
Competing in Information-Intensive Services: Analyzing the Impact of Task Consolidation and Employee Empowerment

ABRAHAM SEIDMANN AND ARUN SUNDARARAJAN

ABRAHAM SEIDMANN is the Xerox Professor and Areas Coordinator of Computers and Information Systems, Management Science and Operations Management at the William E. Simon Graduate School of Business Administration at the University of Rochester, New York. Professor Seidmann is the author of numerous research articles and is a Department Editor on Interdisciplinary Management Research and Applications in Management Science. He is also an Associate or Area Editor for *IIE Transactions*, *International Journal of Flexible Manufacturing Systems*, *Production Planning and Controls*, *Journal of Intelligent Manufacturing*, *Journal of Management Information Systems*, and *Production and Operations Management*. His current research and consulting activities include business process reengineering, strategic manufacturing systems, information economics, health-care management, stochastic processes, and performance modeling for capacity planning. Professor Seidmann has consulted with many of the leading industrial and service corporations and presented research or executive seminars in four continents.

ARUN SUNDARARAJAN is Assistant Professor of Information Systems at the Leonard N. Stern School of Business, New York University. His research analyzes a variety of topics in business process reengineering, interorganizational information sharing, information economics, network economics, and the impact of information technology on the extent and shape of the organization.

ABSTRACT: We analyze the competitive and economic implications of information technology, the allocation of decision rights, and task bundling during business process reengineering. The popular reengineering literature advocates employee empowerment—decentralizing decision authority and consolidating tasks—as complementary strategies. Our analysis reveals, however, that implementing these two changes simultaneously is suboptimal in many cases. Decentralization and consolidation decisions can occur separately or together; the optimal combination depends on the *effectiveness of technology* aimed at skill enhancement and the customers' *sensitivity to time and quality*. We identify those process parameters that can cause decentralization and consolidation to have opposite effects on process performance; we also point to other parameters, such as customer-to-customer variability, which can cause them to complement one another. Finally, we explain why, in a time-based competitive marketplace, firms are more likely to centralize their decision-making process while concentrating their information technology investments on enhancing productivity and intraorganizational communications.

KEY WORDS AND PHRASES: business process reengineering, business process redesign, case manager, decentralization, economics of information systems, mass customization, queueing, strategic and competitive information systems, triage.

OUR RESEARCH IS MOTIVATED BY A SERIES OF PROCESS REENGINEERING projects we have done with a Fortune 50 company. The typical approach to reengineering advocates empowerment through decentralizing decision rights and increasing job scope [12]. However, we found the approach did not always work well. Besides, the nature of the information systems that accompany these organizational changes depends on whether tasks are bundled, decision rights are shifted, or both. For instance, service representatives have been required to answer an increased scope of calls and to perform more tasks on site. However, they are not allowed to make decisions on replacing certain defective parts or formulating service or payment strategies tailored to a customer; they either follow a set of rules dictated by a central engineering team or communicate with experts via their laptops. Although this reduces service duration by allowing one point of contact, the final outcome is not as favorable as one that would result from a thorough analysis of the specific problems presented by the customer. Similarly, in insurance sales, companies expand the decision scope of their sales agents by allowing them to structure a variety of insurance packages. However, they still retain centralized authority for processing all claims on these packages.

This paper offers insights on the options available for reengineering service processes. It addresses the alteration of task boundaries and coordination across them. We contrast sequential and parallel process designs, and specialized and generalized queues, and simultaneously examine when the two facets of empowerment—expanding the scope of an employee's work and expanding their rights to make decisions about their work—are desirable.

When managers reengineer processes, they face the problem of how much to empower their employees. Sometimes, managers may believe that total empowerment is not ideal, but they have no guidance in determining when to bundle tasks and when to allocate decision-making authority. They also face significant difficulties in determining the nature of the new information systems that must be provided to these empowered employees. The design choices made affect timeliness, service quality, profitability, and service volume in different ways. Our research is aimed at giving managers guidelines derived from a generalized framework based on the rigorous analysis of their process design options and the nature of their customers.

Earlier work done in related areas includes, among others, Barua, Lee, and Whinston [1], who analyze the complementarities that may exist between a firm's technology choice and organizational design variables; Buzacott [3], who studies the operational aspects of process design; Clemons, Thatcher, and Row [5], who discuss how the failure of reengineering is related to a firm's lack of understanding of future strategic needs; Clemons and Weber [6], who examine the role of IT in meeting specific customer needs; and Sampler and Short [17], who discuss how the approach to reengineering depends on the longevity of a firm's information and expertise.

Decision Making with Task Separation

WE ANALYZE THE SERVICE-DELIVERY PROCESS OF A FIRM (henceforth, all references to the "firm" indicate the provider of this service). The process consists of two task-processing steps. We assume that the ordering of these tasks is predetermined. Each task has a processing time that is exponentially distributed. Service requests arrive at a Poisson rate λ . In the absence of any additional information processing entailed by making decisions, reading reports, and so on, the processing rate of a worker working on a single task is μ .

There are two primary forms of information exchange that could take place in processes of this kind: communication between two employees and communication between the employee and the customer. These elements are fundamental when deriving the appropriate job design. To understand the complexity of these issues better, consider the following case:

ABC Corporation is an insurance company that provides business insurance coverage for clients in a variety of industries. Their typical sales cycle consists of two steps. In the first step, an industry expert in the specific industrial field (such as aerospace, textiles, or health care) visits the client to determine his or her coverage requirements and to ascertain the risk ABC will bear by covering this client. This information is then transferred to a second employee (typically an actuarial expert), who designs the financial policy and payment schedule for the package to be offered to the customer. The job is therefore completed by two different specialists. The workers in the process are efficient and have high levels of expertise.

However, ABC has realized that the final decision on what product to offer depends on two highly interdependent sets of information from the customer. For instance, the customer may want extensive coverage with a particular payment schedule that also meets specific cash flow capabilities such as deductibles, credit terms, and volume discounts. If ABC cannot meet these needs, the field agent must renegotiate the coverage profile, the payment schedule, or both, and the financial expert must evaluate the new proposal. These information gaps cause a significant amount of rework, mistakes, and delays in many cases.

ABC, therefore, is faced with two alternate approaches. The first involves *bundling* these tasks and empowering the field agent to handle both steps. This approach entails the creation of multifunctional case workers who have the expertise to determine coverage requirements and design financial schedules. It will provide a single point of contact and allow a single worker to understand and process both sets of specifications before making a decision. Since these case workers would not be experts in all aspects of their work, they would have to be provided with standardized templates that would serve as a guideline for risk evaluation and policy decisions. However, lower levels of specialization would also increase the time taken to complete each task, and the templates would lower the level of customization ABC could offer a client. Besides, the case workers would require expensive training and technology support over a broad range of industries, and in a variety of financial evaluation skills.

The second approach involves the creation of *triaged*¹ generalists in each industry. Semispecialized agents for each industry would determine requirements and risk and design the policies for the clients in that particular industry. When clients are triaged by industry, the losses in specialization are reduced. In addition, triage offers higher customization levels relative to the option of using a single case worker for all industries, as

more templates could be tailored for each industry. The triaged approach, however, could result in higher overall net turnaround time for processing an insurance policy, since the applications are not consolidated into a single queue. Having several parallel systems could lead to imbalanced utilization of the available manpower due to the separation of workers across industries.

The issues described above are not unique to this case. Many financial intermediaries suffer from information gaps between their relationship managers and their product specialists; however, when contemplating consolidation, they face the prospect of slower turnaround because of significant losses in expertise.

We model the interdependent information sets as two customer *specifications information* variables, γ_1 and γ_2 . In the case discussed above, γ_1 corresponds to the desired coverage profile, and γ_2 corresponds to the desired payment schedule. We assume that both variables are uniformly distributed from $-v$ to $+v$. The parameter v represents the degree of variability in customer specifications. The choice of the range $-v$ to $+v$ is without loss of generality and simplifies the subsequent algebra; the results would be equivalent if one used a range of 0 to $+v$, or 0 to $+2v$. A low value of v indicates that γ_1 and γ_2 have low variance; a higher value would imply that these variables can take a wider range of values, and that there is therefore greater variability in customer specifications. We limit the values of v to be between 0 and 1. $v = 0$ corresponds to a completely standardized service, while $v = 1$ corresponds to highly variable customer requirements. Interpreting the information contained in a particular specification—that is, determining the value of each γ_i —requires specialized knowledge (in the example above, the field agent has the skills to determine coverage requirements and risk, and the pricing expert had the expertise to determine the suitability of a payment schedule); the class of knowledge required to perform task i is the same as the class of knowledge required to recognize the value of γ_i . Since the skills needed to assess γ_i and the skills needed to perform task i are the same, they cannot be separated readily or without performance degradation.

The specifications variables described above model customer requirements for a particular job. The values of these variables are contained in the specifications information provided by the customer. It may not be possible to meet both specifications simultaneously. In our example, if the insurance policy is based solely on the first criterion and a complete coverage policy is issued, the deviation from the second specification (desired payment schedule) could be extreme. The same holds for decisions based solely on the second criterion. A policy that minimizes the net deviation from customer requirements is normally chosen. We model this decision as one of choosing a *service characteristic* value δ (this is also referred to as *service quality*). The choice of this variable may be made centrally by a manager or by one of the workers performing the task. Customer preferences are characterized by a cost of delay and a cost of deviation from their specifications. More precisely, if the specifications information for a customer is (γ_1, γ_2) , the service characteristic is δ , and the total time taken to provide the service is T , then the price that the customer is willing to pay for service δ is

$$p = p_0 - c(\gamma_1, \gamma_2, \delta, T),$$

where

$$c(\gamma_1, \gamma_2, \delta, T) = c_D T + c_Q [(\gamma_1 - \delta)^2 + (\gamma_2 - \delta)^2] - c_Q \frac{(\gamma_1 - \gamma_2)^2}{2}.$$

The value p_0 is the amount a firm can charge for providing exactly what the customer requires; in many cases this may not be possible, as the specifications may be inconsistent. For instance, a customer may request a low-cost insurance policy with no deductible and extensive coverage. The last term in the function, $c(\cdot)$, is an adjustment to ensure that the optimal (from the customer's perspective) choice of δ (by the firm) will impose a zero deviation cost on the customer and therefore on the firm. Without this term, certain sets of specifications will reduce the price a firm can charge, whatever the level of service; this is not realistic. With this adjustment, a firm that provides optimal service quality can always charge a maximum of p_0 .

Figure 1 illustrates how we model standardized and customized service. The two axes measure the two customer specifications. The diagonal line is the set of possible products the firm can offer. The firm chooses a particular value d on this diagonal (the diagonal is the line on which $\gamma_1 = \gamma_2$). However, the point on the diagonal closest to (γ_1, γ_2) is the best choice the firm could make for the customer (or the best substitute available to the customer). The square of the distance between the customer's actual specifications, and the best possible product is $(\gamma_1 - \gamma_2)^2/2$. The square of the actual distance between the customers' specifications and the standard product (the dashed line joining d and (γ_1, γ_2)) is $[(\gamma_1 - \delta)^2 + (\gamma_2 - \delta)^2]$. The cost borne by the customer is proportional to the square of the distance from the actual product (d) to the product of best fit (b). It is verifiable by observation that this value is $[(\gamma_1 - \delta)^2 + (\gamma_2 - \delta)^2] - [(\gamma_1 - \gamma_2)^2/2]$.

Evidently, the firm can charge a higher price if it reduces the costs imposed on the customer. There are thus two potentially conflicting objectives: reducing cycle time and reducing deviation from customer specifications. Intuitively, a standard service can be provided faster, while customized service will take longer. We investigate these tradeoffs and their associated organizational costs in a process-modeling framework. First, we characterize the firm's optimal choice of δ :

Lemma 1: The value of δ that maximizes the firm's profits is

$$\delta = \frac{\gamma_1 + \gamma_2}{2}.$$

This lemma is contingent on the assumption that providing different choices of δ cost the same to the firm and do not affect the service rates of the workers. While relaxing these assumptions is analytically possible, it adds little value to our research and makes the mathematics considerably messier. This choice of δ minimizes customer costs and therefore maximizes the price the firm can charge. It would be the firm's quality choice under ideal conditions. However, this assumes two conditions: that the specifications information (γ_1, γ_2) can be *accurately* determined by the decision maker, and that the

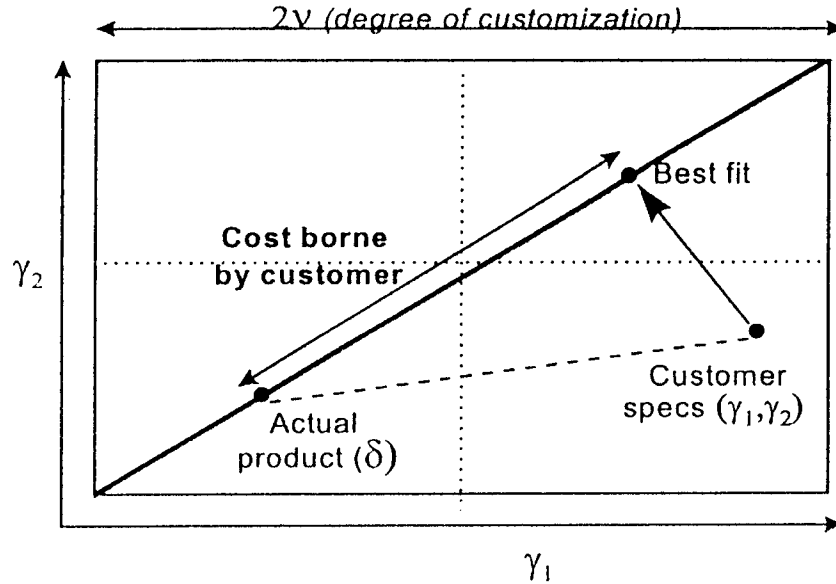


Figure 1. Degree of Customization and Possible Standard Products

determination of these values is costless. There are certain delay-related costs, fixed organizational costs, and information barriers that may preclude the fulfillment of these two conditions, which we also investigate.

Centralized Decision Making with Specialized Task Processing

Our base case is one of centralized decision making and specialized task processing (figure 2a). The service characteristic decision is made centrally, and the separate workers then perform the tasks sequentially. Realistically, a service characteristic decision cannot be made centrally for every instance of the service, as this entails additional information-transmission and decision-making delays. Instead, we assume that a standard decision is mandated; this will be the one that minimizes the expected deviation from customer specifications over the state space of customer specifications information. This corresponds to the use of templates in our case about ABC. The mandated decision that minimizes this deviation is $\delta = 0$ (this is easily verifiable). This model of product specification is fairly common. Clemons and Weber [6] show that, in a wide range of industry, this kind of “one-size-fits-all” product design and pricing is common and encourages successful predatory attack by new entrants. As we shall show below, however, empowering workers to make even rudimentary and sometimes imperfect assessments of customer requirements can produce profound improvements in customer satisfaction and corporate performance.

Since a decision is mandated, there is no additional delay imposed by workers making decisions, and the workers perform their tasks at an exponential rate μ . The expected cost to any customer is therefore:

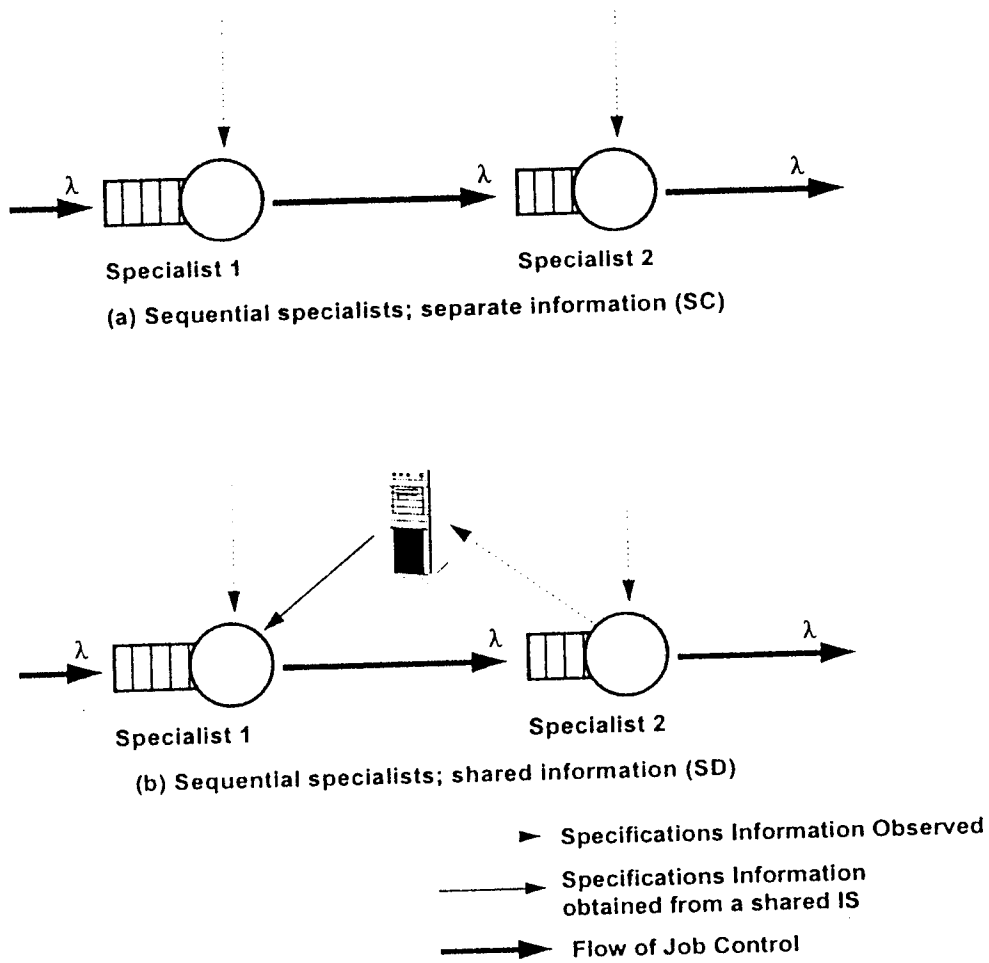


Figure 2. Process Flows and Information Systems for Separated Tasks

$$(1) \quad \frac{2c_D}{\mu - \lambda} + \frac{c_Q}{4v^2} \int_v^{-v} \int_v^{-v} \left[(x-0)^2 + (y-0)^2 - \frac{(x-y)^2}{2} \right] dy dx = \frac{2c_D}{\mu - \lambda} + \frac{c_Q v^2}{3}.$$

The profits to the firm per unit time are therefore

$$(2) \quad \pi_{SC}(\lambda, v) = \lambda \left[p_0 - \frac{2c_D}{\mu - \lambda} - \frac{c_Q v^2}{3} \right] = p_0 \lambda - \frac{2c_D \lambda}{\mu - \lambda} - \frac{c_Q v^2 \lambda}{3}.$$

The following result can be immediately inferred:

Lemma 2: Under centralized decision making and specialized task processing, the profits of the firm are decreasing in the level of customer variability.

This can be proved easily by verifying that $(\partial \pi_{SC})/\partial v$ is negative. Having analyzed our base case, we now investigate the effects of decentralizing decision making, and consolidating tasks.

Decentralized Decisions with Specialized Task Processing

The first common form of process reengineering we consider is the decentralization of decision authority. As mentioned earlier, the knowledge required to perform task i is the same as the knowledge required to observe and interpret γ_i . Therefore, when successive specialists perform their tasks, the specialist who performs task 1 has the skills required to assess the value of γ_1 , and the specialist who performs task 2 has the skills required to assess the value of γ_2 . However, worker i does not possess the knowledge required to assess the value of $\gamma_j, j \neq i$.

Reading and interpreting specifications or requirements are not costless in terms of delays, and customizing a service requires additional processing effort. With these points in mind, we assume that, if worker i determines γ_i and chooses a value of δ based on that observation, then the processing rate of worker i is reduced by a factor α_v , which depends on the degree of customization, v . Since the tasks are sequential, the first worker will have to make the decision.

Suppose the decision making is delegated to the first worker. This worker determines the value of γ_1 but has no information about the value of γ_2 . Assume that incentives are aligned and the worker wishes to make a decision that is optimal for the firm (this can be ensured by tying the worker's performance to service quality). Suppose the decision made is δ . The total cost imposed on the customer during the processing of the first task is

$$c_1 = \frac{c_D}{\alpha\mu - \lambda} + c_Q (\gamma_1 - \delta)^2.$$

Also, the expected cost imposed on the customer in the second stage is

$$c_2 = \frac{c_D}{\mu - \lambda} + \frac{c_Q}{2v} \int_{-v}^v (\delta - x)^2 dx = \frac{c_D}{\mu - \lambda} + \frac{c_Q}{2v} \frac{(\delta + v)^3 - (\delta - v)^3}{3}.$$

The worker chooses the value of δ that minimizes total costs $c = c_1 + c_2$. This value can be determined by solving $\partial c / \partial \delta = 0$, and leads to our next result:

Proposition 1: In the absence of any additional local information sources, in a specialized sequential process, decentralized and uninformed decision making is inferior to centralized decision making if variability is sufficiently low.

The mathematical exposition and proof of this proposition is in the appendix. The natural question one poses at this point is whether a specialized worker can receive and process information not directly related to his or her specialty. This is possible in the presence of an information-sharing system. We build on approaches used by Marschak and Radner [15] and Carter [4] in defining and analyzing the presence of this kind of system.

Decentralized Decision Making with an Information Sharing System

From the preceding analysis, it is evident that some kind of additional information system is required to decentralize decision making in a specialized task-processing

environment. One such system would enable *information sharing* between the two workers. The process and information flows using such a system are illustrated in figure 2b. This system would allow the second worker to transmit the information in γ_2 to the first worker. However, it is likely that some error will be introduced because worker 1 has no specialized knowledge related to γ_2 and the information is observed and used by different individuals. We therefore define an information-sharing system as one that takes as an input a specification information variable γ and returns a report $\gamma + \varepsilon(x)$, where $\varepsilon(x)$ is an error term distributed with mean 0 and variance σ_ε^2 . The value of σ_ε is a reflection of the *quality* of the information-sharing system (a higher variance implies a higher transmission error and thus an inferior system). It is also increasing in v —for simplicity, we assume that it is linear in v . For modeling tractability, we also assume that the process of determining γ_2 is costless² for worker 2. However, the delay generated by reading and interpreting the output of the information system causes a further reduction of α_v in the processing rate³ of worker 1. Worker 1 now knows the values of the two variables as γ_1 and $\gamma_2 + \varepsilon$. The choice of δ is therefore $\delta = (\gamma_1 + \gamma_2 + \varepsilon)/2$, and the expected cost imposed on the customer is

$$\begin{aligned} c &= c_D \left[\frac{1}{\mu - \lambda} + \frac{1}{\alpha_v^2 \mu - \lambda} \right] \\ &+ c_Q \mathbf{E} \left\{ \left[\frac{\gamma_1 + \gamma_2 + \varepsilon(x)}{2} - \gamma_1 \right]^2 + \left[\frac{\gamma_1 + \gamma_2 + \varepsilon(x)}{2} - \gamma_2 \right]^2 - \frac{(\gamma_1 - \gamma_2)^2}{2} \right\} \\ &= c_D \left[\frac{1}{\mu - \lambda} + \frac{1}{\alpha_v^2 \mu - \lambda} \right] + \frac{c_Q \sigma_\varepsilon^2}{2}. \end{aligned}$$

The expected profits of the firm are therefore

$$\pi_{SD}(\lambda, v) = \lambda \left[p_0 - \frac{c_D}{\mu - \lambda} - \frac{c_D}{\alpha_v^2 \mu - \lambda} - \frac{c_Q \sigma_\varepsilon^2}{4v} \right] = p_0 \lambda - \frac{c_D \lambda}{\mu - \lambda} \left(1 + \frac{1}{\alpha_v^2} \right) - \frac{\lambda c_Q \sigma_\varepsilon^2}{2}.$$

This profit function reflects increased losses due to congestion, and losses due to errors in estimation. Examining these expressions leads to our next result:

Proposition 2: Decentralized decision making with information sharing and support for local decisions can outperform centralized decision making, and becomes more desirable as: (a) the degree of customization v increases, and (b) the quality of information-sharing systems increases. However, the demand that a firm can optimally handle declines significantly, and the processing capacity of the non-decision-making worker is underutilized.

A couple of insights that follow immediately have to do with the sensitivity of the profit function to changes in α and σ . As the estimation error σ increases, one will see a move toward centralized decision making. Similarly, as the value of α increases

(reflecting an increase in the complexity or knowledge intensity of the tasks), one will expect centralized decision making to dominate.

Decision Making and Task Bundling

THE POTENTIAL GAINS FROM CONSOLIDATION OF DECISION-RELATED information with a single worker lead to the possibility of consolidating tasks and increasing the scope of the workers' knowledge so that a single worker has the skills required to interpret both specification information signals and perform both tasks. There are two common process designs that allow for this scenario: *parallel generalists with a shared queue* and *triaged generalists with separate queues*. Recall that these were the two options ABC considered in the case described earlier. In both cases, the workers are provided with the skill set required for both tasks; this enables them to make more informed decisions, and it may also reduce the queuing delays of the process by consolidating queues or reducing the variability of processing times. We ignore the latter effect in our current analysis and assume that, at ideal skill levels, the processing of the two consolidated tasks is exponentially distributed⁴ with parameter $\mu/2$. This allows us to use a number of closed-form results for $M/M/s$ queues (see figure 3).

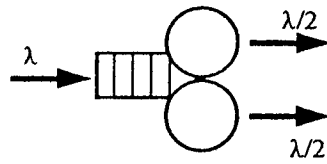
Yet there is also a loss of specialization when tasks are consolidated, which causes a net reduction in the processing rate of a worker performing consolidated tasks. The magnitude of this loss of productivity depends on the *knowledge intensity* of the tasks and the *quality of information systems support*. We model this loss of productivity by the factor β . When tasks are consolidated, the processing rates of a worker is exponentially distributed with parameter $\beta\mu/2$. The value of β is not fixed—it can be influenced by information systems that enhance worker skills such as expert systems and case-based-reasoning tools. We return to this point later and detail our modeling assumptions about the relationship between β and technology.

Centralized Decision Making and Task Consolidation

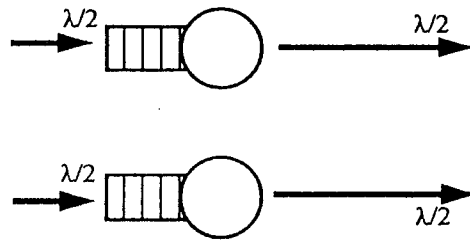
We now examine centralized decision making in a consolidated task setting. The general assumptions outlined in the previous section are still valid, as are those described in the preceding paragraphs. Decentralized information is not used; the firm mandates the service characteristic that minimizes its quality costs. Again, as described in our case, this corresponds to a scenario where nonspecialized workers adhere to predetermined templates corresponding to standard situations.

Parallel Generalists, Shared Queue

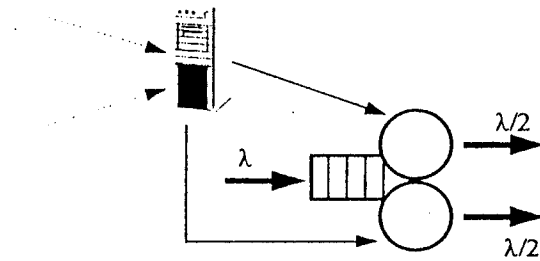
In the case of a shared queue, the centralized decision policy is still $\delta = 0$, as this minimizes expected customer quality costs. The queuing system is an $M/M/2$ with arrival rate λ and service rate $\beta\mu/2$. The expected costs imposed on the customer are given by



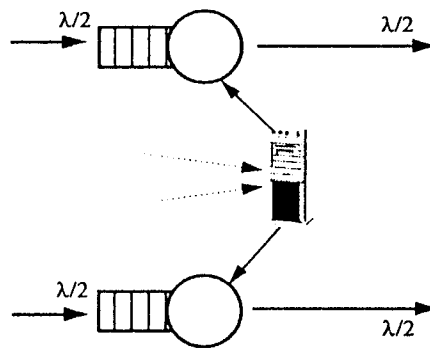
(a) Parallel units, shared queues, centralized decisions (PC)



(b) Triaged units, separate queues, centralized decisions (TC)



(c) Parallel units, shared queues, decentralized decisions (PD)



(d) Triaged units, separate queues, decentralized decisions (TD)

Figure 3. Process Flows and Information Systems for Consolidated Tasks

$$\begin{aligned}
& c_D \left[\frac{1}{\mu(1-\rho_\beta^2)} \right] + \frac{c_Q}{4v^2} \int_{-v}^v \int_{-v}^v \left[(x-0)^2 + (y-0)^2 - \frac{(x-y)^2}{2} \right] dydx \\
&= c_D \left[\frac{1}{\mu(1-\rho_\beta^2)} \right] + \frac{c_Q v^2}{3},
\end{aligned}$$

where

$$\rho_\beta = \frac{\lambda}{\beta\mu}.$$

The expected profits per unit time are therefore

$$\pi_{PC}(\lambda, v) = \lambda \left[p_0 - c_D \frac{1}{\mu(1-\rho_\beta^2)} - \frac{c_Q v^3}{3} \right] = p_0 \lambda - \frac{\lambda c_D}{\mu(1-\rho_\beta^2)} - \lambda \frac{c_Q v^2}{3}.$$

Since there is performance degradation in service rates and no improvement in “quality,” one would expect this case to be strictly worse than the others. However, the shared queue reduces queuing time significantly, thus potentially improving cycle time despite lower service rates. Our analysis provides some idea of when which effect dominates.

Parallel Generalists, Triage System

Triage is normally employed to reduce the variability in processing time [12], but there are other competitive and strategic implications of separating services into parallel separate queues. It could also, for instance, reduce *specifications information variability*. If the jobs are partitioned into two parallel streams, one of which contains requests where $\gamma_1 > 0$, and the other requests where $\gamma_1 < 0$, then a centralized decision-making scheme could be more effective, as the specifications information variability goes down, thereby reducing the quality costs imposed on the customer. Separating jobs based on industry (see our case on ABC) is one example of triage of this kind.

We state without proof that the optimal division policy is to bisect the jobs along one dimension of specifications, that is, to separate them into jobs with $\gamma_1 \geq 0$ and $\gamma_1 \leq 0$. This produces two equal streams of jobs with arrival rates $\lambda/2$. We consider the positive half ($\gamma_1 \geq 0$); by symmetry, the other half has similar results.

Let the optimal decision policy mandated by the centralized decision maker be $z \geq 0$ (it is evident that δ will be positive for the positive stream). The expected quality costs from this policy are:

$$\begin{aligned}
c_1 &= c_Q \left[\frac{1}{2v} \cdot \frac{1}{v} \int_{y=vx=0}^v \int_{y=vx=0}^v (x-z)^2 + (y-z)^2 - \frac{(x-y)^2}{2} \right] dydx \\
&= \frac{c_Q}{2v^2} \left[\frac{2v^4}{3} + 4z^2 v^2 - 2zv^3 \right] = c_Q \left[\frac{v^2}{3} - zv + 2z^2 \right].
\end{aligned}$$

The quality cost minimizing solution is at $\partial c_1/\partial z=0$, or $z=v/4$, which yields $c_1=c_Q(v^2/12)$. This reflects a *net reduction in quality costs of 75 percent*. A similar analysis for the negative stream yields $z=-(v/4)$ and identical quality costs $c_1=c_Q(v^2/12)$. The delay costs in this case are $2c_D/(\beta\mu-\lambda)$. Therefore, the expected profits per unit time are given by

$$\pi_{TC}(\lambda, v) = \lambda \left[p_0 - \frac{2c_D}{\beta\mu - \lambda} - \frac{c_Q v^2}{12} \right] = p_0 \lambda - \frac{2c_D \lambda}{\beta\mu - \lambda} - \frac{c_Q v^2 \lambda}{12}.$$

Performance under this regime will be worse than the base case because the workers are given more tasks to perform, but better than the base case because dividing the customers into two streams reduces the customer dissatisfaction. β , the factor of reduction in processing rate, may be higher (better), since there is scope for specialization within each worker's industry.

Decentralized Decision Making and Consolidation of Tasks

Since the workers have the skill set to perform both tasks, it is natural to assume that they have the ability to determine the values of both γ_1 and γ_2 . This leads to the final process design we consider, where tasks are consolidated and decision making is decentralized. We assume that the supporting information systems enable the workers not only to perform both the tasks, but also to provide decision support by indicating the values of γ_1 and γ_2 . Since there is some attrition in the skill sets of the workers (as compared with a specialist), they cannot observe the true values of γ_1 and γ_2 . Instead, as before, they observe the values $\gamma_1 + \varepsilon_1$ and $\gamma_2 + \varepsilon_2$, where $\varepsilon_i = N(0, \sigma_{ei}^2)$. As before, σ_{ei}^2 depends on the range of customization $[-v, v]$; however, *it also depends on the quality of the supporting information systems* that provide this information and enhance the workers' skills. Evidently, if the technological infrastructure allows for almost complete compensation for losses from reduced specialization, then σ_{ei}^2 will be much lower. Again, as in the case of β , we return to this issue below and detail our modeling assumptions in this regard. Determining and using these values reduces the processing rate of a task by a factor of α_v . For comparability, we assume that this factor is the same as that of the sequential specialist case.

Parallel Generalists, Shared Queue

The system is an $M/M/2$ queue with arrival rate λ and processing rate $\beta\alpha_v\mu/2$. The expected costs imposed on the customer are given by

$$c_D \left[\frac{1}{\mu(1-\rho_{\alpha\beta}^2)} \right] + c_Q \mathbf{E} \left\{ \left[\frac{\gamma_1 + \gamma_2 + \varepsilon_1(x) + \varepsilon_2(x)}{2} - \gamma_1 \right]^2 + \left[\frac{\gamma_1 + \gamma_2 + \varepsilon_1(x) + \varepsilon_2(x)}{2} - \gamma_2 \right]^2 - \frac{(\gamma_1 - \gamma_2)^2}{2} \right\}$$

$$= c_D \left[\frac{1}{\mu(1-\rho_{\alpha\beta})} \right] + c_Q \sigma_\varepsilon^{*2},$$

where

$$\rho_{\alpha\beta} = \frac{\lambda}{\alpha_v \beta \mu}, \text{ and } \sigma_{\varepsilon 1}^2 = \sigma_{\varepsilon 2}^2 = \sigma_\varepsilon^{*2}.$$

The expected net profits per unit time are therefore:

$$\pi_{PD}(\lambda, v) = \lambda \left[p_0 - c_D \left[\frac{1}{\mu(1-\rho_{\alpha\beta})} \right] - c_Q \sigma_\varepsilon^{*2} \right] = p_0 \lambda - \lambda c_D \left[\frac{\mu}{1-\rho_{\alpha\beta}} \right] - \lambda c_Q \sigma_\varepsilon^{*2}.$$

Again, though there is a reduction in processing rate due to task consolidation, the consolidation of queues could counter this, and it is not immediately clear whether cycle time reduces or increases. Customer satisfaction, on the other hand, has increased because of a more accurate assessment of customer specifications by the workers.

Parallel Generalists, Triaged System

We have two parallel $M/M/1$ queues with processing rates $\alpha_v \beta \mu / 2$ and arrival rates $\lambda / 2$. The expected costs imposed on the customer are given by

$$\begin{aligned} & 2c_D \left[\frac{1}{\alpha_v \beta \mu - \lambda} \right] + c_Q \int_{-\infty}^{\infty} \left\{ \left[\frac{\gamma_1 + \gamma_2 + \varepsilon_1(x) + \varepsilon_2(x)}{2} - \gamma_1 \right]^2 + \right. \\ & \left. \left[\frac{\gamma_1 + \gamma_2 + \varepsilon_1(x) + \varepsilon_2(x)}{2} - \gamma_2 \right]^2 - \frac{(\gamma_1 - \gamma_2)^2}{2} \right\} dx \\ & = c_D \left[\frac{2}{\alpha_v \beta \mu - \lambda} \right] + \frac{c_Q (\sigma_{\varepsilon 1}^2 + \sigma_{\varepsilon 2}^2)}{4v} = c_D \left[\frac{2}{\alpha_v \beta \mu - \lambda} \right] + \frac{3c_Q \sigma_\varepsilon^{*2}}{4}. \end{aligned}$$

Therefore,

$$\pi_{TD}(\lambda, v) = \lambda \left[p_0 - c_D \left[\frac{2}{\alpha_v \beta \mu - \lambda} \right] - \frac{3c_Q \sigma_\varepsilon^{*2}}{4} \right] = p_0 \lambda - \lambda c_D \left[\frac{2}{\alpha_v \beta \mu - \lambda} \right] - \lambda \frac{3c_Q \sigma_\varepsilon^{*2}}{4}.$$

It is clear as the variance in estimation error σ^2 goes to zero, the loss due to this error also vanishes. Moreover, separation of customers into high and low values of γ_1 provides significant improvement in customer satisfaction (this is true for any convex loss function). Also, as α goes to its maximum value of 1, the losses due to increased cycle time go to zero.

Strategic and Competitive Implications

THE AVAILABILITY OF MULTIPLE PROCESS DESIGNS HAS IMPORTANT strategic and competitive implications for a corporation. The design of the business process and its

accompanying technology can be chosen based on the strategy the firm wishes to follow. Alternatively, the strategy chosen by a firm may be a consequence of its technological infrastructure and the nature of the service provided. In a competitive market, one such strategy may be profit maximization through rapid service turnaround, if customers have high delay costs. Another strategy, possibly by a market entrant, could be gaining market share (perhaps with a short-term lowering of profits) with the objective of increasing expected future profits. A third could be profit maximization through market segmentation—if customer preferences are highly variable, tailoring services to meet the specific needs of particular market segments may be a successful strategy.

Once a strategy is chosen, the process design that best suits this strategy must be implemented. We examine the strategies outlined above in the context of our service model. Our analysis is based on the assumption that the direction of causality is from markets and services to firms; that is, there is a set of market characteristics, task parameters, and technology choices that a firm does not cause or influence. Instead, it reacts to these parameters by choosing strategies and designing business processes that best suit these strategies. One firm-specific characteristic that we examine is the technological maturity of the organization and workers—how well the firm and its employees can leverage existing technology to their advantage. For instance, a firm with a strong technology culture or a well-established learning infrastructure will get better skill enhancement returns from information systems and smaller losses in productivity from task consolidation.

Information Systems for Specifications Interpretation and Skill Enhancement

As mentioned earlier, technology can influence the losses in productivity faced by workers when tasks are consolidated, and it can facilitate superior interpretation of specifications when specialized workers need to understand diverse customer requests. Information systems such as expert systems and case-based reasoning systems can increase a low value of β and reduce the variance of the error term (σ_e^2) of γ_i . We assume that the firm chooses a technology level θ , which costs them $c_0\theta$, and which influences these parameters as described below.

We have observed in practice that the productivity gains from skill-enhancing technologies are increasing at a decreasing rate (i.e., increasing and concave). Significant gains can be achieved by providing workers with simple case-based-reasoning systems. The amount of additional information support required to go beyond these gains, however, is extremely high. In addition, it is easy to capture superficial rules that apply to broad ranges of cases, but nearly impossible to build a system that goes beyond this form of reasoning to permit understanding of the concepts underlying these rules.

With this functional form in mind, we model the relationship between β and θ as

$$1 - \beta = (1 - \beta_0) e^{-\beta_1 \theta},$$

where β_0 reflects the base knowledge intensity of the tasks ($\beta_0 \in (0,1)$, lower $\beta_0 \Rightarrow$ higher knowledge intensity), and β_1 is the sensitivity of the tasks to increased

technological support ($\beta_1 \in [0, \infty)$, higher $\beta_1 \Rightarrow$ higher returns from technology). β_1 could be a characteristic of the tasks or an attribute of the firm.

Similar reasoning goes into our functional form for σ_ε^2 . However, the variance of ε is also dependent on the range of customer specifications $[-v, v]$. For simplicity, we assume it is linear in v . The exact functional form we assume for the variance of ε is

$$\sigma_\varepsilon^2 = \sigma_{\varepsilon 1}^2 = \sigma_{\varepsilon 2}^2 = \frac{\sigma_{\varepsilon 0}^2}{2} e^{-\varepsilon_1 \theta} [v - (-v)] = v \sigma_{\varepsilon 0}^2 e^{-\varepsilon_1 \theta} (= \sigma_\varepsilon^{*2})$$

for a sequential system or a shared queue and

$$\sigma_{\varepsilon 1}^2 = \frac{\sigma_{\varepsilon 0}^2}{2} e^{-\varepsilon_1 \theta} [v - 0] = \frac{v}{2} \sigma_{\varepsilon 0}^2 e^{-\varepsilon_1 \theta} = \sigma_\varepsilon^{*2} / 2$$

$$\sigma_{\varepsilon 2}^2 = \frac{\sigma_{\varepsilon 0}^2}{2} e^{-\varepsilon_1 \theta} [v - (-v)] = v \sigma_{\varepsilon 0}^2 e^{-\varepsilon_1 \theta} = \sigma_\varepsilon^{*2}.$$

for a triaged system.

The variance of error is linear in the degree of customization and decreasing in the level of technology support. With no customization, there is no error (evidently, since $\gamma_1 = \gamma_2 = 0$), and with very high technology support, there is negligible error. Since the range of values is halved for γ_1 in a triaged system, so is the variance of the error.

In the analysis below, we determine the arrival rate λ and the technology investment θ that maximize firm profits per unit time for each process design. If customers have a relatively higher value of c_D , then we call them *time-sensitive*; if the value of c_Q is relatively higher, we call them *quality-sensitive*.

Strategies That Maximize Single-Period Profits

Consider the case where a service comprises tasks that are knowledge-intensive (β is low) and the returns from information systems that compensate for skill attrition are not high (β_1 is low, ε_1 is low). We analyzed the profit functions for a wide range of parameter values; sample results are shown in figure 4. This figure, and subsequent ones, are samples of our extensive numerical optimization using the closed-form expressions derived above. Our findings were consistent over a wide range of values and may be summarized as follows:

Proposition 3: When a service is knowledge-intensive and a firm's returns from skill-enhancing information systems are low, then a system with separate tasks and centralized decision making is optimal. Decentralization is rarely optimal because the benefits from higher quality are dominated by the cost of information-sharing technology and the revenue lost because of reduced service throughput.

This result was invariant over a fairly wide range of quality and delay costs. It reflects the fact that decentralizing decisions that require knowledge in more than one area (which covers most product customization decisions) are very difficult in a functional

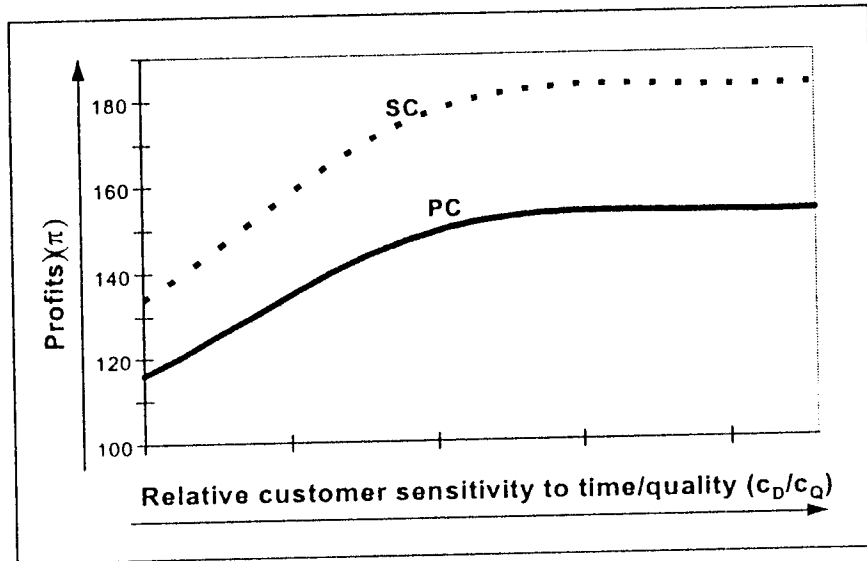


Figure 4. Firm Profits with High Knowledge-Intensity and Low Returns from IT (SC = sequential tasks, centralized decisions; PC = parallel units with consolidated tasks and shared queue, centralized decisions)

organization. There are also services and firms for which the “universal solution” of decentralized decision making by multifunctional workers is not the optimal choice.

This result also illustrates the contradictory nature of some of the textbook suggestions for reengineering. A person with a high level of knowledge related to a particular task is in a better position to interpret customer specifications related to that task, and therefore she or he will do a better job if given decision authority. Besides, customers of services that comprise tasks that are knowledge-intensive are more likely to place a high premium on their specifications being met. However, in such services, although consolidating tasks may improve decision making, it also increases the cycle time of the process significantly. Therefore, we have cases in which task consolidation and decentralization produce *opposite* effects on process performance, rather than *complementing* one another.

There are a large number of services in which technological support can compensate for skill attrition from consolidation (β_1 and ϵ_1 are not low). This is characteristic of services that are not so knowledge-intensive (β_0 and ϵ_0 are not low), and for which it is not difficult to acquire task knowledge through systems support. We investigated systems with these parameters. Our results, summarized in proposition 4, indicate that task consolidation is the ideal solution. Whether queues should be consolidated, and decisions decentralized, however, depends on other process parameters.

Proposition 4: When the returns from skill-enhancing information systems are moderate or high, then task consolidation is optimal.

- a. *When customers are time-sensitive, centralized decision making is optimal.*
- *If specification variability is high, triaged systems are ideal.*

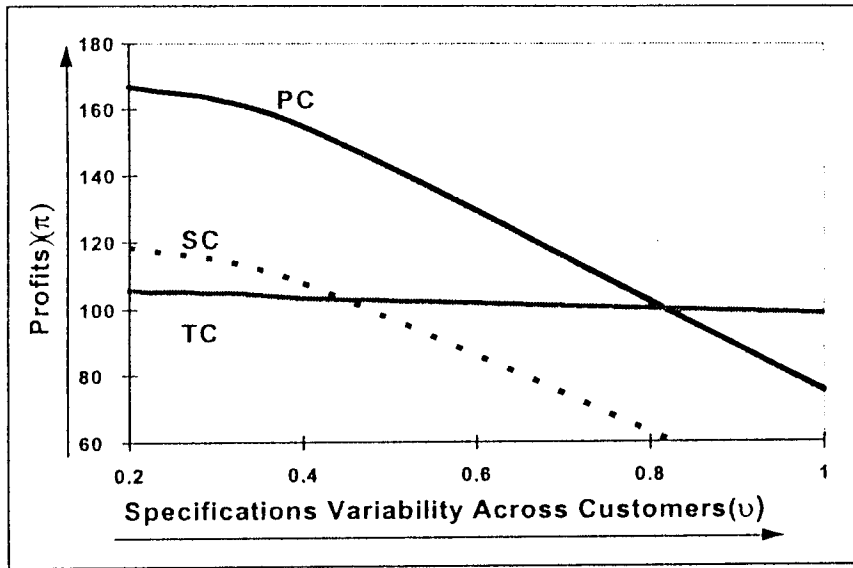


Figure 5. Firm Profits in a Time-Sensitive Market as a Function of Process Design and Customer-to-Customer Variability (SC = sequential tasks, centralized decisions; PC = parallel units with consolidated tasks and shared queue, centralized decisions; TC = triaged units with consolidated tasks and separate queues, centralized decisions)

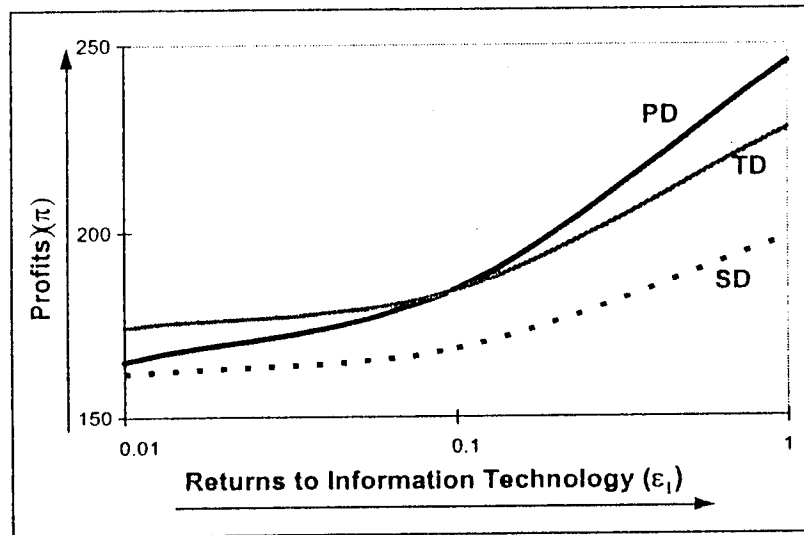


Figure 6. Firm Profits in a Quality-Sensitive Market as a Function of Process Design and Returns to IT (SC = sequential tasks, decentralized decisions; PD = parallel units with consolidated tasks and shared queue, centralized decisions; TD = triaged units with consolidated tasks and separate queues, centralized decisions)

- If specification variability is low, shared queues are optimal.
- b. When customers are quality-sensitive, decentralized decision making is optimal.
- If the returns from information systems that provide specifications interpretation

support are high, then shared queues are optimal.

- *If the returns from these systems are relatively lower, then triaged systems are optimal.*

Figures 5 and 6 show some of the results from our analyses in this region. With only low to moderate skill attrition from consolidation, we expect the optimal process design to be nonsequential. Interestingly, however, task consolidation and decentralization do not always accompany one another. As proposition 4a shows, the sensitivity of customers to delays results in centralized decision making to reduce cycle time. We also see that, in the presence of high customer variability, a triaged work system that reduces quality deviation costs can be preferable to a shared queue system that minimizes queuing delays. This indicates that triage provides a good balance between reducing delays and reducing quality costs, if the customers value time more.

On the other hand, when customers are quality-sensitive, decentralization accompanies task consolidation. The gains from reduced delays can be sacrificed for better quality through triage when ε_1 is low (i.e., when a relatively high technology level is required to reduce the quality of decision making in a consolidated system). The tradeoff is increased delays; however, the quality sensitivity of the customers makes the triaged system ideal in the face of high specifications variability.

Strategies That Maximize Market Share

We have examined situations where a firm desires maximum single-period profits. Yet this may not be the only important criterion; an equally important factor for competitiveness is market share. In our model, the value of λ that maximizes profits can be taken as a measure of the share of the market that a firm captures (the higher the throughput, the larger the number of customers served). Figures 7 and 8 plot the corresponding profit-maximizing values of workload for the different process designs. This approach leads to our next proposition:

Proposition 5: If the objective of a firm is to capture market share, task consolidation is optimal.

a. If the market is uniform in terms of time preferences, then a shared queue system is likely to make a firm more competitive. Centralized decision making tends to be better in this scenario. Short-term profits, however, are lower if the market is quality-sensitive.

b. If the market is nonuniform in terms of time preferences, then it is likely that a triaged system will dominate.

We find that task consolidation and centralized decision making form the ideal strategy for capturing market share in a variety of situations. In a time-sensitive market, the profit-maximizing strategy also maximizes market share, except when variability is high. The triaged process design improves significantly in terms of market share at higher values of variability. Hence, the results in figures 4 and 6 indicate that the

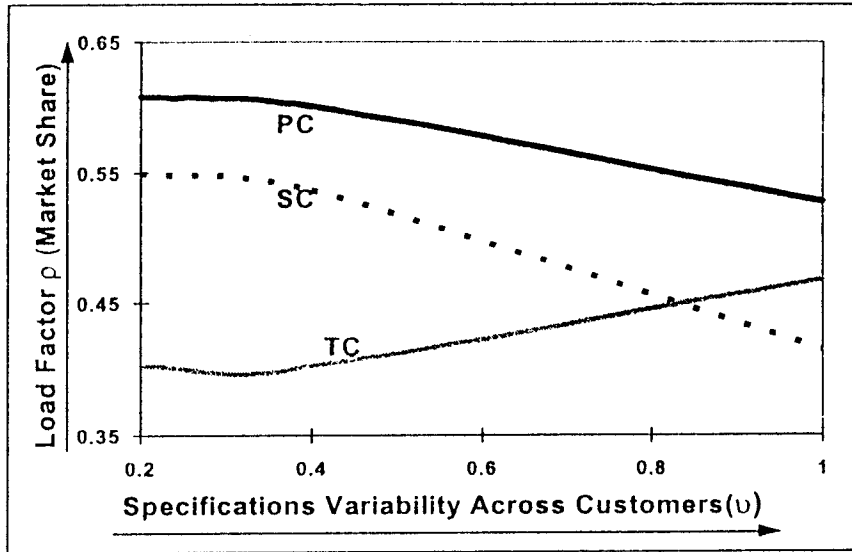


Figure 7. Market Share in a Time-Sensitive Market as a Function of Process Design and Customer-to-Customer Variability (SC = sequential tasks, centralized decisions; PC = parallel units with consolidated tasks and shared queue, centralized decisions; TC = triaged units with consolidated tasks and separate queues, centralized decisions)

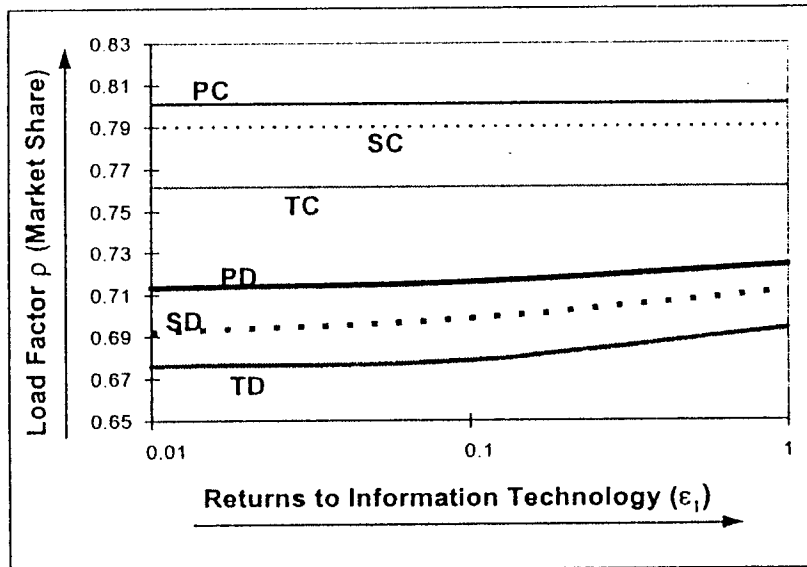


Figure 8. Market Share in a Quality-Sensitive Market as a Function of Process Design and Returns to IT (SD = sequential tasks, decentralized decisions; PD = parallel units with consolidated tasks and shared queue, centralized decisions; TD = triaged units with consolidated tasks and separate queues, centralized decisions)

Decision Rights	Decentralized	<ul style="list-style-type: none"> • Unlikely • Possible if different decisions are independent/ task specific 	<ul style="list-style-type: none"> • Works well in a quality sensitive market • Low technology returns => triaged queues • High technology returns => shared queue • Short term profit maximizing and long term market capturing strategies can conflict
	Centralized	<ul style="list-style-type: none"> • Good for jobs with <ul style="list-style-type: none"> - high knowledge intensity - low technology returns - low spec. variability • Can potentially be a good market capturing strategy 	<ul style="list-style-type: none"> • Works well in a time sensitive market • Low spec. variability => shared queue • High spec. variability => triaged queues • Short term profit maximizing and long term market capturing strategies are aligned
		Separated	Consolidated

Figure 9. Strategic and Competitive Implications of Different Levels of Empowerment

strategy. Figure 7 shows, however, that the decentralized process designs that were optimal for profit maximization did not yield market share maxima; they were dominated by centralized solutions. The competitive implications of this are clear: In a market where customers are quality conscious, firms may sacrifice short-term profits for long-term market share, or vice versa.

Conclusions

WE HAVE INVESTIGATED THE COMPETITIVE AND ECONOMIC IMPLICATIONS of information technology, decentralizing decision rights, and consolidating tasks during process reengineering. Figure 9 summarizes our findings. In a time-based competitive marketplace, centralized decision making and bundled tasks are clearly superior to other process designs; hence, information technology investments should be directed at increasing skill levels (so that loss of expertise from task consolidation is minimized) and improving communications between the line workers and management (so that the centralized decisions are better informed).

As indicated in the lower-right quadrant of figure 9, the work system (shared versus separate queues) depends on the variability in customer demands. On the other hand, a firm facing competition on quality may lose out on market share if it focuses on profit maximization in the short run. It may be necessary to decentralize decision rights

if the variability of customer specifications is high; however, care must be taken that this does not reduce the firm's workload too much. Information technology that enhances information access and sharing is a good choice here. Decision-support and other such systems that minimize information losses and provide timely access to diverse information are likely to be adopted by firms in such markets.

Our ongoing research examines the different process designs that are optimal for a firm facing heterogeneous time preferences in a variety of competitive situations. We are also studying the optimal incentive contracts that a firm needs to implement as a function of the market preferences for timeliness and quality. Our preliminary results have indicated that, independent of the level of information technology, decentralization must be accompanied by performance-based incentives. The choice of performance measures, however, depends on the nature of the customers and the competitive offerings available in the marketplace.

NOTES

Acknowledgments: The authors thank Eric Clemons for his extensive and detailed feedback on an earlier draft of this paper, Bruce Weber and the participants of the Strategic and Competitive Information Systems Mini-track at the Thirtieth Hawaii International Conference on System Sciences for their comments on an earlier version of this paper, and Marshall Freimer for his careful reading of our analysis and proofs.

1. By triage, we mean the separation of customers or work based on a particular distinguishing criterion, such as variability in service time, or customer specifications. This differs somewhat from the usage of the term in emergency rooms; however, it is consistent with how it is used by Hammer and Champy [12].

2. This could be a reflection of the nature of the information system. One can assume that worker 2 is just required to enter a few details about the specifications information and an expert system generates a report understandable to worker 1.

3. One may wonder why the reduction is the same variable αv . Arguably, when much of the information processing has been done by an information system and by worker 2, the value of this parameter should be lower. However, the specifications are not in worker 1's area of specialty, and this can increase the loss in productivity. We assume these effects cancel out.

4. The distribution of the sum of two exponential random variables with parameters μ has a mean $2/\mu$; its variance, however, is lower than that of a random variable with parameter $\mu/2$.

REFERENCES

-
1. Barua, A.; Lee, S.; and Whinston, A.B. The calculus of reengineering. *Information Systems Research*, 70, 4 (December 1996), 409–428.
 2. Brickley, J.; Smith, C.W., Jr.; and Zimmerman, J.L. *Managerial Economics and Organizational Architecture*. Burr Ridge, IL: Richard D. Irwin, 1997.
 3. Buzacott, J. Commonalities in reengineered business processes: models and issues. *Management Science*, 42 (1996), 768–782.
 4. Carter, M. Information and the division of labour: implications for the firm's choice of organisation. *Economic Journal* (March 1995), 385–397.
 5. Clemons, E.K.; Thatcher, M.E.; and Row, M.C. Identifying sources of reengineering failures: a study of the behavioral factors contributing to reengineering risk. *Journal of Management Information Systems*, 12, 2 (Fall 1995), 9–36.
 6. Clemons, E.K., and Weber, B.W. Segmentation, differentiation, and flexible pricing: experiences with information technology and segment-tailored strategies. *Journal of Management Information Systems*, 11, 2 (Fall 1994), 9–36.

7. Davenport, T.H. *Process Innovation*. Boston: Harvard Business School Press, 1993.
8. Davenport, T.H., and Short, J.E. The new industrial engineering: information technology and business process redesign. *Sloan Management Review* (Summer 1990), 11–27.
9. Earl, M.J.; Sampler, J.L.; and Short, J.E. Strategies for business process reengineering: evidence from field studies. *Journal of Management Information Systems*, 12, 1 (Summer 1995), 31–56.
10. Gurbaxani, V., and Whang, S. The impact of information systems on organizations and markets. *Communications of the ACM* (January 1991), 59–73.
11. Hammer, M. Re-engineering work: don't automate, obliterate. *Harvard Business Review* (July–August 1990), 104–112.
12. Hammer, M., and Champy, J. *Reengineering the Corporation: A Manifesto for Business Revolution*. New York: Harper Business, 1993.
13. Huber, G.P. The nature and design of post-industrial organizations. *Management Science*, 30 (1984), 928–951.
14. Markus, M.L., and Robey, D. Information technology and organizational change: causal structure in theory and research. *Management Science*, 34 (1988), 583–598.
15. Marschak, J., and Radner, R. *Economic Theory of Teams*. New Haven, CT: Yale University Press, 1972.
16. Radner, R. The organization of decentralized information processing. *Econometrica* (September 1993), 1109–1146.
17. Sampler, J.L., and Short, J.E. An examination of information technology's impact on the value of information and expertise: implications for organizational change. *Journal of Management Information Systems*, 11, 2 (Fall 1994), 59–73.
18. Seidmann, A., and Sundararajan, A. Business process redesign: case studies and operational models. *Proceedings of the Second Manufacturing and Service Operations Management Conference*, 1996, pp. 270–275.
19. Seidmann, A., and Sundararajan, A. Information systems, incentives and workflow logic: strategic implications for reengineering business processes. In J.I. De Gross, S. Jarvenpaa, and A. Srinivasan (eds.), *Proceedings of the Seventeenth International Conference on Information Systems*, December 1996, pp. 400–420.
20. Seidmann, A., and Sundararajan, A. The effects of task and information asymmetry on business process redesign. *International Journal of Production Economics*, 50 (1997), 117–128.
21. Stoddard, D., and Jarvenpaa, S.L. Business process redesign: tactics for managing radical change. *Journal of Management Information Systems*, 12, 1 (Summer 1995), 81–107.

APPENDIX

Proof of Proposition 1: The cost c_1 at the first stage is

$$c_1 = \frac{c_D}{\alpha\mu - \lambda} + c_Q (\gamma_1 - \delta)^2.$$

Let $\gamma_2 = x$. Then the cost at the second stage is

$$c_2 = \frac{c_D}{\mu - \lambda} + c_Q (x - \delta)^2.$$

The total c_Q related cost is

$$C_Q = c_Q [(\gamma_1 - \delta)^2 + (x - \delta)^2 - \frac{(\gamma_1 - x)^2}{2}].$$

Therefore, the expected cost is:

$$\begin{aligned}
EC_Q &= \frac{c_Q}{2v} \int_{-v}^v \left[(\delta - x)^2 + (\gamma_1 - \delta)^2 - \frac{(\gamma_1 - \delta)^2}{2} \right] dx \\
&= c_Q (\gamma_1 - \delta)^2 + \frac{c_Q}{2v} \left[\frac{(\delta + v)^3 - (\delta - v)^3}{3} + \frac{(\gamma_1 + v)^3 - (\gamma_1 - v)^3}{6} \right] \\
&= c_Q (\gamma_1 - \delta)^2 + \frac{c_Q}{2v} \left[\frac{6\delta^2 v + 2v^3}{3} - \frac{6\gamma_1^2 v + 2v^3}{6} \right] \\
&= c_Q (\gamma_1 - \delta)^2 + c_Q \left[\delta^2 - \frac{\gamma_1^2}{2} + \frac{v^2}{6} \right].
\end{aligned}$$

The optimal choice of d solves $\min EC_Q$, or

$$\begin{aligned}
&= \\
&= \frac{d}{d\delta} c_Q \left[(\gamma_1 - \delta)^2 + \delta^2 - \frac{\gamma_1^2}{2} + \frac{v^2}{6} \right] = 0 \\
&\Rightarrow -2c_Q (\gamma_1 - \delta) + 2c_Q \delta = 0,
\end{aligned}$$

which yields $\delta = 0.5\gamma_1$. Therefore, the expected total cost is

$$\frac{c_D}{\mu - \lambda} + \frac{c_D}{\alpha\mu - \lambda} + c_Q \left[\left(\gamma_1 - \frac{\gamma_1}{2} \right)^2 + \left(\frac{\gamma_1}{2} \right)^2 - \frac{\gamma_1^2}{2} + \frac{v^2}{6} \right] = \frac{c_D}{\mu - \lambda} + \frac{c_D}{\alpha\mu - \lambda} + c_Q \frac{v^2}{6}.$$

The result follows so long as

$$v^2 < \frac{6c_D(1 - \alpha)}{c_Q(\mu - \lambda)(\alpha\mu - \lambda)}.$$