Trends in Business Organization: Do Participation and Cooperation Increase Competitiveness?

International Workshop

Edited by Horst Siebert

Institut für Weltwirtschaft an der Universität Kiel

J.C.B. MOHR (PAUL SIEBECK) TÜBINGEN

Paul Milgrom and John Roberts

Continuous Adjustment and Fundamental Change in Business Strategy and Organization

1. Introduction

Worldwide, firms are attempting major shifts in their strategies and restructurings of their organizations. The most recent wave of such fundamental change began in the United States, Canada, and the United Kingdom during the 1980s, in response to increased international competition and an especially active market for corporate control. Firms restructured to focus more on "core" operations by divesting unrelated operations and businesses and to reduce costs by "downsizing" and flattening their organizational hierarchies (Lichtenberg, 1991). During this period, too, many manufacturing enterprises began attempts to move from a Fordist, mass-production pattern to "modern manufacturing" (Milgrom and Roberts, 1990), introducing high-commitment human resource and supplier relations policies, flexible manufacturing methods, increased quality and response speed, and *kaizen* (continuous improvement). Over the last five years, firms in the transition economies of Eastern Europe and the former Soviet Union have been going through massive changes as they are privatized and then seek to adapt to commercial operations in market economies. Now, in Western Europe and Japan, the pressures of lingering recessions combined with toughened U.S. competition, the emergence of low-cost competitors in Eastern Europe and East Asia, and basic changes in their demographic and economic environments are forcing firms to consider wrenching changes in their strategies and organizational structures. And, with encouragement from international banking organizations, developing countries are privatizing state-owned enterprises, increasing productivity but also losing jobs in the process.

These fundamental changes are hard. The limited evidence available suggests that downsizing often fails to bring the hoped-for gains in productivity and profitability.[1] Even when they succeed, as in the case of General Electric's

1 For example, according to a 1993 Wyatt study reported by Byrne (1994), "fewer than half the companies that downsize meet the financial and operational goals they set for themselves."

refocusing and downsizing under Jack Welch, they bring pain: Welch was dubbed "Neutron Jack" because, like the neutron bomb, he got rid of all the people but left the buildings standing. Firms sometimes avoid undertaking radical changes, even when they seem to be obviously warranted. For example, General Motors president Jack Smith recently described his firm as having been "in denial" when, during the 1980s, it failed to move towards the patterns of "lean manufacturing" that Toyota had shown to be more effective and that Ford and Chrysler were embracing. When fundamental changes are attempted, it is often only in the context of a crisis that threatens the firm's continued existence: this was the case at both Ford and Chrysler.

The difficulties of wholesale change have led some to consider whether slower, partial adjustment, perhaps on an ongoing basis, might be a viable alternative model for adapting strategy and structure to a changing environment. For example, the debate on how best to manage the transition from socialism basically poses these two as substitutes. Indeed, the title that the organizers of the workshop suggested for this paper reflects this view of the two sorts of changes as substitutes: they had "Partial Adjustment *or* Fundamental Change," where we have written "Continuous Adjustment *and* Fundamental Change." One of the purposes of this paper is to show that continuous adjustment and fundamental change are not always substitutes; they can often be complementary. For without the organizational capability to refine and enhance a new system quickly, fundamental changes are hardly likely to succeed.

A central element of our formal analysis of change is that the problem of choosing among strategies and organizational structures is often marked by significant nonconvexities. Correspondingly, there can be multiple *local optima* —values of the choice variables that yield better performance than any other "nearby," similar choices—and multiple *team equilibria*—values for the choice variables which are *coherent* in the sense that no change, however large, in just the (proper) subset of all the firm's decisions that are controlled by an individual manager improves overall company performance. There may even be multiple *decentralized optima*—points from which there is no beneficial change that does not involve changing *all* of the firm's decision variables. Further, the local optima can coincide with team equilibria and can involve quite distinct patterns of strategy and organization, and mixing elements of two such patterns or adopting values for the choice variables that are intermediate between two coherent patterns need not yield coherence or local optimality. Thus, the problem of finding an effective organizational strategy can be conceived of as consisting of two parts: deciding which hill to climb and then climbing it as efficiently as possible.

Which hill to climb—which strategy is best—depends on a variety of environmental factors involving the state of technology, the regulatory environment, firm objectives, and input and final market conditions. This leads to another of our conclusions: there is no universally best strategy and structure.

Consider the example of the automobile industry as sketched in *The Machine that Changed the World* (Womack et al., 1990). The authors of the book describe lean manufacturing as simply being superior to mass production, just as mass production was superior to the craft production system. However, the details of their report seem to tell a different story.

A hundred years ago, automobile manufacturing was a craft industry. Production was in lot sizes of one unit, made to the individual customer's specific order placed directly with the manufacturer. Capital investment was low, and the labor force consisted of highly skilled craftsmen. Beginning with Ford, mass production displaced craft production. The internal logic of the new system, based on the transfer line, interchangeable parts, and the exploitation of economies of scale, was strikingly different from that of craft production. More recently, another new paradigm has emerged in the form of lean and flexible manufacturing as first developed by Toyota. This paradigm, which accentuates flexibility, speed of response, and economies of scope, differs from mass production on a host of identifiable dimensions involving human resource practices, product development strategies, manufacturing methods, supplier and customer relations, and management methods (see Table 1).

Each of these three patterns is distinct from the others and represents a coherent pattern of strategic and organizational decisions. Moreover, *each can be the optimal choice in particular conditions.* Mass production is well suited to situations where educated, disciplined workers are in short supply but unskilled labor is easily available, where customers do not place a high value on variety but are price sensitive, or where input suppliers are operating under the mass production model and cannot easily respond to demands for flexibility that characterize a customer using the modern manufacturing strategy. Arguably, these characteristics described the environment in which the auto industry operated over much of this century, although they probably do not obtain now in the developed countries. Thus, while lean and flexible manufacturing may have become a more rewarding strategy under current conditions, it is not best for all conditions, even in automobile production. Under the conditions prevailing when the automobile was first invented, for example, the craft production system was efficient. Until advances in metallurgy and manufacturing methods allowed the idea of interchangeable parts to become a reality, crafting parts to fit on each car was unavoidable. Craft production remains effective even today in the super-luxury segment of the auto industry, where low volumes make the alternative approaches too costly. Moreover, once one goes beyond the auto industry it is clear that different production systems may be appropriate in different circumstances. For example, *haute couture* is a craft industry; low-fashion clothing is mass-produced; and Benetton and the Gap operate on something approaching the flexible manufacturing model in designing, producing, and marketing moderately priced fashions.

Table 1 — Mass Production versus Lean Manufacturing

DIMENSIONS	MASS PRODUCTION: *Transfer lines, specialization, and economies of scale*	LEAN MANUFACTURING: *Flexibility, rapid responses, and economies of scope*
Capital equipment	Specialized	Flexible (low setup costs)
Production runs	Long runs, large lot sizes	Short runs, small lot sizes
Product changes	Infrequent	Frequent
Markets	Mass markets	Targeted markets
Worker skills	Low or specialized	High, with cross-training
Decision making	Central expertise and hierarchic planning	Local information and self-regulation
Communications	Primarily vertical	Primarily horizontal
Product development	Sequential	Cross-functional teams
Operational focus	Static optimization	Continuous improvement
Day-to-day emphasis	Accent on volume	Accent on quality
Inventories	High	Low
Managing uncertainty	Supply management	Demand management
Customer relations	Make-to-stock, limited communications	Make-to-order, extensive communications
Supplier relations	Short-term, price-based	Long-term, competency-based
Vertical integration	High	Low
Employee relations: Factory workers	Low commitment, confrontational	High commitment, cooperative

Many familiar examples support the idea that successful organizations adopt one of a limited number of distinct organizational and strategic patterns. Lincoln Electric Company, the world leader in the manufacture and sales of arc-welding equipment, has adopted a highly distinctive set of organizational and strategic choices. The firm's record of productivity and profitability for over half a century from the 1930s through the 1980s is legend. In the late 1980s, however, Lincoln acquired overseas production facilities, including several existing plants, to assure its access to markets outside North America. It has had difficulty installing the full range of its distinctive policies and processes at these sites. Caught in the middle with a mix of Lincoln's distinctive practices and traditional, local ones, many of these operations are as yet unsuccessful, and their poor performance actually led the company to incur its first loss since

the Great Depression. Additional examples of systemic change involving a large number of organizational attributes may be found in Chandler's (1962) account of how Standard Oil, DuPont, and Sears Roebuck had to restructure themselves fundamentally to manage their increasingly diverse businesses in the period after World War I, giving birth to the multidivisional form. The remarkable similarities among these independently designed organizations is evidence of the deep forces that account for organizational discreteness.[2]

The auto industry example also illustrates the claim that attempts to mix elements from distinct patterns may result in a strategy or organization that is not viable. Consider an example we will examine later in more detail: the flexibility of the production system and the breadth of the product line. Under mass production, both of these variables are set at low levels, while both are set at high levels under lean, flexible manufacturing: Ford produced a single model of car, in a single color, in a plant that was completely specialized to that task, while Toyota's Tamigo engine plant continually produces hundreds of variants of engine-transmission combinations on a single, highly flexible production line. Each of these alternatives represents a coherent pattern. Yet trying to match a broad product line with inflexible equipment or flexible equipment with a narrow product line does not make sense, and at best would result in much worse performance than either coherent pattern.

When the firm's strategic and organizational choice problem is not concave, and in particular when it has the form suggested by these examples, then changes in the environment that yield little or no change in the local optima or team equilibria can nevertheless markedly alter the comparative advantages of the different strategies. Small-scale adjustment on an ongoing basis can allow the firm to adapt to the same changes, and this can often be successfully done on a decentralized basis with no explicit coordination among the managers. We shall argue that such adjustments can often also lead the firm toward a local optimum even when managers lack a priori knowledge of where the optimum lies. But when there is a shift in the ranking of the different coherent patterns, substantial strategic and organizational changes are warranted, and such changes cannot be accomplished without a coordinated search over a wide terrain.

Of course, even if the firm's problem were concave, as economic analyses so often assume it to be, discontinuous environmental change or past failures to adapt could leave the firm severely maladapted to its current environment, so

2 Other economists have argued similarly that organizational and strategy choices are discrete: Williamson (1975, 1985) poses the multidivisional, M-form, and functional organization as distinct alternatives, and Porter (1980) asserts that there is only a small number of effective generic strategies, with hybrid strategies doomed to failure.

that substantial changes would be warranted. One might think that since gradual adjustment and large-scale change are alternative ways to make parts of the adjustment, the two capabilities are substitutes in the organization's problem. However, close analysis reveals that this intuitive conclusion relies on severely restrictive assumptions. Only in the case of one-dimensional, concave maximization problems do we find that an increase in the firm's ability to make an initial leap toward the new optimum makes it less willing to invest in the capability to make a series of additional, incremental improvements. For both nonconcave problems and problems with at least two dimensions, the reverse conclusion may be true. In nonconcave problems—even in one dimension—a firm with the ability to leap a long way toward the optimum may wish to invest more in a capacity for continuous adjustments that allow it to complete the transition. In contrast, a firm without such a capability is more likely to forego or delay the large adjustment, reducing its incentive to invest in a capacity for additional gradual adjustments. For such cases, the two capabilities are complementary. In higher dimensions, complementarity among the capabilities is even consistent with concave production functions. An ability to make a leap in one dimension is made more valuable by the ability to adapt quickly and easily on complementary dimensions. For example, the ability to lead the market in introducing radically new products is most profitable for a firm that can also quickly modify its production facilities to make the product at low cost. Consequently, regardless of the concavity of the objective, an increased capacity to make large changes on one dimension leads an optimizing firm to increase its investment in the capacity to make complementary changes in other dimensions.

In general, the formal analysis of unstructured, nonconcave maximization problems is extremely difficult. Much of our formal analysis is rooted in the assumption that the firm's problem does have a special structure, namely, that there are complementarities among key elements of the firm's strategy and structure. We have argued elsewhere (Milgrom and Roberts, 1990a, 1992, 1994a, 1994c) that pervasive complementarity relations are a feature of many actual situations. Their presence provides important structure while still being consistent with the firm's problem displaying the nonconcavities we have discussed already. In the next section, we will elaborate on the nature of the structure imposed by complementarity and establish some of its implications. In particular, we show how the firm's search problem of locating a more profitable arrangement is immensely simplified when its choice variables are complementary. As well, complementarity between the environmental variables and the elements of the firm's choice results in unambiguous comparative statics analysis, allowing us to determine the qualitative direction of the optimal response by the firm to environmental changes.

Returning to the automobile example, one of the striking features that emerges from examining the historical development of the lean production model is that it did not appear fully formed but rather developed slowly, by accretion (Toyoda, 1987; Cusumano, 1985). Toyota's leaders in the early postwar period saw that the market conditions it faced and the scale advantages of American manufacturers would not let it succeed using the mass production model, and they then developed the new model over time. The story at Lincoln Electric is similar: the various characteristic features were not adopted all at once, but rather over the course of decades. These observations might seem to cast doubt on our hypothesis that organizational alternatives are discrete. We suggest, however, that they can also be viewed as incremental movements that started from fundamentally different initial conditions than those leading to more standard mass production patterns and that were guided by a different logic and a different vision of the ultimate model. Toyota, for example, began as a manufacturer of textile equipment and entered on a small scale into an automobile industry already dominated by giants and into a local market that, while small, already demanded a variety of products. The initial development of the Toyota system owes much to that heritage. Further, in a context where the firm's choices feed back and affect the evolution of its environment, as in the Toyota example, the pursuit of the different model (modern, lean, and flexible manufacturing) can shape the environment in ways that increase the effectiveness of that model relative to the alternatives. Feedback effects of this sort add momentum to the changes taking place in the firm, leading it to develop an increasingly coherent and refined set of policies (see Milgrom et al., 1991).

Some of our analysis is based on team-theoretic approaches, where the participants in the organization share goals and there is no conflict about objectives. Even here, one can identify reasons why large-scale change should be difficult to achieve successfully. These are investigated in Section 2. No less important as barriers to change are intraorganizational conflicts of interest; people may be expected to resist changes that imperil any rents they are enjoying. One way to avoid this resistance is to avoid entrenching employees in positions that generate rents. Regular job rotations for managers, pay policies that favor equity, and employment guarantees are all policies that can reduce resistance to change. These policies, however, are not without costs. A formal analysis of some of the sources of resistance and policies that alleviate the problem is contained in Section 3. First, though, we turn to a closer study of some of our ideas concerning nonconvexities, complementarities, and the problem of coordinating change.

2. Change in a Team Setting

In this section, we adopt a "team" approach—one that abstracts from the conflicting objectives of the managers and the firm and assumes that all share the common objective of maximizing the firm's profit. Still, there may be a problem of planning and coordinating the team members' decisions.

a. Hill Climbing

To set a formal context for our questions about the need for "fundamental change" and the role of centralized coordination in planning for and implementing change, we begin with a neoclassical model of a firm attempting to adapt to its changing environment. The firm has a production function, say $f(x_I, x_O, \varepsilon)$, where x_I is a vector of purchased inputs, x_O is a vector of other decisions taken by the firm, and ε is a parameter characterizing the environment. If the prevailing input price vector is p and the output price is normalized to unity, then the firm's profits are $\pi(x_I, x_O, \varepsilon, p) = f(x_I, x_O, \varepsilon) - p \cdot x_I$. The issue to be studied is how, and how well, the firm adapts when the environment, ε and p, changes.

We can either assume that the decisions in the firm are centralized or, alternatively, that decisions are controlled by several managers making their choices independently. In either case, suppose that decisions are adjusted in a direction that increases profits. In the first case, this might involve a coordinated, local change on multiple dimensions, while in the latter we think of each individual manager as adjusting his or her choice variables without knowledge of the adjustments that others may be undertaking. Models of this sort are sometimes called "hill-climbing" models, because if one visualizes the firm's choices as lying on a plane and the profits corresponding to any point as the altitude of the terrain, then the ongoing changes always are ones that move the firm uphill, to a point of higher profits. Such a process continues indefinitely upward or until it reaches a local profit maximum. Plausible restrictions on the hill-climbing process ensure that the process eventually approaches a local peak of the profit terrain. Also, with the "standard" assumptions that the profit function π is smooth, bounded, and strictly concave, the only local maximum is also the global profit maximum. With these standard assumptions, it follows that hill-climbing processes, subject to plausible restrictions, lead the firm to its global profit maximum. They may achieve this faster and more directly when the choices are coordinated at each step, but even the decentralized, uncoordinated version reaches the global optimum.

The hill-climbing process that characterizes the firm's behavior does not need to be sophisticated. The process will converge to the maximum even if the

changes selected by the procedure are not far-sighted, that is, even if they do not always move the firm closer (in Euclidean distance) to the actual profit-maximizing combination. There is convergence even if the movements involved are small and the optimum far away and even if the various decisions are made separately and sequentially, with each x_i adjusted by a separate decision maker without regard to the changes to be made later by others. Quite simply, when one is climbing a smooth hill with a single peak, heading always upward and moving at a decent pace, one eventually approaches the top, though the route may be circuitous.

Starting from any initial environment, ε_0 and p_0, suppose that the environment changes to ε_1 and p_1, and that the managers in charge of the various decisions adapt in a continuing attempt to improve profits, but without any explicit coordination. That is a hill-climbing procedure, so, under the "standard" assumptions, the eventual outcome will be a profit maximum. As emphasized in the Introduction, it does not even matter whether the managers know ε_1, that is, whether they know what function they are trying to maximize. All that each manager needs to know is whether there is some local change the manager can make on his or her own to increase the firm's profit.

Here, however, the standard smoothness assumption plays a key role. For, suppose the firm faces a Leontief production function $[\min(x, y)]^{1/2}$ where x and y are costly inputs controlled by different managers. Then many points with $x = y$ can be team equilibria, and uncoordinated hill-climbing will not find an optimum, although coordinated moves that keep x equal to y will succeed.

Concavity has a second implication, too, which was highlighted in the Introduction. When the profit function, π, is concave and the firm's choice can be modeled as being one-dimensional, the ability of a firm to take a large step in the direction of the optimum tends to reduce the value of being able to improve rapidly by taking a series of small uphill steps. They are substitutes because they are alternative ways of approaching the peak, and a firm with an ability to make large leaps will tend to invest less in its ability to make quick small adjustments. Here is a simple formal statement of this idea.

Suppose that by investing I in systems flexibility, the firm can arrange to face an adjustment cost of $c(I)$ per unit of change. Suppose also that the profit-maximizing decision is x^*, that the firm's initial position is x_0, and that the firm's ability to make a "fundamental change" is measured by the maximum distance, A, that it can costlessly jump in the direction of the optimum, x^*. It is natural to assume that the function $c(I)$ is decreasing: one can reduce adjustment costs by investing more in flexibility. Let us also suppose that $x_0 + A \le x^*$, so that increases in A represent strict increases in the firm's ability to jump toward the optimum. Then the firm's profit function is $f(x_0 + A + y) - c(I)|y| - I$; this is the profit rate after the jump of A and the adjustments, y, minus the cost of the adjustments, $c(I)|y|$, minus the investment cost, I. Given our assumptions that $x_0 + A \leq x^*$ and that f is concave, the firm can never benefit by setting $y < 0$, so we can fix $y \geq 0$ and substitute y for $|y|$ without loss of generality. Regarding this whole expression as a function $\pi(y, I, A)$ and using the assumption that f is smooth and concave, one can check directly that the mixed partial derivatives satisfy $\pi_{y,I} \geq 0, \pi_{yA} \leq 0$, and $\pi_{I,A} = 0$. Finally, we assume that I and y are constrained to lie in compact sets.[3] The comparative statics conclusion is then as indicated: As A increases, the optimal level of investment, I^*, falls and with it falls the optimal amount adjustment, y^*, to be made. In that sense, fundamental change and partial adjustment are substitutes. As accentuated in the Introduction, however, this analysis does hinge crucially on the assumption that f is concave[4] and that its domain is one-dimensional.

In general, when f is convex over the relevant portion of the (one-dimensional) domain, a greater ability, A, to make fundamental adjustments will increase the benefit the firm enjoys from its investment in the ability to make adjustments. A firm that can quickly climb whatever local hill it is on may find it more attractive to be an early adopter of a significant new technology, counting on its rapid adjustment capability to make the early leap to the new hill worthwhile. In that case, a firm with the capability of making large "fundamental" changes will be inclined to invest more in developing its ability to make small adjustments than a firm without that capability.[5]

The neoclassical model of the firm with a convex technology has earned its popularity *because* it identifies a set of cases in which it can be proved that decentralized decisions, guided only by prices and local technical information, are logically compatible with attaining efficient outcomes. Nevertheless, the model differs dramatically from most people's understanding of the actual environment of business. Indeed, since Adam Smith's account of the workings of the pin factory, economists have regularly emphasized the practical importance of the economies of specialization and scale. These observations are inconsistent with the hypothesis of a convex technology or concave production function.

3 One may add a constraint in which y is bounded above by any nondecreasing function of I without affecting the conclusions.

4 Milgrom and Roberts (1994b) show that the condition that $\pi_{yA} \geq 0$ is necessary as well as sufficient for the general validity of such conclusions; here it is equivalent to $f'' \leq 0$.

5 In the general case where f is neither concave nor convex, there is no general comparative statics conclusion. The conclusion depends on the particular form of the function $c(\cdot)$ and the particular constraints that apply to the choice of y. See Milgrom and Roberts (1994b).

For a more satisfactory account of the problem of change, therefore, we will need to introduce the possibility of nonconvexities in the firm's production set. Nonconvexities mean that the profit terrain may have many hills and valleys and many local peaks. Reaching the highest peak (the profit-maximizing point) from a different local peak by hill climbing is impossible, even when the adjustments are chosen in a coordinated fashion, because any route has to take the firm through valleys of low profits. Nonconvexities thus imply that adjustments may need to be guided by global considerations. They also create a role for *leaps* across the valleys, that is, for large-scale changes.

Nonconvexities do not automatically imply that changes have to be made by leaps, however. If the costs of adjustment are an increasing convex function of the step size, repeated partial adjustments are less directly costly than a single large change that arrives at the same endpoint. The offsetting drawback of gradual change is that it necessitates operating for more-or-less extended periods in the valleys, so that operating profits are lower than they were at the initial point, and that it further will require at least initially that the adjustments continually lead to ever-worsening performance. We will return to this issue below. Nor do nonconvexities alone imply any benefits from coordination between units. Suppose, for example, that profits in unit i are given by $(1/2)x_i + \sin x_i$ on some bounded range and that firm profits are the sum of individual unit profits. This firm's problem is not concave, and hill-climbing (whether decentralized or coordinated) will not solve it. Yet its separable nature means that it can be solved in a decentralized fashion, one variable at a time, with any improvement on one dimension necessarily improving overall performance. Coordinating change across numerous variables becomes valuable when the *interactions* among the decisions of the different managers are significant. Thus, *fundamental change*—large-scale adjustments on multiple dimensions made more-or-less simultaneously—becomes a major issue only when we combine nonconvexity with nonseparability.

Combining nonconvexity with nonseparability (and perhaps lack of smoothness as well) would seem to leave too little structure on the problem of choosing strategy and organization to hope to be able to say anything except that the problem is hard. However, there are in fact classes of problems with these properties that are both tractable and empirically relevant.

b. Design and Innovation Attributes

In general, as economists since Hayek (1945) have emphasized, central planners and head office executives suffer an important handicap in decision making compared to the managers in the factory or the field: the head office executives do not have direct, continuous access to information about which people and which methods have been most productive, what skills people have to perform new tasks, and so on. Yet, despite that informational disadvantage, there are times when central office involvement in decisions can lead to obvious improvements, either by eliminating losses associated with poor coordination between the departments or divisions or by developing important kinds of information that the managers in the factory or the field office do not have. Both of these are important features in fundamental change.

For example, when a firm plans for the introduction of a new product, the marketing, design, manufacturing, and distribution departments should agree on when the product will be introduced, which consumers to target, and an approximate sales forecast. If the design group targets the wrong group of customers or the factory produces more units than the firm can distribute, or if three departments are ready and waiting to get started while the fourth is still gearing up, losses will be suffered. Similarly, when designing a new car, the separate groups of engineers planning the chassis, brakes, power train, passenger environment, and other subsystems should have prior agreement on the approximate weight and size of the vehicle, as well as on the date of product introduction. A car whose engine cannot carry it up a steep hill, or whose brakes are unreliable for that weight, or whose air conditioning system cannot cool an adequate volume of air or requires too much engine power is hardly likely to be profitable.

As we have argued (Milgrom and Roberts, 1992, p. 91), these "design decisions," in which coordination is especially important, are identified by two key attributes. The first is that there is significant a priori information about how the pieces of the system fit together. For example, in the new product introduction example, we know in advance that the manufacturing and distribution systems should be ready to begin functioning at about the same time and the marketing campaign should begin just a bit earlier, even though we cannot know in advance exactly what those dates should be. The second is that errors of *fit* should be relatively costly compared to errors of design. Thus, if the planned new product introduction date is one month later than the optimal date, that would be less costly than if just the manufacturing facility is one month late, with all the other resources ready on time. Similarly, in designing a car, a mismatch between the components is likely to be a costlier error than having the whole car be slightly heavier or lighter than in the optimal design.

For design decisions, the benefits of central coordination typically outweigh the costs. For while the central plan—especially a highly detailed one—is likely to misspecify what the individual departments should do, it can at least ensure that the plan is coherent, that is, that all the pieces fit together. It can do that because, by definition, a design decision is one in which there is a priori information about what fits together. And while some losses are incurred because the central office decision makers are ill-informed about local circumstances, those costs are, by definition, smaller than the costs of the uncoordinated decisions that would arise in the absence of some central office involvement.

Of course, there are more ways to organize decisions than by uncoordinated decisions among independent decision makers or by planned decisions at the head office. For example, one popular alternative is the use of interdepartmental teams. Nevertheless, our discussion and analysis does make it clear that a design decision cannot be workably implemented by isolated decision makers using just local and price information to guide their decisions. There is significant value created by detailed communications about such matters as weights, sizes, dates, manufacturing tolerances, and so on—by explicit planning of decisions.

In addition to the attributes of design decisions, there is a second kind of decision attribute that favors head office decision making. These are what we have called *innovation attributes*. Innovation attributes arise when the payoffs to alternative strategies depend on information that would be unavailable even if the individual local managers could pool their information, often because they depend on interactions between different functions. For example, the advantages of flexible equipment for manufacturing the current range of products may be something that the manufacturing manager knows today from experiments with existing production facilities. Similarly, customer demand for a certain extended array of new products may be something the marketing manager knows from surveys or field trials. However, the optimal plan for the firm, which involves designing products that can be made effectively on available equipment and that meet customer needs, may involve costs and marketing data that neither the manufacturing manager nor the marketing manager have available and that call for research to support the joint effort in the same way that any innovation calls for research.

	Equipment	
Products	Inflexible ($x_1 = 0$)	Flexible ($x_1 = 1$)
Few ($x_2 = 0$)	10	9
Many ($x_2 = 1$)	8	π

For an example of these ideas, refer to the accompanying table. In this example, suppose the firm currently has inflexible equipment and a narrow product line. Notice that expanding the product line without also adjusting the equipment decision would be too costly to be worthwhile: profits would fall from 10 to 8. Similarly, investing in flexible equipment to manufacture the existing product line would not be worthwhile: profits would fall from 10 to 9. The changes would be worthwhile only if $\pi > 10$, which requires that there be complementarity between the decisions.

In mathematical terms, regardless of π, it is a Nash equilibrium of the game between the two managers for them to choose narrow product lines and inflexible technologies. This is so even though, as portrayed here, both managers have identical objectives and they might both gain by a change (if $\pi > 10$). Notice, too, that in the absence of coordination on a plan, the parties could do strictly worse, winding up with payoffs of 9 or 8. If $\pi \geq 9$, then this is an example of a design decision: coordinating on a single design is more important than having one manager or the other play his or her part of the optimal strategy.

The same example illustrates the role of innovation attributes. Even if the firm has long been employing a strategy involving a narrow product line and inflexible production equipment, the marketing and manufacturing managers might still have learned the consequences of individual changes in their choices by engaging in a process of individual experimentation. However, there is little reason to expect that they could infer π from their individual experiences or experiments. If the environment has changed and π has risen from a low number to one exceeding 10, there may be nothing in the firm's day-to-day operations that make the managers aware of that fact.

c. Complementarities in Business Plans and Organizational Design

In the example given above, the change to a flexible manufacturing, broad product line strategy cannot be worthwhile unless the interaction between the two kinds of changes adds value compared to what the individual changes contribute, that is, unless there are valuable complementarities between the two kinds of changes. Moreover, there are a priori grounds to expect that these changes would be complementary, which makes it plausible that these aspects of the business strategy might well be parts of a design decision.

Complementarities add an especially tractable structure for the analysis of decisions, so we shall pause here to remind the reader of the conditions for, and implications of, complementarity. (A more complete listing of the many implications of complementarity conditions can be found in Milgrom and Roberts [1994a].)

The usual notion of complementarity in mathematical economics is limited to demand theory, where it refers to demand for one input (in the theory of the firm) or good (in the theory of the consumer) falling when the price of a related, complementary input or good rises. The definition used here is equivalent to the usual one when the choice variables are inputs for a firm with a convex

technology. However, our definition is not limited to demand theory problems and allows the methods and intuitions of price theory to be applied rigorously to a much wider range of problems than has been usual.

We define a group of decision variables to be *complementary* when two conditions are satisfied. The first concerns the constraint set, S, in which the choices, x, are constrained to lie: increasing some of the variables must not eliminate any otherwise feasible increase in the remaining variables. (In mathematical terminology, the condition is that S must be a sublattice.) The second condition concerns the payoffs: the (incremental) payoff to increasing any two or more of the variables together must be no less than the sum of the (incremental) payoffs from increasing each variable alone. Formally, if f is the payoff function and $\bar{x} \geq x$ are alternative n-vectors of decisions, then $f(\bar{x}) - f(x) \geq \sum_{i=1}^{n} [f(\bar{x}_i, x_{-i}) - f(x)]$. For a smooth function f, this is equivalent to requiring that for all $i \neq j$, the corresponding mixed partial derivative is nonnegative: $\partial^2 f / \partial x_i \partial x_j \geq 0$. Each of these inequalities reflects the idea that the interaction effects are positive. In the language of mathematics, a function with this property is *supermodular*.

Does the tabulated example given above satisfy the complementarity conditions? There are just two conditions to check. The constraint condition is satisfied by hypothesis: adopting flexible equipment does not preclude broadening the product line (and might even be regarded as necessary if a broad product line is to be produced at all). The second complementarity condition is satisfied as well if $\pi \geq 7$, for then the sum of the changes in profit corresponding to the individual adoption decisions is -3, but the change in profit from adopting both changes together is $\pi - 10 \geq -3$. In that case, the example is indeed one of complementarity.

The presence of complementarities has several implications that are relevant for discussions about organizational change, but we shall discuss just a few of those here. Suppose that $\pi(x, \theta)$ is the firm's profit function with choice variables x and parameter θ. Assume that the vector (x, θ) is complementary. In particular, this requires that increasing the parameter θ raises the marginal returns to some of the activities in x, reduces the marginal returns to none, and does not constrain increases in any component of x. Let $x^*(\cdot)$ denote the profit-maximizing choice[6] as a function of the parameter value θ and let $\hat{x}_i(\theta, x)$ denote the value of x_i that maximizes $f(\cdot, x_{-i}, \theta)$, that is, the optimal value of

[6] Whenever we refer to "the optimum" in an optimization problem, the problems of existence and multiplicity of optima arise. Implicit in our discussion are the background assumptions of continuity of π and compactness of the constraint S, so that existence is assured. When there are multiple optima, then, for specificity, we always focus on the largest one although the smallest would do equally well. In this paragraph, we apply this both to $x^*(\cdot)$ and to $\hat{x}(\cdot)$.

x_i when it is the only choice variable to be changed, with all the other choice variables held fixed. This provides a model of change in which each manager moves without forecasting any change on the part of the others. Suppose that before any change occurs, the environment is θ_B and that, afterwards, it changes to $\theta_A > \theta_B$. Suppose that at time zero, the firm's decisions are optimally adapted to its beginning environment, so that $x(0) = x^*(\theta_B)$. If the managers respond independently of one another, assuming that the others' past choices will not change any further, then their periodic choices will be $x(t) = \hat{x}(x(t-1), \theta_A)$.

THEOREM 1. $x^*(\theta_B) \le x(1) \le x(2) \le \ldots \le x^*(\theta_A)$.

This simple but important theorem asserts three things. First, the change in the optimum will involve increases in all the components of the complementary vector of decision variables.[7] Mathematically, that is the conclusion that $x^*(\theta_B) \le x^*(\theta_A)$. This fact allows a planner who lacks detailed knowledge of the payoff function, π, but does know that it is supermodular to forecast the direction of optimal change and to use that knowledge to guide the firm's decisions. Second, when managers adapt individually without anticipating changes on the parts of other managers, the adjustments they make, while generally in the right direction, tend to be too small.[8] The corresponding mathematical statement is that for all t, $x^*(\theta_B) \le x(t) \le x^*(\theta_A)$. This clearly indicates a value to coordination of fundamental change. Third, in a system of complements, change tends to acquire momentum of its own over time: $x(1) \le x(2) \le \ldots$[9] The pattern of changes may continue until the choices converge to the new optimum (as typically occurs if π is smooth and concave) or it may fall short, but a pattern of continuing change in a recognizable direction does emerge. All of these conclusions take the form of weak inequalities.

To illustrate the theorem, consider again the tabulated 2×2 example and suppose that the relevant change is that the parameter π increases from below 10 (say 8) to above 10 (say 12). In that case, the optimal policy (x_1^*, x_2^*) increases from (0, 0) to (1, 1), consistent with the first conclusion of the theorem. However, the managers, assuming that they each act independently and regard their colleague's strategy as fixed, would make no adjustment at all, consistent with the second conclusion that the managers' adjustments are too small. The third conclusion holds in this case as a weak inequality: the choices remain constant from period to period. If, in addition, the parameter change changed

7 This is a version of a class of monotonicity theorems first developed by Topkis (1978). See Milgrom and Roberts (1994b).

8 This version of the LeChatelier Principle is proven in Milgrom and Roberts (1994b).

9 See Milgrom et al. (1991) for a proof of momentum.

the 8 in the table to 11 (which is consistent with complementarity), then the sequence of (x_1, x_2) choices would be $x(0) = (0, 0)$, $x(1) = (0, 1)$, and $x(t) = (1, 1)$ for $t \ge 2$. In that case, there would be actual momentum for two periods, and the decision would eventually coincide with the optimal one.

A central coordinator might play a constructive role in this problem, either by breaking a logjam to allow the system to move from the point where it is stuck, or by speeding the pace of change, allowing the system to adapt more quickly.

The complementarity structure, when it is known to be present, can also be exploited to simplify the search for better alternatives. In general, when faced with a complex multidimensional problem, a decision maker might be expected to try to simplify the problem by reducing the number of options to study a few leading alternatives. Complementarity structure can help achieve that reduction, as the next theorem illustrates.

THEOREM 2. Let $\pi : S \to R$ be a supermodular function and S a sublattice and let x be an arbitrary point in S. If there is any point $y \in S$ such that $\pi(y) > \pi(x)$, then there is some point $z \in S$ satisfying either $z \ge x$ or $z \le x$ such that $\pi(z) - \pi(x) \ge ½[\pi(y) - \pi(x)]$.

To understand the implications of this theorem,[10] suppose a decision maker is studying a problem with ten variables $z = (z_1, \ldots, z_{10})$, each of which is constrained by $x_i - 1 \le z_i \le x_i + 1$, where x is the organization's initial position. Suppose the coordinator restricts his attention just to changes that involve increasing all the variables $(z \ge x)$ or decreasing all the variables $(z \le x)$. What would be gained and what would be lost by such a restriction? A simple calculation shows what is gained: limiting the search to just coherent plans reduces the number of orthants that need to be studied from 2^n to just 2, reducing the volume of the search region by a factor of more than 500. The cost, of course, is that since the initial point x was arbitrary, the optimum may not lie in the restricted region. Yet, according to the theorem, the improvement in payoff that can be found in the restricted region is *at least* half of what could be found by searching the entire set of possible options. The 50-percent bound actually represents a worst case scenario in which the complementarities are minimal and the initial point is a peculiarly unlucky one.[11] If the complementarities are strong or if the initial point is not chosen so adversely, then the profit im-

10 A proof is given in Milgrom and Roberts (1994a).

11 Indeed, whenever the gain is only 50 percent, one can deduce the global optimum from the information contained in the local search. Let z_L and z_H represent the best points in the orthants $z \le x$ and $z \ge x$. Then the global optimum is $z^* = z_L + z_H - x$. In general, one can use the complementarity structure to guide an efficient search of all the orthants. Along the path followed in this search, the firm's profit converges quickly and monotonically up to the global maximum.

provement to be found among coherent plans will be correspondingly larger. One way to interpret this theorem is as asserting that the complementarity structure gives the decision maker lots of information about the form of a "good" (but not necessarily optimal) decision and allows her to focus attention on that limited set while still enjoying at least half of the returns that would result from a complete search.

d. Complementarity and Fundamental Change

The problem of strategic and organizational choice faced by real firms is highly multidimensional and nonconvex, and there are numerous interactions among the choice variables that might normally be assigned to different managers. The problem is thus sufficiently complex that one might despair of being able to analyze it fruitfully. Yet if the interactions are complementarity relations, there is much to be said, in particular on the issues of partial adjustment and fundamental change.

Consider again the example, from the Introduction, of modern, lean-and-flexible manufacturing versus mass production. Each of these can be interpreted as potential solutions to the organizational design problem. This problem is certainly not convex (for example, there are economies of scale in inventory systems which favor extremal solutions). As we have argued elsewhere (Milgrom and Roberts, 1988, 1990a, 1992, 1994a), it is however a problem marked by widespread complementarity that directly or indirectly links *all* the dimensions of choice in Table 1. In this context, we interpret each pattern as representative of both a local optimum and a team equilibrium of the design problem. This means several things.

First, each pattern is coherent: Changing only some variables while leaving others at their old values cannot be expected to give an accurate indication, even as to sign, of the change in performance that would result from full-fledged adoption of the alternative model. This means, in turn, that changes need to be coordinated, multidimensional, and large, or at least that the negative results from partial changes should not be allowed to deter further change. It also means that half-way measures—partial adjustments in the right general direction—are likely to yield worse results than staying at the original position. These features underlie the General Motors' story of the 1980s: GM invested some $80 billion during the decade in advanced robotics and other flexible capital equipment (several times the market value of Toyota at the time), but failed to make the other changes that were necessary to adopt the new model. The result was an annualized compound rate of return of only 4.0 percent from January 1985 through December 1991, tens of billions in losses in the first third of the 1990s, and the firing of the CEO in a revolt by the board.

Second, any change in the environment of the firm that favors shifting one of the variables from the current model towards the other in fact favors shifting all of them (theorem 1). Thus, for example, a fall in the cost of flexible machinery (such as programmable robots for welding) favors a shift towards the modern model on all dimensions, because all are directly or indirectly linked to the flexibility of machinery. Similarly, factors easing communication with customers and independent suppliers favor wholesale adoption of the new model. Yet the presumed fact that each model is coherent and locally optimal means, too, that there is value to being able to adapt locally to such changes if they are not sufficient to tip the balance between the two alternatives. (We will return to this issue below.)

Third, there is a form of increasing returns to low-dimensional, partial adjustments from one model to the other. As each successive group of decision variables is moved towards the values corresponding to the other model, the payoff to adjusting the others increases. While this means that adjustments get easier over time, it also means that the first steps can be particularly painful. This in turn gives rise to a need for leadership if a fundamental change is to be carried out in more than one step. The leader's role is to enunciate the outlines of the new model and to keep the system moving forward in the face of the worsened performance that results from the initial partial adjustments.

Combining these points yields the conclusion that there is a strong need for global information, because local information or information based only on low-dimensional experiments will not reveal whether a completely different model would be better. The needed information includes the set of relevant dimensions for change and some indication of the nature of the alternative configuration. Theorem 2 suggests that assembling this latter information need not be as daunting a task as might appear, but the former may be significant. Many of the failures of American firms adopting Japanese methods were really failures to understand which complementary changes needed to be considered.

e. Other Factors Favoring Coordinated Planning of Fundamental Change

The elements described above are basic to any explanation of why coordinated change on several dimensions might lead to better outcomes than completely decentralized decision making. Several other elements, however, play important supporting roles. One such element is the need for irreversible investments to implement a new plan. When investments are irreversible, experimentation is costly and it becomes more important to get the plan right the first time. Hill climbing may get one eventually to the top, but with irreversible costs of moving, it may be as important to get wherever one is going in the most direct way as it is to get to the right place. Without interactions among the firm's decisions, these irreversibilities would not be a reason for coordinated decision making. With such interactions, however, irreversibilities increase the benefits from joint planning.

A variation on this theme is associated with learning-by-doing. Many of the processes used in business are improved and refined in unpredictable ways as experience accumulates. This makes it hard to evaluate accurately the success of a new process tried on a small scale or for a short time. When learning-by-doing is important, the direction of change must then be decided before processes can be refined. By the time the actual process performance has been learned, a large investment of time and resources in the process will have been made. Still, learning-by-doing is not itself a reason for coordinated planning; it does not necessarily entail any elements of interdependence among managers. It does, however, entail a significant kind of irreversible investment and so it amplifies the importance of planning when there are interdependencies.

A third factor arises when the interaction takes the form of knowledge or competencies developed in one line of business that will be useful in other lines. In that case, the future beneficiaries of the knowledge acquisition may not even be present when the decision is being taken and the head office may bear the responsibility for representing those interests.

This section has focused on the benefits of central coordination while setting aside the issue of incentives to cooperate in change. The incentive question is the next one to which we turn our attention.

3. Resistance to Change

We now move beyond the team assumption to consider how differing interests among organizational members affect the analysis. Resistance to change in organizations most often comes from individuals who fear that change may bring losses to the organization and, more particularly, to themselves. Only the first of these is an issue under the team assumption, but the latter is crucially important in real companies. Change frequently creates winners and losers, although the extent of this and of the consequent opposition to change can be controlled, because those things depend on many of the details of pay, job security, and the decision processes in organizations.

Why must there be winners and losers from change? Why cannot parties simply bargain efficiently and distribute the gains so that everyone benefits when efficiency is enhanced? In the ordinary course of doing business, even optimal complete contracting implies that, except in very special circumstances, there are good and bad jobs in the firm, with rents or quasi rents accruing to those in the good jobs (Milgrom, 1988). Pay stickiness associated with bounded rationality of management can be an additional reason that pay does not always compensate for differences in job attributes. We use the term *rents* to refer to the net benefits that the employees enjoy in their current job but that they could not expect to duplicate in another job if they were to quit.[12] The fight to hang onto these rents by affecting the firm's decisions is what we have called "influence activities," and the resources wasted in the process are "influence costs" (Milgrom and Roberts, 1990a). We consider some of the sources of such costs in this section and their impact on the issue of ongoing adjustment and fundamental change.

a. Internal Property Rights and Change

Within the firm, property rights play a dual role. On one hand, secure property rights can be an impediment to change, because they give their owners a veto. On the other, protecting worker's rents can be helpful in enlisting their cooperation with proposed changes. Similarly, if jobs are secure, then pay policies that promote equity and a commitment to sharing gains with workers encourages employee cooperation with change (Milgrom and Roberts, 1990b).

As a concrete example, consider the problem of enlisting manufacturing employees cooperation in increasing efficiency in factories. This is a central element in Japanese management systems, where production workers are *trained* to seek improvement in work methods that reduce costs, increase quality, and ease their jobs, are *kept informed* about bottlenecks, quality, and the difficulty of different jobs, and are *rewarded* for seeking improvement. This *kaizen* form of continuous adjustment is regarded as a key element in adjusting to changing conditions as well as in seeking performance improvements in a given context. Employee involvement and empowerment programs, quality circles, and Total Quality Management are some of the ways this has been adapted in Western firms.

Many factors are needed to make such programs work (see below). Perhaps the most important, however, is that employees not fear that any efficiency improvements they generate will result in reduced employment for themselves or their colleagues. Otherwise, the desire to protect job-related rents will overpower whatever incentives are provided for generating changes. The Japanese system of permanent employment achieves this job security as a matter of course, and Western firms have found that to enlist employee involvement they

12 To be precise, these should actually be called *quasi rents*, but that terminology is awkward.

have had to make explicit promises that employee-generated efficiency gains will not be a basis for layoffs.

Even without layoffs, there may be resistance if there is a change from good jobs to bad ones. Pay equity and job rotation can be used to attenuate that problem.

The example of job security also illustrates the tension that exists between the roles of property rights in promoting and blocking change. The permanent employment guaranties that are crucial for the effectiveness of *kaizen* and related practices can prevent large-scale adjustments that would require reducing employment or radically changing the mix of employee skills. The other Japanese practices of consensus decision making and bottom-up policy formation have the same opposing effects, easing the implementation of frequent, small adjustments by ensuring that everyone shares the relevant information, but making adoption of a radical policy change that threatens individuals' rents highly problematic. Japanese firms are wrestling with these contradictions now, as the pressures of the continuing recession there make it increasingly difficult to maintain employment levels. Meanwhile, IBM, which had an explicit no-layoffs policy and a high degree of involvement in decision making through a complex system that was meant to ensure that different units' interests were weighed, found these policies to be an insurmountable barrier to a fundamental restructuring and had to jettison them. A recent example illustrating what happens during downsizing when authority is decentralized comes from Digital Equipment Corporation. Senior management wanted to reduce the number of middle managers in sales. The company's standard procedure was, however, to delegate staffing and personnel decisions to middle management. This precedent was mistakenly followed in this case as well. Reluctant to lay off their immediate colleagues, the managers cut the field sales staff instead (Wilke, 1994).

b. Influence Costs

Influence activities are activities aimed at changing organizational decisions to allow interested parties to capture organizational rents; we have seen several examples of such activities already. These activities are the organization theory analog of the "rent seeking activities" in public sector economics.[13] The costs associated with influence activities are called *influence costs*.

According to the theory of influence activities, there are two main ways that organizations can control influence costs. One is by structuring decision making to limit influence opportunities when the gains to active involvement in decision making are slight. Creating property rights—for example, the right of one division to make certain decisions without others being able to intervene, or the right for a division to veto changes proposed by a second division that adversely affect the first (Milgrom and Roberts, 1990b)—is a common and effective way to limit influence activities, but there are many others. For example, salary decisions that have obvious distributional consequences may be made only once per year or at the time of a promotion, in order to limit the time spent in costly maneuvering over pay levels. Also, the salary setting process itself may be highly structured in order to limit the amount of time and creative energy that workers can effectively devote to raising their salaries.

The second way to reduce influence costs is by adopting policies that reduce the rents for which the parties may contend. For example, to the extent that there are good and bad management jobs associated with different business units, a policy of rotating managers across business units reduces rent differentials. It has the added advantage of making managers more willing to eliminate slack in their units, since they get credit for the savings without having to incur the extra pressures of operating with reduced slack in the indefinite future. Moreover, with frequent job rotation, a manager's career opportunities within a firm are more nearly tied to the firm's success, aligning the manager's interests with those of her employer.

Limiting influence costs is particularly important in declining organizations. Indeed, the desire to limit these costs has been offered as an explanation of the patterns of divestiture of poorly declining units of companies (Meyer et al., 1992). Influence activities have been identified by managers as one of the leading costs of attempts at downsizing.[14]

Here is a simple model that captures some of these effects and highlights the assumptions underpinning the analysis. The variables in the model are i—the efforts devoted to influence by the manager; e, the efforts devoted to reducing waste or "slack" in the business unit; $s(e)$, the level of slack in the unit, $R(s(e), \alpha)$, the rents per period enjoyed by the unit manager; $T(i, \beta)$, the average time in the job until the manager is transferred and the rents are lost; $B(e)$, the bonus received by the manager either in that form or indirectly for efforts devoted to reducing slack; and $C(e, i)$, the private cost incurred by the manager on account of his efforts. The other two symbols, α and β, are parameters that we will use to analyze the effects of various instruments that might be used.

13 Many of the major papers in that literature are reprinted in Buchanan et al. (1980), along with a summary statement by Buchanan of the theory.

14 For example, Wyatt company consultant John Parking asserts that "people become preoccupied with layoffs. People spend more time on internal politics. They become less productive" (as quoted in Byrne, 1994, p. 69).

Our model is not a full-fledged principal-agent model. We make no assumptions about the employer's objectives or what the employer knows and we do not try to optimize the contract. Rather, we want to show how the instruments represented by α and β might affect the agent's behavior regardless of what the employer might want or know. Thus, we focus on the manager's optimization problem, which is

$$\max_{e,i} R(s(e),\alpha)T(i,\beta)+B(e)-C(e,i).$$

We make several assumptions. For convenience, we take the functions to be differentiable, but nothing in this analysis actually depends on that. We make no use of convexity or first-order conditions, so the variables could just as easily be discrete. With this technical assumption as background, the other assumptions are:

(A1) Rents are an increasing function of slack: $R_s > 0$.

(A2) Slack is a decreasing function of efforts at slack reduction: $s' < 0$.

(A3) Job tenure is an increasing function of influence efforts: $T_i > 0$.

(A4) Influence and slack reduction efforts are cost-substitutes for the manager: $C_{ei} > 0$.

Assumption (A1) is justified on the grounds that when a unit has greater slack, it affords its manager more opportunities in a whole variety of ways. She can use the slack to reward favored employees, to reduce the strains and pressures on herself, to indulge in pet projects, and so on. Assumptions (A2) and (A3) characterize the two kinds of efforts: slack reduction effort reduces slack, and influence efforts increase the employee's tenure in the job. Finally, assumption (A4) formalizes the idea that the manager has a limited attention budget to allocate to slack reduction and influence activities, and perhaps other activities as well. This provides one important reason why influence efforts are costly: they substitute for other kinds of productive efforts that the manager might otherwise supply. For example, if influence and slack reduction efforts were perfectly substitutable uses of the manager's time and attention, we could write the cost of effort as $C(e,i) = K(e+i)$. Then, $C_{ei} > 0$ amounts to the assumption that there are increasing marginal costs: $K'' > 0$.

How do changes in the policy parameters α and β affect the manager's behavior? Let us look first at the parameter α.

THEOREM 3. Suppose that assumptions (A1)–(A4) hold. If $R_\alpha \le 0$ and $R_{s\alpha} \le 0$ everywhere, then the agent's optimal choice, $e^*(\alpha), i^*(\alpha)$, has the property that $e^*(\cdot)$ is nondecreasing and $i^*(\cdot)$ is nonincreasing.

Proof. With these assumptions, it is straightforward to verify that the agent's objective function, regarded as a function of $(e,-i,\alpha)$ is supermodular. One

simply verifies that the objective, which we may write as $g(e,-i,\alpha)$, satisfies $g_{ei} \le 0$, $g_{e\alpha} \ge 0$, and $g_{i\alpha} \le 0$. The conclusion then follows from theorem 1.

Thus, policies that reduce rents in the job, $R_\alpha \le 0$, or that attenuate the relationship between rents and slack, $R_{s\alpha} \le 0$, both tend to reduce influence activities and increase the presumably desirable efforts at slack reduction. Reducing rents reduces the direct returns to influence and encourages the worker to devote more time to the substitute activity of cutting slack. Similarly, reducing the relationship between slack reduction and personal rents encourages more effort at slack reduction and diverts efforts away from unproductive influence activities.

THEOREM 4. Suppose that assumptions (A1)–(A4) hold. Then if $T_\beta \le 0$ and $T_{i\beta} \le 0$ everywhere, then the agent's optimal choice, $e^*(\beta), i^*(\beta)$, has the property that $e^*(\cdot)$ is nondecreasing and $i^*(\cdot)$ is nonincreasing.

Proof. The proof is much the same as for theorem 3. One simply verifies that the objective, which we may write as $g(e,-i,\beta)$, satisfies $g_{ei} \le 0$, $g_{e\beta} \ge 0$, and $g_{i\beta} \le 0$. That verifies that the objective is supermodular and the conclusion then follows from theorem 1.

According to theorem 4, reducing the time the manager will spend in this job before being rotated to another or reducing the dependence of that time on subjective factors will reduce influence activities and increase efforts devoted to productive performance such as eliminating slack. The idea is that a manager who anticipates spending a shorter period of time in a particular management assignment has a smaller stake in protecting the rents in that job, leading to a substitution in favor of other kinds of efforts. And, reducing the dependence of job assignments on subjective factors diminishes the payoff to influence activities, again leading to the effort substitution just described. Policies to implement these changes are easy to find. For example, a company might reduce the dependence of job tenure on subjective factors by having a fixed schedule for job rotations.

Another main device to control influence activities within the firm is compensation policy. Pay equity within groups of workers or managers serves to encourage employees to cooperate with each other in promoting change (Milgrom and Roberts, 1990b). Large pay differentials based on comparative performance evaluation can have the reverse effect, leading to costly sabotage among competing workers (Lazear, 1989). Elimination of rents in an organization can be even more important than this analysis suggests, because the agents with the greatest rents are also often the most politically astute and powerful managers.

Affecting the speed of implementation is another way to reduce influence costs. The period from the initial announcement of a decision to change until the change has actually been implemented and become irreversible is one during which internal politics will be most intense and other productive activities will be neglected. Making that period short reduces the costs born during the transition.

4. Conclusion

Taken together, our analyses of the demands of coordination and incentives in implementing change suggest some hard choices for firms facing fundamental change. On the one hand, systemic changes must be centralized to a considerable degree, at least in the guiding vision. But the benefits of such changes will be hard to capture and sustain without the commitment and cooperation of individuals lower in the hierarchy. For without their efforts at refining the new system and their commitment to make it work, there is little chance of long-term success.

As we have seen, these demands are not always in conflict. A company may be able to adopt policies that facilitate some kinds of change. By providing job security, rotating employees to maximize employee breadth and minimize entrenched rents, and emphasizing equity in its pay policies, a company can encourage its employees to cooperate with change. Yet the very need to provide equity and job security can block other kinds of change. Downsizing, especially, necessarily eliminates jobs and rents and undermines the confidence of remaining employees in the security of their jobs. It is virtually impossible to engage in downsizing without doing long-term damage to employee relations.

Other kinds of change may involve different dilemmas. If the change is one dissolving central authority and moving to a more decentralized structure for decisions and action, as in formerly communist countries, and if the set of economic institutions in a system has design attributes (as most complex systems do), then establishing effective and compatible new institutions and practices quickly requires centralized guidance. Yet central guidance destroys the point of the change. The internal contradictions of such transitions make the pain of change unavoidable.

There is no universal analysis of the problem of changing economic organizations and systems. But our analysis suggests that a single set of principles, prominently involving an analysis of the design and innovation attributes associated with the change, the multiple complementary decisions involved, and the prospects for controlling of influence costs, provide a good starting point for such analyses.

Bibliography

Aoki, M. (1986). "Horizontal vs. Vertical Information Structure of the Firm." *American Economic Review*, Vol. 76, pp. 971–983.

—— (1988). *Information, Incentives and Bargaining in the Japanese Economy*. Cambridge.

Buchanan, J., R. Tollison, and G. Tullock (1980). *Toward a Theory of the Rent Seeking Society*. College Station, Texas.

Byrne, J. (1994). "There is an Upside to Downsizing." *Business Week*, May 9, p. 69.

Chandler, A.D., Jr. (1962). *Strategy and Structure: Chapters in the History of the Industrial Enterprise*. Cambridge, Mass.

Cusumano, M.A. (1985). *The Japanese Automobile Industry: Technology and Management at Nissan and Toyota*. Cambridge, Mass.

Hayek, F.A. (1945). "The Use of Knowledge in Society." *American Economic Review*, Vol. 35, pp. 519–530.

Lichtenberg, F. (1991). "Industrial De-Diversification and Its Consequences for Productivity." NBER Working Paper No. 3231, Cambridge, Mass.

Lazear, E. (1989). "Pay Equality and Industrial Politics." *Journal of Political Economy*, Vol. 97, pp. 561–580.

Meyer, M., P. Milgrom, and J. Roberts (1992). "Organizational Prospects, Influence Costs and Ownership Changes." *Journal of Economics and Management Strategy*, Vol. 1, No. 1, pp. 9–35.

Milgrom, P. (1988). "Employment Contracts, Influence Activities and Efficient Organization Design." *Journal of Political Economy*, Vol. 96, pp. 42–60.

Milgrom, P., Y. Qian, and J. Roberts (1991). "Complementarities, Momentum, and the Evolution of Modern Manufacturing." *American Economic Review, Papers and Proceedings*, Vol. 81, No. 2, pp. 85–89.

Milgrom, P., and J. Roberts (1988). "Communication and Inventories as Substitutes in Organizing Production." *Scandinavian Journal of Economics*, Vol. 90, pp. 275–289.

—— (1990a). "Bargaining Costs, Influence Costs and the Organization of Economic Activity." In: J.E. Alt and K.A. Shepsle (eds.), *Perspectives on Positive Political Economy*. Cambridge, pp. 57–89.

Milgrom, P., and J. Roberts (1990b). "The Efficiency of Equity in Organizational Decision Processes." *American Economic Review, Papers and Proceedings*, Vol. 80, No. 2, pp. 154–159.

—— (1990c). "The Economics of Modern Manufacturing: Technology, Strategy and Organization." *American Economic Review*, Vol. 80, pp. 511–528.

—— (1992). *Economics, Organization and Management*. Englewood Ciffs.

—— (1994a). "Complementarities and Fit: Strategy, Structure and Organizational Change." Unpublished.

—— (1994b). "Monotone Methods for Comparative Statics." Unpublished.

—— (1994c). "Complementarities and Systems: Understanding Japanese Economic Organization." *Estudios Economicos*, Vol. 9, pp. 3–42.

North, D., and R. Thomas (1973). *The Rise of the Western World: A New Economic History*. Cambridge.

Porter, M. (1980). *Competitive Strategy*. New York.

Topkis, D.M. (1978). "Minimizing a Submodular Function on a Lattice." *Operations Research*, Vol. 26, pp. 305–321.

Toyoda, E. (1987). *Toyota: Fifty Years in Motion: An Autobiography by the Chairman*. Tokyo.

Wilke, J.R. (1994). "Striving to Adapt—At Digital Equipment, A Resignation Reveals a Key Problem: Selling." *The Wall Street Journal*, April 26.

Williamson, O. (1975). *Markets and Hierarchies: Analysis and Antitrust Implications*. New York.

—— (1985). *The Economic Institutions of Capitalism*. New York.

Womack, J., D. Jones, and D. Roos (1990). *The Machine That Changed the World*. New York.

Comments on Paul Milgrom and John Roberts, "Continuous Adjustment and Fundamental Change in Business Strategy and Organization"

Oliver E. Williamson

Jon Elster, I think, got it right when he advised that "explanations in the social sciences should be organized around (partial) *mechanisms* rather than (general) *theories*" (1994, p. 75). That would be more consistent with the proposed title "Partial Adjustment *or* Fundamental Change . . ." than the "Continuous Adjustment *and* Fundamental Change . . ." title that Paul Milgrom and John Roberts have chosen for their paper. Be that as it may, this paper usefully extends the ambitious research program that the two of them have undertaken, much of it reflected in their recent book on *Economics, Organization and Management* (1992).

My comments deal successively with (1) describing organizations as syndromes of related characteristics, (2) the need for empirical research, (3) the importance of influence costs, and (4) the matter of "how fundamental" is fundamental.

1. Syndromes

Most good ideas have very long histories. The idea of the firm as an internally consistent set of attributes that differs in discrete structural ways from other organizational alternatives is no exception. Recent interest in this condition (and in the development of its comparative institutional importance) has nevertheless moved the idea of the "firm as a system of attributes" well beyond our earlier understanding of this condition.

Relevant contributions include: Chandler (1962) on the transformation of the modern corporation from a unitary to a multidivisional form;[1] Simon, on the clustering of activities within organizations according to principles of near-decomposability (Simon, 1962) and his later enforcement of discrete structural (as against marginal) analysis of comparative economic organization (Simon, 1978); Macneil (1974), who distinguishes classical and relational modes of con-

[1] These organization-form differences are interpreted in Williamson (1970).