Should a Team Be Homogeneous?

Andrea Prat*
London School of Economics, Tilburg University, and CEPR†

August 16, 2000

Abstract

Should an organization hire people with similar backgrounds or with different backgrounds? We formulate this question within the framework of team theory. The team is formed by n agents. The type of each agent is endogenous and determines his information structure and his cost for the team. We show that the sign of complementarity between jobs determines workforce homogeneity. With positive complementarities, the team should be composed of agents of the same type, while, with negative complementarities, workforce heterogeneity is optimal. These results do not rely on restrictions on the way uncertainty is modeled or on the feasible set of agent types: they can be explained in terms of correlation between errors committed by different agents.

Keywords: workforce homogeneity, team theory.

JEL Classification: D21, L20, M12.

1 Introduction

The backgrounds of the people who work in an organization are not exogenously given. The organization (be it a firm, a government agency, a

*This is a revised version of a chapter of my Ph.D. dissertation. I have greatly benefited from discussions with Masahiko Aoki, Kenneth Arrow, and Dimitri Vayanos. I thank Jacques Crémer, Mathias Dewatripont, Joshua Gans, Luis Garicano, Thomas Hellmann, Marcos Lisboa, Paul Milgrom, Dolf Talman, Stef Tijs and audiences at Bocconi, IAE (Barcelona), Pompeu Fabra, Rochester, Stanford, Stockholm School of Economics, Texas A&M, Tilburg, and Toulouse (ESEM 1997) for their helpful comments. I am also grateful to an editor and two anonymous referee for their useful suggestions.

†Address: STICERD, LSE, Houghton Street, London WC2A 2AE, UK; Email: a.prat@lse.ac.uk. Home page: http://econ.lse.ac.uk/staff/prat.

nonprofit, etc.) chooses whom to hire. This paper is concerned with one dimension of the hiring policy: the degree of variety in the backgrounds of the people who are hired.

On this dimension, different organizations can adopt strikingly different policies. Some organizations pursue a policy of hiring people with homogeneous backgrounds, while other organizations actively seek a degree of background diversity in their workforce. What characteristics of an organization determine its optimal degree of background homogeneity? Clearly, this is a complex question and one could approach it from several angles. Motivational factors, incentive issues, the need for secrecy all play a role in this choice. This paper, however, will restrict its attention to one factor: informational efficiency.

Even when examined in isolation, informational efficiency constitutes a highly complex problem. Crémer [8, 9] has pioneered the idea of applying *team theory* to issues of corporate culture. Team theory, developed by Marschak and Radner [15], examines the problem of several decision makers who maximize a joint objective but have different information when they choose their actions. The agents share a common prior but may receive different signals about the state of the world. The challenging aspect is that each agent makes his decision without knowing the information of the other agents, and thus uses his own information to infer both the state of the world and the information received by the others. For a given information structure, a team problem can be seen as a game of incomplete information in which players have the same payoffs. The solution to the team problem for a fixed information structure (a decision rule for each agent) corresponds to a Bayes-Nash equilibrium of the game, with the restriction that the equilibrium must be Pareto efficient (that is, there are no coordination problems). However, the question that team theory usually asks is not only what the optimal decision rules are for a given information structure but also how the information structure of the team affects its performance.

This paper uses team theory to model the problem of workforce heterogeneity. The section on Related Literature will describe the precise contribution of this paper with respect to the existing team-theoretical literature. The remainder of this section summarizes the results of the paper in simple terms.

An organization is made of n agents (n is given exogenously). The type of an agent, which is endogenous, determines his information structure – the most important concept of this paper. The information structure is the grid through which the agent observes the world. Mathematically, it is a mapping from the state of the world (a random variable not directly observed) into a signal available to the agent. For instance, an agent with the type "doctor" has an information structure which, when confronted with a patient, provides him with a signal on the patient's health. The type of an agent also determines the cost of that agent for the team. The endogeneity of agents' types has two interpretations. The direct one is that the team spends resources to endow the existing agents with information structures. A less direct interpretation is that the team is made of n slots that are filled by hiring agents from a labor market. The cost of an information structure is the market wage of an agent with that information structure.

Given the information structures, each agent is given a decision function, that tells the agent how to respond to each signal he may receive.[1] Of course, the decision function can vary from agent to agent. When the state of the world is realized, agents observe their signals through their information structures and choose their actions through their decision functions. The gross payoff to the organization is a function of the state of the world and of the actions taken by the agents. The team problem consists of selecting a type and a decision function for each of the n agents in order to maximize the expected value of the gross payoff minus the cost of information structures. Figure 1 depicts the problem for an organization in which $n = 2$.

We assume that the payoff function is symmetric in the agents' actions. This means that the total payoff does not depend on the job labels the agents carry but only on the actions they choose. With a sport analogy, this assumption implies that the number on a player's shirt is immaterial in determining the outcome of the player's actions (true in basketball; false in soccer because of the goalkeeper's special status). An intrinsically asymmetric payoff function would clearly bias results in favor of heterogeneous agent types.[2]

With this assumption, we prove the two central results of the paper. First, if the agents' actions are complements in the payoff function, then the set of optimal solutions contains a solution in which all agents have the same type. In that case, the team problem reduces to looking at configurations with only one type of agents. Second, if two additional assumptions hold (concavity and nonuniqueness of the single-type optima), we prove that,

[1] In contrast to most of the recent economic literature on organizations, this paper is not directly concerned with incentive issues. We assume that there is no moral hazard on the part of agents: if an agent is given a decision function, he will follow it.

[2] An earlier version of this paper shows that the main results hold also in a particular asymmetric environment (Prat [22]).

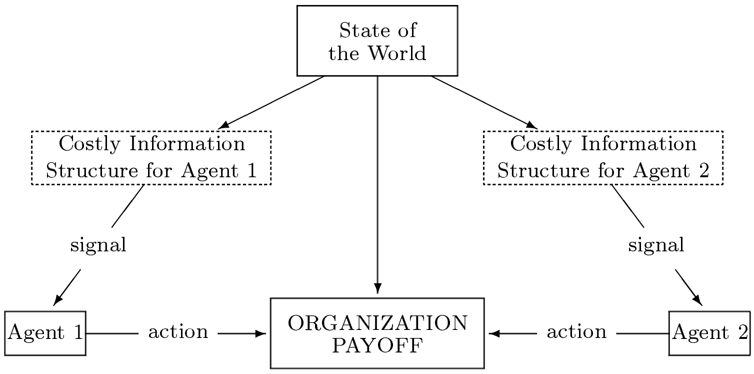

Figure 1: An Organization with Two Agents

if the agents' actions are substitutes in the payoff function, then the set of optimal solution contains a solution with at least two distinct types of agents.

Complementarities are represented through the lattice-theoretic notions of supermodular and submodular functions (Topkis [25], Vives [27], Milgrom and Roberts [17]). The main advantage of lattice theory is that its definitions are applicable also when continuity and differentiability are not assured, which is the case in this paper. Moreover, proofs are made simpler and statements more intuitive. Indeed, the intuition for the main results for this paper is captured by the following lattice-theoretic result, also proven in the paper. Consider a function the argument of which is a vector of random variables. We show that if the function is supermodular, then the expected value of the function is higher if the random variables are perfectly correlated rather than stochastically independent. On the other hand, if the function is submodular, the expected value is higher if the random variables are stochastically independent. In the problem at hand, agents do not in general have perfect information. Thus, they are bound to deviate from the full-information solution. If their actions are complements, it is optimal for them to deviate in a coordinated manner, which occurs if their information structures are identical. Thus, having agents of the same type is optimal. An analogous line of reasoning can be followed when the agents' actions are substitutes.

As the intuition is general, one would expect the results to hold in a very general setting. Indeed they do. In this paper, thanks to the use of lattice theory, no particular functional form is assumed for the payoff function. Moreover, the set of possible states of the world and the probability distribution on it are defined in a general way. Finally, no assumption is made on the set of agents' type, on information structures, or on decision functions.

The results are illustrated by two examples. In the first, a product is made of two components produced by two different divisions of the same firm. We show that the presence of complementarities in the production function pushes the team to have division managers with the same type. The second example is a search problem with two researchers choosing in which direction to devote their search efforts. We show that the team objective is submodular and hence researchers should have different types.

The plan of the paper is as follows. Section 2 introduces the model. Section 3 reports the main results. Section 4 provides intuition for the main results based on the idea of error correlation. Section 5 concludes.

Related Literature As stated above, we adopt the team-theoretical framework. Team theory was developed in the Sixties by Marschak and Radner [15]. In the Seventies it gave rise to a literature on the possibility of decentralizing decision-making, such as Groves and Radner [12] and Arrow and Radner [4]. After more than a decade of limited use – which coincided with the development of principal-agent theory – team theory has been experiencing a renewed interest. Several authors have applied it to problems in organization theory that do not seem to find a satisfactory answer within the principal-agent framework. Examples are Aoki [1, 2], Crémer [7, 8, 9], Geanakoplos and Milgrom [11], Li [14], Ponssard, Steinmetz, and Tanguy [20], and Qian, Roland, and Xu [23, 24].

The closest source of inspiration for this paper is Crémer [8, 9], who uses team theory to study corporate culture. Crémer considers a team with a quadratic objective function. The coefficient of the linear term is unknown and represents the state of the world. The state of the world is a normally distributed random variable. Each agent observes the state of the world plus a normally distributed disturbance. Crémer considers two cases: (1) the disturbances are identical across agents (shared knowledge) and (2) the disturbances are uncorrelated across agents (diversified knowledge). He finds that if the sign of the second-order cross-derivative in the agent actions is positive, then shared knowledge is optimal, while if the sign is negative than diversified knowledge is optimal.

The original contribution of the present paper is to extend Crémer's problem beyond a particular formulation. This generalization is of interest in itself, but is especially valuable because it allows us to identify the sign of complementarities as the main driving force in the choice between a homogeneous and a heterogeneous workforce. This finding appears to be new in economics and management science.[3]

The assumptions made by Crémer – quadratic payoff function and normally distributed signals – are common to all the recent team-theoretical literature. Although those assumptions are quite restrictive, they are made because they allow for a closed-form solution. An incidental contribution of the present paper paper is to show that those assumptions are not needed if team theory is combined with lattice theory. We formulate an organizational problem in a general way. Finding a closed-form solution is impossible and is not attempted. Instead, we apply lattice theory concepts and we study the set of optimal solution. This is sufficient to answer the question we are interested in and to generate testable implications. Hopefully, this methodology can be used to study other questions in organization theory that are still open.[4]

2 The Model of a Team

The state of the world x belongs to a finite set X. The state of the world is not observed directly. Every agent has common prior distribution $\phi : X \to \Re$.

A team is composed of n agents and its payoff depends on the decisions taken by the agents and on the state of the world. The payoff function is

[3] A note by Prat [21] presents some results on complementarity and workforce homogeneity. However, that work does not consider agent types but only information structures and therefore cannot directly relate to workforce homogeneity. Moreover, the scope of the present work is much broader.

[4] This paper differs from the industrial organization literature on information sharing in oligopoly, like Gal-Or [10] and Vives [26]. Those works examine the incentives of oligopolists to communicate to each other the private signals they have received. The choice between sharing or not sharing information is dictated by strategic consideration. In contrast, there are no strategic considerations in the present paper. As agents do not have conflicting interests, they *always* have an incentive to share information with each other.

given by

$$\omega(a_1, \ldots, a_n, x) \tag{1}$$

where $a_i \in A$ represents the action taken by agent i. A is an ordered set. No assumption is made on the form of ω except the following:

Assumption 1 (Symmetry) *For all $x \in X$,*

$$\omega(a_1, \ldots, a_n, x) = \cdots = \omega(a_n, \ldots, a_1, x) \tag{2}$$

for any rearrangement of the action indices.

Agent i is endowed with an information structure $\theta_i \in \Theta$, where $\theta : X \to Y$. We will often refer to an agent's information structure as his "type". The assumption that Y does not depend on θ is without loss of generality.[5] The function θ_i induces a partition $\mathcal{P}_\theta$ on the set X. This corresponds to the standard definition of information structure (see Marschak and Radner [15, p. 48-49]).

Information structures are costly. Endowing an agent with structure θ costs the team c_θ, so that the total cost incurred is $\sum_{i=1}^{n} c_{\theta_i}$.

We denote Agent i's decision function by $\alpha_i : Y \to A$. The decision function must be taken from a set of feasible decision functions denoted by $\mathcal{A}$. Then, the action taken by i in state x is $a_i = \alpha_i[\theta_i(x)]$.

To summarize, the givens of the *team problem* are: a set of states of the world (X), the prior distribution (ϕ), the payoff function (ω), a set of actions (A), a set of information structures ($\mathcal{H}$), a set of agent types (Θ), a set of decision functions ($\mathcal{A}$), and the cost function (c). The team must select, for each agent i, an information structure θ_i and a decision function α_i (which we will refer to as the team's *configuration*). The goal of the team is to maximize the expected payoff less the cost of information structures:

$$\max_{\{\theta_i \in \Theta, \alpha_i \in \mathcal{A}\}_{i=1,\ldots,n}} E\{\omega[\alpha_1(\theta_1(x)), \ldots, \alpha_n(\theta_n(x)), x]\} - \sum_{i=1}^{n} c_{\theta_i}$$

In the remainder of the paper, we assume that the team problem has at least one solution.

[5] Suppose that Y_θ denotes the set of possible signals received by an agent of type θ. We just need to let $Y = \bigcup_{\theta \in \Theta} Y_\theta$.

3 General Results

3.1 Defining Complementarities

To represent complementarities, we adapt the general definition of supermodular and submodular functions to the problem at hand (See for instance Milgrom and Shannon [19] for general definitions):

Definition 1 *The payoff function ω is supermodular in the agents' actions if, for any two vectors $(\hat{a}_1, \ldots, \hat{a}_n) \in A^n$ and $(\check{a}_1, \ldots, \check{a}_n) \in A^n$ and for all $x \in X$, the following holds*

$$\begin{aligned} &\omega(\hat{a}_1, \ldots, \hat{a}_n, x) + \omega(\check{a}_1, \ldots, \check{a}_n, x) \\ &\leq \omega[\min(\hat{a}_1, \check{a}_1), \ldots, \min(\hat{a}_n, \check{a}_n), x] + \omega[\max(\hat{a}_1, \check{a}_1), \ldots, \max(\hat{a}_n, \check{a}_n), x]. \end{aligned}$$

Conversely, ω is submodular in the agents' actions if, given any two vectors $(\hat{a}_1, \ldots, \hat{a}_n) \in A^n$ and $(\check{a}_1, \ldots, \check{a}_n) \in A^n$, for all $x \in X$, the following holds

$$\begin{aligned} &\omega(\hat{a}_1, \ldots, \hat{a}_n, x) + \omega(\check{a}_1, \ldots, \check{a}_n, x) \\ &\geq \omega[\min(\hat{a}_1, \check{a}_1), \ldots, \min(\hat{a}_n, \check{a}_n), x] + \omega[\max(\hat{a}_1, \check{a}_1), \ldots, \max(\hat{a}_n, \check{a}_n), x]. \end{aligned}$$

How does supermodularity relate to the notion of complementarity based on cross-derivatives? The latter definition is applicable only if the function is twice-differentiable. Topkis [25, Th. 3.2] shows that, if a function is twice-differentiable, then the function is supermodular if and only if the second-order cross derivatives are all nonnegative (in the present case: $\partial^2 \omega / \partial a_i \partial a_j \geq 0$ for $i \neq j$), while the function is submodular if and only if the cross derivatives are all nonpositive. Thus, supermodularity is a generalization of the traditional notion of complementarity. Its use derives from the fact that in many problems the second-order cross derivative is not well-defined (for instance, because the agent's action is a discrete variable).

The following result (proven in the appendix) will be used repeatedly in the paper.

Lemma 1 *Given an ordered set A and a vector $a = (a_1, \ldots, a_n) \in A^n$, define $P(a)$ as the set of all vectors obtained by permuting elements of a. Consider $f : A^n \to \Re$. If f is supermodular (submodular), then*

$$\frac{1}{n!} \sum_{(p_1, \ldots, p_n) \in P(a)} f(p_1, \ldots, p_n) \leq (\geq) \frac{1}{n} \sum_{i=1, \ldots, n} f(a_i, \ldots, a_i)$$

Consider a function, the arguments of which are all defined on the same ordered set. Take a particular vector of arguments. If the function is supermodular, then the average value of the function for all possible permutations of the vector is smaller or equal to the average value of the function of vectors in which all elements are equal to one of the arguments of the initial vector. The converse holds if the function is submodular.

The following is immediate:[6]

Corollary 1 *Suppose f is such that $f(a_1, \ldots, a_n) = \cdots = f(a_n, \ldots, a_1)$ for any rearrangement of the vector a. If f is supermodular (submodular), then*

$$f(a_1, \ldots, a_n) \leq (\geq) \frac{1}{n} \sum_{i=1,\ldots,n} f(a_i, \ldots, a_i)$$

3.2 Sufficient Condition for the Optimality of Workforce Homogeneity

It is now possible to state the main result of this paper. Provided that the team's problem has at least one optimal solution and provided that Assumption 1 holds, we have the following:

Proposition 1 *If ω is supermodular in the agents' actions, then the set of optimal solutions to the team problem contains at least one configuration in which $\theta_1 = \ldots = \theta_n$.*

Proof Suppose ω is supermodular in the agents' actions. We will prove that, for any configuration in which not all agents have the same type, we can find a configuration which gives a greater or equal expected payoff and in which all agents have the same type.

Consider a feasible choice of types, denoted by $\theta_1, \ldots, \theta_n$, and a feasible choice of decision functions, denoted by $\alpha_1(\cdot), \ldots, \alpha_n(\cdot)$. Notice that

$$\omega[\alpha_1(\theta_1(x)), \ldots, \alpha_n(\theta_n(x)), x]$$

[6] Corollary 1 has been proven in a direct way by Meyer and Mookherjee [16, Proposition 1]. To the best of my knowledge, Lemma 1 – which is of independent interest – is new. The Corollary is sufficient to prove the results in Sections 3.2 and 3.3. The Lemma is needed for Section 4.

is supermodular in $\alpha_1(\theta_1(x)), \ldots, \alpha_n(\theta_n(x))$ for all $x \in X$. By Corollary 1 and Assumption 1, we have that, for all $x \in X$,

$$\begin{aligned}
&\frac{1}{n}\{\omega[\alpha_1(\theta_1(x)), \ldots, \alpha_1(\theta_1(x)), x] \\
&+ \ldots + \omega[\alpha_n(\theta_n(x)), \ldots, \alpha_n(\theta_n(x)), x]\} \\
&\geq \omega[\alpha_1(\theta_1(x)), \ldots, \alpha_n(\theta_n(x)), x]
\end{aligned}$$

Take expectations over the set of possible states and add the total cost of wages on both sides

$$\begin{aligned}
&\frac{1}{n}(E\{\omega[\alpha_1(\theta_1(x)), \ldots, \alpha_1(\theta_1(x)), x]\} + nc_{\theta_1} \\
&+ \ldots + E\{\omega[\alpha_n(\theta_n(x)), \ldots, \alpha_n(\theta_n(x)), x]\} + nc_{\theta_n}) \\
&\geq E\{\omega[\alpha_1(\theta_1(x)), \ldots, \alpha_n(\theta_n(x)), x]\} + \sum_{i=1}^{n} c_{\theta_i}
\end{aligned}$$

Then for at least one of the types, say θ_k,

$$\begin{aligned}
&E\{\omega[\alpha_k(\theta_k(x)), \ldots, \alpha_k(\theta_k(x)), x]\} + nc_{\theta_k} \\
&\geq E\{\omega[\alpha_1(\theta_1(x)), \ldots, \alpha_n(\theta_n(x)), x]\} + \sum_{i=1}^{n} c_{\theta_i}
\end{aligned}$$

Thus, for any feasible choice of types and of decision functions, there exists a feasible choice of types and decision functions in which $\theta_1 = \ldots = \theta_n$ who does at least as well. Then, if the set of solutions to the team problem is not empty, as we have assumed, it must contain a solution in which $\theta_1 = \ldots = \theta_n$.

The intuition behind Proposition 1 is that if the payoff function is supermodular then the team is better off if agents commit correlated errors rather than uncorrelated errors. We will explore this theme in Section 4.[7]

Proposition 1 says that, among all the possible solutions to the team's problem, there is at least one in which all agents have the same type. It

[7] Suppose instead that the payoff function is *strictly* supermodular, that is, that '$<$' replaces '$\leq$' in the definition of supermodular function given in Section 4. Then, it is easy to see that, for any solution in which $\theta_1 = \ldots = \theta_n$ does not hold, it is almost always possible to find at least one solution in which $\theta_1 = \ldots = \theta_n$ and which yields a strictly higher expected payoff. Then, the set of optimal solutions will generically contain *only* configurations in which $\theta_1 = \ldots = \theta_n$.

does not identify which, among all the solutions of that kind, is the optimal one. However, a proposition like 1 greatly simplifies the task of designing an optimal organizational structure. The organization designer can, without loss of generality, focus on solutions in which all agents have the same type. For example, if there are 20 agents and 4 information structures, the number of possible configurations – counting symmetric structures only once – is 8855. However, by applying Proposition 1, the organization designer knows that the number of configurations she needs to check is just 4.

The following example illustrates the use of Proposition 1:

Example 1 (A Product Made of Two Components): Consider a firm made of two divisions. Agent 1 is the manager of Division 1 and Agent 2 is the manager of Division 2. The final product of the firm is obtained by assembling a component produced by Division 1 and a component produced by Division 2. Agent 1 decides a_1, the quantity produced by Division 1, and Agent 2 decides a_2, the quantity produced by Division 2. Because each product needs both components, the number of items produced is $\min(a_1, a_2)$. The firm faces an inelastic demand curve. It can sell up to x products at a unit price p. If it produces more than x products, the excess will be unsold. Therefore, the number of products sold is the minimum between the number of products produced and the number of products demanded: $\min(a_1, a_2, x)$. Demand depends on the state of the world represented by the real random variable x with a given probability distribution $p(x)$. The unit cost is the same for both components: k (let us assume that $k < 0.5p$). The payoff function of the firm is

$$\omega(a_1, a_2, x) = p \min(a_1, a_2, x) - k(a_1 + a_2) \tag{3}$$

The structure of the problem suggests that the agents' actions are complements. Indeed, it can be verified that the payoff function (3) is supermodular in a_1 and a_2 for all x (See the Appendix for a formal verification). Therefore, we can apply Proposition 1: the set of optimal solutions contains at least one solution in which Agent 1 and Agent 2 have the same information structure. This result is independent of the distribution of x and of the feasible set of types.

3.3 Sufficient Condition for the Optimality of Workforce Heterogeneity

This subsection presents a partial parallel of Proposition 1 for workforce heterogeneity. Two additional assumptions are necessary. First,

Assumption 2 (Concavity of the Payoff Function) *The set A is convex and the payoff function ω is concave in $a_1, a_2, \ldots, a_n$.*

Assumption 2 guarantees that the team is risk-averse. A risk-loving manager may want to hire homogeneous agents in order to coordinate on riskier actions.

The second additional assumption excludes situations in which one type of worker is superior to all other types:

Definition 2 *A one-type optimum is a solution to*

$$\max_{\theta^*, \{\alpha_i^*\}_{i=1,\ldots,n}} E\{\omega[\alpha_1^*(\theta^*(x)), \ldots, \alpha_n^*(\theta^*(x)), x]\} - n\theta^*$$

Assumption 3 (Nonuniqueness of One-Type Optima) *There exist two distinct values θ^* and θ^{**} and two sets of decision functions $\{\alpha_i^*\}_{i=1,\ldots,n}$ and $\{\alpha_i^{**}\}_{i=1,\ldots,n}$ such that $(\theta^*, \{\alpha_i^*\}_{i=1,\ldots,n})$ and $\theta^{**}, \{\alpha_i^{**}\}_{i=1,\ldots,n})$ are both one-type optima.*

Assumption 3 considers a restricted problem. Suppose the team can only hire agents of one type: which type of agents would it hire? The assumption requires that there are at least two optimal types. Without this assumption, it could be the case that a type of agent is strictly 'better' than the others. This would imply that workforce homogeneity is optimal in a trivial way. On the contrary, in reality, the labor supply is heterogeneous and, for any given profile, it comprises several types of workers, none of which clearly dominates the other.[8]

Of course, Assumptions 2 and 3 do not imply that heterogeneity is the optimal solution. In particular, if the payoff function is strictly supermodular, then all optimal configurations still require full homogeneity, as predicted by Proposition 1.

With Assumptions 1 through 3, the following holds:

[8]For instance, if a department wants to hire a faculty, it can choose from graduates of various graduate schools. There will be several schools with similar rankings. However, within the set of schools in the same ranking, there may be large differences in terms of focus or style.

Proposition 2 *If ω is submodular, then the set of optimal solutions contains at least one solution in which it is not true that $\theta_1 = \ldots = \theta_2$.*

To illustrate the scope of Proposition 2, consider the following:

Example 2 (A Search Problem): Consider a team of two researchers: 1 and 2. There are two possible fields of research: Left and Right. Only one of the fields is promising, but the researchers do not know which. Let z be a random variable which can assume the values 0 and 1 with equal probability. If $z = 0$, Left is promising. If $z = 1$, Right is promising. Each agent works for a unitary amount of time. Agent i chooses $a_i \in [0, 1]$, which represents the percentage of his work time that he devotes to searching Left. The percentage of time that he devotes to searching Right is given by $1 - a_i$.

Within a research field, there are decreasing returns to scale. For instance, assume that the probability of success is proportional to the square root of the total research time spent in that field. If the research is successful, the team receives a payoff of Q. The team's payoff can be written as:

$$\omega(a_1, a_2, z) = Qz\frac{\sqrt{2 - a_1 - a_2}}{\sqrt{2}} + Q(1 - z)\frac{\sqrt{a_1 + a_2}}{\sqrt{2}} \tag{4}$$

where the denominator $\sqrt{2}$ is used to normalize the probabilities.

Suppose that there are two types of researchers: those educated in university H (with information structure $\theta = h$) and those educated in university K (with $\theta = k$). Both types observe y with some error, but the errors are uncorrelated across types. Let the state of the world be $x = (y, \epsilon_H, \epsilon_K)$, where ϵ_H and ϵ_K are independently uniformly distributed on $[0, 1]$. The two possible information structures are defined as follows:

$$h(z) = \begin{cases} z & \text{if } \epsilon_H \geq p \\ 1 - z & \text{if } \epsilon_H \geq p \end{cases}$$

and

$$k(z) = \begin{cases} z & \text{if } \epsilon_K \geq p \\ 1 - z & \text{if } \epsilon_K \geq p \end{cases}$$

where $p \in (0.5, 1]$ denotes the precision of the signal.

It is easy to check that the function in (4) is submodular in a_1 and a_2. Moreover, Assumptions 2 and 3 hold as well. Therefore, by Proposition 2, the search problem described here always has an optimal solution in which one agent is of type h and the other is of type k.

Search problems typically entail submodular payoff functions. The more one agent searches in a direction, the more the other agents should search in other directions. Of course, the team can always order the agents to spread equally on all possible directions. However, this decision function is clearly not optimal because it does not take into account the agents' signals. The best thing the team designer can do is to hire agents who receive uncorrelated signals, so that agents can spread on different directions without renouncing their signals.[9]

4 Error Coordination

This section does not directly refer to the central theme of the paper. However, it illustrates a property of supermodular and submodular functions that is useful in interpreting the results of this paper. Consider a function of random variables. Suppose the random variables can be either perfectly codependent or mutually independent. This section proves that, if the function is supermodular, the expected value of the function is higher when the variables are perfectly codependent, while, if the function is submodular, it is higher when the variables are mutually independent. In the light of this result, we can provide some intuition on the results presented in Section 3.

Consider two random vectors:

$$\begin{aligned} y^I &= (y_1, y_2, \cdots, y_n) \\ y^D &= (y_0, y_0, \cdots, y_0) \end{aligned}$$

where $y_0, y_1, y_2, \cdots, y_n$ are identically distributed, mutually independent random variables. Consider a function $f : \Re^n \to \Re$. The following can be proven.[10]

Proposition 3 *If f is supermodular, then $E[f(y^D)] \geq E[f(y^I)]$, while if f is submodular, then $E[f(y^D)] \leq E[f(y^I)]$.*

If a function is supermodular, then the expected value is higher in the case of correlated errors than in the case of uncorrelated errors. The opposite

[9]For a discussion of the role of uncorrelated information in search problems, see also Bassan and Scarsini [5]. They consider a class of multi-agent search problems and demonstrate the value of heterogeneity based on the idea of experimentation externalities.

[10]Proposition 3 was conjectured by Milgrom and Roberts [18]. The proof is in the Appendix.

holds if the function is submodular. Although Proposition 3 is not used to prove Propositions 1 and 2, it provides intuition for those results.

Suppose the agents of an organization are bound to commit errors. The y's can be interpreted as the actions of the agents. The actions are random because the agent's signal has a random disturbance. Assume that the organization designer cannot reduce the entity of errors, but can choose whether the errors are perfectly codependent or mutually independent. If the team payoff function is supermodular, the organization designer will want the errors to be perfectly codependent. Workforce homogeneity is a device to make the errors perfectly codependent. On the other hand, if the team payoff function is submodular, the organization designer will want the errors to be mutually independent and heterogeneity is a device to make errors mutually independent.

5 Conclusion

We have considered the problem of a team that must decide the information structure of its agents. This paper has established a general connection between complementarities across agents and the opportunity of hiring agents with similar characteristics. If the payoff function is supermodular, agents should be homogeneous. If the payoff function is submodular, agents should be heterogeneous.

While the analysis presented here has been purely theoretical, its main ideas can be applied to important organizational issues. This paper predicts that the workforce homogeneity of a company is determined by the type of interaction between its agents. Therefore, activities for which good fit between various units is the first concern will have a homogeneous workforce in order to maximize coordination. On the other hand, activities that revolve around exploitation of new opportunities will have a more heterogeneous workforce in order to maximize the chance of developing successful innovations.

Perhaps the most limiting assumption of this paper is that agents cannot communicate with each other between the time they observe their signals and the time they take their actions. Of course, if there is an exogenous level of communication, the present model can easily be extended to apply to the that part of information which has not been communicated. However, the real challenge is to let communication be endogenous. Arrow [3, p. 56-59] noted that each organization develops its *code* – a set of channels of intraorganizational communication. How organizations develop their codes is a problem which is central to real organizations but has not yet been studies in economic theory. Future research might use a model similar to the present one to study coding.

6 Appendix: Proofs

Proof of Lemma 1 Suppose that S is an ordered set and that $q : S^m \to \Re$ is supermodular. Then, it is immediate from the definition of supermodularity that, for any $t \in S$ and $w \in S$,

$$q(t, \overbrace{w, \ldots, w}^{m-1 \text{ times}}) + q(w, \overbrace{t, \ldots, t}^{m-1 \text{ times}}) \leq q(\overbrace{w, \ldots, w}^{m \text{ times}}) + q(\overbrace{t, \ldots, t}^{m \text{ times}})$$

Consider now $\tilde{q} : S^{l+m} \to \Re$ and assume that $\tilde{q}$ is supermodular in all its arguments. Then, for any vector $(z_1, \ldots, z_l) \in S^l$, any $t \in S$, and any $w \in S$,

$$\begin{aligned} & \tilde{q}(z_1, \ldots, z_l, t, \overbrace{w, \ldots, w}^{m-1 \text{ times}}) + \tilde{q}(z_1, \ldots, z_l, w, \overbrace{t, \ldots, t}^{m-1 \text{ times}}) \\ \leq \quad & \tilde{q}(z_1, \ldots, z_l, \overbrace{t, \ldots, t}^{m \text{ times}}) + \tilde{q}(z_1, \ldots, z_l, \overbrace{w, \ldots, w}^{m \text{ times}}) \end{aligned} \tag{5}$$

If we apply (5) to the problem at hand, we have that for any $k = 2, 3, \ldots, n$ and for any $(p_1, \ldots, p_k) \in A^k$,

$$\begin{aligned} & f(p_1, \ldots, p_{k-2}, p_{k-1}, \overbrace{p_k, \ldots, p_k}^{n-k+1 \text{ times}}) + f(p_1, \ldots, p_{k-2}, p_k, \overbrace{p_{k-1}, \ldots, p_{k-1}}^{n-k+1 \text{ times}}) \\ & \leq f(p_1, \ldots, p_{k-2}, \overbrace{p_k, \ldots, p_k}^{n-k+2 \text{ times}}) + f(p_1, \ldots, p_{k-2}, \overbrace{p_{k-1}, \ldots, p_{k-1}}^{n-k+2 \text{ times}}) \end{aligned} \tag{6}$$

Let $P_k(a) = \{(p_1, \ldots, p_k) | (p_1, \ldots, p_n) \in P(a)\}$. Let $\bar{p}_k = (p_1, \ldots, p_k)$. By (6),

$$\begin{aligned} & \sum_{\bar{p}_k \in P_k(a)} [f(p_1, \ldots, p_{k-2}, p_{k-1}, \overbrace{p_k, \ldots, p_k}^{n-k+1 \text{ times}}) + f(p_1, \ldots, p_{k-2}, p_k, \overbrace{p_{k-1}, \ldots, p_{k-1}}^{n-k+1 \text{ times}})] \\ & \leq \sum_{\bar{p}_k \in P_k(a)} [f(p_1, \ldots, p_{k-2}, \overbrace{p_k, \ldots, p_k}^{n-k+2 \text{ times}}) + f(p_1, \ldots, p_{k-2}, \overbrace{p_{k-1}, \ldots, p_{k-1}}^{n-k+2 \text{ times}})] \end{aligned} \tag{7}$$

However, it is easy to see that

$$\sum_{\bar{p}_k \in P_k(a)} f(p_1, \ldots, p_{k-2}, p_{k-1}, \overbrace{p_k, \ldots, p_k}^{n-k+1 \text{ times}})$$
$$= \sum_{\bar{p}_k \in P_k(a)} f(p_1, \ldots, p_{k-2}, p_k, \overbrace{p_{k-1}, \ldots, p_{k-1}}^{n-k+1 \text{ times}})$$

and

$$\sum_{\bar{p}_k \in P_k(a)} f(p_1, \ldots, p_{k-2}, \overbrace{p_k, \ldots, p_k}^{n-k+2 \text{ times}})$$
$$= \sum_{\bar{p}_k \in P_k(a)} f(p_1, \ldots, p_{k-2}, \overbrace{p_{k-1}, \ldots, p_{k-1}}^{n-k+2 \text{ times}})$$

so that (7) becomes:

$$\sum_{\bar{p}_k \in P_k(a)} f(p_1, \ldots, p_{k-1}, \overbrace{p_k, \ldots, p_k}^{n-k+1 \text{ times}})$$
$$\leq \sum_{\bar{p}_k \in P_k(a)} f(p_1, \ldots, p_{k-2}, \overbrace{p_{k-1}, \ldots, p_{k-1}}^{n-k+2 \text{ times}}) \quad (8)$$

Notice, however, that now p_k does not appear in the right-hand side of (8). Thus, for any $(p_1, \ldots, p_{k-1}) \in P_{k-1}(a)$, the summation contains $n - k + 1$ identical elements corresponding to the possible values of p_k. Hence, (8) becomes

$$\sum_{\bar{p}_k \in P_k(a)} f(p_1, \ldots, p_{k-1}, \overbrace{p_k, \ldots, p_k}^{n-k+1 \text{ times}})$$
$$\leq (n-k+1) \sum_{\bar{p}_{k-1} \in P_{k-1}(a)} f(p_1, \ldots, p_{k-2}, \overbrace{p_{k-1}, \ldots, p_{k-1}}^{n-k+2 \text{ times}}) \quad (9)$$

By applying (9) recursively, we have

$$\sum_{\bar{p}_n \in P_n(a)} f(p_1, \ldots, p_{n-2}, p_{n-1}, p_n)$$

$$
\begin{aligned}
&\leq \quad 1 \cdot \sum_{\bar{p}_{n-1} \in P_{n-1}(a)} f(p_1, \ldots, p_{n-2}, p_{n-1}, p_{n-1}) \\
&\vdots \qquad\qquad \vdots \\
&\leq \quad 1 \cdot 2 \cdot \cdots \cdot (n-k) \sum_{\bar{p}_k \in P_k(a)} f(p_1, \ldots, p_{k-1}, \overbrace{p_k, \ldots, p_k}^{n-k+1 \text{ times}}) \\
&\vdots \qquad\qquad \vdots \\
&\leq \quad (n-1)! \sum_{\bar{p}_1 \in P_1(a)} f(\overbrace{p_1, \ldots, p_1}^{n \text{ times}})
\end{aligned} \tag{10}
$$

Notice that the left-hand side of (10) is equal to

$$
\sum_{(p_1, \ldots, p_n) \in P(a)} f(p_1, \ldots, p_n)
$$

and that

$$
\sum_{i=1,\ldots,n} f(a_i, \ldots, a_i) = \sum_{\bar{p}_1 \in P_1(a)} f(\overbrace{p_1, \ldots, p_1}^{n \text{ times}})
$$

Hence, (10) becomes

$$
\sum_{(p_1, \ldots, p_n) \in P(a)} f(p_1, \ldots, p_n) \leq (n-1)! \sum_{i=1,\ldots,n} f(a_i, \ldots, a_i)
$$

By dividing both sides of (6) by $n!$, the proposition is proven for f supermodular. If f is submodular, the proof goes through in the same fashion with switched inequality signs.

Verification That the Payoff Function in Eq. (3) Is Supermodular Consider any two vectors of strategies (a_1', a_2') and (a_1'', a_2''). By Definition 1, we have to prove that, for all x,

$$
\begin{aligned}
&\omega[\min(a_1', a_1''), \min(a_2', a_2''), x] + \omega[\max(a_1', a_1''), \max(a_2', a_2''), x] \\
&\geq \omega(a_1', a_2', x) + \omega(a_1'', a_2'', x)
\end{aligned} \tag{11}
$$

Assume without loss of generality that $a_1'' \geq a_1'$. If also $a_2'' > a_2'$, then (11) holds as an equality. Hence, suppose that $a_2'' \leq a_2'$. Then, (11) becomes

$$
\min(a_1'', a_2', x) + \min(a_1', a_2'', x) \geq \min(a_1', a_2', x) + \min(a_1'', a_2'', x) \tag{12}
$$

Given that $a_1'' \geq a_1'$ and $a_2'' \leq a_2'$, without loss of generality, we can assume that $a_1'' \geq a_2'$ (if it happens that $a_1'' < a_2'$, we can switch the suffixes $'$ and $''$ and also switch the indexes 1 and 2). There are three possible cases:

$$a_1'' \geq a_2' \geq a_2'' \geq a_1'$$

$$a_1'' \geq a_2' \geq a_1' \geq a_2''$$

$$a_1'' \geq a_1' \geq a_2' \geq a_2''$$

For all three cases it is easy to verify that (12) holds. Therefore, the payoff function in Eq. (3) is supermodular.

To prove Proposition 2 the following is useful

Lemma 2 *Under Assumption 2, the set of one-type optima includes a solution in which*

$$\alpha_1(y) = \ldots = \alpha_1(y)$$

Lemma 2 refers to the problem in which agents are artificially restricted to be of the same type. In that case, convexity of the action space and concavity of the payoff function are together a sufficient condition to have an optimal solution in which agents all agents have the same decision function.

Proof of Lemma 2 Suppose all agents have the same type θ. For all $y \in Y$, the team chooses a rule of action which maximizes

$$g[\alpha_1(y), \ldots, \alpha_n(y)] = E\{\omega[a_1, \ldots, a_n, x] | \theta(x) = y\} \tag{13}$$

Suppose ω is symmetric and concave in a. Because the expectation is a linear operator, also g is symmetric and concave in a.

Assume the rule of action $a^* = (a_1^*, a_2^*, \cdots, a_n^*)$ is a maximum of $g(a, y)$. By symmetry, all reorderings of a^* are maxima, too. If we define

$$\bar{a}^* = \left(\frac{\sum_{i=1}^n a_i^*}{n}, \frac{\sum_{i=1}^n a_i^*}{n}, \ldots, \frac{\sum_{i=1}^n a_i^*}{n} \right)$$

then, by concavity of g, we get $g(\bar{a}^*, y) \geq g(a^*, y)$. This holds for all possible signals y. Moreover, $\bar{a}^*$ is a symmetric rule of action. Then, if there exists an optimal rule of action – as has been assumed throughout this paper – then there also exists a symmetric optimal rule of action.

Proof of Proposition 2 By Assumption 3, there exist distinct types θ' and θ'' such that

$$(\theta', \theta'') \in \text{argmax}_{\alpha_1,\alpha_2,\theta_1,\theta_2} \omega[\alpha_1(\theta_1(x)), \ldots, \alpha_n(\theta_n(x)), x] - c_{\theta_1} - c_{\theta_2}$$

Suppose that α' is the optimal rule of action associated to θ' and α'' is the optimal rule of action associated to θ''. By Lemma 2, for θ' there exists an optimal set of decision functions $\alpha'_1(\cdot) = \cdots = \alpha'_n(\cdot) = \alpha'(\cdot)$ and for θ' there exists an optimal set of decision functions $\alpha''_1(\cdot) = \cdots = \alpha''_n(\cdot) = \alpha''(\cdot)$. Suppose ω is submodular. By Corollary 1,

$$\begin{aligned}
&\omega\{\overbrace{\alpha'[\theta'(x)], \ldots, \alpha'[\theta'(x)]}^{k \text{ agents}}, \overbrace{\alpha''[\theta''(x)], \ldots, \alpha''[\theta''(x)]}^{n-k \text{ agents}}\} - [kc_{\theta'} + (n-k)c_{\theta''}] \\
\geq\ & \frac{k}{n}\omega\{\overbrace{\alpha'[\theta'(x)], \ldots, \alpha'[\theta'(x)]}^{n \text{ agents}}\} - kc_{\theta'} \\
&+ \frac{n-k}{n}\omega\{\overbrace{\alpha''[\theta''(x)], \ldots, \alpha''[\theta''(x)]}^{n \text{ agents}}\} - (n-k)c_{\theta'}
\end{aligned}$$

for all $x \in X$ and for any $k = 1, 2, \ldots, n$. Take expectations and recall that θ' and θ'' yield the same expected net profit. Then, for $k = 1, 2, \ldots, n-1$,

$$\begin{aligned}
&E(\omega\{\overbrace{\alpha'[\theta'(x)], \ldots, \alpha'[\theta'(x)]}^{k \text{ agents}}, \overbrace{\alpha''[\theta''(x)], \ldots, \alpha''[\theta''(x)]}^{n-k \text{ agents}}\}) - [kc_{\theta'} + (n-k)c_{\theta''}] \\
\geq\ & E\{\omega[\overbrace{\alpha'[\theta'(x)], \ldots, \alpha'[\theta'(x)]}^{n \text{ agents}}]\} - nc_{\theta'} \\
=\ & E\omega[\overbrace{\alpha''[\theta''(x)], \ldots, \alpha''[\theta''(x)]}^{n \text{ agents}}] - nc_{\theta''}
\end{aligned}$$

It follows that diversified knowledge is optimal.

Proof of Proposition 3 Suppose f is supermodular. Because the y's are identically distributed, the expected value of f is invariant to permutations of the y's:

$$E[f(y_1, \cdots, y_n)] = \cdots = E[f(y_n, \cdots, y_1)]$$

and, obviously,

$$E[f(y_i, \cdots, y_i)] = E[f(y_0, \cdots, y_0)] \text{ for } i = 1, 2, \ldots, n$$

If f is supermodular, its expected value is supermodular as well. Consider the arithmetic average of all the possible permutation of the y's. By Lemma 1,

$$\frac{\sum_{(y_{j_1},\ldots,y_{j_n})\in P(y)} f(y_{j_1},\ldots,y_{j_n})}{N[P(y)]} \leq (\geq) \frac{1}{n}[f(y_1,\ldots,y_1)+f(y_2,\ldots,y_2)+\cdots+f(y_n,\ldots,y_n)]$$

implying

$$E[f(y_1,\cdots,y_n)] \leq E[f(y_0,\cdots,y_0)]$$
$$E[f(y^I)] \leq E[f(y^D)]$$

and conversely when f is submodular.

References

[1] Masahiko Aoki. Horizontal vs. vertical information structure of the firm. *American Economic Review*, 76:971–983, 1986.

[2] Masahiko Aoki. An evolving diversity of organizational mode and its implications for transitional economies. *Journal of the Japanese and International Economies*, 9(4):330–353, 1995.

[3] Kenneth J. Arrow. *The Limits of Organizations.* Norton, New York, 1974.

[4] Kenneth J. Arrow and Roy Radner. Allocation of resources in large teams. *Econometrica*, 47:361–386, 1979.

[5] Bruno Bassan and Marco Scarsini. On the value of information in multi-agent decision theory. *Journal of Mathematical Economics*, 24:557–576, 1995.

[6] Jacques Crémer. A partial theory of the optimal organization of bureaucracy. *Bell Journal of Economics*, 11:683–93, 1980.

[7] Jacques Crémer. Common knowledge and the coordination of industrial activities. In Aoki et al., editor, *The Firm as a Nexus of Treaties.* Sage Publications, 1990.

[8] Jacques Crémer. Corporate culture and shared knowledge. *Industrial and Corporate Change*, 2(3):351–386, 1993.

[9] Esther Gal-Or. Information sharing in oligolopoly. *Econometrica*, 53:329–343, 1985.

[10] John Geanakoplos and Paul Milgrom. A theory of hierarchies based on limited managerial attention. *Journal of the Japanese and Interanational Economies*, 5:205–225, 1991.

[11] Theodore Groves and Roy Radner. Allocation of resources in a team. *Journal of Economic Theory*, 3:415–444, 1972.

[12] Michael Kremer. The O-ring theory of economic development. *Quarterly Journal of Economics*, 108(3):551–575, 1993.

[13] Bo Li. The optimal design of communication networks: M-form, u-form, and matrix management. mimeo, Stanford University, 1995.

[14] Jakob Marschak and Roy Radner. *Economic Theory of Teams*. Yale University Press, New Haven, Connecticut, 1972.

[15] Margaret A. Meyer and Dilip Mookherjee. Incentives, compensation, and social welfare. *Review of Economic Studies*, 54:209–226, 1987.

[16] Paul Milgrom and John Roberts. Rationalizibility, learning, and equilibrium in games with strategic complementarities. *Econometrica*, 58(6):1255–1277, 1990.

[17] Paul Milgrom and John Roberts. Complementarities and fit: Strategy, structure and organizational change in manufacturing. *Journal of Accounting and Economics*, 19:179–208, 1994.

[18] Paul Milgrom and Christina Shannon. Monotone comparative statics. *Econometrica*, 62(1):157–180, 1994.

[19] Jean-Pierre Ponssard, Sébastien Steinmetz, and Hervé Tanguy. Team models as a framework to analyze coordination problems within the firm. mimeo, Laboratoire D'Econométrie de l'Ecole Polytechnique, 1996.

[20] Andrea Prat. Shared knowledge vs. diversified knowledge in teams. *Journal of the Japanese and International Economies*, 10(2):181, 1996.

[21] Andrea Prat. How homogeneous should a team be? Center Discussion Paper n. 9845. Tilburg University. 1998.

[22] Yingyi Qian, Gérard Roland, and Chenggang Xu. Information and coordination in U-form and M-form organizations. mimeo, Stanford University, 1995.

[23] Yingyi Qian, Gérard Roland, and Chenggang Xu. Why is China different from Eastern Europe? Perspectives from organization theory. *European Economic Review*, papers and proceedings, 43:1085–1094, 1999.

[24] Donald Topkis. Minimizing a submodular function. *Operations Research*, 26(2):305–321, 1978.

[25] Xavier Vives. Duopoly information equilibrium: Cournot and Bertrand. *Journal of Economic Theory*, 34:71–94, 1985.

[26] Xavier Vives. Nash equilibrium with strategic complementarities. *Journal of Mathematical Economics*, 19(3):305–21, 1990.