WARWICK

ECONOMICS

CRETA

Centre for Research in Economic Theory and its Applications

Discussion Paper Series

Designing Communication Hierarchies

Dimitri Migrow

October 2018 No: 44

CRETA
Centre for Research in Economic Theory and its Applications
Department of Economics
University of Warwick, Coventry,
CV4 7AL, United Kingdom

warwick.ac.uk/fac/soc/economics/research/centres/creta/papers

Designing Communication Hierarchies*

Dimitri Migrow†

Abstract

A manager aims to elicit employees' information by designing a hierarchical communication network. She decides who communicates with whom, and in which order, where communication takes the form of "cheap talk" (Crawford and Sobel, 1982) and the information structure is beta-binomial. 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 on biases and a uniform prior, the optimal network features a single intermediary who communicates directly to the manager.

Keywords: organizational design, strategic communication, information aggregation

JEL Classification: C72, D82, D83

*I thank Ricardo Alonso, Dan Bernhardt, Ken Binmore, Vincent Crawford, Bhaskar Dutta, Alexander Jakobsen, Marina Halac, Andrew Harkins, Sergiu Hart, Navin Kartik, Jacob Leshno, Niko Matouschek, Meg Meyer, Rob Oxoby, Motty Perry, Herakles Polemarchakis, Heikki Rantakari, Debraj Ray, Phil Reny, Ariel Rubinstein, Mehdi Shadmehr, Alastair F. Smith, Francesco Squintani and Julia Wirtz. I thank the seminar participants at Autonoma (Barcelona), Calgary, HEC Paris, Kellogg School of Business (Northwestern), Maastricht, Manchester, Oxford (Nuffield), Vienna and Warwick and the conference participants at SIOE 2017.

†University of Calgary, Department of Economics. E-mail: dimitri.migrow@ucalgary.ca

1 Introduction

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. For the uniform prior, 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 influential studies of Dalton (2017), Crozier (2009) and Cyert, March et al. (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 different preferences over this decision; specifically, the experts can 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 conditional on reports 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

[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, Fourné, and Elfring (2015).

[2]A manager cannot contract upon transfers or any information received in Dessein (2002), Alonso, Dessein, and Matouschek (2008), Alonso, Dessein, and Matouschek (2015), and Grenadier, Malenko, and Malenko (2016). See also the literature discussion in Gibbons and Roberts (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.

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 the information and communication structure in Morgan and Stocken (2008) and - for the case of the uniform prior - upon Galeotti, Ghiglino, and Squintani (2013) who study simultaneous communication in a similar setting. The crucial difference to Galeotti et al. (2013) - apart from a more general prior structure in my model - is that I study the optimal sequential structure from the 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 (Garicano, 2000; Radner, 1993; Sah and Stiglitz, 1986) 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 (Ahuja, 2000; Oberg and Walgenbach, 2009).

I show that an optimal communication network resolves a trade-off between two competing forces. On the one hand, the intermediation force (illustrated in section 3) 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 (illustrated in section 4) 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 the pooling of 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.

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 (Jackson, 2010). My analysis shows that for a wide range of model parameters a star communication network is 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 - under some conditions on the players' prior distribution of the payoff-relevant information - 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 it they were 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 incentivizing 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.

My third main result assumes a uniform prior - which has been studied widely in the literature - and shows that either if the experts' biases are sufficiently close to one another while being large enough relative to the decision maker's preferences, or sufficiently close to the decision maker, 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 the incentivizing of highly-biased experts to reveal their private information truthfully.

In section 5 I examine the effect of allowing players to adapt mixed strategies in a setting with two experts. I show that for small biases the decision maker prefers simultaneous communication whereas for large biases she prefers intermediation by one of the experts. The intuition for the latter result is that a strongly biased expert communicates to the decision maker only if he is able to aggregate and strategically coarsen both signals. Interestingly, if the biases are small and intermediation is not optimal, the strategy profile requires one of the experts to communicate very precise information allowing the other expert to communicate sufficiently coarse information making a less tight constraint possible in the latter case.

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 et al. (1963) observe that managerial decisions are lobbied by groups of employees that provide distorted information to the authority. Similarly, Dalton (2017) and Crozier (2009) 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), Tushman 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 subsection discusses the related literature. Section 2 illustrates the main idea with a simple example. Section 3 presents the model and the main results. Section 4 provides specific results assuming a uniform prior and discusses the case of three experts. Section 5 studies mixed strategies with a uniform prior. Section 6 studies the case of experts with opposing biases, and the benefits and limitations of using non-hierarchical networks. Section 7 concludes.

Related Literature. This paper is motivated by the literature studying how to design organizations in the absence of full commitment power by the manager. The literature discusses several optimal policies such as closing down communication channels and rewarding those employees who focus on productive activities (Milgrom and Roberts, 1988) or delegating decisions to better informed parties (Alonso et al., 2008, 2015; Dessein, 2002; Rantakari, 2008). This paper studies a different incentive instrument: the design of communication hierarchies in the presence of arbitrarily many players with strategic communication motives.

My focus on hierarchies is motivated by the extensive literature on the optimality of hierarchical organizations. For example, Bolton and Dewatripont (1994), Van Zandt (1999) and Van Zandt and Radner (2001) show that hierarchies are optimal as they minimize information processing costs. Friebel and Raith (2004) show the optimality of hierarchies as they can prevent conflicts between different organizational layers. Empirically, the importance of hierarchical communication in firms is identified in Ahuja and Carley (1999) and Oberg and Walgenbach (2009).

I model communication as cheap talk (Crawford and Sobel, 1982). In the famous extensions of Crawford and Sobel (1982) to multiple senders (Battaglini, 2002; Krishna and Morgan, 2001a,b) full revelation of private information is possible if experts' reports are cleverly played against one another. However, such equilibrium construction is not possible in my model because the experts' signals are conditionally independent and so it is possible to receive different signals on-path. As a result, the decision maker cannot credibly threaten to punish the experts due to "incompatibility" of their reports as it is the case, for example, in Ambrus

[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).

and Lu (2014) and Mylovanov and Zapechelnyuk (2013) where the experts receive the same signal with probability 1 (or close to 1).

I model imperfectly informed experts that are also studied in Austen-Smith (1993) and Battaglini (2004). Both papers focus on how to improve the informativeness of decision-making either by comparing simultaneous and sequential communication (Austen-Smith, 1993) or by simultaneously consulting many experts (Battaglini, 2004). My focus is rather related to Austen-Smith (1993) but I allow for many arbitrarily biased experts and thus consider a large family of communication structures.

The following papers are most closely related to my paper. First, an important study by Wolinsky (2002) compares simultaneous communication to a partitioning of experts into groups when the decision maker lacks commitment power. Wolinsky (2002) shows that the decision maker can align the experts' objectives with her own goals by splitting the experts into groups and making them reveal information in instances where individual communication fails to do so. I differ from Wolinsky (2002) in two main respects. First, while he focuses on how partitioning the experts into groups enhances the informational content of their reports, I explicitly pose the question of the optimal design of communication structures. Second, he assumes that experts have the same preferences while some of my results exploit the strategic conflict between the experts: for example, in section 6.1 I show how the decision maker benefits from partitioning the experts into groups according to the sign of their biases.

Second, I borrow from Morgan and Stocken (2008) who study strategic information transmission through polling. Their information structure is Beta-binomial model and they study the properties of signal aggregation among the constituents who simultaneously report to a decision maker. As their paper naturally focuses on the question of simultaneous information aggregation, they do not study the optimal design of communication structures which is the main focus of my paper.

Finally, Galeotti et al. (2013) and Hagenbach and Koessler (2011) study equilibrium communication networks with one round of simultaneous communication. The case of the uniform prior in my model resembles Galeotti et al. (2013) with the difference that I pose the question of the optimal communication design whereas Galeotti et al. (2013) look at the properties of a single-round communication where each player can communicate to every other player. Hagenbach and Koessler (2010) study a related setting but introduce coordination motives between the 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 such as complete networks (Galeotti et al., 2013; Hagenbach and Koessler, 2010) or sequential communication with a single or multiple intermediaries arranged in a line (Ambrus, Azevedo, and Kamada, 2013; Ivanov, 2010).

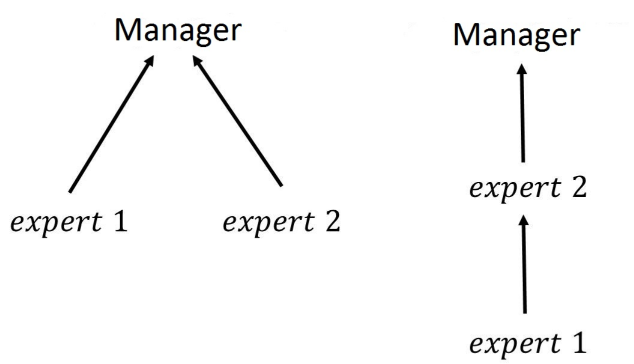
2 The main idea

The following example shows how a manager benefits from sequential communication in the simplest possible environment.

There is a manager ("she") and two experts ("he"). The manager chooses an action $y \in \mathbb{R}$ with the goal to minimize the loss function $-(y - \theta)^2$ where $\theta \in [0, 1]$ is the unobserved state. Each expert wants to minimize his loss function $-(y - (\theta + b))^2$ where b, called *bias*, parameterizes the conflict of interest between the experts and the manager and is known to all players. Notice that in this example there are no strategic motives between the experts. Assume $b > 0$ so that each expert wants the manager to take an action $\theta + b$ to the right of the state, whereas the manager wants to match the state.

Each expert receives a (conditionally independent) binary signal $s \in \{0, 1\}$ with $Prob(s = 1) = \theta$. The manager wants to elicit experts' signals by choosing a communication structure. Suppose that she can only choose between a star where the experts send simultaneous messages to the manager (Figure 1, left) or a line where expert 2 first listens to expert 1 and only then reports to the manager (Figure 1, right). Since the manager is unable to commit to side payments, communication is cheap talk. All model components are common knowledge. Which communication structure does the manager prefer?

Figure 1: Illustration of the main idea



I focus on pure strategy equilibria with the highest payoff for the manager.[5] Without loss of generality we can assume that each expert's message space is equivalent to his signal

[5]Mixed strategies are studied in section 5.

space. In the star network each expert has exactly two message strategies. He can either reveal his signal to the manager or send signal-independent messages (known as babbling). The manager's optimal action for $n \in \{0, 1, 2\}$ truthful messages, with $k \in \{0, .., n\}$ denoting the sum of the messages, is her posterior of θ:

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

Suppose, first, that only one expert (say, expert 1) reveals his signal. As $b > 0$, he wants to reveal his signal "1" resulting in $y(1, 1) = \frac{2}{3}$. If he truthfully reveals his signal "0" the manager chooses $y(0, 1) = \frac{1}{3}$. If he lies given his signal "0" (and the manager believes him), she chooses $\frac{2}{3}$ while expert 1 expects the state to be $\frac{1}{3}$. It is easy to show that expert 1 is truthful only if $b \leq \frac{1}{6}$.

Suppose, next, that the manager expects both experts to reveal their signals. Consider expert 1 with the signal "0". He believes that expert 2 has a signal "0" with probability $\frac{2}{3}$ and a signal "1" otherwise. Thus, if he communicates his signal truthfully, he expects the manager to choose $y(0, 2) = \frac{1}{4}$ with probability $\frac{2}{3}$ and $y(1, 2) = \frac{1}{2}$ otherwise. By deviation, he expects the manager to choose $y(1, 2) = \frac{1}{2}$ with probability $\frac{2}{3}$ and $y(2, 2) = \frac{3}{4}$ otherwise. As in both cases the shift in the manager's action is $\frac{1}{4}$, expert 1 does not lie if $b \leq \frac{1}{8}$ (the same is true for expert 2).

Third, consider the line network where expert 1 reveals his signal to expert 2 and so the latter observes two signals from the set $\{00, 01, 10, 11\}$. Suppose that expert 2 uses the following *strategic pooling* in his message strategy - he either informs the manager that the sum of both signals is "00", or that the actual signal combination is *within the set* $\{01, 10, 11\}$. The second message is "noisy". Notice the difference to how noise is modeled in Myerson (1991) and Blume, Board, and Kawamura (2007). In their studies, conditional on some signals a message perfectly arrives to a receiver with some probability, and is drawn from a fixed error distribution otherwise. In the current model the noise does not come from an error distribution but from a strategic decision of the sender to pool multiple signals into a single message.

Upon receiving the first message the manager chooses $\frac{1}{4}$, and upon receiving the second message she chooses $y = \frac{5}{8}$ that is the midpoint between her action $\frac{1}{2}$ were she able to exactly observe the signals "01" or "10", and her action $\frac{3}{4}$ were she able to exactly observe the signals 11. Since the experts are positively biased, the only relevant deviation has to be conditioned on both signals being 0. It is easy to show that the experts do not have incentives to deviate from the above message strategies if b does not exceed the half-distance between the two equilibrium actions of the manager, which implies $b \leq \frac{3}{16}$.

Notice that due to the pooling of the signals in the line the difference between the two equilibrium actions of the manager is larger compared to the star with one truthful report. This is why it is costlier for the experts to misreport in the line compared to the star. As a consequence, the line supports information transmission where the star does not (which is the case for $b > \frac{1}{6}$). As I show in section 3, strategic pooling by an intermediary benefits the decision maker for a large family of priors and not only for the uniform prior in the current example.

Furthermore - as shown in the Appendix - the above equilibrium in the line strictly dominates both the babbling equilibrium and the equilibrium with a single truthful signal. As a result, for

$$b \in \left(\frac{1}{8}, \frac{3}{16}\right]$$

the line results in a strictly higher expected payoff for the manager compared to the star. As shown in section 5, allowing for mixed strategies does not change this result showing its robustness.

3 Model

An organization consists of $n \geq 2$ experts ("he") and a decision maker ("she", denoted by DM). There is an unobserved state θ that can represent a parameter of the market demand, a state of technology or a profit-maximizing capital allocation. The decision maker wants to match the state with her action $y \in \mathbb{R}$ that might represent, for example, the amount of an internal budget allocated to a particular project the profitability of which depends on θ. Her payoff function $-(y-\theta)^2$ indicates progressively higher losses the further away is the action from the actual state. Each of the experts from the set $N^e := \{1, .., n\}$ has a systematic bias relative to the true state, for example due to career concerns. The payoff function of an expert $i \in N^e$ is $-(y - \theta - b_i)^2$. I assume that the biases can take arbitrary values, and that the entire bias vector is common knowledge. This parameterization of players' preferences is attractive due to the linearity of best responses and is prominent in organizational economics (Gibbons, Matouschek, and Roberts, 2013).

The state is unobserved and is commonly known to be distributed on $[0, 1]$ according to a Beta distribution with the parameters (α, β):[6]

$$f_{\alpha,\beta}(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}, \ \alpha > 0, \beta > 0.^7$$

[6] To my knowledge, Morgan and Stocken (2008) is the first paper using a Beta-binomial model in the context of information aggregation.

[7] The two key properties of the gamma function $\Gamma(\cdot)$ are: 1) for an integer c we have $\Gamma(c) = (c-1)!$ and 2) since $\alpha > 0, \beta > 0$ we have $\Gamma(c) = (c-1)\Gamma(c-1)$.

The set of prior distributions is a family of continuous functions with the full support on $[0, 1]$. The Beta prior is a natural choice: first, it comes from the conjugate family (as shown below) and second, depending on the distributional parameters (α, β) it can take one of many functional forms featured in the literature such as the uniform case ($\alpha = \beta = 1$), linearly increasing or decreasing (e.g. $\alpha = 2, \beta = 1$) or bell-shaped (e.g. $\alpha = \beta = 2$). To save on notation I omit (α, β) when referring to functional notations.

Each expert receives a private binary signal $s_i \in \{0, 1\}$ with $Prob(s_i = 1) = \theta$. If a player knew the exact sum k of n signals, his posterior would be:

$$f(\theta|k, n) = \frac{\Gamma(\alpha + \beta + n)}{\Gamma(\alpha + k)\Gamma(\beta + n - k)} \theta^{\alpha+k-1}(1 - \theta)^{\beta+n-k-1}.$$

As we see later, this simple structure of discrete signals results in a tractable summary statistic. The decision maker chooses a communication network that determines both who talks to whom and in which order. The network is represented by a directed graph $Q = (N, E)$ with the set of nodes N and an $(n+1)\times(n+1)$ adjacency 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. I use a notion of a *path* in a network Q, $H_{i_1 i_L}(Q)$, that is a sequence of nodes $i_1, i_2, .., i_L$ such that $e_{i_l i_{l+1}} = 1$ for each $l = 1, .., L - 1$. Building upon a large literature on hierarchies (in the paper I use the terms hierarchy and network interchangeably) I focus on directed tree networks where the decision maker is located at the "root" of the tree. Specifically, the underlying set of networks $\mathbb{Q}$ satisfies the following properties:[8]

1. each expert has a single outgoing link and (possibly) multiple incoming links,

2. there are no cycles, and

3. the decision maker has at least one incoming link and no outgoing links.

Further, a network determines the order of play: if there are two links $e_{ij} = 1$ and $e_{jv} = 1$, then j first listens to i and only then talks to v. Each expert's i *type* is his own private signal and - depending on the network - the information that he receives from other experts.

In the main body of the paper I study pure strategy equilibria (see section 5 for mixed strategies). To define players' strategies let $P(S)$ denote a partition of an arbitrary set S. Each expert's communication strategy is a partition of his type space. For each $i \in N$, let $N_i(Q) := \{j \in N^e : e_{ji} = 1\}$ define the set of experts who can send messages to i in a network Q. If an expert $i \in N^e$ babbles, it means that he sends messages independent of his private

[8]See the appendix for the formal definition

information, his communication strategy is denoted by $P_i^b(Q)$. The *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

$$P_j : \ T_j(Q) \rightarrow P(T_j(Q)),$$

where

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

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

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

The decision maker's beliefs in a network Q are determined by the information and 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

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

Most communication networks give rise to multiple equilibria which is a common feature of strategic communication games. I focus on the best equilibrium for the decision maker in any given network.[9] When referring to optimality I use two concepts:

1. for an unrestricted domain of biases I say that a network Q weakly dominates a different network Q' if, for any players' biases, the best equilibrium payoff for the decision maker

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

in Q is at least as high as in Q', and for some biases it is strictly higher,

2. for a restricted domain of biases I say that Q dominates a different network Q' if – for the specified biases – Q results in a strictly higher expected payoff for the decision maker than Q'.

A Perfect Bayesian Equilibrium in pure strategies is a tuple

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

such that 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 $p_i' \neq p_i$.

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

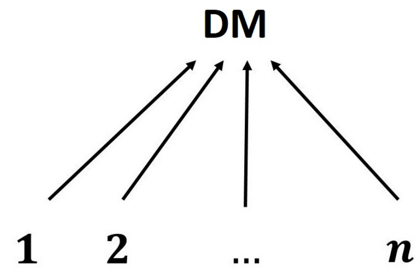
Given the equilibrium conditions, the decision maker aims to match the state and chooses $y(\cdot) = \mathbb{E}_{DM}(\theta|p)$ for $p \in P_{DM}(Q)$.

3.1 Simultaneous versus sequential communication

A natural benchmark is a network called a *star* where each $i \in N^e$ has a link $e_{iDM} = 1$ (Figure 2). The equilibria in a star are easy to characterize (see also Morgan and Stocken (2008)):

Proposition 1: *Fix any (α, β). Take any number of experts, n, with arbitrary biases. An equilibrium in a star network with $n' \leq n$ experts communicating their signals truthfully to the decision maker exists iff for every expert $i \in \{1, .., n'\}$, $|b_i| \leq \frac{1}{2(\alpha+\beta+n')}$. The expected payoff of the decision maker is $-\frac{\alpha\beta}{(\alpha+\beta)(1+\alpha+\beta)(n'+\alpha+\beta)}$.*

Figure 2: A star with n experts



To understand Proposition 1 notice, first, that for n' equilibrium truthful signals and the sum of the truthful signals k, the posterior of the decision maker is

$$f(\theta|k, n') = \frac{\Gamma(\alpha + \beta + n')}{\Gamma(\alpha + k)\Gamma(\beta + n' - k)}\theta^{\alpha+k-1}(1 - \theta)^{\beta+n'-k-1}.$$

Given the posterior, the decision maker chooses the action

$$y(k, n') = E_{DM}(\theta|k, n') = \int_0^1 \theta f(\theta|k, n')d\theta = \frac{\alpha + k}{\alpha + \beta + n'}.$$

Consider any expert i who is believed (correctly in equilibrium) to truthfully reveal his signal. If $s_i = 0$ and i lies about his signal, he expects a shift in the decision maker's action to the *right* of his expected state by the distance $\frac{1}{\alpha+\beta+n'}$. Similarly, if $s_i = 1$ and i lies about his signal, he expects a shift in the decision maker's action to the *left* of his expected state by $\frac{1}{\alpha+\beta+n'}$. Since the expert's payoff function is symmetric around his ideal point, he does not deviate for $|b_i| \leq \frac{1}{2(\alpha+\beta+n')}$. Further, notice that for any (α, β) the decision maker's posterior increases in precision with larger n'. As we see from her expected payoff in Proposition 1, she benefits from a larger number of truthful signals. However, since her posterior is concave in the number of truthful signals, a larger n' makes it more profitable for each expert to deviate. As a result, a larger n' corresponds to smaller biases (in absolute value).

Proposition 1 further reveals that for any given (α, β) the experts' communication incentives depend on the equilibrium number of truthful messages and not on the exact realization of the signals. In other words, even if an expert observes some of other experts' signals, his deviation incentives would not change. This has a direct implication on how an optimally organized hierarchy can result in the same number of truthful signals and generate the same payoff profile as the star network.

Proposition 2: *Fix any* (α, β) *and take any number* n *of experts with* F *biases. Take any network* Q *that is not a star with the restriction that for any* i *connected to* j, $e_{ij} = 1$, *it must be true that* $|b_j| \leq |b_i|$. *Then any equilibrium number of truthful signals and the corresponding payoff profile in the star is also implementable in* Q.

The optimal ordering requirement in Proposition 2 states that the experts with smaller biases (in absolute value) have to be closer to the decision maker. To illustrate the intuition suppose that this condition is violated. Suppose there is an expert with a high enough bias ($b > \frac{1}{2(\alpha+\beta+1)}$) who never reveals his signal in a star whatever the number n' of truthful experts. Proposition 1 tells us that he would not reveal his signal even if he knew the signal realizations of the other experts since for any given n' his incentives do not depend on k. But then, if the decision maker expects the same information in a (multi-layer) hierarchy, as in the star, she should not make this expert an intermediary.

Does it mean that a hierarchy can never incentivize a highly biased expert to reveal his signal? The answer is no as the introductory example shows. In the example the pooling of signals connected to higher states and separating them from signals connected to lower states, increases experts' deviation costs. As a result, highly biased experts reveal some information in the line where the star does not transmit any information.

The next proposition generalizes the introductory example for two experts and shows that an *arbitrarily* ordered line transmits coarse information from the experts to the decision maker where the star does not, if one of the following conditions holds: either the experts are positively biased and the prior distribution assigns sufficiently high beliefs to higher states, or the experts are negatively biased and the prior assigns sufficiently high beliefs to lower states.

Proposition 3: *Take two experts with the biases strictly above* $|\frac{1}{2(\alpha+\beta+1)}|$ *such that the only equilibrium in the star is babbling. An arbitrarily ordered line dominates the star if either* $\alpha > \frac{1}{2}(\sqrt{1+\beta(\beta+6)} - 1 - \beta)$, $\beta > 0$ *and*

$$b_{1,2} \in \Big(\frac{1}{2(\alpha+\beta+1)}, \frac{1+\alpha+\beta}{(2+\alpha+\beta)(1+\alpha+2\beta)}\Big]$$

resulting in the decision maker's expected payoff

$$-\frac{\alpha\beta(4+\alpha^2+\beta(5+2\beta)+\alpha(5+3\beta))}{(\alpha+\beta)(1+\alpha+\beta)(2+\alpha+\beta)^2(1+\alpha+2\beta)}$$

or $\beta > \frac{1}{2}(\sqrt{1+\alpha(\alpha+6)} - 1 - \alpha)$, $\alpha > 0$ *and*

$$b_{1,2} \in \Big[-\frac{1+\alpha+\beta}{(2+\alpha+\beta)(1+2\alpha+\beta)}, -\frac{1}{2(\alpha+\beta+1)} \Big)$$

resulting in the decision maker's expected payoff

$$-\frac{\alpha\beta(4+\beta^2+\alpha(5+2\alpha)+\beta(5+3\alpha))}{(\alpha+\beta)(1+\alpha+\beta)(2+\alpha+\beta)^2(1+2\alpha+\beta)}.$$

To understand the equilibrium construction in Proposition 3 (that I use throughout the paper), think of two positively biased experts with biases strictly exceeding $\frac{1}{2(\alpha+\beta+1)}$. From Proposition 1 we know that the only equilibrium in the star is a non-informative one (babbling). Denote the bottom expert by 1 and the intermediary by 2: $e_{12} = 1$ and $e_{2DM} = 1$. Consider a strategy profile in which expert 1 truthfully reveals his signal and expert 2 either truthfully reveals the signals 00 or sends a message $\tilde{p}$ if both signal realizations are either 01, 10 or 11.

Upon receiving the first message the decision maker forms a posterior

$$f(\theta|0,2) = \frac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha)\Gamma(\beta+2)}\theta^{\alpha-1}(1-\theta)^{\beta+1}$$

with the expected value of θ, $\frac{\alpha}{\alpha+\beta+2}$. Upon receiving the second message, she forms the posterior via Bayes' rule:

$$f(\theta|\tilde{p},2) = \frac{\sum_{k=1}^{2}\frac{2!}{k!(2-k)!}\theta^k(1-\theta)^{2-k}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}}{\int_0^1\sum_{k=1}^{2}\frac{2!}{k!(2-k)!}\theta^k(1-\theta)^{2-k}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}d\theta} =$$

$$\frac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+1)\Gamma(\beta)}\frac{\theta^{\alpha}(1-\theta)^{\beta-1}(2-\theta)}{(2\beta+\alpha+1)}$$

and the corresponding expected value is $\int_0^1 \theta f(\theta|\tilde{p},2)d\theta = \frac{(\alpha+2\beta+2)(\alpha+1)}{(\alpha+2\beta+1)(\alpha+\beta+2)}$.

Consider, next, the experts' communication incentives. Since both experts are positively biased we only need to check deviations to messages resulting in a higher expected posterior of the decision maker. In particular, given the above message strategy the incentives for an upward deviation for *both* experts have to be conditioned on both signals being 0. As I show in the appendix, an expert does not deviate if his bias is weakly below the half-distance between his expected value of θ if both signals are 0 (which is $\frac{\alpha}{\alpha+\beta+2}$) and the decision maker's action if the signals 00 are not truthfully revealed (which is $\frac{(\alpha+2\beta+2)(\alpha+1)}{(\alpha+2\beta+1)(\alpha+\beta+2)}$). Therefore, expert 1 does not deviate from truthfully revealing his signal 0 and expert 2 doesn't deviate from truthfully revealing the signals 00 if

$$b_i \leq \frac{1+\alpha+\beta}{(2+\alpha+\beta)(1+\alpha+2\beta)},\ i=1,2.$$

Clearly, the above message profile yields a higher expected payoff to the decision maker compared to the babbling equilibrium even though communication contains noise.[10] From Proposition 1 we know that for $b_{1,2} > \frac{1}{2(\alpha+\beta+1)}$ the only equilibrium in the star is babbling. Therefore, the decision maker strictly prefers the line over the star if $\frac{1+\alpha+\beta}{(2+\alpha+\beta)(1+\alpha+2\beta)} > \frac{1}{\alpha+\beta+1}$ or

$$\alpha > \frac{1}{2}(\sqrt{1+\beta(\beta+6)}-1-\beta)$$

where $\frac{1}{2}(\sqrt{1+\beta(\beta+6)}-1-\beta) < 1$ and monotonically converges to 1 (from below) as β goes to infinity. As the prior is proportional to $\theta^{\alpha-1}(1-\theta)^{\beta-1}$, a sufficently high α means that the players are sufficiently "optimistic" about the state (the prior assigns sufficiently high beliefs to higher states). As we see from the above discussion, if $\alpha \geq 1$, even for an arbitrarily large β there still exists a range for positive experts' biases where the line outperforms the star that only has a babbling outcome.

The case for negative biases is similar. The equilibrium construction uses the fact that an intermediary communicating to the decision maker either sends a message containing the signals 00, 01 and 10, or reveals the signals 11. Thus, in terms of the equilibrium construction the case of negative biases is the mirror image of the case with positive biases.[11]

Which network should be favored by the decision maker who has access to more than two experts? The next proposition characterizes a network with at least three experts that results in a coarse information transmission where a star (with arbitrarily many experts) or a line with two experts do not transmit any information. This happens if either for positively biased experts the prior assigns sufficiently high beliefs to higher states or, for negatively biased experts, the prior assigns sufficiently high beliefs to lower states. The corresponding network should allow one of the experts to collect information from two of the other experts.

[10] My hypothesis is that the construction in Proposition 3 is also applicable to more general signals if the signal precision is concave in the number of signals and – whenever the biases are sufficiently positive (negative) – if the prior assigns sufficiently high probability to higher (lower) states. The reason is that for, say, positive biases, by pooling the high states there is enough probability mass attached to higher messages in order to generate a large a enough shift in the posterior of the decision maker by deviation. This would make deviations costly even for large biases.

[11] The insight that noise facilitates communication is, of course, not new (Blume et al., 2007; Myerson, 1991). The novelty of the above approach is that noise is created by connecting the experts which results in strategic pooling of aggregated signals.

Proposition 4: *Take $n \geq 3$ experts with all absolute values of biases strictly above $\frac{1+\alpha+\beta}{(2+\alpha+\beta)(1+\alpha+2\beta)}$ such that the only equilibrium in a star and any line with two experts is babbling. Take any Q which has an expert $i \in N^e$ connected to the decision maker, $e_{iDM} = 1$, an expert $j \in N^e$ connected to i, $e_{ji} = 1$, and another expert $j' \in N^e$ either connected to j, $e_{j'j} = 1$, or connected to i, $e_{j'i} = 1$. For $\alpha \geq 1,\ \beta > 0$ and*

$$b_i, b_j, b_{j'} \in \Big(\frac{1+\alpha+\beta}{(2+\alpha+\beta)(1+\alpha+2\beta)}, \frac{3(1+\alpha+\beta)(2+\alpha+\beta)}{2(3+\alpha+\beta)(2+\alpha^2+3\alpha(1+\beta)+3\beta(2+\beta))}\Big]$$

or $\beta \geq 1,\ \alpha > 0$ and

$$b_i, b_j, b_{j'} \in \Big[-\frac{3(1+\alpha+\beta)(2+\alpha+\beta)}{2(3+\alpha+\beta)(2+\beta^2+3\beta(1+\alpha)+3\alpha(2+\alpha))}, -\frac{1+\alpha+\beta}{(2+\alpha+\beta)(1+2\alpha+\beta)}\Big)$$

the network Q results in a strictly higher payoff for the decision maker compared to a star or any line with two experts.

The equilibrium construction is similar to the one in Proposition 3 and uses three signals.[12] In the case of positive biases the expert i who is connected to the decision maker, aggregates information about three signals and sends only two messages in equilibrium. He either informs the decision maker that the sum of the three signals is 0, or that the sum of the three signals is within the set $\{1, 2, 3\}$. This message strategy has a minimal informational content: the decision maker only knows if all signals are 0, or not. Since almost all signal combinations are pooled together in a higher message, a sufficiently large α ensures that a deviation from the lower message results in a large shift of decision maker's action from the expert's expected value of the state. Thus, even the experts with large biases are incentivized to stick to the above message strategy. The equilibrium construction for negative biases is similar - an intermediary either informs the decision maker that the sum of all three signals is 3, or not.

A consequence of Propositions 2 and 3 is stated in the next Corollary. It shows when, for $n \geq 2$, an optimally designed hierarchy dominates the star:

Corollary 1: *Take any network Q that is not a star, with at least two experts such that if expert i communicates to expert j, $e_{ij} = 1$, then $|b_j| \leq |b_i|$. For either $\alpha > \frac{1}{2}(\sqrt{1+\beta(\beta+6)} - 1 - \beta)$ or $\beta > \frac{1}{2}(\sqrt{1+\alpha(\alpha+6)} - 1 - \alpha)$ the network Q results in a higher expected payoff for the decision maker than the star.*

We see the power of information aggregation in hierarchies - an intermediary can pool signals

[12]A more general argument using a similar equilibrium construction for n signals can be shown using simulation results. Its qualitative predictions are similar to Propositions 3 and 4.

and increase the costs of deviation from the message strategies of *all* experts participating in a non-babbling communication. Thus, aggregation of signals helps to overcome incentive problems faced by experts with sufficiently large biases.

4 Uniform prior

In this section I assume a uniform prior ($\alpha = \beta = 1$: see, e.g., Galeotti et al., 2013 or Argenziano et al., 2015) that describes a situation in which the experts do not possess any prior knowledge about the underlying parameter of interest (e.g. the quality of the project). The corresponding optimal network in case of large and similar biases requires (1) a single intermediary to be connected to the decision maker, (2) to aggregate all signals and (3) to report his information in the coarsest possible way.

Proposition 5: *Assume that the biases of all experts are either within the interval* $(\frac{n}{4(n+1)}, \frac{n+1}{4(n+2)}]$ *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 5 goes as follows. Given the parameter restrictions, there exists an equilibrium in which a single intermediary connected to the decision maker is informed about all signals. This enables him to send messages according to the coarsest possible message profile: in the case of positive biases, for example, this expert either sends a message containing the sum of the signals 0, or a second message that pools together all the other signals.[13] As a result, all experts condition their upward deviation on all n signals being 0. Whenever an expert with a summary statistic 0 deviates to a message with a higher summary statistic, and the actual sum of all signals is 0, the decision maker chooses a policy *far to the right* of the experts' expected value of the state. In particular, the upper bound $\frac{n+1}{4(n+2)}$ in Proposition 5 is determined by an upward deviation constraint of the partition $\{\{0\}, \{1, .., n\}\}$. The lower

[13]For negative biases an intermediary either informs the decision maker that the sum of the signals is n or that the sum of the signals is not n.

bound is determined by the corresponding upward deviation constraint with $n-1$ experts where the single intermediary communicates to the decision maker according to the partition $\{\{0\},\{1,..,n-1\}\}$. This results in the bound $\frac{(n-1)+1}{4((n-1)+2)} = \frac{n}{4(n+1)}$.

Figure 3: Examples of optimal networks in Proposition 5

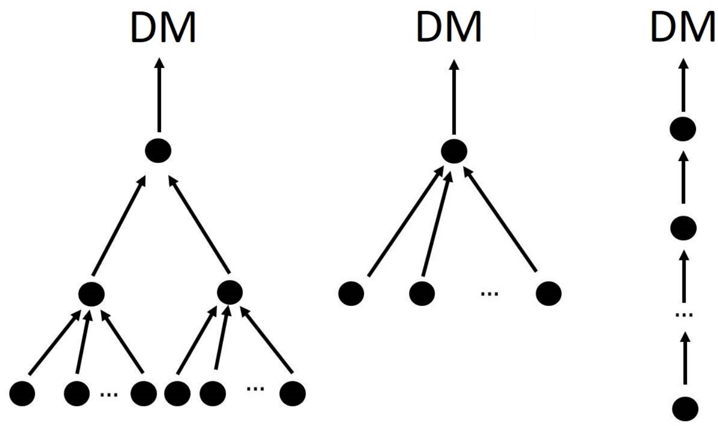


Figure 3 shows the examples of possible optimal networks characterized in Proposition 5. The ordering of the experts can be arbitrary. This is because the intervals specified in Proposition 5 are not large enough to create incentive conflicts between the experts. A direct consequence of Propositions 5 is

Corollary 2: *Whenever the absolute values of biases of all $n \geq 2$ experts are strictly above $\frac{1}{6}$ and there are either $m \geq 2$ experts with $\frac{1}{6} < b_i \leq \frac{n+1}{4(n+2)}$, $i = 1,..,m$, or $m' \geq 2$ experts with $-\frac{n+1}{4(n+2)} \leq b_j < -\frac{1}{6}$, $j = 1,..,m'$, then the communication strategies in the optimal network feature strategic coarsening of information.*

Consider now the case of sufficiently small biases (but not too small to completely alleviate the players' conflict of interest). The optimal network requires the expert with the smallest $|b|$ to communicate to the decision maker and all other experts to be directly connected to that expert.

Proposition 6: *Take any number $n \geq 2$ of experts such that*

1. *either there is an expert i with the bias $0 < b_i \leq \frac{1}{2(n+2)}$ and the bias of each expert $j = 1,..,n$, $j \neq i$, satisfies $\frac{1}{2(n+2)} < b_j \leq \frac{3+n(n+1)}{2n(2+3n+n^2)}$ or*

2. *there is an expert i with the bias $-\frac{1}{2(n+2)} \leq b_i < 0$ and the bias of each expert $j = 1,..,n$, $j \neq i$, satisfies $-\frac{3+n(n+1)}{2n(2+3n+n^2)} \leq -\frac{1}{2(n+2)}$.*

Then, the optimal network requires i to be connected to the decision maker, $e_{iP} = 1$, and each other expert to be connected to i, $e_{ji} = 1$. The corresponding expected payoff for the decision maker is $-\frac{n^2+3n+5}{6(n+1)(n+2)^2}$.

The underlying logic is related to Proposition 5 with the difference that the message profile in Proposition 6 features the *smallest possible* incentive compatible coarsening of signals. For example, for positively biased experts the partition according to which the decision maker receives her information is

$$\{\{0\}, \{1\}, ..\{n-2\}, \{n-1, n\}\}.$$

Notice that Proposition 5 only requires the decision maker to receive information from a single expert whereas Proposition 6 provides an additional restriction as it requires the intermediary to have the smallest bias and all other experts to have a directed link to the intermediary. The reason for the additional restriction is to maximize the aggregate uncertainty among the experts about exactly which element of the decision maker's partition is communicated to her. We next study how the uncertainty of experts about the signals of other experts affects communication incentives.

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.[14]

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, $e_{ji} = 1$.*

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

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, in an optimal network he receives this information unmediated meaning from each of those experts directly.

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

4.1 Optimal network for 3 positively biased experts

Using the insights of the previous sections, this section characterizes the optimal networks and corresponding equilibria for 3 positively biased experts where none of the experts babbles.[15] 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 $\mathbb{E}U_{DM} \simeq -0.033$.

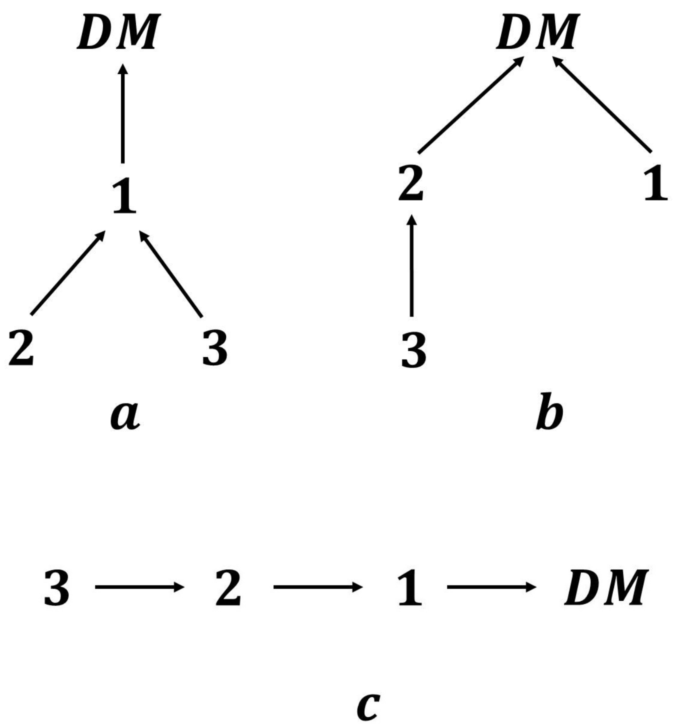
For the following assume $b > 0.1$ for at least one of the experts. The optimal network depends on the biases as follows

1. *If the biases are small, the optimal network is depicted in Figure 4(a) (denoted by $\underline{Q_a}$).* 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 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. *If the biases are in the intermediate range, the optimal network is depicted in Figure 4(b) (denoted by $\underline{Q_b}$).* 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.

[15] If, for example, one of the experts babbles, we are back to the case of two communicating experts covered in the introductory example.

Figure 4: Networks for 3 experts



3. *If the biases are large, the optimal networks are depicted in Figure 4(a) ($\underline{Q_a}$) and 4(c) (the latter is denoted by $\underline{Q_c}$).* This is a direct consequence of Proposition 5. Fix $0.1875 < b_1 \leq 0.2$ 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 $\mathbb{E}U_{DM} \simeq -0.05$.

5 Mixed strategies with a uniform prior and two experts

In this section I allow for mixed strategies while only looking at two experts. Studying mixed strategies is important for at least three (related) reasons. First, in the context of discrete type spaces it is well known that mixing allows for a richer strategy space potentially resulting in more information transmitted (see, e.g., an example in Myerson (1991)). Second, it is known from a related context featuring information aggregation (e.g. Austen-Smith (1993)) that acquiring information by a decision maker from different sources can incentivize experts to distort information in order to "outweigh" the decision maker's information. But this implies that partial revelation of information by some experts can incentivize truthtelling by other experts which implies that some of the mixed strategy equilibria can payoff-dominate pure strategy equilibria. Finally, Ivanov (2010) and Ambrus et al. (2013) show how biased intermediaries can improve communication which means that a strongly biased expert who uses mixed strategies could be efficient in information transmission. The current model puts a strong constraint on the tractability of general results with mixed strategies as both the possibility of strategic coarsening of discreet signals and the randomizations over the partitions of available signals typically result in a very large type space. For this reason, I only look at the two-experts case.

For simplicity, I assume that the each expert's message space is countable. Since the equilibrium definition is standard, I relegate it to the appendix. Suppose that both experts are biased in the same direction (biases of different signs are analyzed in the appendix). A mixed strategy equilibrium in the star has the following feature:

Lemma 3: *In a mixed strategy equilibrium in a star with two positively (negatively) biased experts, each expert always truthfully reveals his signal 1 (signal 0).*

To understand this result, suppose that a positively biased expert randomizes given his signal 1, and suppose that in this case he either sends m or m'. As one of the messages results in a higher posterior of the decision maker, assume $E(\theta|m, m_{-i}) < E(\theta|m', m_{-i})$ where m_{-i} denotes a realization of the mixed strategy of the other expert. For any message $m_{-i} \in M$ from the other expert, the expert's posterior is larger than the decision maker's posterior after she receives the message profile (m, m_{-i}). But then the expert always prefers to send m' instead of m and therefore a mixed strategy that violates Lemma 3 is not incentive-compatible.

To see how allowing for mixed strategies can make the star optimal for the decision maker I focus on the simplest case with $|M| = 2$. In words, the players have two messages at their disposal, and given Lemma 3, if the experts are positively (negatively) biased then each expert uses both messages when his signal realization is 0 (is 1) and a single message if his signal realization is 1 (is 0). Thus, the mixed strategy profile for positively biased experts goes as follows:[16] for each $i = 1, 2$, expert i with signal $s_i = 0$ sends $m_i = 0$ with probability $p_i \in [0, 1]$, and $m_i = 1$ otherwise. Moreover, he sends $m_i = 1$ if $s_i = 1$. For this strategy profile, we obtain the following result:

Proposition 7: *A star dominates an arbitrarily ordered line with two positively biased experts if the biases satisfy*

$$b_1 = f(p_1, p_2) \text{ and } b_2 = g(p_1, p_2)$$

such that[17]

[16]The case of negatively biased experts is symmetric.

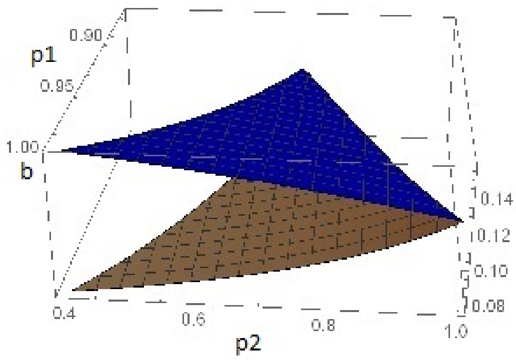
[17]The exact functions are $f(p_1, p_2) = \frac{(4(14-5p_1)p_1-39)p_2^3+2(p_1(55p_1-159)+114)p_2^2-3(p_1(61p_1-180)+132)p_2+24(3-2p_1)^2}{8(2p_1-3)(2p_2-3)(p_1(2p_2-3)-3p_2+6)(p_1(3p_2-4)-4p_2+6)}$

1. $p_2 \geq \frac{18p_1(9p_1-23)-\sqrt{3}\sqrt{4p_1(p_1(p_1(11p_1-102)+312)-387)+675}+261}{4(p_1(32p_1-81)+51)}$, *and*

2. *either* $b_1 > \frac{1}{8}$ *or* $b_2 > \frac{1}{8}$.

Although the result is somewhat technical, the intuition is straightforward. First, condition 1 in the above proposition indicates that both probabilities p_1 and p_2 have to be sufficiently high to ensure information transmission to the decision maker is not too noisy. This is because for each $i = 1, 2$ a higher p_i corresponds to a higher likelihood that expert i will truthfully reveal his signal 0. Moreover, the consequence of condition 2 is that only one bias can exceed $\frac{1}{8}$. It implies that the other bias has to be sufficiently small to ensure a sufficiently high informativeness of communication.

Figure 5: Illustration of Proposition 7



The corresponding values of biases as a function of (p_1, p_2) are illustrated in Figure 5. The dark (blue) area shows the range for b_1, and the lighter (pale brown) area shows the range for b_2 depending on (p_1, p_2) in the region where the star dominates the line. In the figure, $b_2 \leq \frac{1}{8} \leq b_1$ indicates that one of the experts is required to have a relatively small bias. This finding highlights a trade-off where a coarser communication by one of the experts (expert 1) has to be compensated by a finer information transmission by another expert (expert 2) in order for the decision maker to prefer the star over any mixed strategy profile in the line.

Finally, given the calculations in the appendix it is straightforward to show that for symmetric mixed strategies with two experts the star never outperforms the line if both biases strictly exceed $\frac{1}{8}$ which can be interpreted as the robustness check of the introductory example.

and $g(p_1, p_2) = \frac{p_1^3(4(14-5p_2)p_2-39)+2p_1^2(p_2(55p_2-159)+114)-3p_1(p_2(61p_2-180)+132)+24(3-2p_2)^2}{8(2p_1-3)(2p_2-3)((2p_1-3)p_2-3p_1+6)((3p_1-4)p_2-4p_1+6)}$

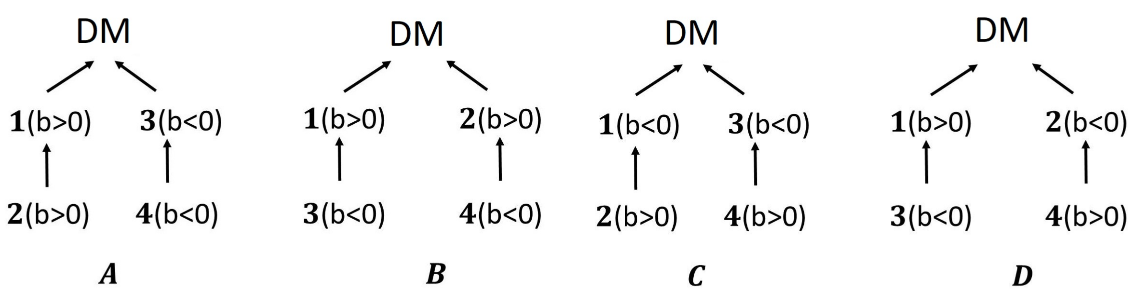
6 Further results

6.1 Biases of opposing signs

Suppose that a decision maker consults four experts two of which have the same positive bias, $b_1 = b_2 = b^+ > 0$, and the other two 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 6. 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 when the experts have biases of opposite signs. I show that it is optimal to put experts with the same biases into a single line and therefore to separate them from the experts with biases of the opposite sign. The reason for this is twofold. First, having the same biases within each group prevents an incentive conflict so that signals are shared. Second, within each line the experts distort their reports in the direction opposite to the distortion of reports in the other line. As the decision maker wants to minimize the residual variance, she benefits from having two reports biased in opposite directions.

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

Figure 6: Possible group arrangements



In case A (Figure 6A) the experts are partitioned into two lines according to the sign of their biases. In cases B, C and D (Figure 6B, 6C, 6D) 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

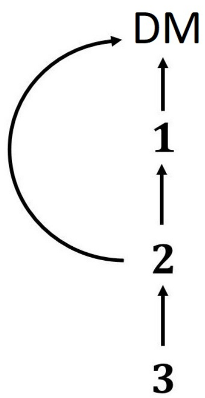
$$y(p_1, p_3) = \frac{2}{9},\ y(p_1, p_3') = y(p_1', p_3) = \frac{1}{2},\ y(p_1', p_3') = \frac{7}{9}.$$

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 6A dominates all other networks.[18]

6.2 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.

Figure 7: Expert 2 has two outgoing links



Example: Consider three positively biased experts organized in the network depicted in Figure 7, 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 case if p_2' is sent) the decision maker receives additional information about expert 1's signal.

[18] 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.

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 incentive 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 7 but not incentive compatible in the line.

7 Conclusions

This paper studies the optimal design of intra-organizational communication where private information is a strategic asset due to a conflict of interest within an organization and the lack of full commitment on the side of the management.

I show how an intermediation force in the optimal hierarchy leads to aggregation and incentive-compatible pooling of signals in the instances where no information transmission would be possible were the experts assigned to smaller groups. Moreover, an uncertainty force familiar from the mechanism design setting separates the experts into groups and relaxes their communication incentives. An optimally designed hierarchy can always generate the same outcomes as simultaneous communication. More importantly, if the positively (negatively) biased experts assign sufficiently high weight to higher (lower) states, *any* optimally ordered hierarchy outperforms the star. Since an optimal hierarchy partitions the experts into groups where the experts exchange and coarsen their signals, the equilibrium communication resonates with communication structures discussed in Cyert et al. (1963), Crozier (2009) and Dalton (2017).

In real-life organizations communication patterns can be very complex. A hierarchy is a useful first step. For future research it will be important to look at other communication structures such as networks accommodating for cycles, communication to multiple audiences and communication links of varying "strength".

Moreover, communication is a dynamic activity typically featuring 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 single round communication (Krishna and Morgan, 2004). It is interesting to see how the equilibria in optimal hierarchies are affected by communicating dynamics.

References

AHUJA, G. (2000): "Collaboration networks, structural holes, and innovation: A longitudinal study," *Administrative science quarterly*, 45, 425–455.

AHUJA, M. K. AND K. M. CARLEY (1999): "Network structure in virtual organizations," *Organization Science*, 10, 741–757.

ALLEN, T. J. (1977): "Managing the flow of technology: Technology transfer and the dissemination of technological information within the R & D organization(Book)," *Research supported by the National Science Foundation. Cambridge, Mass., MIT Press, 1977. 329 p.*

ALONSO, R., W. DESSEIN, AND N. MATOUSCHEK (2008): "When does coordination require centralization?" *American Economic Review*, 98, 145–79.

——— (2015): "Organizing to adapt and compete," *American Economic Journal: Microeconomics*, 7, 158–87.

AMBRUS, A., E. M. AZEVEDO, AND Y. KAMADA (2013): "Hierarchical cheap talk," *Theoretical Economics*, 8, 233–261.

AMBRUS, A. AND S. E. LU (2014): "Almost fully revealing cheap talk with imperfectly informed senders," *Games and Economic Behavior*, 88, 174–189.

AUSTEN-SMITH, D. (1993): "Interested experts and policy advice: Multiple referrals under open rule," *Games and Economic Behavior*, 5, 3–43.

BATTAGLINI, M. (2002): "Multiple referrals and multidimensional cheap talk," *Econometrica*, 70, 1379–1401.

——— (2004): "Policy advice with imperfectly informed experts," *Advances in theoretical Economics*, 4.

BLUME, A., O. J. BOARD, AND K. KAWAMURA (2007): "Noisy talk," *Theoretical Economics*, 2, 395–440.

BOLTON, P. AND M. DEWATRIPONT (1994): "The firm as a communication network," *The Quarterly Journal of Economics*, 109, 809–839.

CRAWFORD, V. P. AND J. SOBEL (1982): "Strategic information transmission," *Econometrica: Journal of the Econometric Society*, 1431–1451.

CROZIER, M. (2009): *The bureaucratic phenomenon*, vol. 280, Transaction Publishers.

Cyert, R. M., J. G. March, et al. (1963): "A behavioral theory of the firm," *Englewood Cliffs, NJ*, 2, 169–187.

Dalton, M. (2017): *Men who manage: Fusions of feeling and theory in administration*, Routledge.

Dessein, W. (2002): "Authority and communication in organizations," *The Review of Economic Studies*, 69, 811–838.

Friebel, G. and M. Raith (2004): "Abuse of authority and hierarchical communication," *RAND Journal of Economics*, 224–244.

Galeotti, A., C. Ghiglino, and F. Squintani (2013): "Strategic information transmission networks," *Journal of Economic Theory*, 148, 1751–1769.

Garicano, L. (2000): "Hierarchies and the Organization of Knowledge in Production," *Journal of political economy*, 108, 874–904.

Gibbons, R., N. Matouschek, and J. Roberts (2013): "Decisions in organizations," *The handbook of organizational economics*, 373–431.

Gibbons, R. and J. Roberts (2013): "Economic theories of incentives in organizations," *The handbook of organizational economics*, 56–99.

Glaser, L., S. P. Fourné, and T. Elfring (2015): "Achieving strategic renewal: The multi-level influences of top and middle managers' boundary-spanning," *Small Business Economics*, 45, 305–327.

Grenadier, S. R., A. Malenko, and N. Malenko (2016): "Timing decisions in organizations: Communication and authority in a dynamic environment," *American Economic Review*, 106, 2552–81.

Hagenbach, J. and F. Koessler (2010): "Strategic communication networks," *The Review of Economic Studies*, 77, 1072–1099.

Ivanov, M. (2010): "Communication via a strategic mediator," *J. Economic Theory*, 145, 869–884.

Jackson, M. O. (2010): *Social and economic networks*, Princeton university press.

Krishna, V. and J. Morgan (2001a): "Asymmetric information and legislative rules: Some amendments," *American Political Science Review*, 95, 435–452.

——— (2001b): "A model of expertise," *The Quarterly Journal of Economics*, 116, 747–775.

Milgrom, P. and J. Roberts (1988): "An economic approach to influence activities in organizations," *American Journal of sociology*, 94, S154–S179.

Morgan, J. and P. C. Stocken (2008): "Information aggregation in polls," *American Economic Review*, 98, 864–96.

Myerson, R. (1991): "Game theory: analysis of conflict," *The President and Fellows of Harvard College, USA*.

Mylovanov, T. and A. Zapechelnyuk (2013): "Decision rules revealing commonly known events," *Economics Letters*, 119, 8–10.

Oberg, A. and P. Walgenbach (2009): "Hierarchical structures of communication in a network organization," in *Methoden in der Betriebswirtschaftslehre*, Springer, 229–259.

Radner, R. (1993): "The organization of decentralized information processing," *Econometrica: Journal of the Econometric Society*, 1109–1146.

Rantakari, H. (2008): "Governing adaptation," *The Review of Economic Studies*, 75, 1257–1285.

Sah, R. K. and J. E. Stiglitz (1986): "The architecture of economic systems: Hierarchies and polyarchies," *The American Economic Review*, 716–727.

Schilling, M. A. and C. Fang (2014): "When hubs forget, lie, and play favorites: Interpersonal network structure, information distortion, and organizational learning," *Strategic Management Journal*, 35, 974–994.

Tushman, M. L. and T. J. Scanlan (1981): "Boundary spanning individuals: Their role in information transfer and their antecedents," *Academy of management journal*, 24, 289–305.

Van Zandt, T. (1999): "Decentralized information processing in the theory of organizations," in *Contemporary economic issues*, Springer, 125–160.

Van Zandt, T. and R. Radner (2001): "Real-time decentralized information processing and returns to scale," *Economic Theory*, 17, 545–575.

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

8 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 cell $p := \{1,2\}$, 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:

Denote by N the set of all players, $N := N^e \cup DM$. 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 $j' \in N$, $j' \neq j$ with $e_{ij'} = 1$. 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.

For a given n, the probability of $k \in \{0, .., n\}$ is:

$$Pr(k) = \int_0^1 Pr(k|\theta) f(\theta) d\theta = \frac{n!}{k!(n-k)!} \theta^{k+\alpha-1} (1-\theta)^{n-k+\beta-1} \frac{(\alpha+\beta-1)!}{(\alpha-1)!(\beta-1)!} =$$

$$\frac{n!}{k!(n-k)!} \frac{(\alpha+k-1)!(\beta-k+n-1)!}{(\alpha+\beta+n-1)!} \frac{(\alpha+\beta-1)!}{(\alpha-1)!(\beta-1)!}$$

Finally, when a partition P is adopted, the expected value of θ for a communicated $p \in P$ is

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

where $Pr(p|\theta) = \sum_{k \in p} Pr(k|\theta) Pr(k|p)$ where $Pr(k|p)$ reflects the fact that the elements of $\{0, 1\}^n$ with the same sum of the signals can be featured in different cells of a partition. Thus:

$$Pr(\theta|p, n) = \frac{\sum_{k \in p} \frac{n!}{k!(n-k)!} \theta^k (1-\theta)^{n-k} Pr(k|p) \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{a-1} (1-\theta)^{b-1}}{\int_0^1 \sum_{k \in p} \frac{n!}{k!(n-k)!} \theta^k (1-\theta)^{n-k} Pr(k|p) \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{a-1} (1-\theta)^{b-1} d\theta}$$

Proof of Proposition 1: Fix any (α, β). Denote a star network 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 (the sum of the signals) $\{\{0\}, \{1\}, .., \{n'\}\}$. For $k \in \{0, .., n'\}$, the sequentially rational decision maker chooses

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

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

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

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

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

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

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

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

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

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

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

which can be rewritten as

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

The above inequality holds for

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

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

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

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

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

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

which can be rewritten as:

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

The above inequality holds for

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

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

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

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

The expected payoff of the decision maker is

$$-\sum_{k=0}^{n'} Pr(k|n')\int_0^1\Big(\frac{\alpha+k}{\alpha+\beta+n'}-\theta\Big)^2\frac{\Gamma(\alpha+\beta+n')}{\Gamma(\alpha+k)\Gamma(\beta+n-k)}\theta^{\alpha+k-1}(1-\theta)^{\beta+n'-k-1}d\theta=$$

$$-\frac{\alpha\beta}{(\alpha+\beta)(1+\alpha+\beta)(\alpha+\beta+n')}.$$

Q.E.D.

Proof of Proposition 2: Fix any (α,β) 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. Fix 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 $|b_i|\leq\frac{1}{2(\alpha+\beta+n')}$ for each $i\in N'$.

Fix any tree network $Q\in\mathbb{Q}$ that is not a star with the set N' of experts such that the experts are ordered monotonically according to the absolute value of their biases: if expert i reports to expert j, then $|b_i|\geq|b_j|$. If we show that Q generates the same outcome as the star, so does any $Q'\in\mathbb{Q}$ with n experts which is not a star and where the experts are ordered optimally. This is because the babbling experts at the "bottom" of the hierarchy can always be ignored.

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 each expert 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 $y(k, n') = \frac{\alpha+k}{\alpha+\beta+n'}$, 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 $T_j(Q) = \{0, 1, .., \tilde{n}\}$.[19] Expert j does not observe the signals of $n - \tilde{n}$ experts. The type set of those experts, denoted by $T_{-j}(Q)$, can be expressed in terms of the summary statistic as $T_{-j}(Q) = \{0, .., n - \tilde{n}\}$.

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

$$\sum_{l=0}^{n'-\tilde{n}} Pr(t_{-j} = l | t_j = k') \Big(y(k' + 1 + l, n') - y(k' + l, n') \Big)$$

$$\Big(y(k' + 1 + l, n') + y(k' + l, n') - 2\mathbb{E}(\theta | k' + l, n') - 2b_j) \Big) \geq 0$$

Using the decision maker's best response we have

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

and

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

Since $\mathbb{E}(\cdot) \geq 0$, the incentive constraint implies

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

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

$$\sum_{l=0}^{n'-\tilde{n}} Pr(t_{-j} = l | t_j = k') \Big(y(k' - 1 + l, n') - y(k' + l, n') \Big)$$

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

$$\left(y(k'-1+l,n')+y(k'+l,n')-2\mathbb{E}(\theta|k'+l,n')-2b_j)\right)\geq 0$$

Since

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

and

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

$$-\frac{1}{\alpha+\beta+n'}.$$

we derive

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

which proves the Proposition.

Q.E.D.

Proof of Proposition 3: Denote the expert connected to the decision maker by 2, and fix the following strategy profile: expert 1 reveals his signal to expert 2, and expert 2 sends one of the two messages in equilibrium - either he informs the decision maker that the sum of both signals is 0 (message p_2), or that the sum of the signals is either 1 or 2 (message p_2').

After receiving p_2 the decision maker's posterior is

$$f(\theta|0,2)=\frac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha)\Gamma(\beta+2)}\theta^{\alpha-1}(1-\theta)^{\beta+1},$$

and therefore she chooses the optimal action

$$y(p_2)=E(\theta|0,2)=\int_0^1\theta f(\theta|0,2)d\theta=\frac{\alpha}{\alpha+\beta+2}.$$

After receiving p_2' the decision maker's posterior distribution is

$$f(\theta|p_2',2)=\frac{\sum_{k=1}^{2}\frac{2!}{k!(2-k)!}\theta^k(1-\theta)^{2-k}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}}{\int_0^1\sum_{k=1}^{2}\frac{2!}{k!(2-k)!}\theta^k(1-\theta)^{2-k}\frac{(\alpha+\beta-1)!}{(\alpha-1)!(\beta-1)!}\theta^{\alpha-1}(1-\theta)^{\beta-1}d\theta}=$$

$$\frac{\theta^{\alpha}(1-\theta)^{\beta-1}(2-\theta)\Gamma(\alpha+\beta+2)}{(2\beta+\alpha+1)\Gamma(\alpha+1)\Gamma(\beta)}$$

so that she chooses the action

$$y(p_2') = E(\theta|p_2', 2) = \int_0^1 \theta f(\theta|p_2', 2) d\theta = \frac{(\alpha + 2\beta + 2)(\alpha + 1)}{(\alpha + 2\beta + 1)(\alpha + \beta + 2)}$$

each expert conditions upward deviation on the same event that both signals are 0. The incentive constraint of an expert $i \in \{1, 2\}$ is therefore:

$$-\int_0^1 \Big(y(p_2) - \theta - b_i\Big)^2 f(\theta|0, 2) d\theta \geq -\int_0^1 \Big(y(p_2') - \theta - b_i\Big)^2 f(\theta|0, 2) d\theta$$

which implies:

$$b_i \leq \frac{1}{2}\Big(y(p_2) + y(p_2') - 2E(\theta|0, 2)\Big) = \frac{1 + \alpha + \beta}{(2 + \alpha + \beta)(1 + \alpha + 2\beta)}.$$

In a similar way we obtain the downward deviation constraint: for both experts the downward deviation has to be conditioned on the sum of the signals $k = 1$. Thus, an expert $i \in \{1, 2\}$ does not deviate downward for

$$-\int_0^1 \Big(y(p_2') - \theta - b_i\Big)^2 f(\theta|1, 2) d\theta \geq -\int_0^1 \Big(y(p_2) - \theta - b_i\Big)^2 f(\theta|1, 2) d\theta$$

which yields:

$$b_i \geq -\frac{\beta}{(2 + \alpha + \beta)(1 + \alpha + 2\beta)}.$$

To calculate the expected payoff of the decision maker notice that $Pr(k = 0) = \frac{\beta(\beta+1)}{(\alpha+\beta)(\alpha+\beta+1)}$. Therefore, the expected payoff is:

$$-\frac{\beta(\beta + 1)}{(\alpha + \beta)(\alpha + \beta + 1)} \int_0^1 (y(p_2) - \theta)^2 f(\theta|p_2, 2) d\theta -$$

$$\Big(1 - \frac{\beta(\beta + 1)}{(\alpha + \beta)(\alpha + \beta + 1)}\Big) \int_0^1 (y(p_2') - \theta)^2 f(\theta|p_2', 2) d\theta =$$

$$-\frac{\alpha\beta(4 + \alpha^2 + \beta(5 + 2\beta) + \alpha(5 + 3\beta))}{(\alpha + \beta)(1 + \alpha + \beta)(2 + \alpha + \beta)^2(1 + \alpha + 2\beta)}.$$

Given Proposition 1 we know that without any signals, the expected payoff of the decision maker is:

$$-\frac{\alpha\beta}{(\alpha + \beta)^2(1 + \alpha + \beta)}.$$

Since

$$-\frac{\alpha\beta(4+\alpha^2+\beta(5+2\beta)+\alpha(5+3\beta))}{(\alpha+\beta)(1+\alpha+\beta)(2+\alpha+\beta)^2(1+\alpha+2\beta)}-\Big(-\frac{\alpha\beta}{(\alpha+\beta)^2(1+\alpha+\beta)}\Big)=$$

$$\frac{4\alpha\beta(1+\beta)}{(\alpha+\beta)^2(2+\alpha+\beta)^2(1+\alpha+2\beta)}>0,$$

the decision maker always benefits from receiving information according to the coarse partition above compared to no information.

Finally, notice that the above message profile can support larger biases compared to the largest bias for any non-babbling equilibrium in a star since

$$\frac{1+\alpha+\beta}{(2+\alpha+\beta)(1+\alpha+2\beta)}>\frac{1}{2(1+\alpha+\beta)}$$

which holds for

$$\alpha\geq\frac{1}{2}(\sqrt{1+\beta(6+\beta)}-1-\beta).$$

For the negative biases, the equilibrium strategy profile is as follows: expert 1 reveals his signal to expert 2, and expert 2 sends one of the two messages in equilibrium - either he informs the decision maker that the sum of the signals is 2 (message $\hat{p}_2$) or that the sum of the signals is either 0 or 1 (message $\hat{p}'_2$).

After receiving $\hat{p}_2$ the decision maker's posterior is:

$$f(\theta|\hat{p}_2,2)=\frac{\sum_{k=0}^{1}\frac{2!}{k!(2-k)!}\theta^k(1-\theta)^{2-k}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}}{\int_0^1\sum_{k=0}^{1}\frac{2!}{k!(2-k)!}\theta^k(1-\theta)^{2-k}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}d\theta}=$$

$$\frac{(\alpha+\beta)(1+\alpha+\beta)(1+\theta)\theta^{\alpha-1}(1-\theta)^{\beta}\Gamma(\alpha+\beta)}{\beta(1+2\alpha+\beta)\Gamma(\alpha)\Gamma(\beta)}$$

and the expected value is

$$\int_0^1\theta f(\theta|\hat{p}_2,2)d\theta=\frac{\alpha(3+2\alpha+\beta)}{(2+\alpha+\beta)(1+2\alpha+\beta)}$$

Conditional on $\hat{p}'_2$, the decision maker's expected value is $\frac{\alpha+2}{\alpha+\beta+2}$.

The corresponding upward deviation by the experts is conditioned on the event in which the sum of both signals is 1. The corresponding deviation constraint is:

$$-\int_0^1 \Big(y(\hat{p}_2') - \theta - b_i\Big)^2 f(\theta|1,2)d\theta \geq -\int_0^1 \Big(y(\hat{p}_2) - \theta - b_i\Big)^2 f(\theta|1,2)d\theta$$

which implies

$$b_{1,2} \leq \frac{\alpha}{(2+\alpha+\beta)(1+2\alpha+\beta)}$$

The downward deviation for both experts is conditioned on the event in which both signals are 1 ($k = 2$):

$$-\int_0^1 \Big(y(\hat{p}_2) - \theta - b_i\Big)^2 f(\theta|2,2)d\theta \geq -\int_0^1 \Big(y(\hat{p}_2') - \theta - b_i\Big)^2 f(\theta|2,2)d\theta$$

which implies

$$b_{1,2} \geq -\frac{1+\alpha+\beta}{(2+\alpha+\beta)(1+2\alpha+\beta)}.$$

It turns out that

$$-\frac{1+\alpha+\beta}{(2+\alpha+\beta)(1+2\alpha+\beta)} < -\frac{1}{2(\alpha+\beta+1)}$$

for

$$\beta > \frac{1}{2}(\sqrt{1+\alpha(6+\alpha)} - 1 - \alpha).$$

To calculate the expected payoff notice that $Pr(2) = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}$, and therefore the expected value for the decision maker is:

$$-\frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} \int_0^1 (y(\hat{p}_2') - \theta)^2 f(\theta|\hat{p}_2', 2)d\theta -$$

$$\Big(1 - \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}\Big) \int_0^1 (y(\hat{p}_2) - \theta)^2 f(\theta|\hat{p}_2, 2)d\theta =$$

$$-\frac{\alpha\beta(4+\beta^2+\alpha(5+2\alpha)+\beta(5+3\alpha))}{(\alpha+\beta)(1+\alpha+\beta)(2+\alpha+\beta)^2(1+2\alpha+\beta)}.$$

Q.E.D.

Proof of Proposition 4: Denote the expert connected to the decision maker by 1, and fix the following strategy profile: two experts connected to expert 1 reveal their signals, and expert 1 sends one of the two messages in equilibrium - either he informs the decision maker that the sum of both signals is 0 (message p_1), or that the sum of the signals is either 1, 2 or 3 (message p_1').

After receiving p_1 the decision maker's posterior is:

$$f(\theta|p_1,3)=\frac{\Gamma(\alpha+\beta+3)}{\Gamma(\alpha)\Gamma(\beta+3)}\theta^{\alpha-1}(1-\theta)^{\beta+2},$$

and so the decision maker chooses the action

$$y(p_1)=\mathbb{E}_{DM}(\theta|p_1,2)=\int_0^1\theta f(\theta|p_1,n=3,\alpha,\beta)d\theta=\frac{\alpha}{\alpha+\beta+3}.$$

After receiving p_2' the decision maker's posterior distribution is

$$f(\theta|p_1',3)=\frac{\sum_{k=1}^{3}\frac{3!}{k!(3-k)!}\theta^k(1-\theta)^{3-k}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}}{\int_0^1\sum_{k=1}^{3}\frac{3!}{k!(3-k)!}\theta^k(1-\theta)^{3-k}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}d\theta}=$$

$$\frac{((\frac{1}{1-\theta})^3-1)\theta^{\alpha-1}(1-\theta)^{\beta+2}}{\Gamma(\alpha)\left(\frac{\Gamma(\beta)}{\Gamma(\alpha+\beta)}-\frac{\Gamma(\beta+3)}{\Gamma(\alpha+\beta+3)}\right)}$$

so that she chooses the action

$$y(p_1')=E(\theta|p_1',3)=\int_0^1\theta f(\theta|p_1',3)d\theta=$$

$$\frac{(1+\alpha)(\alpha^2+3(1+\beta)(2+\beta)+\alpha(5+3\beta)}{(3+\alpha+\beta)(2+\alpha^2+3\alpha(1+\beta)+3\beta(2+\beta)}$$

For an upward deviation, the binding constraint (for any of the three communicating experts) has to be conditioned on the sum of three signals $k=0$. An expert $i\in\{1,2,3\}$ does not deviate if

$$-\int_0^1\left(y(p_1)-\theta-b_i\right)^2f(\theta|0,3)d\theta\geq-\int_0^1\left(y(p_1')-\theta-b_i\right)^2f(\theta|0,3)d\theta$$

which implies:

$$b_i\leq\frac{3(1+\alpha+\beta)(2+\alpha+\beta)}{2(3+\alpha+\beta)(2+\alpha^2+3\alpha(1+\beta)+3\beta(2+\beta))}$$

Notice that this bound is strictly higher than the upper bound for the biases in case of a single truthfuly revealed signal. For $\alpha \geq 1$ it is easy to show that

$$\frac{3(1+\alpha+\beta)(2+\alpha+\beta)}{2(3+\alpha+\beta)(2+\alpha^2+3\alpha(1+\beta)+3\beta(2+\beta))} > \frac{1+\alpha+\beta}{(2+\alpha+\beta)(1+\alpha+2\beta)}.$$

In terms of the downward deviation, the binding constraint is conditioned on the case in which the sum of the signals is $k=1$. An expert $i \in \{1,2,3\}$ does not deviate downwards if

$$-\int_0^1 \Big(y(p_1') - \theta - b_i\Big)^2 f(\theta|1,3)d\theta \geq -\int_0^1 \Big(y(p_1) - \theta - b_i\Big)^2 f(\theta|1,3)d\theta$$

which implies that

$$b_i \geq \frac{2+\alpha(3+\alpha)-3\beta(1+\beta)}{2(3+\alpha+\beta)(2+\alpha^2+3\alpha(1+\beta)+3\beta(2+\beta))}$$

which is strictly smaller than $\frac{1+\alpha+\beta}{(2+\alpha+\beta)(1+\alpha+2\beta)}$.

To calculate the expected payoff of the decision maker notice that $Pr(k=0) = \frac{\beta(\beta+1)(\beta+2))}{(\alpha+\beta)(\alpha+\beta+1)(\alpha+\beta+2)}$ so that the expected payoff for the decision maker for the above equilibrium is

$$-\frac{\beta(\beta+1)(\beta+2)}{(\alpha+\beta)(\alpha+\beta+1)(\alpha+\beta+2)}\int_0^1 (y(p_1)-\theta)^2 f(\theta|p_1,3)d\theta -$$

$$\Big(1-\frac{\beta(\beta+1)(\beta+2)}{(\alpha+\beta)(\alpha+\beta+1)(\alpha+\beta+2)}\Big)\int_0^1 (y(p_1')-\theta)^2 f(\theta|p_1',3)d\theta =$$

$$-\frac{\alpha\beta(31+\alpha^3+\alpha^2(9+4\beta)+\beta(29+3\beta(9+2\beta)))}{(\alpha+\beta)(1+\alpha+\beta)(3+\alpha+\beta)^2(2+\alpha^2+3\alpha(1+\beta)+3\beta(2+\beta))}.$$

This equilibrium yields a strictly higher payoff to the decision maker compared to a babbling equilibrium in a star since

$$-\frac{\alpha\beta(31+\alpha^3+\alpha^2(9+4\beta)+\beta(29+3\beta(9+2\beta)))}{(\alpha+\beta)(1+\alpha+\beta)(3+\alpha+\beta)^2(2+\alpha^2+3\alpha(1+\beta)+3\beta(2+\beta))} -$$

$$\Big(-\frac{\alpha\beta}{(\alpha+\beta)^2(1+\alpha+\beta)}\Big) =$$

$$\frac{9\alpha\beta(2+3\beta+\beta^2)}{(\alpha+\beta)^2(3+\alpha+\beta)^2(2+\alpha^2+6\beta+3\beta^2+3\alpha(1+\beta))} > 0.$$

The case for the negative biases is calculated analogously for an equilibrium in which expert 1 receives truthful information from the other two experts and sends one of the two messages in equilibrium: he either informs the decision maker that the sum of the signals is within the set $\{0,1,2\}$ or that the sum of the signals is $k=3$.

Q.E.D.

Proof of Corollary 1: Given Proposition 2, every optimally designed hierarchy is able to generate same equilibria as a star. If Q is not a star, there exists $i \in N^e$ with $e_{iDM} = 1$ such that there is $j \in N^e$, $j \neq i$ with $e_{ji} = 1$. Therefore, given that either $\alpha > \frac{1}{2}(\sqrt{1+\beta(\beta+6)} - 1 - \beta)$ or $\beta > \frac{1}{2}(\sqrt{1+\alpha(\alpha+6)} - 1 - \alpha)$ is true, there exists a range of biases for which an equilibrium from Proposition 3 can be implemented in Q and therefore Q strictly dominates the star.

Q.E.D.

Proof of Proposition 5: Consider positively biased experts and a network Q with only one expert $j \in N^e$ communicating directly to the decision maker, $e_{jP} = 1$, while all other experts have directed links to expert j. Think of a strategy profile in Q where each $i \neq j,\ i \in N^e$, reveals his signal and so expert j observes all n signals. Suppose that j sends one of two messages - either he reveals that the sum of all signals is 0 (m), or that the sum of all signals is not 0 (m'). The decision maker's optimal actions are

$$y(m) = \frac{1}{n+2},\ y(m') = \frac{1}{n}\sum_{k=1}^{n}\frac{k+1}{n+2} = \frac{3+n}{4+2n}.$$

The incentive constraint of the experts who condition their deviation on all signals being 0 is

$$\frac{1}{n+2} + \frac{3+n}{4+2n} - 2\frac{1}{n+2} - 2b_i = 0,\ i = 1,..,n$$

which results in

$$b_i \leq \frac{1+n}{4(2+n)},$$

which is increasing in n and has a minimal value for $n = 2$ which is $\frac{3}{16}$. On the other hand, the downward deviation implies

$$\frac{1}{n+2} + \frac{3+n}{4+2n} - 2\frac{2}{n+2} - 2b_i = 0,\ i = 1,..,n$$

which results in

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

Next, I show that for any number n of signals, among all partitions of $\{0,1\}^n$, the partition

$\{\{0\}, \{1, .., n\}\}$ is supported by the largest biases.

First, I show that the largest biases are supported by a partition of $\{0,1\}^n$ that consists of two cells. Suppose not, and consider a partition of $\{0,1\}^n$ of the form $P' = \{p_1, p_2, .., p_k\}$ with $k \geq 3$ where p_i, $i = 1, .., k$, denotes a cell of the partition P'. Suppose that this partition is implementable in some network Q with $P_{DM}(Q) = P'$ and take some $i \in N^e$ participating in non-babbling communication. Fix a type of this expert i, t_i, such that there exist two different cells $p, p' \in P'$ with $p \in supp(\Delta(P')|t_i)$ and $E(\theta|p) < E(\theta|p')$. Denote the largest bias that prevents t_i's deviation from p to p' by $\bar{b}$. We can construct a (possibly) different partition $\hat{P} \neq P'$ that: (a) contains a cell $\hat{p}$ with the property that $k = 0 \in \hat{p}$, (b) has another cell denoted by $\hat{p}'$ where $max_k \in \hat{p}' := k' \leq n$ and (c) there is an element of $\Delta(\hat{p})$ such that if an expert had beliefs corresponding to this probability distribution, then the upper bias preventing deviation from $\hat{p}$ to $\hat{p}'$ is exactly $\bar{b}$. But since $max_k \in \hat{p}' := k' \leq n$, we can always include signal realizations $k \geq k'$ into the cell $\hat{p}'$ (if $k' < n$) such that the deviation of expert i from $\hat{p}$ to $\hat{p}'$ for a fixed $\Delta(\hat{p})$ results in an upper bound for the bias $\bar{b}_i^{'} \geq \bar{b}_i$. Thus, we conclude that the largest upper bias among all partitions of $\{0,1\}^n$ is generated by a partition with two cells.

Next, I show that this partition has the form $\{\{0\}, \{1, .., n\}\}$. Take any partition of $\{0,1\}^n$ which consists of two cells and is denoted by $P = \{p_1, p_2\}$. Denote the *highest* sum of the signals in the first cell by k_1 and denote the *lowest* sum of the signals in the second cell by k_2. It has to be the case that $k_1 \geq k_2$ (otherwise P is not a partition of $\{0,1\}^n$). First, assume $k_1 > k_2$. Fix any conditional belief of any expert i on p_1, denoted by $\tilde{t}_i$, with $supp(\Delta(P)|\tilde{t}_i) \cap supp(p_1) \neq \emptyset$. Now, take another partition with the difference to the previous partition such that all $k \in [k_1, k_2]$ are now elements of the first cell. For this new partition, use the notation $\tilde{P} = \{\tilde{p}_1, \tilde{p}_2\}$. Since

$$y(p_1) < y(\tilde{p}_1) \text{ and } y(p_2) < y(\tilde{p}_2),$$

for the given conditional belief $\tilde{t}_i$ the resulting upward deviation in $\tilde{P}$ results in a higher upper bound for the biases compared to a corresponding deviation in P. Per construction, $\tilde{P}$ is a partition of the *sum* of n signals.

For the next step notice, first, that multiple types of an expert can send messages which result in the decision maker being informed that the relevant cell is $\tilde{p}_1$. The binding constraint for an upward deviation uses a type of an expert with the highest posterior over θ among all types which result in the decision maker being informed that the relevant cell is $\tilde{p}_1$. Notice that the posterior of this type cannot be lower than $E_{DM}(\theta|\tilde{p}_1)$. I show next that for a two-cell partition of the type $\tilde{P} = \{\tilde{p}_1, \tilde{p}_2\}$ (to remind, this is a partition of the sum of the signals), the highest possible bias is attained if the posterior of the expert is exactly $\tilde{p}_1$. The upper bound for the expert's bias is determined by

$$y(\tilde{p}_1) + y(\tilde{p}_2) - 2E(\theta|t_i) - 2b_i = 0$$

for some type t_i of expert i. But since the posterior of expert's type cannot be lower than $y(\tilde{p}_1)$, for any given $y(\tilde{p}_1)$ and $y(\tilde{p}_2)$ the largest possible bias is generated if when the experts' posterior is exactly at the lower bound, it means if $E(\theta|t_i) = y(\tilde{p}_1)$. Which $y(\tilde{p}_1)$ and $y(\tilde{p}_2)$ generate the largest possible bias given that $E(\theta|t_i) = y(\tilde{p}_1)$? To remind, k_1 denotes the highest sum of the signals included into $\tilde{p}_1$. Sequential rationality of the decision maker dictates that

$$y(\tilde{p}_1) = \frac{\frac{1}{n+1}\sum_{i=0}^{k_1}\frac{i+1}{n+2}}{\frac{k_1+1}{n+1}} = \frac{2+k}{4+2n} \text{ and } y(\tilde{p}_2) = \frac{\frac{1}{n+1}\sum_{i=0}^{k_1}\frac{i+1}{n+2}}{\frac{n-k}{n+1}} = \frac{3+k_1+n}{4+2n},$$

and so the upper bound for the bias, $\bar{b}$, is determined by

$$\frac{3+k_1+n}{4+2n} + \frac{2+k_1}{4+2n} - 2\Big(\frac{2+k_1}{4+2n}\Big) - 2\bar{b} = 0,$$

which is independent of k_1! Thus, the only relevant constraint for a two-cell partition of the sum of the signals is that the expert's type that determines the binding upward deviation constraint, denoted by t_i', is such that $E(\theta|t_i') = E(\theta|y(\tilde{p}_1))$. Since the partition in question consists of two cells, the decision maker has to receive information only from a single expert. But then, $E(\theta|t_i') = E(\theta|y(\tilde{p}_1))$ is only possible if $y(\tilde{p}_1) = 0$. Thus, the partition that results in the largest possible experts' bias given n signals is of the form

$$\{\{0\}, \{1, .., n\}\}.$$

Now, for $n-1$ signals, the partition of $\{0,1\}^{n-1}$ that supports the largest possible bias for each expert is $\{\{0\}, \{1, .., n-1\}\}$ resulting in the bias range

$$\Big[\frac{n-4}{4(n+1)}, \frac{n}{4(n+1)}\Big].$$

Clearly, any partition of $\{0,1\}^{n'}$ with $n' < n-1$ cannot be supported by biases larger than $\frac{n}{4(n+1)}$. Thus, a network characterized in the proposition can implement information transmission according to the partition $\{\{0\}, \{1, .., n\}\}$ for the biases in the range

$$\Big(\frac{n}{4(n+1)}, \frac{n+1}{4(n+2)}\Big].$$

Moreover, no information transmission is possible for biases strictly above $\frac{n+1}{4(n+2)}$. The proof for negative biases is symmetric.

Q.E.D.

Proof of Corollary 2: According to Proposition 1, whenever the absolute values of all experts' biases are strictly above $\frac{1}{6}$, there is no informative equilibrium without informational pooling. According to the proof construction in Proposition 2 and the statement of Proposition 3, the equilibrium with the largest pooling features either an equilibrium partition $P = \{\{0\}, \{1, .., n\}\}$ or $P' = \{\{0, 1, .., n-1\}, \{n\}\}$. Suppose there are $m \geq 2$ experts $j_1, .., j_m$ with the biases strictly above $\frac{1}{6}$. Then, P is implementable whenever for each $i \in \{1, .., m\}$, $b_i \leq \frac{n+1}{4(n+2)}$. Suppose there are $m' \geq 2$ experts with the biases strictly below $\frac{1}{6}$. Then, P' is implementable whenever for each $j \in \{1, .., m'\}$, $b_j \geq -\frac{n+1}{4(n+2)}$.

Q.E.D.

Proof of Proposition 6: Consider, first, the partition $P_1 = \{\{0\}, \{1\}, .., \{n-2\}, \{n-1, n\}\}$. The corresponding expected utility of the decision maker is

$$-\frac{1}{3} + \frac{1}{(n+1)(n+2)^2}\Big(\sum_{i=1}^{n-1} i^2 + \frac{(n+(n+1))^2}{2}\Big) = -\frac{n^2+3n+5}{6(n+1)(n+2)^2}$$

where I use the fact that $y(n-1, n) = \frac{2n+1}{2(n+2)}$. The implementation of the partition $P_{DM} = P_1$ requires a strategy profile where a single expert i aggregates all signals and communicates directly to the decision maker. First, it is easy to show that the corresponding optimal network requires that all experts communicate with the expert i directly. This is because in any other network that can potentially implement the above partition, their incentive constraints will be tighter. The bias constraints of every other expert $j = 1, ., , n$, $j \neq i$, in such a network featuring a single intermediary satisfy the inequality

$$\sum_{k=0}^{n-3} \frac{1}{n+1} \frac{\binom{n-1}{k}}{\binom{n}{k}} \Big(\frac{k+1}{n+2} - \frac{k}{n+2}\Big)\Big(\frac{k+1}{n+2} - \frac{k}{n+2} - 2b_j\Big) +$$

$$\frac{1}{n+1} \frac{\binom{n-1}{n-2}}{\binom{n}{n-2}} \Big(\frac{2n+1}{2(n+2)} - \frac{n-1}{n+2}\Big)\Big(\frac{2n+1}{2(n+2)} - \frac{n-1}{n+2} - 2b_j\Big) \geq 0$$

where I use the fact that $y(n-1, n) = \frac{2n+1}{2(n+2)}$. This inequality implies

$$b_j \leq \frac{3 + n(n+1)}{2n(2+3n+n^2)}.$$

Next, I show that the above partition maximizes the decision maker's expected payoff and is incentive compatible given the bias requirements stated in the proposition. For this, notice that the partition according to which the decision maker receives her information can – depending on network – be a product of multiple partitions. Suppose that she receives information according to

$$P_2 = \{\{0\}, .., \{n-3\}, \{n-2, n-1\}\} \times \{\{0\}, \{1\}\}$$

To calculate the expected payoff notice that $y(\{n-2, n-1\} \times \{0\}) = \frac{3n-2}{3(n+2)}$, and $y(\{n-2, n-1\} \times \{1\}) = \frac{2n^2}{3n+2n^2-2}$. Thus, the expected utility is

$$-\frac{1}{3} + \frac{1}{n+1}\Big[\sum_{k=0}^{n-3} \frac{\binom{n-1}{k}}{\binom{n}{k}} \Big(\frac{k+1}{n+2}\Big)^2 + \sum_{k=0}^{n-3} \frac{\binom{n-1}{k}}{\binom{n}{k+1}} \Big(\frac{k+2}{n+2}\Big)^2 +$$

$$\frac{3}{n(n+1)}\Big(\frac{3n-2}{3(n+2)}\Big)^2 + \frac{2n-1}{n(n+1)}\Big(\frac{2n^2}{(n+2)(2n+1)}\Big)^2\Big] = -\frac{2n^3+3n^2+4n-4}{6n(n+2)^2(2n-1)}$$

The difference in the expected payoffs between P_1 and P_2 is

$$-\frac{n^2+3n+5}{6(n+1)(n+2)^2} - \Big(-\frac{2n^3+3n^2+4n-4}{6n(n+2)^2(2n-1)}\Big) = \frac{5n-4}{6n(n+2)^2\,(2n^2+n-1)} > 0$$

Now, consider the case in which the decision maker receives information according to a product of partitions where one of the subpartitions has $n-t$ signals, $t > 1$. Then, whenever there is a smallest coarsening with $n-t$ signals compatible with positive biases, and all other signals are truthfully communicated to the decision maker, the decision maker receives information according to the partition

$$P_t = \{\{0\}, .., \{n-t-2\}, \{n-t-1, n-t\}\} \times \{\{0\}, \{1\}, .., \{t\}\},\ t \geq 2.$$

We can rewrite P_t as

$$\{\{0\}, .., \{n-2\}, \{n-t-1, n-t\}, .., \{n-2, n-1\}, \{n-1, n\}\}.$$

However, P_t is dominated by a partition

$$P_t' = \{\{0\}, .., \{n-2\}, \{n-t-1, n-t\}, .., \{n-2\}, \{n-1, n\}\}$$

where the only difference of P_t' to P_t is that P_t' assigns the sum of the $n-1$ signals to the last cell.

To see why, denote by z the fraction of elements of $\{0,1\}^n$ with the sum of the signals $n-1$ which are inside the cell $\{n-2, n-1\}$ in P_t. Then,

$$y(n-2, n-1) = \frac{1}{1+z}\frac{n-1}{n+2} + \frac{z}{1+z}\frac{n}{n+2} = \frac{n+nz-1}{(n+2)(1+z)}$$

$$y(n-1, n) = \frac{1-z}{2-z}\frac{n}{n+2} + \frac{1}{2-z}\frac{n+1}{n+2} = \frac{n(2-z)+1}{(n+2)(2-z)}$$

Then, the expected payoff from $\{n-2, n-1\}, \{n-1, n\}$ is proportional to

$$(1+z)\Big(\frac{n+nz-1}{(n+2)(1+z)}\Big)^2 + (2-z)\Big(\frac{n(2-z)+1}{(n+2)(2-z)}\Big)^2 = \frac{2}{(n+2)^2}(n^2 + \frac{1}{2+z-z^2})$$

which is maximized either for $z=0$ or $z=1$. Thus, partition P_t is dominated by P'_t. It is easy to see that P'_t is dominated by P_1. Notice that any further coarsening of P_t results in even lower expected payoff for the decision maker. Therefore, whenever an expert i communicates according to P_1 and has $b_i \leq \frac{1}{2(n+2)}$, and every other expert $j \neq i,\ j = 1, .., n$, satisfies

$$\frac{1}{2(n+2)} < b_j \leq \frac{3+n(n+1)}{2n(2+3n+n^2)}.$$

the partition P_1 dominates any other partition which involves coarsening of information, it is incentive compatible and is implemented by a network which features i communicating directly to the decision maker according to P_1, and all other experts communicating their signals truthfully and directly to the expert i.

Finally, the best partition for $n-1$ signals is $\{\{0\}, .., \{n-1\}\}$ and yields the expected payoff for the decision maker $-\frac{1}{6(n+1)}$. The difference

$$-\frac{n^2+3n+5}{6(n+1)(n+2)^2} - \Big(-\frac{1}{6(n+1)}\Big) = \frac{n-1}{6(n+1)(n+2)^2} > 0 \text{ for } n > 1.$$

Q.E.D.

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.

$$b'_j = \operatorname*{argmin}_{b_j \in \mathbb{R}} \Big\{ -\sum_{\tilde{p}_i \in \tilde{P}_i(Q)} Pr(\tilde{p}_i|t_j) \sum_{p \in P_{DM}(Q)} \Pr(p|p_i(\tilde{p}_i, p_j), P_{-(i,j)}(Q)) \sum_{k \in p} Pr(k|t_i)$$

$$\int_0^1 (y(p) - \theta - b_j)^2 f(\theta|k, n) d\theta =$$

$$-\sum_{\tilde{p}_i \in \tilde{P}_i(Q)} Pr(\tilde{p}_i|t_j) \sum_{p \in P_{DM}(Q)} \Pr(p|p_i(\tilde{p}_i, p'_j), P_{-(i,j)}(Q)) \sum_{k \in p} Pr(k|t_i)$$

$$\int_0^1 (y(p) - \theta - b_j)^2 f(\theta|k, n) d\theta \ |\text{for all } t_j \in T_j(Q), t_j \in p_j, t_i \in T_i(Q), t_i \in p_i \Big\}. \tag{8.1}$$

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

$$b'_i = \operatorname*{argmin}_{b_i \in \mathbb{R}} \Big\{ \sum_{p \in P_{DM}(Q)} \Pr(p|p_i(\tilde{p}_i, \overline{p}_j), P_{-(i,j)}(Q)) \sum_{k \in p} Pr(k|t_i)$$

$$\int_0^1 (y(p) - \theta - b_i)^2 f(\theta|k, n) d\theta =$$

$$\sum_{p \in P_{DM}(Q)} \Pr(p|p'_i, P_{-(i,j)}(Q)) \sum_{k \in p} Pr(k|t_i)$$

$$\int_0^1 (y(p) - \theta - b_i)^2 f(\theta|k, n) d\theta \mid t_i \in T_i(Q), t_i \in p_i \Big\}, \tag{8.2}$$

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

$$b'_{j'} = \underset{b_i \in \mathbb{R}}{\text{argmin}} \Big\{ \sum_{p \in P_{DM}(Q')} \Pr(p|p_{j'}(t_{j'}), P_j(Q'), P_{-(i,j)}(Q')) \sum_{k \in p} Pr(k|t_i)$$

$$\int_0^1 (y(p) - \theta - b_{j'})^2 f(\theta|k, n) d\theta =$$

$$\sum_{p \in P_{DM}(Q')} \Pr(p|p'_i, P_{-(i,j)}(Q')) \sum_{k \in p} Pr(k|t_i)$$

$$\int_0^1 (y(p) - \theta - b'_{j'})^2 f(\theta|k, n) d\theta \ |\text{for all } t_{i'} \in T_{j'}(Q'), t_{j'} \in p_{j'} \Big\}. \tag{8.3}$$

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 $\bar{b}_{j'}$, is such that

$$\bar{b}_{j'} = \underset{b_i \in \mathbb{R}}{\text{argmin}} \Big\{ \sum_{p \in P_{DM}(Q')} \Pr(p|\hat{p}_{j'}(t_{j'}, p_j), P_{-(i,j)}(Q')) \sum_{k \in p} Pr(k|t_i)$$

$$\int_0^1 (y(p) - \theta - b_{j'})^2 f(\theta|k, n) d\theta =$$

$$\sum_{p \in P_{DM}(Q')} \Pr(p|\hat{p}'_{j'}, P_{-(i,j)}(Q')) \sum_{k \in p} Pr(k|t_i)$$

$$\int_0^1 (y(p) - \theta - b_{j'})^2 f(\theta|k, n) d\theta \ |\text{for all } t_{i'} \in T_{j'}(Q'), t_{j'} \in p_{j'} \Big\}, \tag{8.4}$$

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 $\bar{b}_{j'} \leq b'_{j'}$.

Q.E.D.

9 Appendix B: mixed strategies

9.1 Equilibrium definition

9.1.1 Equilibrium in the star

Define a collection of beliefs of expert i, β_i, by a probability distribution on the (Borel-measurable) subsets of $\{0,1\} \times [0,1]$ as a collection of conditional probability distributions $\beta_i(s_i)$ for $s_i \in \{0,1\}$. A collection of beliefs of the decision maker, β_{DM}, is defined by a probability distribution on the (Borel-measurable) subsets of $\{0,1\}^2 \times [0,1] \times M^2$ as a collection of conditional probability distributions $\beta_{DM}(m_1, m_2)$ for all $m_i \in M$, $i = 1, 2$.

A strategy of expert i is

$$\mathbf{m}_i : \; \{0,1\} \; \rightarrow \; \Delta(M)$$

where $m_i \in M$ and $\mathbf{m}_i(m_i|s_i)$ denotes the conditional probability that expert i chooses m_i given his signal s_i. A strategy of the decision maker is

$$y : \; M \times M \; \rightarrow \; \mathbb{R}.$$

A strategy profile $(\mathbf{m}_1, \mathbf{m}_2, y)$ and a belief system $(\beta_1, \beta_2, \beta_{DM})$ constitute a perfect Bayesian equilibrium if the following conditions are satisfied:

1. The decision maker chooses an optimal action given her beliefs

$$y \in \operatorname*{argmax}_{y' \in \mathbb{R}} \sum_{s_1} \sum_{s_2} Pr(s_1, s_2 | m_1, m_2) \int_0^1 (y(m_1, m_2) - \theta)^2 f(\theta | s_1, s_2) d\theta$$

2. Each expert chooses optimal actions given his beliefs:

$$m_i \in \arg\max_{m_i' \in M} \sum_{s_{-i}} Pr(s_{-i}|s_i) \sum_{m_{-i}} Pr(m_{-i}|s_{-i}) \int_0^1 (y(m_1, m_2) - \theta - b_i)^2 f(\theta|s_1, s_2) d\theta \;\forall i = 1, 2.$$

3. The beliefs are consistent with the actions meaning that for each $i = 1, 2$, β_i is a probability distribution on $\{0,1\} \times [0,1]$ generated by the prior distribution and the mapping from the signals to the state. The decision maker's belief β_{DM} is a conditional probability distribution over distributions on $\{0,1\}^2 \times [0,1] \times M^2$ generated by the prior, the mapping from the state to the signals and the strategies of both experts. Moreover, the beliefs have to be consistent between the players.

9.1.2 Equilibrium in the line

Denote by 1 the bottom expert and by 2 the intermediary. Define a belief of expert 1, β_1, by a probability distribution on the (Borel-measurable) subsets of $\{0,1\} \times [0,1]$ as a collection of conditional probability distributions $\beta_1(s_1)$ for $s_1 \in \{0,1\}$. Define a belief of expert 2, β_2, by a probability distribution on the (Borel-measurable) subsets of $\{0,1\} \times [0,1] \times M$ as a collection of conditional probability distributions $\beta_2(s_2, m_1)$ for $s_2 \in \{0,1\}$ and all $m_1 \in M$. A collection of beliefs of the decision maker, β_{DM}, is defined by a probability distribution on the (Borel-measurable) subsets of $\{0,1\}^2 \times [0,1] \times M^2$ as a collection of conditional probability distributions $\beta_{DM}(m_1, m_2)$ for all $m_i \in M$, $i = 1, 2$.

The strategies are

$$\mathbf{m}_1 : \{0,1\} \rightarrow \Delta(M),$$

$$\mathbf{m}_2 : \{0,1\} \times M \rightarrow \Delta(M).$$

A strategy of the decision maker is

$$y : M \rightarrow \mathbb{R}.$$

A strategy profile $(\mathbf{m}_1, \mathbf{m}_2, y)$ and a belief system $(\beta_1, \beta_2, \beta_{DM})$ constitutes a perfect Bayesian equilibrium if the following conditions are satisfied:

1. The decision maker chooses an optimal action given her beliefs

$$y \in \underset{y' \in \mathbb{R}}{\operatorname{argmax}} \sum_{s_1} \sum_{s_2} Pr(s_1, s_2|m_2) \int_0^1 (y(m_2) - \theta)^2 f(\theta|s_1, s_2) d\theta$$

2. Each expert chooses optimal actions given his beliefs:

$$m_1 \in \arg\max_{m_1' \in M} \sum_{s_2} Pr(s_2|s_1) \sum_{m_2} Pr(m_2|s_2) \int_0^1 (y(m_2) - \theta - b_1)^2 f(\theta|s_1, s_2) d\theta,$$

$$m_2 \in \arg\max_{m_2' \in M} \sum_{s_1} Pr(s_1|s_2, m_1) \int_0^1 (y(m_2) - \theta - b_2)^2 f(\theta|s_1, s_2) d\theta \ \forall i = 1, 2.$$

3. The beliefs are consistent with the actions meaning that β_1 is a probability distribution on $\{0,1\} \times [0,1]$ generated by the prior distribution and the mapping from the signals to the state, β_2 is a conditional probability distribution on the distributions over $\{0,1\} \times [0,1] \times M$ and β_{DM} is a conditional probability distribution over distributions of $\{0,1\}^2 \times [0,1] \times$

M generated by the prior, the mapping from the state to the signals and the strategies of both experts. Moreover, the beliefs have to be consistent between the players.

Proof of Lemma 3: Take two positively biased experts and denote the (mixed) strategy of expert i by $\mathbf{m}_i$, $i = 1, 2$. Denote by $M' \subseteq M$ the support of expert 1's mixed strategy conditional on $s_1 = 1$, it means: $\forall m \in M'$, $m \in supp(\mathbf{m}_1(s_1 = 1, \cdot))$. If for any $m' \in M'$ it is not true that $m' \in supp(\mathbf{m}_1(s_1 = 0, \cdot))$, then given m' the decision maker knows that $s_1 = 1$; so suppose that there are some messages within M' that are also used by expert 1 when playing a mixed strategy conditional on $s_1 = 0$. But then, given that each message profile of the experts results in a different posterior of the decision maker (otherwise the same message profiles can be pooled without changing the equilibrium allocations), for each $m_2 \in M$ there is $\hat{m} \in M'$ such that $E_{DM}(\theta|\hat{m}, m_2) > E_{DM}(\theta|\hat{m}', m_2)$ for any $\hat{m}' \in M'$ and $\hat{m}' \neq \hat{m}$. Moreover, if expert 1 with $s_1 = 1$ would observe m_2, his posterior would be $E_1(\theta|\hat{m}, m_2) > E_1(\theta|\hat{m}', m_2)$ for any $m_2 \in M$. But then it must be true that

$$\sum_{m_2 \in M} Pr(m_2|s_1 = 1)E_1(\theta|\hat{m}, m_2) > \sum_{m_2 \in M} Pr(m_2|s_1 = 1)E_1(\theta|\hat{m}', m_2)$$

and therefore given $s_1 = 1$, a positively biased expert 1 wants to deviate to $\hat{m}$. Thus, his mixed strategy is degenerate for $s_1 = 1$. Given the symmetry between the experts, the same is true whenever expert 2 is positively biased. Given the payoff symmetry, if both experts are negatively biased, none of them randomizes for $s_i = 0$, $i = 1, 2$.

Q.E.D.

Proof of Proposition 7: Consider two positively biased experts in the star, denoted by Q^s, and the following mixed strategy for each expert $i = 1, 2$. If $s_i = 0$, an expert i sends a message $m_i = 0$ with probability p_i and $m_i = 1$ with probability $1 - p_i$. If $s_i = 1$, expert i sends the message $m_i = 1$. Denote the message profile received by the decision maker by (m_1, m_2). We start with the decision maker's posterior's and expected values of θ depending on the received messages :

1. Suppose that the decision maker receives a message $(0, 0)$. Since a type of an expert with signal 1 never sends message 0, the decision maker's posterior distribution over θ is simply $3(1-\theta)^2$ with the expected value $\int_0^1 \theta 3(1-\theta)^2 d\theta = \frac{1}{4}$.

2. Suppose, next, that the decision maker receives a message profile (0,1). Then, her poste-

rior of θ is

$$f(\theta|m_1=0,m_2=1)=\frac{p_1(1-p_2)f(s_1=s_2=0|\theta)+\frac{1}{2}p_1f(s_1=0,s_2=1|\theta)}{\int_0^1\Big[p_1(1-p_2)f(s_1=s_2=0|\theta)+\frac{1}{2}p_1f(s_1=0,s_2=1|\theta)\Big]d\theta}=$$

$$\frac{(1-p_2)(1-\theta)^2+\frac{1}{2}2\theta(1-\theta)}{\int_0^1\Big[(1-p_2)(1-\theta)^2+\frac{1}{2}2\theta(1-\theta)\Big]d\theta}=\frac{6(\theta-1)(p_2(\theta-1)+1)}{2p_2-3},$$

and so the expected value is

$$\int_0^1\theta\frac{6(\theta-1)(p_2(\theta-1)+1)}{2p_2-3}d\theta=\frac{2-p_2}{6-4p_2}.$$

3. Next, suppose that the decision maker receives the message (1,0). Given the previous result, the posterior, expected value and the ex ante probability of this message profile is

$$f(\theta|(1,0))=\frac{6(\theta-1)(p_1(\theta-1)+1)}{2p_1-3},\ E(\theta|(1,0))=\frac{2-p_1}{6-4p_1},\ Pr((m_1,m_2)=(1,0))=\frac{p_1(3-2p_1)}{6}.$$

4. Finally, suppose that the decision maker receives the message profile (1,1). Her posterior in this case is

$$f(\theta|m_1=1,m_2=1)=$$

$$\frac{(1-p_1)(1-p_2)f(s_1=s_2=0|\theta)+[2-p_1-p_2]f(s_1=0,s_2=1|\theta)+f(s_1=s_2=1|\theta)}{\int_0^1\Big[(1-p_1)(1-p_2)f(s_1=s_2=0|\theta)+[2-p_1-p_2]\frac{1}{2}f(s_1=0,s_2=1|\theta)+f(s_1=s_2=1|\theta)\Big]d\theta}$$

where the above formulation uses the fact that $f(\theta|s_1=0,s_2=1)=f(\theta|s_1=1,s_2=0)$. Using the prior distributions for the signals, we can rewrite the above expression as

$$\frac{(1-p_1)(1-p_2)(1-\theta)^2+[2-p_1-p_2]\theta(1-\theta)+\theta^2}{\int_0^1\Big[(1-p_1)(1-p_2)(1-\theta)^2+[2-p_1-p_2]\theta(1-\theta)+\theta^2\Big]d\theta}=$$

$$\frac{6(p_1(\theta-1)+1)(p_2(\theta-1)+1)}{p_1(2p_2-3)-3p_2+6}.$$

Decision maker's payoff in the star

Applying the same method as throughout the paper, we obtain

$$EU_{DM}(Q^s, \mathbf{m}_1(p_1,p_2), \mathbf{m}_2(p_1,p_2)) = \frac{p_1^2((141-52p_2)p_2-96)+3p_1(p_2(47p_2-134)+96)-24(3-2p_2)^2}{48(2p_1-3)(2p_2-3)(p_1(2p_2-3)-3p_2+6)}$$

Expert's biases in the star

$$-\frac{2}{3}p_2\int_0^1\Big(\frac{1}{4}-\theta-b_1\Big)^2 3(1-\theta)^2 d\theta-\frac{2}{3}(1-p_2)\int_0^1\Big(\frac{2-p_2}{6-4p_2}-\theta-b_1\Big)^2 3(1-\theta)^2 d\theta-$$

$$\frac{1}{3}\int_0^1\Big(\frac{2-p_2}{6-4p_2}-\theta-b_1\Big)^2 6\theta(1-\theta)d\theta,$$

whereas if he sends $m_1=1$, his expected payoff is

$$-\frac{2}{3}p_2\int_0^1\Big(\frac{2-p_1}{6-4p_1}-\theta-b_1\Big)^2 3(1-\theta)^2 d\theta-\frac{2}{3}(1-p_2)\int_0^1\Big(\frac{p_1(p_2-2)-2p_2+6}{4p_1p_2-6p_1-6p_2+12}-\theta-b_1\Big)^2 3(1-\theta)^2 d\theta-$$

$$\frac{1}{3}\int_0^1\Big(\frac{p_1(p_2-2)-2p_2+6}{4p_1p_2-6p_1-6p_2+12}-\theta-b_1\Big)^2 6\theta(1-\theta)d\theta,$$

and so the incentive constraint guaranteeing indifference of expert 1 can be written as:

$$\frac{2}{3}p_2\Big(\frac{2-p_1}{6-4p_1}-\frac{1}{4}\Big)\Big(\frac{2-p_1}{6-4p_1}+\frac{1}{4}-2\frac{1}{4}-2b_1\Big)+$$

$$\frac{2}{3}(1-p_2)\Big(\frac{p_1(p_2-2)-2p_2+6}{4p_1p_2-6p_1-6p_2+12}-\frac{2-p_2}{6-4p_2}\Big)\Big(\frac{p_1(p_2-2)-2p_2+6}{4p_1p_2-6p_1-6p_2+12}+\frac{2-p_2}{6-4p_2}-2\frac{1}{4}-2b_1\Big)+$$

$$\frac{1}{3}\Big(\frac{p_1(p_2-2)-2p_2+6}{4p_1p_2-6p_1-6p_2+12}-\frac{2-p_2}{6-4p_2}\Big)\Big(\frac{p_1(p_2-2)-2p_2+6}{4p_1p_2-6p_1-6p_2+12}+\frac{2-p_2}{6-4p_2}-2\frac{1}{2}-2b_1\Big)=0.$$

This results in the bias:

$$b_1=\frac{(4(14-5p_1)p_1-39)p_2^3+2(p_1(55p_1-159)+114)p_2^2-3(p_1(61p_1-180)+132)p_2+24(3-2p_1)^2}{8(2p_1-3)(2p_2-3)(p_1(2p_2-3)-3p_2+6)(p_1(3p_2-4)-4p_2+6)}.$$

And therefore:

$$b_2=\frac{p_1^3(4(14-5p_2)p_2-39)+2p_1^2(p_2(55p_2-159)+114)-3p_1(p_2(61p_2-180)+132)+24(3-2p_2)^2}{8(2p_1-3)(2p_2-3)((2p_1-3)p_2-3p_1+6)((3p_1-4)p_2-4p_1+6)}.$$

Next, consider the line where the bottom expert is denoted by 1 and the intermediary by 2. First, notice that the above equilibrium outcome in the star can in principle be implemented in the line. However, the incentive constraints for the intermediary are tighter compared to the corresponding constraints in the star as in the line the intermediary observes the exact message of the bottom expert. Thus, for the same strategy profile the line cannot dominate the star. However, the line allows for pooling of signals received by the intermediary. Consider the profile where expert 1 reveals his signal. From the previous discussion we know that if expert 2 communicates according to the partition (of the sum of signals) $\{\{0\},\{1,2\}\}$, the biases supporting this strategy profile are within $[-1/8, 3/16]$ and the expected payoff of the decision maker is $-\frac{5}{96}$. In the following we show that any randomization involving pooling cannot result in the payoff strictly higher than $-\frac{5}{96}$.

First, if the intermediary randomizes given the signals 00, and sends the same message whenever his signals are 01,10 and 11, the expected payoff of the decision maker is strictly below $-\frac{5}{96}$ due to more noise in communication.

Second, suppose that the intermediary randomizes if his summary statistic is 1. In particular, consider the strategy profile where the bottom expert reveals his signal the the intermediary sends $m_2 = m$ if both signals are 0, $m_2 = m'$ if both signals are 11, and if the sum of both signals is 1, he sends $m_2 = m$ with probability π and $m_2 = m'$ otherwise. The decision maker's posteriors and expected values of θ depending on expert 2's message are

$$f(\theta|m) = \frac{(1-\theta)^2 + \pi 2\theta(1-\theta)}{\int_0^1 [(1-\theta)^2 + \pi 2\theta(1-\theta)]d\theta} = -\frac{3(\theta-1)((2\pi-1)\theta+1)}{\pi+1}$$

resulting in the expected value $\frac{1+2\pi}{4(1+\pi)}$. Further,

$$f(\theta|m') = \frac{3\theta(-2\pi(\theta-1)+\theta-2)}{\pi-2}$$

resulting in the expected value $\frac{5-2\pi}{4(2-\pi)}$. The decision maker's expected payoff is

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

$$\frac{2-\pi}{3}\int_0^1 \Big(\frac{5-2\pi}{4(2-\pi)}-\theta\Big)^2 \frac{3\theta(-2\pi(\theta-1)+\theta-2)}{\pi-2}d\theta = -\frac{5+4\pi(1-\pi)}{48(1+\pi)(2-\pi)}.$$

Since

$$argmax_{\pi\in[0,1]} - \frac{5+4\pi(1-\pi)}{48(1+\pi)(2-\pi)}$$

yields $\pi = 0$ or $\pi = 1$. Thus, for $b \leq \frac{3}{16}$ the decision maker does not benefit from randomization in the line, and therefore the best partition is simply $\{\{0\},\{1,2\}\}$ resulting in the payoff $-\frac{5}{96}$.

Comparison of the star and the line

Suppose $p_1 = 1$. Then:

$$b_1(p_2 = 1) = \frac{3p_2^3 - 20p_2^2 + 39p_2 - 24}{8(p_2 - 3)(p_2 - 2)(2p_2 - 3)}$$

where b_1 monotonically decreases in p_2 and reaches its minimum at $p_2 = 1$. Take any $\epsilon > 0$ that is sufficiently small. For $p_2 = 1 - \epsilon$, the bias b_1 is slightly higher than $\frac{1}{8}$. Further, for $p_1 = 1$ the critical bias b_2 is

$$\frac{3p_2^2 - 10p_2 + 9}{-16p_2^3 + 104p_2^2 - 216p_2 + 144}$$

that is monotonically increasing in p_2 and reaches its maximum $\frac{1}{8}$ at $p_2 = 1$. Therefore, for $p_1 = 1$ and $p_2 = 1-\epsilon$ the critical bias b_2 is slightly lower than $\frac{1}{8}$. However, because the expected value difference between perfect truthtelling with 2 experts and a line with $\{\{0\}, \{1, 2\}\}$ is strictly positive, for sufficiently small ϵ, the mixed strategy equilibrium dominates the line with pooling using pure strategies.

10 Biases of different signs

Assume $b_1 > 0$ and $b_2 < 0$. Think about the following strategy profile: expert 1 randomizes for $s_1 = 0$: he sends $m_1 = 0$ with probability p_1 and $m_1 = 1$ otherwise. Moreover, he perfectly reveals his signal 1. Expert 2 randomizes for $s_2 = 1$: he sends $m_2 = 1$ with probability p_2 and $m_1 = 0$ otherwise. Moreover, he perfectly reveals his signal 0.

10.1 Posteriors and expected values

Suppose the decision maker receives the message profile (0,0). Then her posterior of θ is

$$f(\theta|(0,0)) = \frac{p_1(1-\theta)^2 + p_1(1-p_2)\theta(1-\theta)}{\int_0^1 p_1(1-\theta)^2 + p_1(1-p_2)\theta(1-\theta)d\theta} = \frac{6(1-\theta)(1-p_2\theta)}{3-p_2},$$

and so the expected value is

$$E(\theta|(0,0)) = \int_0^1 \theta\frac{6(1-\theta)(1-p_2\theta)}{3-p_2}d\theta = \frac{2-p_2}{2(3-p_2)}$$

and the ex ante probability of the message profile $(0,0)$ is

$$Pr((m_1, m_2) = (0,0)) = \int_0^1 p_1(1-\theta)^2 + p_1(1-p_2)\theta(1-\theta)d\theta = p_1\frac{3-p_2}{6}.$$

Suppose, next, that the decision maker receives the message profile (0,1). Since the decision maker knows $m_1 = 0$ ($m_2 = 1$) only comes from expert 1 (expert 2) with the signal $s_1 = 0$ ($s_2 = 1$), the posterior is

$$f(\theta|(0,1)) = \frac{p_1 p_2 \theta(1-\theta)}{\int_0^1 p_1 p_2 \theta(1-\theta) d\theta} = 6\theta(1-\theta),$$

and the expected value is $\int_0^1 6\theta(1-\theta) = \frac{1}{2}$, and the ex ante probability of the message profile (0,1) is $\int_0^1 p_1 p_2 \theta(1-\theta) d\theta = \frac{p_1 p_2}{6}$.

Suppose, next, that the decision maker receives the message profile (1,0). The decision maker's posterior over θ is

$$f(\theta|(1,0)) = \frac{(1-p_1)(1-\theta)^2 + (1-p_1)(1-p_2)\theta(1-\theta) + \theta(1-\theta) + (1-p_2)\theta^2}{\int_0^1 (1-p_1)(1-\theta)^2 + (1-p_1)(1-p_2)\theta(1-\theta) + \theta(1-\theta) + (1-p_2)\theta^2 d\theta} =$$

$$-\frac{6(p_1(\theta-1)+1)(p_2\theta-1)}{p_1(p_2-3)-3p_2+6},$$

and so the expected value is

$$\int_0^1 \theta\Big(-\frac{6(p_1(\theta-1)+1)(p_2\theta-1)}{p_1(p_2-3)-3p_2+6}\Big) d\theta =$$

$$\frac{p_1(p_2-2)-4p_2+6}{2(p_1(p_2-3)-3p_2+6)},$$

and the ex ante probability of the message profile (1,0) is

$$\int_0^1 \Big[(1-p_1)(1-\theta)^2 + (1-p_1)(1-p_2)\theta(1-\theta) + \theta(1-\theta) + (1-p_2)\theta^2 d\Big]\theta =$$

$$\frac{1}{6}(p_1(p_2-3)-3p_2+6).$$

Finally, suppose that the decision maker receives the message profile (1,1). Her expectation of θ is

$$f(\theta|(1,1)) = \frac{(1-p_1)p_2\theta(1-\theta) + p_2\theta^2}{\int_0^1 (1-p_1)p_2\theta(1-\theta) + p_2\theta^2 d\theta} = \frac{6(\theta - \theta p_1(1-\theta)}{3-p_1},$$

and so the expected value is

$$E(\theta|(1,1)) = \int_0^1 \theta\frac{6(\theta-\theta p_1(1-\theta)}{3-p_1} d\theta = \frac{4-p_1}{2(3-p_1)}$$

and the ex ante probability of receiving the message profile (1,1) is

$$\int_0^1 (1-p_1)p_2\theta(1-\theta) + p_2\theta^2 d\theta = \frac{1}{6}p_2(3-p_1).$$

As a result, the expected payoff for the decision maker is

$$-p_1\frac{3-p_2}{6}\int_0^1 \Big(\frac{2-p_2}{2(3-p_2)} - \theta\Big)^2 \frac{6(1-\theta)(1-p_2\theta)}{3-p_2} d\theta -$$

$$\frac{p_1p_2}{6}\int_0^1 \Big(\frac{1}{2} - \theta\Big)^2 6\theta(1-\theta) d\theta -$$

$$\frac{1}{6}(p_1(p_2-3)-3p_2+6)\int_0^1 \Big(\frac{p_1(p_2-2)-4p_2+6}{2(p_1(p_2-3)-3p_2+6)} - \theta\Big)^2 \Big(-\frac{6(p_1(\theta-1)+1)(p_2\theta-1)}{p_1(p_2-3)-3p_2+6}\Big) d\theta -$$

$$\frac{1}{6}p_2(3-p_1)\int_0^1 \Big(\frac{4-p_1}{2(3-p_1)} - \theta\Big)^2 \frac{6(\theta-\theta p_1(1-\theta)}{3-p_1} d\theta =$$

$$\frac{p_1^2((21-4p_2)p_2-24)+3p_1(p_2(7p_2-34)+36)-12(p_2-3)(2p_2-3)}{24(p_1-3)(p_2-3)(p_1(p_2-3)-3p_2+6)}.$$

The bias of expert 1 with $s_1 = 0$ solves the following problem (making the player indifferent between sending a message $m_1 = 0$ and $m_1 = 1$.

$$\frac{2}{3}\Big(\frac{p_1(p_2-2)-4p_2+6}{2(p_1(p_2-3)-3p_2+6)} - \frac{2-p_2}{2(3-p_2)}\Big)\Big(\frac{p_1(p_2-2)-4p_2+6}{2(p_1(p_2-3)-3p_2+6)} + \frac{2-p_2}{2(3-p_2)} - \frac{1}{2} - 2b_1\Big)+$$

$$\frac{1}{3}(1-p_2)\Big(\frac{p_1(p_2-2)-4p_2+6}{2(p_1(p_2-3)-3p_2+6)} - \frac{2-p_2}{2(3-p_2)}\Big)\Big(\frac{p_1(p_2-2)-4p_2+6}{2(p_1(p_2-3)-3p_2+6)} + \frac{2-p_2}{2(3-p_2)} - 1 - 2b_1\Big)+$$

$$\frac{1}{3}p_2\Big(\frac{4-p_1}{2(3-p_1)} - \frac{1}{2}\Big)\Big(\frac{4-p_1}{2(3-p_1)} + \frac{1}{2} - 1 - 2b_1\Big) = 0$$

Finally, the bias of expert 2 with $s_2 = 1$ solves the following problem (making the player indifferent between sending a message $m_2 = 0$ and $m_2 = 1$):

$$\frac{2}{3}\Big(\frac{p_1(p_2-2)-4p_2+6}{2(p_1(p_2-3)-3p_2+6)} - \frac{4-p_1}{2(3-p_1)}\Big)\Big(\frac{p_1(p_2-2)-4p_2+6}{2(p_1(p_2-3)-3p_2+6)} + \frac{4-p_1}{2(3-p_1)} - \frac{6}{4} - 2b_2\Big)+$$

$$\frac{1}{3}(1-p_1)\Big(\frac{p_1(p_2-2)-4p_2+6}{2(p_1(p_2-3)-3p_2+6)} - \frac{4-p_1}{2(3-p_1)}\Big)\Big(\frac{p_1(p_2-2)-4p_2+6}{2(p_1(p_2-3)-3p_2+6)} + \frac{4-p_1}{2(3-p_1)} - 1 - 2b_2\Big)+$$

$$\frac{1}{3}p_2\Big(\frac{1}{2} - \frac{2-p_2}{2(3-p_2)}\Big)\Big(\frac{1}{2} + \frac{2-p_2}{2(3-p_2)} - 1 - 2b_2\Big) = 0$$

Now, suppose that $p_2 = 1$. Then, the critical biases are:

$$b_1 = \frac{3p_1^2 - 10p_1 + 9}{-16p_1^3 + 104p_1^2 - 216p_1 + 144},$$

$$b_2 = \frac{-8p_1^3 + 49p_1^2 - 87p_1 + 48}{8(p_1 - 3)(2p_1 - 3)(3p_1 - 4)}$$

and the decision maker's expected payoff is

$$-\frac{-7p_1^2 + 27p_1 - 24}{48(3 - 2p_1)(p_1 - 3)}.$$

To show that a star can dominate a line with the partition $\{\{0\}, \{1, 2\}\}$, notice that at $p_1 = 1$ the critical bias is $b_1 = \frac{1}{8}$ and $b_2 = -\frac{1}{8}$. Moreover, both continuously decrease in p_1. Further, the expected payoff for $p_1 = 1$ is $-\frac{1}{24}$, and the corresponding best payoff the line is it $-\frac{5}{96}$. Take a small enough $\epsilon > 0$ and take $p_1' = 1 - \epsilon$. Then b_1 is slightly below $\frac{1}{8}$ and b_2 is slightly below $-\frac{1}{8}$. Thus, the line cannot implement a perfectly revealing equilibrium, and the mixed strategy equilibrium in the star is arbitrary close to $-\frac{1}{24}$ (and strictly above $-\frac{5}{96}$). As a result, the proposed mixed strategy equilibrium in the star strictly dominates the line.