The Hierarchical Organization of Management in the Firm: Economic Reasons and Efficient Structures*

Jacek A. Cukrowski

April 1995

Abstract

Based on the analysis of organizational aspects of data processing in decision making, an economic explanation of the hierarchical organization of management in the firm is provided. Contrary to the prevailing opinion that 'the explanation of hierarchy may in many cases be more sociological and psychological than purely economic in the mainstream sense' (Radner, 1992, p.1384) the paper shows that decreasing returns to scale in information processing is a necessary condition for hierarchical organization of management in business firms. However, decentralization of decision-making is desirable only if an additional (sufficient) condition is satisfied, i.e., if the information workload of the decision-making sector is sufficiently large. Moreover, the paper shows that specific features of human information-processing in the firm such as disagreement about the goals of data analysis or the possibility of random errors do not imply a hierarchical organization of management, but could change forms of the efficient decision-making structures. Finally, contrary to the results recently presented in economic literature, the paper shows that in the firm, unlike in computer systems, there is no unique architecture for the efficient information-processing structures, but a number of various efficient forms can be observed.

Keywords: Information-processing, organization of the firm, decentralization, hierarchy.

Abstrakt

Na základě analýzy organizačních aspektů zpracování dat v rozhodovacím procesu se podává ekonomické vysvětlení hierarchické organizace managementu v firmě. Na rozdíl of převlaádajícího názoru, podle něhož může být vysvětlení hierarchie v mnoha případech spíše sociologoické a psychologické než čisté ekonomické ve smyslu hlavního proudu (Radner 1992, str. 1384), článek ukazuje, že klesající výnosy z rozsahu ve zpraconání informací jsou nutnou podmínkou pro hierarchickou organizaci managementu ve firmě. Decentralizace v rozhodovacím procesu je však vhodná, jen když je splněna další (postačující) podmínka, t.j. když je informační zatížení rozhodovacího sektoru dostatečně velké. Čtaánek navíc ukazuje, že specifické rysy zpracování informací člověkem ve firmě, jakými jsou například neshoda ohledně cílů analýzy dat nebo možnost náhocných chyb, neimplikují hierarchickou organizaci managementu, ale mohly by změnit formy efektivních rozhodovacích struktur. Článek nakonec ukazuje, oproti výsledkům nedávno prezentovaným v ekonomické literatuře, že ve

* Most of this work was done during the author's stay at Tinbergen Institute in Amsterdam and at the Department of Economics of New York University.

firmě neexistuje (na rozdíl od počítačových systémů) jediná stavba efektivních struktur zpracování informací, nýbrž že je možné sledovat množství různých efektivních forem.

Klíčová slova: Zpracování informací, organizace firmy, decentralizace, hierarchie.

1. Introduction

In classical microeconomic theory, the firm is viewed as a ‘black box’ transforming inputs into outputs according to a rule described by a production function. In fact, firms have complex multidivisional, usually hierarchical, structures which determine their economic performance (see, for instance, Sah and Stiglitz, 1986; Alchian and Demsetz, 1972; Milgrom and Roberts, 1990; Keren and Levhari, 1983, or Cyert, 1988). Especially large corporations are widely perceived to be organized hierarchically (see, e.g., Williamson, 1975 and 1981; or Radner, 1992). Not surprisingly therefore, the analysis of various multilevel organizational forms of business firms and their implications on different aspects of economic theory has appeared recurrently in the literature (see, for example, Simon, 1957; Lydall, 1968; Williamson, 1967; or Calvo and Wellisz, 1978 and 1979). The economic significance of the hierarchical organization of enterprises has been investigated, among others, by Sah and Stiglitz (1986), Milgrom and Roberts (1990), Williamson (1986) and Radner (1992). There also exists a notable body of economic literature concerned with organizational aspects of internal parts of large corporations such as, for example, information-processing, monitoring and control systems or decision-making sectors (see, for instance, Daft and Lengel, 1990; Van Zandt, 1990; Radner, 1992 and 1993; Radner and Van Zandt, 1992 and 1993; or Bolton and Dewatripont, 1994).

The overview of the contributions of recent research to understanding the economic significance of the hierarchical organization of information processing in the management of the firm is presented by Radner (1992). In this paper, however, he emphasizes that ‘*research to date has not provided an adequate explanation on economic grounds alone of the conditions under which one expects to see a hierarchical organization of business firms*’ (Radner, 1992, p. 1385). His analysis of data processing in decision making focuses on associative operations and so-called ‘skip-level reporting’ structures, derived by computer scientists investigating issues of parallel computing, and proven to be efficient for the computation of associative operations (see Gibbons and Rytter, 1988). However, despite the formal proof of efficiency, Radner stresses that ‘*reporting through skipped levels is not unheard of in corporate hierarchies (in fact, at AT&T this is called "skip-level reporting"), but the practice does not seem to be (...) widespread ...* ’(Radner, 1992, p. 1396; Radner, 1993, p. 1121).

The purpose of the present paper is to provide an economic justification of the hierarchical organization of information processing in the management of the firm, to explain the reasons why information processing in the firm differs from that in an idealized computer, and to show how these differences affect the efficient organizational forms of information-processing structures.

The paper is organized as follows: in Section 2, the management sector of the firm is considered. Special attention is paid to the managerial activities and particularly to information processing in decision making. Consideration focuses on associative operations[2] because the number of decision-making paradigms such as, for instance, linear decision rule, pattern matching, project selection or finding a maximum involve primarily operations of such kind (see Radner, 1992 and 1993).

In Section 3, the dynamic parallel processing model of associative computation in the firm is presented. Information processing in the firm is described as in an idealized computer, whose processors are members of the decision-making team. However, unlike in the original model, it is assumed that computations are made with the help of the capital and labor allocated to the number of distributed processing elements (computational centers). The value of the computational service is represented as a function of the capital and labor used in data analysis, and the objective of the firm in information processing is specified.

In Section 4, the concept of the returns to scale in data processing is considered, and the conditions for increasing, decreasing and constant returns to scale in information processing are formally defined.

Section 5 focuses on the decentralization of information processing in decision making. It is shown that decreasing returns to scale in data processing is the necessary condition for a decentralized (hierarchical) organization of management in business firms. However, for decentralization of decision making, an additional (sufficient) condition should also be satisfied, i.e. the information workload of the decision-making sector has to be sufficiently large. Moreover, the analysis shows that the decentralization process leads to specific hierarchical forms ('skip-level reporting' structures) which have been proven to be efficient for associative computations (see Gibbons and Rytter, 1988; or Radner, 1992).

In Section 6, the implications of human information-processing on the forms of hierarchical structures are considered. The dynamic parallel processing model of associative computation is adjusted to the modelling of a project selection in the team of decision makers, where the value of the computational service depends not only upon the delay in information processing, but also upon the error in data analysis. It is shown that specific features of human information-processing such as disagreement about the goals of data analysis or the possibility of random errors, do not imply a hierarchical organization of the

[2] A binary operation (*) is associative, if (A*B)*C=A*(B*C).

management, but significantly change the organizational forms of decision-making teams.

Section 7 focuses on the decision-making process where the value of the computational service is determined by the delay in information processing and the error in data analysis. The conditions for the efficient organization of information processing in decision making are formally defined, and the relationship between the delay in information processing, the expected value of the error in data analysis, and the architecture of efficient structures is considered. Based on the analysis of the project selection in the firm it is found that the skip-level reporting structures are efficient for computation of associative operations only if (1) the error in data analysis is not matter of concern for the firm, or (2) all the projects analyzed are identical, or (3) all processing elements have the same objectives in information processing. Moreover, it is shown that, in general, efficient organizational forms of data processing in decision making have to be determined individually, for (1) each particular decision problem, and (2) each possible combination of inputs to information processing, and, consequently, that a number of different organizational forms of information-processing can be observed in real enterprises.

2. Management in the Modern Firm

Empirical studies of the labor market in industrialized countries show that the fraction of the labor force devoted to management in economy has permanently increased since the beginning of the century (see fig.2.1). Currently, in these countries much more than one-third of employees work full time in activities not directly connected with the production process such as processing and communicating information, monitoring actions of the other members of the firm, analyzing the market, planning, training employees, making decisions, etc. (see Radner, 1992, for a detailed analysis of this issue). The common feature of all these activities is information processing, i.e., collecting and aggregating information, transforming data, presenting them in the appropriate form, etc. The majority of the activities associated with data processing in the firm is carried out by managers with the help of staff, secretaries or clerks using computation and telecommunication equipment, buildings, electricity, etc. Thus, the management sector of the firm can be described as a complex system, which takes signals from the environment and transforms them (with the help of labor and capital) into actions to be taken by the ‘real workers’.

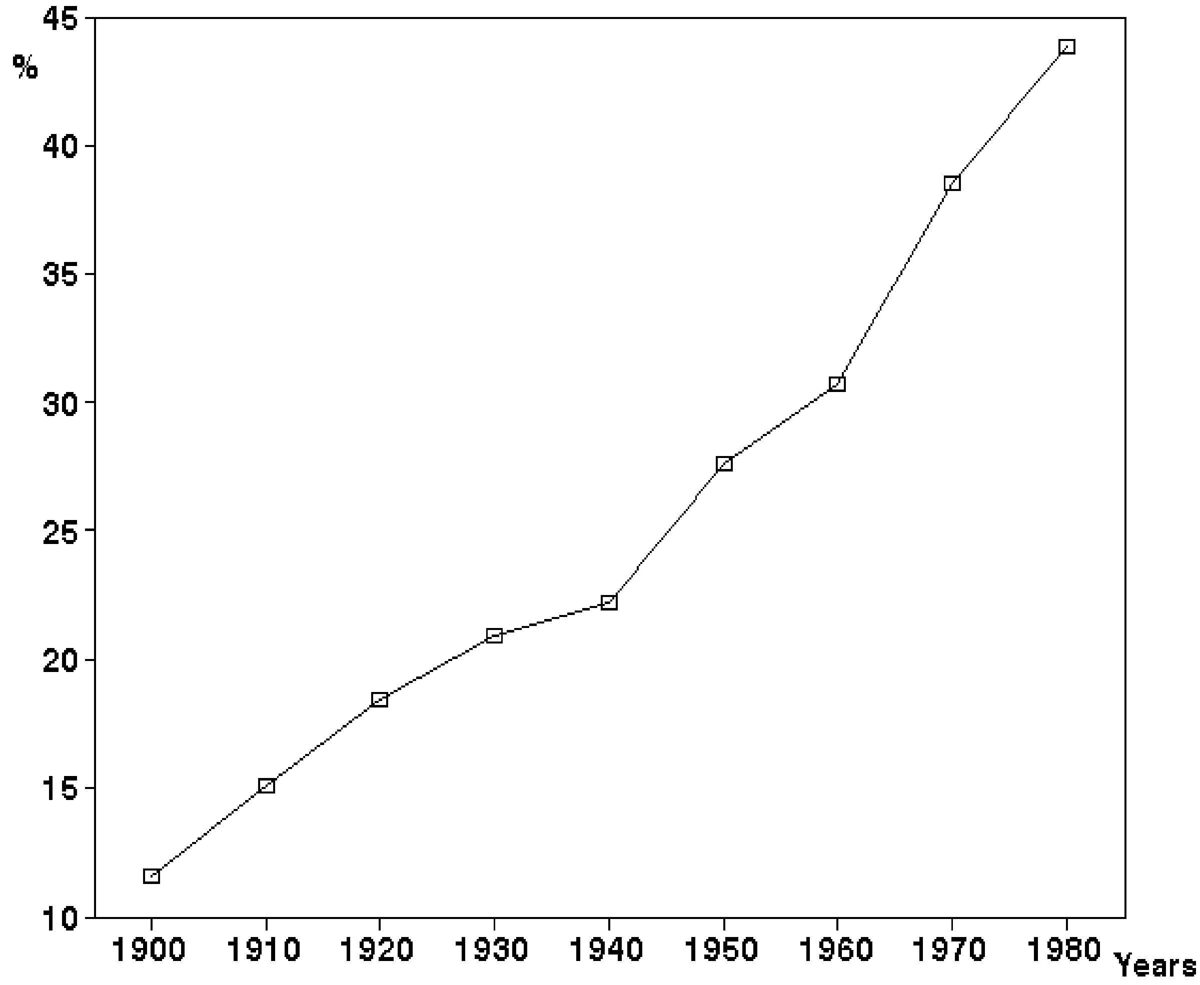


Figure 2.1.: Employment in management in the US economy[3]

The simplest example of the transformation of data from the environment into decisions is a linear decision rule, where the value of the linear function:

$$\sum_{i=1}^{N} c_i x_i,$$

is computed (c_i is a coefficient of conversion to a common unit, and x_i is a numerical data item, i=1,2,...,N). In practice, the items aggregated may not be just numbers but vectors or matrices. Computations of such a kind are

[3] Source: US Bureau of the Census 1975, 1984, Vol. 1, Ch. D, Part 1, Sec. A, Tables 253-310.

commonly used in the methods of statistical prediction or statistical control (see Marschak and Radner, 1972; Aoki, 1986; or Radner and Van Zandt, 1992 and 1993). Based on the analysis of computational processes for the purpose of predicting demand in the firm Radner and Van Zandt (1992) show that the values of the decisions made according to the linear decision rule depend on the quality of the result computed, measured by the delay in information processing.

Another widespread decision-making paradigm used in the management of business firms is a selection of the best project (see Radner 1992, 1993). In this case the projects (or the signals of a different kind) coming from the environment are evaluated, and compared with the attributes of the project (correspondingly, the signal) considered as the best by the entrepreneur. The purpose of information processing is to determine the project n^*, such that $n^*=\arg\{\mathrm{Min}\,|Q^*-Q_n|\}$, where Q_n denotes a numerical (aggregate) characteristic of the project, n (n=1,2,...,N, where N is the number of projects analyzed), and Q^* is a numerical attribute of the project wanted by the entrepreneur (see section 6 for details). If the analysis is carried out in the team of decision-makers then (due to the misunderstanding of the purpose of analysis or to the private incentives of the decision-makers) the values of the attributes of the projects considered as the best by individual decision makers can deviate from Q^*. Moreover, through random mistakes, the characteristics of the projects analyzed, Q_n (n=1,2,...,N), could be evaluated imprecisely. Consequently, the value of the computational service in project selection is determined not only by the delay in information processing, but also by the error in data analysis (i.e., the absolute value of the difference between a numerical characteristic of the project wanted by the entrepreneur, Q^*, and the project selected, Q_{n^*}).

It has been shown that other common decision-making paradigms such as pattern matching or finding a maximum (or a minimum) are similar to the simple addition or to the project selection in the sense that they all involve primarily associative operations (see Radner, 1992), and, consequently, all of them can be described in the conceptual framework of the parallel processing model of associative computation used in economic literature for the modelling of information processing in the firm (Radner, 1992 and 1993; Radner and Van Zandt, 1992 and 1993).

3. The Model of Information Processing in Decision Making

Consider the decision-making sector in the firm in which decisions are made based on data analysis. The value of the decisions, and, consequently, the value of the computational service provided depend on how good the resulting decisions are compared to how good they would be without the service.

Assuming that the linear decision rule is used in the decision-making process, the value of the computational service depends only upon the delay in data analysis (more precisely they are inversely proportional to the delay in information processing).[4]

To simplify the analysis of the delay in the computational process, assume that the linear decision rule requires summations of cohorts of N items of data (conversion to a common unit is not required, i.e., $c_i=1$, for $i=1,2,...,N$), and that the decision-making system considered works in a one-shot regime, i.e., delays between subsequent cohorts of data coming into the system are greater (or at least equal) to the time of a single cohort processing (it ensures that queues of data in the information-processing structure cannot arise).

Following Radner (1992 and 1993), represent the computational process in the decision-making sector of the firm as in idealized computer, i.e., assume that each processing element (a computational center) is modelled as a processor which contains an infinite memory where data are stored (called a buffer) and a register where summations are made. Each processor can read a single item of data from its memory and add the value to the register, resetting it equal to the resulting sum (errors in computation are not allowed). Loading and adding a single datum to the contents of the register is called an operation. The time is assumed to be the same whatever the values of data added are, or when a datum is added to the cleared register (i.e., to zero). A processor can send the contents of its register to an output or to the buffer of any other processor (through a communication channel) in zero time, i.e., it is assumed that communication does not need time (see Radner and Van Zandt, 1992, for details).

Each processor has a limited capacity (i.e. a maximum computational power), in that there is a maximum number of operations it can compute per unit of time. In business firms, however, the computational power of each processing element depends upon the capital and labor allocated to it. The relationship between the resources allocated to the processing element and the number of operations it can compute in a unit of time is determined by the existing technology of information processing, and can be written in functional form as $F(k,l)$[5]: $R_+ x R_+ \rightarrow R_+$, where $F(k,l)$ is continuous, twice differentiable and strictly concave in k and l.

Consequently, the duration of a single operation (d) is also a function of the

[4] See Radner and Van Zandt (1992) for a detailed analysis of this issue.

[5] F(k,l) is called an information-processing function, and is understood as a ’production function’ in information- processing.

capital (k) and labor (l) employed in the processing element ($d(k,l)=1/F(k,l)$).

In any information-processing structure, the delay in the summation of N items of data (D_N) is proportional to the duration of individual operations, and, consequently, is a decreasing function of the resources allocated to the computational structure, i.e., $\delta D_N(K,L)/\delta K<0$ and $\delta D_N(K,L)/\delta L<0$.

Assuming that data items are not costly, the total cost of the computational process, C(K,L), is determined as

$$C(K,L) = r\,K + w\,L,$$

where

- wL is a cost of labor involved in the computation (w denotes the price of labor);
- rK is a cost of capital (r is the price of capital).

The consideration above implies that the objective of the firm in decision making is to maximize the difference between the value of the computational service, $V(D_N(K,L))$, and the cost of the resources used in computation, C(K,L).

4. Economies of Scale in Information Processing

Several authors, notably Keren and Levhari (1989 and 1983), Radner and Van Zandt (1992 and 1993) and Bolton and Dewatripont (1994) have been analyzed the problem of returns to scale in data processing. A detailed analysis of this issue, in the framework of the dynamic parallel processing model of associative computation, has been presented by Radner and Van Zandt (1992 and 1993). The authors understood returns to scale in information processing as an inquiry, whether, by multiplying the size of the information workload (N) as well as all inputs to information processing (where the processors were considered as the only scarce resources in information processing) by a certain constant, the same quality of the result (measured by the delay in information processing) can be obtained.

In the model analyzed in this paper, neither the processors nor the computational centers, but the capital and labor allocated to them are considered as inputs to information processing. Thus, in the framework of the model under study, one can say that if the capital (K) and labor (L) allocated to data processing as well as the information workload (N) are multiplied by the same constant, say $\alpha>1$ (or $0<\beta<1$), then the information-processing system faces:

(a) increasing returns to scale, if the quality of the result computed increases, i.e., $D_N(K,L)>D_{\alpha N}(\alpha K,\alpha L)$, (correspondingly, decreases, if $0<\beta<1$, i.e.,$D_N(K,L)<D_{\beta N}(\beta K,\beta L)$);

(b) constant returns to scale, if the quality of the result computed does not change, i.e., $D_N(K,L)=D_{\alpha N}(\alpha K,\alpha L)$, (correspondingly, $D_N(K,L)=D_{\beta N}(\beta K,\beta L)$);

(c) decreasing returns to scale, if the quality of the result computed decreases, i.e., $D_N(K,L)<D_{\alpha N}(\alpha K,\alpha L)$, (correspondingly, increases, if $0<\beta<1$, i.e.,$D_N(K,L)>D_{\beta N}(\beta K,\beta L)$).

5. The Decentralization of Information-Processing in the Firm

Consider the decision-making sector in which decisions are made according to the linear decision rule (based on the summation of cohorts of N items of data), and assume that the amounts of resources allocated to information processing are fixed (the cost of computation is fixed as well). In this case the objective of the firm is to organize data analysis in decision making in a way in which it maximizes the values of the decisions made based on the computational service ($V(D_N)$), or in other words, in which it minimizes the delay in information processing (D_N).

Suppose, in the beginning, that all the resources used in decision-making are allocated to a single computational center ($P=1=2^0$). If the technology of information processing is such that $D_N(K,L)>D_{\beta N}(\beta K,\beta L)$, i.e., if the firm faces decreasing returns to scale in information processing, then a better quality result can be computed if the inputs to information processing (K and L), as well as the information workload (N), are reduced. Thus, sums of N/2 data items[6] can be computed in two separate computational centers (with equally divided resources, $\beta=1/2$) with less delay than the sum of N data items in the original structure. However, a linear decision rule requires the sum of N data items. Consequently, the computational centers have to be connected and one additional operation has to be made in order to summarize the partial results computed. Therefore, the decentralization of information-processing is desirable only if $D_N(K,L)>D_{N/2}(K/2,L/2)+d(K/2,L/2)$, where $d(K/2,L/2)$ is the duration of the last operation, i.e., if

[6] For simplicity it is assumed that the number of data items processed (N) is such that ($N \bmod 2^m$)=0, for $m=0,1,...,\log_2(N/2)$.

$$Nd(K,L)=\frac{N}{F(K,L)}\ >\ \frac{N}{2}d(\frac{K}{2},\frac{L}{2})+d(\frac{K}{2},\frac{L}{2})=\frac{\frac{N}{2}+1}{F(\frac{K}{2},\frac{L}{2})}\quad .$$

The inequality above is satisfied if

$$N>\frac{F(K,L)}{F(\frac{K}{2},\frac{L}{2})-\frac{1}{2}F(K,L)}\quad .$$

If the decentralized computational system (with $P=2^1$ processing elements) faces decreasing returns to scale in information processing, then the delay in the summation of N/2 data items in the structure, in which the resources K/2 and L/2 are allocated to $P=2^1$ processing elements, is smaller than the delay in the computation of the sum of N data items in the structure with $P=2^1$ processing elements and the entire resources. If the information workload (N) and the resources (K,L) are divided in two equal parts, and sums of N/2 items of data are computed in two identical structures, then the top-level computational centers of these structures have to be connected and one additional operation has to be made in order to add the partial sums. The duration of this operation equals $1/F(K/2^2,L/2^2)$. Consequently, the decentralization of the structure with $P=2^1$ processing elements is desirable only if

$$\frac{\frac{N}{2}+1}{F(\frac{K}{2},\frac{L}{2})}\ >\ \frac{\frac{N}{2^2}+1}{F(\frac{K}{2^2},\frac{L}{2^2})}+\frac{1}{F(\frac{K}{2^2},\frac{L}{2^2})}=\frac{\frac{N}{2^2}+2}{F(\frac{K}{2^2},\frac{L}{2^2})}\quad ,$$

i.e., when

$$N>\frac{2[2F(\frac{K}{2},\frac{L}{2})-F(\frac{K}{2^2},\frac{L}{2^2})]}{F(\frac{K}{2^2},\frac{L}{2^2})-\frac{1}{2}F(\frac{K}{2},\frac{L}{2})}\quad .$$

The first two steps of the decentralization process are presented in fig. 5.1.

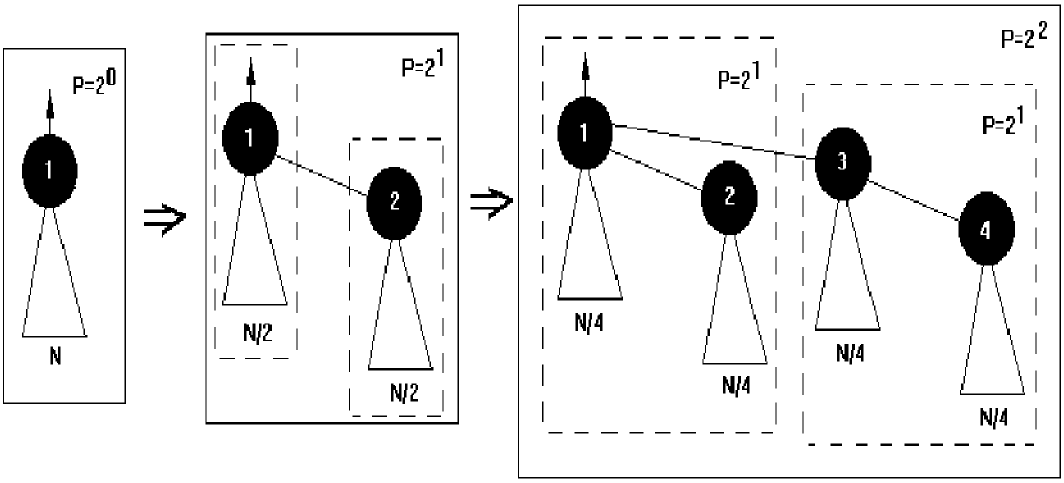


Figure 5.1.: The first steps of the decentralization process (circles denote computational centers, and triangles represent the information workload)

Decentralization can be continued (if the corresponding conditions are satisfied) until the number of the processing elements in the structure equals P=N/2 (P is bounded because at least two data items have to be assigned to each computational center).

One can see that the decentralization process produces so-called 'skip-level reporting' structures. Hierarchical forms of such kind contain P (where P is a power of 2) processing elements organized in hierarchical (multilevel) formations where each computational center has one immediate subordinate at every lower level[7] of the hierarchy. It has been proven that these structures minimize the delay in information processing for a given number of processing elements, i.e., that they are efficient for the computations of associative operations (see, Gibbons and Rytter, 1988, or Radner, 1993).

After m ($m \leq \log_2(N/2)$) steps of the decentralization process, the skip-level reporting structure contains $P=2^m$ processing elements, and $N/2^m$ data items are assigned to each of them. The delay in the summation of N data items in such

[7] The processor belongs to the level

$$X = \begin{cases} 0, \text{ if it does not have any subordinate processors;} \\ x+1, \text{ otherwise;} \end{cases}$$

where x denotes the highest level of the hierarchy to which one of its immediate subordinate processors belongs.

a structure (if N is a multiple of P) is determined as[8]

$$D_N(K,L)=\left(\frac{N}{P}+\log_2 P\right)d\left(\frac{K}{P},\frac{L}{P}\right)=\frac{\frac{N}{2^m}+m}{F\left(\frac{K}{2^m},\frac{L}{2^m}\right)}.$$

Therefore, if the system with $P=2^m$ ($m<\log_2(N/2)$) processing elements faces decreasing returns to scale in information processing (it is a necessary condition), and if

$$\frac{\frac{N}{2^m}+m}{F\left(\frac{K}{2^m},\frac{L}{2^m}\right)} > \frac{\frac{N}{2^{m+1}}+m}{F\left(\frac{K}{2^{m+1}},\frac{L}{2^{m+1}}\right)}+\frac{1}{F\left(\frac{K}{2^{m+1}},\frac{L}{2^{m+1}}\right)}=\frac{\frac{N}{2^{m+1}}+m+1}{F\left(\frac{K}{2^{m+1}},\frac{L}{2^{m+1}}\right)},$$

then it is desirable to expand the information-processing sector and allocate the resources to $P=2^{(m+1)}$ computational centers.

The inequality above is satisfied if the number of data items processed (N) is such that

$$N > \frac{2^m\left[(m+1)\,F\left(\frac{K}{2^m},\frac{L}{2^m}\right)-m\,F\left(\frac{K}{2^{m+1}},\frac{L}{2^{m+1}}\right)\right]}{F\left(\frac{K}{2^{m+1}},\frac{L}{2^{m+1}}\right)-\frac{1}{2}F\left(\frac{K}{2^m},\frac{L}{2^m}\right)}.$$

The expression above describes a sufficient condition for decentralization of the skip-level reporting structure with $P=2^m$ processing elements ($m=0,1,...,\log_2(N/2)-1$).

As already emphasized, despite the formal proof of efficiency, it is difficult to find such organizational forms of data processing in real firms. The architecture of information-processing structures in enterprises differs from the one described above, because (1) in real firms computations are usually much more complicated than a simple addition of numerical data, and (2) in the real firms, unlike in the computer systems, people (not electronic processors) form the information-processing structures. The implications of human information-processing on the forms of hierarchical organization of decision-making are considered below.

[8] See Radner (1992 and 1993).

6. Implications of Human Information-Processing on Organizational Forms of Decision Making in the Firm

In the preceding sections information processing in the firm has been represented as in an idealized parallel computer because, in the case when numerical data are simply added up, there is no significant difference between the computation in a firm and in a computer system. However, this is not the case when decentralized information processing involves operations in which some freedom of choice, based on individual judgement, is given to each individual member of the decision-making team.

To clarify the statement above, consider the process of selecting the best project (out of N projects submitted) in a team of P decision makers. Without loss of generality, assume that project n (n=1,2,...,N) is fully characterized by the value of one (aggregated) numerical parameter, Q_n, determined based on the analysis of the entire project. Therefore, one of the projects considered, say n^*, should be selected as the best one if $n^* = \arg\{\min | Q^* - Q_n |\}$, where Q_n is a numerical characteristic of the project n (n=1,2,...,N), and Q^* is an attribute of the project wanted by the entrepreneur (the best for the firm).

As discussed in section 2, the values of the decisions decrease with the delay in information processing. Consequently, in order to reduce the time of data analysis, the process of selecting the best project can be decentralized. However, if each processing element of the decentralized structure represents a single member of the decision-making team, then each individual decision maker, p (p=1,2,...,P), computes (and compares) the absolute values of the differences: $| Q_p^* - Q_{p,n} |$, where $Q_{p,n}$ denotes his subjective evaluation of the project n ($n \in \mathbf{N_p}$, $\mathbf{N_p}$ is a set of projects analyzed by member p), and Q_p^* is an attribute of the project considered as the best by decision-maker p. Divergences between decision makers[9] in attitudes,[10] in perceptual abilities, or in their ability to concentrate, and also some random factors such as, for instance, emotions, frustrations or stresses, imply that subjective evaluations of the same project, say n, by different members of the team, $Q_{p,n}$ (p=1,2,...,P), could not be the same, i.e., $Q_{1,n} \neq Q_{2,n} \neq ... \neq Q_{P,n} \neq Q_n$. Moreover, the possibility of misinterpretating the target of data analysis (i.e., of the goal of the entrepreneur, Q^*) and divergences among the members' individual goals in information processing imply that the

[9] See, for example, O'Reilly III (1990) for a detailed analysis of differences in information use between decision makers.

[10] An attitude consists of feelings, beliefs and predispositions to behave in certain ways, it is understood as 'an unseen force that people presume exists in order to explain certain behavior' (Organ and Bateman, 1986).

understanding of which is the best project could be different for each individual decision maker, i.e., $Q_1^* \neq Q_2^* \neq ... \neq Q_P^* \neq Q^*$. Consequently, if all the projects submitted would be considered by all the members of the decision-making team, then each decision maker could choose a different project (also different from the project that would be selected by the entrepreneur). Therefore, the decentralization of the process of project selection in the firm implies that the result of data analysis could be determined with error, measured by the absolute value of the difference between the numerical characteristics of the project wanted by the entrepreneur (Q^*), and the project selected (Q_{n^*}), $|Q^*-Q_{n^*}|$.

To represent the divergences between the members of the information-processing team, and to describe the possible variability in subjective evaluations of the information analyzed in the framework of the dynamic parallel processing model of associative computation, assume that decision-makers (i.e., the processing elements) do not make errors in the evaluations of the projects analyzed (i.e., $Q_{1,n}=Q_{2,n}=...=Q_{P,n}=Q_n$, n=1,2,...,N), but each decision maker p (p=1,2,...,P) computes results according to his individual understanding of the goal of the analysis, Q_p^* (the possibilities of random mistakes in evaluations of projects can be represented as random shifts in Q_p^*).

Thus, a numerical characteristic of the projects considered as the best by the members of the team ($Q_1^* \neq Q_2^* \neq ... \neq Q_P^* \neq Q^*$) can be described by the random variables distributed around a numerical characteristic of the project wanted by the entrepreneur, Q^*. For the sake of simplicity assume that this distribution is normal, with mean Q^* and variance σ_*^2.

Assuming that all the projects submitted are not identical, but all of them satisfy more or less the expectations of the entrepreneur, we can presume that their characteristics, Q_n (n=1,2,...,N), are distributed around the numerical characterisctic of the project wanted by the enterpreneur (Q^*). For the sake of simplicity, assume that this distribution is normal, with mean Q^* and variance σ^2.

The random factors in data analysis imply that the selection process should be organized in the decentralized structure (in order to minimize the delay in information processing) which minimizes the expected value of the error in data analysis, $E=\mathbf{E}(|Q^*-Q_{n^*}|)$, where **E** denotes the operator of expectation.

If Q_n (n=1,2,...N) and Q_p^* (p=1,2,...,P) are normally distributed random variables, with mean Q^* and variances σ^2 and σ_*^2, respectively, then the random variable characterizing the project selected (Q_{n^*}) in the arbitrary decision-making structure can be represented as follows:
where P_n denotes the probability that the project n will be selected (n=1,2,...,N).

$$Q_{n^*} = \sum_{n=1}^{N} Q_n P_n \quad ,$$

This implies that, in an arbitrary information-processing structure, the random variable Q_{n^*} is distributed normally, with mean Q^* and variance $\sigma_{n^*}^2$. Moreover, the random variable Q^*-Q_{n^*} is normally distributed, with zero mean and variance $\sigma_{n^*}^2$. Consequently, the expected value of the error in data analysis can be determined as

$$\mathrm{E}\left(\left|Q^* - Q_{n^*}\right|\right) = \int_0^{\infty} x f_{|Q^*-Q_{n^*}|}(x)\, dx = \int_0^{\infty} \frac{2x}{\sigma_{n^*}\sqrt{2\pi}} e^{-\frac{x^2}{\sigma_{n^*}^2}} dx =$$

$$= -\frac{\sigma_{n^*}}{\sqrt{2\pi}} e^{-\frac{x^2}{\sigma_{n^*}^2}} \Big|_0^{\infty} = \frac{\sigma_{n^*}}{\sqrt{2\pi}} \quad .$$

This means that the information-processing structure which minimizes the expected value of the error in data analysis also minimizes the variance of the random variable Q_{n^*}. In an arbitrary information-processing structure, this variance can be computed as

$$\sigma_{n^*}^2 = \sigma^2 \sum_{n=1}^{N} P_n^2 \quad ,$$

where P_n denotes the probability that the project n (n=1,2,...,N) will be selected.

The values of the probabilities P_n^* (n=1,2,...,N) minimizing variance $\sigma_{n^*}^2$, and, consequently, minimizing the expected value of the error in data analysis, can be determined by finding the solution to the following optimization problem:

$$\underset{P_1, P_2, \ldots, P_N}{Min} \quad \sigma^2 \sum_{n=1}^{N} P_n^2$$

such that

$$\sum_{n=1}^{N} P_n = 1 \quad .$$

The probabilities P_n^* (n=1,2,...,N) equal $P_1^* = P_2^* = \ldots = P_N^* = 1/N$ and, consequently, the minimum expected value of the error in data analysis equals $E_{min}= \sigma/(2\pi\ N)^{1/2}$.

If the characteristics of the projects considered as the best by each individual decision-maker p, Q_p^* (p=1,2,...,P), are normally distributed random variables with mean Q^*, then the values of the probabilities P_n (n=1,2,...,N) are determined only by the architecture and the information workload of the computational structure. Thus, the result obtained above has the following implications concerning the form of the structures minimizing the expected error in data analysis:

1. The structures minimizing the expected value of the error in data analysis have to be regular[11] with equally loaded processing elements (this ensures that all the projects analyzed are selected with the same probability).
2. The expected value of the error in data analysis does not depend upon the number of processing elements in the structure, i.e., it is the same for one processing element (centralized structure) as for any decentralized-regular structure with an equalized workload of computational centers.
3. The expected value of the error in data analysis (E) increases with the variance of the projects submitted (σ^2).
4. In an arbitrary information-processing structure, the variance of characteristic of the project selected is inversely proportional to the number of the projects analyzed, and, therefore, the expected value of the error in analysis decreases if the number of projects considered (N) increases.
5. For any number of projects analyzed (N) there exists at least one structure of data processing (possibly more) which minimizes the expected value of the error in data analysis (this one is a centralized structure).

The analysis above shows that the expected value of the error in data processing is the same in centralized as in any equally loaded decentralized-regular structure. This means that the specific features of human information-processing such as disagreement about the goals of data analysis or the possibility of random errors, do not imply a hierarchical organization of the management. However, if the value of the computational service depends not only upon the delay in information processing, but also upon the error in data analysis, then the forms of information-processing structures could be more regular than those derived for idealized parallel computers.

[11] The hierarchy is called regular if (1) all the immediate subordinates of any processor are at the next lower level, and (2) all processors of the same level have the same number of immediate subordinates.

7. Efficient Organizational Forms for Associative Computation in the Firm

As discussed in section 3, the value of the computational service depends on how good the resulting decisions are compared to how good they would be without the service. Consequently, the expected value of the computational service in arbitrary decision-making structure can be represented as

$$V(D_N(K,L),E) = V_{\max} - V^o_{\max} - (\rho D_N(K,L) + \gamma E) = V_{\max} - V^o_{\max} - (\rho D_N(K,L) + \gamma \frac{\sigma_{n^*}}{\sqrt{2\pi}}),$$

where V_{max} is the maximum value of the decision with the computational service, $V°_{max}$ is the maximum value of the decision if no information is processed, D_N(K,L) is the delay in information processing, E is the expected value of the error in computation, σ_{n^*} is the standard deviation of the random variable describing a numerical characteristic of the project selected, ρ and γ denote the unit costs of the delay in information processing and the error in data analysis, respectively.

Thus, the information-processing structure is said to be efficient, for a given:

(1) information-processing function (F(k,l)),
(2) information workload (N),
(3) variance of the characteristics of the projects submitted (σ^2),
(4) form of the relationship between the delay in information processing, the error in data analysis and the value of the computational service $V(D_N,E)$,

if it is not possible to get the same the expected value of the computational service using less of one input to information processing (i.e., labor or capital) and no more of the other.

Taking into account that in an arbitrary information-processing structure a standard deviation (σ_{n^*}) of a random variable describing a numerical characteristic of the project selected can be represented as $\sigma_{n^*} = A_N\sigma$ (where A_N is the coefficient characterizing the structure considered, and σ is the standard deviation of the numerical characteristics of the projects analyzed), the expected value of the computational service is determined solely by the delay in information processing, if (1) σ^2=0, i.e., if all the projects submitted are identical, or (2) γ=0, i.e., when the error in data analysis is not a matter of concern for the firm. Moreover, if all members of the decision-making team have the same objectives and do not make errors in computation, i.e., if $Q_1^*=Q_2^*=...=Q_P^*$, then the expected value of the error in data analysis (E) does not depend on the architecture of the information-processing structure. In these particular cases the computational process in the firm can be analyzed similarly to an idealized parallel computer, and, consequently, the skip-level reporting structures can be considered as efficient for the computation of associative operations in the firm; otherwise, the architecture of the efficient information-processing structures should be determined taken into account not only the delay in information processing (D_N), but also the expected value of the error in data analysis (E).

To see how the delay in information processing and the expected value of the error in data analysis depend on the architecture of information-processing structures, consider the example of the reduction of the regular structure with P=7 processing elements presented in fig.7.1 (the characteristics of the projects analyzed, Q_n, n=1,2,...,8, and the projects considered as the best by individual decision makers, Q_p^*, p=1,2,...,P, are described by normally distributed random variables, with mean Q^* and variances σ^2 and σ_*^2, respectively). The reduction process under study makes each subsequent structure less regular (see fig.7.1), and, consequently, increases the expected value of the error in computation (E), but at the same time decreases the delay in information processing (D_N).

This confirms the theoretical result that the expected value of the error in data analysis is minimized in regular structures, while the delay in information processing is minimized in irregular, so-called skip-level reporting structures (see Gibbons and Rytter, 1988, or Radner, 1993). Moreover, it implies that the expected value of the error in data analysis and the delay in information processing can be minimized in the single structure only if the decentralization of data analysis doesn't decrease[12] the delay in information processing (in this case a centralized structure is efficient). Otherwise, the forms of the efficient information-processing structures should be determined individually for each particular decision problem, taking into account: the information workload of the structure (N), the information-processing function (F(k,l)), the variance of the characteristics of the projects submitted (σ^2), and the form of the relationship between the delay in information processing, the error in data analysis and the value of the computational service ($V(D_N,E)$).

[12] It corresponds to the case when the necessary and sufficient conditions for the decentralization of a centralized structure are not satisfied.

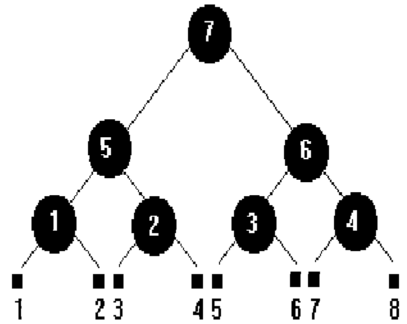
The considerations above imply that:

-if the firm is concerned about the delay in information- processing ($\rho>0$) and doesn't care about the error in computation ($\gamma=0$), then the skip-level reporting structures of information processing are efficient for computation of associative operations in the firm;

- if the firm is concerned much more about the error in data analysis than about the delay in information processing (i.e., if $\gamma>0$ and $\rho\geq0$, but $\gamma>>\rho$), then the efficient structures of information processing in the firm are regular with equally loaded processing elements;

- if γ and ρ are of the same order, then (1) the relationship between the error in data analysis (E) and the variability of the numerical characteristics of the projects submitted ($E\sim\sigma_{n^*}\sim\sigma$, where σ is the standard deviation of the numerical characteristics of the projects submitted) implies that, for small values of σ, more irregular structures are expected to be efficient than otherwise (in the extreme case, if $\sigma=0$ then the skip-level reporting structures are efficient);
(2) the relationships between the delay in data processing (D_N), the expected value of the error in data analysis and the information workload ($D_N\sim N$ and $E\sim1/N$, where N denotes a number of data items analyzed), imply that, for large numbers of data items analyzed, more irregular structures are expected to be efficient than otherwise (if all other parameters do not change).

a.

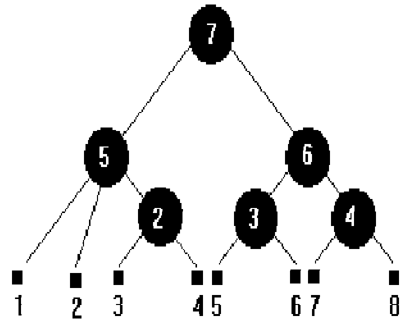


Regular structure (P=7)

D=6 d(K/7,L/7),

E=$\sigma/(4\pi^{1/2})$

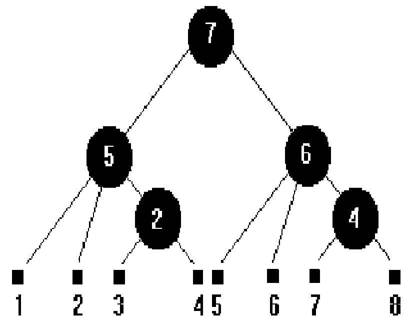
b.



D=6 d(K/6,L/6) (↓),

E=$\sigma/(288\pi/19)^{1/2}$ (↑)

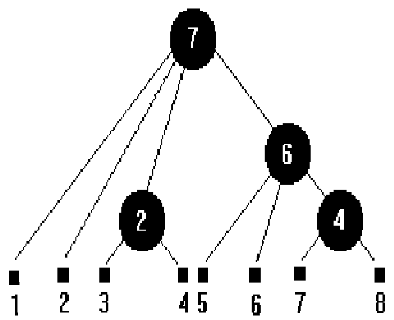
c.



D=5 d(K/5,L/5) (↓),

E=$\sigma/(288\pi/20)^{1/2}$ (↑)

d.



Skip-level reporting structure

D=4 d(k/4,L/4) (↓),

E=$\sigma/(288\pi/25)^{1/2}$ (↑)

Figure 7.1.: The reduction of the regular structure with P=7 processing elements (the projects analyzed are represented as rectangles, and the processing elements as circles)[13]

---

[13] The expected values of the errors in analysis are determined in Appendix 1.

Moreover, changes in the resources allocated to various information-processing structures affect differently the loss due to non-instantaneous and imprecise information processing ($\rho D+\gamma E$), and, consequently, the value of the computational service. This implies that the architecture of the efficient structures depends also upon the amounts of resources used in data processing.

As an example, the relationship between the loss due to non-instantaneous and imprecise information processing ($\rho D+\gamma E$) and the resources ($K=L\in[0.01,0.2]$) allocated to various computational structures (designed for the selection of the best project out of N=8 projects submitted; $\rho=0.01$, $\gamma=1$, $\sigma^2=1$, and the information-processing function $F(k,l)=k^{0.1}l^{0.1}$) is presented in fig.7.2. The intersection of curves F and H implies that if equal amounts of labor and capital are used in information processing (i.e., K=L) and $K=L\in[0.01,0.2]$ then the efficient information-processing structure (in the example considered) is

- skip-level reporting with P=4 processing elements (fig. 7.1.d) if $0.01<K=L<0.042$,
- irregular with P=5 processing elements (fig. 7.1.c) if $0.042<K=L<0.2$.

Therefore, contrary to the results presented in the literature (see, e.g., Radner, 1993; or Radner and Van Zandt, 1992), one can conclude that, in the firm (unlike in the computer systems) there is no single form of an efficient information-processing structure, but a number of various architectures of efficient structures can be observed.

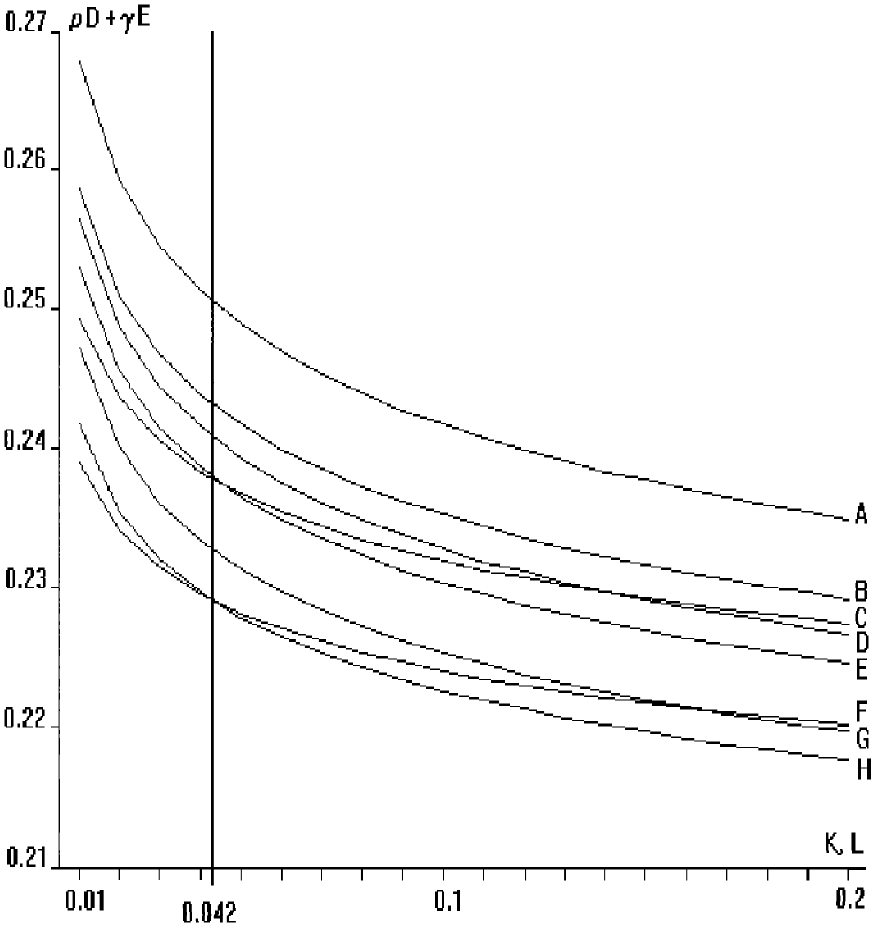


Figure 7.2.: The relationship between the loss due to non-instantaneous and imprecise information processing ($\rho D+\gamma E$) and the resources used in computation (K,L), N=8, $\rho=0.01$, $\gamma=1$, $\sigma^2=1$, $F(k,l)=k^{0.1}l^{0.1}$:

A: centralized structure, P=1, D=8d(K,L), $E=\sigma/(4\pi^{1/2})$,
B: irregular structure (fig.7.1.b),
C: skip-level reporting structure, P=2, D=5d(K/2,L/2), $E=\sigma/(200\pi/17)^{1/2}$
D: regular structure (fig.7.1.a),
E: regular structure, P=4, D=6d(K/4,L/4), $E=\sigma/(324\pi/21)^{1/2}$
F: skip-level reporting structure (fig.7.1.d),
G: regular structure, P=3, D=6d(K/3,L/3), $E=\sigma/(4\pi^{1/2})$
H: irregular structure (fig.7.1.c)

8. Conclusion

The studies of organizational aspects of decision-making and information processing in business firms have appeared recurrently in the economic literature, but the arguments justifying the existence of the hierarchies have been more psychological and sociological than economic. Contrary to these opinions, the analysis presented in this paper shows that the explanation of the hierarchies is purely economic. It is proven that decreasing returns to scale in information-processing is a necessary condition for a hierarchical (decentralized) organization of the management in the firm; however, decentralization is desirable only if an additional (sufficient) condition is satisfied, i.e., if the information workload of the decision-making sector is sufficiently large. Moreover, the paper shows that the decentralization of information processing leads to irregular hierarchical forms (so-called ‘skip-level reporting’ structures).

Furthermore, based on the analysis of the model of project selection in a firm, it is shown that specific features of human information-processing such as disagreement about the goals of data analysis or the possibility of random errors in computation do not imply a hierarchical organization of management, but affect organizational forms of hierarchical structures (make them more regular).

Finally, the architecture of the efficient information-processing structures is analyzed. It is shown that skip-level reporting structures (considered in economic literature as efficient for associative computation in enterprises) are efficient only in some particular cases of data processing in the firm. This explains why it is extremely difficult to find such structures in real enterprises. Moreover, contrary to the results recently presented in the economic literature, the analysis shows that, in the firm, unlike in a computer system, there is no single form of efficient information-processing structures, but a number of various architectures of efficient structures can be observed.

APPENDIX

DERIVATION OF EXPECTED VALUES OF THE ERROR IN DATA ANALYSIS IN SELECTED INFORMATION-PROCESSING STRUCTURES

Suppose, that the characteristics of the projects analyzed, (Q_n, n=1,2,...,8) and the projects considered as the best by individual decision makers (Q_p^*,p=1,2,...,P) are described by normally distributed random variables, with mean Q^* and variances σ_2 and σ_{*2}, respectively; and consider the information-processing structures presented in fig.7.1.

The structure presented in fig. 7.1.a is regular therefore each project considered is selected with the same probability (1/8). Consequently, the random variable describing the numerical characteristic of the project selected in this structure (Q_{7^*}) can be represented as

$$Q_{7^*} = 1/8Q_1 + 1/8Q_2 + 1/8Q_3 + 1/8Q_4 + 1/8Q_5 + 1/8Q_6 + 1/8Q_7 + 1/8Q_8.$$

The variance of this random variable equals $\sigma_{(7)^{*2}}$ =1/8 σ^2, and the expected value of the error in data analysis is

$$E = \sigma_{(7)^*} /(2\pi)^{1/2} = \sigma/(4\pi^{1/2}).$$

To analyze the structure presented in fig. 7.1.b, consider two of its substructures with the processing elements 5 and 6 at the tops. The substructure with the processing element 6 at the top is regular, consequently, projects 5,6,7 and 8 are selected in the processing element 6 with the same probability (1/4). The substructure with the processing element 5 at the top is irregular (skip-level reporting). Therefore, project 1 is selected in the processing element 5

(1) if it is selected from the projects 1,2 and 3 in the processing element 5 (random variables Q_1,Q_2 and Q_3 are idendependent and identically distributed, thus, the probalility of this event equals 1/3), and project 3 is selected in the processing element 2 (the probability of this event equals 1/2); or

(2) if it is selected from the projects 1,2, and 4 in the processing element 5 (random variables Q_1,Q_2 and Q_4 are idendependent and identically distributed, thus, the probalility of this event equals 1/3), and project 4 is selected in the processing element 2 (the probability of this event equals 1/2);

Finally, the project 1 is selected in the processing element 5 with the probability 1/2 1/3+1/2 1/3=1/3. Analogously, project 2 is selected in the processing element 5 with the probability 1/3.

Project 3 is selected in the processing element 5 if it is selected in the processing element 2 (probability of this event equals 1/2), and if it is selected in the processing element 5 (the probalility of this event equals 1/3). Consequently, the probability that project 3 is selected in the processing element 5 equals 1/2 1/3 = 1/6. Analogously, the probability that project 4 is selected in the processing element 5 equals 1/6.

Project 1 is selected as the best (in the processing element 7) if it is selected in the processing element 5 (probability of this event equals 1/3), and if it is selected in the processing element 7. Project 1 is selected in the processing element 7 if it is better than projects 5, or 6, or 7, or 8 on the condition that the corresponding project (i.e., 5,6,7, or 8) is selected in the processing element 6. The probabilities that projects 5,6,7, or 8 are selected in the processing element 6 equal 1/4. Thus, the probability that project 1 is selected in the processing element 7 equals 4(1/3 1/2 1/4)= 1/6.

Analogously, the probabilities that projects 2,3,...,8 are selected in the processing element 7 equal 1/6, 1/12, 1/12, 1/8, 1/8, 1/8, 1/8, respectively. Consequently, the random variable characterizing the project selected in the structure under study (Q_{7^*}) is

$$Q_{7^*} = 1/6Q_1 + 1/6Q_2 + 1/12Q_3 + 1/12Q_4 + 1/8Q_5 + 1/8Q_6 + 1/8Q_7 + 1/8Q_8.$$

The variance of this random variable equals $\sigma_{(7)^*}{}^2 = 19/144\ \sigma^2$, and the expected value of the error in data analysis is

$$E = \sigma_{(7)^*}/(2\pi)^{1/2} = \sigma/(288\pi/19)^{1/2}.$$

The structure presented in fig. 7.1.c contains two identical irregular (skip-level reporting) substructures with the processing elements 5 and 6 at the tops. The probabilities that the projects 1,2,3 and 4 are selected in the processing element 5 equal 1/3, 1/3, 1/6 and 1/6, respectively (see the consideration above). Analogously, the probabilities that projects 5,6,7 and 8 are selected in the processing element 6 equal 1/3, 1/3, 1/6 and 1/6, respectively. Consequently, probabilities that projects 1,2,3,...,8 are selected in the processing element 7 equal 1/3 1/2=1/6, 1/3 1/2=1/6, 1/6 1/2=1/12, 1/6 1/2=1/12, 1/3 1/2=1/6, 1/3 1/2=1/6, 1/6 1/2=1/12, 1/6 1/2=1/12, respectively (see the analysis above). The random variable characterizing the project selected in the structure under study (Q_{7^*}) is

$$Q_{7^*} = 1/6Q_1 + 1/6Q_2 + 1/12Q_3 + 1/12Q_4 + 1/6Q_5 + 1/6Q_6 + 1/12Q_7 + 1/12Q_8.$$

The variance of this random variable equals $\sigma_{(7)^*}{}^2 = 20/144\ \sigma^2$, and the expected value of the error in data analysis is

$$E = \sigma_{(7)^*}/(2\pi)^{1/2} = \sigma/(288\pi/20)^{1/2}.$$

In the structure presented in fig. 7.1.d, project 1 is selected as the best in the processing element 7 if it is better than

(1) project 2,

and it is better than

(2) project 3, if project 3 is selected in the processing element 2 (the probability that this project is selected in the processing element 2 equals 1/2), or project 4, if project 4 is selected in the processing element 2 (probability that this project is selected in the processing element 2 equals 1/2);

and it is better than

(3) projects 5, or 6, or 7, or 8, on the condition that the corresponding project (i.e., 5, 6, 7, or 8) is selected in the processing element 6 (probabilities that these projects are selected in the processing element 6 equal 1/3, 1/3, 1/6 and 1/6, respectively).

This implies that the probability of project 1 being selected in the processing element 7 equals 1/4. Analogously, the probabilities that projects 2,3,...,8 are selected as the best (in the processing element 7) equal 1/4, 1/8, 1/8, 1/12, 1/12, 1/24, 1/24, respectively. Consequently, the random variable characterizing the project selected in the structure under study (Q_{7^*}) is

$$Q_{7^*}=1/4Q_1 +1/4Q_2 +1/8Q_3 +1/8Q_4 +1/12Q_5 +1/12Q_6 +1/24Q_7 +1/24Q_8.$$

The variance of this random variable equals $\sigma_{(7)^*}^2 = 25/144\, \sigma^2$, and the expected value of the error in data analysis equals

$$E = \sigma_{(7)^*}/(2\pi)^{1/2} = \sigma/(288\pi/25)^{1/2}.$$

REFERENCES

Alchian, A. A., and H. Demsetz. 1972. "Production, Information Costs, and Economic Organization." *American Economic Review*, 62:777-795.

Aoki, M. 1986. "Horizontal vs. Vertical Information Structure of the Firm." *American Economic Review*, 76:971-983.

Bolton, P., and M. Dewatripont. 1994. "The Firm as a Communication Network." *The Quarterly Journal of Economics*, Vol. CIX, no. 4:809-840.

Calvo, G., and S. Wellisz. 1978. "Supervision, Loss of Control and the Optimal Size of the Firm." *Journal of Political Economy*, no. 86:943-952.

Calvo, G., and S. Wellisz. 1979. "Hierarchy, Ability and Income Distribution." *Journal of Political Economy*, 87, no. 5:991-1010.

Cyert, R. M. 1988. *The Economic Theory of Organization and the Firm*. Hertfordshire: Harvester Wheatsheaf.

Daft, R. L., and R. H. Lengel. 1990. "Information Richness: A new Approach to Managerial Behavior and Organization Design." In *Information and Cognition in Organizations*. Edited by L. L. Cummings and B. M. Staw. London: JAI Press Inc..

Gibbons, A., and W. Rytter. 1988. *Efficient Parallel Algorithms*. Cambridge: Cambridge University Press.

Keren, M., and D. Levhari 1979. "The Optimum Span of Control in a Pure Hierarchy." *Management Science*, no. 11:1162-1172.

Keren, M., and D. Levhari. 1983. "The Internal Organization of the Firm and the Shape of Average Cost." *The Bell Journal of Economics*, no. 14:474-486.

Lydall, H. F. 1968. *The Structure of Earnings*. Oxford: The Clarendon Press.

Marschak, T., and R. Radner. 1972. *Economics theory of teams*. New Haven, CT: Yale University Press.

Milgrom, P., and J. Roberts. 1990. "The Economics of Modern Manufacturing: Technology, Strategy, and Organization." *American Economic Review*, 80 (3):511-528.

Organ, D. W., and T. Bateman. 1986. *Organizational Behavior. An Applied Psychological Approach*, Homewood, Illinois: Business Publications, Inc..

O'Reilly III, Ch. A. 1990. "The Use of Information in Organizational Decision Making." In *Information and Cognition in Organizations*. Edited by L. L. Cummings and B. M. Staw. London: JAI Press Inc..

Radner, R. 1992. "Hierarchy: The Economics of Managing." *Journal of Economic Literature*, vol. 30:1382-1415.

Radner, R. 1993. "The Organization of Decentralized Information Processing." *Econometrica* 61:1109-1146.

Radner, R., and T. Van Zandt. 1992. "Information Processing in Firms and Return to Scale" *Annales d'Economie et de Statistique*, 25/26:265-298.

Radner, R., and T. Van Zandt. 1993. *Information Processing and Returns to Scale in Statistical Decision Problems*. Murray Hill, New Jersey: AT&T Bell Laboratories.

Sah, R. K., and J. E. Stiglitz. 1986. "The Architecture of Economic Systems: Hierarchies and Polyhierarchies." *American Economic Review*, 76 (4):716-727.

Simon, H. A. 1957. "The Compensation of Executives." *Sociometry*, 20, no. 1:32-35.

Williamson, O. E. 1967. "Hierarchical Control and Optimum Firm Size." *Journal of Political Economy*, 75, no. 2:123-138.

Williamson, O. E. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications.* New York: Free Press.

Williamson, O. E. 1981. "The Modern Corporation: Origins, Evolution, Attributes." *Journal of Economic Literature*, vol. 19:1537-68.

Williamson, O. E. 1986. *Economic Organization. Firms, Markets and Policy Control*. New York: New York University Press.

Van Zandt, T. 1990. *Efficient Parallel Addition*. Murray Hill, New Jersey: AT&T Bell Laboratories.