j}$. The first two are trivial, while the third is a result of lemma 4. However, jointly, these three inequalities imply that:

$$\left(1 - r \sum_{j \notin D} g_{c,j} \beta_{-c,j}\right)^2 \left(\frac{\sigma^2}{\sigma^2 + K_c - d - 1} - \frac{\sigma^2}{\sigma^2 + K_c - 1}\right) < \left(1 - r \sum_{j \notin D} g_{a,j} \beta_{-a,j}\right)^2 \left(\frac{\sigma^2}{\sigma^2 + K_a - 1} - \frac{\sigma^2}{\sigma^2 + K_a + d - 1}\right)$$

and also

$$2 \frac{r\sigma^2}{\sigma^2 + K_c - 1} \left(1 - r \sum_{j \notin D} g_{c,j} \beta_{-c,j}\right) \sum_{j \in D} \beta_{-c,j} < 2 \frac{r\sigma^2}{\sigma^2 + K_a - 1} \left(1 - r \sum_{j \notin D} g_{a,j} \beta_{-a,j}\right) \sum_{j \in D} \beta_{-a,j}$$

Finally, we can work with the last term so that:

$$\begin{aligned}
& \left[\sum_{j \in D} \beta_{-c,j}^2 - \frac{1}{\sigma^2 + K_c - 1} \left(\sum_{j \in D} \beta_{-c,j}\right)^2 \right] - \left[\sum_{j \in D} \beta_{-a,j}^2 - \frac{1}{\sigma^2 + K_a + d - 1} \left(\sum_{j \in D} \beta_{-a,j}\right)^2 \right] \\
& = \sum_{j \in D} \beta_{-c,j}^2 - \sum_{j \in D} \beta_{-a,j}^2 + \frac{1}{\sigma^2 + K_a + d - 1} \left(\sum_{j \in D} \beta_{-a,j}\right)^2 - \frac{1}{\sigma^2 + K_c - 1} \left(\sum_{j \in D} \beta_{-c,j}\right)^2 \\
& = \sum_{j \in D} \beta_{-c,j}^2 - \sum_{j \in D} \beta_{-a,j}^2 + \frac{1}{d_z} \left[\left(\sum_{j \in D} \beta_{-a,j}\right)^2 - \left(\sum_{j \in D} \beta_{-c,j}\right)^2 \right]
\end{aligned}$$

where $\frac{1}{d_z} = \frac{\frac{1}{\sigma^2 + K_a + d - 1} (\sum_{j \in D} \beta_{-a,j})^2 - \frac{1}{\sigma^2 + K_c - 1} (\sum_{j \in D} \beta_{-c,j})^2}{(\sum_{j \in D} \beta_{-a,j})^2 - (\sum_{j \in D} \beta_{-c,j})^2} > 0$.

Dividing by d_z the expression, we have:

$$\begin{aligned}
& \frac{1}{d_z} \sum_{j \in D} (\beta_{-c,j}^2 - \beta_{-a,j}^2) - \frac{1}{d_z^2} \left[\left(\sum_{j \in D} \beta_{-c,j} \right)^2 - \left(\sum_{j \in D} \beta_{-a,j} \right)^2 \right] \\
&= \left[\frac{1}{d_z} \sum_{j \in D} \beta_{-c,j}^2 - \left(\frac{1}{d_z} \sum_{j \in D} \beta_{-c,j} \right)^2 \right] - \left[\frac{1}{d_z} \sum_{j \in D} \beta_{-a,j}^2 - \left(\frac{1}{d_z} \sum_{j \in D} \beta_{-a,j} \right)^2 \right] \\
&= \left[\frac{1}{d_z} \sum_{j \in D} \left(\beta_{-c,j} - \frac{1}{d_z} \sum_{s \in D} \beta_{-c,s} \right)^2 \right] - \left[\frac{1}{d_z} \sum_{j \in D} \left(\beta_{-a,j} - \frac{1}{d_z} \sum_{s \in D} \beta_{-a,s} \right)^2 \right] \\
&= \left[\frac{1}{d_z} \sum_{j \in D} \left(\beta_{-c,j} - \frac{1}{d_z} \sum_{s \in D} \beta_{-c,s} \right)^2 \right] - \left[\frac{1}{d_z} \sum_{j \in D} \left(\beta_{-c,j} + \frac{1}{n-1} \lambda_{c,j} - \frac{1}{d_z} \sum_{s \in D} \left(\beta_{-c,s} + \frac{1}{n-1} \lambda_{c,s} \right) \right)^2 \right] \\
&= \left[\frac{1}{d_z} \sum_{j \in D} \left(\beta_{-c,j} - \frac{1}{d_z} \sum_{s \in D} \beta_{-c,s} \right)^2 \right] - \left[\frac{1}{d_z} \sum_{j \in D} \left(\beta_{-c,j} - \frac{1}{d_z} \sum_{s \in D} \beta_{-c,s} + \frac{1}{n-1} \lambda_{c,j} - \frac{1}{d_z} \sum_{s \in D} \left(\frac{1}{n-1} \lambda_{c,s} \right) \right)^2 \right] \\
&= \left[\frac{1}{d_z} \sum_{j \in D} \left(\beta_{-c,j} - \frac{1}{d_z} \sum_{s \in D} \beta_{-c,s} \right)^2 \right] - \left[\frac{1}{d_z} \sum_{j \in D} \left(\beta_{-c,j} - \frac{1}{d_z} \sum_{s \in D} (\beta_{-c,s}) \right)^2 \right] \\
&\quad - \left[\frac{1}{d_z} \sum_{j \in D} \left(\frac{1}{n-1} \lambda_{c,j} - \frac{1}{d_z} \sum_{s \in D} \left(\frac{1}{n-1} \lambda_{c,s} \right) \right)^2 \right] - 2 \left[\frac{1}{d_z} \sum_{j \in D} \left(\beta_{-c,j} - \frac{1}{d_z} \sum_{s \in D} (\beta_{-c,s}) \right) \left(\frac{1}{n-1} \lambda_{c,j} - \frac{1}{d_z} \sum_{s \in D} \left(\frac{1}{n-1} \lambda_{c,s} \right) \right) \right] \\
&= - \left[\frac{1}{d_z} \sum_{j \in D} \left(\frac{1}{n-1} \lambda_{c,j} - \frac{1}{d_z} \sum_{s \in D} \left(\frac{1}{n-1} \lambda_{c,s} \right) \right)^2 \right] - 2 \left[\frac{1}{d_z} \sum_{j \in D} \left(\beta_{-c,j} - \frac{1}{d_z} \sum_{s \in D} (\beta_{-c,s}) \right) \left(\frac{1}{n-1} \lambda_{c,j} - \frac{1}{d_z} \sum_{s \in D} \left(\frac{1}{n-1} \lambda_{c,s} \right) \right) \right] \\
&\leq -2 \left[\frac{1}{d_z} \sum_{j \in D} \left(\beta_{-c,j} - \frac{1}{d_z} \sum_{s \in D} (\beta_{-c,s}) \right) \left(\frac{1}{n-1} \lambda_{c,j} - \frac{1}{d_z} \sum_{s \in D} \left(\frac{1}{n-1} \lambda_{c,s} \right) \right) \right] \\
&= -2 \frac{d}{d_z} \text{Cov}_{j \in D} \left(\beta_{-c,j}, \frac{1}{n-1} \lambda_{c,j} \right)
\end{aligned}$$

< 0 by Lemma 5

This finally characterizes our contradiction.

E Origins of Leadership

In this subsection, we proof proposition 6, which characterizes how heterogeneity among agents determines the agent’s position on the information structure. In the main text we considered three dimensions in which agents may differ: (i) precision of individual signal, (ii) cost structure, and (iii) resoluteness.

We provide first the proof for resoluteness, which is easily extended to different cost structures. As discussed in the main text, we focus on the case in which the cost of making more than one connection is prohibitive, thus agents will tap into at most one other signal. Let the cost of tapping into one signal be c . For simplicity, we consider all agents to have the same resoluteness, $r = r_h$, except one of them, who is more resolute $r = r_l < r_h$. The proof for this case can be expanded to a general distribution of r_i , with the introduction of a lot of notation. Also, we consider the case in which each individual signal is equally precise, that is σ is the same for all agents. For simplicity, we present the proof for the case of $\sigma = 1$. The proof can be extended in the obvious way.

There are two information configurations in which a leader emerges. In the first, all players tap into the signal of one player—player 1—which does not tap into any other player’s signal. In the second, all players tap into the signal of one player—player 1—who also taps into the signal of another player. Call the first one \mathcal{A} , and the second one \mathcal{B} . For each configuration, there are two possibilities, (1) either the central player is not the more resolute player, or (2) it is. We can now restate the theorem: If there exists an equilibrium of type $\mathcal{A}1$ or $\mathcal{B}1$, then it exists one of type $\mathcal{A}2$ or $\mathcal{B}2$.

The proof will follow a non-obvious path. We will first show that if the cost is large, the existence of an equilibrium of type $\mathcal{A}1$ implies one of type $\mathcal{A}2$. All players have coordinated to observe a player, thus we show that if all were to coordinate on observing the more resolute player no one would have incentives to deviate. The proof only holds for large enough costs because we cannot determine the effect over the more resolute player. But there is always a high enough cost such that if all other players were to coordinate into tapping into her signal, she would not choose to tap into another signal.

We will then show that for very small costs c , the existence of an equilibrium of type $\mathcal{B}1$ implies one of type $\mathcal{B}2$. Given that all players are tapping into some signal, all what’s left to show is that if all players were to coordinate in taping the more resolute agent’s signal instead, no one would have incentives to deviate. This concern only small costs, because we cannot determine the effect over the more resolute player. That is, if all other players were to coordinate and tap into her signal, would she still be interested in taping into a signal? For sufficiently small costs, the answer is always yes.

The hard part of the proof resides in the case when the cost is neither so high nor so low. When the cost is in this endogenously defined mid-range we have to follow a more creative path with the proof. We show that the cost values for which there exists an equilibrium of type $\mathcal{A}2$ or $\mathcal{B}2$ is convex. We first establish that starting from a configuration $\mathcal{B}2$, as cost increases the central player is the first to choose to deviate and break a link. We also have that starting from a configuration $\mathcal{A}2$, as cost decreases the central player is the first to choose to deviate and form a link. The final step is to consider the more-resolute-player’s incentives to tap into a signal when she is the center. We show that if all players are considering that she will tap into another signal, then she has more incentives to do so than if the players are not. That is, for certain costs,

there exist both $\mathcal{A}2$ and $\mathcal{B}2$ equilibria.

First, consider configuration \mathcal{A} . Note that there are two types of agents, the center—player 1—and the periphery—denoted as player s . Notice that how much an agent considers the other agent's signal, $\lambda_{i,j}$ depends on the position of the other agent on the network as well as on the resoluteness.

	$\mathcal{A}1$	$\mathcal{A}2$
$\lambda_{1,0}$	$\frac{1}{2}$	$\frac{1}{2}$
$\lambda_{1,1}$	$\frac{1}{2}$	$\frac{1}{2}$
$\lambda_{s,1}$	$\frac{1}{3} + \frac{r_s}{3}\beta_{-s,1}$	$\frac{1}{3} + \frac{r_s}{3}\beta_{-s,1}$
$\lambda_{s,0}$	$\frac{1}{3} + \frac{r_s}{3}\beta_{-s,1}$	$\frac{1}{3} + \frac{r_s}{3}\beta_{-s,1}$
$\lambda_{s,s}$	$\frac{1}{3} - \frac{2r_s}{3}\beta_{-s,1}$	$\frac{1}{3} - \frac{2r_s}{3}\beta_{-s,1}$

Given that the signal of player 1 is observed by all players, and that is common knowledge, it works as a public signal. In this case it has the same precision of the prior, and thus shares the same consideration. Any difference in precision between the prior and an individual signal would simply mean a proportional adjustment between the considerations. The difference between the two columns resides on the r_s . In the first, the more resolute player is one of the s -type players, while on the second all players of type s have the same resoluteness. This difference should be clear in the recursively formula for the influence of an agent: $(n-1)\beta_{-i,j} = \sum_{k \neq i} \lambda_{k,j}$.

Observe that, given $\sigma = 1$, $\beta_{-i,0} = \beta_{-i,1}$, $\forall i$. Again, any difference in precision between the equilibrium public signal, e_1 , and the prior will translate into a proportionality factor. Also, note that $\beta_{-s,s} = 0$ for any non-center player s . For notation simplicity, in situation $\mathcal{A}1$ let the more resolute player be player n .

	$\mathcal{A}1$	$\mathcal{A}2$
$(n-1)\beta_{-1,1}$	$\frac{n-1}{3} + \frac{(n-2)r_h}{3}\beta_{-s,1} + \frac{r_l}{3}\beta_{-n,1}$	$\frac{n-1}{3} + \frac{(n-1)r_h}{3}\beta_{-s,1}$
$(n-1)\beta_{-s,1}$	$\frac{n-2}{3} + \frac{1}{2} + \frac{(n-3)r_h}{3}\beta_{-s,1} + \frac{r_l}{3}\beta_{-n,1}$	$\frac{n-2}{3} + \frac{1}{2} + \frac{(n-2)r_h}{3}\beta_{-s,1}$
$(n-1)\beta_{-n,1}$	$\frac{n-2}{3} + \frac{1}{2} + \frac{(n-2)r_h}{3}\beta_{-s,1}$	—

This presents us with two systems of linear equations. Solving it gives us:

	$\mathcal{A}1$	$\mathcal{A}2$
$\beta_{-s,1}$	$\frac{(3(n-1)+r_l)(n-\frac{1}{2})}{3(n-1)(3(n-1)-(n-3)r_h)-r_l r_h(n-2)}$	$\frac{(3(n-1)+r_h)(n-\frac{1}{2})}{3(n-1)(3(n-1)-(n-3)r_h)-r_h r_h(n-2)}$

To show that the term on the second column is strictly larger than the term on the first, it suffices to show that the derivative of the term on the first column with respect to r_l is strictly positive. Thus $\beta_{-s,1}^{\mathcal{A}1} < \beta_{-s,1}^{\mathcal{A}2}$. Notice that the incentive for player s to look at the central player is given by $\Pi_s(\text{look}) - \Pi_s(\text{not look})$.

$$\Pi_s(\text{look}) - \Pi_s(\text{nl}) = - \left[\frac{1}{3}(1 - 2r_s\beta_{-s,1})^2 + (n-2)r_s^2\beta_{-s,k}^2 + r_s^2\beta_{-s,n}^2 \right] + \left[\frac{1}{2}(1 - r_s\beta_{-s,1})^2 + r_s^2\beta_{-s,k}^2 + (n-2)r_s^2\beta_{-s,k}^2 + r_s^2\beta_{-s,n}^2 \right] - c$$

$$\Pi_s(\text{look}) - \Pi_s(\text{nl}) = \left(\frac{1}{6} + \frac{1}{3}r_s + \frac{1}{6}r_s^2 \right) \beta_{-s,1} - c$$

Thus, the incentive to look to the central player is higher under configuration $\mathcal{A}2$ than under configuration $\mathcal{A}1$. This establishes the first part of the proof. If the cost is large enough that the central player does not want to form a link, then if there exists an equilibrium in which all players tap into the signal of a leader there exists an equilibrium in which the more resolute player is the leader.

Let us now focus on the \mathcal{B} configuration. There are three different positions an agent can have in this configuration. A player can be the leader, to whom all other players pay attention, call it player 1; the player the center is observing, player n , or she can be observing the center while the center is not observing her. In configuration $\mathcal{B}2$, the more resolute player is in the leader, while in configuration $\mathcal{B}1$ she is the n^{th} player. We proceed to construct the influences of each player in each configuration, but first note that again $\lambda_{i,1} = \lambda_{i,0}$ and $\beta_{-i,1} = \beta_{-i,0}$, $\forall i$, and also that for any player s whose signal is not observed by any other player, $\beta_{-s,s} = 0$. Finally, note that $\beta_{-1,s} = \beta_{-n,s} = \beta_{-k,s} \quad \forall k \neq s$.

	$\mathcal{B}1$	$\mathcal{B}2$
$\lambda_{1,1}$	$\frac{1}{3} + \frac{r_h}{3}\beta_{-1,1} - \frac{r_l}{3}\beta_{-1,n}$	$\frac{1}{3} + \frac{r_l}{3}\beta_{-1,1} - \frac{r_l}{3}\beta_{-1,n}$
$\lambda_{1,n}$	$\frac{1}{3} - 2\frac{r_h}{3}\beta_{-1,1} + 2\frac{r_h}{3}\beta_{-1,n}$	$\frac{1}{3} - 2\frac{r_l}{3}\beta_{-1,1} + 2\frac{r_l}{3}\beta_{-1,n}$
$\lambda_{s,1}$	$\frac{1}{3} + \frac{r_l}{3}\beta_{-s,1}$	$\frac{1}{3} + \frac{r_l}{3}\beta_{-s,1}$
$\lambda_{s,s}$	$\frac{1}{3} - 2\frac{r_l}{3}\beta_{-s,1}$	$\frac{1}{3} - 2\frac{r_l}{3}\beta_{-s,1}$
$\lambda_{n,1}$	$\frac{1}{3} + \frac{r_h}{3}\beta_{-n,1} - \frac{r_l}{3}\beta_{-n,n}$	$\frac{1}{3} + \frac{r_h}{3}\beta_{-n,1} - \frac{r_h}{3}\beta_{-n,n}$
$\lambda_{n,n}$	$\frac{1}{3} - 2\frac{r_l}{3}\beta_{-n,1} + 2\frac{r_l}{3}\beta_{-n,n}$	$\frac{1}{3} - 2\frac{r_h}{3}\beta_{-n,1} + 2\frac{r_h}{3}\beta_{-n,n}$

We can now proceed to construct the influence recursively, by using the fact that $(n-1)\beta_{-i,j} = \sum_{k \neq i} \lambda_{k,j}$.

	$\mathcal{B}1$	$\mathcal{B}2$
$(n-1)\beta_{-1,1}$	$\frac{n-1}{3} + (n-2)\frac{r_h}{3}\beta_{-s,1} + \frac{r_l}{3}(\beta_{-n,1} - \beta_{-n,n})$	$\frac{n-1}{3} + (n-2)\frac{r_h}{3}\beta_{-s,1} + \frac{r_h}{3}(\beta_{-n,1} - \beta_{-n,n})$
$(n-1)\beta_{-1,n}$	$\frac{1}{3} - 2\frac{r_l}{3}(\beta_{-n,1} - \beta_{-n,n})$	$\frac{1}{3} - 2\frac{r_h}{3}(\beta_{-n,1} - \beta_{-n,n})$
$(n-1)\beta_{-1,s}$	$\frac{1}{3} - 2\frac{r_h}{3}\beta_{-s,1}$	$\frac{1}{3} - 2\frac{r_h}{3}\beta_{-s,1}$
$(n-1)\beta_{-s,1}$	$\frac{n-1}{3} + (n-3)\frac{r_h}{3}\beta_{-s,1} + \frac{r_h}{3}(\beta_{-1,1} - \beta_{-1,n}) + \frac{r_l}{3}(\beta_{-n,1} - \beta_{-n,n})$	$\frac{n-1}{3} + (n-3)\frac{r_h}{3}\beta_{-s,1} + \frac{r_l}{3}(\beta_{-1,1} - \beta_{-1,n}) + \frac{r_h}{3}(\beta_{-n,1} - \beta_{-n,n})$
$(n-1)\beta_{-n,1}$	$\frac{n-1}{3} + (n-2)\frac{r_h}{3}\beta_{-s,1} + \frac{r_h}{3}(\beta_{-1,1} - \beta_{-1,n})$	$\frac{n-1}{3} + (n-2)\frac{r_h}{3}\beta_{-s,1} + \frac{r_l}{3}(\beta_{-1,1} - \beta_{-1,n})$
$(n-1)\beta_{-n,n}$	$\frac{1}{3} - 2\frac{r_h}{3}(\beta_{-1,1} - \beta_{-1,n})$	$\frac{1}{3} - 2\frac{r_l}{3}(\beta_{-1,1} - \beta_{-1,n})$

Once more, this gives us two systems of linear equations, now with six equations each. Solving the systems gives us:

	$\mathcal{B}1$	$\mathcal{B}2$
$\beta_{-s,1}$	$\frac{3(n-1)(n-1-\frac{r_l r_h}{n-1})+(n-2)[r_h+r_l+2\frac{r_n r_l}{n-1}]}{3(n-1-\frac{r_l r_h}{n-1})[3(n-1)-(n-3)r]-(n-2)r_l[r_l+r_h+2\frac{r_n r_l}{n-1}]}$	$\frac{3(n-1)(n-1-\frac{r_l r_h}{n-1})+(n-2)[r_h+r_l+2\frac{r_n r_l}{n-1}]}{3(n-1-\frac{r_l r_h}{n-1})[3(n-1)-(n-3)r]-(n-2)r_l[r_l+r_h+2\frac{r_n r_l}{n-1}]}$

Yes, it is indeed the same expression in both columns. This result is quite surprising. For a player that is not being looked by anyone, it is irrelevant whether the leader is more or less resolute. The intuition behind this result is that, whether the leader is more or less resolute, she is taping into the signal of the more resolute one, and the more resolute is taping into the signal of the less resolute. Thus both players will be influenced by the signal of the leader averages out, independently of who the leader actually is.

This shows that $\beta_{-s,1}^{\mathcal{B}1} \leq \beta_{-s,1}^{\mathcal{B}2}$. Note that the incentive for player s to look at the central player is given by $\Pi_s(\text{look}) - \Pi_s(\text{not look})$, calculated above. Finally, this implies that for sufficiently small costs such that the central player wants to tap into another signal, if there exists an equilibrium of type $\mathcal{B}1$, then there must exist one of type $\mathcal{B}2$.

All what's left now is to work with costs c in the mid-range not covered by the two cases above. For that, we need to compare $\mathcal{A}2$ with $\mathcal{B}2$.

First, we simplify the expressions for $\beta_{-s,1}^{\mathcal{A}2}$ and $\beta_{-s,1}^{\mathcal{B}2}$ provided above. First, $\beta_{-s,1}^{\mathcal{A}2}$, simplifies to:

$$\beta_{-s,1}^{\mathcal{A}2} = \frac{n - 1 + \frac{1}{2}}{3(n - 1) - (n - 3)r_h - r_h}$$

While for $\beta_{-s,1}^{\mathcal{B}2}$, the expression is a bit more complicated:

$$\beta_{-s,1}^{\mathcal{B}2} = \frac{n - 1 + (n - 2) \frac{(n-1)(r_l+r_h)+2r_l r_h}{3(n-1)^2-3r_l r_h}}{3(n - 1) - (n - 3)r - r(n - 2) \frac{(n-1)(r_l+r_h)+2r_l r_h}{3(n-1)^2-3r_l r_h}}$$

Let $(n - 2) \frac{(n-1)(r_l+r_h)+2r_l r_h}{3(n-1)^2-3r_l r_h} := z$, then we can compare $\beta_{-s,1}^{\mathcal{B}2}$ with $\beta_{-s,1}^{\mathcal{A}2}$.

$$z < \frac{1}{2} + \frac{r_h}{6} \frac{n - \frac{1}{2}}{n - 1 + \frac{r_h}{2}} \implies \beta_{-s,1}^{\mathcal{B}2} < \beta_{-s,1}^{\mathcal{A}2}$$

We will show then that $z \leq \frac{1}{2} + \frac{r_h}{6} < \frac{1}{2} + \frac{r_h}{6} \frac{n - \frac{1}{2}}{n - 1 + \frac{r_h}{2}}$.

$$(n - 2) \frac{(n - 1)(r_l + r_h) + 2r_l r_h}{3(n - 1)^2 - 3r_l r_h} < \frac{3 + r_h}{6}$$

$$\frac{(n - 2)(n - 1)(r_l + r_h) + 2(n - 2)r_l r_h}{(n - 1)^2 - r_l r_h} < \frac{3 + r_h}{2}$$

$$2(n - 2)(n - 1)(r_l + r_h) + 4(n - 2)r_l r_h < 3(n - 1)^2 - 3r_l r_h + (n - 1)^2 r_h - r_l r_h^2$$

$$2(n - 2)(n - 1)(r_l + r_h) + 4(n - 2 + \frac{3}{4})r_l r_h + r_l r_h^2 - (n - 1)^2 r_h < 3(n - 1)^2$$

Since $n \geq 3$, $2(n - 2) \geq n - 1$, and thus $2(n - 2)(n - 1) > (n - 1)^2$. This implies that the expression on the left-hand-side is increasing in both r_l and r_h . Evaluating the expression at $r_l = r_h = 1$ gives us.

$$4(n - 2)(n - 1) + 4(n - 2 + \frac{3}{4}) + 1 \leq 4(n - 1)^2$$

$$4(n - 2)(n - 1) + 4(n - 1) \leq 4(n - 1)^2$$

$$4(n - 2)(n - 1) + 4(n - 1) \leq 4(n - 1)^2$$

Note that the expression on both sides is the same. Thus $\beta_{-s,1}^{\mathcal{B}2} < \beta_{-s,1}^{\mathcal{A}2}$.

So far, we have only shown a relationship between $\beta_{-s,1}$ in both configurations. It is clear however that the incentives for player 1 to tap into another agent are not a function of $\beta_{-s,1}$. One would guess that it is a function of $\beta_{-1,1}$ and $\beta_{-1,n}$.

$$\Pi_1(l) - \Pi_1(nl) = - \left[\frac{1}{3} (1 - 2r_l \beta_{-1,1} - r_l \beta_{-1,n})^2 + (n-2)r_l^2 \beta_{-1,k}^2 \right] + \left[\frac{1}{2} (1 - 2r_l \beta_{-1,1})^2 + r_l^2 \beta_{-1,n}^2 + (n-2)r_l^2 \beta_{-1,k}^2 \right] - c$$

$$\Pi_1(l) - \Pi_1(nl) = \frac{1}{6} (1 - 2r_l (\beta_{-1,1} - \beta_{-1,n}))^2 - c$$

Thus we can compare $\beta_{-s,1}$ in both configurations, but need actually to compare $\beta_{-1,1} - \beta_{-1,n}$. The solution to our systems of linear equations comes to our rescue.

$$\begin{array}{l|l} \beta_{-n,1} - \beta_{-n,n} & \begin{array}{c} \mathcal{A2} \\ \frac{n-2+\frac{3}{2}+(n-2)r_h \beta_{-s,1}^{\mathcal{A2}}}{3(n-1)} \end{array} \\ (n-1)\beta_{-1,1} & \frac{1}{3} + (n-2)\frac{r_h}{3}\beta_{-s,1}^{\mathcal{A2}} + \frac{r}{3}(\beta_{-n,1}^{\mathcal{A2}} - \beta_{-n,n}^{\mathcal{A2}}) \\ (n-1)\beta_{-1,n} & \frac{1}{3} - 2\frac{r}{3}(\beta_{-n,1}^{\mathcal{A2}} - \beta_{-n,n}^{\mathcal{A2}}) \end{array} \quad \begin{array}{l|l} \mathcal{B2} \\ \frac{n-2+(n-2)r_h \beta_{-s,1}^{\mathcal{B2}}}{3(n-1)+r_l \frac{n-1+r_h}{n-1+r_l}} \\ \frac{1}{3} + (n-2)\frac{r_h}{3}\beta_{-s,1}^{\mathcal{B2}} + \frac{r}{3}(\beta_{-n,1}^{\mathcal{B2}} - \beta_{-n,n}^{\mathcal{B2}}) \\ \frac{1}{3} - 2\frac{r}{3}(\beta_{-n,1}^{\mathcal{B2}} - \beta_{-n,n}^{\mathcal{B2}}) \end{array}$$

Given that $\beta_{-s,1}^{\mathcal{B2}} < \beta_{-s,1}^{\mathcal{A2}}$, we have that $\beta_{-n,1}^{\mathcal{B2}} - \beta_{-n,n}^{\mathcal{B2}} < \beta_{-n,1}^{\mathcal{A2}} - \beta_{-n,n}^{\mathcal{A2}}$. These two facts imply that the incentives for player 1 to tap into player n 's signal is smaller at $\mathcal{A2}$ than at $\mathcal{B2}$. This result is quite intuitive. If other players expect player 1 to tap into player n 's signal, they will consider player 1 to be less influenced by her own signal. This makes her own signal less influential and less informative of the average action, thus increasing the payoff for player 1 to tap into player n 's signal.

The final step is to show that as cost increases, player 1 is the first one to deviate and stop tapping into another player's signal. We have already computed $\Pi_1(l) - \Pi_1(nl)$, and can see that $\Pi_1(l) - \Pi_1(nl) < \frac{1}{6} - c$. Let's now proceed to compute the payoff of player n deviating and not observing player 1's signal. Player n 's incentive to do so is obviously higher than any other player not in the center, since player's n signal is more informative of the average action.

$$\Pi_n(l) - \Pi_n(nl) = - \left[\frac{1}{3} (1 - 2r_h \beta_{-n,1} - r_h \beta_{-n,n})^2 + (n-2)r_h^2 \beta_{-n,k}^2 \right] + \left[\frac{1}{2} (1 - r_h \beta_{-n,1} - r_h \beta_{-n,n})^2 + r_h^2 \beta_{-n,1}^2 + (n-2)r_h^2 \beta_{-n,k}^2 \right] - c$$

$$\Pi_n(l) - \Pi_n(nl) = \frac{1}{6} (1 + 2r_h (\beta_{-n,1} - \beta_{-n,n}))^2 - c$$

$$\Pi_n(l) - \Pi_n(nl) > \frac{1}{6} - c$$

Which completes the proof.

Different Costs The proof extends naturally to consider different costs. Indeed, when considering that players differ only regarding their communication skills the analysis is much simpler. In that case the benefit of forming a link is equal to all players—up to the additive cost term.

Let $c_{i,j}$ be the cost of player i observing player j 's signal. In the main text we focus on two particular cases: (1) the good listener $\forall i c_{i,j} = c_i \forall j$, and (2) the good communicator, $\forall j c_{i,j} = c_j \forall i$. While the first imply

that agents differ in their individual costs of acquiring information, the second implies that different information has different costs to be acquired by all agents.

It is clear that whenever a leader emerges in equilibrium, there exists also an equilibrium in which the leader will be the agent who is a good communicator. Even though if agents were to coordinate in any leader the benefit would be the same, to coordinate in the good communicator is cheaper. This equilibrium will always exist when the cost is low enough, and furthermore for any parameter specification (n, σ, r) there always exists a cost such that it is the unique equilibrium to coordinate in the good communicator.

Finally, note that being a good listener does not change the incentive of others to coordinate on your signal. That is, the benefit for other players to coordinate on your information is the same as to coordinate in any other, thus it is not true that a good listener will always be selected as the leader.