ELSEVIER

Journal of Economic Behavior and Organization
Vol. 25 (1994) 37–51

JOURNAL OF
Economic Behavior
& Organization

Information processing and organization design

Peter W. Kennedy

University of Victoria, P.O. Box 3050, Victoria, B.C., Canada V8W 3P5

Received March 1993, final version received December 1993

Abstract

This paper examines the optimal structure of an organization in which analysts are hired to process information on behalf of a principal decision-maker whose attention is limited. I focus on the case where information processing exhibits *declining complexity*. This means that information processing becomes less complex as it progresses. The optimal organization design is determined endogenously as an optimal response to the limitations of the principal decision-maker in her attention to communication and supervision, and to the limited processing ability of the analysts. I examine serial and parallel processing structures. I show that the optimal serial structure is ordered by ability. This ordering reflects specialization according to comparative advantage in processing. The choice between a serial structure and a parallel structure involves a tradeoff between the benefits of specialization in the serial structure and lower communication costs in the parallel structure.

Keywords: Organizational behavior; Information processing

JEL classification: D23; D83

1. Introduction

It is an accepted fact that economic agents are limited in the amount of attention they can devote to any particular activity. This has been recognized at least since the time of Adam Smith [1]. A first formal treatment of the issue is provided by Gary Becker (1965) who examines the implications of time constraints for consumer behavior. More recently an extensive literature has developed on the implications of limited attention for the design of organizations. One

[1] Sharon Gifford (1992a) provides a citation in her survey of the literature on limited attention.

SSDI 0167-2681(94)00049-K

area of this literature has attempted to explain organizational structure as an optimal response to the limited information processing ability of decision-makers [2]. It is to this area of the literature that my paper contributes. The purpose of my paper is to present some new ideas on optimal organization design in response to information processing limitations in the context of a more complete framework than has typically been applied in the existing literature.

My paper attempts to shed some light on two questions. First, to what extent can differences in commonly observed organizational structures be attributed to differences in the nature of the information processing problems faced by the decision-makers in those organizations? Consider some example structures. Law firms typically are 'flat' organizations in which each member of the firm works more or less independently on a case and retains that case for its duration. We usually do not observe law firms with extensive hierarchical structures in which a single case is handled by a series of lawyers. Similarly, accounting and auditing firms are ordinarily flat organizations in which each client firm is handled exclusively by a single accountant. Apart from the screening role played by senior partners, these firms are usually organized as teams of autonomous units. In contrast, many firms and government departments are organized as hierarchical structures: information processing problems are dealt with in a serial manner. For example, a broad government policy objective originates with an elected representative or senior administrator and is then passed down sequentially through the layers of the bureaucracy where it is researched and refined into an implementable policy. Similarly, in large corporations, entrepreneurial ideas and strategies are transformed into implementable business policy via serial processing through the corporate hierarchy. My paper asks whether this difference in the way organizations are structured can be at least partially explained by differences in the nature of the information processing involved.

A second issue on which my paper may be illuminating relates to organizational restructuring in response to changing technology. It has been widely argued by management analysts that advances in information processing technology is leading to the demise of deep hierarchical corporate structures (such as the old IBM) and the advent of flatter more autonomous structures [3]. The widespread elimination of middle management layers during the 1980s appears to be consistent with this view. However, the theory behind this argument is rather thin. While it seems reasonable to suppose that technological improvements will lead firms to substitute out of human information processing and into information processing technology, there is no *prima facie* reason to expect that optimal organizational structures will also become less hierarchical. 'Leaner' does not necessarily mean

[2] There is also an extensive literature that examines organization design in the context of a non-optimizing framework.

[3] For example, see Schein (1989) and Business Week (1989).

'flatter'. My paper offers some theoretical insights into when leaner may also mean flatter.

The framework I use to examine these questions can be summarized as follows. A decision-maker must divide her limited attention across many different decision problems that require information processing. She can hire *analysts* (whose attention is also limited) to assist with information processing and thereby bring more attention to bear on those decision problems. The decision-maker must communicate with her analysts and supervise their activities. These tasks are themselves time-consuming and the decision-maker can devote only limited attention to them. This limits the size of the organization, which in turn limits the number of decision problems to which it will optimally devote attention. Two types of organizational structure are considered: parallel structures and serial structures. Analysts in a parallel structure work autonomously on processing problems assigned exclusively to them. Only one analyst works on any given problem. In a serial structure information is processed sequentially through a series of analysts. Every analyst in the series works on part of every problem.

The main insights gained from my analysis are the following. The serial structure permits analyst specialization according to comparative advantage in processing at different stages of the processing. This is not possible in the parallel structure. However, the serial structure necessarily involves more inter-analyst communication than the parallel structure. The choice between the two structures involves a tradeoff between the benefits of specialization and the costs of more communication.

I model the potential for analyst specialization in a very simple way. I allow analysts to differ in processing ability and I focus on situations where information processing exhibits *declining complexity*. By this I mean that information processing comprises a series of steps that build on each other in such a way that each step is less time-consuming than the previous one for an analyst of given ability. I provide a more precise definition in section 2 but an analogy should help to clarify the idea at this point. Think of decision-making as solving a jigsaw puzzle. The jumbled pieces represent information that when combined or processed in the appropriate way provides a clear picture of the best decision. The completeness of the picture determines the quality of the decision. As anyone who has ever solved a jigsaw puzzle will know, the task becomes less complicated as it progresses [4].

[4] It should be noted that my jigsaw puzzle interpretation of information processing is quite different from the view of information processing as prior belief revision on the basis of sampling from a probability distribution. Geanakoplos and Milgrom (1991) treat information processing in this way. In that view information quality is interpreted as the precision of the posterior distribution. For conjugate families of distributions (those for which there is a closed analytical relationship between the prior distribution, the sampling distribution and the resulting posterior distribution) the precision of the posterior distribution is linear in the sample size. Normal sampling problems, for example, have this property. This implies, in my terminology, that there is constant complexity in processing. In many instances this may be an appropriate way to view information processing but in other instances the jigsaw analogy is likely to be more appropriate.

There are two reasons for this. First, as more pieces are put in place, there are less remaining pieces to sort through; and second, as the picture begins to evolve it becomes easier to see where the remaining pieces fit into the puzzle [5]. Information processing exhibits declining complexity in many instances for precisely these reasons. I examine the structure of organizations designed to solve decision-problems that exhibit this declining complexity property. However, it should be noted that the main insights of the paper are not tied to this focus on declining complexity in processing. The marriage of heterogeneous ability and declining complexity is simply a convenient way of modeling the potential for analyst specialization in processing.

My paper builds on and borrows from the existing literature on the optimal allocation of attention and optimal organization design. The problem of allocating limited attention across many different decision problems was first examined by Radner (1975) and Radner and Rothschild (1975). The decision-maker in their model undertakes all information processing herself. The possibility of hiring analysts to assist with information processing has been examined by Keren and Levhari (1983), Geanakoplos and Milgrom (1991), and others. In Keren and Levhari (1983) the time taken for the directives of a firm's chief executive to be relayed to production units increases with the span of his control. This costly time lag can be reduced by hiring parallel managers who concurrently relay directives to a smaller span of control. In Geanakoplos and Milgrom (1991) managers must divide their attention between processing cost information on the production units under their control. Appointing more managers permits a finer partition of the information space and thereby permits a better allocation of resources across the production units. In most attention-allocation models of this type the number of targets for attention (usually production units) is fixed exogenously. An exception is Gifford (1992b) who examines a model in which an entrepreneur must allocate her limited attention between improving the operation of ongoing projects, and introducing new projects that demand ongoing attention if introduced. My model also yields an endogenous determination of targets for attention but this endogeneity stems from communication and supervision requirements. A number of papers have examined the problem of limited supervisory attention in firms with many production workers. They include Williamson (1967), Beckmann (1977), Calvo and Wellisz (1978 and 1979), Rosen (1982) and Oi (1983). These papers focus on the optimal design of supervisory structures when there are an exogenously fixed number of workers to be supervised.

The rest of the paper is arranged as follows. Section 2 presents the model. Section 3 examines the optimal parallel structure. Section 4 examines the optimal serial structure and compares it with the parallel structure. Section 5 discusses the

[5] The analogy can clearly be extended to capture prior information acquisition (jigsaw pieces must first be purchased) and information overload (the jumble of pieces may include extraneous pieces that are not relevant to the puzzle).

results in relation to the empirical issues raised at the beginning of the paper. Section 6 concludes. An appendix contains one proof.

2. The model

Time is divided into periods of length one. A decision-maker (henceforth called the principal) receives over time a continuous stream of decision problems called projects. The expected payoff from a project is equal to the quality of the information on which the decision is based. All projects have an initial quality of information equal to zero. The quality of information is improved by information processing but information processing is time-consuming. This means that a better decision on one project necessarily implies a worse decision on some other project. The principal can hire analysts to undertake information processing and so bring higher quality information to bear on her decisions. An analyst of processing ability α can be hired from a competitive market at a wage $w(\alpha)$ per period. This wage is increasing and convex in ability: $w'(\alpha) > 0$ and $w''(\alpha) \geqslant 0$.

An analyst of ability α who receives information of quality q and devotes time t to information processing produces information of quality $q + f(t,\alpha,q)$. Thus, $f(t,\alpha,q)$ is the increment to information quality produced by the analyst. This increment is increasing in t and α: $f_1 > 0$ and $f_2 > 0$. Returns to processing time eventually diminish although there may be some initial finite range over which returns are increasing. That is, $f_{11} < 0$ beyond some finite value of t. I assume that $f_{22} \leqslant 0$. That is, there are no increasing returns to ability [6]. Declining complexity in processing is captured by $f_3 > 0$. That is, the increment to information quality achieved in a given time is increasing in the amount of processing that has already been accomplished (as reflected in q). I assume that $f_{32} \leqslant 0$. This means that the effect of declining complexity in processing is at least as large for lower ability analysts as for higher ability analysts. This seems reasonable since lower ability analysts are likely to be more limited by processing complexity than higher ability analysts. In any case, the assumption is needed only to rule out the possible dominance of second-order effects in the choice of ability.

Analysts can be organized in either a serial or parallel structure. In a parallel structure a project is assigned exclusively to one analyst who is responsible for the entire processing on that project. After a specified processing time t he reports information of quality $q = f(t,\alpha,0)$ to the principal. There is a communication time cost of $\delta/2$ incurred by both the principal and the analyst when the project is assigned and again when the report is made. A serial structure on the other hand

[6] This assumption is more strict than is necessary. I require only that f_{22} be not too strongly positive relative to $w''(\alpha)$. This ensures that it is never optimal to hire an infinitely able analyst.

resembles an assembly line. Each analyst in the series spends some time on each project before transferring it to the next analyst for further processing. Each transfer requires a communication time cost of $\delta/2$ from both analysts [7]. The i^{th} analyst in the series adds quality $x_i = f(t_i, \alpha_i, q_{i-1})$ to the information, where $q_{i-1} \equiv \Sigma_{j=1}^{i-1} x_j$ is the cumulative quality added by analysts preceding analyst i^{th} in the series. (Note that the $(i-1)^{th}$ analyst precedes the i^{th} analyst). The last analyst in a series of n analysts reports information of quality $q = \Sigma_{i=1}^{n} x_i$ to the principal. A communication time cost of $\delta/2$ is incurred when the project is assigned to the first analyst in the series and again when a report is made by the last analyst.

Splitting the processing time on a project between two or more analysts of equal ability is assumed to have no effect on the quality of information produced other than through the time lost to communication. That is, information processing is temporally additive across analysts of equal ability. Formally, $\forall \alpha, q$ and $\forall t, \tau > 0$, $f(\tau + t, \alpha, q) = f(\tau, \alpha, q) + f(t, \alpha, q + f(\tau, \alpha, q))$. This means, for example, that there are no gains to work-sharing associated with the stimulation provided by task variety; analysts do not get bored. In reality there may be some gains of this nature associated with a serial structure but I wish to abstract from them in order to focus on a potentially more important effect related to specialization in processing.

An analyst's productivity depends on the amount of supervision he receives from the principal. If the principal spends time s supervising n analysts then each analyst will spend a fraction p(s,n) of his work time engaged in productive processing. The remaining time is wasted. This productivity factor is increasing in s and decreasing in n ($p_1 > 0$ and $p_2 < 0$) and $p(0,n) = 0$. I assume that the returns to supervision are high enough to make it worthwhile for the principal to hire at least some analysts. I further assume initially that the principal specializes completely in supervision and communication, and undertakes no processing herself. I later examine the implications of relaxing this last assumption [8].

It should be noted that I confine consideration to optimal organizational structures in *steady state*. I ignore the transition to that steady state and in particular, the question of optimal sequential hiring at the beginning of the organization's life. Moreover, I have not examined the issue of whether the optimal start-up process will have persistent implications for the design of the

[7] This assumes that the communication time cost is independent of the quality of information transferred. In reality the communication time cost might conceivably increase or decrease along the series. Incorporating either of these possibilities complicates the analysis but leaves the main results unchanged.

[8] It should be noted that this representation of supervision abstracts from the question of how supervision time should be optimally allocated across analysts. I have implicitly constrained the principal to provide equal supervision to all analysts.

organization once a steady state is reached. This is a potentially important issue but I have left it for future work [9].

3. Parallel processing

There are three parts to the organization design problem. The first is to choose the number of analysts and their ability types. The second is to decide how many projects to accept for processing per period. The third is to decide how much processing time an analyst of given ability should allocate to an accepted project. The principal's objective in the design problem is to maximize per period profit [10]. It is best to address the three aspects of the design problem separately. First, note that any number of analysts of a particular ability can be hired at the same wage and all accepted projects have the same payoff function. So if a particular ability type is optimal for one project then it is optimal for all projects and all positions in the structure. Consideration can therefore be restricted to structures with analysts of identical ability. Next consider the time allocation problem. All accepted projects assigned to an analyst of ability α have the same payoff function $q = f(t, \alpha, 0)$ and this exhibits eventual diminishing returns to processing time. It follows that every accepted project should receive the same (finite) processing time. Let T be the time spent on an accepted project inclusive of communication time. An amount of time $\delta/2$ is lost in both the assignment and report communication. This leaves $(T - \delta)$ available for processing. Only a fraction $p(s,n)$ of this time is spent in productive processing, so the payoff from a project processed by an analyst of ability α is $q = f(p(s,n)(T - \delta), \alpha, 0)$.

Next consider the choice of how many projects to accept for processing each period. Suppose there are n analysts and m accepted projects. A single analyst can process $1/T$ projects per period. If $n > mT$, then there are more analysts than are needed to process the accepted projects and the wage bill could be reduced by firing idle analysts with no reduction in output. If $n < mT$ then there will be some accepted projects that go unanalyzed. These excess projects could instead be declined with no change in output or costs. It can therefore be assumed without loss of generality that $m = n/T$ in the optimal structure. This means that the principal will spend time $n\delta/T$ in communication in each period, leaving $(1 - n\delta/T)$ available for supervision.

The design problem can now be characterized as choosing n, α and T to maximize profit per period,

$$\Pi = nf(p(1 - n\delta/T, n)(T - \delta), \alpha, 0)/T - nw(\alpha). \quad (1)$$

[9] I am grateful to an anonymous referee for raising this issue.

[10] This is equivalent to wealth maximization because all periods are identical.

This problem should properly be formulated as an integer-programming problem because n is not a continuous variable. However, useful insights can be gained by treating n as continuous and examining the first-order conditions for an interior maximum:

$$f(.)/T = w(\alpha) + n(T-\delta)[(\delta p_1/T) - p_2]f_1/T. \tag{2}$$

$$f_2/T = w'(\alpha). \tag{3}$$

$$nf_1[p(.) + (T-\delta)n\delta p_1/T^2]/T = nf(.)/T^2. \tag{4}$$

These conditions equate marginal benefit and marginal cost. Consider condition (2). Hiring one more analyst permits the processing of $1/T$ more projects each worth f(.). The cost of hiring that analyst comprises the direct wage cost plus the cost of reduced productivity associated with reduced supervision. Supervision is reduced because the principal must devote δ/T extra time to communication and because available supervision time must be spread across one more analyst. The reduced productivity associated with reduced supervision represents a 'loss of control' effect. This effect has been widely identified in various forms in the existing literature [11]. Condition (3) dictates that the value marginal product of ability be equated to its marginal wage cost. Condition (4) equates the marginal cost and benefit of increasing processing time on accepted projects. The LHS represents the marginal benefit. On n/T projects per period there is a gain comprising the productivity of the extra processing time itself plus the boost to productivity associated with the increased supervision time available when fewer projects are accepted. The marginal cost of increased processing time is the reduction in the number of projects accepted.

The above conditions for an interior optimum provide some insight into the nature of the tradeoffs within the organization. However, it should be stressed that there are limits to the size and scope of the organization even if the second-order conditions for an interior solution are not satisfied. In particular, there will always be strictly less than $1/\delta$ analysts hired or else communication will consume all of the principal's time, and the resulting absence of supervision will completely undermine productivity.

4. Serial processing

The design problem for the serial structure comprises the same three stages as for the parallel structure problem: choosing a time allocation rule, choosing the

[11] See for example Williamson (1967), Simon (1976), Calvo and Wellisz (1978 and 1979), Rosen (1982) and Keren and Levhari (1989).

project acceptance rate, and choosing the number of analysts and their ability types. Consider first the time allocation problem. All accepted projects have the same payoff function $q = \Sigma_{i=1}^{n} f(t_i, \alpha_i, q_{i-1})$ and there are eventual diminishing returns to processing time. It follows that all accepted projects will receive the same processing time. Let T be the time spent on an accepted project inclusive of communication time.

Now consider how that time should be divided across analysts. Every accepted project has the same payoff function so if a particular allocation rule is optimal for one project then it is optimal for all projects. Consideration can therefore be restricted to allocation rules in which all accepted projects are treated identically. Consider an arrangement in which the k^{th} analyst in the series spends more time on a project than does the $(k+1)^{th}$ analyst. Then the $(k+1)^{th}$ analyst will transfer each project before he receives a new project and will therefore spend some time idle. This cannot be optimal in the steady state. Similarly, if the k^{th} analyst in the series spends less time on a project than does the $(k+1)^{th}$ analyst then the $(k+1)^{th}$ analyst will receive new projects more quickly than he dispatches current projects. This means that a backlog will develop and eventually some accepted projects will receive no attention. This too cannot be optimal in the steady state. It follows that all analysts should spend the same time on each project. Thus, if there are n analysts in the series then each will spend T/n on each accepted project, inclusive of communication time. Since there are n analysts each spending T/n on accepted projects it follows that there will be n/T projects accepted for processing in each period. The principal will therefore spend $n\delta/T$ time in communication each period leaving $1 - n\delta/T$ available for supervision. The payoff from each accepted project will therefore be $q = \Sigma_{i=1}^{n} f(p(1 - n\delta/T, n)[(T/n) - \delta], \alpha_i, q_{i-1})$.

Now consider the choice of n and $\{\alpha\}$. It will be pedagogically useful to begin by examining a structure in which all analysts have the same ability. The payoff from a project processed in such a structure will be $q = \Sigma_{i=1}^{n} f(\hat{t}, \alpha, q_{i-1})$ where $\hat{t} = p(1 - n\delta/T, n)[(T/n) - \delta]$. Since information processing is temporally additive across analysts of equal ability (by assumption), this payoff can be rewritten as $\Sigma_{i=1}^{n} f(\hat{t}, \alpha, q_{i-1}) = f(n\hat{t}, \alpha, 0)$. Profit per period will therefore be

$$nf(np(1 - n\delta/T, n)[(T/n) - \delta], \alpha, 0)/T - nw(\alpha). \tag{5}$$

Comparing (1) and (5) yields the following result:

Lemma 1. A serial structure with uniform ability is payoff-equivalent to the parallel structure if $\delta = 0$.

Proof. Set $\delta = 0$ in (1) and (5) and the result is obvious.

This result follows more or less by construction. There is no difference between parallel processing and serial processing with agents of uniform ability except for the amount of communication involved. In a parallel structure there are only two communication exchanges for each project. In a serial structure there are n communication exchanges. It is straightforward to show and intuitively obvious that for this reason the parallel structure is superior to a serial structure with uniform ability when $\delta > 0$. But when $\delta = 0$ the two structures are equivalent.

Our interest in lemma 1 lies in its usefulness as a benchmark for comparing the parallel structure with the optimal serial structure. The next result shows that optimal serial structure generally does not involve uniform ability.

Proposition 1. In the optimal serial structure, ability decreases along the series. Projects are transferred to progressively lower ability analysts.
Proof. See the appendix.

This result reflects the benefits to be reaped by matching ability with the evolving complexity of the processing task. The initial stages of processing are the most complex and should be assigned to higher ability analysts. Progressively lower ability analysts are assigned to the progressively less complex processing stages. It is important to note that this result does not reflect exogenously imposed gains to specialization. The specialization here arises endogenously through the optimal exploitation of comparative advantage.

A number of other papers have shown that hierarchical structures may be optimally ordered according to ability, but for quite different reasons. Calvo and Wellisz (1979) argue that the most productive manager is placed at the top of a supervisory hierarchy because there is a multiplicative productivity effect through the hierarchy. Similarly, Geanakoplos and Milgrom (1991) find that a hierarchy in which decisions are made sequentially should be ordered according to ability because a decision made quickly at an upper level can have a multiplicative effect on lower levels of the hierarchy. In a rather different vein, Rosen (1982) argues that the most able manager should be placed at the top of a supervisory hierarchy because of a scale economy in management. The most able manager should be given the widest span of control so that his high quality decisions can affect the greatest number of subordinates. My proposition 1 is closest in spirit to Calvo and Wellisz (1979) and Geanakoplos and Milgrom (1991). Matching high ability analysts with the early, more complex processing tasks boosts the productivity of analysts working on the later stages of the processing. In this sense there is a multiplicative productivity effect.

The next result indicates that the benefits of matching ability with processing complexity give the serial structure a potential advantage over the parallel structure.

Proposition 2. If δ is sufficiently small then the optimal serial structure yields a higher payoff than the parallel structure.
Proof. The uniform ability serial structure is payoff equivalent to the parallel structure when $\delta = 0$ (by lemma 1). But the optimal serial structure does not involve uniform ability (by proposition 1) and so must yield a strictly higher

payoff than the parallel structure when $\delta = 0$. By continuity, the optimal serial structure will also yield a higher payoff when δ is positive but small.

The serial structure allows processing ability to be matched with processing complexity. The parallel structure does not. At most stages of processing the analyst in the parallel structure will be less suited to his task than an analyst in the serial structure. The serial structure permits the exploitation of comparative advantage through specialization; in contrast, analysts in the parallel structure are generalists. However, there are limits to the net gains from specialization in processing, due to communication costs. If communication costs are too high then the time lost at each transfer in the serial structure will more than offset the specialization gains, and a parallel structure will be superior. The optimal structure therefore involves a tradeoff between the benefits of specialization in the serial structure and the benefits of lower communication costs in the parallel structure.

The analysis has so far assumed that the principal does not participate in information processing but instead devotes all of her time to communication and supervision. This need not necessarily be the case. The principal may not completely specialize in supervision if the returns to supervision are steeply diminishing. She may instead choose to divide her time between supervision and processing. Proposition 2 continues to hold in this case. Because the principal can place herself in the serial structure at a point that best matches her processing ability with processing complexity, she must be able to derive a higher product from a given amount of processing time than she could in the parallel structure. Of course if communication is very costly then this may no longer be true. In particular, if communication is sufficiently costly then it may be optimal for the principal to forego the benefits of matching her ability with task complexity in order to avoid duplication in communication. She would instead place herself either in the first or last position in the series because at any other position she would have to engage in four rounds of communication rather than two. If she has relatively high processing ability then she will choose to occupy the first position in the series, even if she is less able at processing than some of her analysts. On the other hand, if she has relatively poor processing ability then she will choose to occupy the last position in the series, even if she is more able at processing than some of her analysts. In either case she cannot fully exploit her own comparative advantage in processing. Of course if communication costs are high enough then it will not be worthwhile to exploit comparative advantage for any of the analysts, and the serial structure will be dominated by the parallel structure.

5. Discussion

I believe the foregoing results can help to shed some light on the two empirical issues raised in the introduction. My results indicate that different organizational forms can be at least partly explained in terms of the nature of the information processing undertaken by those organizations. Recall the law firm example from the introduction. My results suggest that law firms tend to be organized as parallel structures because there are limited gains from specialization at different stages in the conduct of a case and/or because there are high communication costs associated with a serial handling of a case. This is not inconsistent with casual observation. It is undoubtedly difficult to communicate the subtleties of a law case. While it may be relatively straightforward (though nonetheless time-consuming) to communicate the 'facts' of the case, less objective information – such as a 'hunch' – may be impossible to communicate perfectly. These high communication costs are likely to dominate any gains that might exist from specialization at different stages of preparing a case. In contrast, the gains to specialization in an entrepreneurial corporation are likely to outweigh the higher communication costs associated with a hierarchical structure. The skills required to identify a profitable opportunity are quite different from those needed for figuring out the details of implementing a chosen strategy. There is clear scope for specialization here. Moreover, inter-analyst communication will for the most part require only the transfer of facts or broad directives and so communication costs are likely to be relatively low.

I make no claim that these casual observations constitute empirical support for my hypothesis about organizational structure. Nonetheless, I believe there is enough consistency between my results and the stylized evidence to suggest that the tradeoff between specialization and communication requirements is an important determinant of the structure of organizations. Just how important this factor might be can only be ascertained through proper empirical research.

Now consider the second empirical issue raised in the introduction. The argument put forth by some management analysts is that developments in information processing and communication technology is giving rise to flatter organizations. This argument would be consistent with my hypothesis about organizational form if those technological developments led to higher communication costs and/or reduced gains to specialization in processing. The second possibility is more immediately plausible than the first. The use of increasingly sophisticated computer software is likely to broaden the skill sets of analysts in a way that reduces the specificity of comparative advantage. The relative demise of actuaries in the wake of computer developments is a clear example of the computer displacing the information processing specialist. Developments like these tend to reduce the gains to specialization and the benefits of a serial processing structure. But what about communication costs? It seems natural to presume that computer developments have reduced the costs of communication, which would tend to favor serial structures. However, this presumption may be quite wrong. Computers have certainly made it easier to physically transfer data from one agent to another, but there is more to communication than the simple transfer of data. Communication in my model involves the transfer of a full understanding of the transferred data. It is not clear that computers have helped agents to more easily absorb and comprehend information. Indeed, the ease with which data can now be transferred may actually have caused communication costs to *rise*: information of marginal significance that may once have been omitted from a communicated report might now be included, thereby increasing the costs of sorting for the receiver. Most managers complain of 'information overload'.

In summary, I believe there is a plausible link between computer developments and a shift towards flatter organizational structures that is consistent with my hypothesis about organizational form. However, the fact that hierarchical organizations are becoming leaner is not necessarily reflective of a flattening effect. The observed restructuring could also be attributable to a computer-induced improvement in the absolute productivity of all analysts. It may have nothing to do with reduced gains to specialization at all. A natural way to distinguish between these two possible effects is to examine the *relative* extent to which serial and parallel structures have become leaner. The absolute productivity effect should arise in both serial and parallel organizations more or less to the same degree. Conversely, if my hypotheses are correct, then computer-induced reduced gains to specialization will only impact on serial structures.

6. Conclusion

This paper has examined optimal organizational structures for processing information. The existence of the organization arises endogenously as a response to the limited attention of the principal decision-maker. The optimal size and scope of the organization also reflect this limited attention through the explicit recognition that the principal must communicate with and supervise the work of her analysts. These activities are time-consuming and become more burdensome as the organization grows.

The main results are: (1) a serial structure permits an optimal matching of task complexity with the abilities of analysts and is therefore optimally ordered by ability; and (2) the choice between a serial structure and a parallel structure involves a tradeoff between the benefits of specialization in the serial structure and lower communication costs in the parallel structure. If communication costs are sufficiently small then the serial structure is superior. These results are independent of whether or not the principal shares in the information processing or specializes completely in communication and supervision activities. If the principal does participate in information processing then the existence of communication costs means that her optimal position in the serial structure need not be commensurate with her relative processing ability.

Acknowledgements

This is a revised version of Chapter 4 of my doctoral dissertation from Queen's University. Valuable comments were provided at various stages in the evolution of this paper by Bentley MacLeod, Dan Bernhardt, Curtis Eaton and participants at a seminar at the University of British Columbia. I am particularly grateful to an anonymous referee for many valuable suggestions. Any errors are of course my own. This research has been partly funded by SSHRC Grant No. 410-93-1528.

Appendix

Proof of Proposition 1. Consider a serial structure with a given number of analysts and given project acceptance rate, and with a pair of adjacent analysts k and k + 1 such that $\alpha_k \leqslant \alpha_{k+1}$. I will show that profit can be increased by raising α_k and reducing α_{k+1}. Let $d\alpha_k$ be the change in each. The associated change in profit per period is

$$d\Pi = (n/T)(\partial q_n/\partial q_{k+1})dq_{k+1} + n[w'(\alpha_{k+1}) - w'(\alpha_k)]d\alpha_k, \qquad (A1)$$

where

$$dq_{k+1} = [f_2(\alpha_k) - f_2(\alpha_{k+1})]d\alpha_k + f_2(\alpha_k)f_3(\alpha_{k+1})d\alpha_k.$$

Note that $dq_{k+1} > 0$ since $f_{22} \leqslant 0$, $f_{23} \leqslant 0$, $f_2 > 0$ and $f_3 > 0$. The last term in (A1) is non-negative since $w''(\alpha) \geqslant 0$. It remains only to be shown that $(\partial q_n/\partial q_{k+1}) > 0$. By temporal additivity of the information-quality increment function,

$$q_n = q_{k+1} + \sum_{i=k+2}^{n} f(\hat{t}, \alpha_i, q_{i-1}).$$

Since $f_3 > 0$, it follows that for two ability profiles $\{\alpha\}$ and $\{\alpha'\}$ such that $\alpha_i = \alpha_i' \forall i \in \{k+2, \ldots, n\}$, $q'_{k+1} > q_{k+1} \Rightarrow q'_n > q_n$. Thus, $(\partial q_n/\partial q_{k+1}) > 0$.

References

Becker, Gary S., 1965, A theory of the allocation of time, Economic Journal 75, 493–517.

Beckmann, Martin J., 1977, Management production functions and the theory of the firm, Journal of Economic Theory 14, 1–18.

Business Week, 1989, Special report: a new era for management, April 25, 50–86.

Calvo, Guillmero A. and Stanislaw Wellisz, 1978, Supervision, loss of control and the optimal size of the firm, Journal of Political Economy 86, 943–952.

Calvo, Guillmero A. and Stanislaw Wellisz, 1979, Hierarchy, ability, and income distribution, Journal of Political Economy 87, 991–1010.

Geanakoplos, John and Paul Milgrom, 1991, A theory of hierarchies based on limited managerial attention, Journal of the Japanese and International Economies 5, 205–225.

Gifford, Sharon, 1992a, Allocation of entrepreneurial attention, Journal of Economic Behavior and Organization 19, 265–284.

Gifford, Sharon, 1992b, Innovation, firm size and growth in a centralized organization, RAND Journal of Economics 23, 284–298.

Keren, Michael and David Levhari, 1983, The internal organization of the firm and the shape of average costs, Bell Journal of Economics 14, 474–486.

Keren, Michael and David Levhari, 1989, Decentralization, aggregation, control loss and costs in a hierarchical model of the firm, Journal of Economic Behavior and Organization 11, 213–236.

Oi, Walter Y., 1983, Heterogeneous firms and the organization of production, Economic Inquiry 21, 147–171.

Radner, Roy, 1975, A behavioral model of cost reduction, Bell Journal of Economics 6, 196–215.

Radner, Roy and Michael Rothschild, 1975, On the allocation of effort, Journal of Economic Theory 10, 358–376.

Rosen, Sherwin, 1982, Authority, control and the distribution of earnings, Bell Journal of Economics 13(2), 311–323.

Schein, Edgar H., 1989, Reassessing the divine rights of managers, Sloan Management Review, Winter, 36–43.

Simon, Herbert, 1976, Applying information technology to organization design, reprinted in Administrative Behavior, 3rd ed., (MacMillan, New York).

Williamson, Oliver, 1967, Hierarchical control and optimum firm size, Journal of Political Economy 75, 123–138.