Optimal Organizational Structure

Ashley Perry

June 14, 2020

Abstract

This paper is concerned with the relative efficiency of two organizational structures, which I call simple and complex. I model the organization as a communication network where the structure governs who communicates with whom. Workers receive a noisy signal, they meet a decision maker (DM) and communicate their information as a report. The DM must then estimate the true state of the world based on these reports. The clarity of the report is dependent on the length of the meeting and the communication ability of the worker. The DM is assumed to have a greater value of opportunity cost than the workers. I show that even in the absence of strategic incentives that different communication structures can be preferred.

1 Introduction

In order for a decision maker of an organization to decide on a course of action where she is not fully informed, she can extract information from other potentially better informed employees. In a Marschak and Radner (1972) framework where the firms' objective and that of its' employees are aligned, the organizational structure of the firm, by dictating who interacts with whom, can determine what and how this information is used. Yet, even in this team theoretic case when employees have a common objective, individually they may disagree on how to best achieve it. Understanding this trade off and the situations in which one may be more preferable to the other, is important for deciding on how best to organize employees so that the optimal decisions for the firm are taken.

Arrow (1974) argues that organizations exist to allow economic agents to coordinate on collective actions when there is uncertainty about the state of the world. This uncertainty could be due to a lack of information. For instance, a manager may need to decide on an output level for their good. Although they may be knowledgeable about the production process and so the potential capacity constraint of their firm. They may not be as informed about their competitors production decisions and so what the possible demand for their product may be. A potential solution to this problem is for the manager to gather information from other employees whose job it is to know about such matters. There are two potentially complementary mechanisms by which the production manager could make the best choice. Either they speak to a large number of employees leveraging the wisdom of crowds argument (Surowiecki (2005)), or they speak to more informed employees. In the first instance, the manager may want to spend her time meeting as many people as possible. In the second, she may only contact a few employees but might invest large amounts of time in getting the most out of them.

Clearly, how a manager spends her time and who she spends it with will have an impact on the quality of the decision she makes. However, she might also have other responsibilities, such as raising outside investment in the firm or bidding for new contracts. If she spends all her time in meetings then there is clearly a cost in terms of the time she could have spent doing on some other productive activity.

There are two structure's which differ in the degrees of separation or layers of the organization. The first structure only has 1 degree of separation between the decision maker and the lowest ranking employee. Whereas, the second structure has 2 degrees of separation. I will refer to the first organization as simple and the second one as complex. I introduce the idea that the value of the DM's opportunity cost is higher. I show that if the environment is symmetric, in terms of processing and communication ability that even if the value of opportunity cost goes to ∞ the simple structure is always preferred to the complex. However, if I introduce heterogeneous communication abilities such that the worker directly linked to the DM is better at communicating. Then I can show that there exists a point where the value of the opportunity cost is large enough for the preference to switch to the complex structure.

Section 2 reviews the related literature and how this model contributes to it. Section 3 will detail the models for two particular organizational structures, which I will call simple and complex. Section 4 gives details for the value of opportunity cost. Section 5 will detail the main results of the paper, with more general results in section 6. I conclude in section 7. Finally, the proofs of the propositions of this paper can be found in appendix A.

2 Related Literature

The underlying framework in the model of this paper is based on the Bayesian approach to signal extraction (see Kay (1993)). A number of papers have worked with a particular functional form that is consistent with the Bayesian approach, the quadratic loss function (see Morris and Shin (2002)). Working with this environment Calvó-Armengol et al. (2015) study agents coordinating their action on an unknown (local to the agent) state of the world and the actions of other agents. The agents make decisions endogenously on how much active communication (speaking or writing) and passive communication (reading or listening) to in invest in, taking into account how others will choose. They show there are strategic complementarities between the two forms of communication. That is if one agent believes another agent will listen to them, then they will invest time in improving their speaking and vice versa. As in Calvó-Armengol et al. (2015), I will also endogenize communication flows, but I shut down the consideration of coordinating on other agents actions and focus instead on the relative payoffs of different communication networks.

A particular application of this Bayesian setup is to political parties by Dewan and Myatt (2009). They look at how political leaders can influence the beliefs and actions of party members (followers). A noisy signal of the true policy is revealed to the leaders who have certain predetermined abilities such as "sense of direction" and "clarity". The followers want to make both the right choice (match the policy) and the choice that is inline with other followers (match the average action of the followers). They show that in equilibrium, a follower will listen to leaders who have already captured the attention of other followers. This is determined by the underlying "sense of direction" of the leader. Dewan and Myatt (2009) are concerned with understanding how a group of candidate political party leaders can be selected by the members of their party after they have assessed the information the leaders communicate. Whereas, this paper is not concerned with an organization working out which single employee they should listen to among their employees based on their ability. Rather, it is determining how to best organize the employees a firm has to extract relevant information, given their abilities.

Dessein et al. (2016) study optimal communication flows in an organization when solving tasks. They utilize the idea that for a given task the firm's loss function is dependent on how well the actions of the employees match the underlying state of the world and how well they coordinate with each other. However, in their setup hierarchy plays no role in the firm and so cannot answer if a particular organizational design is preferred to another.

There are a number of other methods to model the optimal structures of firms. Bolton and Dewatripont (1994) approach this problem by arguing that a firm's objective is to minimize the time it's employees spend processing information. They show that particular network structures arise as employees specialize in handling certain types of information and then communicate it to each other. This approach focuses on reducing the time an agent requires to process information and highlights which communication networks of employees aide this. In contrast, the approach I take is to hold time as fixed and then understand how to allocate that time effectively within various organizational structures.

Garicano (2000) posits that the aim of the firm is not necessarily to solely process information, but instead to acquire the appropriate knowledge to solve the problems it faces. He finds that in response to heterogeneity in the difficulty of the tasks, the optimal structure is a pyramidal hierarchy. This approach focuses on how, in response to the uncertainty of the types of problems a firm faces, should that firm invest in its employees skills and then organize them. This implicitly assumes a long term horizon for the firm and the pyramidal structure is a result of learning the frequency of particular problems. In contrast, the approach in this paper is to instead look at a "snap shot" of the firm and understand how to best organize itself in response to a particular problem given the resources it has.

Migrow (2018) argues that particular structure arise to overcome misaligned objectives between it's employees. As with the setup of this paper, he has two types of agents; experts who receive the information and a single decision maker. But unlike this paper, he assumes that the objectives of the agents are different. Specifically, the experts try to strategically manipulate the action of the decision maker towards their "biased" action by the information they reveal. The main finding is that it is optimal to group the experts by their biases with one expert aggregating and and relaying all the information to the decision maker. He also finds that if the biases are sufficiently close then all that is needed is one group.

This approach sheds light on what structures help create incentives for employees to act in the interest of their firm when they don't necessarily align with their own. Whilst this is undoubtedly relevant for many contexts. This approach does not help explain, as this paper attempts to do, why hierarchies might exist when there is a strong organizational goal which aligns employees objectives.

3 Model

Agents and Structure. I consider two organizational structures. Since the type of agents are common to both I will outline them briefly. There are two types of agents; A decision maker (DM) who is the head of the organization and workers who are subordinates to the DM. I assume, for now, that there are only two workers. This small structure is sufficient to demonstrate my main findings. In a later section I will show that the model can be extended to a more general setting with n workers.

I first consider a simple structure, where the DM is directly connected to the workers (see figure 1). I then consider a complex structure (a line), where the DM is directly connected to a single worker who then is also connected to the other worker (see figure 2). The arrow head indicates the direction in which information flows.

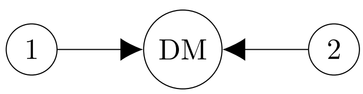

Figure 1: Simple structure

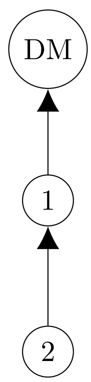

Figure 2: Complex structure

3.1 Simple Structure

This section outlines the baseline setting for this model and how agents interact.

Information and meeting time. Each agent is endowed with a fixed amount of meeting time κ. The DM has a prior about the unknown state of the world θ. This prior is normally distributed as follows

$$\theta \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2)$$

The workers receive an individual signal about the state of the world with some normally distributed noise ϵ_i for $i = \{1, 2\}$. The parameter $\rho_i \geq 0$ captures the the processing ability of the worker, the higher ρ_i the lower the variance of the error term ϵ_i and the more accurate the signal is. Hence, the more able the worker is at processing the signal they receive, the clearer the signal is. The $\rho_i's$ are taken as exogenous. The information contained within each signal for worker i is given by

$$s_i = \theta + \epsilon_i \qquad \text{where} \qquad \epsilon_i \sim \mathcal{N}\left(0, \frac{1}{\rho_i}\right)$$

Given the simple structure, the workers can interact with the DM, but can not do so with each other. As the workers are assumed to have the same objective as the DM, the workers don't try to bias the DM's decision. However, their signal is communicated to the DM along with some additional normally distributed noise η_i. I will call this the report of the meeting. The DM can allocate varying amounts of meeting time $m_i \geq 0$, to determine the clarity of the report by affecting the variance of the noise term η_i. For instance, the interaction between the DM and worker 1 can be thought of as a meeting and m_1 captures the length of time of that meeting. However, given that the maximum time the DM has for meeting is κ it must be that $\sum_1^2 m_i \leq \kappa$. This constraint will also ensure that the length of time worker 1 and 2 are in a meeting doesn't exceed their endowment of κ either. Along with the endogenously determined meeting time the noise will be exogenously affected by an innate communicating ability ξ_i^2. This will play an important role in determining the optimal communication structure. The reports are of the following form

$$r_i = s_i + \eta_i \qquad \text{where} \qquad \eta_i \sim \mathcal{N}\left(0, \frac{\xi_i^2}{m_i}\right)$$

The random variables $\{\theta, \epsilon_i, \eta_i\}$ are all independent of each other. For instance, for workers 1 and 2 the co-variance of the error terms in their signals would be zero.

Action and payoffs. Given the vector of reports $\mathbf{r}$ the DM has, she aims to take an action a that corresponds to the true state of the world θ. The setting here is a team theoretic one. The objective of the workers and DM are common. This means that the payoff of the DM can be thought of as the same as the firm. This payoff is given by the following utility function

$$u_{dm} = \bar{u} - (\theta - a)^2$$

where the DM maximizes her utility by minimizing the quadratic loss function.

Given this payoff and taking the constraints of the meeting time available the optimization problem the DM faces is the following

$$\begin{aligned} \max_{m_1, m_2} \bar{u} - \mathbb{E}\big[(\theta - a)^2 | \mathbf{r}\big] \\ \text{s.t.} \quad m_1 + m_2 \leq \kappa \end{aligned} \tag{1}$$

Note that there are constraints on the time the workers can meet e.g. for worker 1 it must be that $m_1 < \kappa$. But, this is always satisfied by the constraint in (1).

Timing. The DM decides on the length of the meetings $\mathbf{m}$. The workers receive their signal, whose accuracy is determined by the given ρ_i. The workers then meet the DM and a report is produced, for which the clarity depends on the chosen m_i and given ξ_i^2. The DM observes the vector of reports $\mathbf{r}$ and then decides on an action a which maximizes her utility u_{dm}.

3.1.1 Optimal action and expected utility for an arbitrary meeting allocation

The expected payoff of the DM is $u_{dm} = \bar{u} - \mathbb{E}\big[(\theta - a)^2|\mathbf{r}\big]$ and assuming an arbitrary meeting allocation $\mathbf{m}$ and reports $\mathbf{r}$, the first order condition with respect to a yields $a^* = \mathbb{E}[\theta|\mathbf{r}]$. Given the linear structure of the messages, the normally distributed error terms and the posterior PDF $p(\theta|\mathbf{r})$ is Gaussian the optimal action a^* will be a linear action. The linear structure of a^* and the independence of the error terms simplifies her objective function, as given in the following proposition.

Proposition 1. *Let there be an arbitrary meeting time allocation $\{m_1, m_2\}$ and diffuse prior $\sigma_\theta^2 = \infty$. Then, given reports $\{r_1, r_2\}$ the optimal action and the expected payoff of the DM is given by*

$$a^* = \sum_1^2 \omega_i r_i \qquad \textit{and} \qquad u_{dm} = \bar{u} - \frac{1}{\sum_1^2 h_i}$$

Where $\omega_i = \frac{\frac{\rho_i m_i}{\xi_i^2 \rho_i + m_i}}{\sum_{i=1}^n \frac{\rho_i m_i}{\xi_i^2 \rho_i + m_i}}$ *and* $h_i = \frac{\rho_i m_i}{\xi_i^2 \rho_i + m_i}$ $\quad \forall i = \{1, 2\}$.

This shows that the optimal action for the DM is a weighted linear combination of the reports she receives. Where the weight ω_i is the relative precision the DM assigns to a given report. The utility of the DM is increasing in the precision weight h_i of each workers report. The more workers she meets with, the greater the number of reports and so greater the utility.

3.2 Complex structure

This section outlines how agents interact in the complex structure. That is the DM has a single connection to worker 1, who then has another connection to worker 2. The DM and worker 2 have no direct connection between them (see figure 2).

Information and meeting time. In the complex structure the DM can no longer meet with worker 2, but only with worker 1. Thus, she now decides on how long she should meet, m_1, with worker 1 and how long worker 1 should meet, m_2, with worker 2. The constraint on the time is still $\sum_1^2 m_i \le \kappa$, but now applies to worker 1. Information is passed through the organization as follows, worker 1 processes the information available to him, after his meeting with worker 2, and then transmits this to the DM in their meeting. The other details, such as the independence of the error terms, are kept the same.

Action and payoffs. The report from worker 2's meeting with worker 1 $r_{2,1}$ will be worker 2's signal and some normally distributed noise ϵ_2 dependent of their processing ability ρ_2, plus some normally distributed noise η_2 dependent on the amount of time they were allocated m_2 and their communicating ability ξ_2^2. Due to the independence of the error terms this report has the following structure

$$r_{2,1} = \theta + \mathcal{N}\big(0, \frac{1}{\rho_2} + \frac{\xi_2^2}{m_2}\big)$$

Worker 1's information set will be their own signal s_1 and report $r_{2,1}$. He can then estimate the state of the world conditional on this information

$$\mathbb{E}\left[\theta|\{s_1, r_{2,1}\}\right]$$

Worker 1 creates precision measures for each piece of information (which can be thought of as observations). In this example, he treats the signal as the first observation and the report as the second observation. The precision measures would be

$$h_{1,1} = \rho_1, \qquad h_{1,2} = \frac{m_2 \rho_2}{m_2 + \xi_2^2 \rho_2}$$

He then uses Bayesian updating to estimate a posterior belief

$$\mathbb{E}\left[\theta | \{s_1, r_{2,1}\}\right] = q_{1,1} s_1 + q_{1,2} r_{2,1}$$

Given the assumption of a diffuse prior the relative weights are defined as below.

$$q_{1,1} = \frac{h_{1,1}}{h_{1,1} + h_{1,2}} \text{ and } q_{1,2} = \frac{h_{1,2}}{h_{1,1} + h_{1,2}}$$

Where $q_{1,1} + q_{1,2} = 1$. He then sends a message to the DM, where the message consists of their posterior belief of the state of the world plus an error term η_1. This error term is the dependent on the length of the meeting m_1 between the DM and worker 1's communication ability. The report from that meeting has the following form

$$r_{1,dm} = \mathbb{E}\left[\theta | \{s_i, r_{2,1}\}\right] + \eta_1$$

The DM's maximization problem is the same as 1.

Timing. The DM decides on the length of the meetings $\mathbf{m}$. The workers receive their signal, whose accuracy is determined by the given ρ_i. The worker 1 and 2 meet and produce a report. Then worker 1 meets the DM and produces a report based on the information available. These reports depend on the chosen m_i and given ξ_i^2. The DM observes the vector of reports $\mathbf{r}$ and then decides on an action a which maximizes her utility.

3.2.1 Optimal action and expected utility for an arbitrary meeting allocation

The expected payoff of the DM is the same as before except it will be denoted as $\hat{u_{dm}}$, to help differentiate it from the simple structure. With a diffuse prior assumed the only information the DM has is the report she has from the meeting with worker 1. Hence the best she can do is to treat that as their optimal action. The following proposition outlines the form the DM's expected payoff takes as a result of her action.

Proposition 2. *Let there be an arbitrary meeting time allocation $\{m_1, m_2\}$ and diffuse prior $\sigma_\theta^2 = \infty$. Then, given reports $\{r_{2,1}, r_{1,dm}\}$ the optimal action and the expected payoff of the DM is given by*

$$a^* = r_{1,dm} = \theta + q_{1,1}\epsilon_1 + q_{1,2}\left(\epsilon_2 + \eta_2\right) + \eta_1 \qquad \textit{and} \qquad \hat{u_{dm}} = \bar{u} - \frac{1}{\rho_1 + \frac{m_2 \rho_2}{m_2 + \rho_2}} - \frac{1}{m_1}$$

4 Opportunity Cost

I will use the concept of opportunity cost to capture the idea that employees, in this case a single DM and two workers, have different values of how they spend their time. The DM has relatively less time than the workers for the meetings. It seems reasonable to argue that the higher up the organization an employee is the more likely they are to have a greater number of responsibilities than their junior colleagues. Hence, they may face greater trade-offs in how they spend their time.

I will keep the same setup from section 3, but will introduce heterogeneous values of opportunity cost. The DM will have a higher value of opportunity cost than the workers. In this model that corresponds to a smaller endowment of meeting time. The maximum time the DM can spend in meetings will be $\bar{m} \leq \kappa$. Whilst the workers can still spend κ.

Opportunity cost is the time that an agent could engage with other activities other than their primary one. Such as, raising investment or signing a new deal with a supplier. The activity of estimating the true θ (section 3) can be thought if as the primary activity, such a production decisions and internal promotions. This is not meant to be a comprehensive characterization, as clearly there could be some interdependence between these activities. It is only meant to illustrate that a manager has a fixed amount of time to spend on all the activities her job may entail and how she allocates this time has consequences for the firm.

This means that the constraint in (1) is now $m_1 + m_2 \leq \bar{m}$. Whereas, for the complex there are now two relevant constraints, given it is now worker 1 who is involved in two meetings, $m_1 \leq \bar{m}$ and $m_1 + m_2 \leq \kappa$.

5 Results

This section outlines the main findings of this paper.

Proposition 3. *Assume symmetric processing ability, $\rho_1 = \rho_2 = \rho$. Also assume the DM has higher value opportunity cost, $\bar{m} < \kappa$. If communication ability is symmetric, $\xi_1^2 = \xi_2^2 = \xi^2$ then the simple structure is preferred to the complex structure.*

The implication of proposition 3 is that in the complex setting, even if worker 1 and worker 2 meet for a *very* long time to minimize the communication noise between them, the performance of this structure is still worse than that of the simple. This is surprising since in the complex setting the sum of the time spent in meetings is greater than in the simple structure and so it might seem that the communication noise would be less in the complex structure. As more time can be spent by worker 1 and worker 2 in their meeting, to produce as clear a report as possible.

Figure 3 graphs the difference in payoff between the two structures when communication ability is the same. It illustrates that as κ increases the value of the opportunity cost difference between the two structures decreases although it is always positive. However, If I alter the

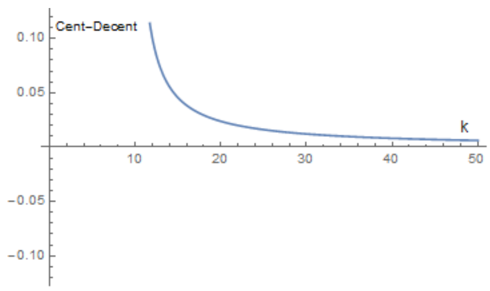

Figure 3: $\{\bar{m}, \rho, \xi_1^2, \xi_2^2\} = \{10, 1, 1, 1\}$

communication ability so that worker 1 is relatively better at communicating than worker 2. Then, there exists a threshold where the complex structure is preferred to the simple as the following proposition states.

Proposition 4. *Assume symmetric processing ability, $\rho_1 = \rho_2 = \rho$. Also assume the DM has higher value opportunity cost, $\bar{m} < \kappa$. If worker 1 is a better communicator than worker 2, $\xi_1^2 < \xi_2^2$, then there exists a threshold of meeting time $\bar{\kappa}$ such that the complex is preferred to the simple structure.*

The intuition for both these results is as follows. For proposition 3, the environment is symmetric in all respects ($\rho_1 = \rho_2$ and $\xi_1^2 = \xi_2^2$). Hence, given their identical communication abilities of the workers, in the simple structure the DM gives equal attention to them both.

This means that she is equally weighting the underlying signals of the reports, which given their identical processing ability is optimal. However, in the complex structure the DM is sub-optimally weighting the signals as she is giving less weight to worker 2's signal. This is true even as the workers are given relatively more time to meet and thus produce a clearer report. Since there is noise in the report between worker 1 and 2 that is unaffected by the DM meeting worker 1. The noise to signal ratio is always higher.

For proposition 4, the environment is no longer symmetric, since worker 1 is better at communicating than worker 2 ($\xi_1^2 < \xi_2^2$). Now, in the simple structure he DM will pay more attention to worker 1 than to worker 2. This means the DM is not equally weighting the underlying signals. This is sub-optimal since (a) the processing ability has not changed. and (b) the total noise generated in the simple structure is the same as under the symmetric case. The DM will meet both workers for a total of $\bar{m}$ amount of time but that will not be equally split.

However, this sub-optimal weighting can be overcome in the complex structure. If worker 1 and 2 can meet for a sufficient amount of time to re-balance the weighting of the signals. Then the weighting of signals can be optimized and the DM can extract more information. That is, there is point at which the preference for a structure flips from simple to complex. Demonstrating that when time is scarce for the head of an organization they would do well to have their best communicator directly speak to them once they have gathered all the relevant information. Rather than meeting everyone individually to gather the information.

Figure 4 graphs the difference in payoff between the two structures when worker 1 has a better communication ability than worker 2. It also shows that as κ increases the difference between the two structures decreases, but now there exits a point $\bar{\kappa}$ where the complex is preferred to the simple.

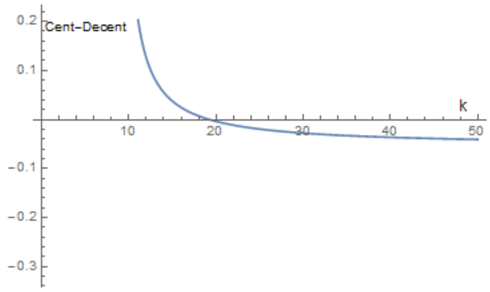

Figure 4: $\{\bar{m}, \rho, \xi_1^2, \xi_2^2\} = \{10, 1, 1, 1.5\}$

Proposition 3 applies to knife edge case but the result but it can be shown to apply more generally. That is if the assumptions are still met but that $\xi_1^2 \geq \xi_2^2$ then the simple structure is preferred to the complex structure. Figure 5 shows that if $\xi_1^2 < \xi_2^2$ as $\xi_1^2 \to \xi_2^2$ the threshold $\bar{\kappa}$ increases. This is seen by the shaded blue region corresponding to a threshold of $\bar{\kappa} = 10$ whereas the smaller orange region corresponding to a threshold of $\bar{\kappa} = 20$. The white region is the parameter space where there is no threshold and this is the region to the right of the 45° line e.g. $\xi_1^2 \geq \xi_2^2$.

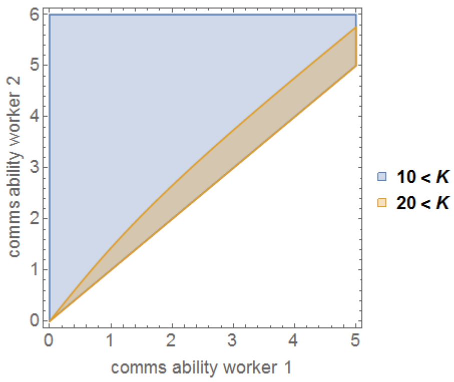

Figure 5: $\{\bar{m}, \rho, \xi_1^2, \xi_2^2\} = \{10, 1, 1, 1.5\}$

6 Extension

In this section I will extend the model to allow for n workers. Figure 6 shows that for the generalization in the simple structure the DM now meets with n workers. Whereas, figure 7 shows that in the complex the DM still only meets with worker 1 but that they now meet with $n-1$ workers.

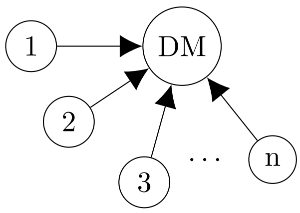

Figure 6: General simple structure

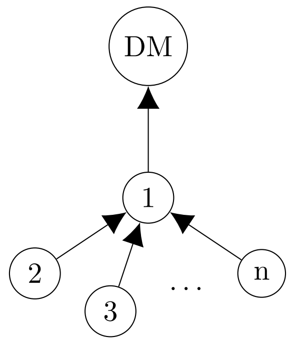

Figure 7: General complex structure

Introducing the extra agents in this way allows for this model to be applied to larger organizations or even teams within the organization. The simple structure can be thought of as a team where the decision making is concentrated, in the sense that the DM has to gather all the information dispersed within her team. Whereas, in the complex structure the DM "outsources" some of this information gathering activity to a single team member. As with the 3 agent version in this setting the results will rely on the heterogeneity in the workers communication skills.

Proposition 5. *Assume symmetric processing ability, $\rho_1 = \rho_2 = \rho$. Let workers $i \in \{2, \dots, n\}$ have the same communication ability ξ^2. Also assume the DM has higher opportunity cost, $\bar{m} < \kappa$. If all the other workers are the same as communicating as worker 1, $\xi_1^2 = \xi^2$. Then, the simple structure is preferred to the complex. However, if worker 1 is a strictly better communicator $\xi_1^2 < \xi^2$ then there exists a threshold $\bar{\kappa}$ such that the complex communication structure is preferred to the simple one.*

The intuition for this result is the same as for the 3 agent case. When the communication ability is homogeneous then the DM is optimally weighting the underlying signals in the simple structure. But, in the complex she is over weighting the signal from worker 1. This situation does not hold once worker 1 has better communicating ability than his co-workers and the value of the opportunity cost is high enough. Now in the simple structure the DM over weights the signal of worker 1, since they are better able to communicate. However, in the complex structure as the value of the opportunity cost gets bigger $k - \bar{m}$, the organization will reach a point where worker 1 can produce reports with the other workers such that the DM is optimally weighting the underlying signals.

7 Conclusion

In this paper I outline a framework to model the interactions amongst employees in a firm. I use a Marschak and Radner (1972) team theoretic setting, where all the employees have a common objective. I initially consider a firm with three employees a single decision maker (DM) and two workers. Then I extend the model to n workers. All the employees are endowed with a fixed amount of time κ. The DM has is to estimate an unknown state of the world based on information elicited from the workers.

I consider two organizational structures. The first I call simple, the DM separately meets both workers. The second I call complex, the DM only meets one worker and not the other but the workers meet with each other. The structures outline the communication protocols, that is who can communicate with whom. During these meeting a report is produced. The DM must choose the length of meetings between her and the workers given a particular structure. The clarity of the report is dependent on (a) the exogenous abilities of the worker and (b) the length of meeting chosen, given the time constraint κ.

There is a value of opportunity cost, which captures that the DM has valuable other activities they could be doing rather than being in a meeting given by $\kappa - \bar{m}$. I show that if the environment is symmetric, in the sens of processing ability and communication ability. Then, the simple structure is preferred to the complex. Even as $\kappa \to \infty$. The DM optimally weights the signals underlying the report. However, if the environment is asymmetric, in particular the communication ability of the worker 1 is better than worker 2. Then, the DM over weights the signal of worker 1 in the simple structure. Yet, since the DM only meets with worker 1 in the complex, there is a high enough value of opportunity. Where the DM is able to optimally weigh the signals. The complex structure is preferred to the simple. I extend this result to a more general setting with n workers.

References

Arrow, K. J. (1974): *The Limits of Organization*, W.W. Norton and Company, New York.

Bolton, P. and M. Dewatripont (1994): "The Firm as a Communication Network," *The Quarterly Journal of Economics*, 109, 809–839.

Calvó-Armengol, A., J. De Martí, and A. Prat (2015): "Communication and Influence," *Theoretical Economics*, 10, 649–690.

Dessein, W., A. Galeotti, and T. Santos (2016): "Rational Inattention and Organizational Focus," *American Economic Review*, 106, 1522–1536.

Dewan, T. and D. P. Myatt (2009): "The Qualities of Leadership: Direction, Communication, and Obsfuscation," *American Political Science Review*, 102, 351–368.

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

Kay, S. (1993): *Fundamentals of Statistical Signal Processing, Vol. I - Estimation Theory*, Prentice Hall.

Marschak, J. and R. Radner (1972): *Economic Theory of Teams*, Cowles Foundation and Yale University Press, New Haven.

Migrow, D. (2018): "Designing Communication Hierarchies," *Working Paper.*

Morris, S. and H. S. Shin (2002): "The Social Value of Public Information," *American Economic Review*, 92, 1521–1534.

Surowiecki, J. (2005): *The Wisdom of Crowds*, Anchor.

A Proofs

This section details the proofs for the propositions of the paper.
Proposition 1

Proof. Given the observed reports $\mathbf{r}$ are linear and have normally distributed error terms this means that the ith report can be written as follows.

$$\begin{aligned} r_i &= \theta + \epsilon_i + \eta_i \\ &= \theta + \mathcal{N}\left(0, \frac{1}{\rho_i}\right) + \mathcal{N}\left(0, \frac{\xi_i^2}{m_i}\right) \\ &= \theta + \mathcal{N}\left(0, \frac{m_i + \xi_i^2 \rho_i}{\rho_i m_i}\right) \end{aligned}$$

Where the precision weight h_i is the inverse of the variance.

$$h_i = \frac{\rho_i m_i}{\xi_i^2 \rho_i + m_i} \qquad \forall i = \{1, 2\}$$

I define h_0 as $\frac{1}{\sigma_\theta^2}$. With bayesian updating the posterior mean is equal to the precision weighted sum of the prior and sample observations over the total sum of the precision weights.

$$\mathbb{E}\big[\theta|\mathbf{m}\big] = \frac{h_o \mathbb{E}[\theta] + \sum_{i=1}^2 h_i r_i}{h_o + \sum_{i=1}^2 h_i}$$

This can re-arranged to

$$\begin{aligned} \mathbb{E}\big[\theta|\mathbf{r}\big] &= \frac{h_o \mathbb{E}[\theta] + \sum_{i=1}^2 h_i r_i}{h_o + \sum_{i=1}^2 h_i} \\ &= \mathbb{E}[\theta] - \mathbb{E}[\theta] + \frac{h_o \mathbb{E}[\theta] + \sum_{i=1}^2 h_i r_i}{h_o + \sum_{i=1}^2 h_i} \\ &= \mathbb{E}[\theta] - \frac{\mathbb{E}[\theta](h_o + \sum_{i=1}^2 h_i r_i)}{h_o + \sum_{i=1}^2 h_i} + \frac{h_o \mathbb{E}[\theta] + \sum_{i=1}^2 h_i r_i}{h_o + \sum_{i=1}^2 h_i} \\ &= \mathbb{E}[\theta] + \frac{-\sum_{i=1}^2 h_i r_i \mathbb{E}[\theta] - h_o \mathbb{E}[\theta] + h_o \mathbb{E}[\theta] + \sum_{i=1}^2 h_i r_i}{h_o + \sum_{i=1}^2 h_i} \qquad x \\ &= \mathbb{E}[\theta] + \frac{-\sum_{i=1}^2 h_i r_i \mathbb{E}[\theta] + \sum_{i=1}^2 h_i r_i}{h_o + \sum_{i=1}^2 h_i} \\ &= \mathbb{E}[\theta] + \frac{\sum_{i=1}^2 h_i (r_i - \mathbb{E}[\theta])}{h_o + \sum_{i=1}^2 h_i} \\ &= \mathbb{E}[\theta] + \frac{h_1 (r_1 - \mathbb{E}[\theta])}{h_o + \sum_{i=1}^n h_i} + \frac{h_2 (r_2 - \mathbb{E}[\theta])}{h_o + \sum_{i=1}^n h_i} \\ &= \mathbb{E}[\theta] + \omega_1 (r_1 - \mathbb{E}[\theta]) + \omega_2 (r_2 - \mathbb{E}[\theta]) \\ &= \mu_\theta + \omega_1 (r_1 - \mu_\theta) + \omega_2 (r_2 - \mu_\theta) \end{aligned}$$

Hence we get that

$$\mathbb{E}\big[\theta|\mathbf{r}\big] = \mu_\theta + \omega_1 \left(r_1 - \mu_\theta\right) + \omega_2 \left(r_2 - \mu_\theta\right) \tag{2}$$

Note that

$$\omega_i = \frac{\frac{\rho_i m_i}{\xi_i^2 \rho_i + m_i}}{\frac{1}{\sigma_\theta^2} + \sum_{i=1}^n \frac{\rho_i m_i}{\xi_i^2 \rho_i + m_i}}$$

Here ω_i is the relative precision assigned by the DM to a particular report.

If we think of the prior μ_θ as the 0^{th} report then (2) can be re-written as

$$\begin{aligned}
\mathbb{E}\left[\theta|\mathbf{r}\right] = a^* &= \mu_\theta + \omega_1 \left(r_1 - \mu_\theta\right) + \omega_2 \left(r_2 - \mu_\theta\right) \\
&= r_0 + \omega_1 \left(r_1 - r_0\right) + \omega_2 \left(r_2 - r_0\right) \\
&= r_0 + \omega_1 r_1 - \omega_1 r_0 + \omega_2 r_2 - \omega_2 r_0 \\
&= r_0 - \omega_1 r_0 - \omega_2 r_0 + \omega_1 r_1 + \omega_2 r_2 \\
&= r_0 (1 - \omega_1 - \omega_2) + \omega_1 r_1 + \omega_2 r_2 \\
&= r_0 \left(1 - \sum_{i=1}^{2} \omega_i\right) + \sum_{i=1}^{2} \omega_i r_i \\
&= r_0 \underbrace{\left(1 - \sum_{i=1}^{2} \omega_i\right)}_{\omega_0} + \sum_{i=1}^{2} \omega_i r_i
\end{aligned}$$

This means that optimal action of the DM can be written as $a^* = \sum_{i=0}^{2} \omega_i r_i$. Note, that it is also possible to arrive at this solution by assuming that the optimal strategy is some linear combination of the observed messages. Then plugging this into the ex-ante expected payoff of the DM and asking the question, “what are the optimal weights that would minimize the ex-ante expected payoff?”. The result would be a weight which is proportional to the precision weights for an individual message, the same result as above.

Plugging $a^* = \sum_{i=0}^{2} \omega_i r_i$ into the quadratic loss function yields

$$\begin{aligned}
\mathbb{E}\left[(\theta - a^*)^2\right] &= \mathbb{E}\left[(\theta - \sum_{i=0}^{3}\omega_i r_i)^2\right] \\
&= \mathbb{E}\left[(\theta - \sum_{i=1}^{3}\omega_i r_i - \omega_0 r_0)^2\right] \\
&= \mathbb{E}\left[(\theta - \sum_{i=1}^{3}\omega_i(\theta + \eta_i + \epsilon_i) - \omega_0 r_0)^2\right] \\
&= \mathbb{E}\left[(\theta - \sum_{i=1}^{3}\omega_i(\theta + \eta_i + \epsilon_i) - \left(1 - \sum_{i=1}^{3}\omega_i\right)\mu_0)^2\right] \\
&= \mathbb{E}\left[(\theta - \sum_{i=1}^{3}\omega_i(\theta + \eta_i + \epsilon_i) - \left(1 - \sum_{i=1}^{3}\omega_i\right)\mu_0 + \left(1 - \sum_{i=1}^{3}\omega_i\right)\theta - \left(1 - \sum_{i=1}^{3}\omega_i\right)\theta)^2\right] \\
&= \mathbb{E}\left[(\theta - \sum_{i=1}^{3}\omega_i(\theta + \eta_i + \epsilon_i) - \left(1 - \sum_{i=1}^{3}\omega_i\right)(\mu_0 - \theta) - \left(1 - \sum_{i=1}^{3}\omega_i\right)\theta)^2\right] \\
&= \mathbb{E}\left[(\left(1 - \sum_{i=1}^{3}\omega_i\right)\theta - \sum_{i=1}^{3}\omega_i(\eta_i + \epsilon_i) - \left(1 - \sum_{i=1}^{3}\omega_i\right)(\mu_0 - \theta) - \left(1 - \sum_{i=1}^{3}\omega_i\right)\theta)^2\right] \\
&= \mathbb{E}\left[(-\sum_{i=1}^{3}\omega_i(\eta_i + \epsilon_i) - \left(1 - \sum_{i=1}^{3}\omega_i\right)(\mu_0 - \theta))^2\right] \\
&= \mathbb{E}\left[(-1)^2(\sum_{i=1}^{3}\omega_i(\eta_i + \epsilon_i) + \left(1 - \sum_{i=1}^{3}\omega_i\right)(\mu_0 - \theta))^2\right] \\
&= \mathbb{E}\left[(\sum_{i=1}^{3}\omega_i(\eta_i + \epsilon_i) + \left(1 - \sum_{i=1}^{3}\omega_i\right)(\mu_0 - \theta))^2\right] \\
&= \mathbb{E}\left[\left[\sum_{i=1}^{3}\omega_i(\eta_i + \epsilon_i)\right]^2 + \sum_{i=1}^{3}\omega_i(\eta_i + \epsilon_i)\left(1 - \sum_{i=1}^{3}\omega_i\right)(\mu_0 - \theta) + \left(1 - \sum_{i=1}^{3}\omega_i\right)^2(\mu_0 - \theta)^2\right] \\
&= \mathbb{E}\left[\left[\sum_{i=1}^{3}\omega_i(\eta_i + \epsilon_i)\right]^2 + \underbrace{\sum_{i=1}^{3}\omega_i(\eta_i + \epsilon_i)\left(1 - \sum_{i=1}^{3}\omega_i\right)(\mu_0 - \theta)}_{=0 \text{ in expectation}} + \left(1 - \sum_{i=1}^{3}\omega_i\right)^2(\mu_0 - \theta)^2\right] \\
&= \sum_{i=1}^{3}\omega_i^2\mathbb{E}\left[(\eta_i + \epsilon_i)^2\right] + \left(1 - \sum_{i=1}^{3}\omega_i\right)^2\mathbb{E}\left[(\mu_0 - \theta)^2\right] \\
&= \sum_{i=1}^{3}\omega_i^2\underbrace{\mathbb{E}\left[(\eta_i + \epsilon_i)^2\right]}_{\text{Var}(\eta_i + \epsilon_i)} + \left(1 - \sum_{i=1}^{3}\omega_i\right)^2\underbrace{\mathbb{E}\left[(\mu_0 - \theta)^2\right]}_{\text{Var}(\theta)} \\
&= \sum_{i=1}^{3}\omega_i^2\left(\frac{1}{\rho_i} + \frac{1}{r_i}\right) + \left(1 - \sum_{i=1}^{3}\omega_i\right)^2\sigma_\theta^2 \\
&= \sum_{i=1}^{3}\frac{\omega_i^2}{h_i} + \left(1 - \sum_{i=1}^{3}\omega_i\right)^2\sigma_\theta^2
\end{aligned}$$

Assuming a diffuse prior, that is $\sigma_\theta^2 = \infty$ then $h_0 = \frac{1}{\sigma_\theta^2} = 0$. Essentially, the precision of the

prior is nil due to the large variance. This assumption helps simplify the algebra of the problem. As a result, $\omega_0 = 0$ which, implies that $\sum_{i=1}^{2} \omega_i = 1$ as $\omega_0 = 1 - \sum_{i=1}^{2} \omega_i$.
This means that with the diffuse prior the quadratic loss function reduces to

$$\begin{aligned}
\mathbb{E}\left[(\theta - a^*)^2\right] &= \sum_{i=1}^{2} \frac{\omega_i^2}{h_i} + \left(1 - \sum_{i=1}^{2} \omega_i\right)^2 \sigma_\theta^2 \\
&= \sum_{i=1}^{2} \frac{\omega_i^2}{h_i} \\
&= \sum_{i=1}^{2} \frac{1}{h_i} \left(\frac{h_i}{\sum_{i=1}^{2} h_i}\right)^2 \\
&= \frac{\sum_{i=1}^{2} h_i}{\left(\sum_{i=1}^{2} h_i\right)^2} \\
&= \frac{1}{\sum_{i=1}^{2} h_i}
\end{aligned}$$

Yielding the following form for the expected utility of the DM

$$u_{dm} = \bar{u} - \mathbb{E}\left[(\theta - a^*)^2\right] = \bar{u} - \frac{1}{\sum_{i=1}^{2} h_i} \tag{3}$$

□

Proposition 2

Proof. The report of worker 1 can be written as

$$\begin{aligned}
r_{1,dm} &= \mathbb{E}\left[\theta | \{s_1, r_{2,1}\}\right] + \eta_1 \\
&= q_{1,1} s_1 + q_{1,2} r_{2,1} + \eta_1 \\
&= q_{1,1} (\theta + \epsilon_1) + q_{1,2} (\theta + \epsilon_2 + \eta_2) + \eta_1 \\
&= \underbrace{(q_{1,1} + q_{1,2})}_{=1} \theta + q_{1,1}\epsilon_1 + q_{1,2} (\epsilon_2 + \eta_2) + \eta_1 \\
&= \theta + q_{1,1}\epsilon_1 + q_{1,2} (\epsilon_2 + \eta_2) + \eta_1
\end{aligned}$$

This means that, given the independence of the error terms, the objective function can be re-arranged as follows

$$\begin{aligned}
\mathbb{E}\left[(\theta - a^*)^2\right] &= \mathbb{E}\left[(\theta - r_{1,dm})^2\right] \\
&= \mathbb{E}\left[(\theta - (\theta + q_{1,1}\epsilon_1 + q_{1,2} (\epsilon_2 + \eta_2) + \eta_1))^2\right] \\
&= \mathbb{E}\left[(q_{1,1}\epsilon_1)^2 + q_{1,2}^2 \left(\epsilon_2^2 + \eta_2^2\right) + (\eta_1)^2\right] \\
&= q_{1,1}^2 \frac{1}{\rho_1} + q_{1,2}^2 \left(\frac{1}{\rho_2} + \frac{\xi_2^2}{m_2}\right) + \frac{\xi_1^2}{m_1} \\
&= \left[\frac{\rho_1}{\rho_1 + \frac{m_2\rho_2}{m_2+\xi_2^2\rho_2}}\right]^2 \frac{1}{\rho_1} + \left[\frac{\frac{m_2\rho_2}{m_2+\xi_2^2\rho_2}}{\rho_1 + \frac{m_2\rho_2}{m_2+\xi_2^2\rho_2}}\right]^2 \left(\frac{m_2 + \xi_2^2\rho_2}{\rho_2 m_2}\right) + \frac{\xi_1^2}{m_1} \qquad (4) \\
&= \frac{\rho_1}{\left[\rho_1 + \frac{m_2\rho_2}{m_2+\xi_2^2\rho_2}\right]^2} + \frac{\frac{m_2\rho_2}{m_2+\xi_2^2\rho_2}}{\left[\rho_1 + \frac{m_2\rho_2}{m_2+\xi_2^2\rho_2}\right]^2} + \frac{\xi_1^2}{m_1} \\
&= \frac{1}{\rho_1 + \frac{m_2\rho_2}{m_2+\xi_2^2\rho_2}} + \frac{\xi_1^2}{m_1}
\end{aligned}$$

□

Proposition 3

Proof. Firstly, I will determine the optimal meeting time for the simple and complex structures. Then I will allow the organizations time to be unconstrained $\kappa \to \infty$ and compare the optimal actions of the two structures. I will show that in the limit they are identical and that for any bounded κ it must be the case that the simple performs better than the complex.

Simple structure. To determine the optimal allocation of meeting time the DM maximize their expected utility found in (3) subject to the meeting time constraint $\sum_{i=1}^{2} m_i \leq \bar{m}$ since the DM's constraint is the one that binds. As the utility is increasing in m_i, the meeting time is exhausted $\sum_{i=1}^{2} m_i = \bar{m}$. For now I will keep the communication and processing ability heterogeneous.
This is equivalent to the following maximization problem

$$\begin{aligned} \max_{\mathbf{r} \in \mathbb{R}} \quad & \sum_{i=1}^{2} h_i \\ \text{subject to} \quad & \sum_{i}^{2} m_i = \bar{m} \end{aligned}$$

Plugging in the constraint into the objective function to get it in terms of m_1 gives

$$V(m_1) = \frac{m_1 \rho_1}{m_1 + \rho_1 \xi_1^2} + \frac{(\kappa - m_1)\rho_2}{\kappa - m_1 + \xi_2^2 \rho_2}$$

The function is strictly concave by the second order condition

$$\frac{\partial^2 V(m_1)}{\partial m_1^2} = -2 \left[\frac{\xi_1^2 \rho_1^2}{(\xi_1^2 \rho_1 + m_1)^2} + \frac{\xi_2^2 \rho_2^2}{(\kappa - m_1 + \xi_1^2 \rho_2)^2} \right] < 0 \quad \forall m_1$$

This means that any critical point found in the following maximization problem will be at least a local maximum and there is at most one global maximum.
The Lagrangian is the following

$$\mathcal{L}(\mathbf{m}) = \sum_{i=1}^{2} h_i - \lambda(\sum_{i}^{2} m_i - \bar{m})$$

The first order condition with respect to m_i is as follows

$$\frac{\partial \mathcal{L}}{\partial m_i} = 0 \Leftrightarrow \frac{\partial h_i}{\partial m_i} = \lambda \qquad (5)$$

Where

$$
\begin{aligned}
\frac{\partial h_i}{\partial m_i} &= \frac{\partial \frac{\rho_i m_i}{\xi_i^2 \rho_i + m_i}}{\partial m_i} \\
&= \frac{\partial (\rho_i m_i)}{\partial m_i} \frac{1}{\xi_i^2 \rho_i + m_i} + \frac{\partial (\xi_i^2 \rho_i + m_i)^{-1}}{\partial m_i} \\
&= \frac{\rho_i}{\xi_i^2 \rho_i + m_i} - \frac{\rho_i m_i}{(\xi_i^2 \rho_i + m_i)^2} \\
&= \frac{\rho_i (\xi_i^2 \rho_i + m_i)}{(\xi_i^2 \rho_i + m_i)^2} - \frac{\rho_i m_i}{(\xi_i^2 \rho_i + m_i)^2} \\
&= \frac{\xi_i^2 \rho_i^2}{(\xi_i^2 \rho_i + m_i)^2} \\
&= \frac{m_i^2 \xi_i^2 \rho_i^2}{m_i^2 (\xi_i^2 \rho_i + m_i)^2} \\
&= \frac{h_i^2 \xi_i^2}{m_i^2}
\end{aligned}
$$

Thus from (5)

$$
\begin{aligned}
\frac{h_i^2 \xi_i^2}{m_i^2} &= \lambda \\
\frac{h_i^2 \xi_i^2}{\lambda} &= m_i^2 \\
m_i &= \frac{\xi_i h_i}{\sqrt{\lambda}} \\
m_i &= \frac{\xi_i \rho_i m_i}{(\xi_i^2 \rho_i + m_i)\sqrt{\lambda}} \\
1 &= \frac{\xi_i \rho_i}{(\xi_i^2 \rho_i + m_i)\sqrt{\lambda}} \\
(\xi_i^2 \rho_i + m_i)\sqrt{\lambda} &= \xi_i \rho_i \\
m_i \sqrt{\lambda} &= \xi_i \rho_i - \xi_i^2 \rho_i \sqrt{\lambda} \\
m_i &= \frac{\xi_i \rho_i (1 - \xi_i \sqrt{\lambda})}{\sqrt{\lambda}}
\end{aligned}
\tag{6}
$$

Using the above and $\sum_{i=1}^{2} m_i = \bar{m}$ to solve for λ in terms of the primitives of the model

$$\begin{aligned}
m_i &= \frac{\xi_i\rho_i(1-\xi_i\sqrt{\lambda})}{\sqrt{\lambda}} \\
\sum_{i=1}^{2} m_i &= \frac{\sum_{i=1}^{2}\xi_i\rho_i(1-\xi_i\sqrt{\lambda})}{\sqrt{\lambda}} \\
\bar{m} &= \frac{\sum_{i=1}^{2}\xi_i\rho_i(1-\xi_i\sqrt{\lambda})}{\sqrt{\lambda}} \\
\bar{m}\sqrt{\lambda} &= \sum_{i=1}^{2}\xi_i\rho_i(1-\xi_i\sqrt{\lambda}) \\
\bar{m}\sqrt{\lambda} + \sum_{i=1}^{2}\xi_i^2\rho_i\sqrt{\lambda} &= \sum_{i=1}^{2}\xi_i\rho_i \\
\sqrt{\lambda}\left(\bar{m} + \sum_{i=1}^{2}\xi_i^2\rho_i\right) &= \sum_{i=1}^{2}\xi_i\rho_i \\
\sqrt{\lambda} &= \frac{\sum_{i=1}^{2}\xi_i\rho_i}{\bar{m} + \sum_{i=1}^{2}\xi_i^2\rho_i}
\end{aligned}$$

Now plugging $\sqrt{\lambda}$ back into (6) to yield

$$\begin{aligned}
m_i^* &= \frac{\xi_i\rho_i}{\sqrt{\lambda}}(1-\xi_i\sqrt{\lambda}) \\
&= \xi_i\rho_i\frac{\left(\bar{m} + \sum_{i=1}^{2}\xi_i^2\rho_i\right)}{\sum_{i=1}^{2}\xi_i\rho_i}\left(1 - \frac{\xi_i\sum_{i=1}^{2}\xi_i\rho_i}{\bar{m} + \sum_{i=1}^{2}\xi_i^2\rho_i}\right) \\
&= \xi_i\rho_i\frac{\left(\bar{m} + \sum_{i=1}^{2}\xi_i^2\rho_i\right)}{\sum_{i=1}^{2}\xi_i\rho_i}\left(\frac{\bar{m} + \sum_{i=1}^{2}\xi_i^2\rho_i - \xi_i\sum_{i=1}^{2}\xi_i\rho_i}{\bar{m} + \sum_{i=1}^{2}\xi_i^2\rho_i}\right) \\
&= \xi_i\rho_i\frac{\left(\bar{m} + \sum_{i=1}^{2}\xi_i^2\rho_i - \xi_i\sum_{i=1}^{2}\xi_i\rho_i\right)}{\sum_{i=1}^{2}\xi_i\rho_i}
\end{aligned}$$

The optimal allocation of meeting time is thus

$$m_i^* = \xi_i\rho_i\frac{\left(\bar{m} + \sum_{i=1}^{2}\xi_i^2\rho_i - \xi_i\sum_{i=1}^{2}\xi_i\rho_i\right)}{\sum_{i=1}^{2}\xi_i\rho_i} \tag{7}$$

Given that $h_i = \frac{\rho_i m_i}{\xi_i^2\rho_i + m_i}$ this can be expressed in terms of the primitives of the model

$$\begin{aligned}
h_i^* &= \frac{\xi_i\rho_i^2}{\xi_i^2\rho_i + \xi_i\rho_i\frac{\left(\bar{m}+\sum_{i=1}^{2}\xi_i^2\rho_i - \xi_i\sum_{i=1}^{2}\xi_i\rho_i\right)}{\sum_{i=1}^{2}\xi_i\rho_i}}\frac{\left(\bar{m} + \sum_{i=1}^{2}\xi_i^2\rho_i - \xi_i\sum_{i=1}^{2}\xi_i\rho_i\right)}{\sum_{i=1}^{2}\xi_i\rho_i} \\
&= \frac{\xi_i\rho_i^2}{\xi_i^2\rho_i\sum_{i=1}^{2}\xi_i\rho_i + \xi_i\rho_i\left(\bar{m} + \sum_{i=1}^{2}\xi_i^2\rho_i - \xi_i\sum_{i=1}^{2}\xi_i\rho_i\right)}\left(\bar{m} + \sum_{i=1}^{2}\xi_i^2\rho_i - \xi_i\sum_{i=1}^{2}\xi_i\rho_i\right) \\
&= \frac{\xi_i\rho_i^2}{\xi_i^2\rho_i\sum_{i=1}^{2}\xi_i\rho_i + \xi_i\rho_i\bar{m} + \xi_i\rho_i\sum_{i=1}^{2}\xi_i^2\rho_i - \xi_i^2\rho_i\sum_{i=1}^{2}\xi_i\rho_i}\left(\bar{m} + \sum_{i=1}^{2}\xi_i^2\rho_i - \xi_i\sum_{i=1}^{2}\xi_i\rho_i\right) \\
&= \frac{\rho_i\left(\bar{m} + \sum_{i=1}^{2}\xi_i^2\rho_i - \xi_i\sum_{i=1}^{2}\xi_i\rho_i\right)}{\bar{m} + \sum_{i=1}^{2}\xi_i^2\rho_i}
\end{aligned}$$

Now as $\kappa \to \infty$, that is the workers have an unconstrained amount of time to meet since they are both meeting the DM in the simple structure the binding constraint has not changed it is still $\bar{m}$. Given that processing and communicating ability are assumed to be symmetric, $\rho_1 = \rho_2 = \rho$ and $\xi_1^2 = \xi_2^2 = \xi^2$ plugging this into (7) yields

$$\begin{aligned}
m_i^* &= \xi\rho \frac{\left(\bar{m} + \sum_{i=1}^2 \xi^2\rho - \xi \sum_{i=1}^2 \xi\rho\right)}{\sum_{i=1}^2 \xi\rho} \\
&= \xi\rho \frac{\left(\bar{m} + \xi^2\rho \sum_{i=1}^2 - \xi^2\rho \sum_{i=1}^2\right)}{\xi\rho \sum_{i=1}^2} \\
&= \xi\rho \frac{\left(\bar{m} + 2\xi^2\rho - 2\xi^2\rho\right)}{2\xi\rho} \\
&= \xi\rho \frac{(\bar{m})}{2\xi\rho} \\
&= \frac{\bar{m}}{2}
\end{aligned}$$

The DM optimally spends an equal amount of time meeting both workers. This leads to equal precision weights $h_1^* = h_2^*$ and so the DM equally weights both reports $\omega_1^* = \omega_2^* = 1/2$. The optimal action is given by the following

$$a^* = \frac{1}{2} r_1 + \frac{1}{2} r_2$$

plugging in $r_i = s_i + \eta_i$

$$a^* = \frac{1}{2} s_1 + \frac{1}{2} s_2 + \frac{1}{2} \eta_1 + \frac{1}{2} \eta_2$$

Given the linearity of a^* and that the error terms η are normally distributed. the above can be re-written as

$$\begin{aligned}
a^* &= \frac{1}{2} s_1 + \frac{1}{2} s_2 + \frac{1}{2} (\eta_1 + \eta_2) \\
&= \frac{1}{2} s_1 + \frac{1}{2} s_2 + \frac{1}{2} \left[\mathcal{N}\left(0, \frac{\xi^2}{m_1^*}\right) + \mathcal{N}\left(0, \frac{\xi^2}{m_2^*}\right)\right] \\
&= \frac{1}{2} s_1 + \frac{1}{2} s_2 + \frac{1}{2} \left[\mathcal{N}\left(0, \frac{\xi^2}{m_1^*} + \frac{\xi^2}{m_2^*}\right)\right] \\
&= \frac{1}{2} s_1 + \frac{1}{2} s_2 + \frac{1}{2} \left[\mathcal{N}\left(0, 2\frac{\xi^2}{\bar{m}} + 2\frac{\xi^2}{\bar{m}}\right)\right] \\
&= \frac{1}{2} s_1 + \frac{1}{2} s_2 + \frac{1}{2} \left[\mathcal{N}\left(0, 4\frac{\xi^2}{\bar{m}}\right)\right] \\
&= \frac{1}{2} s_1 + \frac{1}{2} s_2 + \left[\mathcal{N}\left(0, \frac{1}{4} 4\frac{\xi^2}{\bar{m}}\right)\right]
\end{aligned}$$

This gives the optimal action for the simple structure as $\kappa \to \infty$ as

$$a^* = \frac{1}{2} s_1 + \frac{1}{2} s_2 + \left[\mathcal{N}\left(0, \frac{\xi^2}{\bar{m}}\right)\right] \tag{8}$$

Where I have used the property that if a random variable X follows a standard normal distribution $\mathcal{N}(0, 1)$, then X times a scalar c has the following normal distribution $\mathcal{N}\left(0, c^2\right)$

Complex structure. For the complex structure since worker 2 does not meet the DM and only meets worker 1 this leads to the following constraints on meeting time.

$$\begin{aligned} &m_1 \leq \bar{m} && \text{DM time constraint} \\ &m_1 + m_2 \leq \kappa && \text{Worker 1 time constraint} \\ &m_2 \leq \kappa && \text{Worker 2 time constraint} \end{aligned}$$

The report produced from worker 1 and 2 meeting is the following

$$r_{2,1} = \theta + \left[\mathcal{N}\left(0, \frac{1}{\rho} + \frac{\xi^2}{m_2}\right)\right]$$

As $\kappa \to \infty$ the constraint of worker 2 becomes unbounded. This means that worker 2 can effectively perfectly communicate their signal to worker 1

$$m_2 \to \infty \Rightarrow \lim_{m_2 \to \infty} r_{2,1} = \theta + \left[\mathcal{N}\left(0, \frac{1}{\rho}\right)\right] = s_2$$

Worker 1 now has two pieces of information their own signal s_1 and worker 2's signal s_2. They calculate their posterior belief of θ. Since the processing ability of both workers is the same worker 1 equally weights both pieces of information

$$\mathbb{E}\left[\theta | \{s_1, s_2\}\right] = \frac{1}{2} s_1 + \frac{1}{2} s_2$$

Now the DM and worker 1 meet and produce a report, in which worker 1 communicates their information with noise

$$r_{1,dm} = \frac{1}{2} s_1 + \frac{1}{2} s_2 + \eta_1$$

Given the diffuse prior (the DM's belief is uninformative) the optimal action is for the DM to take an action according to the above report. The constraint on the DM is binding and so they optimally meet for as long as possible. Given the normally distributed error it can be written as follows

$$a^* = \frac{1}{2} s_1 + \frac{1}{2} s_2 + \left[\mathcal{N}\left(0, \frac{\xi^2}{\bar{m}}\right)\right]$$

Comparing the above with (8) it is clear to see that in the limit ($\kappa \to \infty$) that the optimal action is the same. This means that the payoffs in both structures will be equivalent. However, for any $\kappa < \infty$ there will clearly be a higher noise to signal ratio in the complex structure.

$$r_{1,dm} \neq s_2$$

Hence, the complex structure will perform worse to the simple structure and there is a clear preference for the simple structure. □

Proposition 4

Proof. **Simple structure.** The maximization problem for the simple structure remains the same as before. However now there is heterogeneous communication where worker 1 is better at communicating than worker $\xi_1^2 < \xi_2^2$. The optimal amount of meeting time is given below (from (7))

$$m_i^* = \xi_i \frac{\left(\bar{m} + \rho \sum_{i=1}^2 \xi_i^2 - \rho \xi_i \sum_{i=1}^2 \xi_i\right)}{\sum_{i=1}^2 \xi_i}$$

Plugging in $\xi_1^2 < \xi_2^2$ yields the following

$$m_1^* = \xi_1 \frac{\left(\bar{m} + \rho \xi_2 \left(\xi_2 - \xi_1\right)\right)}{\sum_{i=1}^2 \xi_i} \qquad m_2^* = \xi_2 \frac{\left(\bar{m} - \rho \xi_1 \left(\xi_2 - \xi_1\right)\right)}{\sum_{i=1}^2 \xi_i}$$

Hence $m_1^* \neq m_2^*$. from the above equations the optimal precision weights h_i^* are calculated and given below

$$h_1^* = \frac{\rho\left(\bar{m} + \rho\xi_2\left(\xi_2 - \xi_1\right)\right)}{\bar{m} + \rho\sum_{i=1}^{2}\xi_i^2} \qquad h_2^* = \frac{\rho\left(\bar{m} - \rho\xi_1\left(\xi_2 - \xi_1\right)\right)}{\bar{m} + \rho\sum_{i=1}^{2}\xi_i^2}$$

As worker 1 is a better communicator his report has a higher precision. Thus more weight is given to worker 1's report than to worker 2's report $\omega_1 > \omega_2$. Plugging the precision weights into the optimal utility given in proposition 1 gives

$$u_{dm} = \bar{u} - \frac{\bar{m} + \rho\sum_{i=1}^{2}\xi_i^2}{\rho\left(2\bar{m} + \rho\left(\xi_1 - \xi_2\right)^2\right)}$$

Complex structure. For the complex structure there are the following binding constraints

$$\begin{aligned} m_1 &= \bar{m} && \text{DM time constraint} \\ m_1 + m_2 &= \kappa && \text{Worker 1 time constraint} \end{aligned}$$

This is a system of equations with two equations and two unknowns, so there exists a unique solution given by

$$m_1^* = \bar{m} \qquad m_2^* = \kappa - \bar{m}$$

Plugging these into the optimal utility given in proposition 2 gives

$$\hat{u_{dm}} = \bar{u} - \frac{\bar{m}\left(\kappa - \bar{m} + \rho\xi_2^2\right) + \xi_1^2\rho\left(2(\kappa - \bar{m}) + \rho\xi_2^2\right)}{\bar{m}\rho\left(2(\kappa - \bar{m}) + \rho\xi_2^2\right)}$$

The point at which the two payoffs are equal $u_{dm} = \hat{u_{dm}}$ is given by the threshold $\bar{k}$

$$\bar{\kappa} = \frac{\bar{m}^2\left(3\xi_1^2 - 2\xi_1\xi_2 - 2\xi_2^2\right) + 2\bar{m}\rho\xi_1\left(\xi_1^3 - 2\xi_1^2\xi_2 - \xi_2^3\right) - \rho^2\xi_1^2\xi_2^2\left(\xi_1 - \xi_2\right)^2}{\left(\xi_1 - \xi_2\right)\left[\bar{m}\left(3\xi_1 + \xi_2\right) + 2\rho\xi_1^2\left(\xi_1 - \xi_2\right)\right]}$$

□

Proposition 5

Proof. The first section will correspond to homogeneous communication and the second section will correspond to heterogeneous communication.

A.1 ξ_i^2 is the same for all workers

Simple structure. Performing the same steps as I did for the proof of proposition 1 but with n workers will lead to the following expected utility. Where h_i corresponds to a precision weight

$$u_{dm} = \bar{u} - \mathbb{E}\left[(\theta - a^*)^2\right] = \bar{u} - \frac{1}{\sum_{i=1}^{n} h_i} \tag{9}$$

The higher the precision of each report the smaller the quadratic loss and the better the performance of the firm in estimating θ.

As $\kappa \to \infty$, the workers have an unconstrained amount time but the binding constraint is still $\bar{m}$. Given that processing and communicating ability are assumed to be symmetric, it will be optimal for the DM to meet each worker for the same amount of time

$$m_i^* = \frac{\bar{m}}{n} \ \forall \ i$$

This leads to equal precision weights $h_i^* = h$ and so the DM equally weights all the reports $\omega_i^* = 1/n$. The optimal action is given by the following

$$a^* = \frac{1}{n}\sum_{i=1}^{n} r_i$$

plugging in $r_i = s_i + \eta_i$

$$a^* = \frac{1}{n}\sum_{i=1}^{n} s_i + \frac{1}{n}\sum_{i=1}^{n} \eta_i$$

Given the linearity of a^* and that the error terms η are normally distributed. the above can be re-written as

$$\begin{aligned}
a^* &= \frac{1}{n}\sum_{i=1}^{n} s_i + \frac{1}{n}\left(\sum_{i=1}^{n} \eta_i\right) \\
&= \frac{1}{n}\sum_{i=1}^{n} s_i + \frac{1}{n}\left[\mathcal{N}\left(0, \sum_{i=1}^{n} \frac{\xi^2}{m_i^*}\right)\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} s_i + \frac{1}{n}\left[\mathcal{N}\left(0, \sum_{i=1}^{n} n\frac{\xi^2}{\bar{m}}\right)\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} s_i + \frac{1}{n}\left[\mathcal{N}\left(0, n^2\frac{\xi^2}{\bar{m}}\right)\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} s_i + \left[\mathcal{N}\left(0, \frac{1}{n^2}n^2\frac{\xi^2}{\bar{m}}\right)\right]
\end{aligned}$$

This gives the optimal action for the simple structure as $\kappa \to \infty$ as

$$a^* = \frac{1}{n}\sum_{i=1}^{n} s_i + \left[\mathcal{N}\left(0, \frac{\xi^2}{\bar{m}}\right)\right] \tag{10}$$

Where I have used the property that if a random variable X follows a standard normal distribution $\mathcal{N}(0,1)$, then X times a scalar c has the following normal distribution $\mathcal{N}\left(0, c^2\right)$

Complex structure. For the complex structure since worker $i \neq 1$ does not meet the DM and only meets worker 1 this leads to the following constraints on meeting time.

$$\begin{aligned}
&m_1 \leq \bar{m} && \text{DM time constraint} \\
&m_1 + \sum_{i=2}^{n} m_i \leq \kappa && \text{Worker 1 time constraint} \\
&m_i \leq \kappa && \text{Worker } i \text{ time constraint for } i \geq 2
\end{aligned}$$

The report produced from worker 1 and i meeting is the following

$$r_{i,1} = \theta + \left[\mathcal{N}\left(0, \frac{1}{\rho} + \frac{\xi^2}{m_i}\right)\right]$$

Clearly as $\kappa \to \infty$ the constraint of worker i becomes unbounded. This means that worker i can perfectly communicate their signal to worker 1

$$m_i \to \infty \Rightarrow \lim_{m_i \to \infty} r_{i,1} = \theta + \left[\mathcal{N}\left(0, \frac{1}{\rho}\right)\right] = s_i \;\forall\; i \neq 1$$

After worker 1 meets all the other worker he has n pieces of information their own signal s_1 and the signals of all the other workers. They calculate their posterior belief of θ. Clearly since the processing ability of all workers is the same worker 1 equally weights all pieces of information

$$\mathbb{E}\left[\theta|\{s_1, s_2\}\right] = \frac{1}{n}\sum_{i=1}^{n} s_i$$

Now the DM and worker 1 meet and produce a report to which worker 1 communicates their information with noise

$$r_{1,dm} = \frac{1}{n}\sum_{i=1}^{n} s_i + \eta_1$$

Given the diffuse prior (the DM's belief is uninformative) the optimal action is for the DM to take an action according to the above report. The constraint on the DM is binding and so they optimally meet for as long as possible. Given the normally distributed error it can be written as follows

$$a^* = \frac{1}{n}\sum_{i=1}^{n} s_i + \left[\mathcal{N}\left(0, \frac{\xi^2}{\bar{m}}\right)\right]$$

as with the 3 agent case it is clear to see that in the limit ($\kappa \to \infty$) that the optimal action is the same in both structures. This means that the payoffs in both structures will be equivalent. However, for any $\kappa < \infty$ there will clearly be a higher noise to signal ratio in the complex structure.

$$r_{1,dm} \neq s_2$$

Hence, the complex structure will perform worse to the simple structure and there is a clear preference for the simple structure given homogeneous communication abilities.

A.2 $\xi_1^2 < \xi^2$

Simple structure. For the simple structure the optimal precision weight is derived from the maximization problem in the proof of proposition 3

$$h_i^* = \frac{\rho\left(\bar{m} + \rho\sum_{i=1}^{n}\xi_i^2 - \rho\xi_i\sum_{i=1}^{n}\xi_i\right)}{\bar{m} + \rho\sum_{i=1}^{n}\xi_i^2}$$

Given that all workers $i \in \{2, \ldots, n\}$ have the same communication ability ξ^2 and $\xi_1^2 < \xi^2$. This means that the optimal precision becomes

$$h_1^* = \frac{\rho\left(\bar{m} + \rho\left(n-1\right)\xi\left(\xi - \xi_1\right)\right)}{\bar{m} + \rho\left(\xi_1^2 + (n-1)\xi^2\right)},\ h_i^* = \frac{\rho\left(\bar{m} + \rho\xi_1\left(\xi_1 - \xi\right)\right)}{\bar{m} + \rho\left(\xi_1^2 + (n-1)\xi^2\right)}\ \forall\, i \neq 1$$

Plugging this into $u_{dm} = \bar{m} - 1/\sum h_i$ after some algebra gives an optimal payoff of

$$u_{dm} = \bar{u} - \frac{\bar{m} + \rho\left(\xi_1^2 + (n-1)\xi^2\right)}{\rho\left(n\bar{m} + \rho(n-1)\left(\xi_1 - \xi\right)^2\right)}$$

Complex structure. For the complex structure following similar steps to the proof for proposition 2 the report between worker 1 and the DM will be

$$\begin{aligned} r_{1,dm} &= \mathbb{E}\left[\theta|\{s_1, r_{2,1}, r_{3,1}, \ldots, r_{n,1}\}\right] + \eta_1 \\ &= \theta + q_{1,1}\epsilon_1 + \sum_{i=2}^{n} q_{1,i}\left(\epsilon_i + \eta_i\right) + \eta_1 \end{aligned}$$

This means that, given the independence of the error terms, the objective function can be re-arranged as follows

$$
\begin{aligned}
\mathbb{E}\left[(\theta-a^*)^2\right] &= \mathbb{E}\left[(\theta-r_{1,dm})^2\right] \\
&= \left[\frac{\rho}{\rho+\sum_{i=2}^{n}\frac{m_i\rho}{m_i+\xi_i^2\rho}}\right]^2\frac{1}{\rho}+\left[\frac{\frac{m_2\rho}{m_2+\xi_2^2\rho}}{\rho+\sum_{i=2}^{n}\frac{m_i\rho}{m_i+\xi_i^2\rho}}\right]^2\left(\frac{m_2+\xi_2^2\rho}{\rho m_2}\right)+ \\
&\quad \ldots+\left[\frac{\frac{m_n\rho}{m_n+\xi_n^2\rho n}}{\rho+\sum_{i=2}^{n}\frac{m_i\rho}{m_i+\xi_i^2\rho}}\right]^2\left(\frac{m_n+\xi_n^2\rho}{\rho m_n}\right)+\frac{\xi_1^2}{m_1} \\
&= \frac{1}{\rho+\sum_{i=2}^{n}\frac{m_i\rho}{m_i+\xi_i^2\rho}}+\frac{\xi_1^2}{m_1}
\end{aligned}
$$

The relevant constraints are the following

$$
m_1=\bar{m}
$$

$$
\sum_{i=1}^{n} m_i=\kappa
$$

Plugging the first constraint into the second means that $m_2+\ldots+m_n=\kappa-\bar{m}$. From the perspective of worker 1, they have $n-1$ meetings with workers who all have the same communication ability ξ^2. This means it is optimal for him spend an equal amount of time with each. That is $m_i=m\ \forall\ i\neq 1$. Hence

$$
m_i=\frac{\kappa-\bar{m}}{n-1}\ \forall\ i\neq 1
$$

plugging this into the payoff of the DM

$$
\begin{aligned}
\hat{u_{dm}} &= \bar{u}-\frac{1}{\rho+\sum_{i=2}^{n}\frac{\frac{\kappa-\bar{m}}{n-1}\rho}{\frac{\kappa-\bar{m}}{n-1}+\xi^2\rho}}-\frac{\xi_1^2}{\bar{m}} \\
&= \bar{u}-\frac{1}{\rho+(n-1)\frac{\frac{\kappa-\bar{m}}{n-1}\rho}{\frac{\kappa-\bar{m}}{n-1}+\xi^2\rho}}-\frac{\xi_1^2}{\bar{m}} \\
&= \bar{u}-\frac{1}{\rho+\frac{(\kappa-\bar{m})\rho(n-1)}{\kappa-\bar{m}+\xi^2\rho(n-1)}}-\frac{\xi_1^2}{\bar{m}} \\
&= \bar{u}-\frac{\kappa-\bar{m}+\xi^2\rho\,(n-1)}{\rho\,(\kappa-\bar{m}+\xi^2\rho\,(n-1))+(\kappa-\bar{m})\,\rho\,(n-1)}-\frac{\xi_1^2}{\bar{m}} \\
&= \bar{u}-\frac{\kappa-\bar{m}+\xi^2\rho\,(n-1)}{\rho\,(\xi^2\rho\,(n-1)+n\,(\kappa-\bar{m}))}-\frac{\xi_1^2}{\bar{m}} \\
&= \bar{u}-\frac{\bar{m}\left(\kappa-\bar{m}+\xi^2\rho\,(n-1)\right)+\xi_1^2\rho\left(\xi^2\rho\,(n-1)+n\,(\kappa-\bar{m})\right)}{\bar{m}\rho\,(\xi^2\rho\,(n-1)+n\,(\kappa-\bar{m}))}
\end{aligned}
$$

Equating the two payoffs I can use numerical solver such as mathematica to find the threshold as a function of the parameters.

□