Designing Communication Hierarchies to Elicit Information*

Dimitri Migrow†

Abstract

This paper studies how a manager can elicit employees' information by designing a hierarchical communication network. The manager decides who communicates with whom, and in which order, where communication takes the form of cheap talk (Crawford and Sobel, 1982). I show that the optimal network is shaped by two competing forces: an *intermediation* force that calls for grouping employees together and an *uncertainty* force that favours separating them. The manager optimally divides employees into groups of similar bias. Under simple conditions, the optimal network features a single intermediary who communicates directly to the manager.

*I am grateful to my advisors - Francesco Squintani, Dan Bernhardt, Marina Halac and Peter Hammond. I benefited from discussions with Ricardo Alonso, Ken Binmore, Vincent Crawford, Bhaskar Dutta, Andrew Harkins, Sergiu Hart, Navin Kartik, Jacob Leshno, Niko Matouschek, Meg Meyer, Motty Perry, Herakles Polemarchakis, Debraj Ray, Phil Reny, Ariel Rubinstein and Julia Wirtz. I thank the seminar participants in Autonoma (Barcelona), Calgary, HEC Paris, Kellogg School of Business (Northwestern), Maastricht, Manchester, Oxford (Nuffield), Vienna and Warwick. All mistakes are my own.

†University of Manchester, Department of Economics. E-mail: dimitri.migrow@gmail.com

1 Introduction

> *"Information in an organization, particularly decision-related information, is rarely innocent, thus rarely as reliable as an innocent person would expect. Most information is subject to strategic misrepresentation..."* James G. March, 1981.

Much of the information relevant for decision making in organizations is typically dispersed among employees. Due to time, location or qualification constraints, management is unable to observe this information directly. Managers aim to collect decision-relevant information from their subordinates, but employees often have their own interests and hence communicate strategically to influence decision making in their favour. In this paper, I study how a manager optimally elicits information by designing a communication structure within the organization. The manager commits to a hierarchical network that specifies who communicates with whom, and in which order. Her objective is to maximize information transmission.[1]

My analysis shows that the optimal communication network is shaped by two competing forces: an *intermediation force* that calls for grouping employees together and an *uncertainty force* that favours separating them. The manager optimally divides employees into groups of similar bias. Each group has a group leader who collects information directly from the group members and communicates this information in a coarse way to either another group leader or the manager. If employees' biases are sufficiently close to one another and far away from the manager's, the optimal network consists of a single group. My results resonate with the classic studies of Dalton (1959), Crozier (1963), and Cyert and March (1963), who observe that groups — or "cliques" — collect decision-relevant information in organizations and distort this information before communicating it to organization members outside the group.

The model I present considers a decision maker and a set of employees whom I call experts. Each expert observes a noisy signal of a parameter that is relevant for a decision to be made by the decision maker. The decision maker and the experts have

[1] Evidence suggests that the communication structure within an organization indeed affects employees' incentives to reveal their private information. See discussions in Schilling and Fang (2014) and Glaser et al. (2015).

different preferences over this decision; specifically, the experts have biases of arbitrary sign and magnitude over the decision maker's choice. The decision maker does not observe any signal of the relevant parameter and relies on communication with the experts. As committing to transfers or to decisions as a function of the information transmitted is often difficult in an organizational context, I rule these out.[2] The decision maker instead commits to a communication network, which specifies who communicates with whom, and in which order.[3] Communication is direct and costless, i.e. it takes the form of "cheap-talk" as in Crawford and Sobel (1982). I focus on the best equilibrium payoffs for the decision maker in any given communication network and characterize the optimal network for the decision maker.

My model builds upon Galeotti et al. (2013) who study simultaneous communication in a similar setting. The crucial difference is that my model studies the optimal sequential structure from decision maker's perspective, where they focus on the properties of simultaneous communication in different network structures. In particular, I restrict attention to tree communication networks, or "hierarchies." This type of network is a natural starting point in the study of communication in organizations. In the theoretical literature, hierarchies are regarded as the optimal formal organization for reducing the costs of information processing (Sah and Stiglitz, 1987; Radner, 1993; and Garicano, 2000) and for preventing conflicts between subordinates and their superiors (Friebel and Raith, 2004). In practice, hierarchies have been identified as a prominent communication structure in organizations, even in those that aim to have non-hierarchical communication and decision rights allocation (see Ahuja, 2000 and Oberg and Walgenbach, 2008).

I begin my analysis of optimal communication networks by identifying a trade-off between two competing forces. On the one hand, the intermediation force pushes in favour of grouping experts together, in order to enable them to pool privately held information and have more flexibility in communicating to the decision maker. On the other hand, the uncertainty force pushes in favour of separating the experts, in order to increase their uncertainty about the information held by other experts and relax their incentive constraints. As in other contexts, uncertainty allows to pool incentive constraints, so a less informed expert can be better incentivized because fewer constraints have to be satisfied compared to the case of a more informed expert.

[2] A manager cannot contract upon transfers or any information received in Dessein (2002), Alonso et al. (2008), Alonso et al. (2015), and Grenadier (2015). See also the literature discussion in Gibbons et al. (2013).

[3] The design of communication structures appears as a more natural form of commitment. For example, if a party commits not to communicate with an agent, she will ignore any reports from the agent so long as they are not informative, and the agent in turn will not send informative reports as he expects them to be dismissed.

Building upon the interaction between the intermediation force and the uncertainty force, I derive three main results. My first main result concerns star networks — those in which each expert communicates directly to the decision maker. Star networks are a simple and a prominent benchmark in the social network literature (see Jackson, 2008). My analysis shows, however, that a star communication network is always dominated by an optimally-designed sequential communication network. Sequential communication between the experts can generate as much information transmission to the decision maker as a star network, and sometimes strictly more. The improvement arises because coordination in reports gives experts the possibility to report pooled information in a coarse way. This is strictly beneficial for the decision maker whenever the experts would send a less informative report were they unable to coarsen information.

My second main result shows that an optimal communication network consists of "groups" of experts. In a group, a single expert — the group leader — receives direct reports from all other members of the group and then communicates the aggregated information in a coarse way either to another group leader or directly to the decision maker. The coarsening of information by a group leader is key to incentivize the experts to reveal their signals truthfully. As for the optimal composition of a group, I show that group members who only observe their own private signals have identical ranges of biases that support their equilibrium strategies; the reason is that they have the same expected uncertainty about the signals of other experts and their reports are treated symmetrically by their group leader. Consequently, the decision maker benefits from grouping similarly biased experts together.

Finally, my third main result shows that if the experts' biases are sufficiently close to one another while large enough (relative to the decision maker's preferences), then the optimal network consists of a single group. The group leader acts as a single intermediary who aggregates all the information from the other experts and sends a coarse report to the decision maker. Aggregation of the entire information allows this intermediary to send a report with minimal information content. As a consequence, from the perspective of each expert, any deviation from truth-telling results in the largest possible shift in the decision maker's policy from the expected value of the state. This allows to incentivize highly-biased experts to reveal their private information truthfully.

As noted, my findings are in line with work on the modern theory of the firm, which emphasizes the importance of coordination between employees for intra-firm information transmission. Cyert and March (1963) observe that managerial decisions are lobbied by groups of employees that provide distorted information to the authority. Similarly, Dalton (1959) and Crozier (1963) view an organization as a collection of cliques that aim to conceal or distort information in order to reach their goals. Dalton claims that having cliques as producers and regulators of information is essential for the firm, and provides examples of how central management influences the composition of such groups through promotions and replacements.[4] Group leaders in my model also resemble the *internal communication stars* identified in the sociology and management literature. Allen (1977), Tuchman and Scanlan (1981), and Ahuja (2000) describe these stars as individuals who are highly connected and responsible for a large part of information transmission within an organization, often acting as informational bridges between different groups.

The next section discusses the related literature. Section 2 describes the model. Section 3 illustrates the main ideas with a simple example, provides a characterization of the intermediation and uncertainty forces, and derives the main results. Section 4 provides additional results: I study the optimal ordering of biases, the case of experts with opposing biases, the value of commitment, and the benefits and limitations of using non-hierarchical networks. Section 5 concludes.

Related Literature. An early strand of literature on communication within organizations takes a team-theoretic approach which assumes same objectives for all players. Communication is non-strategic and the optimal mechanism minimizes communication and/or information processing costs (starting with Marschak and Radner, 1962 and more recently Bolton and Dewatripont, 1994; Radner and van Zandt, 1992; Calvó-Armengol and de Mart, 2007 and 2009).

This paper fits into a different strand of literature which models a conflict of interest within organizations such that the players disclose their information strategically. In the absence of complete contracts the revelation principle does not work and the management acts in the world of second-best. The literature discusses multiple organizational responses to strategic communication motives. For example, decision maker(s) can close down communication channels and reward those agents who focus on productive activities (Milgrom and Roberts, 1988) or optimally delegate decisions to better informed parties (Dessein, 2002; Alonso, Dessein and Matouschek, 2008; and Rantakari, 2008 and Alonso, Dessein and Matouschek, 2015). This paper studies a different incentive instrument: the design of communication hierarchies in the presense of arbitrary many players with strategic communication motives, who have arbitrary ideal points regarding the policy domain.

My focus on hierarchies is motivated by the extensive literature on hierarchies

[4]See p. 65-67. Dalton describes a case in which the new members of a clique were instructed about the "distinction between their practices and *official misleading instructions*" (italics are from the original text).

within organizations. Hierarchies minimize information processing costs (Bolton and Dewatripont, 1994; van Zandt, 1999a and 1999b; and Garicano, 2000) or prevent conflicts between different organizational layers (Friebel and Raith 2004). More generally, optimally designed hierarchies can reduce moral hazard on "lower" level of organizations and thus minimize resource misallocation.[5] Furthermore, empirical literature shows how organizations adopt hierarchical structures even if they publicly emphasise non-hierarchical decision-making (Oberg and Walgenbach 2008, Ahuja and Carley 1998).

The literature on communication in firms has a counterpart in sociology and management. Early studies of Cyert and March (1963), Dalton (1959) and Crozier (1963) identify intra-organizational groups which lobby decisions by distorting the information provided to the management (see, e.g., Glaser et al. 2015). Tuchman and Scalan (1981) and Allen (1977) show that much of intra-organizational information is mediated by internal communication stars — individuals who act as informational bridges between different groups. Both findings resonate with my result on "group leaders" within optimal communication hierarchies who collect information from the members of their groups, optimally distort it and then communicate it either to another group leader or directly to the management.[6]

I model communication as cheap talk. The workhorse model is Crawford and Sobel (1982) (CS) who study costless and direct communication between a perfectly informed sender and an uninformed receiver who is unable to commit to choices contingent on sender's messages. There are extensions of CS to multiple senders: Krishna and Morgan (2001a) study simultaneous communication and show that the decision maker benefits from consulting two experts with opposing biases; Krishna and Morgan (2001b) study information revelation with two experts from a mechanism design perspective; and Battaglini (2002) shows that perfect revelation with multiple dimensions of the state space is generically possible even for large experts' biases.

All those papers achieve full revelation constructing an equilibrium that plays the experts' reports against one another. Such equilibrium construction is not possible in my model. Since experts' signals are conditionally independent, the on-path signal realizations are unrestricted and the decision maker cannot credibly threaten to punish the experts due to "incompatibility" of their reports as in Ambrus and Lu (2014) or Mylovanov and Zapechelnyk (2013).

[5]For the literature on incentives in hierarchies *without* strategic communication see an excellent overview in Mookherjee (2006).

[6]More recently, the importance of such communication stars is documented in Wadell (2012). Strategic distortions of communicated information are documented in Dhanaraj and Parke (2006) and Schilling and Fang (2013).

I model imperfectly informed experts. Austen-Smith (1993) first studies two imperfectly informed experts in a cheap-talk environment and compares simultaneous and sequential communication. Battaglini (2004) extends the analysis to many experts who communicate simultaneously to the decision maker. Different to my paper, these papers do not study the optimal communication design.

The following three papers are most closely related to my paper. First, Wolinsky (2002) studies a team of imperfectly informed experts and a single decision maker who is unable to commit to decisions based on experts' reports. Wolinsky compares simultaneous communication to a division of experts into groups. He provides examples that show that the decision maker prefers to split the experts into groups rather than to consult each expert individually, and characterizes the lower bound of the group size. There are three major differences of Wolinsky's approach to my model: First, and most importantly, Wolinsky does not look for the optimal communication network whereas my paper optimizes over a large family of communication structures. Second, he assumes the same preferences of the experts and so abstracts from strategic information transmission motives between them. My model, however, allows for arbitrary ideal points of experts with respect to the final decision and shows the importance of the conflict of interest between the experts for the optimal communication. Finally, Wolinsky assumes partial verifiability and non-correlated information, which is both not the case in my model.

The two other papers are Galeotti et al. (2013) and Hagenbach and Koessler (2011) who study equilibrium communication networks with one round of simultaneous communication. My model is based on Galeotti et al. (2013)[7] with the major difference that I pose a question of the optimal communication design whereas Galeotti et al. (2013) look at the properties of one-round communication where each player can read messages to all other players.

Generally, to the best of my knowledge the question of optimal network design in the presence of cheap talk for a broad class of networks and an arbitrary number of players with arbitrary ideal points in the policy space has not been studied yet. However, the literature provides valuable insights for a restricted class of networks like complete networks (Galeotti et al., 2013; Hagenbach and Koessler, 2011) and sequential communication with a single or multiple intermediaries arranged in a line (see, e.g., Ivanov, 2010; Ambrus et al. 2013).

[7]This framework is based on Morgan and Stocken (2008) and is currently used in many political economy and informational economics applications, see e.g. Argenziano et al. (2015) or Dewan (2014).

2 The main idea

Think of two experts ("he") and a decision maker ("she") who wants to obtain payoff-relevant information from the experts. All players share a common prior over an unobserved state θ that is uniformly distributed on $[0,1]$. Each expert receives a costless signal which is either "0" or "1": signal "1" is received with the probability equal to the state and, conditional on the state, the signals are independent. The decision maker chooses a policy $y \in \mathbb{R}$ and wants to minimize the quadratic loss function $-(y-\theta)^2$. Each expert wants to minimize $-(y-\theta-b)^2$ where $b \in \mathbb{R}$ denotes each expert's bias. The decision maker is uninformed and relies on communication with the experts. I assume that each expert can lie at no cost.

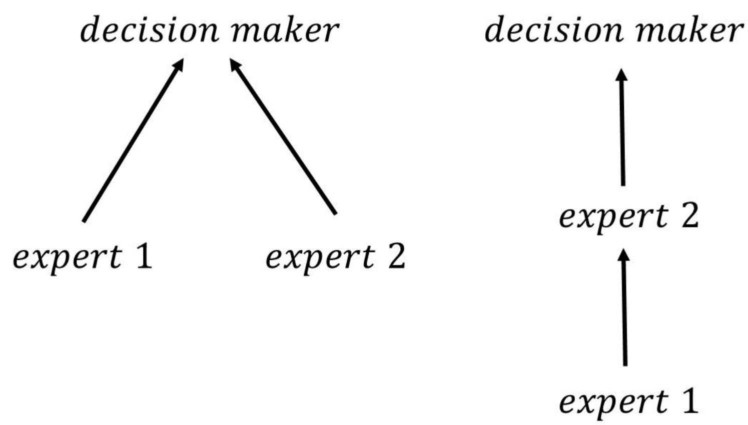
Prior to communication, the decision maker chooses a hierarchical communication network that specifies which player communicates with whom, and in which order. Assume that the decision maker can only choose between a star where the experts send simultaneous messages to the decision maker (Figure 1, left), or a line where expert 1 sends a message to expert 2, and then expert 2 sends a message to the decision maker (Figure 1, right). I assume that the experts can send any messages. Which network maximizes decision maker's payoffs if we focus on the best pure strategy equilibria for the decision maker?[8]

Without loss of generality assume that each expert's message space is equivalent to the space of his possible signals. Before studying the incentives of the players, notice that in the star network each expert has only two message strategies. He can either send signal-independent messages which contain no information for the decision maker, or he can truthfully communicate his signal. In the line expert 1 has the same two message strategies. However, expert 2 has *more* message strategies. If expert 1 reveals his signal truthfully to expert 2, expert 2's information is one of the elements within the set $\{00, 01, 10, 11\}$. Therefore, expert 2's message strategy can pool multiple signal combinations into a single message - for example, expert 2 can inform the decision maker if the realization of both signals is $\{00\}$ or within the set $\{10, 01, 11\}$. Notice that such *strategic pooling* of signals is not possible in the star. But would expert 2 use a pooling message strategy in equilibrium, and - if yes - how does it affect decision maker's preference over the two communication networks?

In the star network, the decision maker's payoff increases in the number of truthfully communicated signals: 0, 1 and 2 signals result in the decision maker's expected payoff of $-\frac{1}{12}$, $-\frac{1}{18}$ and $-\frac{1}{24}$ respectively. Since the decision maker wants to match the state, only a small enough bias of the experts can guarantee that they are truthful in the star network. In particular, an equilibrium in which only one expert commu-

[8] I analyze mixed strategy equilibria in section 4.4.

Figure 1: Illustration of the main idea



nicates his signal truthfully requires $b \leq \frac{1}{6}$, and an equilibrium in which both experts communicate their signals truthfully requires $b \leq \frac{1}{8}$. For $b > \frac{1}{6}$ there is a unique equilibrium in the star in which every expert sends signal-independent messages.[9]

Now, think about the following strategy profile in the line that was hinted above. Expert 1 reports his signal truthfully to expert 2, who, in turn, observes one of the signal combinations from the set $\{00, 01, 10, 11\}$. Expert 2 informs the decison maker if the two signals are $\{00\}$ or within the set $\{10, 01, 11\}$. The latter message is "coarse": upon receiving the message the decison maker only knows that the combined signals are not $\{00\}$. The decision maker's expected payoff is $-\frac{5}{96}$ and this strategy profile is an equilibrium for $b \leq \frac{3}{16}$. Notice that for $b > \frac{1}{6}$ the star does not transmit any information whereas in the line both experts communicate according to the coarse message profile described above. The key is that strategic coarsening increases the deviation costs of the experts. In the star, a deviation from a truthful message results in a smaller shift of decision maker's policy compared to a shift from the truthful message $\{00\}$ to the higher message $\{10, 01, 11\}$ in the line. A larger shift in decision maker's policy means that there is an "overshooting" effect in terms of the experts' objectives: by deviation to a higher message, the policy y moves too far away from experts' ideal points. This means that the message profile in the line

[9]The uninformative equilibrium is usually called *babbling* and it is a standard feature of cheap talk games. There is a literature on refinements of cheap talk equilibria that - among other things - aims to exclude babbling, see e.g. Farrell (1993).

can be supported by large biases.[10] As a result, for $\frac{1}{8} < b < \frac{3}{16}$ the decision maker strictly prefers the line over the star.[11]

3 Model

There are n experts, labelled 1,..,n, and a single decision maker denoted by DM. Let $N^e := \{1, .., n\}$ denote the set of all experts, and $N := N^e \cup \{DM\}$ denote the set of all players. Each player $i \in N$ has a payoff function

$$u_i = -(y - \theta - b_i)^2,$$

where $y \in \mathbb{R}$ is the policy chosen by the decision maker and $b_i \in \mathbb{R}$ is the bias of player i. Without loss of generality I set $b_{DM} = 0$. Each player's payoff depends on the state θ which is commonly known to be uniformly distributed on $[0, 1]$.

The state is unobserved and every expert receives a costless conditionally independent signal $s \in \{0, 1\}$ with $Prob(s = 1) = \theta$. Player's updating follows the Beta-binomial model: if the sum of n signals is k, the expected value of the state is $\frac{k+1}{n+2}$.

The decision maker chooses a communication network which is represented by a directed graph $Q = (N, E)$ with the set of nodes N and an adjacency $n \times n$ matrix $E = [e_{ij}]_{i,j \in N}$ with $e_{ij} \in \{0, 1\}$ representing the availability of a directed link from i to j. A *path* in a network Q, $H_{i_1 i_L}(Q)$, is a sequence of nodes $i_1, i_2, .., i_L$ such that $e_{i_l i_{l+1}} = 1$ for each $l = 1, .., L - 1$. I only study communication networks with the following properties (see the Appendix for the formal definition):

1. every expert has only one outgoing link but can have multiple incoming links,
2. there are no cycles in the graph, and
3. the decision maker has at least one incoming link, but no outgoing links.[12]

[10] There is parallel to the cheap-talk communication in CS who show that pooling of multiple signals into a single message is necessary to support influential communication between a receiver and a biased sender.

[11] The example shows that the decision maker strictly prefers to receive her information according to the partition $P := \{\{00\}, \{01, 10, 11\}\}$ rather than according to the partition $P' := \{\{0\}, \{1\}\}$ in which only one experts communicates his signal truthfully. Notice that P *is not* more informative than P' in the Blackwell sense. Think of a payoff function $-(y - \theta)^3$ for the decision maker. Then, the expected payoff of the decision maker is $-\frac{1}{640}$ for P and 0 for P'.

[12] I only look at deterministic mechanisms.

Let $\mathbb{Q}$ denote the family of communication networks which satisfy the above properties. Notice that those networks are essentially communication hierarchies (in the paper I use the terms hierarchy and network interchangeably).

The entire payoff-relevant information of expert $i \in N^e$, or his *type*, consists of his own private signal and the information communicated by the experts directly connected to i in a network Q. Denote by $P(S)$ a partition of an arbitrary set S. Each expert's communication strategy is a partition of his type space. To define expert's type space and strategies we need an additional definition and a notation. First, for any player $i \in N$, let $N_i(Q) := \{j \in N^e : e_{ji} = 1\}$ define the set of those experts who can send messages to i in a network Q. Second, if some expert $i \in N^e$ babbles, which means that he sends messages independent of his private information, his communication strategy is denoted by $P_i^b(Q)$.

A *type set* of expert $i \in N^e$ is

$$T_i(Q) := \Pi_{j \in N_i(Q)} P_j(\{0,1\}^{\tilde{N}_j(Q)}) \times \{0,1\},$$

and a *communication strategy* of expert $j \in N^e$ is

$$P_j(Q) := T_j(Q) \to P(T_j(Q)),$$

where

$$\tilde{N}_j(Q) := \{i' \in N^e : \exists\ H_{i'j}(Q) \text{ and for all } j' \in H_{i'j}(Q)\ P_{j'}(Q) \neq P_{j'}^b(Q)\}$$

defines the set of all experts located on all paths leading to expert i in Q who are not babbling, and whose successors on the entire path leading to expert i are not babbling either.

Notice the recursive structure of expert's beliefs and strategies: expert's communication strategy is a partition of his type space, and his type space is a function of communication strategies of experts directly connected to him, whereas the communication strategies of those experts are partitions of their type spaces where each of those type spaces depends on the communication strategies of experts directly connected to them etc. Naturally, for experts at the bottom of the hierarchy the type space is $\{0,1\}$ and their communication strategy is a partition of $\{0,1\}$.

Decision maker's beliefs in a network Q are determined by communication strategies of the experts directly connected to her and can be represented by a partition $P_{DM}(Q) := \Pi_{i \in N_{DM}(Q)} P_i(Q)$. The strategy of the decision maker consists, first, of choosing and committing a network $Q \in \mathbb{Q}$ and, second, of choosing an action

$$y(Q): \ P_{DM}(Q) \to \ \mathbb{R}.$$

Most communication networks give rise to multiple equilibria which is typical for strategic communication games. I focus on the best equilibrium for the decision maker in any given network.[13] In particular, when referring to optimality I use the notion of a weak dominance: a network Q weakly dominates a different network Q' if, for any players' biases, the best equilibrium payoff for the decision maker in Q is at least as high as in Q', and for some biases it is strictly higher.[14]

A Perfect Bayesian Equilibrium specifies strategies and beliefs for each player $i \in N$ and is a tuple

$$\Big(Q, \{T_i(Q)\}_{i=1,..,n}, \{P_i(Q)\}_{i=1,..,n,DM}, y(P_{DM}(Q))\Big).$$

The following conditions are satisfied in equilibrium:

1. $y(\cdot)$ must be sequentially rational. For $k \in \{0,..,n\}$ it means that if $p' \in P_{DM}(Q)$ is reported to the decision maker, she chooses

$$y \in \underset{y\in\mathbb{R}}{\operatorname{argmax}} \ -\sum_{k\in p'} Pr(k) \int_0^1 (y-\theta)^2 f(\theta|k,n)d\theta.$$

2. For every $t_i \in T_i(Q)$, the partition $P_i(Q)$ is incentive compatible if, for $t_i \in p_i$, $p_i \in P_i(Q)$:

$$-\sum_{p\in P_{DM}(Q)} \Pr(p|p_i, P_{-i}(Q)) \sum_{k\in p} Pr(k|t_i) \int_0^1 (y(p)-\theta-b_i)^2 f(\theta|k,n)d\theta \geq$$

$$-\sum_{p\in P_{DM}(Q)} \Pr(p|p_i', P_{-i}(Q)) \sum_{k\in p} Pr(k|t_i) \int_0^1 (y(p)-\theta-b_i)^2 f(\theta|k,n)d\theta$$

for $p_i' \in P_i(Q)$ and $p_i' \neq p_i$.

[13] Such an equilibrium is Pareto optimal in the ex-ante sense: before experts receive their signals each player aims to minimize the residual variance $\mathbb{E}[-(y-\theta)^2]$.

[14] A similar optimality criterion is used in Austen-Smith (1993).

3. Finally, Q maximizes the expected payoff of the decision maker:

$$Q \in \underset{Q \in \mathbb{Q}}{\operatorname{argmax}} - \sum_{p \in P_{DM}(Q)} Pr(p|P(Q)) \sum_{k \in p} Pr(k) \int_0^1 (y(p) - \theta)^2 f(\theta|k', n) d\theta,$$

where $P(Q) := \Pi_{i \in N^e} P_i(Q)$.

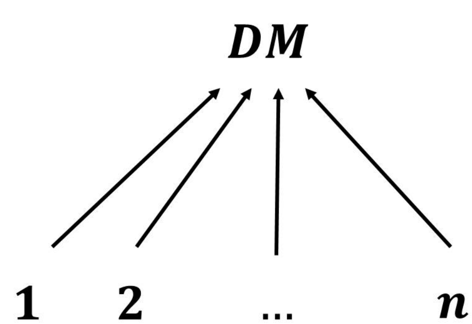
Given the equilibrium conditions, the decision maker aims to match the state and chooses $y(\cdot) = \mathbb{E}_{DM}(\theta|p)$ for $p \in P_{DM}(Q)$.

3.1 Simultaneous versus sequential communication

We start with a natural benchmark where all experts send their messages simultaneously to the decision maker. Such a star network is shown in Figure 2. It turns out that equilibria in a star have a simple characterization (see also Morgan and Stocken, 2008; and Galeotti et al., 2013):

Proposition 1: *Take any number of experts, n, with arbitrary biases. An equilibrium in a star network in which $n' \leq n$ experts communicate their signals truthfully to the decision maker exists iff for every expert $i \in \{1, .., n'\}$, $|b_i| \leq \frac{1}{2(n'+2)}$.*

Figure 2: A star with n experts



Thus, a smaller difference in biases between the experts and the decision maker results in more experts revealing their signals truthfully. The influence of every truthful message on decision maker's policy gets smaller with a larger number of equilibrium truthful messages. This makes a deviation more profitable for every expert. Therefore, in order to sustain a larger number of truthful messages in equilibrium, the biases of the experts have to be small enough. Conversely, a smaller number of truthful messages corresponds to a larger impact of each truthful message on decision maker's policy making a deviation less profitable. As a result, such an equilibrium can be supported by larger biases.

A quick intuition might suggest that a star network is beneficial for the decision maker since no expert can distort the messages of the other experts. The next Proposition shows that this intuition is wrong: a star network is weakly dominated by an optimally designed intermediation where at least one expert has access to a signal of at least one other expert.

Proposition 2: *Take any number n of experts with arbitrary biases. Take any network Q that is not a star, such that if expert i communicates to expert j, $e_{ij} = 1$, then $|b_j| \leq |b_i|$. Then:*

1. *Any equilibrium outcome in a star network is also implementable as an equilibrium outcome in network Q.*

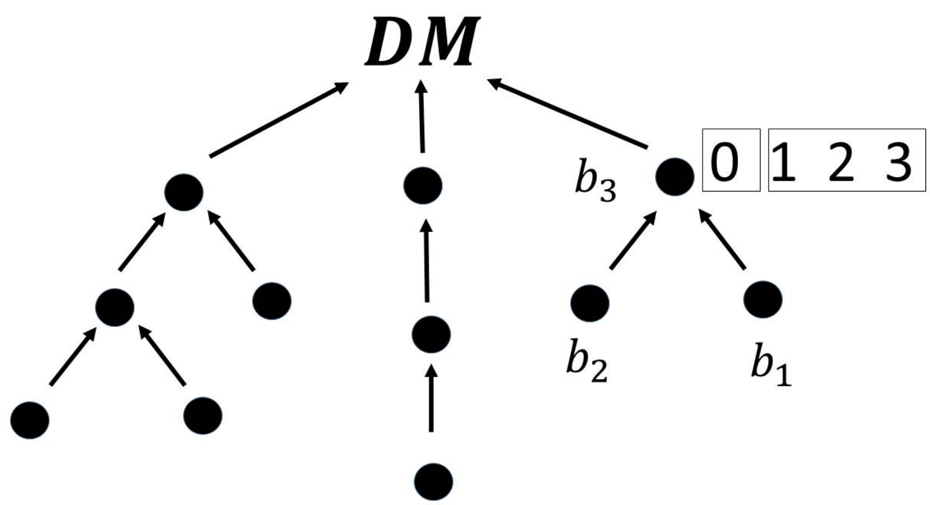
2. *There is a range of experts' biases for which the best equilibrium in Q strictly dominates the best equilibrium in a star. Further, this equilibrium in Q involves strategic coarsening of information by at least one of the experts.*

Proposition 2 provides an important insight into optimal communication hierarchies. It shows that an optimally designed intermediation can always generate the same equilibrium payoffs for the decision maker as the star network. Moreover, it can sometimes generate strictly higher payoffs for the decision maker. As I show in the proof, it happens due to informational coarsening by intermediaries, it means, by the experts who receive and forward messages of at least one other expert.

The proof of the first part requires three steps. First, fix an equilibrium with $n' \leq n$ truthful experts in a star. The decision maker receives information according to a partition $\{\{0\}, \{1\}, .., \{n'\}\}$ where the sum of the signals $k \in \{0, .., n'\}$ is the summary statistic. Upon receiving experts' messages, a sequentially rational decision maker chooses $y = \frac{k+1}{n'+2}$. Second, take any truthful expert i and suppose that the sum of the signals of all the other truthful experts is $k' \in \{0, .., n'-1\}$. If i receives $s_i = 0$ and reveals it truthfully, then the decision maker chooses $y = \frac{k'+1}{n'+2}$. However, if i deviates to the higher message "1", the decision maker chooses $y' = \frac{k'+2}{n'+2}$. The difference $(y' - y)$ is independent of k' and therefore the incentive of expert i to deviate to the higher message does not depend on the exact number k'. The same is true for expert i's downward deviation in case $s_i = 1$. With other words, conditional on expert i's private signal, his deviation only depends on the equilibrium number of truthful signals but not on the realization of other experts' signals.

Third, from Proposition 1 we know that in the star, for any number of truthful signals communicated to the decision maker, there always exists an equilibrium in which the largest absolute value of the bias among all truthful experts is weakly smaller than the smallest absolute value of the bias among all non-truthful experts.[15] The optimal ordering in Proposition 2 ensures that all such truthful experts are "closest" to the decision maker. Since in any equilibrium in which decision maker's partition is $\{\{0\}, \{1\}, .., \{n'\}\}$, the deviation incentives of the experts are independent of their beliefs about the signals of the other truthful experts, such an equilibrium can be generated by an optimally ordered hierarchy.

Figure 3: Example of an informational coarsening by expert 3



The proof of the second part uses the fact that within an optimally ordered hierarchy each expert has at least as many message strategies as in the star. Figure 3 illustrates this point. If experts 1 and 2 truthfully reveal their signals, the summary

[15]Think, for example, of three experts with $b_1 = b_2 = \frac{1}{8}$ and $b_3 = \frac{1}{10}$. Given Proposition 1, the most informative equilibrium in the star features only - and *any* - two experts. Thus, it can be the case that in the most informative equilibrium experts 1 and 2 communicate truthfully and expert 3 babbles, although expert 3 has the smallest bias. But what matters for Proposition 2 is that there exists an equilibrium with two truthful experts one of whom is expert 3.

statistic of expert 3 has one of the realizations from the set $\{0, 1, 2, 3\}$. Thus, expert 3 can use a message strategy which pools some of his information sets into a single message. In particular, he can either communicate to the decision maker that his summary statistic is 0, or everything else (as shown in Figure 3). As I show in the proof, such a message strategy is optimal for the decision maker if all experts' biases are sufficiently large. The decision maker benefits from coarse information transmission since otherwise no information is transmitted in the star. The reason is that the informational coarsening increases the "costs" of deviation from truthtelling and makes the experts more trustworthy than in the star.

Propositions 1 and 2 further imply that if the biases of all n experts are weakly below $\frac{1}{2(n+2)}$, the shape of an optimal communication network does not matter as long as every expert has a directed link to one of the other players. In this case the conflict of interest among the players is so small that all signals are truthfully revealed to the decision maker. However, once the bias of at least one of the experts exceeds $\frac{1}{2(n+2)}$, a star network can never outperform an optimally designed intermediation.

The next Proposition shows that for high enough positive (or low enough negative) biases the optimal communication network has two simple requirements. First, a decision maker has to receive messages from a single expert, no matter which one. Second, each of the other experts must have a directed link to any of the other experts (examples of optimal networks are depicted in Figure 4).

Proposition 3: *Take any number $n \geq 2$ of experts such that all their biases are either within the interval $(\frac{n}{4(1+n)}, \frac{n+1}{4(n+2)}]$ or within the interval $[-\frac{n+1}{4(n+2)}, -\frac{n}{4(n+1)})$. Then:*

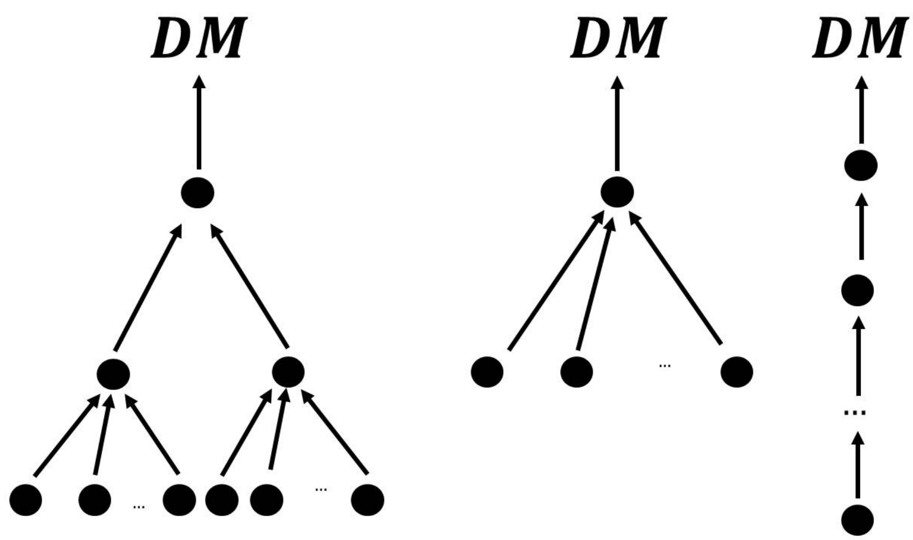
1. *In the optimal network the decision maker is connected to a single expert $i \in \{1, .., n\}$, $e_{iDM} = 1$, and each expert apart from i is connected to some other expert $j \in \{1, .., n\}$, no matter which one.*

2. *Any other network does not transmit any information from the experts to the decision maker.*

Moreover, If the biases of all experts are above $\frac{n+1}{4(n+2)}$ or below $-\frac{n+1}{4(n+2)}$, then no information is transmitted in equilibrium.

The explanation for Proposition 3 goes as follows. If the stated conditions are satisfied, there exists an equilibrium in which an expert who is directly connected to the decision maker knows the signals of all other experts. This enables him to send messages according to the coarsest possible message profile. In the case of positive biases, this expert sends either a message containing the sum of the signals 0, or a second message which pools together all the other signals. This message profile implies that all experts condition their upward deviation on the same event that all n signals are 0. Whenever an expert with a summary statistic 0 deviates to a higher message, then, conditional on all signals being 0, the decision maker chooses a policy which is a large distance away from the experts' ideal points. This prevents deviations to higher messages even by experts with very large biases.

Figure 4: Examples of optimal networks in Proposition 3



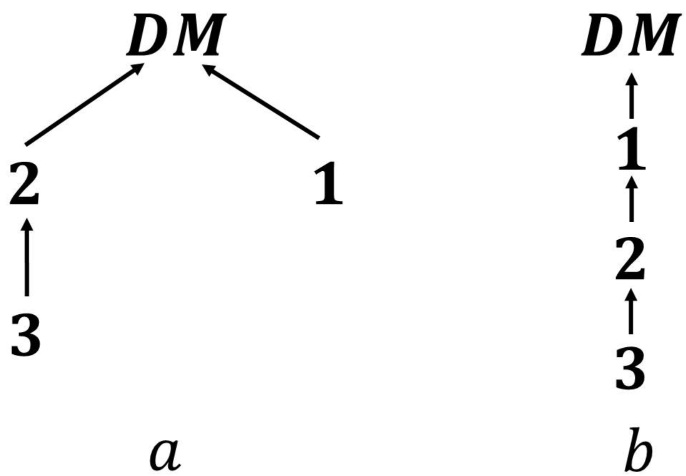
Examples of possible optimal networks from Proposition 3 are depicted in Figure 4. The ordering of the experts can be arbitrary. This is because the experts' biases are within a narrow bias range and, therefore, the experts do not face incentive problems when informing each other about their signals.

The proof in case of negative biases is similar: the corresponding downward deviation which leads to the largest possible downward shift of decision maker's policy away from the experts' ideal points, requires a communication strategy which either informs the decision maker that the sum of the signals is n or that the sum of the signals is everything else.

3.2 Uncertainty and incentives

Example: Think of three positively biased experts, labelled 1, 2 and 3. Consider, first, the network depicted in Figure 5a in which the experts 1 and 2 are uncertain about each other's signals. Think of a strategy profile in which expert 3 reveals his signal to expert 2 who, then, informs the decision maker if the combined signals are 00, or within the set $\{10, 01, 11\}$. Expert 1 reveals his signal to the decision maker. The decision maker, then, matches the policy y to her posterior of the state. As I show in the appendix, this strategy profile is an equilibrium for $b_2, b_3 \leq \frac{74}{525}$ and $b_1 \leq \frac{293}{2550}$. Moreover, for $b_1 \leq \frac{293}{2550}$ and $\frac{1}{8} < max\{b_2, b_3\} \leq \frac{74}{525}$ the network in Figure 5a is optimal.

Figure 5: Illustration of the example (p 17)



The decision maker can implement the same payoff profile in the line (Figure 5b) but for the above bias range the line is not an optimal network. To see why, think of the following strategy profile: experts 2 and 3 have the same message strategies as in the network in Figure 5a. Expert 1, however, communicates to the decision maker *both* his own signal *and* the message of expert 2. The decision maker receives information according to the same partition of experts' signals as before, and therefore the expected payoff allocation of all players remains the same. Moreover, the incentive constraints for experts 2 and 3 remain unchanged. However, the line requires a strictly smaller bias for expert 1: it has to be weakly below $\frac{1}{30}$. This is because in the line expert 1 conditions his deviation on the exact message of expert 2 and therefore has to satisfy more incentive constraints than in the other network.

An important feature of the above example is the strategic coarsening of information by expert 2. If an equilibrium does not feature informational coarsening by at least one of the experts, Proposition 2 tells us that the incentive constraints of the experts are the same across all optimally ordered hierarchies. The next Lemmas show that this is no longer true in equilibria with strategic coarsening of information. In such equilibria, the less information an expert has about the overall signals, the larger is the slack in his incentive constraints.[16]

Definition: Expert i receives full information from some other expert j in a network Q if i perfectly observes every $p \in P_j(Q)$ communicated by expert j.[17]

Lemma 1: *Take any equilibrium in an optimal network Q that involves strategic coarsening of information. If some expert j truthfully communicates all his signals, $P_j(Q) = T_j(Q)$, and some other expert i receives full information from j, then the range of biases supporting the equilibrium strategy of i is weakly included in the range of biases supporting the equilibrium strategy of j.*

The next Lemma shows the implication of Lemma 1 for the structure of optimal networks.

Lemma 2: *Take any number of experts with arbitrary biases. In the corresponding optimal network Q, if an equilibrium involves strategic coarsening of information and there is a player $i \in N$ who receives full information from some expert $j \in N^e$, $j \neq i$, then j has to be connected to i, $e_{ji} = 1$.*

The proof of Lemma 2 uses the insight that if, contrary to the Lemma, there are some experts between i and j, those experts face tighter constraints compared to the case in which j is directly connected to i, without changing the expected payoff allocation. Thus, the Lemma shows that if an expert receives full information from some other experts and coarsens the aggregated information, then in an optimal network he receives this information unmediated, it means from each of those experts directly.

As a result, an optimal network consists of "information coordination units" - I call them *groups* - which are essentially stars. At the center of a group there is an expert - I call him a *group leader* - who receives direct messages from all other group members, strategically coarsens the aggregated information and communicates it up the hierarchy.

Definition: In an optimal network, a group G is a non-empty subset of a communi-

[16]This effect is well known in mechanism design where in general the interim incentive constraints are easier to satisfy than the ex post constraints.

[17]Formally, the partition $P_i^b(Q)$ is finer than the partition $P_j(Q)$.

cation network consisting of some expert $i \in N^e$ and a non-empty subset of experts $\tilde{N} \subseteq N^e \setminus \{i\}$ such that each $j \in \tilde{N}$ is directly connected to i. Expert i is a group leader and I denote him by i^G.

Notice that I am not excluding the possibility that a *group leader* can be directly connected to a leader of another group if the latter strategically coarsens information in his message strategy. Thus, the "depth" of an optimal network has a direct connection to the number of informational coarsening rounds. The next Lemma provides insights into the intra-group structure.

Lemma 3: *Suppose that an equilibrium in an optimal network Q involves strategic coarsening of information. Then:*

1. *if two distinct experts $j' \in N^e$ and $j \in N^e$ belong to the same group G, none of them is a group leader of G, both j and j' have no incoming links, $\sum_{k \in N^e} e_{kj} = \sum_{k \in N^e} e_{kj'} = 0$ and both of them do not babble, $P_j(Q) \neq P_j^b(Q)$ and $P_{j'}(Q) \neq P_j^b(Q)$, then their message strategies are supported by the same range of biases, and*

2. *the equilibrium strategy of i^G is supported by a weakly smaller range of biases, compared to j and j'.*

To understand Lemma 3, think of a group in an optimal hierarchy which features experts with no incoming links. Their only information are their private signals. Lemma 3 shows that, conditional on a group, in equilibrium all those experts face the same incentive constraints. As shown in the proof of Lemma 3, this is, first, because the experts with no incoming links face the same expected uncertainty about the overall signals and, second, because their signals enter the message strategy of their group leader symmetrically. In this sense, groups in an optimal network include experts with similar biases.

Further, a *group leader* has a better knowledge about the overall signals compared to those experts in his group who only observe their own signals. In an equilibrium with strategic coarsening of information, the range of biases supporting the equilibrium strategy of a group leader is weakly included into the range of biases supporting the equilibrium strategy of those experts, which is a direct consequence of Lemma 1.

Lemmas 2 and 3 show that due to the uncertainty force an optimal communication hierarchy should not be "too deep". An optimal hierarchy consists of groups of experts with similar biases. A *group leader* collects information from the other group members and communicates it in a coarse way either to another *group leader* or directly to the decision maker. This result resonates with the findings on information processing in organizations by Cyert and March (1963) who show that groups include organizational members with similar objectives who aggregate information within a group before communicating it in a systematically distorted way further up the hierarchy.

3.3 Optimal networks

This section uses the results of the previous sections and characterizes the optimal networks and corresponding equilibria for three positively biased experts such that none of them babbles.[18] When referring to the experts I label them 1, 2 and 3 such that $b_1 \leq b_2 \leq b_3$.

Propositions 1 and 2 imply that for experts' biases below $\frac{1}{10}$, any network leads to complete revelation of all signals and yields $\mathbb{E}U_{DM} \simeq -0.033$.

For the following assume $b > 0.1$ for at least one of the experts. The optimal network depends on the biases as follows

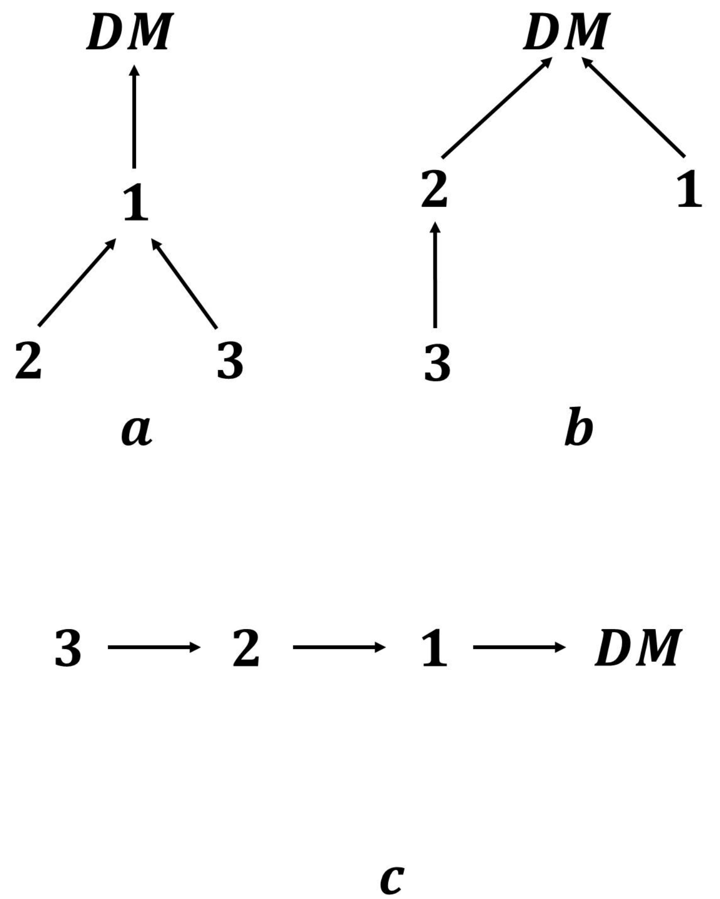
1. *The optimal network is depicted in Figure 6(a) (denoted by Q_a) if the biases are small.* Assume $b_1 \leq 0.1$ and $0.1 < b_3 \leq 0.125$. In equilibrium experts 2 and 3 communicate their signals truthfully to expert 1. Expert 1 sends one of the three messages: if the sum of the signals is 0, he sends p_1, if the sum of the signals is 1, he sends p_1', and if the sum of the signals is either 2 or 3, he sends p_1''. Decision maker's choices are $y(p_1) = \frac{1}{5}$, $y(p_1') = \frac{2}{5}$ and $y(p_1'') = \frac{7}{10}$. This strategy profile yields $\mathbb{E}U_{DM} \simeq -0.038$.

2. *The optimal network is depicted in Figure 6(b) (denoted by Q_b) if all biases are in the intermediate range.* The biases $b_1 \leq 0.115$ and $0.125 < b_3 \leq 0.14$ support the following equilibrium strategy profile: expert 3 communicates his signal truthfully to expert 2. Expert 2 sends p_2 if both his private signal and the message of expert 3 are 0, and sends p_2' otherwise. Expert 1 sends p_1 if his signal is 0, and p_1' otherwise. The decision maker's choices are $y(p_1, p_2) = \frac{1}{5}$, $y(p_1', p_2) = \frac{2}{5}$, $y(p_1, p_2') = \frac{7}{15}$ and $y(p_1', p_2') = \frac{18}{25}$ resulting in $\mathbb{E}U_{DM} \simeq 0.04$. We know from the section 3.2 that the same outcome can be implemented in a line, but only for a strictly smaller range of expert 1's bias.

3. *The optimal networks are depicted in Figure 6(a) and 6(c) (the latter is denoted by Q_c) if all biases are large.* This is a direct consequence of Proposition 3. Fix

[18] If, for example, one of the experts babbles, we are back to the case with two communicating experts which has been covered in the introductory example.

Figure 6: Networks for 3 experts



$0.1875 < b_1 \leq 0.2$ and $b_3 \leq 0.2$. In equilibrium experts 2 and 3 communicate their signals truthfully to expert 1. Expert 1 sends p_1 if the sum of all signals is 0. Otherwise he sends p_1'. The decision maker chooses $y(p_1) = \frac{1}{5}$ and $y(p_1') = \frac{3}{5}$ with the resulting $\mathbb{E}U_{DM} \simeq -0.05$.

4 Further results

4.1 Opposing signs of biases

Suppose that a decision maker consults four experts: two experts have the same positive bias, $b_1 = b_2 = b^+ > 0$, and the other two experts have the same negative bias, $b_3 = b_4 = b^- < 0$. Suppose that the decision maker can only split the experts into two lines, each one including two experts. The available networks (up to relabelling of the experts) are depicted in Figure 7. What is the optimal assignment of the experts into both lines? Although the question is more restrictive than before, it offers additional insights into optimal networks with experts having opposite signs of biases. I show that it is optimal to put the biases of the same sign into a single line, and therefore to separate them from the biases of the opposite sign. The reason is that in the optimal network the experts within each line bias their messages in the direction opposite to the other line. The decision maker wants to minimize the residual variance and benefits from having two reports biased in the opposite directions.

Assume $max\{|b^-|, |b^+|\} > \frac{1}{12}$ as otherwise Proposition 2 states that all experts communicate their signals truthfully in any network.

In case A (Figure 7A) the experts are partitioned into two lines according to the sign of their biases. In cases B, C and D (Figure 7B, 7C, 7D) the experts have mixed signs within the lines and differ by the signs of the experts communicating to the decision maker.

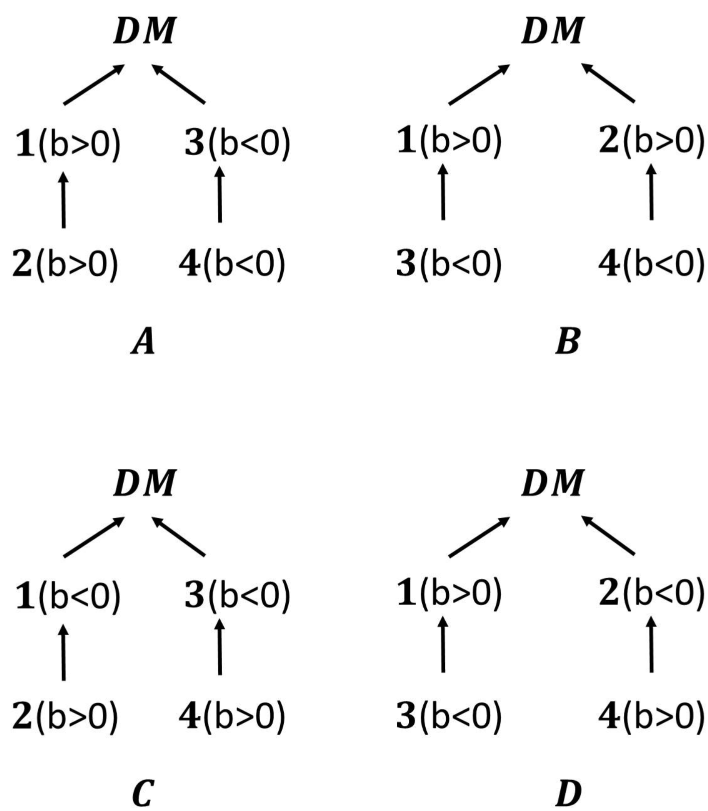
Think of the following message strategies in case A: experts 2 and 4 communicate their signals truthfully to experts 1 and 3. If the sum of the signals observed by expert 1 is 0, he sends p_1. Otherwise he sends p_1'. If the sum of the signals which expert 3 observes is either 0 or 1, he sends p_3. Otherwise he sends p_3'. As a result, the decision maker receives messages according to the partition $\{\{0\}, \{1, 2\}\} \times \{\{0, 1\}, \{2\}\}$. Notice that expert 1's communication strategy pools together the two of the largest sums of signals whereas expert 3's message strategy pools together the two of the lowest sums of signals. Thus, the experts within each group bias their communication strategy in a direction opposite to the other group. The decision maker chooses

$$y(p_1, p_3) = \frac{2}{9},\ y(p_1, p_3') = y(p_1', p_3) = \frac{1}{2},\ y(p_1', p_3') = \frac{7}{9}.$$

The expected payoff of the decision maker is $\mathbb{E}U_{DM} \simeq -0.037$. This strategy profile is an equilibrium for $b^+ \leq 0.13$ and $b^- \geq -0.13$. As I show in the appendix, for $b^+ \in (0.1, 0.13]$ and $b^- \in [-0.13, -0.1)$ the network depicted in Figure 7A dominates all other networks.[19]

[19]For $max\{|b^-|, |b^+|\} \leq \frac{1}{10}$ any of the networks is optimal and involves any of the three experts revealing their signals perfectly to the decision maker.

Figure 7: Possible group arrangements



4.2 Commitment

Suppose that the decision maker can commit to a mechanism which implements allocations contingent on experts' messages. According to the revelation principle we can focus on a direct mechanism. Denote by σ_i a (pure) strategy of expert $i \in N^e$ in a direct mechanism and by σ_{-i} the strategy profile of all other experts. Formally, a direct mechanism is a rule q that maps experts' types into the final decision $y \in \mathbb{R}$:

$$q : \{0,1\}^n \to \mathbb{R},$$

and which satisfies

$$\mathbb{E}U_i(\sigma_i = t_i, \sigma_{-i} | t_i \in \{0,1\}) \geq \mathbb{E}U_i(\sigma_i = 1 - t_i, \sigma_{-i} | t_i \in \{0,1\}), \text{ for every } i \in N^e$$

meaning that an expert of type $t_i \in \{0,1\}$, $i \in N^e$, has no incentive to communicate $1 - t_i$ instead of t_i if all other experts communicate their types truthfully.

Proposition 4: *Every equilibrium outcome generated by an optimal network $Q \in \mathbb{Q}$ is also an equilibrium outcome of the direct mechanism q. The converse is false.*

The intuition for Proposition 4 goes as follows. Each equilibrium outcome in a communication network can be summarized by a partition of $\{0,1\}^n$ according to which the decision maker receives her information. Per definition, such an equilibrium partition is incentive-compatible for every player and can be implemented by an incentive-compatible direct mechanism q.

The converse is not true. In a communication network which is not a star at least one of the experts observes the message of at least one other expert. According to Lemma 1, in equilibria which involve strategic coarsening of information, observing the signals of the experts tightens the incentive constraints of an expert. In the direct mechanism, however, no expert observes the signals of the other experts. Therefore, a direct mechanism can in general implement an allocation for a larger range of biases compared to an optimal network which is not a star.

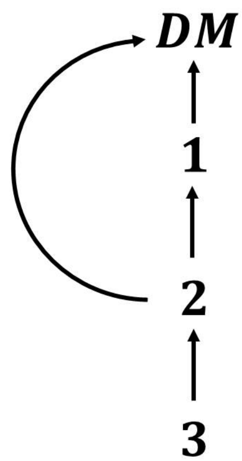
In equilibria that do not involve strategic coarsening of information (as in the star), the incentives to communicate information are independent of players' beliefs about the signals of other experts, as Proposition 2 shows. In this case the optimal network and a direct mechanism generate the same outcome.

4.3 Beyond Tree Networks

How limiting is the focus on trees compared to alternative communication networks with only one round of communication? I show how a network with multiple outgoing links can outperform an optimal (tree) hierarchy.

Example: Consider three positively biased experts organized in a network depicted in Figure 8, and the following strategy profile: expert 3 truthfully reveals his signal to expert 2. Expert 2 sends the *same* message to the decision maker and to expert 1: if both expert 3's message and his private signal are 0 he sends p_2, otherwise he sends p_2'. Expert 1 communicates as follows: if he receives p_2, he sends p_1 irrespective of his private signal. If he receives p_2', he sends p_1' if his private signal is 0 and p_1'' if his private signal is 1. Thus, if the decision maker receives p_2 from expert 2, she disregards expert 1's message. Otherwise she can distinguish between different types of expert 1. As a result, in $\frac{1}{3}$ of cases (which is the case if p_2 is sent) the decision maker receives coarse information only from experts 2 and 3, and in $\frac{2}{3}$ of cases (which is the

Figure 8: Expert 2 has two outgoing links



case if p_2' is sent) the decision maker receives additional information about expert 1's signal. The expected utility of the decision maker is -0.044. It turns out that for $0.125 < b_2, b_3 \leq 0.14$ and $0.115 < b_1 \leq 0.13$, the network depicted in Figure 8 dominates any tree network.

Intuitively, only some (but not all) types of expert 1 want to reveal themselves to the decision maker. The additional communication channel from expert 2 informs the decision maker if a type of expert 1 has an incentives to truthfully reveal itself. If the network were a line, the decision maker would not be able to distinguish whether expert 1's type is within a "truthful" subset or not, since all types of expert 1 would want to appear to be in this subset. The above message strategy is therefore incentive compatible in the network in Figure 8 but not incentive compatible in the line.

4.4 Symmetric mixed strategies in the star with two experts

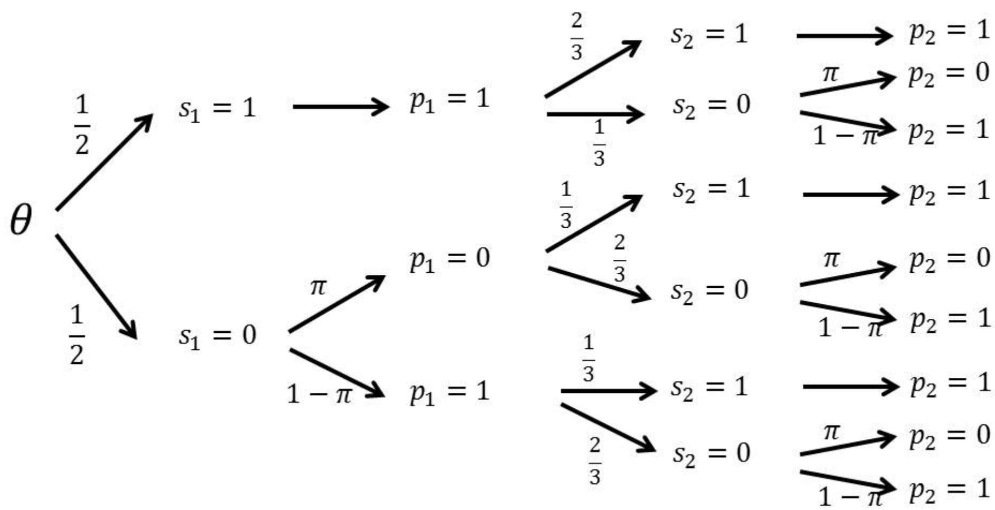
4.4.1 Symmetric equilibrium for positively biased experts

Since the experts are positively biased (their biases are not necessarily equal), they reveal signal 1 truthfully. If an expert $i \in \{1, 2\}$ receives signal 0, assume that he communicates $p_i = 0$ with probability $\pi < 1$ and $p_i = 1$ otherwise.[20] The

[20] One can think of a strategy profile in which only one expert mixes and the other uses pure strategies. I disregard this case since the bias of the mixing expert has to be exactly $\frac{1}{8}$ which is a non-generic case.

corresponding game tree is shown in Figure 9.

Figure 9: Symmetric mixed strategies with positively biased experts.



The decision maker's posteriors of the state - and so her choices of y - are:

$$y_0 := \mathbb{E}(\theta|p_1 = p_2 = 0) = \frac{1}{4}.$$

$$y_1 := \mathbb{E}(\theta|p_i = 0, p_j = 1) = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{2}{3}(1-\pi)}\left(\frac{1}{2}\right) + \frac{\frac{2}{3}(1-\pi)}{\frac{1}{3} + \frac{2}{3}(1-\pi)}\left(\frac{1}{4}\right) = \frac{2-\pi}{6-4\pi}$$

for $i, j \in \{1, 2\},\ i \neq j$, and

$$y_2 := \mathbb{E}(\theta|p_1 = p_2 = 1) = \frac{\frac{2}{3}}{\frac{2}{3} + 2\frac{1}{3}(1-\pi) + \frac{2}{3}(1-\pi)^2}\left(\frac{3}{4}\right) + \frac{2\frac{1}{3}(1-\pi)}{\frac{2}{3} + 2\frac{1}{3}(1-\pi) + \frac{2}{3}(1-\pi)^2}\left(\frac{1}{2}\right) +$$

$$\frac{\frac{2}{3}(1-\pi)^2}{\frac{2}{3}+2\frac{1}{3}(1-\pi)+\frac{2}{3}(1-\pi)^2}\left(\frac{1}{4}\right)=\frac{6-4\pi+\pi^2}{4(3-3\pi+\pi^2)}.$$

It turns out that this strategy profile puts stronger requirements on the biases of the experts compared to the case of complete signal revelation by both experts. Think of expert 1 who receives $s_1 = 0$. If he sends $p_1 = 0$, his expected payoff is

$$-\frac{1}{3}\int_0^1 (y_1-\theta-b_1)^2 6\theta(1-\theta)d\theta-\frac{2}{3}\Big(\pi\int_0^1 (y_0-\theta-b_1)^2 3(1-\theta)^2 d\theta+(1-\pi)\int_0^1 (y_1-\theta-b_1)^2 3(1-\theta)^2 d\theta\Big)$$

If he deviates to $p_1 = 1$, his expected payoff is:

$$-\frac{1}{3}\int_0^1 (y_2-\theta-b_1)^2 6\theta(1-\theta)d\theta-\frac{2}{3}\Big(\pi\int_0^1 (y_1-\theta-b_1)^2 3(1-\theta)^2 d\theta+(1-\pi)\int_0^1 (y_2-\theta-b_1)^2 3(1-\theta)^2 d\theta\Big)$$

Expert 1 is indifferent between sending both messages for

$$b_1^*(\pi)=\frac{-5\pi^4+34\pi^3-84\pi^2+90\pi-36}{8(2\pi-3)\left(\pi^2-3\pi+3\right)\left(3\pi^2-8\pi+6\right)}<\frac{1}{8}$$

where $\frac{\partial b_1^*(\pi)}{\partial \pi} > 0$ and $b_1^*(\pi) < \frac{1}{8}$. Since we focus on the best equilibrium for the decision maker, we conclude that in the star the experts with positive biases never play a symmetric mixed strategy equilibrium since for $b_i \leq \frac{1}{8}$, $i \in \{1, 2\}$, the best equilibrium in the star features both experts revealing their signals perfectly to the decision maker.

4.4.2 Equilibrium with opposing biases

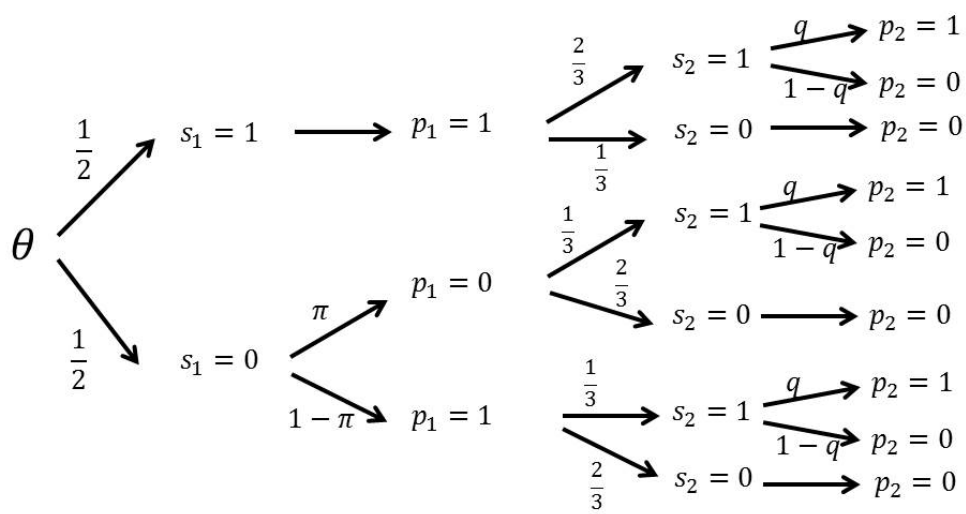
Assume $b_1 = b > 0$ and $b_2 = -b$.[21] Expert 1 truthfully reveals his signal $s_1 = 1$. Conditional on $s_1 = 0$, he sends $p_1 = 0$ with probability $\pi < 1$, and $p_1 = 1$ otherwise. Expert 2 truthfully reveals $s_2 = 0$ and, conditional on $s_2 = 1$, he sends $p_2 = 1$ with probability q and $p_2 = 0$ otherwise. The corresponding game tree is shown in Figure 10.

The sequentially rational decision maker matches y to her expectation of the state depending on the message profile as follows:

$$y_0' := \mathbb{E}(\theta|p_1=0,p_2=0)=\frac{\frac{1}{3}(1-q)}{\frac{1}{3}(1-q)+\frac{2}{3}}\frac{1}{2}+\frac{\frac{2}{3}}{\frac{1}{3}(1-q)+\frac{2}{3}}\frac{1}{4}=\frac{2-q}{2(3-q)}.$$

[21]The bias restriction is assumed for tractability.

Figure 10: Symmetric mixed strategies with opposing biases.



$$y_1' := \mathbb{E}(\theta|p_1 = 1, p_2 = 0) = \frac{\frac{3}{4}[\frac{2}{3}(1-q)] + \frac{1}{2}[\frac{1}{3} + \frac{1}{3}(1-\pi)(1-q)] + \frac{1}{4}[(1-\pi)\frac{2}{3}]}{\frac{2}{3}(1-q) + \frac{1}{3} + (1-\pi)\frac{1}{3}(1-q) + (1-\pi)\frac{2}{3}} =$$

$$\frac{6 - \pi(2-q) - 4q}{2(6 - \pi(3-q) - 3q)}.$$

$$y_2' = \mathbb{E}(\theta|p_1 = 0, p_2 = 1) = \frac{1}{2}$$

$$y_3' = \mathbb{E}(\theta|p_1 = 1, p_2 = 1) = \frac{\frac{2}{3}q}{\frac{2}{3}q + (1-\pi)\frac{1}{3}q}\frac{3}{4} +$$

$$\frac{(1-\pi)\frac{1}{3}q}{\frac{2}{3}q + (1-\pi)\frac{1}{3}q}\frac{1}{2} = \frac{4-\pi}{2(3-\pi)}.$$

The expected payoff of the decision maker is

$$-\frac{1}{2}\Big(\frac{2}{3}\Big[q\int_0^1(y_3'-\theta)^2 3\theta^2 d\theta+(1-q)\int_0^1(y_1'-\theta)^2 3\theta^2 d\theta\Big]+\frac{1}{3}\int_0^1(y_1'-\theta)^2 6\theta(1-\theta)d\theta\Big)-$$

$$\frac{1}{2}\pi\Big(\frac{1}{3}\Big[q\int_0^1(\frac{1}{2}-\theta)^2 6\theta(1-\theta)d\theta+(1-q)\int_0^1(y_0'-\theta)^2 6\theta(1-\theta)d\theta\Big]+\frac{2}{3}\int_0^1(y_0'-\theta)^2 3(1-\theta)^2 d\theta\Big)-$$

$$\frac{1}{2}(1-\pi)\Big(\frac{1}{3}\Big[q\int_0^1(y_3'-\theta)^2 6\theta(1-\theta)d\theta+(1-q)\int_0^1(y_1'-\theta)^2 6\theta(1-\theta)d\theta\Big]+\frac{2}{3}\int_0^1(y_1'-\theta)^2 3(1-\theta)^2 d\theta\Big).$$

Now, think of the incentive constraints of expert 1 who receives $s_1 = 0$. His expected payoff when truthfully revealing his signal is

$$-\frac{2}{3}\int_0^1(y_0'-\theta-b)^2 3(1-\theta)^2 d\theta-\frac{1}{3}q\int_0^1(\frac{1}{2}-\theta-b)^2 6\theta(1-\theta)d\theta-\frac{1}{3}(1-q)\int_0^1(y_0'-\theta-b)^2 6\theta(1-\theta)d\theta$$

whereas if he deviates to $p_1 = 1$ he expects the payoff

$$-\frac{2}{3}\int_0^1(y_1'-\theta-b)^2 3(1-\theta)^2 d\theta-\frac{1}{3}q\int_0^1(y_3'-\theta-b)^2 6\theta(1-\theta)d\theta-\frac{1}{3}(1-q)\int_0^1(y_1'-\theta-b)^2 6\theta(1-\theta)d\theta.$$

Next, the expected payoff of expert 2 from sending $p_2 = 1$ is

$$-\frac{1}{3}\pi\int_0^1(\frac{1}{2}-\theta+b)^2 6\theta(1-\theta)d\theta-\frac{1}{3}(1-\pi)\int_0^1(y_3'-\theta+b)^2 6\theta(1-\theta)d\theta-\frac{2}{3}\int_0^1(y_3'-\theta+b)^2 3\theta^2 d\theta$$

whereas by deviation to $p_2 = 0$ expert 2 expects the payoff

$$-\frac{1}{3}\pi\int_0^1(y_0'-\theta-b_2)^2 6\theta(1-\theta)d\theta-\frac{1}{3}(1-\pi)\int_0^1(y_1'-\theta-b_2)^2 6\theta(1-\theta)d\theta-\frac{2}{3}\int_0^1(y_1'-\theta-b_2)^2 3\theta^2 d\theta.$$

The strategy profile which satisfies the indifference conditions is:

$$p^*(b) = q^*(b) = \frac{12b-1}{4b}$$

and the corresponding expected utility of the decision maker is:

$$b - \frac{1}{6}.$$

Can the mixed strategy equilibrium in the star be strictly better than the equilibrium in the line from the introductory example? The answer is no for the following reason. First, notice that the range of b supporting equilibrium strategies has to be such that $p^*, q^* \in [0, 1]$. This implies that a mixed strategy equilibrium exists for $b \in [\frac{1}{12}, \frac{1}{8}]$. But we know that for all $b \leq \frac{1}{8}$, both experts reveal their signals in the star and the in line. Therefore, the above mixed strategy equilibrium in the star cannot strictly dominate the line.

5 Conclusions

This paper shows how an organization should optimally design its communication hierarchies if the players communicate their information strategically. The optimal hierarchies resemble the ones discussed in now classic contributions to the theory of organizations by Cyert and March (1963), Dalton (1959) and Crozier (1963).

I show that an optimal communication hierarchy is shaped by two competing forces. On the one hand, an intermediation force brings experts together and enables strategic coarsening of information. On the other hand, an uncertainty force separates the experts and relaxes their incentives to reveal their signals. Perhaps surprising, simultaneous communication is always dominated by an optimally designed sequential communication where the experts closer to the decision maker are the ones whose bias is closer to the decision maker's bias. Optimal sequential communication, in general, separates the experts into groups of similar bias. If the biases of experts are sufficiently close to one another, and sufficiently different to the bias of the decision maker, the optimal network has a simple requirement of a single intermediary communicating to the decision maker given that all the other experts are interconnected in some way.

In real-life organizations, communication patterns can be very complex. Communication hierarchies are a useful first approximation. For future research it will be important to look at other communication structures where the experts can talk to multiple audiences and where communication networks allow for cycles or varying "strengths" of links.

Communication is a dynamic acticity that usually features multiple rounds of informational exchange. Literature on strategic communication shows that even adding a second round of communication between an informed sender and an uninformed receiver can enlarge the set of equilibrium outcomes compared to a signle round communication (Krishna and Morgan, 2004). It will be interesting to see how the equilibria in optimal hierarchies are affected by communicating dynamics.

Further, it will be important to combine the optimal communication design with the other incentive tools such as transfers, delegation or monitoring.

6 Appendix

Calculations for the leading example:

For n and k successes, the posterior on θ is $f(\theta|k,n) = \frac{(n+1)!}{k!(n-k)!}\theta^k(1-\theta)^{n-k}$. The corresponding expected value is $\mathbb{E}(\theta|k,n) = \frac{k+1}{n+2}$.

1. Equilibrium in the star with one truthful expert.

   If an expert $i \in \{1,2\}$ reveals his signal truthfully, the decision maker's choices are

   $$y(0) = \mathbb{E}_{DM}(\theta|0,1) = \frac{1}{3},\ y(1) = \mathbb{E}_{DM}(\theta|1,1) = \frac{2}{3},$$

   Expert i does not deviate from the truthful revelation of his signal 0 if:

   $$-\int_0^1 (\frac{1}{3} - \theta - b_i)^2 f(\theta|0,1)d\theta \geq -\int_0^1 (\frac{2}{3} - \theta - b_i)^2 f(\theta|0,1)d\theta,$$

   where $f(\theta|0,1) = 2(1-\theta)$. Further, he does not deviate from the truthful revelation of his signal 1 if

   $$-\int_0^1 (\frac{2}{3} - \theta - b_i)^2 f(\theta|1,1)d\theta \geq -\int_0^1 (\frac{1}{3} - \theta - b_i)^2 f(\theta|1,1)d\theta,$$

   where $f(\theta|1,1) = 2\theta$. The above inequalities hold for $|b_i| \leq \frac{1}{6}$.

   Decision maker's utility is

   $$-\frac{1}{2}\int_0^1 (\frac{1}{3} - \theta)^2 2(1-\theta)d\theta - \frac{1}{2}\int_0^1 (\frac{2}{3} - \theta)^2 2\theta d\theta = -\frac{1}{18}.$$

2. Equilibrium in the star with two truthful experts.

   The decision maker receives messages according to the partition $\{\{0\},\{1\},\{2\}\}$, where the sum of successes $k \in \{0,1,2\}$ is the summary statistic. Her optimal choices are

$$y(0) = \frac{1}{4},\ y(1) = \frac{1}{2},\ y(2) = \frac{3}{4}.$$

   Suppose, an expert $i \in \{0,1\}$ receives a signal $s_i = k'$. Then, he assigns probability $\frac{2}{3}$ to the other expert having the same signal k', and probability $\frac{1}{3}$ to the other expert having signal $1-k'$.

   Thus, an expert i does not deviate from the truthful revelation of his signal 0 if:

$$-\frac{2}{3}\int_0^1 (\frac{1}{4}-\theta-b_i)^2 f(\theta|0,2)d\theta - \frac{1}{3}\int_0^1 (\frac{1}{2}-\theta-b_i)^2 f(\theta|1,2)d\theta \geq$$
$$-\frac{2}{3}\int_0^1 (\frac{1}{2}-\theta-b_i)^2 f(\theta|0,2)d\theta - \frac{1}{3}\int_0^1 (\frac{3}{4}-\theta-b_i)^2 f(\theta|1,2)d\theta,$$

   where $f(\theta|0,2) = 3(\theta)^2$ and $f(\theta|1,2) = 6\theta(1-\theta)$. Similarly, an expert i does not deviate from the truthful revelation of his signal 1 if:

$$-\frac{1}{3}\int_0^1 (\frac{1}{2}-\theta-b_i)^2 f(\theta|1,2)d\theta - \frac{2}{3}\int_0^1 (\frac{3}{4}-\theta-b_i)^2 f(\theta|2,2)d\theta \geq$$
$$-\frac{1}{3}\int_0^1 (\frac{1}{4}-\theta-b_i)^2 f(\theta|1,2)d\theta - \frac{2}{3}\int_0^1 (\frac{1}{2}-\theta-b_i)^2 f(\theta|2,2)d\theta,$$

   where $f(\theta|2,2) = 3\theta^2$. Both incentive constraints hold for $|b_i| \leq \frac{1}{8}$.

   To calculate the expected utility of the decision maker notice that the summary statistic $k \in \{0,1,2\}$ is uniformly distributed:

$$Prob(k|n=2) = \int_0^1 Pr(k|\theta, n=2)d\theta = \int_0^1 \frac{2!}{k!(2-k)!}\theta^k(1-\theta)^{2-k}d\theta = \frac{1}{3}.$$

   Therefore, the expected utility of the decision maker is

$$-\frac{1}{3}\int_0^1(\frac{1}{4}-\theta)^2 3(1-\theta)^2 d\theta-\frac{1}{3}\int_0^1(\frac{1}{2}-\theta)^2 6\theta(1-\theta)d\theta-$$

$$-\frac{1}{3}\int_0^1(\frac{3}{4}-\theta)^2 3\theta^2 d\theta=-\frac{1}{24}.$$

3. Equilibrium in the line with two truthful experts perfectly revealing their entire information.

   Think of the following strategy profile: expert 1 truthfully reveals his signal to expert 2, who, in turn, truthfully reveals both his signal and the message of expert 1 to the decision maker. The decision maker receives information according to the partition $\{\{0\},\{1\},\{2\}\}$, where the sum of successes $k\in\{0,1,2\}$ is the summary statistic. Her optimal choices are the same as in the case of two truthful experts in the star:

$$y(0)=\frac{1}{4},\ y(1)=\frac{1}{2},\ y(2)=\frac{3}{4}.$$

   The incentive constraints for expert 2 depend on his signal and the message from expert 1. Expert 2's information sets can be represented by the summary statistic $\{0,1,2\}$ which reflects the sum of successes of his private signal and the message of expert 1. Without loss of generality, assume that his message set is $\{0,1,2\}$ such that in equilibrium his message truthfully reveals his summary statistic. If his summary statistic is 0, he has no incentives to deviate to the next highest message 1 (and therefore not to deviate to an even higher message 2) if

$$-\int_0^1(\frac{1}{4}-\theta-b_1)^2 f(\theta|0,2)d\theta\geq-\int_0^1(\frac{1}{2}-\theta-b_1)^2 f(\theta|0,2)d\theta.$$

   If his summary statistic is 1, he has no incentives to deviate upwards to the message 2 if

$$-\int_0^1(\frac{1}{2}-\theta-b_1)^2 f(\theta|1,2)d\theta\geq-\int_0^1(\frac{3}{4}-\theta-b_1)^2 f(\theta|1,2)d\theta$$

   and no incentives to deviate downwards to the message 0 if

$$-\int_0^1 (\frac{1}{2}-\theta-b_1)^2 f(\theta|1,2)d\theta \geq -\int_0^1 (\frac{1}{4}-\theta-b_1)^2 f(\theta|1,2)d\theta.$$

Finally, if his summary statistic is 2, he has no incentives to deviate to a lower message 1 (and, thus, no incentives to deviate to an even lower message 0) if

$$-\int_0^1 (\frac{3}{4}-\theta-b_1)^2 f(\theta|2,2)d\theta \geq -\int_0^1 (\frac{1}{2}-\theta-b_1)^2 f(\theta|2,2)d\theta.$$

Combining all incentive constraints for expert 1 we obtain:

$$|b_2| \leq \frac{1}{8}.$$

Finally, the incentive constraints for expert 1 are the same as in the star with two truthful experts: in both cases he has the same expectation over the decision maker's summary statistic. This implies $|b_1| \leq \frac{1}{8}$.

The expected utility of the decision maker is the same as in the star with both experts truthfully revealing their signals, which is $-\frac{1}{24}$.

4. Equilibrium in the star in which expert 2 strategically coarsens his information.

   In this equilibrium the decision maker receives her information according to the partition $\{\{0\},\{1,2\}\}$ where the sum of successes $k \in \{0,1,2\}$ is the summary statistic of experts' signals. If the decision maker is informed that the signals are contained within the pool $p := \{1,2\}$, his posterior of θ is formed by Bayes rule:

$$f(\theta|k \in p, n) = \frac{Pr(p|\theta)}{\int_0^1 Pr(p|\theta)d\theta} = \frac{\sum_{k\in p} Pr(k)f(k|\theta,n)}{\int_0^1 \sum_{k\in p} Pr(k)f(k|\theta,n)d\theta}.$$

   where $f(k|\theta,n) = \frac{n!}{k!(n-k)!}\theta^k(1-\theta)^{n-k}$. Using the fact that $\int_0^1 \theta^k(1-\theta)^{n-k} = \frac{k!(n-k)!}{(n+1)!}$, we obtain:

$$f(\theta|k \in p, n=2) = \frac{3}{2}\theta(2-\theta),$$

   and therefore

$$\mathbb{E}(\theta|k \in p, n) = \int_0^1 \frac{3}{2}\theta^2(2-\theta)d\theta = \frac{5}{8}.$$

Thus, decision maker's choices contingent on the received messages are
$y(0) = \frac{1}{4}$ an $y(1,2) = \frac{5}{8}$.

Given truthful communication from expert 1 to expert 2, the latter knows the sum of successes of both signals. I denote expert 2's first message by 0, and his second message by $(1,2)$. Suppose, first, that expert 2 believes that the sum of successes is 0. He has no incentives to deviate from the truthful message 0 to the other message $(1,2)$ if

$$-\int_0^1 (\frac{1}{4} - \theta - b_1)^2 f(\theta|0,2)d\theta \geq -\int_0^1 (\frac{5}{8} - \theta - b_1)^2 f(\theta|0,2)d\theta.$$

If his summary statistic is 1, truthful communication requires that he sends a message $(1,2)$. He has no incentives to deviate to the lower message 0 if

$$-\int_0^1 (\frac{5}{8} - \theta - b_1)^2 f(\theta|1,2)d\theta \geq -\int_0^1 (\frac{1}{4} - \theta - b_1)^2 f(\theta|1,2)d\theta.$$

Finally, if his summary statistic is 2, truthful communication requires that he sends message $(1,2)$. He has no incentives to deviate to the lower message 0 if

$$-\int_0^1 (\frac{5}{8} - \theta - b_1)^2 f(\theta|2,2)d\theta \geq -\int_0^1 (\frac{1}{4} - \theta - b_1)^2 f(\theta|2,2)d\theta.$$

The above incentive constraints hold for

$$-\frac{1}{16} \leq b_2 \leq \frac{3}{16}.$$

Finally, assuming that expert 2 communicates according to the partition $\{\{0\},\{1,2\}\}$ (which is part of the equilibrium strategy profile which we fixed above), expert 1 has no incentives to deviate from truthfully communicating his signal 0 if

$$-\frac{2}{3}\int_0^1 (\frac{1}{4} - \theta - b_2)^2 f(\theta|0,2)d\theta - \frac{1}{3}\int_0^1 (\frac{5}{8} - \theta - b_2)^2 f(\theta|1,2)d\theta \geq$$

$$-\frac{2}{3}\int_0^1(\frac{5}{8}-\theta-b_2)^2 f(\theta|0,2)d\theta-\frac{1}{3}\int_0^1(\frac{5}{8}-\theta-b_2)^2 f(\theta|1,2)d\theta.$$

Finally, expert 1 has no incentives to deviate from truthfully communicating his signal 1 if

$$-\{\frac{2}{3}\int_0^1(\frac{5}{8}-\theta-b_1)^2 f(\theta|1,2)d\theta+\frac{1}{3}\int_0^1(\frac{5}{8}-\theta-b_1)^2 f(\theta|2,2)d\theta\}\geq$$
$$-\{\frac{2}{3}\int_0^1(\frac{5}{8}-\theta-b_1)^2 f(\theta|1,2)d\theta+\frac{1}{3}\int_0^1(\frac{5}{8}-\theta-b_1)^2 f(\theta|2,2)d\theta\}.$$

The above incentive constraints hold for

$$-\frac{1}{16}\leq b_1\leq\frac{3}{16}.$$

The expected utility of the decision maker is

$$-\frac{1}{3}\int_0^1(\frac{1}{4}-\theta)^2 3(1-\theta)^2 d\theta-\frac{2}{3}(\frac{5}{8}-\theta)^2\frac{3}{2}\theta(2-\theta)d\theta=-\frac{5}{96}.$$

Remaining definitions from the model section:

In this paper I study directed graphs with the following properties:

(1) for each $i\in N^e$, there is $j\in N$, $j\neq i$, with $e_{ij}=1$, and there is no other $j'\in N$, $j'\neq j$, such that $e_{ij'}=1$. This means that every expert has a single outgoing link.

(2) Take any $i\in N^e$. There is no path $H_{ij}(Q)$, $j\in N$ with $j\neq i$, such that $e_{ji}=1$. This means that there are no cycles.

(3) $\sum_{j\in N^e}e_{jDM}\geq 1$ and $\sum_{j\in N^e}e_{DMj}=0$ which means that the decision maker has at least one incoming link but no outgoing links.

Bayesian updating follows the Beta-binomial model: given n observations and k successes (sum of the signals), the conditional pdf is:

$$f(k|\theta, n) = \frac{n!}{k!(n-k)!}\theta^k(1-\theta)^{n-k}.$$

The distribution of k's is uniform:

$$Prob(k|n) = \int_0^1 Prob(k|\theta, n)d\theta = \int_0^1 \frac{n!}{k!(n-k)!}\theta^k(1-\theta)^{n-k}d\theta = \frac{1}{n+1}.$$

The posterior is

$$f(\theta|k, n) = \frac{(n+1)!}{k!(n-k)!}\theta^k(1-\theta)^{n-k},$$

Thus, $\mathbb{E}(\theta|k, n) = \frac{k+1}{n+2}$.

For n trials with k successes the probability of having j successes from m additional trials is:

$$P(j|m, n, k) = \int_0^1 P(j|m, \theta)P(\theta|k, n)d\theta =$$

$$\int_0^1 \frac{m!}{j!(m-j)!}\theta^j(1-\theta)^{m-j}\frac{(n+1)!}{k!(n-k)!}\theta^k(1-\theta)^{n-k}d\theta =$$

$$\int_0^1 \frac{m!}{j!(m-j)!}\frac{(n+1)!}{k!(n-k)!}\theta^{k+j}(1-\theta)^{m-j+n-k}d\theta =$$

$$\frac{m!}{j!(m-j)!}\frac{(n+1)!}{k!(n-k)!}\frac{(k+j)!(m-j+n-k)!}{(n+m+1)!}.$$

Finally, when a partition P is adoped, the expected utility of type $t_i \in T_i(Q)$ of player $i \in N$ can be written as

$$-\sum_{p\in P} Pr(p|t_i)\int_0^1 (y(p) - \theta - b_i)^2 f(\theta|p)d\theta,$$

where $Pr(p) = \sum_{k\in p} Pr(k)$.
The posterior $f(\theta|p)$ is calculated by Bayes rule:

$$f(\theta|p) = \frac{Pr(p|\theta)}{\int_0^1 Pr(p|\theta)d\theta} =$$

$$\frac{\sum_{k\in p} Pr(k)\frac{n!}{k!(n-k)!}\theta^k(1-\theta)^{n-k}}{\int_0^1 \sum_{k\in p} Pr(k)\frac{n!}{k!(n-k)!}\theta^k(1-\theta)^{n-k}d\theta}.$$

Given $\int_0^1 \theta^k(1-\theta)^{n-k}d\theta = \frac{k!(n-k)!}{(n+1)!}$, we have

$$f(\theta|p) = \frac{\sum_{k\in p} Pr(k)\frac{n!}{k!(n-k)!}\theta^k(1-\theta)^{n-k}}{\sum_{k\in p} Pr(k)\frac{n!}{k!(n-k)!}\frac{k!(n-k)!}{(n+1)!}} = \frac{\sum_{k\in p} Pr(k)\frac{(n+1)!}{k!(n-k)!}\theta^k(1-\theta)^{n-k}}{\sum_{k\in p} Pr(k)}$$

But since $f(\theta|k,n) = \frac{(n+1)!}{k!(n-k)!}\theta^k(1-\theta)^{n-k}$, we can rewrite the above expression as

$$f(\theta|p) = \frac{\sum_{k\in p} Pr(k) f(\theta|k,n)}{\sum_{k\in p} Pr(k)}.$$

Thus, the expected utility can be expressed as:

$$-\sum_{p\in P}\sum_{k\in p} Pr(k|t_i)\int_0^1 (y(p)-\theta-b_i)^2 f(\theta|k,n)d\theta,$$

where $y(p) = \mathbb{E}(\theta|p)$ which is

$$\mathbb{E}(\theta|p) = \int_0^1 \theta f(\theta|p)d\theta = \frac{1}{\sum_{k\in p} Pr(k)}\sum_{k\in p} Pr(k)\frac{(n+1)!}{k!(n-k)!}\frac{(k+1)!(n-k)!}{(n+2)!} =$$

$$\frac{1}{\sum_{k\in p} Pr(k)}\sum_{k\in p} Pr(k)\mathbb{E}(\theta|k,n)$$

and $\mathbb{E}(\theta|k,n) = \frac{k+1}{n+2}$.

Proof of Proposition 1:

Take a star network with n experts and denote it by Q^s. Fix a strategy profile $P_1(Q^s)\times .. \times P_{n'}(Q^s)$ in which $n' \leq n$ experts communicate their signals truthfully such that $P_i(Q^s) = \{\{0\},\{1\}\}$ for each $i \in \{1,..,n'\}$. We derive conditions on experts' biases for the above strategy profile to be an equilibrium.

Given the above strategy profile, the decision maker receives her information according to the partition of the summary statistics (which is the sum of the signals), which is $\{\{0\},\{1\},..,\{n'\}\}$. For $k \in \{0,..,n'\}$, the sequentially rational decision maker chooses

$$y(k) = \mathbb{E}(\theta|k,n') = \frac{k+1}{n'+2}.$$

Denote by t_{-i} the vector of types of all truthful experts rather than i, expressed in terms of their summary statistic, $t_{-i} \in \{0,..,n'-1\}$. Let $y(p_{-i},p_i)$ be the action profile of the decision maker if she receives message $p_i \in \{0,1\}$ from expert i and the messages of all other experts, $p_{-i} = t_{-i}$. If $t_i = 0$, expert i truthfully reveals his signal 0 if

$$-\Big(\mathbb{E}_i(t_{-i}=0|t_i=0)\int_0^1 (y(0,0)-\theta-b_i)^2 f(\theta|k=0,n)d\theta + ...$$
$$+\mathbb{E}_i(t_{-i}=n'-1|t_i=0)\int_0^1 (y(n'-1,0)-\theta-b_i)^2 f(\theta|k=n'-1,n')d\theta\Big) \geq$$
$$-\Big(\mathbb{E}_i(t_{-i}=0|t_i=0)\int_0^1 (y(0,1)-\theta-b_i)^2 f(\theta|k=0,n')d\theta + ...$$
$$+\mathbb{E}_i(t_{-i}=n'-1|t_i=0)\int_0^1 (y(n'-1,1)-\theta-b_i)^2 f(\theta|k=n'-1,n')d\theta\Big).$$

Using decision maker's optimal choices, the above inequality can be written as

$$-\Big(\mathbb{E}_i(t_{-i}=0|t_i=0)\int_0^1 (\frac{1}{n'+2}-\theta-b_i)^2 f(\theta|k=0,n)d\theta + ...$$
$$+\mathbb{E}_i(t_{-i}=n'-1|t_i=0)\int_0^1 (\frac{n'}{n'+2}-\theta-b_i)^2 f(\theta|k=n'-1,n')d\theta\Big) \geq$$
$$-\Big(\mathbb{E}_i(t_{-i}=0|t_i=0)\int_0^1 (\frac{2}{n'+2}-\theta-b_i)^2 f(\theta|k=0,n')d\theta + ...$$
$$+\mathbb{E}_i(t_{-i}=n'-1|t_i=0)\int_0^1 (\frac{n'+1}{n'+2}-\theta-b_i)^2 f(\theta|k=n'-1,n')d\theta\Big),$$

which can be rewritten as

$$\sum_{k=0}^{n'-1} \mathbb{E}(t_{-i}=k|t_i=0)\frac{1}{n'+2}\,(\frac{k+1}{n'+2}+\frac{k+2}{n'+2}-2\frac{k+1}{n'+2}-2b_i) \geq 0.$$

The above inequality holds for

$$b_i \leq \frac{1}{2(n'+2)}.$$

Similarly, if $t_i = 1$, expert i communicates his signal truthfully if

$$-\Big(\mathbb{E}_i(t_{-i} = 0|t_i = 1)\int_0^1 (\frac{2}{n'+2} - \theta - b_i)^2 f(\theta|k = 1, n')d\theta + ...$$

$$+\mathbb{E}_i(t_{-i} = n' - 1|t_i = 1)\int_0^1 (\frac{n'+1}{n'+2} - \theta - b_i)^2 f(\theta|k = n', n')d\theta\Big) \geq$$

$$-\Big(\mathbb{E}_i(t_{-i} = 0|t_i = 1)\int_0^1 (\frac{1}{n'+1} - \theta - b_i)^2 f(\theta|k = 1, n')d\theta + ...$$

$$+\mathbb{E}_i(t_{-i} = n' - 1|t_i = 1)\int_0^1 (\frac{n'}{n'+2} - \theta - b_i)^2 f(\theta|k = n', n')d\theta\Big),$$

which can be rewritten as:

$$\sum_{k=0}^{n'-1} \Pr(t_{-i} = k|t_i = 1)\frac{1}{n'+2}\ (\frac{k+1}{n'+2} + \frac{k+2}{n'+2} - 2\frac{k+2}{n'+2} - 2b_i) \geq 0.$$

The above inequality holds for

$$b_i \geq -\frac{1}{2(n'+2)}.$$

Summing up, there is an equilibrium with $n' \leq n$ truthful experts iff

$$|b_i| \leq \frac{1}{2(n'+2)}$$

for each $i \in \{1, .., n'\}$.
Q.E.D.

Proof of Proposition 2:

Here is the proof of the first part.

Fix a set of experts $N = \{1, .., n\}$, $n \geq 2$, and a bias profile $b(n) = (b_1, .., b_n)$ such that there exists at least one equilibrium in the star in which at least one of the experts is not babbling. Take any of the non-babbling equilibria and denote the set of truthful experts by N' with $|N'| = n'$. From Proposition 1 we know that it implies $|b_i| \leq \frac{1}{2(n'+2)}$ for each $i \in N'$.

Fix any tree network $Q \in \mathbb{Q}$ which is not a star such that the experts are ordered monotonically according to the absolute value of their biases: if expert i reports to expert j, then $|b_i| \geq |b_j|$. In the following I show that Q has the same equilibrium outcomes as the star with n' truthful experts. Think of a strategy profile in Q in which every experts within N' perfectly reveals his type to his successor in the network. It means, for any $t_i \in T_i(Q)$, $i \in N'$, $P_i(Q) = T_i(Q)$. Upon observing experts' messages, the decision maker chooses $y = \frac{k+1}{n+2}$, where $k \in \{0, .., n'\}$ is the decision maker's summary statistic.

Take any expert $j \in N'$. Using the notation $\tilde{n} = |\tilde{N}_j(Q)|$, the type space of j can be represented as $T_j(Q) = \{0, 1, .., \tilde{n} + 1\}$.[22] Expert j does not observe the signals of $n - (\tilde{n} + 1)$ experts. The type set of those experts, denoted by $T_{-j}(Q)$, can be expressed in terms of the summary statistic as $T_{-j}(Q) = \{0, .., n - (\tilde{n} + 1)\}$.

The incentive constraint which ensures that a type $t_j = k' \in \{0, .., \tilde{n} + 1\}$ sends a truthful message $p_j = k'$ instead of deviating (upward) to the next highest message $k' + 1$ is

$$\sum_{l=0}^{n'-(\tilde{n}+1)} \mathbb{E}(t_{-j} = l | t_j = k') \Big(y(k' + 1 + l) - y(k' + l) \Big)$$
$$\Big(y(k' + 1 + l) + y(k' + l) - 2\mathbb{E}(\theta | k' + l, n') - 2b_j) \Big) \geq 0$$

Notice that

$$y(k' + 1 + l) - y(k' + l) = \frac{k' + l + 2}{n' + 2} - \frac{k' + l + 1}{n' + 2} = \frac{1}{n' + 2},$$

and

$$y(k'+1+l) + y(k'+l) - 2\mathbb{E}(\theta | k'+l, n') = \frac{k' + l + 2}{n' + 2} + \frac{k' + l + 1}{n' + 2} - 2\frac{k' + l + 1}{n' + 2} = \frac{1}{n' + 2}.$$

Given that $\sum_{l=0}^{n'-(\tilde{n}+1)} \mathbb{E}(t_{-j} = l | t_j = k')$, the incentive constraint implies

$$b_j \leq \frac{1}{2(n' + 2)}.$$

[22] Remember that $\tilde{N}_j(Q)$ is defined as the set of all non-babbling experts on all paths in Q leading to j.

Similarly, the incentive constraint which ensures that a type $t_j = k' \in \{0,..,n'\}$ sends truthful message $p_j = k'$ instead of deviating (downward) to the next lower message $k'-1$ is

$$\sum_{l=0}^{n'-(\tilde{n}+1)} \mathbb{E}(t_{-j}=l|t_j=k')\Big(y(k'-1+l)-y(k'+l)\Big)$$
$$\Big(y(k'-1+l)+y(k'+l)-2\mathbb{E}(\theta|k'+l,n')-2b_j)\Big)\geq 0$$

Notice that

$$y(k'-1+l)-y(k'+l)=\frac{k'+l}{n'+2}-\frac{k'+l+1}{n'+2}=-\frac{1}{n'+2}$$

and

$$y(k'-1+l)+y(k'+l)-2\mathbb{E}(\theta|k'+l,n')=\frac{k'+l}{n'+2}+\frac{k'+l+1}{n'+2}-2\frac{k'+l+1}{n'+2}=-\frac{1}{n'+2}.$$

This implies

$$b_j \geq -\frac{1}{2(n'+2)}.$$

which proves the first part of the Proposition.

Here is the proof of the second part.

Fix a set of experts $N^e = \{1,..,n\}$, $n \geq 2$, and a bias profile $b(n) = (b_1,..,b_n)$ such that for each $i \in N^e$, $b_i > \frac{1}{6}$. Proposition 1 tells us that the only equilibrium in the star is the babbling equilibrium which results in $\mathbb{E}U_{DM} = -\int_0^1(\frac{1}{2}-\theta)^2 d\theta = -\frac{1}{12}$.

Take any network $Q \in \mathbb{Q}$ which is not a star and denote the number of experts which belong to the longest path in Q, starting with some expert $j \in N^e$ and ending with the decision maker, $H_{jDM}(Q)$, by $r+1$. Since Q is not a star, it must be $r \geq 2$.

Denote the expert on $H_{jDM}(Q)$ who is directly connected to the decision maker by i', $e_{i'DM} = 1$. Think of a strategy profile in Q in which all experts on $H_{jDM}(Q)$ apart from i' perfectly reveal their types, expert i' communicates to the decision maker according to the partition $P_{i'}(Q) = \{\{0\},\{1,..,r\}\}$, and all the remaining experts babble.

The actions of the decision maker depend on the communicated pool of $P_{i'}(Q)$ as follows:

$$y(0) = \frac{1}{r+2},\ y(1,..,r) = \Big(\frac{2}{r+2} + .. + \frac{r+1}{r+2}\Big)\frac{1}{r} = \frac{r+3}{2(r+2)}.$$

Notice that in case of an upward deviation, all experts on $H_{jDM}(Q)$ condition their deviation on the same event that the sum of all signals is 0. The corresponding incentive constraint for each $i \in H_{jDM}(Q)$ has a simple form:

$$-\int_0^1 (y(0) - \theta - b_i)^2 f(\theta|k=0,r)d\theta \geq \int_0^1 (y(1,..,r) - \theta - b_i)^2 f(\theta|k=0,r)d\theta$$

which implies $b_i \leq \frac{r+1}{4(r+2)}$. Notice that:

$$\frac{r+1}{4(r+2)} > \frac{1}{6} \text{ for } r \geq 2,$$

In a similar way we obtain that the incenive preventing a deviation of every $i \in H_{jDM}(Q)$ to a lower message results in $b_i \geq \frac{r-3}{4(r+2)}$. It proves that $P_{i'}(Q)$ can be supported for biases strictly larger than $\frac{1}{6}$.

It remains to show that this partition results in a higher expected utility for the decision maker, compared to the strategy profile in which every expert babbles. The expected utility of the decision maker in Q is:

$$\mathbb{E}U_{DM}(P_{DM}(Q) = P_{i'}(Q)) = -\frac{1}{3} + \frac{1}{(r+1)(n+2)^2}\Big(1 + \frac{(2+..+(r+1))^2}{r}\Big) = -\frac{4+r+r^2}{12(2+r)^2}$$

which is strictly higher than $-\frac{1}{12}$ for $r \geq 2$. We conclude that if the biases of all experts on $H_{jDM}(Q)$ are within the interval $[\frac{r-3}{4(r+2)}, \frac{r+1}{4(r+2)}] \cap (\frac{1}{6}, \infty]$, then the network Q has at least one equilibrium which strictly dominates the (uninformative) equilibrium in the star.

Q.E.D.

Proof of Proposition 3:

The proof is lengthy but follows a straightforward structure. First, I show that among all implementable two-pool partitions of $\{0,1\}^n$ according to which the decision maker receives her information, it is the partition $\{\{0\},\{1,..,n\}\}$ (where

$k \in \{0, .., n\}$ is the decision maker's summary statistic) which results in the largest shift in decision maker's policy upon an upward deviation by any of the experts. Therefore, it can be supported by the largest possible biases. For this partition, I obtain the corresponding upper bias threshold $\bar{b}_1 = \frac{n+1}{4(n+2)}$. Second, I show that any other decision maker's two-pool partition of $\{0,1\}^n$ is supported by the upper bias of at most $\frac{(n+3)(n-1)}{4(n+1)(n+2)} < \bar{b}_1$. Third, I show that the largest shift in decision maker's policy with $n' < n$ experts, conditional on decision makers' two-pool partitions of $\{0,1\}^n$, results in the upper threshold $\bar{b}_2 = \frac{n}{4(n+1)}$, with $\frac{(n+3)(n-1)}{4(n+1)(n+2)} < \bar{b}_2 < \bar{b}_1$. The partition $\{\{0\}, \{1, .., n\}\}$ can only be implemented in a network specified in the Proposition. Forth, I show that a decision maker's partition $\{0,1\}^n$ with at least three pools cannot generate an equilibrium which supports biases higher than $\bar{b}_2$ (Claim 3). It follows that, if the biases of all n experts are in the interval $(\bar{b}_2, \bar{b}_1]$, then only a network which satisfies the requirements of the Proposition can transmit any information from the experts to the decision maker.

At the start of the proof, I prove two Claims which provide a useful characterization of experts' message strategies.

Claim 1: *Take an expert $i \in N^e$ in a network Q who communicates his information to a player $j \in N$ according to an equilibrium partition $P_i'(Q)$ consisting of two pools, p_1, p_2, with $\mathbb{E}_j(\theta|p_2) > \mathbb{E}_j(\theta|p_1)$. Take any $t_i \in p_1$ and any $t_i' \in p_2$. Then, it must be true that $\mathbb{E}_i(\theta|t_i) < \mathbb{E}_i(\theta|t_i')$.*

Proof of Claim 1: Suppose not such that $\mathbb{E}_i(\theta|t_i) \geq \mathbb{E}_i(\theta|t_i')$. But then, if t_i' has no incentives to deviate from sending p_2, the type t_i would deviate from sending p_1 and send p_2 instead, a contradiction. *Q.E.D.*

The next Claim shows that in any $Q \in \mathbb{Q}$, the sums of the signals in any two-pool partition according to which the decision maker receives her signals are "nicely" ordered.

Claim 2: Take any optimal network $Q \in \mathbb{Q}$ with $n' \leq n$ non-babbling experts such that $|N_{DM}(Q)| = 1$. Denote the single expert who communicates to the decision maker by i. If an equilibrium partition of i, $P_i(Q)$, consists of two pools p_1, p_2 with $\mathbb{E}_{DM}(\theta|p_1) < \mathbb{E}_{DM}(\theta|p_2)$, and p_2 contains an element of $\{0,1\}^{n'}$ with the sum of the signals k, $k < n'$, then p_2 contains at least one element of $\{0,1\}^{n'}$ with the sum of the signals $k+1$.

Proof of Claim 2: First, suppose that i's type space can be represented by $T_i = \{0, 1, .., n'\}$ which means that i observes the exact sum of n' signals. According to Claim 1, all the elements of $\{0, 1\}^{n'}$ with the same sum of the signals can only be part of the same pool. With other words, if expert i is the only expert who pools together his information sets, it has to be the case that all elements with the same sum of the signals are either in p_1 or in p_2. Therefore, the only way that the elements of $\{0, 1\}^{n'}$ with the same sum of the signals are distributed between two pools is that there exists a path $H_{ji}(Q)$, $j \in N^e$, such that:

1. there is $i' \in H_{ji}(Q)$ (which can be i) such that $|H_{ji'}(Q)| > 2$,
2. for every $j' \in H_{ji}(Q)$, $P_{j'}(Q) \neq P^b_{j'}(Q)$ which means that none of the experts on the path $H_{ij}(Q)$ babbles,
3. for every $\tilde{l} \in N_{i'}(Q)$, $T_{\tilde{l}}(Q) = \{0, 1, .., |\tilde{N}_{\tilde{l}}(Q)|\}$ which means that all experts within $N_{i'}(Q)$ receive uncoarsed information, and
4. there is at least one $l \in N_{i'}(Q)$ who's type space can be represented by $\{0, 1, .., |\tilde{N}_l(Q)|\}$ - which simply means that l knows the summary statistic of the signals of all non-babbling experts on all possible paths in Q leading to l - and that there exists $p \in P_l(Q)$ and two different types $t_l, t'_l \in T_l(Q)$, $\mathbb{E}(\theta|t_l) \neq \mathbb{E}(\theta|t'_l)$, such that $t_l \in p$ and $t'_l \in p$, which means that l strategically coarsens his information.

Notice that the type space of i', $T_{i'}(Q) = \{0, 1\} \times \Pi_{i \in N_{i'}(Q)} P_i(Q)$, has two properties. First, there are at least two different types $t_{i'}, t'_{i'} \in T_{i'}(Q)$ such that both of them include at least one element of $\{0, 1\}^{|\tilde{N}_{i'}(Q)|+1}$ with the same sum of the signals. Second, if a type $t \in T_{i'}(Q)$ includes at least one element of $\{0, 1\}^{|\tilde{N}_{i'}(Q)|+1}$ with the sum of the signals k, and at least one element of $\{0, 1\}^{|\tilde{N}_{i'}(Q)|+1}$ with the sum of the signals $k+2$, then it has to include at least one element of $\{0, 1\}^{|\tilde{N}_{i'}(Q)|}$ with the sum of the signals $k+1$. This happens because the type space of i' is a product of partitions with the property that for any $\hat{i} \in N_{i'}(Q)$, either $P_{\hat{i}}(Q) = T_{\hat{i}}(Q) = \{0, .., |\tilde{N}_{\hat{i}}(Q)|\}$ or $P_{\hat{i}}(Q)$ is an incentive-compatible coarsening of $\{0, .., |\tilde{N}_{\hat{i}}(Q)|\}$.

Therefore, it must be true that the type space of each $l' \in H_{i'DM}$ is such that if there is a type $t \in T_{l'}(Q)$ which includes at least one element of $\{0, 1\}^{|\tilde{N}_{l'}(Q)|+1}$ with the sum of the signals k, and at least one element of $\{0, 1\}^{|\tilde{N}_{l'}(Q)|+1}$ with the sum of 1's $k+2$, then it has to include at least one element of $\{0, 1\}^{|\tilde{N}_{l'}(Q)|+1}$ with the sum of the signals $k+1$. But then, since $P_i(Q)$ has to be incentive-compatible, the ordering of the sums of the signals within $p_1, p_2 \in P_i(Q)$ has to satisfy conditions stated in the Claim. *Q.E.D.*

For the next steps I assume that a partition of $\{0,1\}^n$ according to which the decision maker receives her information consists of two pools. Given Claims 1 and 2, the set of such two-pool partitions can be represented by a family of partitions

$$P^k_{DM}(Q) = \{\{0,..,k\},\{k+1,..,n\}\},\ 0 \le k \le n-1,$$

and

$$\begin{gathered} P^{z,k}_{DM}(Q,\omega) := \{\{0,..,(k-z)\omega_0,(k-z+1)\omega_1,..,(k)\omega_z\}, \\ \{(k-z)(1-\omega_0),(k-z+1)(1-\omega_1),..,(k)(1-\omega_z),..,n\}\}, \\ 0 \le z \le k-1,\ 1 \le k \le n-1,\ \omega_i = \frac{n_{k-z+i}}{\binom{n}{k-z+i}},\ \omega = (\omega_0,..,\omega_z), \end{gathered}$$

where $P^k_{DM}(Q)$ is the partition of $\{0,1\}^n$, n_{k-z+i}, $i \in \{0,..,z\}$, is the number of the elements of $\{0,1\}^n$ with the sum of the signals $k-z+i$ which is featured in the first pool. Since $1 \le n_{k-z+i} \le \binom{n}{k-z+i}$, for every $i \in \{0,..,z\}$, $\omega_i \le 1$ and there is at least one $j \in \{0,..,z\}$ for which $\omega_j < 1$. Further, $Q \in \mathbb{Q}$ is such that $|N_{DM}(Q)| = 1$ since otherwise the decision maker's partition contains strictly more than 2 pools (assuming that all experts connected to the decision maker are not babbling). In the following I denote the expert who is directly connected to the decision maker by i'.

For tractability I write P^k and $P^{z,k}$ instead of $P^k_{DM}(Q)$ and $P^{z,k}_{DM}(Q,\omega)$ unless necessary. In the next step I show that among all decision maker's two-pool partitions of $\{0,1\}^n$ that are implementable in any $Q \in \mathbb{Q}$, the partition which supports the largest possible experts' biases is $P^0 = \{\{0\},\{1,..,n\}\}$. For P^0 the corresponding choices of the decision maker are:

$$y(p^0_1) = \frac{1}{n+2},\ y(p^0_2) = \frac{1}{n}\sum_{i=1}^{n}\frac{i+1}{n+2} = \frac{3+n}{2(n+2)},\ p^0_1 = \{0\},\ p^0_2 = \{1,..,n\}$$

(1) First, I show that P^0 has a higher upper bias threshold compared to P^k for any $k > 0$. In case of P^k the corresponding choices of the decision maker are:

$$y(p^k_1) = \frac{1}{k+1}\sum_{i=0}^{k}\frac{i+1}{n+2} = \frac{2+k}{2(n+2)},\ y(p^k_2) = \frac{1}{n-k}\sum_{i=k+1}^{n}\frac{i+1}{n+2} = \frac{3+k+n}{2(n+2)},$$

where $p_1^k = \{0, .., k\}$ and $p_2^k = \{k+1, .., n\}$. In case of P^0 (P^k) every expert conditions his upper deviation on all signals being 0 (k). As a result, in case of P^0 (P^k) every experts' strategy is constrained by the upper bias threshold denoted by $\overline{b}^0$ ($\overline{b}^k$). I aim to show that $\overline{b}^0 > \overline{b}^k$.

If $\overline{b}^0 > \overline{b}^k$, the following must be true for every $i \in N^e$

$$\int_0^1 (y_2(p_2^0) - \theta - \overline{b}^0) f(\theta|0,n) - \int_0^1 (y_1(p_1^0) - \theta - \overline{b}^0) f(\theta|0,n) \geq$$
$$\int_0^1 (y_2(p_2^k) - \theta - \overline{b}^k) f(\theta|k,n) - \int_0^1 (y_1(p_1^k) - \theta - \overline{b}^k) f(\theta|k,n).$$

The inequality is satisfied if the following in true:

$$y(p_1^0) + y(p_2^0) - 2\mathbb{E}(\theta|0,n) > y(p_1^k) + y(p_2^k) - 2\mathbb{E}(\theta|k,n). \tag{1}$$

Substituting the policy decisions and the expected values of θ we get:

$$\frac{1}{n+2} + \frac{3+n}{2(n+2)} - 2\frac{1}{n+2} - \Big(\frac{2+k}{2(n+2)} + \frac{3+k+n}{2(n+2)} - 2\frac{k+1}{n+2}\Big) = \frac{k}{n+2} > 0$$

which shows that (1) is true.

(2) Next, fix any $P_{DM}^{z,k}(Q,\omega)$ and - with a bit of notational abuse - denote the corresponding two pools by $\hat{p}_1$ and $\hat{p}_2$. The corresponding choices of the decision maker are:

$$y(\hat{p}_1) = \frac{\frac{1}{n+2}\Big(1 + .. + (k-z) + (k-z+1)\omega_0 + ..(k+1)\omega_z\Big)}{1 + .. + \omega_0 + .. + \omega_z}$$

$$y(\hat{p}_2) = \frac{\frac{1}{n+2}\Big((k-z+1)(1-\omega_0) + ..(k+1)(1-\omega_z) + (k+2) + .. + (n+1)\Big)}{(1-\omega_0) + .. + (1-\omega_z) + (n-k)}$$

If the decision maker receives message $\hat{p}_1$, he forms a posterior over the elements of $\hat{p}_1$ denoted by $\Delta_{DM}(\hat{p}_1)$. Take the expert i' who is connected to the decision maker in Q. His message strategy is either to send the message $\hat{p}_1$ or the other message $\hat{p}_2$. Denote the type of i' which determines the upper bias threshold of expert i' by $t'_{i'}$. Suppose that $\hat{p}_1$ consists of a single type. Then $t'_{i'} = \hat{p}_1$ and $\Delta_{i'}(t_{i'}) = \Delta_{DM}(\hat{p}_1)$ where $\Delta_{i'}(t'_{i'})$ is the probability distribution of $t_{i'}$ over the elements of $\hat{p}_1$. This implies

$$\sum_{k\in\hat{p}_1} Pr(k|t'_{i'})\mathbb{E}(\theta|k,n) = \sum_{k\in\hat{p}_1} Pr^{DM}(k)\mathbb{E}(\theta|k,n),\ \ Pr^{DM}(k\in\hat{p}_1)\in\Delta_{DM}(\hat{p}_1). \tag{2}$$

Suppose that there is a set of types $T_1 := \{t_{i'}\ :\ \ t_{i'} \in T_{i'}(Q),\ t_{i'} \in \hat{p}_1\}$ with $|T_1| > 1$. Take a type within T_1 with the highest conditional expectation of θ, which is $t'_{i'} \in T_1$, such that there is no $\tilde{t}_{i'} \in T_1$ with $\mathbb{E}(\theta|\tilde{t}_{i'}) > \mathbb{E}(\theta|t'_{i'})$. Notice that in this case $\Delta_{i'}(t'_{i'})$ first order stochastically dominates $\Delta_{DM}(\hat{p}_1)$ implying

$$\sum_{k\in\hat{p}_1} Pr(k|t'_{i'})\mathbb{E}(\theta|k,n) > \sum_{k\in\hat{p}_1} Pr^{DM}(k)\mathbb{E}(\theta|k,n),\ \ Pr^{DM}(k\in\hat{p}_1)\in\Delta_{DM}(\hat{p}_1). \tag{3}$$

Denote the upper bias of expert i' supporting his strategy $P_{i'}(Q) = P^{z,k}_{DM}(Q,\omega)$ by $\overline{b}_{i'}$. I aim to show that $\overline{b}^{1,0} > \overline{b}_{i'}$, where $\overline{b}_{i'}$ is implicitly defined by

$$-\sum_{k\in\hat{p}_1} Pr(k|t'_{i'})\int_0^1 (y(\hat{p}_1)-\theta-\overline{b}_{i'})^2 f(\theta|k,n) = -\sum_{k\in\hat{p}_1} Pr(k|t'_{i'})\int_0^1 (y(\hat{p}_2)-\theta-\overline{b}_{i'})^2 f(\theta|k,n)$$

which can be rewritten as

$$-\sum_{k\in\hat{p}_1} Pr(k|t'_{i'})(y(\hat{p}_1)+y(\hat{p}_2)-2\mathbb{E}(\theta|k,n)-2\overline{b}_{i'}) = 0,$$

and therefore

$$\frac{1}{2}(y(\hat{p}_1)+y(\hat{p}_2)) - \sum_{k\in\hat{p}_1} Pr(k|t'_{i'})\mathbb{E}(\theta|k,n) = \overline{b}_{i'}$$

Further, define a new variable $\overline{b}$

$$\frac{1}{2}(y(\hat{p}_1)+y(\hat{p}_2))-\sum_{k\in\hat{p}_1}Pr^{DM}(k)\mathbb{E}(\theta|k,n):=\overline{b},\ Pr^{DM}(k\in\hat{p}_1)\in\Delta_{DM}(\hat{p}_1).$$

Given (2) and (3), $\overline{b}\geq\overline{b}_{i'}$. In the following I show that $\overline{b}^{1,0}>\overline{b}$. Then, it follows that $\overline{b}^{1,0}>\overline{b}_{i'}$. It is useful to focus on $\overline{b}$ since $\sum_{k\in\hat{p}_1}Pr^{DM}(k)\mathbb{E}(\theta|k,n)=y(\hat{p}_1)$, and therefore $\overline{b}^{1,0}>\overline{b}$ is satisfied if

$$y(p_2^0)-y(p_1^0)>y(\hat{p}_2)-y(\hat{p}_1)$$

is satisfied. Remember that $y(p_2^0)-y(p_1^0)=\frac{3+n}{2(n+2)}-\frac{1}{n+2}=\frac{n+1}{2(n+2)}$.

In the following I express $y(\hat{p}_1)$ and $y(\hat{p}_2)$ in a convenient form. Notice that $y(\hat{p}_1)$ can be written as

$$\frac{\frac{1}{n+2}\left(\frac{1}{2}(k-z)(k-z+1)+(k-z)\sum_{i=0}^{z}\omega_i+\sum_{k=0}^{z}\sum_{i=k}^{z}\omega_i\right)}{(k-z)+\sum_{i=0}^{z}\omega_i}$$

and $y(\hat{p}_2)$ can be written as

$$\frac{\frac{1}{n+2}\left((k-z)\sum_{i=0}^{z}(1-\omega_i)+\sum_{k=0}^{z}\sum_{i=k}^{z}(1-\omega_i)+\sum_{i=k+2}^{n+1}i\right)}{(1-\omega_0)+..+(1-\omega_z)+(n-k)}$$

Using the following two rearrangements:

$$\sum_{i=0}^{z}(1-\omega_i)=(z+1)-\sum_{i=0}^{z}\omega_i,$$

and

$$\sum_{k=0}^{z}\sum_{i=k}^{z}(1-\omega_i)=\sum_{i=1}^{z+1}i-\sum_{k=0}^{z}\sum_{i=k}^{z}\omega_i=\frac{1}{2}(z+1)(z+2)-\sum_{k=0}^{z}\sum_{i=k}^{z}\omega_i$$

we can express $y(\hat{p}_2)$ as

$$\frac{\frac{1}{n+2}\Big((k-z)[(z+1)-\sum_{i=0}^{z}\omega_i]+\frac{1}{2}(z+1)(z+2)-\sum_{k=0}^{z}\sum_{i=k}^{z}\omega_i+\frac{1}{2}(n-k)(k+n+3)\Big)}{(z+1)-\sum_{i=0}^{z}\omega_i+(n-k)}.$$

With those rearrangements, we have

$$y(p_2^0)-y(p_1^0)-(y(\hat{p}_2)-y(\hat{p}_1))=\frac{(n+1)(\sum_{i=0}^{z}\omega_i+(\sum_{i=0}^{z}\omega_i)^2-2\sum_{k=0}^{z}\sum_{i=k}^{z}\omega_i)}{2(n+2)(k+\sum_{i=0}^{z}\omega_i-z)(k-n-1+\sum_{i=0}^{z}\omega_i-z)}.$$

Notice that the denominator is strictly smaller than 0, because $(k+\sum_{i=0}^{z}\omega_i-z)>0$ as $k\geq z+1$ and $(k-n-1+\sum_{i=0}^{z}\omega_i-z)<0$ since $n>k$ and $\sum_{i=0}^{z}\omega_i<(z+1)$ as every for every $i\in\{0,..,z\}$, $\omega_i<0$. Thus, in order to show that $y(p_2^{1,0})-y(p_1^{1,0})-(y(\hat{p}_2)-y(\hat{p}_1))>0$, we need to show that

$$(\sum_{i=0}^{z}\omega_i+(\sum_{i=0}^{z}\omega_i)^2-2\sum_{k=0}^{z}\sum_{i=k}^{z}\omega_i)<0.$$

The above inequality can be rewritten as

$$(\sum_{i=0}^{z}\omega_i)^2+\sum_{i=0}^{z}\omega_i<2\sum_{k=0}^{z}\sum_{i=k}^{z}\omega_i. \tag{4}$$

First, we can rewrite $(\sum_{i=0}^{z}\omega_i)^2$ as

$$\omega_0(\omega_0+..+\omega_z)+\omega_1(\omega_1+..+\omega_z)+\omega_0\omega_1+\omega_2(\omega_2+..+\omega_z)+\omega_2(\omega_0+\omega_1)+..=$$

$$(\omega_0+..+\omega_z)\omega_0+(\omega_1+..+\omega_z)\omega_1+(\omega_1+..+\omega_z)\omega_0+(\omega_2+..+\omega_z)\omega_2+(\omega_2+..+\omega_z)\omega_1+..=$$

$$(\omega_0+..+\omega_z)\omega_0+(\omega_1+..+\omega_z)[\omega_1+\omega_0]+(\omega_2+..+\omega_z)[\omega_2+\omega_1]+...$$

Therefore, the LHS of (4) is

$$(\omega_0 + .. + \omega_z)(1 + \omega_0) + (\omega_1 + .. + \omega_z)[\omega_0 + \omega_1] + (\omega_2 + .. + \omega_z)[\omega_2 + \omega_1] + ...$$

The RHS of (4) is

$$2(\omega_0 + .. + \omega_z) + 2(\omega_1 + .. + \omega_z) + 2(\omega_2 + .. + \omega_z) + ...$$

Since for every $i \in \{0, .., z\}$, $\omega_i \leq 1$ and for some i $\omega_i < 1$, the last two expressions reveal that the inequality (4) is true. Thus, we have shown that for i', the largest upper bias is supported by the strategy $P_{i'}(Q) = \{\{0\}, \{1, .., n\}\}$. But then the same upper threshold is true for every other expert as well, since $P_{i'}(Q)$ implies that every expert conditions his upper deviation incentive on the same event that all experts' signals are 0.

There is no partition of $\{0, 1\}^{n'}$, for any $n' \leq n$, which can accommodate experts' biases larger than $\overline{b}^{1,0}$. Suppose not. Thus, there is an expert $j' \in N^e$ with $b_{j'} > \overline{b}^{1,0}$. But then there is no equilibrium with n experts where the decision maker receives information according to $\{\{0\}, \{1, .., n\}\}$ since expert j' has an incentive to deviate to a higher message. Without j', the partition which ensures the largest bias of the expert connected to the decision maker (where a single expert is connected to the decision maker which is a necessary condition for implementing a partition with largest possible upper biases for the experts) is $\{\{0\}, \{1, .., n-1\}\}$, and the upper bias is given by

$$\frac{1}{2}(\frac{1}{n+1} + \frac{2+n}{2(n+1)}) - \frac{1}{n+1} = \frac{n}{4(n+1)}$$

which is strictly smaller than $\overline{b}^{1,0} = \frac{n+1}{4(n+2)}$.

Next I calculate the lower bound for experts' biases that support P^0. The binding downward deviation by any expert $i \in N^e$ conditions on the event in which any one of experts' signals is 1, and all the other signals are 0. The corresponding incentive constraint is

$$-\int_0^1 (\frac{n+3}{2(n+2)} - \theta - b_i)^2 f(\theta|1, n) d\theta \geq -\int_0^1 (\frac{1}{n+2} - \theta - b_i)^2 f(\theta|1, n) d\theta,$$

which holds for:

$$b_i \geq \frac{n-3}{4(n+2)}.$$

Although I do not know the implementable two-pool decision maker's partition of $\{0,1\}^n$ with the second highest shift in decision maker's policy (implying the second highest upper bias threshold), we know that the difference between the two pools cannot exceed $(\frac{n+3}{2(n+2)} - \frac{1}{n+2})$. Moreover, the first pool of such a partition includes at least one element of $\{0,1\}^n$ with the sum of the signals 1. Therefore, for any decision maker's two-pool partition which is not P^0, the lower bound on expert's beliefs that determines his upward deviation constraint cannot be smaller than assigning the posteriors $\frac{n}{n+1}$ to $k=0$ and $\frac{1}{n+1}$ to $k=1$. Therefore, we know that the second highest bias cannot exceed $\hat{b}$ which solves the following problem:

$$-\frac{n}{n+1}\int_0^1 (\frac{1}{n+2} - \theta - \hat{b})^2 f(\theta|0,n)d\theta - \frac{1}{n+1}\int_0^1 (\frac{1}{n+2} - \theta - \hat{b})^2 f(\theta|1,n)d\theta =$$

$$-\frac{n}{n+1}\int_0^1 (\frac{n+3}{2(n+2)} - \theta - \hat{b})^2 f(\theta|0,n)d\theta - \frac{1}{n+1}\int_0^1 (\frac{n+3}{2(n+2)} - \theta - \hat{b})^2 f(\theta|1,n)d\theta.$$

The equation holds for $\hat{b} = \frac{(n+3)(n-1)}{4(n+1)(n+2)}$. Notice that $\hat{b} < \bar{b}^{1,0}$ which is intuitive since in the case of $\hat{b}$ a positive weight is assigned to the sum of the signals 1.

We, further, know that for any $n' < n$, the largest upper bias is weakly below $\frac{n}{4(n+1)}$. Since

$$\frac{(n+3)(n-1)}{4(n+1)(n+2)} < \frac{n}{4(n+1)}, \text{ for } n \geq 2,$$

we can state that, conditional on decision maker's two-pool partitions of $\{0,1\}^n$ implementable in any $Q \in \mathbb{Q}$, and given that the biases of all n experts are within the interval $(\frac{n}{4(1+n)}, \frac{n+1}{4(n+2)}]$, only a network Q which satisfies the requirements stated in the Proposition can lead to information transmission from the experts to the decision maker. The corresponding equilibrium is characterized by a partition $P^0_{DM}(Q)$ according to which the decision maker receives her information.

Finally, I show that it is without loss of generality to focus on two-pool partitions.

Claim 3: A partiton of $\{0,1\}^n$ according to which the decision maker receives her information in a network $Q \in \mathbb{Q}$ and which supports the largest possible upper bias threshold for each expert, has to consist of two pools.

Proof of Claim 3: Suppose not. First, take any equilibrium in $Q \in \mathbb{Q}$ with $|N_{DM}(Q)| = 1$ and $P_{DM}(Q) = \{p_1, .., p_l\}$ such that $l \geq 3$.

1. Take any $i \in N^e$. There is a type of expert i, denoted by t_i', which determines the binding upward incentive constraint of expert i and therefore the upper bias threshold b_i' which supports $P_i(Q)$ is therefore implicitly determined by:

$$-\sum_{p \in P_{DM}(Q)} \Pr(p|p_i, P_{-i}(Q)) \sum_{k \in p} Pr(k|t_i') \int_0^1 (y(p) - \theta - b_i')^2 f(\theta|k, n) d\theta -$$

$$-\sum_{p \in P_{DM}(Q)} \Pr(p|p_i', P_{-i}(Q)) \sum_{k \in p} Pr(k|t_i') \int_0^1 (y(p) - \theta - b_i')^2 f(\theta|k, n) d\theta \quad (5)$$

for $t_i' \in p_i$ and $p_i \neq p_i'$ where $p_i' \in P_i(Q)$ is the next highest message to p_i.

2. Next, choose $p_{j'} \in P_{DM}(Q)$ such that

$$p_{j'} = \max_{j \in \{1,..,l-1\}} \{p_{j+1} - p_j\}.$$

Define the bias $\hat{b}'$ which implicitly solves

$$-\int_0^1 (y(p_j) - \theta - \hat{b}')^2 f(\theta|k \in p_j, n) d\theta = -\int_0^1 (y(p_{j+1}) - \theta - \hat{b}')^2 f(\theta|k \in p_j, n) d\theta,$$

where $f(\theta|k \in p_j, n)$ is the posterior according to which the decision maker forms her beliefs about the signals when her state-relevant information is summarized by p_j. It follows that $\hat{b}' > b_i'$.

3. Next, construct another strategy profile which is the same as before for every expert apart from an expert $i' \in N^e$ who is directly connected to the decision maker. The latter communicates to the decision maker according to the partition $P_{i'}(Q) = \{p_1', p_2'\}$ where the first pool is $p_1' := \cup_{i \in \{1,..,j\}} p_i$, and the second pool is $p_2' := \cup_{i \in \{j+1,..,l\}} p_i$.

Notice that $y(p_2') - y(p_1')$ is larger than $y(p_{j+1}) - y(p_j)$. Therefore, the bias $\hat{b}''$ which solves

$$-\int_0^1 (y(p'_1) - \theta - \hat{b}'')^2 f(\theta|k \in p'_1, n)d\theta = -\int_0^1 (y(p'_2) - \theta - \hat{b}'')^2 f(\theta|k \in p'_1, n)d\theta$$

is such that $\hat{b}'' > \hat{b}'$.

4. We know from the previous considerations in the proof of Proposition 3 that if in Q the decision maker would receive information according to $P_{DM}(Q) = \{\{0\}, \{1, .., n\}\}$ then

$$y(p'_2) - y(p'_1) < y(1, .., n) - y(0).$$

Therefore, for $p_1 = \{0\},\ p_2 = \{1, .., n\}$ the bias $\tilde{b}$ which solves

$$-\int_0^1 (y(p_1) - \theta - \tilde{b})^2 f(\theta|k = 0, n)d\theta = -\int_0^1 (y(p_2) - \theta - \tilde{b})^2 f(\theta|k = 0, n)d\theta$$

is such that $\tilde{b} > \hat{b''}$. But since $\tilde{b}$ determines each experts' upper bias in case the decision maker receives her information according to the partition $\{\{0\}, \{1, .., n\}\}$, we conclude that it cannot be that decision maker's partition which consists of 3 pools generates an equilibrium that supports experts' biases larger than $\frac{(n+3)(n-1)}{4(n+1)(n+2)}$.

Now, suppose that $|N_{DM}(Q)| > 1$, which means that the decision maker is connected to multiple experts. Take any expert $i \in N_{DM}(Q)$ and delete the links from all experts $j \in N_{DM}(Q)$, $j \neq i$. Next create an additional link for every such expert going from j to i. It means for every $j \in N_{DM}(Q)$, $j \neq i$, there is $e_{ji} = 1$. Denote the new network by Q'. Now use the same argument as before to show that in equilibrium with $P_{DM}(Q') = \{\{0\}, \{1, .., n\}\}$ the upper bias threshold for every expert is larger than in the former equilibrium represented by $P_{DM}(Q) = \{p_1, .., p_l\}$. *Q.E.D.*

This shows that the Proposition is true.
Q.E.D.

Example on p. 17:

Equilibrium in network in Figure 5a:

The strategy profile specified in the example implies that the message strategies of experts can be represented by the following partitions: $P_1 = \{\{0\},\{1\}\}$, $P_2 = \{\{0\},\{1,2\}\}$ and $P_3 = \{\{0\},\{1\}\}$. Denote by $p_j \in P_j$ a typical element of a partition $P_j,\ j \in \{1,2,3\}$.

The choices of the decision maker depend on communicated pools of the partitions $P_1,\ P_2$ as follows:

$$y(p_1 = 0, p_2 = 0) = \frac{1}{5},\ y(p_1 = 1, p_2 = 0) = \frac{2}{5},\ y(p_1 = 0, p_2 = 1,2) = \frac{7}{15},$$
$$y(p_1 = 1, p_2 = 1,2) = \frac{18}{25}.$$

We start with the incentive constraints for the types of expert 2: $t_2 = 0$ assigns the posteriors $\frac{3}{4}$ to $t_1 = 0$ and $\frac{1}{4}$ to $t_1 = 1$. Thus, the incentive constraint for $t_2 = 0$ is:

$$-\frac{3}{4}\int_0^1 (\frac{1}{5} - \theta - b_2)^2 f(\theta|0,3)d\theta - \frac{1}{4}\int_0^1 (\frac{2}{5} - \theta - b_2)^2 f(\theta|1,3)d\theta \geq$$
$$-\frac{3}{4}\int_0^1 (\frac{7}{15} - \theta - b_2)^2 f(\theta|0,3)d\theta - \frac{1}{4}\int_0^1 (\frac{18}{25} - \theta - b_2)^2 f(\theta|1,3)d\theta,$$

so that $b_2 \leq \frac{74}{525} \approx 0.14$.

$t_2 = 1$ assigns the posteriors $\frac{1}{2}$ to $t_1 = 0$ and $\frac{1}{2}$ to $t_1 = 1$. The incentive constraint of $t_2 = 1$ is:

$$-\frac{1}{2}\int_0^1 (\frac{7}{15} - \theta - b_2)^2 f(\theta|1,3)d\theta - \frac{1}{2}\int_0^1 (\frac{18}{25} - \theta - b_2)^2 f(\theta|2,3)d\theta \geq$$
$$-\frac{1}{2}\int_0^1 (\frac{1}{5} - \theta - b_2)^2 f(\theta|1,3)d\theta - \frac{1}{2}\int_0^1 (\frac{2}{5} - \theta - b_2)^2 f(\theta|2,3)d\theta,$$

so that $b_2 \geq -\frac{43}{825} \approx -0.05$ which is satisfied since we assumed that all experts are positively biased. Therefore, the binding constraint implies $b_2 \leq \frac{74}{525}$.

Since expert 3 conditions pivotality on the same information sets as expert 2, the incentive constraints for expert 3 imply as well $b_3 \leq \frac{74}{525}$.

Next, $t_1 = 0$ assigns the posteriors $\frac{1}{2}$ to $t_2 = 0$, $\frac{1}{3}$ to $t_2 = 1$ and $\frac{1}{6}$ to $t_2 = 2$. The corresponding incentive constraint for $t_1 = 0$ is:

$$-\frac{1}{2}\int_0^1(\frac{1}{5}-\theta-b_1)^2 f(\theta|0,3)d\theta-\frac{1}{3}\int_0^1(\frac{7}{15}-\theta-b_1)^2 f(\theta|1,3)d\theta-$$
$$\frac{1}{6}\int_0^1(\frac{7}{15}-\theta-b_1)^2 f(\theta|2,3)d\theta\geq$$
$$-\frac{1}{2}\int_0^1(\frac{2}{5}-\theta-b_1)^2 f(\theta|0,3)d\theta-\frac{1}{3}\int_0^1(\frac{18}{25}-\theta-b_1)^2 f(\theta|1,3)d\theta-$$
$$\frac{1}{6}\int_0^1(\frac{18}{25}-\theta-b_1)^2 f(\theta|2,3)d\theta,$$

resulting in $b_1 \leq \frac{293}{2550} \approx 0.115$.

Finally, $t_1 = 1$ assigns the posteriors $\frac{1}{6}$ to $t_2 = 0$, $\frac{1}{3}$ to $t_2 = 1$ and $\frac{1}{2}$ to $t_2 = 2$. The corresponding incentive constraint for t_1, assuming that the other players stick to the specified strategy profile, is:

$$-\frac{1}{6}\int_0^1(\frac{2}{5}-\theta-b_1)^2 f(\theta|1,3)d\theta-\frac{1}{3}\int_0^1(\frac{18}{25}-\theta-b_1)^2 f(\theta|2,3)d\theta-$$
$$\frac{1}{2}\int_0^1(\frac{18}{25}-\theta-b_1)^2 f(\theta|3,3)d\theta\geq$$
$$-\frac{1}{6}\int_0^1(\frac{1}{5}-\theta-b_1)^2 f(\theta|1,3)d\theta-\frac{1}{3}\int_0^1(\frac{7}{15}-\theta-b_1)^2 f(\theta|2,3)d\theta-$$
$$\frac{1}{2}\int_0^1(\frac{7}{15}-\theta-b_1)^2 f(\theta|3,3)d\theta,$$

resulting in $b_1 \geq -\frac{203}{1650} \approx -0.12$ which is satisfied per assumption.

To summarize, the above strategy profile constitutes an equilibrium for $b_2, b_3 \leq 0.14$, and for $b_1 \leq 0.115$. The expected utility of the DM is $= -0.0396$.

Implementation of the same equilibrium outcome in the line (Figure 5b):

Consider the following strategy profile. Expert 3 communicates his signals to expert 2. If expert 2's summary statistic is 0, he sends p_2 to expert 1; otherwise he sends p_2'. Expert 1 sends one out of the four messages to the decision maker: p_1 if $(s_1 = 0, p_2)$, p_1' if $(s_1 = 0, p_2')$, p_1'' if $(s_1 = 1, p_2)$ and p_1''' if $(s_1 = 1, p_2')$. The decision maker chooses:

$$y(p_1)=\frac{1}{5},\ y(p_1')=\frac{7}{15},\ y(p_1'')=\frac{2}{5}\text{ and }y(p_1''')=\frac{18}{25}.$$

The incentive constraints for experts 3 and 2 are the same as in network in Figure 5a because both experts remain pivotal for the same information sets of the decision maker. However, the bias constraints for expert 1 get tighter. In particular, his tightest constraint for the upward deviation is defined for $t_1 = (s_1 = 1, p_2)$, which prevents deviation to the message $(s_1 = 0, p_2')$:

$$-\int_0^1 (\frac{2}{5} - \theta - b_1)^2 f(\theta|1,3)d\theta \geq -\int_0^1 (\frac{7}{15} - \theta - b_1)^2 f(\theta|1,3)d\theta,$$

or:

$$\frac{2}{5} + \frac{7}{15} - 2\frac{2}{5} - 2b_1 \geq 0,$$

which results in $b_1 \leq 0.033$.

The tightest constraint for the downward deviation of expert 1 prevents deviation from $t_1 = (s_1 = 0, p_2')$ to $(s_1 = 1, p_2)$:

$$-\frac{2}{3}\int_0^1 (\frac{7}{15} - \theta - b_1)^2 f(\theta|1,3)d\theta - \frac{1}{3}\int_0^1 (\frac{7}{15} - \theta - b_1)^2 f(\theta|2,3)d\theta \geq$$
$$-\frac{2}{3}\int_0^1 (\frac{2}{5} - \theta - b_1)^2 f(\theta|1,3)d\theta - \frac{1}{3}\int_0^1 (\frac{2}{5} - \theta - b_1)^2 f(\theta|2,3)d\theta,$$

which can be written as:

$$\frac{2}{3}\Big(\frac{7}{15} + \frac{2}{5} - 2\frac{2}{5} - 2b_1\Big) + \frac{1}{3}\Big(\frac{7}{15} + \frac{2}{5} - 2\frac{2}{5} - 2b_1\Big) \geq 0,$$

or:

$b_1 \geq -0.033$. Therefore, the binding constraint for expert 1 shows that he is incentivized to communicate according to the specified strategy profile in a line for a strictly smaller range of biases, compared to the network in Figure 5a.

Proof of Lemma 1:

Let $P_{-(i,j)}(Q)$ denote the strategy profile of all experts apart from the experts i and j, which is defined as $P_{-(i,j)} = \Pi_{l \in N^e} P_l(Q)$, $l \neq i, l \neq j$.

Since we assume that in Q expert i is informed about expert j's type, we can express the type space of i as $T_i(Q) = \tilde{P}_i(Q) \times P_j(Q)$, where $\tilde{P}_i(Q)$ captures the private signal of i and potentially some further information depending on the connections in Q (the exact information which i receives beyond his own private signal and p_j is inessential). The incentive constraints of $t_j \in T_j(Q)$ can be expressed in terms of his expectation over the types of expert i, and therefore in terms of i's strategy given $t_i \in p_i$ and $t_j \in p_j$:

$$- \sum_{\tilde{p}_i \in \tilde{P}_i(Q)} Pr(\tilde{p}_i|t_j) \sum_{p \in P_{DM}(Q)} \Pr(p|p_i(\tilde{p}_i, p_j), P_{-(i,j)}(Q)) \sum_{k \in p} Pr(k|t_i) \int_0^1 (y(p) - \theta - b_j)^2 f(\theta|k, n) d\theta \geq$$

$$- \sum_{\tilde{p}_i \in \tilde{P}_i(Q)} Pr(\tilde{p}_i|t_j) \sum_{p \in P_{DM}(Q)} \Pr(p|p_i(\tilde{p}_i, p_j'), P_{-(i,j)}(Q)) \sum_{k \in p} Pr(k|t_i) \int_0^1 (y(p) - \theta - b_j)^2 f(\theta|k, n) d\theta,$$

where $p_j' \neq p_j$, and p_j' is the next highest message to p_j such that by deviation j expects a change in decision maker's policy (otherwise the deviaion leaves j's payoffs unaffected). The upper bias threshold of j which supports $P_j(Q)$ is denoted by b_j' and solves the following problem.

$$b_j' = \operatorname*{argmin}_{b_j \in \mathbb{R}} \Big\{ - \sum_{\tilde{p}_i \in \tilde{P}_i(Q)} Pr(\tilde{p}_i|t_j) \sum_{p \in P_{DM}(Q)} \Pr(p|p_i(\tilde{p}_i, p_j), P_{-(i,j)}(Q)) \sum_{k \in p} Pr(k|t_i)$$
$$\int_0^1 (y(p) - \theta - b_j)^2 f(\theta|k, n) d\theta =$$
$$- \sum_{\tilde{p}_i \in \tilde{P}_i(Q)} Pr(\tilde{p}_i|t_j) \sum_{p \in P_{DM}(Q)} \Pr(p|p_i(\tilde{p}_i, p_j'), P_{-(i,j)}(Q)) \sum_{k \in p} Pr(k|t_i)$$
$$\int_0^1 (y(p) - \theta - b_j)^2 f(\theta|k, n) d\theta \ |\text{for all } t_j \in T_j(Q), t_j \in p_j, t_i \in T_i(Q), t_i \in p_i \Big\}. \quad (6)$$

The upper bias threshold of i which supports $P_i(Q)$ is denoted by b_i' and solves the following problem for a given $\overline{p}_j$:

$$b_i' = \operatorname*{argmin}_{b_i \in \mathbb{R}} \Big\{ \sum_{p \in P_{DM}(Q)} \Pr(p|p_i(\tilde{p}_i, \overline{p}_j), P_{-(i,j)}(Q)) \sum_{k \in p} Pr(k|t_i)$$
$$\int_0^1 (y(p) - \theta - b_i)^2 f(\theta|k, n) d\theta =$$
$$\sum_{p \in P_{DM}(Q)} \Pr(p|p_i', P_{-(i,j)}(Q)) \sum_{k \in p} Pr(k|t_i)$$
$$\int_0^1 (y(p) - \theta - b_i)^2 f(\theta|k, n) d\theta \mid t_i \in T_i(Q), t_i \in p_i \Big\}, \quad (7)$$

where $p'_i \in P_i(Q)$ is the next highest message to $p_i(\tilde{p}_i)$ such that by deviation i expects a change in decision maker's policy (otherwise the deviaion leaves j's payoffs unaffected). Before comparing b'_i and b'_j notice that

1. In equilibria with strategic coarsening of information by at least one of the experts, and an equilibrium partition $P_{DM}(Q) = \{p_1, .., p_l\}$ with $v \geq 3$, the difference $(y(p_{v+1}) - y(p_v))$, $v \in \{1, .., l-1\}$, is in general different for every distinct v.

2. Expert j is uncertain about the exact realization of $\tilde{p}_i$. Different realizations of $\tilde{p}_i$ can result in different $y(p_v)$ on path.

As a result, it has to be true that $b'_i \leq b'_j$ since i conditions his deviation on a particular tuple $(\tilde{p}_i, p_j)$ whereas expert j only observers $t_j = p_j$. By a symmetric argument, $b''_i \geq b''_j$, where b''_i (b''_j) is the lower bias threshold for expert i (j) which makes the strategy $P_i(Q)$ $(P_j(Q))$ incentive compatible.

Q.E.D.

Proof of Lemma 2:

Fix an equilibrium strategy profile in a network $Q \in \mathbb{Q}$, $P(Q)$, which involves strategic coarsening of information such that there is an expert $i \in N^e$ who receives full information from some other expert $j \in H_{ji}(Q)$. If $e_{ji} = 1$, then the Lemma is satisfied.

If $|H_{ji}(Q)| > 2$, there is at least one other expert on the path $H_{ji}(Q)$, strictly between j and i. Denote the set of those experts by $\hat{N}$. The communication strategy of every $j' \in \hat{N}$ can be written as $P_{j'}(Q) = P_j(Q) \times \{P_{j'}(Q) \setminus P_j(Q)\}$. This is because by assumption j' truthfully communicates the message sent by expert j, p_j.

Fix any $j' \in \hat{N}$ and suppose that $P_{j'}(Q)$ is supported by the biases within an interval $[\underline{b}_{j'}, \overline{b}_{j'}] \subset \mathbb{R}$.

Next, take a link in Q going out from j, delete it, and create a new directed link from j to i, $e_{ji} = 1$. Denote the new network by Q' (clearly $Q = Q'$ if j is directly connected to i in Q), and assume the following strategy profile: for every $i' \in \{N^e \setminus \hat{N}\}$, $P_{i'}(Q') = P_{i'}(Q)$, and expert j' uses the strategy $P_{j'}(Q') = P_{j'}(Q) \setminus P_j(Q)$. In words, experts on the path $H_{ji}(Q)$ which were strictly between i and j in Q, communicate the same information in the new network Q' compared to Q, only without the message of expert j. Moreover, all the other experts have the same communication strategies as before. Clearly, the incentive constraints of the experts outside of $\hat{N}$ remain the same.

However, the experts in $\hat{N}$ face more relaxed incentive constraints in Q' compared to Q, since in the former case they do not observe the message of expert j. To see why, notice that the upper bias of j' in Q', denoted by $b'_{j'}$, is determined by the following minimization program. For $t_{j'} \in p_{j'}$ and $p_{j'}, p'_{j'} \in P_{j'}(Q')$, $p_{j'} \neq p'_{j'}$, where $p'_{j'}$ is the next highest message to $p_{j'}$ (such that by deviation j' expects a change in decision maker's policy), the bias threshold $b'_{j'}$ is such that

$$b'_{j'} = \operatorname*{argmin}_{b_i \in \mathbb{R}} \Big\{ \sum_{p \in P_{DM}(Q')} \Pr(p|p_{j'}(t_{j'}), P_j(Q'), P_{-(i,j)}(Q')) \sum_{k \in p} Pr(k|t_i)$$
$$\int_0^1 (y(p) - \theta - b_{j'})^2 f(\theta|k,n) d\theta =$$
$$\sum_{p \in P_{DM}(Q')} \Pr(p|p'_i, P_{-(i,j)}(Q')) \sum_{k \in p} Pr(k|t_i)$$
$$\int_0^1 (y(p) - \theta - b'_{j'})^2 f(\theta|k,n) d\theta \;| \text{for all } t_{i'} \in T_{j'}(Q'), t_{j'} \in p_{j'} \Big\}. \tag{8}$$

For $p_j \in P_j(Q)$ and $\hat{p}_{j'} \in P_{j'}(Q)$ the upper bias threshold supporting the strategy $P_{j'}(Q)$, denoted by $\overline{b}_{j'}$, is such that

$$\overline{b}_{j'} = \operatorname*{argmin}_{b_i \in \mathbb{R}} \Big\{ \sum_{p \in P_{DM}(Q')} \Pr(p|\hat{p}_{j'}(t_{j'}, p_j), P_{-(i,j)}(Q')) \sum_{k \in p} Pr(k|t_i)$$
$$\int_0^1 (y(p) - \theta - b_{j'})^2 f(\theta|k,n) d\theta =$$
$$\sum_{p \in P_{DM}(Q')} \Pr(p|\hat{p}'_{j'}, P_{-(i,j)}(Q')) \sum_{k \in p} Pr(k|t_i)$$
$$\int_0^1 (y(p) - \theta - b_{j'})^2 f(\theta|k,n) d\theta \;| \text{for all } t_{i'} \in T_{j'}(Q'), t_{j'} \in p_{j'} \Big\}, \tag{9}$$

where $\hat{p}'_{j'} \in P_{j'}(Q)$ is the next highest message to $\hat{p}_{j'}$ such that by deviation j' expects a change in decision maker's policy (otherwise the deviaion leaves the payoff of j' unaffected).

Notice that in equilibria with strategic coarsening of information by at least one of the experts, and an equilibrium partition $P_{DM}(Q) = \{p_1, .., p_l\}$ with $v \geq 3$, the difference $(y(p_{v+1}) - y(p_v))$, $v \in \{1, .., l-1\}$, is in general different for every distinct

v. Since an Q expert j' is uncertain about the message of j, and in Q' expert j' observes the exact message of j, comparing (8) and (9) we see that $\overline{b}_{j'} \leq b'_{j'}$.

Q.E.D.

Proof of Lemma 3:

Suppose that an optimal network Q features $n' \neq n$ non-babbling experts and a group G with $|G| > 3$ which includes $j, j' \in N'$ who have no incoming links and have outgoing links to the group leader, i^G. Since j and j' do not babble, $P_j(Q) = P_{j'}(Q) = \{\{0\},\{1\}\}$. Take any message profile of the experts within G other than j, j' and denote it by $p \in \Pi_{i \in \{N_i(Q) \setminus \{j,j'\}\}} P_i(Q)$. For a given private signal of the group leader, $s_{i^G} \in \{0,1\}$, and any $k \in \{0,1\}$ and $k' \in \{0,1\}$, then it must be true for a type of i^G, $t_{i^G} \in T_{i^G}(Q)$:

$$t_{i^G}(s_{i^G}, p_j = k, p_{j'} = k', p) = t_{i^G}(s_{i^G}, p_j = k', p_{j'} = k, p),$$

Therefore, the communication strategy of i^G must be such that

$$p_{i^G}(s_{i^G}, p_j = k, p_{j'} = k', p) = p_{i^G}(s_{i^G}, p_j = k', p_{j'} = k, p),\ p_{i^G} \in P_{i^G}(Q)$$

since otherwise two types of i^G with the same beliefs over $\{0,1\}^{n'}$ would take two different equilibrium actions which is not incentive compatible.

Since $T_j(Q) = T_{j'}(Q) = \{\{0\},\{1\}\}$, j and j' have the same ex ante beliefs over the entire signal space $\{0,1\}^n$. As their messages are treated symmetrically in the communication strategy of i^G, their communication strategies are supported by the same range of biases.

Finally, since in equilibrium i^G receives truthful information about the exact types of j and j', conditions of Lemma 1 apply. It means that the range of biases supporting $P_{i^G}(Q)$ is weakly included into the range of biases supporting $P_j(Q)$ or $P_{j'}(Q)$.

Q.E.D.

Proof of Proposition 4:

Fix any network $Q \in \mathbb{Q}$ and any equilibrium decision maker's partition of $\{0,1\}^n$, $P_{DM}(Q)$. Notice that per definition the decision maker can commit to implement any partition P_{DM} of $\{0,1\}^n$. Each pool of the partition consists of a subset of $\{0,1\}^n$, such that for every element of $\{0,1\}^n$ within any given pool, each of the digits can be mapped back to each experts' private signals. Thus, if every expert truthfully reveals his private signals, this generates partition P_{DM}.

Notice that in a direct mechanism each expert only observes his own private signal whereas in a network Q which is not a star, there is at least one expert which observes at least one further signal additional to his own private signal. According to Lemma 2, if the experts' strategy profile generating the partition $P_{DM}(Q)$ involves strategic coarsening of information by at least one expert, the bias range supporting the equilibrium strategy of each expert in a network Q is weakly included into the bias range supporting the expert's equilibrium strategy in a network Q. Therefore, the outcome of a network Q is implementable in a direct mechanism but the converse is not true.

Q.E.D.

7 References

Ahuja, Gautam (2000): "Collaboration networks, structural holes, and innovation: A longitudinal study." Administrative Science Quarterly 45(3), 425-455.

Ahuja, Manju and Kathleen Carley (1998): "Network Structure in Virtual Organizations." Journal of Computer-Mediated Communication, Vol 3. Issue 4.

Allen, Tomas J. (1977): "Managing the Flow of Technology." MIT Press, Cambridge, MA.

Alonso, Ricardo, Wouter Dessein, and Niko Matouschek (2015): "Organizing to Adapt and Compete." American Economic Journal: Microeconomics, 7(2), 158-87.

Alonso, Ricardo, Wouter Dessein, and Niko Matouschek (2008): "When Does Coordination Require Centralization?" American Economic Review, 98(1), 145-79.

Ambrus, Atilla and Shih En Luz (2014): "Almost Fully Revealing Cheap Talk with Imperfectly Informed Senders." Working paper.

Ambrus, Atilla, Eduardo Azevedo and Yuichiro Kamada (2013): "Hierarchical cheap talk." Theoretical Economics, 8, 223-261.

Ambrus, Attila, and Satoru Takahashi (2008): "Multi-sender cheap talk with restricted state spaces." Theoretical Economics, 3(1), 1-27.

Argenziano, Rossella, Sergei Severinov, and Francesco Squintani (2015): "Strategic Information Acquisition and Transmission." American Economic Journal: Microeconomics, forthcoming.

Austen-Smith, David (1993): "Interested Experts and Policy Advice: Multiple Referrals under Open Rule." Games and Economic Behavior, V, 1-43.

Aumann, Robert J, and Sergiu Hart (2003): "Long Cheap Talk." Econometrica, 71(6), 1619-60.

Battaglini, Marco (2004): "Policy Advice with Imperfectly Informed Experts." Advances in Theoretical Economics: Vol. 4: Article. 1.

Battaglini, Marco (2002): "Multiple Referrals and Multidimensional Cheap Talk." Econometrica, 70, 1379 - 1401.

Bolton, Patrick and Dewatripont, Mathias (1994): "The Firm as a Communication Network." The Quarterly Journal of Economics, MIT Press, vol. 109(4), 809-39.

Calvó-Armengol, Antoni and Joan de Mart (2007): "Communication Networks: Knowledge and Decisions." American Economic Review, Papers and Proceedings, 97(2), 86-91.

Calvó-Armengol, Antoni and Joan de Mart (2009): "Information Gathering in Organizations: Equilibrium, Welfare and Optimal Network Structure." Journal of the European Economic Association, 7(1), 116-161.

Crawford, Vincent and Joel Sobel (1982): 'Strategic Information Transmission." Econometrica, 50, 1431-1452.

Crozier, Michel (1963): "Le Phenomene Bureaucratique". Paris: Editions du Seuil. Translated by Crozier as "The Bureaucratic Phenomenon." Chicago: University of Chicago Press.

Cyert, Richard and James March (1963): "A Behavioral Theory of the Firm." Englewood Cliffs, NJ: Prentice-Hall.

Dalton, Melville (1959): "Men Who Manage." New York: Wiley and Sons.

Dewan, Torun, Galeotti, Andrea, Ghiglino, Christian and Francesco Squintani (2014): "Information Aggregation and Optimal Selection of the Executive". American Journal of Political Science, 59(2): 475-494.

Dessein, Wouter (2002): "Authority and Communication in Organizations." Review of Economic Studies 69, 811-838.

Friebel, Guido and Michael Raith (2004): "Abuse of Authority and Hierarchical Communication." RAND Journal of Economics, The RAND Corporation, vol. 35(2), 224-244.

Galeotti, Andrea, Christian Ghiglino and Francesco Squintani (2013): "Strategic Information Transmission Networks." Journal of Economic Theory, 148(5), 1751-1769.

Garicano, Luis (2000): "Hierarchies and the Organization of Knowledge in Production." Journal of Political Economy, 108:874-904.

Gibbons, Robert and John Roberts (2013): "Economic Theories of Incentives in Organizations." In Handbook of Organizational Economics, eds Robert Gibbons and John Roberts, Princeton University Press.

Glaser, Lotte, Sebastian P.L. Fourne and Tom Elfring (2015): "Achieving strategic renewal: the multi-level influences of top and middle managers boundary-spanning." Small Business Economics, 45, 305-327.

Grenadier, Steven R, Andrey Malenko, and Nadya Malenko (2015): "Timing Decisions in Organizations: Communication and Authority in a Dynamic Environment." Stanford University Graduate School of Business Research Paper No. 15-1.

Hagenbach, Jeanne and Frédéric Koessler (2011): "Strategic communication networks." Review of Economic Studies, 77(3), 1072-1099.

Ivanov, Maxim (2010): "Information control and organizational design." Journal of

Economic Theory, 145, 869-884.

Ishida, Junichiro and Takashi Shimizu (2012): "Can More Information Facilitate Communication?" ISER Discussion Paper 0839, Institute of Social and Economic Research, Osaka University.

Jackson, Matthew (2008): "Social and Economic Networks." Princeton University Press, Princeton.

Krishna, Vijay and John Morgan (2001)a: "A model of expertise." Quarterly Journal of Economics 116, 74775.

Krishna, Vijay and John Morgan (2001)b: "Asymmetric information and legislative rules: some amendments." American Political Science Review 95, 43552.

Marschak, Jacob and Roy Radner (1972): "Economic Theory of Teams." Yale University Press, New Haven, Connecticut.

Milgrom, Paul and John Roberts (1988): "An Economic Approach to Influence Activities in Organizations." American Journal of Sociology 94, 154-179.

Mookherjee, Dilip (2006): "Decentralization, Hierarchies, and Incentives: A Mechanism Design Perspective." Journal of Economic Literature, 44(2), 367-390.

Morgan, John and Philip Stocken (2008): "Information Aggregation in Polls." American Economic Review, 2008, Vol. 98, No. 3, 864-896.

Mookherjee, Dilip (2013): "Incentives in Hierarchies." In Handbook of Organizational Economics, eds. Robert Gibbons and John Roberts, Princeton University Press.

Mylovanov, Timofyj and Andriy Zapechelnyuk (2013): "Decision rules revealing commonly known events." Economics Letters, 119, 8-10.

Oberg, Achim and Peter Walgenbach (2008): "Hierarchical structures of communication in a network organization." Scandinavian Journal of Management, 24, 183-198.

Radner, Roy (1993), "The organization of decentralized information processing."

Econometrica, 61, 1109-1146.

Radner, Roy (1962): "Team Decision Problems." Annals of Mathematical Statistics 33, 857-881.

Radner, Roy and Timothy van Zandt (1992): "Information Processing in Firms and Returns to Scale." Annales d'Economie et de Statistique, No. 25-26, 265-298.

Rantakari, Heikki (2008): "Governing Adaptation." Review of Economic Studies 75(4), 1257-1285.

Sah, Raaj Kumar and Joseph Stiglitz (1986): "The Architecture of Economic Systems: Hierarchies and Polyarchies." American Economic Review 76(4), 716727.

Schilling, Melissa and Christian Fang (2013): "When Hubs forget, lie, and play favorites: Interpersonal Network Structure, Information Distortion, and Organizational Learning." Strategic Management Journal, 35(7), 974-994.

Tushman, Michael and Thomas Scanlan (1981): "Boundary spanning individuals: Their role in information transfer and their antecedents." Acad. Management J. 24(2), 289305.

Van Zandt, Timothy (1999a): "Decentralized Information Processing in the Theory of Organizations." Contemporary Economic Issues, Vol. 4: Economic Design and Behavior (Murat Sertel, ed.), 125160, MacMillan Press Ltd, London.

Van Zandt, Timothy (1999b): "Real-time decentralized information processing as a model of organizations with boundedly rational agents." Review of Economic Studies, 66, 633-658.

Wolinsky, Asher (2002): "Eliciting information from multiple experts." Games and Economic Behavior, 41, 141-160.