

Speed and Quality of Collective Decision Making: Incentives for Information Provision*

Hans Peter GRÜNER

University of Mannheim and CEPR, London

Elisabeth SCHULTE

University of Mannheim

June 2009

Abstract

We study an information aggregation problem in which agents have to provide effort in order to understand the information they are supposed to process. Showing that any problem which is incentive compatible for a single information processor is incentive compatible for a decentralized organization, but not vice versa, we derive a new rationale for decentralized information processing. For a class of problems, the fastest organization—the reduced tree proposed by Radner (1993)—yields also the best incentives for information processing.

Keywords: Information processing, hierarchies, incentives for information provision.

JEL D23, D70, D83, L22, P51.

*We thank Enno Mammen, Thomas Gall, Nicola Persico, Jörg Oechssler, Frank Riedel, Benny Moldovanu, Michel LeBreton and participants of workshops and seminars at Erasmus University Rotterdam, Gerzensee, Bonn University, Mannheim University, UCLA, at ZEI Bonn, Université Toulouse and Aachen University, and participants of the 2004 ENTER Jamboree in Barcelona and of the Econometric Society World Congress 2005 in London, and an anonymous referee for helpful comments and suggestions. Elisabeth Schulte acknowledges financial support through the DFG research group on the institutionalization of international negotiation systems (IINS). Correspondence address: University of Mannheim, Department of Economics, L7, 3-5, D-68131 Mannheim, Germany; email: hgruener@rumms.uni-mannheim.de; elisabeth.schulte@vwl.uni-mannheim.de.

1 Introduction

This paper shows that speed and quality of information aggregation need not be conflicting goals. It identifies economic environments in which the best (quality maximizing) hierarchy also minimizes decision delay.

The aggregation of information is a key task of almost all complex economic organizations (Radner, 1993). Assuming that information processing takes time, one might be tempted to believe that there is always a trade-off between the speed and the quality of decision making. Indeed, the quickest decision can be made instantaneously by disregarding all the information. However, disregarding information is not the only option to speed up the decision making process. One can also use parallel processing structures (see Radner, 1993 or Radner and Van Zandt, 1992). Delegated information processing has one easily available shirking option: instead of studying the information, one may just guess its content. In this paper, we explore the effect of decentralization on the incentives to provide effort when processing information. We will study the incentives for information provision in hierarchies, in which agents are motivated to provide effort only by its effect on the decision.

Our paper studies the optimal design of institutions that process complex and decentralized information. Various real world institutions deal with the kind of information processing tasks that we consider in this paper. One example is a financial intermediary who has to identify the best investment opportunity for new deposits. The evaluation of an investment opportunity is a complex and time consuming task and it is natural to delegate this task to several individuals (experts) in order to save on processing time. Under such a decentralized arrangement incentive problems naturally arise. These incentive problems are likely to affect the quality of the overall outcome. Many other large organizations (firms or government agencies) which decide upon the proper allocation of a given budget face similar problems. Incentive problems of the sort studied in this paper may also arise in firms which operate in several regional markets when headquarters have to rely on aggregates of regional customer data. The organization of information processing will determine both the delay of a decision and the individuals' incentives for information provision.

Papers like Mukhopadaya (2003), Gerardi and Yariv (2008) or Persico (2004) suggest that the delegation of information acquisition to more than one individual dilutes individual incentives. In these papers, agents are supposed to access the same (or at least correlated) information, which creates a strong incentive to free-ride, because individual

information acquisition investments are strategic substitutes. In the present paper, instead, we are interested in information aggregation problems in which concrete tasks need to be performed in order to uncover a particular part of decision-relevant state of the world. Examples are the assessment of a pool of investment opportunities in order to choose the best of them, the aggregation of regional customer data in order to optimize the production plan, or the collection of data regarding a firm's financial standing by a team of rating analysts. We present a formal analysis of some examples for our framework in the appendix. What these problems have in common is that there are strong complementarities among processing tasks.

We derive a new rationale for decentralization of information processing in such decision problems: Whenever a single decision maker provides effort when performing the information processing tasks, any member of an organization within which the tasks are distributed provides effort (in the Pareto-dominant equilibrium). The opposite is not necessarily true. Hence, delegating information processing to multiple agents may increase both, speed and quality of decision making. Moreover, for a class of information aggregation problems, it is the fastest organization (the reduced tree proposed by Radner (1993)) which yields the best incentives for information processing. While there is still a trade-off when deciding how much information to process, there is no such trade-off when organizing the processing of a given set of information.

Besides addressing the specific problem of the relationship between the speed and the quality of information processing, the present paper also makes a more general contribution to the literature on information aggregation in organizations. So far the literature on delay in information processing can be seen as disjoint from the literature on incentives. The seminal contribution by Radner (1993), and the literature on information processing in hierarchies which followed, disregards that agents who are supposed to perform assigned tasks need incentives to do so properly. On the other hand, the game theoretical literature ignores that the design of hierarchies and incentives schemes may also have an impact on the delay. Extending Radner's original analysis of information processing for a game structure enables us to treat both problems in the same model.

The paper is organized as follows: In the next section, we discuss the related literature. Then, we present a toy model that illustrates the main intuition for the benefits of decentralized information processing. In Section 4, we present our model of information processing. We study the conditions for incentive compatibility in Section 5 and present our conclusions for efficient organization design in Section 6. In Section 7, we conclude.

2 Related literature

This paper is most closely related to a recent literature on organization design which draws on insights from computer science, starting with Radner (1992, 1993), Radner and Van Zandt (1992). The reduced tree is designed for one-shot problems (to which we restrict attention). These are problems in which there is only one set of data to be processed, or the processing of the data is finished before another calculation task occurs. Van Zandt (1997, 1998) and Meagher, Orbay and Van Zandt (2003) study the case when new data comes in before the processing of the old set is finished. Orbay (2002) adds the frequency with which new data arrives as a new dimension to the analysis of efficient hierarchies. Prat (1997) studies hierarchies in which some managers are able to work faster than others, and the wage a manager is paid is a function of his ability. It turns out that with these modifications—except for the one made by Prat (1997)—the reduced tree is still (close to) efficient. Bolton and Dewatripont (1994) allow for specialization, which reduces the time an agent needs to understand information he handles frequently. In their model, the trade-off between specialization and communication costs determines the extent and the form of decentralization.

In Radner’s model—and in most of the information processing literature which followed—individuals are thought of as machines, perfectly doing what they are programmed to do. There are at least two important features of information processing by human agents which may require modifications of this basic model. First, an individual’s calculation ability might not be perfect, i.e. occasionally individuals may make mistakes. Second, when delegating tasks, one has to make sure that the agent has an incentive to perform them. Thus, there emerges another dimension for the evaluation of hierarchies: the quality of the decision.

Meagher (2003) is the only paper in this literature which takes account of incentive problems and processing delay in a unified framework. He considers a firm which pays efficiency wages to agents on all levels in a balanced hierarchy, and studies the effect of improved communication technology on the size of the hierarchy. An agent’s task is either to work or to supervise (which is both costly), whereas in our paper, the focus is on task assignment, allowing for more general hierarchy structures than in Meagher’s paper.

The joint analysis of speed and quality of hierarchical decision processes has previously been carried out in Jehiel (1999) and in Schulte and Grüner (2007). Jehiel considers the case where some signals get lost in the hierarchy with an exogenous probability, depending on the size of the groups of which the hierarchy consists. Schulte and Grüner study the

role of the hierarchy design when individuals make mistakes with an exogenously given probability. In the present paper, the quality of collected information is endogenously determined by the actions of self-interested agents.

Our paper is also related to a huge game theoretic literature that studies incentives in hierarchies, in particular Aghion and Tirole (1997), Mookherjee and Reichelstein (1997), and Melhumad, Mookherjee and Reichelstein (1995). These papers consider problems in which certain tasks as well as authority have to be delegated (and sub-delegated) to (and by) agents whose interests diverge from that of the principal. Delegation involves a loss of control for the principal, but strengthens the incentives for the agent. In our model, all agents have the same objective: The available information shall be processed as accurately as possible in order to make an informed decision.

Winter (2004) studies incentive provision in a hierarchy via a transfer scheme. In his paper, the n tasks are assigned to n agents right from the beginning. He does not allow for the possibility to assign tasks differently, nor that one agent performs all of them. Unlike Winter (2004), we are not interested in a transfer scheme that ensures effort provision, but rather in the effect of task assignment on effort provision. A related moral-hazard-in-team problem is studied in Dewatripont and Tirole (2005). They consider a sender-receiver game, where the sender has payoff-relevant information for the receiver, and both must invest costly and unobservable effort for the receiver to understand the information. In our model, communication is costless for the sender, but he has to provide effort to acquire the information in the first place.

Other papers derive decentralized (hierarchical) organizations from technology. Crémer (1980) considers a problem of resource allocation under constraints on managerial time and finds that hierarchical organizations increase the amount of information that can be applied to a particular decision. Rosen (1982) has a paper in which a hierarchical structure emerges due to the need to supervise production (and supervision).

The paper is also closely related to the recent seminal work by Garicano (2000) and Garicano and Rossi-Hansberg (2004, 2007) on the microeconomic and distributional consequences of IT innovations. Garicano (2000) introduces an alternative model of hierarchical information aggregation. In this model agents may either specialize in problem-solving or in providing advice to other problem solvers. The major difference between both approaches is that our paper applies game theory in the analysis of programmed networks. This enables us to analyze the interplay between the computing structure and incentive problems within hierarchies.

3 A toy model

Our model of information processing is essentially one of team production. However, it has some features that are special to information aggregation problems: (i) strategies do not only specify whether to provide effort or not, but also what to report to the superior, and (ii) involving one more agent into the process creates a new task because the partial result produced by that agent needs to be incorporated. For the toy model, in order to deliver the intuition for the benefits of decentralization, we abstract from these issues and consider a simple production problem in which n tasks need to be completed. Suppose that each task is completed only if the agent to which it is assigned provides effort, where effort is a binary variable, either 0 or 1. Let $e_k \in \{0, 1\}$ represent the effort provided on task k , $k = 1, \dots, n$, and let $f(e_1, \dots, e_n)$ denote the gross payoff for each agent.

The problem is to assign the tasks to P agents such that effort is provided on each task. Let N_i be the set of tasks assigned to agent i , where $\{N_i\}_{i=1, \dots, P}$ is a partition of $\{1, \dots, n\}$. We will call an organization A *more decentralized* than organization B if the task assignment in A is a finer partition than that in B, i.e. if for at least one agent, a strict subset of the tasks assigned to him in B is assigned to another agent in organization A without adding additional tasks to those agents in B who are assigned non-empty sets of tasks in B.¹ Agent i suffers a cost of effort provision, $\sum_{k \in N_i} e_k \cdot c$. Given task assignments and effort decisions (e_1, \dots, e_n) , agent i 's payoff is

$$f(e_1, \dots, e_n) - \sum_{k \in N_i} e_k \cdot c.$$

A given task assignment defines a game. We are interested in the incentive compatibility constraints for full effort provision, i.e. the conditions under which it is optimal for each agent to provide effort on all the tasks assigned to him given that all other agents do the same. We will establish the following:

1. If a task assignment is incentive compatible, then any more decentralized task assignment (any finer partition) is incentive compatible.
2. The converse is not necessarily true if $f(\cdot)$ exhibits strict complementarities among the tasks.

¹We can think of other measurements of decentralization, for instance in terms of the size of the largest task. However, any other useful definition of decentralization should be in accordance with the above definition.

Item 1 holds because the incentive compatibility constraints for a set of tasks which are assigned to two individuals are implied by the incentive compatibility constraints for one individual facing the entire set of tasks. To see this, suppose that task k is assigned to agent i and assume $|N_i| = n_i > 1$. One of the incentive compatibility constraints demands agent i to provide effort on all his tasks rather than to shirk only on task k , given that all other agents provide effort on all other tasks. That is:

$$c \leq f(1, \dots, 1, 1, 1, \dots, 1) - f(1, \dots, 1, 0, 1, \dots, 1). \quad (1)$$

Now, consider a more decentralized version of the task assignment. Add agent j and assign task k to him. Keep the rest of the task assignment fixed. Note that incentive compatibility constraints for all agents but i and j remain the same. The incentive compatibility constraint for task k is still:

$$c \leq f(1, \dots, 1, 1, 1, \dots, 1) - f(1, \dots, 1, 0, 1, \dots, 1). \quad (2)$$

Hence, if agent i provides effort when performing task k in the old organization, then agent j provides effort in the new organization. Supposing that agent j provides effort on task k , agent i 's remaining incentive compatibility constraints become (weakly) weaker than if he himself has to provide that effort.

To understand why item 2 holds, assume that the output depends only on the number of tasks on which agents provided effort. We elaborate on some relevant examples for which this is the case in the appendix. Suppose that tasks are strict complements. The loss in output from shirking on m tasks simultaneously is smaller than the sum of losses from shirking on each of the tasks (one at a time). Denote the loss in output from shirking on m tasks $\Delta(m)$. Given the other players' behavior, the incentives to provide effort depend only on the number of assigned tasks. If tasks are strict complements, $\Delta(m)/m$ is decreasing in m , which implies that an agent who prefers to provide effort on m tasks to shirking on all m tasks also prefers to provide effort on m tasks to shirking on a subset of tasks. The necessary and sufficient condition for the existence of an equilibrium with full effort provision is $c \cdot n_{\max} \leq \Delta(n_{\max})$, where $n_{\max} = \max\{|N_i|\}$. Hence, incentive compatibility constraints are weaker in more decentralized organizations. Another implication for the case of complementarities is that for a given number of agents, a symmetric assignment of tasks yields the strongest incentives for effort provision.

If tasks are strict substitutes, the organizational form has no impact on the incentive structure, because $c \leq \Delta(1)$ is a necessary and sufficient condition for full effort provision

in any organization. Any decentralized task assignment is incentive compatible if and only if the centralized task assignment is incentive compatible.

In our toy model, there does not arise a trade-off between speed and quality: Delegating to as many agents as possible and performing tasks simultaneously speeds up the process without compromising with respect to quality. On the other hand, in an economy with complementarities, slower (less decentralized) organizations actually provide less incentives for individual effort.

4 A model of information processing

4.1 Relation to the toy model

In this section, we illustrate our model of information processing. To understand why the toy model does not entirely capture the information processing framework, consider the following simple information processing task. The task is to calculate the sum of n variables x_k , $k = 1, \dots, n$, $x_k \in \{0, 1\}$. This task—like all others which we consider in this paper—has the property that only the number of items on which individuals provide effort counts for the quality of the outcome. If the entire task is assigned to a single agent, it consists of n subtasks, each of which is to read a variable and to add it to the partial result stored in the memory (in the beginning of the procedure, the memory is set to 0). If the task is assigned to $P > 1$ agents, the task consists of $n + P - 1$ subtasks, because the final result is only obtained if all partial results are summed up. For that, it is necessary to read the partial results from $P - 1$ agents.

Hence, the problem cannot be represented by a production function with effort decisions on the n original tasks as inputs anymore. We also need to make sure incentive compatibility on higher levels, i.e. for the aggregation of partial results. Moreover, agents' strategy spaces are more complex because agents do not only decide whether to provide effort or not, but also what to report to the superior.

Further, in the case of an information aggregation problem, the “loss in production” from shirking stems from an inherent interest in the decision to be made contingent on the aggregate. Shirking means to guess the information rather than to process it properly, which induces uncertainty about the aggregate and therefore about the best decision. Given the inherent interest in the decision, an agent who shirks will formulate the best guess about the partial result he is supposed to provide. The best guess depends on the concrete problem. Hence, also $\Delta(m)$, the loss in utility due to shirking on m tasks,

depends on the problem, i.e. the distribution of the information variables, the aggregation operation, and on how much agents suffer from deviations of the actual from the optimal decision.

Consider again the above summation problem. It may represent the problem of a host who has to decide how many fish to buy for his dinner party. For that, he wants to know how many guests are coming for dinner. The information variables are the messages of the individuals he invited, containing either a “0” meaning “Unfortunately, I cannot come” or a “1” meaning “I am happy to join the party”. Our host would like best to match the number of fishes with the number of guests. We represent his preferences by the utility function:

$$u(d, x) = -|d - x|,$$

where $x = \sum_{k=1, \dots, n} x_k$. Suppose that $\text{prob}\{x_k = 1\} = \text{prob}\{x_k = 0\} = 1/2$. For this problem, the best guess for the sum of m variables is the median of their distribution, $m/2$ in this case.²

In the appendix of this paper (Example 3) we derive the cost of guessing, which is a weakly increasing step-function with diminishing increments. We can conclude from our toy model that incentives for raw data processing become better the more agents are involved, i.e. the smaller the workload of the individual agent is. However, in addition to guaranteeing incentive compatibility for raw data processing, we need to make sure incentive compatibility for the processing of partial results.

In the remainder of this section, we present a more general formulation of the problem. In the following sections, we formally derive the superiority of decentralized organization structures and the efficient form of decentralization.

4.2 The problem

There is a collective decision $d \in D$ to be made which generates a state-dependent payoff $u(d, x)$ for each member of society. The state of the world $x \in X$ is the aggregate of n information variables³ x_i , where $x_i \in X$, $i = 1, \dots, n$. The vector of information variables $\mathbf{x} = (x_1, \dots, x_n)$, where $\mathbf{x} \in X^n$, is transformed into the aggregate x via a function $f_n : X^n \rightarrow X$.⁴

²See DeGroot (1970), p. 232.

³We use the terms “information variables” and “data items” interchangeably.

⁴See the appendix for concrete examples.

The information variables x_i are drawn from commonly known distribution functions $\phi(x_i)$, $i = 1, \dots, n$. The realization can be learned only if costly effort is provided. In order to learn the aggregate state of the world x , the information variables have to be processed, that is $f_n(\mathbf{x})$ has to be calculated. The aggregation operations are commutative and associative, hence information processing can be decentralized i.e. raw data can be aggregated to partial results⁵ by some agents which can be further aggregated by other agents. Since there is common knowledge about the distributions $\phi(x_i)$ of the information variables and the processing task, the distributions of partial aggregates and the distribution of the aggregate x are common knowledge as well. Decentralized information processing takes place in a programmed network.⁶ The program describes how information processing tasks are distributed within the network. There are n data items that have to be aggregated by P agents called managers.

The program consists of assignments $a : \{1, \dots, n\} \rightarrow \{1, \dots, P\}$ of raw data items to managers, and a reporting structure $p : \{1, \dots, P-1\} \rightarrow \{1, \dots, P\}$. Moreover, it specifies reporting spaces R_i , $i = 1, \dots, P$. Let agent P represent the “top manager”, i.e. the head of the organization. The top manager’s report r_P is the final result, which is used as an input for the decision rule. The assignments $a(\cdot)$ generate a partition $\{N_i\}_{i=1, \dots, P}$ of the raw data set $\{1, \dots, n\}$. The reporting structure $p(\cdot)$ connects the members of the organization and generates a tree. We denote with r_j the report send by j , and with S_i the set of agents reporting to i (i ’s subordinates). Hence, agent i is supposed to handle the items in the set $N_i \cup \{r_j : j \in S_i\}$.

We assume that a manager only understands the content of an item he is supposed to process if he provides unobservable, costly effort. Denote with $e_k \in \{0, 1\}$ the effort provided on item k , $k \in \{1, \dots, n\} \cup \{r_j : j \in S_i\}_{i=1, \dots, P}$. Given his effort decisions, agent i suffers a cost $\sum_{k \in N_i \cup \{r_j : j \in S_i\}} e_k \cdot c$.⁷ If an agent does not provide effort when reading an information variable, he does not receive any information about its realization. However, if the realization is not learned, it may be guessed. This is the moral hazard problem we study in this paper.

On top of the assignment structure, the program includes a time schedule: It instructs

⁵We use the term “partial result” for an information variable that is the outcome of an aggregation operation and “raw data (item)” for an information variable that has not yet been combined with another one.

⁶See Radner (1993) for a more detailed description of a programmed network.

⁷We focus on the case of constant marginal cost of effort provision. Note that a convex cost function makes decentralization even more attractive.

each agent when to read an item, which operation to perform, and when to send a report. Each manager is endowed with an inbox, a processing unit and a memory. We assume that reading an object takes one unit of time, and that the other tasks do not cause any delay. Hence, in one unit of time, a manager is able to perform the following tasks: (i) reading an object from his inbox into his processing unit, (ii) aggregating it to what is stored in his memory, and (iii) sending a report.

4.2.1 Preferences

All individuals in the network are interested in the quality of the overall result, but they also suffer from an individual cost of providing effort. Moreover, they suffer from delay in decision making.⁸

All agents have identical preferences, which can be represented by the following utility function:

$$U_i = u(d, x) - \sum_{k \in N_i \cup \{r_j: j \in S_i\}} e_k \cdot c,$$

where $\sum_{k \in N_i \cup \{r_j: j \in S_i\}} e_k$ measures how many times individual i provides effort. Utility is additively separable in the utility derived from the decision and the cost of effort provision. For each x , there exists a $d^*(x) \in D$ maximizing $u(d, x)$.

Alternative to processing the whole set of information, agents may guess part or all of the information. If information processing takes place in a decentralized manner, that is if at some stage agent i reports r_i to another agent, then the information contained in the report is learned by that agent only if he provides effort. We call a data item k *processed properly* if every agent who is supposed to read an information variable in which the item is contained provides effort.

Definition 1 *Data item $k \in \{1, \dots, n\}$ is processed properly if and only if $e_k = e_{r_j} = 1$ for all j on the path from object k to the top.*

Let m denote the number of information variables which are not processed properly, and let $f_{m'}(x^{m'})$ denote the aggregate of $m' = n - m$ properly processed items. We assume that there exists a decision rule that maximizes agents' expected utility given the information they have, i.e. there exists $d^{m'}$ maximizing $E[u(d, x) | f_{m'}(x^{m'})]$, where

⁸We handle speed and quality of collective decision making as orthogonal dimensions. The time dimension has been explored previously by Radner (1993). We will briefly recall his results in Section 6, where we present our results on efficient organization design along both dimensions. In the rest of the paper, we focus on the quality dimension.

$0 \leq m' \leq n$.⁹ The cost of guessing is the expected loss of utility due to a possible deviation from the optimal decision.

Definition 2 *The cost of guessing m out of n information variables is*

$$\Delta(m) = E[u(d^*, x) - u(d^{m'}, x) | f_{m'}(x^{m'})]. \quad (3)$$

For information processing problems into which the original data items enter as complements, the cost of guessing exhibits diminishing increments. These are problems with the property that the more information is already guessed, the less additional guessing hurts. If raw data items are substitutes in that sense, the cost of guessing exhibits increasing increments. These are problems for which the marginal return of acquiring additional information is decreasing.

4.2.2 Properties of the problem

We restrict our analysis to problems for which (i) all data items enter symmetrically into the processing problem, (ii) the best guess of an information variable does not depend on previously acquired information, and (iii) $u(d^*, x)$ is the same for all x . These restrictions are imposed for convenience.¹⁰ Therewith, the cost of guessing (3) depends only on the number of original data items which are not processed properly.

The information aggregation problems to which we restrict attention always permit the best possible use of the information that has been processed properly. This means that if $n - m$ objects are processed properly, there exists a communication profile such that agents' expected gross payoff $u(d^*, x) - \Delta(m)$ is maximal.

In the appendix of this paper, we provide an analysis of relevant examples for information processing problems which have the properties stated in this paragraph. For instance,

⁹This requires that r_P , the final result provided by the top manager, incorporates the best guess of the aggregate of the information variables that are not processed properly. In the case of decentralized information processing, this in turn requires that agents report the best guess of information variables that are not processed properly and independence of best guesses. We elaborate on this issue in Section 4.2.2.

¹⁰Relaxing (i) would require to consider more incentive compatibility constraints and to identify the most binding ones to insure the existence of an equilibrium with full effort provision. Relaxing (ii) would require to consider more complex reporting spaces. In order to incorporate a possible correlation into the best guess, agents should be allowed to report on which of their assigned tasks they provided effort in addition to reporting from the set of possible partial results. Relaxing (iii) would introduce additional uncertainty into the individual optimization problems. Agents would have to form expectations over the cost of guessing and to update them with any additional information they acquire.

in Example 2, reporting expected values of information variables (which means to report the realization in case of effort provision) is a communication profile which minimizes the cost of guessing.

4.3 Classical surplus and efficiency

We evaluate an organization along two dimensions: (i) the delay within which a decision is reached and (ii) classical surplus. Let m denote the number of objects that are not processed properly, where $m \in \{0, 1, \dots, n\}$. The *classical surplus* of the economy (the sum of utilities) is defined by

$$v := s \cdot (u(d^*, x) - \Delta(m)) - \sum_{i=1}^P \sum_{k \in N_i \cup \{r_j: j \in S_i\}} e_k \cdot c, \quad (4)$$

where s denotes population size, and P denotes the number of agents involved in information processing.¹¹ Agents who are not involved in information processing do not contribute any effort. Note that generally adding agents to the hierarchy generates a cost: If every hierarchy member provides effort, the total cost of effort provision is $(n + P - 1)c$. There are two possible reasons for decentralizing information processing: (i) incentives or (ii) the speed of decision making.

4.4 Equilibrium concept and equilibrium selection

The specification of preferences, the moral hazard problem and the underlying programmed network fully describe a game. In this game a strategy is a plan that fixes (i) when to provide effort (possibly history dependent) and (ii) what to report (again possibly history dependent). We call part (i) an action plan and part (ii) a communication plan. Correspondingly, a strategy profile can be divided into an action profile and a communication profile. The equilibrium concept we use is Nash equilibrium.

Since agents' preferences concerning the decision are identical, we consider only equilibria which have the following property. If a random variable is guessed, then the "best guess" is formulated in the sense that it maximizes expected utility when used as input for subsequent aggregation operations. In the case that an information variable x_i is guessed, the input to the operation $f_2(\cdot)$ is the best guess.¹²

¹¹We consider the total cost of effort provision as a measure of cost. See Meagher and Van Zandt (1998) for a discussion of appropriate cost measures for decentralized information processing tasks.

¹²Note that the best guess need not be unique, e.g. when guessing which of two objects has a higher value, both of them are best guesses, if the objects' values are identically distributed.

We use the Pareto-criterion for equilibrium selection where applicable.¹³ Under our assumption of effective communication, if there exists an equilibrium in which $n - m$ objects are processed properly, and in which agents' expected payoff is smaller than $u(d^*, x) - \Delta(m)$, then this equilibrium is Pareto-dominated by an equilibrium with the same action profile and effective communication such that the expected payoff is $u(d^*, x) - \Delta(m)$. We can hence focus our equilibrium analysis on the incentives to provide effort and may thereby represent incentive compatibility constraints in terms of the cost function $\Delta(m)$.

The Pareto-criterion selects the equilibrium with maximal effort provision.

Lemma 1 *Consider a programmed network. Let $m_1 > m_2$. Suppose that two equilibria 1 and 2 coexist in which m_1 and m_2 objects are processed properly, respectively. Assume that in the first equilibrium all agents provide effort on at least as many items as in the second equilibrium. The equilibrium in which m_1 objects are processed properly is Pareto-dominant.*

Proof. Consider first the individuals who provide the same effort and hence face the same cost of effort in both equilibria. If m_1 objects are processed properly, they face a cost of guessing $\Delta(m_1) < \Delta(m_2)$. They are better off in the equilibrium in which m_1 objects are processed properly.

Consider next those individuals who provide more effort in the equilibrium in which m_1 objects are processed properly. For them, shirking on some of their tasks is an available deviation which they do not choose. Hence, they must be better off in the equilibrium with more effort provision as well. Q.E.D.

5 Equilibrium Analysis

We begin our equilibrium analysis by studying the optimization problem of a single decision maker who is supposed to handle the entire set of data. We obtain a simple condition for incentive compatibility that will be used in the rest of the paper. Next, we turn to the conditions for effort provision in the reduced trees. The conclusions can then be used to prove our main results.

¹³Note also that the game is a potential game with the potential function $f(e_1, \dots, e_n) - (n - m)c$. The equilibria that we consider maximize this potential function.

5.1 A single decision maker

In order to understand the role of hierarchies for information processing, consider first the fully centralized version, i.e. one individual processing the entire set of data. The following conditions are necessary for full effort provision by the single decision maker. Together, they are sufficient.

$$n \cdot c \leq \Delta(n), \tag{5}$$

$$m \cdot c \leq \Delta(m) \quad \forall m \in \{1, \dots, n-1\}. \tag{6}$$

(5) demands that full effort is better than no effort, and (6) makes sure that full effort provision is better than shirking on m items. If $\Delta(\cdot)$ exhibits weakly diminishing increments, (5) is a sufficient condition for full effort provision as it implies that (6) holds for all $m \in \{1, \dots, n-1\}$. If $\Delta(\cdot)$ exhibit weakly increasing increments, a sufficient condition for full effort provision is that (6) holds for $m = 1$ as this implies that it holds for all $m \in \{2, \dots, n-1\}$ and that (5) holds.

5.2 The reduced tree

We now turn to the incentive compatibility constraints in the reduced tree, and we will show in the next section that this organizational form is efficient for a class of decision problems.¹⁴ We first briefly recall Radner's (1993) result regarding the optimal organization with respect to the dimension speed (for a given number of managers P and a given number of items $n \geq 2P$). Radner has shown that the following organization yields minimum delay: Assign the n items to P managers as evenly as possible. Therewith, raw data is processed with minimum delay as maximum parallel processing is implemented. The P partial results are aggregated quickest by having one half of the agents (up to integers) report to the other half on the next level. Continue that reporting procedure on all other levels, reducing the number of partial results by half at each step. The final result is obtained (neglecting integers) in $n/P + \log_2 P$ units of time. The resulting organization is called a reduced tree.

We restrict attention to information processing problems in which integer problems do not arise, i.e. we assume $n = bP, b \in \mathbb{N} \setminus \{0, 1\}$. We impose this assumption in order

¹⁴The working paper version of this paper contains results on other organizational forms and a longer discussion of the set of equilibria.

to avoid case distinctions in the proofs which do not add any further insights to the analysis.¹⁵

We now turn to the conditions for full effort provision in the reduced tree. First, we consider the case (i) that the cost of guessing exhibits weakly increasing increments. Next, we discuss the case (ii) that it exhibits weakly diminishing increments.

(i) It should be clear from the previous section that if the cost of guessing exhibits weakly increasing increments, the necessary condition for full effort provision in any organization

$$c \leq \Delta(1) \tag{7}$$

is sufficient. This is so because the last raw data item which is processed properly (given that all other items are processed properly) involves the smallest cost of guessing. If it is incentive compatible to provide effort when reading that item, effort is incentive compatible for all tasks. In this case, there is no trade-off between speed and quality of organization design, as the fastest organization has an equilibrium in which full effort is provided if any other organization has such an equilibrium.

(ii) If the cost of guessing exhibits weakly diminishing increments,

$$b \cdot c \leq \Delta(b) \tag{8}$$

is a necessary condition for full effort provision, as it guarantees proper incentives for those only working on raw data items not to shirk (given that all other items are processed properly). For the agents who handle only raw data, the above condition is sufficient for effort provision, because it is better to provide effort on all raw items rather than to shirk on a subset of raw data items if it is better to provide effort on all raw data items than to shirk on the whole set.

Next, we consider the incentives for effort provision on higher levels of information processing. Consider first the agents who are assigned one partial result in addition to the b raw data items. The partial result contains b raw data items. Hence, for providing effort once more than the raw data processing agents, agents on this level of information processing provide twice as much information. Necessary conditions for effort provision by these players are (8) and

$$(b + 1) \cdot c \leq \Delta(2b), \tag{9}$$

which is implied by (8) if $c \leq \Delta(2b) - \Delta(b)$.

¹⁵The full analysis is available from the authors upon request.

It is a general feature of the reduced tree, that on each higher level of information processing, agents working on that level have to provide effort one more time than the agents on the next lower level and therewith provide twice as much information as those agents. In addition to the incentive compatibility constraints faced by his subordinates, an agent processing x partial results faces the constraint

$$(b + x) \cdot c \leq \Delta(2^x b), \quad (10)$$

which is implied by the lower level incentive compatibility constraints if $c \leq \Delta(2^x b) - \Delta(2^{x-1}b)$. Hence, (8) implies (10) for all $x \geq 1$ if the cost of guessing satisfies the following property, which demands that the costs are not “too concave”.

Property (*)

$\frac{\Delta(m)}{m} \leq \Delta(2m) - \Delta(m) \forall m \in \{2, \dots, b\}$ and $\Delta(2^x m) - \Delta(2^{x-1}m)$ is non-decreasing in x .

Proposition 1 *Let $\Delta(m)$ fulfill Property (*). Consider a reduced tree with P managers, where $\frac{n}{P} = b \in \mathbb{N}$. A full effort equilibrium exists if and only if*

$$c \leq \Delta(m)/m \forall m \in \{1, \dots, b\}. \quad (11)$$

Proof. It remains to be shown that (11) is sufficient for the existence of a full effort equilibrium. Consider first the case that the cost of guessing exhibits weakly increasing increments. $c \leq \Delta(1)$ is necessary and sufficient for effort provision. We prove the proposition for the case of diminishing increments by induction. If the cost of guessing has Property (*), $c \leq c(b)$ implies that all players on all levels prefer to provide full effort rather than to shirk on all their assigned tasks. However, we also need to consider a possible deviation on a subset of tasks. If so, a player would choose to shirk on a subset of tasks for which the cost of guessing is lowest, i.e. those with the least information content. Hence, as he prefers to provide full effort to no effort, he will at least provide effort on the last item he is supposed to read. Given that he will provide effort on the last task, the incentive compatibility constraints for the remaining tasks are implied by those of his subordinates. Then, given that raw data processing agents provide effort, their direct superiors provide effort and so do theirs, till the top manager. Q.E.D.

Thus, in order to induce all players in the reduced tree to provide effort, one only has to take care of the incentives of lowest level players. An immediate consequence is stated in Corollary 1.

Corollary 1 *If $\Delta(m)$ exhibits diminishing increments and has Property (*), the parameter range for which a full effort equilibrium exists can be increased by increasing the number of information-processing managers in the reduced tree, until $\frac{n}{P} = 2$. If $\Delta(m)$ exhibits weakly increasing increments, the parameter range for which a full effort equilibrium exists is not affected by increasing the number of managers.*

Hence, if the cost of guessing exhibits weakly diminishing increments and has Property (*), by decentralizing information processing tasks in the way proposed by Radner (1993), proper incentives can be provided by involving enough agents. It is easy to show that this result holds even more generally. Any more decentralized organization, in the sense that tasks are reassigned to new members who report to the old members, has a full effort equilibrium if the original organization has, but not necessarily vice versa.

Note that for information processing problems which satisfy Property (*), if it is not possible to provide proper incentives to the members of a reduced tree, then there exists no hierarchy design with which this is possible. This is due to the fact that in the reduced tree it suffices to provide incentives for raw data processing only, which is a necessary condition for effort provision in any network. Since this (sub)set of individual tasks (for a given number of managers) is smallest in the reduced tree, it is not possible to provide better incentives.

6 Efficient organization design

An important consequence of our analysis concerns the efficient size of an organization. Radner (1993) points out that some hierarchy sizes, i.e. the number of managers P , are inefficient for the aggregation of n data items, since the same delay may be achieved with a smaller number of managers. For example, if 10 objects are to be processed, the reduced tree yields a delay of 5 units of time for any $P \in \{3, 4, 5\}$. Hence, a hierarchy employing more than 3 managers is inefficient for this task in Radner's framework. However, when taking into account incentives for information provision, employing 5 managers may be efficient. This is the case if $c \in (\Delta(3)/3, \Delta(2)/2]$, that is if the reduced tree with 3 managers does not have a full effort equilibrium, but the reduced tree with 5 managers does.

In our extended framework, in an organization in which every player provides full effort, the number of managers P is inefficiently high, only if there is a smaller number of managers achieving the same delay and yielding a full effort equilibrium. That is the

case if (i) $\lfloor \frac{n}{P} \rfloor + \lceil \log_2 (P + n \bmod P) \rceil = \lfloor \frac{n}{P-1} \rfloor + \lceil \log_2 (P-1 + n \bmod (P-1)) \rceil$ and (ii) $\Delta \left(\left\lceil \frac{n}{P-1} \right\rceil \right) / \left\lceil \frac{n}{P-1} \right\rceil \geq c$.

Proposition 1 enables us to characterize the set of efficient organizations. It is important to note that if the cost function $\Delta(m)$ exhibits weakly diminishing increments, every organization which does not have a full effort equilibrium does not have an equilibrium in which any effort is provided. Therefore, it can be replaced by the trivial organization in which the aggregate state of the world x is guessed without delay. Hence, we may restrict the analysis of efficient organizations to those which have a full effort equilibrium.¹⁶

Our main result is that any efficient outcome can be achieved with a reduced tree.

Proposition 2 *Let $\Delta(m)$ have Property (*) and let $P = \frac{n}{b}, b \in \mathbb{N}^+$. If the reduced tree with P managers does not have a full effort equilibrium for the processing task, then there exists no programmed network with P managers which does.*

Every reduced tree produces at least the same speed and classical surplus as any other organization with P managers.

Proof. We will show that any programmed network \tilde{N} which has a full effort equilibrium for an information processing task which satisfies Property (*) can be replaced by a reduced tree $R(\tilde{N})$ with the same number of processors that also has a full effort equilibrium for this task. Therefore $R(\tilde{N})$ generates the same classical surplus and a (weakly)¹⁷ faster decision. Again, we distinguish the cases that (i) the cost of guessing exhibits weakly increasing increments, and (ii) it exhibits weakly diminishing increments. We again restrict attention to the case that n is a multiple of P , $n = bP, b \in \mathbb{N}^+ \setminus \{0, 1\}$.

(i) For the case that the cost of guessing exhibits weakly increasing increments, the result follows straightforward from the fact that the necessary and sufficient condition for effort provision is identical for all organizations.

(ii) Let there be a programmed network \tilde{N} of P managers which has a full effort equilibrium for processing $n = bP$ objects. In this network, there must be at least one agent handling at least b raw data items. Since the cost function $\Delta(m)$ exhibits weakly diminishing increments, the maximum benefit for these tasks is $\Delta(b)$. Hence, it must be true that $c \leq \Delta(b)/b$. This in turn is the sufficient condition for the existence of a full effort

¹⁶A time constraint may force an organization, for which a full effort equilibrium exists, to limit the amount of information which is processed properly to $n' < n$. We briefly discuss this issue at the end of this section.

¹⁷The reduced tree is in general not the unique delay-minimizing organization.

equilibrium in a reduced tree $R(\tilde{N})$ with the same number of managers, if the information processing problem has Property (*). Q.E.D.

Hence, any organization producing classical surplus \bar{v} coexists with a reduced tree that also produces \bar{v} with a weakly lower delay. However, these organizations are not efficient if the same delay can be achieved with a smaller number of managers without destroying incentives.

Definition 3 *Consider an information processing problem which satisfies Property (*). An organization with P managers is called P -efficient if there exists no $P' < P$ such that (i) $\lfloor \frac{n}{P'} \rfloor + \lceil \log_2(P' + n \bmod P') \rceil \leq \lfloor \frac{n}{P} \rfloor + \lceil \log_2(P + n \bmod P) \rceil$ and (ii) $c \leq \frac{\Delta(\lceil \frac{n}{P'} \rceil)}{\lceil \frac{n}{P'} \rceil}$.*

We have that:

Corollary 2 *Let $\Delta(m)$ have Property (*).*

- (i) *Every P -efficient reduced tree with a full effort equilibrium is efficient.*
- (ii) *Every organization which is not P -efficient is dominated by a P -efficient reduced tree.*

In order to determine the part of the efficiency frontier for which the whole set of information is processed properly, we may restrict attention to P -efficient reduced trees with a full effort equilibrium. A P -efficient reduced tree may coexist with other organizations producing the same classical surplus with the same delay. However, there is a class of problems for which the efficient organization is unique.

Corollary 3 *Consider an information aggregation problem which has Property (*), $n = b2^a$, $b, a \in \mathbb{N}^+ \setminus \{0, 1\}$, and $c \leq \Delta(b2^{a-j})/b2^{a-j}$, for some $j \in \{1, \dots, a\}$. An organization with 2^j managers is efficient if and only if it is a reduced tree.*

Proof. From Radner (1993) we know there exists no organization with 2^j managers except for the reduced tree which achieves the minimum delay. It remains to be shown that the reduced tree with 2^j managers is P -efficient. It is enough to show that $\lfloor \frac{b2^a}{2^j-1} \rfloor + \lceil \log_2(2^j - 1 + b2^a \bmod (2^j - 1)) \rceil > b2^{a-j} + j$. Note that $\lfloor \frac{b2^a}{2^j-1} \rfloor \geq b2^{a-j}$ and that $\lceil \log_2(2^j - 1 + b2^a \bmod (2^j - 1)) \rceil = j$ only if $b2^a \bmod (2^j - 1) \leq 1$ and $> j$ else. If $b2^a \bmod (2^j - 1) \leq 1$, then $\lfloor \frac{b2^a}{2^j-1} \rfloor > b2^{a-j}$. Q.E.D.

A time constraint may force the organization to disregard some of the information. The minimum delay within which a set of n items can be processed is $1 + \lceil \log_2 n \rceil$.¹⁸ If a decision

¹⁸See Radner (1993).

has to be made within a shorter period of time, the organization is forced to only partially (if at all) process the set of relevant information. If the cost function $\Delta(m)$ exhibits diminishing increments, incentives are diluted by restricting the information processing to a subset of the relevant information.

Suppose the decision must be made till time unit $T < 1 + \lceil \log_2 n \rceil$. If society is large enough such that the welfare gain of better decisions always outweigh the processing cost, it is clear that in the limited time span as much information as possible should be processed. In T units of time, 2^{T-1} information items can be processed properly in a reduced tree (with 2^{T-2} managers) if the cost of effort provision is sufficiently small such that a full effort equilibrium for this task exists. This defines the efficiency frontier for delays smaller than $1 + \lceil \log_2 n \rceil$.

Figure 1 depicts the efficiency frontier for an information aggregation problem with Property (*). In this example, there are $n = 16$ data items to be aggregated. The maximal classical surplus v^* is obtained if a single manager works on the whole set of data items. Adding a manager decreases the delay of decision making to 9, but also reduces the classical surplus to $v^* - c$, since there is an additional variable (the report) to be read. The delay cannot be reduced below 5 by increasing the number of managers beyond 7 if all data items shall be processed. In order to obtain a smaller delay, the organization is forced to guess part of the data and limit information processing activities to the data which can be handled within the given period of time. Note that if $\Delta(\cdot)$ exhibits diminishing increments, c must be small enough to sustain effort provision as an equilibrium, because the marginal benefit of effort provision is smaller if part of the data is not processed properly.

7 Conclusion

Parallel processing is widely used in practice. Many organizations, firms and government agencies delegate subtasks to different individuals who simultaneously process information. One advantage of decentralization is that it decreases the time span within which the organization fulfills its tasks. The present paper shows that for a large class of information aggregation problems, parallel processing also yields desirable high quality decisions.

Our analysis merges two recent strands of the literature on organization design: the game-theoretical analysis of institutions and the theory of the efficient design of programmed networks. We identify an elementary advantage of decentralized structures:

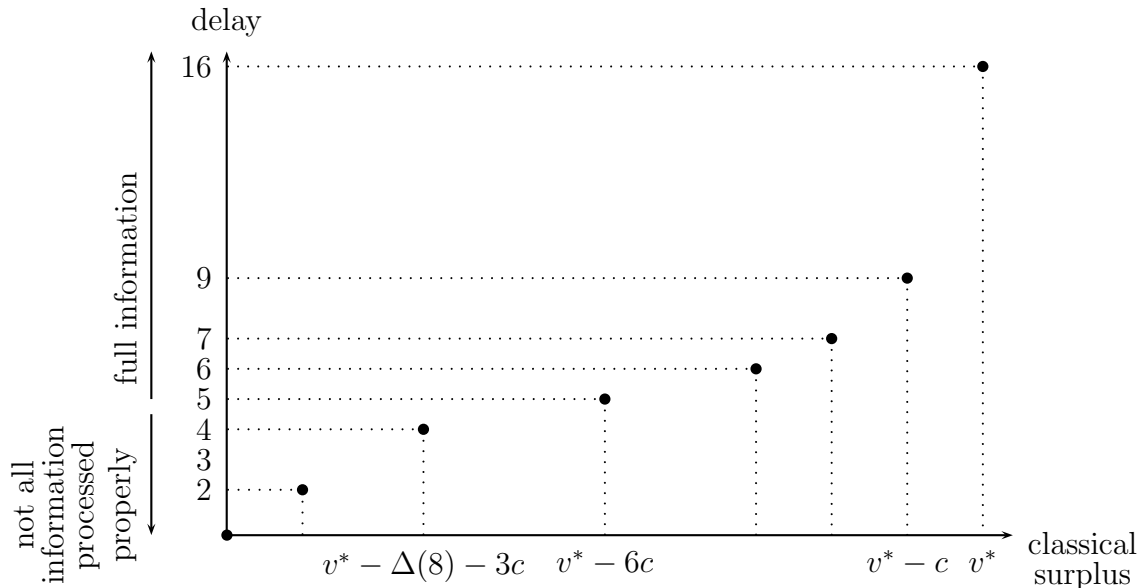


Figure 1: Efficiency frontier for an information processing problem with Property (*), $n = 16$.

these structures provide better incentives for self-interested managers. For a large class of managers' preferences and stochastic structures of the information aggregation problem, the division of tasks favors equilibria where all players provide effort.

Our second main result is that the speed of a decision procedure and the quality of the decision need not be conflicting objectives when it comes to the evaluation of organization designs. We find that reduced trees à la Radner (1993) outperform other arrangements along both dimensions. Information is aggregated in parallel. Therefore, in the course of a manager's processing activity, the information content of the data he is supposed to read is increasing. Those managers who have to work most also have the most important jobs. Hence, Radner's efficiency result is robust with respect to more complex behavioral assumptions (see also Schulte and Grüner, 2007).

An interesting extension of our work would be to study a framework in which the cost of guessing is not independent of information previously acquired, e.g. a framework with Bayesian updating. In such an information aggregation problem, the value of an information item is path-dependent, and it can be beneficial to approach the agents sequentially. It has been shown in Smorodinsky and Tennenholtz (2005) and in Gershkov and Szentes (2007) that limiting the agents' information (concerning their position in the sequence and the values of partial results) enhances their incentives for information acquisition. In both papers incentive problems only arise when learning the initial data

whereas the aggregation operations are not subject to incentive constraints. Moreover, decision delay is not costly per se. It is certainly worth studying this class of information aggregation problems in our extended framework, therewith introducing the dimension time also into the literature on committee decision making.

There are several further useful extensions of the present framework. One would be to include different conflicts of interest. In the present paper the only conflict of interest among agents arises from the individual disutility from providing effort. In many interesting applications there is also some disagreement about the best decision even if there is perfect information. A second approach would be to look at alternative verifiability structures. In this paper, we considered problems for which it is difficult (or impossible) to implement a monetary transfer scheme to insure effort provision. If some data is verifiable ex post, then monetary incentives may be used and the set of efficient organizations might look different. One could also link the cost of information aggregation to the complexity of the task. Finally, it would be interesting to study operations that cannot be permuted without altering the results. All these extensions may prove useful for extending the range of applications of the present framework.

Appendix

In the appendix, we provide concrete examples for general framework presented in the main part of the paper. They correspond to tasks such as assessing the aggregate willingness to pay for a public project or the total demand of a multinational firm (Examples 1-3), or to choosing an alternative out of a finite set of alternatives (Example 4). We derive the cost functions $\Delta(m)$ for these problems and show that they have Property (*).

Example 1: Addition of normally distributed random variables

Consider a problem in which the utility-maximizing decision is to match the aggregate state of the world, $d = x$. Agents dislike deviations from the optimal decision. The aggregate state of the world is given by the sum of n i.i.d. normally distributed random variables. We have: $X = D = \mathbb{R}$. Moreover, we may restrict the reporting spaces R_j , where $j = 1, \dots, P$, to \mathbb{R} as well. The aggregation operation is given by $f_2(x_i, x_j) = x_i + x_j$, and $f_n(\mathbf{x}) = \sum_{i=1}^n x_i$.

Agents' preferences are represented by the following utility function:

$$u(d, x) = -|d - x|.$$

The aggregate x is the sum of n normally distributed random variables, $x_i \sim N(0, \sigma_{x_i}^2)$, $i = 1, \dots, n$. We have $x \sim N(0, \sigma^2)$, where $\sigma^2 = n\sigma_{x_i}^2$. The best guess of the total (or a partial) sum is the median of its distribution.¹⁹

Hence, the cost of guessing m numbers is:

$$\Delta(m) = \int_{-\infty}^{\infty} |x| \cdot \frac{1}{\sigma(2\pi)^{\frac{1}{2}}} e^{-\frac{1}{2\sigma^2}x^2} dx = 2\frac{\sigma}{(2\pi)^{\frac{1}{2}}}. \quad (12)$$

Using $\sigma^2 = m\sigma_{x_i}^2$, we get

$$\Delta(m) = \sqrt{\frac{2m}{\pi}}\sigma_{x_i}, \quad (13)$$

which is a strictly increasing, concave function in m . Property (*) holds for $m \geq 3$:

$$\begin{aligned} 2\Delta(2m) &\leq \Delta(4m) + \Delta(m) \\ \Leftrightarrow 2\sqrt{\frac{4m}{\pi}}\sigma_{x_i} &\leq \sqrt{\frac{8m}{\pi}}\sigma_{x_i} + \sqrt{\frac{2m}{\pi}}\sigma_{x_i} \\ \Leftrightarrow 4 &\leq 3\sqrt{2}, \end{aligned}$$

and

$$\begin{aligned} \frac{\Delta(m)}{m} &\leq \Delta(2m) - \Delta(m) \Leftrightarrow \sqrt{\frac{2}{\pi m}}\sigma_{x_i} \leq \sqrt{\frac{4m}{\pi}}\sigma_{x_i} - \sqrt{\frac{2m}{\pi}}\sigma_{x_i} \\ \Leftrightarrow \frac{1}{\sqrt{2}-1} &\leq m. \end{aligned}$$

Example 2: Quadratic utility

Consider next the above problem of adding random variables but with quadratic utilities,

$$u(d, x) = -(d - x)^2.$$

In a setup with quadratic losses the best guess is the expected value and the loss equals the variance of the error. Given that the x_i are i.i.d. $\Delta(m)$ is linear in m :

$$\Delta(m) = m\sigma_{x_i}^2. \quad (14)$$

It is easy to verify that this specification also satisfies Property (*).

¹⁹See DeGroot (1970), p. 232.

Example 3: Binary distribution

Now consider a task of adding n information variables $x_i \in \{0, 1\}$, where the information variables are independent random variables, and each outcome is equally likely. Again, let the utility function assume the form

$$u(d, x) = -|d - x|.$$

The expected utility-maximizing guess of the sum of m numbers is the median of its distribution, $\frac{m}{2}$. The cost of guessing m numbers are:

$$\Delta(m) = E\left(\left|\frac{m}{2} - \sum_{i=1}^m x_i\right|\right). \quad (15)$$

Note that the true result assumes the value $i < \frac{m}{2}$ with probability $\frac{1}{2^m} \binom{m}{i}$. The loss of utility in this case is $\frac{m}{2} - i$. Note that deviations are symmetric. Hence:

$$\Delta(m) = 2 \sum_{i=0}^{\lfloor \frac{m}{2} \rfloor} \frac{1}{2^m} \binom{m}{i} \left(\frac{m}{2} - i\right) \quad (16)$$

$$= \frac{1}{2^{m-1}} \sum_{i=0}^{\lfloor \frac{m}{2} \rfloor} \binom{m}{i} \left(\frac{m}{2} - i\right). \quad (17)$$

$\Delta(m)$ is a weakly increasing step-function with diminishing increments. One can show that $\Delta(m)$ has Property (*). The proof involves some algebra and is available upon request.

Example 4: Project selection

Our next example relates to situations in which a project has to satisfy several criteria in order to be useful. Consider the task of selecting an object out of a set of objects O , where $|O| = o$. Each object is either useful or not. An object is useful if and only if it has n properties, $n < o$. There are n information items available, each of which identifies $x_i \subset O$, the set of objects which have property i . The reporting spaces are given by the powerset of O . The aggregation task is to exclude objects which are not useful, i.e. $f_2(x_i, x_j) = x_i \cap x_j$. The aggregate state of the world x is given by the set of useful objects, $x = \bigcap_{i=1, \dots, n} x_i$. The utility from selecting object j is:

$$u(j) = \begin{cases} 1, & \text{if } j \in x \\ 0, & \text{else.} \end{cases}$$

If m out of n information variables are not processed properly, the best guess is a random choice out of the set $x_1 \cap x_2 \dots \cap x_{n-m}$. It is easy to specify the information environment such that the cost of guessing has Property (*). For example, assume that each object j has at least $n - 1$ properties, and that each property is satisfied by all but one object. The probability that the randomly chosen object is useful is $(o - n)/(o - n + m)$. Hence, the cost of guessing is

$$\Delta(m) = \frac{m}{o - n + m}.$$

Simple calculations verify that $\Delta(m)$ satisfies Property (*), if o is sufficiently larger than n .

References

- [1] Aghion, Philippe and Jean Tirole (1997): Formal and Real Authority in Organizations, *Journal of Political Economy* 105(1):1-29.
- [2] Bolton, Patrick and Mathias Dewatripont (1994): The Firm as a Communicating Network, *Quarterly Journal of Economics* 109:809-839.
- [3] Crémer, Jacques (1980): A Partial Theory of the Optimal Organization of a Bureaucracy, *Bell Journal of Economics* 11:683-693.
- [4] DeGroot (1970): Optimal Statistical Decisions, McGraw-Hill Book Company, New York.
- [5] Dewatripont, Mathias and Jean Tirole (2005): Modes of Communication, *Journal of Political Economy* 113:1217-1238.
- [6] Garicano, Luis (2000): Hierarchies and the Organization of Knowledge in Production, *Journal of Political Economy* 108(5), 874-904.
- [7] Garicano, Luis and Esteban Rossi-Hansberg (2004): Inequality and the Organization of Knowledge, *American Economic Review*, 94, 197-202.
- [8] Garicano, Luis and Esteban Rossi-Hansberg (2006): Organization and Inequality in a Knowledge Economy, *Quarterly Journal of Economics* 121:1383-1435.
- [9] Gerardi, Dino and Leeat Yariv (2008): Information Acquisition in Committees, *Games and Economic Behavior* 62:436-459.

- [10] Gershkov, Alex and Balász Szentes (2007): Optimal Voting Schemes with Costly Information Acquisition, *Journal of Economic Theory*, forthcoming.
- [11] Jehiel, Philippe (1999) Information Aggregation and Communication in Organizations, *Management Science*, 45:659-669.
- [12] Meagher, Kieron, Hakan Orbay and Timothy Van Zandt (2003): Hierarchy Size and Environmental Uncertainty, 2003, in: *Advances in Economic Design*, M.R. Sertel and S. Koray (Eds.), Springer-Verlag.
- [13] Meagher, Kieron and Timothy Van Zandt (1998): Managerial Costs for One-shot Decentralized Information Processing, *Review of Economic Design* 3:329-345.
- [14] Meagher, Kieron (2003): “Generalizing incentives and loss of control in an optimal hierarchy: the role of information technology”, *Economics Letters* 78(2):273-280.
- [15] Melumad, Nahum D., Dilip Mookherjee and Stefan Reichelstein (1995): Hierarchical Decentralization of Incentive Contracts, *RAND Journal of Economics*, 26:654-672.
- [16] Mookherjee, Dilip and Stefan Reichelstein (1997): Budgeting and Hierarchical Control, *Journal of Accounting Research* 35(2):129-155.
- [17] Mukhopadaya, Kaushik (2003): Jury Size and the Free Rider Problem, *The Journal of Law, Economics, and Organization* 19(1):24-44.
- [18] Orbay, Hakan (2002): Information Processing Hierarchies, *Journal of Economic Theory* 105:370-407.
- [19] Persico, Nicola (2004): Committee Design with Endogenous Information, *Review of Economic Studies* 71(1):165-194
- [20] Prat, Andrea (1997): Hierarchies of Processors with Endogenous Capacity, *Journal of Economic Theory* 77:214-222.
- [21] Radner, Roy (1992): Hierarchy: The Economics of Managing, *Journal of Economic Literature* 30:1382-1415.
- [22] Radner, Roy (1993): The Organization of Decentralized Information Processing, *Econometrica* 61(5):1109-1146.

- [23] Radner, Roy and Timothy Van Zandt (1992): Information Processing in Firms and Returns to Scale, *Annales d'économie et de statistique* 25/26:265-298.
- [24] Rosen, Sherwin (1982): Authority, control, and the distribution of earnings, *Bell Journal of Economics* 13:311-323.
- [25] Schulte, Elisabeth and Hans Peter Grüner (2007): Speed and Quality of Collective Decision Making: Imperfect Information Processing, *Journal of Economic Theory* 134:138-154.
- [26] Smorodinsky, Rann and Moshe Tennenholtz (2005): Overcoming Free Riding in Multi-Party Computations—The Anonymous Case, *Games and Economic Behavior* 55:385-406.
- [27] Van Zandt, Timothy (1997): The Scheduling and Organization of Periodic Associative Computation: Essential networks, *Review of Economic Design* 3:15-27.
- [28] Van Zandt, Timothy (1998): The Scheduling and Organization of Periodic Associative Computation: Efficient networks, *Review of Economic Design* 3:93-127.
- [29] Winter, Eyal (2004): Incentives and Discrimination, *American Economic Review* 94:764-773.