ELSEVIER Journal of Accounting and Economics 19 (1995) 179–208

JOURNAL OF Accounting & Economics

Complementarities and fit
Strategy, structure, and organizational change in manufacturing

Paul Milgrom[a], John Roberts*,[b]

[a]*Department of Economics, Stanford University, Stanford, CA 94305-6072, USA*
[b]*Graduate School of Business, Stanford University, Stanford, CA 94305-5015, USA*

(Received September 1993; final revision received August 1994)

Abstract

The theories of supermodular optimization and games provide a framework for the analysis of systems marked by complementarity. We summarize the principal results of these theories and indicate their usefulness by applying them to study the shift to 'modern manufacturing'. We also use them to analyze the characteristic features of the Lincoln Electric Company's strategy and structure.

Key words: Methodology; Supermodularity; Organizational strategy; Lincoln Electric Co.

JEL classification: D20; C60; L23

*Corresponding author.

This paper was begun while the authors were Fellows of the Center for Advanced Study in the Behavioral Sciences at Stanford. The hospitality of the Center, the research assistance of Susan Athey, Joshua Gans, and Scott Schaefer, the financial support of Booz, Allen, and Hamilton, Inc. and the National Science Foundation, the comments of Lorne Carmichael, Michael Gibbs, and Edward Lazear, and the suggestions of the referees and Glenn M. MacDonald (the editor) are gratefully acknowledged.

SSDI 016541019400382 F

1. Introduction

At least since the publication of Alfred Chandler's *Strategy and Structure* (Chandler, 1962), students of business policy and organizations have argued, largely on inductive and experiential grounds, that a firm's strategy, its structure, and its managerial processes have to 'fit' with one another. They have also accentuated the difficulties in achieving fit and, especially, the problems of changing an organization's design and processes to fit new environments or strategies. More recently, many elements of business strategy, structure, and process have come within the purview of economic research, and important advances have been made in understanding these using economic theory. Industrial organization economics (both pre- and post-game theoretic) has provided a logical foundation and method for studying market strategy, while transaction cost economics, the economics of information, and incentive and contract theories have elucidated issues of organizational structure and managerial processes. Yet, despite these advances in the study of strategy and structure, we do not seem to have made much headway on understanding the relations between them, or even in making formal sense of the intuitive notion of fit.

Our purpose in this essay is to suggest that the ideas of complementarity and supermodularity in optimization and games may be quite useful in this regard. As we show, these ideas give substance to previously elusive notions such as 'fit' or 'systems effects', provide some basis for interpreting claims such as the need for strategy and structure to fit one another, give an approach to modeling such issues formally, clarify some ambiguities and enrich our understanding concerning directions of causation, and also suggest reasons why fit may be hard to achieve and change may be slow, painful, and uncertain.

To show how these methods and ideas apply, we will use them to analyze two complicated and quite different manufacturing systems. The first is a formal model that captures many of the elements of the shift from mass production to 'modern', 'lean', or 'flexible' manufacturing, a new paradigm that various authors have described (see, e.g., Milgrom and Roberts, 1990; Womack, Jones, and Roos, 1990). Our treatment of this model, which builds on that in Milgrom and Roberts (1990), illustrates the ways these formal methods can be used to draw rigorous conclusions about situations that previously might have seemed completely intractable because they involve many choice variables and important nonconvexities. Using the mathematics of complementarity, we are able to obtain clear comparative statics conclusions that enable us to interpret observed changes in the strategies and structures of manufacturing firms as optimizing responses to environmental changes. Second, we apply these same ideas to analyze a case study of a manufacturing firm, the Lincoln Electric Company. The case discussion shows how these ideas can be used informally but still rigorously to structure one's thinking about complex strategic and organizational phenomena. As a prelude to these analyses, however, we first need to develop briefly the main mathematical ideas.

2. Complementarity

The notion of complementarity we use is due to Edgeworth: activities are *Edgeworth complements* if doing (more of) any one of them increases the returns to doing (more of) the others. In the differentiable framework that Edgeworth employed, this idea corresponds to positive mixed-partial derivatives of some payoff function: the marginal returns to one variable are increasing in the levels of the other variables. However, for many of the problems one wants to address, it is unnatural or extremely restrictive to assume even divisibility of choice variables, let alone smoothness of objective functions. Fortunately, however, those conditions are also unnecessary.

Looking at the definition above, we see that Edgeworth complementarity is a matter of order – 'doing *more* of one thing *increases* the returns to doing *more* of another'. Moreover, the comparisons and predictions that we typically seek in economic analysis are also a matter of order – we seek to show that a higher level of an exogenous variable leads to higher (or lower) levels of the endogenous variables. The importance of order leads us to focus our formal theorizing on choices from sets of objects that are (partially) ordered. This is the subject of a branch of mathematics known as *lattice theory*, and much of what we report here carries over to general lattices. To avoid unfamiliar concepts, however, we largely limit attention in this essay to the Euclidean lattice $\mathbb{R}^N$ and its subsets. Even on $\mathbb{R}^N$, letting the language and concepts of lattice theory shape and direct the analysis leads both to changes in emphasis and to important new content. Lattice theory spotlights complementarities, casts returns to scale in a supporting role as one special but important source of complementarities, relegates less important ideas (like smoothness) to minor supporting roles, and shows that there is no important role at all for conditions like concavity which have often been featured players in neoclassical economic models.

Formally, a *lattice* $(X, \geq)$ is a set X with a partial order $\geq$ with the property that for any x and y in X, X also contains a smallest element under the order that is larger than both x and y and a largest element that is smaller than both. We write $x \vee y$ (read 'x join y') to denote the smallest element larger than x and y, and $x \wedge y$ (read 'x meet y') to denote the largest element smaller than x and y. The real numbers with the usual (total) order thus is a lattice, and any subset of the real line is also a lattice. (In fact, since each such set is totally ordered, it is a *chain*.) For the Euclidean space $\mathbb{R}^N$ together with the component-wise order, the meet and join operations are given by $x \wedge y = (\min\{x_1, y_1\}, \ldots, \min\{x_N, y_N\})$ and $x \vee y = (\max\{x_1, y_1\}, \ldots, \max\{x_N, y_N\})$, as in Fig. 1. Another example is provided by the set of subsets of some set, with set inclusion defining the partial

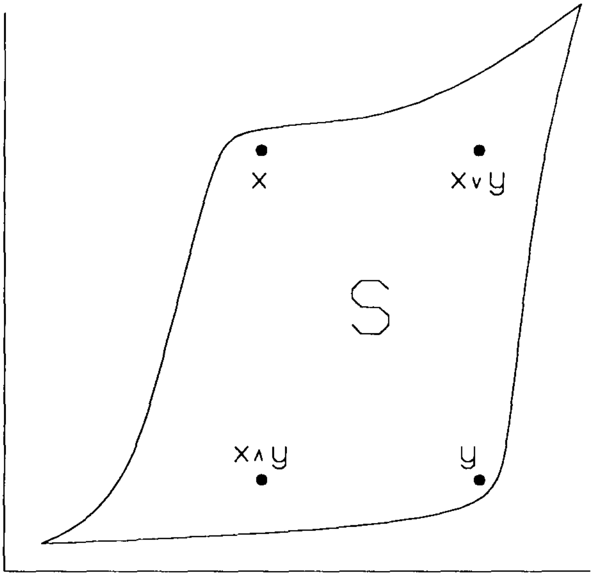

Fig. 1. The sets S, $\{x, y, x \wedge y, x \vee y\}$, $\{x \wedge y, x \vee y\}$, and the four singletons are all sublattices of $\mathbb{R}^2$.

order. In this context, $x \wedge y$ is simply the intersection of the sets x and y, and $x \vee y$ *is* their union. This example is handy in helping recall the meaning of $\wedge$ and $\vee$ in terms of intersection and union, and it also indicates that lattices can be fairly complex entities.

Given a lattice $(X, \geq)$, a *sublattice* is a subset S of X that is closed under the operations of meet and join as defined in terms of the original order $\geq$ on X. For example, any subset of the real line with the usual order is a sublattice. A subset of $\mathbb{R}^2$ (with the component-wise order) gives a sublattice if and only if its boundaries involve no 'downward-sloping' portions. Thus, the set S in Fig. 1 is a sublattice of $\mathbb{R}^2$, as are each of the sets $\{x, y, x \wedge y, x \vee y\}$, $\{y, x \wedge y, x \vee y\}$, $\{x, x \wedge y, x \vee y\}$, $\{x \wedge y, x \vee y\}$, $\{x, x \vee y\}$, $\{y, x \vee y\}$, $\{x, x \wedge y\}$, $\{y, x \wedge y\}$, and the four singleton sets. In contrast, $\{x, y, x \vee y\}$, $\{x, y, x \wedge y\}$, and $\{x, y\}$ are not sublattices. More generally, for Euclidean spaces, the two-dimensional sublattices play an especially important role: Topkis (1976) showed that every sublattice of $\mathbb{R}^N$ can be expressed as by a collection of $N(N-1)/2$ restrictions of the form $(x_i, x_j) \in S_{ij}$, where each S_{ij} is a sublattice of $\mathbb{R}^2$.

The reason for being interested in sublattices is that constraining a choice x to lie in a sublattice expresses a kind of technical complementarity: it says that increasing the value of some variables never prevents one from increasing the others as well (although it may actually require increasing some), and similarly that decreasing some variables never prevents decreasing others. For example,

Table 1

	Low y	High y
Low x	5	4
High x	3	$4+\theta$

a sublattice constraint in $\mathbb{R}^N$ could be used to model the idea that investing in more flexible equipment and a more broadly trained factory work force never prevents a firm from widening its product line, and may be a necessary prerequisite for such a change.

The second element of complementarity is expressed not through the constraints but through the objective function. Given a real-valued function f on a lattice X, we say that f is *supermodular* and its arguments are *(Edgeworth) complements* if and only if for any x and y in X,

$$f(x) - f(x \wedge y) \leq f(x \vee y) - f(y).$$

In the $\mathbb{R}^2$ example, this says that the change in f going from the coordinate-wise minimum, $x \wedge y$, to x (or y) is less than that associated with the 'parallel' move from y (or x) to the maximum, $x \vee y$ (see Fig. 1 again): Raising one of the variables increases the return to raising the other. Note that complementarity is symmetric: If doing more of activity a raises the value of increases in activity b, then increasing b also raises the value of increasing a.

Any function of a single real variable is trivially supermodular. If f is twice continuously differentiable, the defining condition is equivalent to nonnegative mixed-partial derivatives: The marginal returns to increasing any one argument are increasing in the level of any other argument. Thus the Cobb–Douglas function $ax^\alpha y^\beta$ is supermodular on $\mathbb{R}^2_+$ if $a\alpha\beta \geq 0$. If g: $\mathbb{R} \to \mathbb{R}$ is convex, then $g(x+y)$ is supermodular, while if g is concave, then $g(x-y)$ is supermodular in x and y. The sum of supermodular functions is supermodular, as is the product of nonnegative, nondecreasing supermodular functions. If g is increasing and convex and if f is supermodular and increasing (or decreasing) in all its arguments, then $h(x) = g(f(x))$ is supermodular. If $f(x, y)$ is supermodular, so is $h(y) = \max_{x \in X} f(x, y)$. An example to which we will return later is the function given in Table 1. It is supermodular if $\theta \geq -2$.

The theories of optimization of supermodular functions and of noncooperative games in which the payoff functions are supermodular originated in the 1960s in the unpublished work of Donald Topkis and Arthur Veinott. The first published results are those of Topkis (1978, 1979). Extensions of the theories and applications in economics and management have proliferated recently: See, for example, Bagwell and Ramey (1994), Gates, Milgrom, and Roberts (1994), Holmstrom and Milgrom (1994), Meyer, Milgrom, and Roberts (1992), Meyer and Mookherjee (1987), Milgrom, Qian, and Roberts (1991), Milgrom and Roberts (1988, 1990, 1990a, 1991, 1992, 1994, 1994a, 1994b), Milgrom and Shannon (1992), Shannon (1990, 1992), Topkis (1987, 1994), and Vives (1990). A brief, very informal survey of some of the key properties and results will suggest some of the reasons for this interest.[1]

First, supermodularity provides a way to formalize the intuitive idea of synergies and systems effects – the idea that 'the whole is more than the sum of its parts'. To see this in a simple context, let x and y be any two points in $\mathbb{R}^n$ with x strictly larger than y. Supermodularity is mathematically equivalent to the statement that for every such x and y, the gains from increasing every component from y_i to x_i is more than the sum of the gains from the individual increases:

$$f(x) - f(y) \geq \sum_{i=1}^{n} [f(x_i, y_{-i}) - f(y)].$$

Moreover, the implications of supermodularity described below do not depend on the usual kinds of specialized assumptions that economists make for reasons of tractability but that seem so implausible in the business strategy context. For example, we do not need any divisibility or concavity assumptions, so increasing returns are easily encompassed. Indeed, the existence of strong and widespread complementarities among sufficiently many choices will itself imply that the objective cannot be concave. Further, choices might be over such 'messy' things as business strategies and organizational policies, provided we can order each of them in some useful way. In this context, it is worth noting that when there are only two options for each choice variable, then assigning such an ordering is often easy. Moreover, the possibility of assigning the reverse of the 'natural' order to some variables (essentially, of looking instead at their inverses) is very helpful in this regard. For example, in a two-variable problem where the variables are substitutes because the mixed partial derivative is negative, reversing the ordering on one variable reverses the sign of the mixed derivative to yield a system of complements. This trick, used by Vives (1990) in the context of Cournot duopoly, will work whenever one variable is a substitute for all the others.

Second, if x and y maximize a supermodular f on a sublattice S, then so do $x \wedge y$ and $x \vee y$. Thus, the maximizers have a nice pattern and structure: If there is not a unique maximizer, either the maximizers are strictly ordered (with all choices being low in one solution and all being high in another), or else for any unordered pair of maximizers there are other maximizers that are strictly greater and strictly less than both the given ones.

[1] We provide proofs only for those results that have not been shown elsewhere. Interested readers should consult the references listed in this paragraph for the missing details.

Third, a decision maker attempting to verify whether a particular choice x maximizes a supermodular function on a sublattice S can restrict the search for improvements to just those points that are strictly higher or strictly lower than x. If none of these points has a higher payoff, then no point does. Further, optimizing over just this limited domain of alternatives is assured even in the worst case of getting at least half the gains that are potentially attainable from an unrestricted optimization.[2] (If x is not optimal, this worst case arises only when the complementarities are zero, and even then generally it requires an unlucky choice of the starting point.) The ability to restrict search to only 2 out of 2^n orthants with a less than 50% loss in performance could be important in problems with large numbers of choice variables n, for which finding the actual maximum may be too demanding of informational and computational resources.

Fourth, if the domain of a supermodular function $f(x, \theta)$ is a sublattice consisting of vectors of choice variables x and vectors of parameters θ, then the comparative statics on the maximizers are unambiguous: (some selection from) the maximizers $x^*(\theta)$ will be monotone nondecreasing in the parameters θ. For example, in the 2×2 example in the box above, if θ increases from 0 to 2, the optimizer rises from 'low' on both variables to 'high'. More generally, the choice variables tend to move up or down *together* in a systematic, coherent fashion in response to environmental changes, and a change that favors increasing any one variable leads to increases in all the variables.[3] In cross-sectional statistical studies where the various parameters are independently distributed, any two endogenous variables $x_i(\theta)$ and $x_j(\theta)$ will be positively correlated.[4]

Moreover, supermodularity is not merely sufficient for such monotone comparative statics results, it is also necessary if the monotonicity conclusion is to be preserved when one includes additional terms in the objective. Such terms might represent effects and features that are not included in the basic model but which might still be present in actual situations to which the model might be applied. Clearly, we want comparative statics conclusions that remain valid when these effects are recognized. As an example of a robustness restriction, if the

---

[2] *Proof*: Let x^* be an optimum of f and let v be the maximum of $f(y)$ subject to $y \geq x$ or $y \leq x$. Since $x \wedge x^* \leq x$ and $x \vee x^* \geq x$, $2[v - f(x)] \geq [f(x \wedge x^*) - f(x)] + [f(x \vee x^*) - f(x)] \geq f(x^*) - f(x)$, where the last inequality is a rearrangement of the definition of supermodularity.

[3] When the parameter θ is multidimensional, supermodularity of f is actually stronger than is needed for comparative statics. It is enough that f is supermodular in x and each of the components of θ individually; we do not need to control interactions among the components of the parameter to conclude that if the parameter increases, so does the optimizing value of x.

[4] More strongly, the whole vector of endogenous variables will satisfy the statistical criterion known as *association*. This latter condition also has the advantage of being well defined even when x does not take values in $\mathbb{R}^N$.

monotonicity of $x^*(\theta)$ is to obtain for all objectives derived from f by the addition of concave quadratic functions of the individual x_i's, then f must be supermodular. Thus, constructing models whose comparative statics are not sensitively dependent on a detailed specification of all the additive terms in the profit function essentially requires assuming that the objective is supermodular.

Fifth, if the payoff can be written as $f(x_1, \ldots, x_n) + \sum g^i(x_i, y^i)$ for some n disjoint sets of variables y^i and if f is supermodular, then so too is the function $\hat{f}(x_1, \ldots, x_n) \equiv \sup_y f(x_1, \ldots, x_n) + \sum g^i(x_i, y^i)$ obtained by maximizing out the y^i variables. Note that while each y^i is allowed to interact with only one of the components of x, there are no restrictions in this formulation on the nature of the variables y^i – they need not be vectors or numbers or ordered variables. This result allows the theory to be extended to situations where the overall objective function is not supermodular, perhaps because some of the choice variables are substitutes for one another. So long as the firm's objective can be divided up among a set of complementary effects that extend across subunits through the strategic choice variables x and other effects that enter only through the local variables y, the conclusions about complementary choices and their comparative statics are unaffected.

Combining these last two observations suggests that a firm adapting to environmental change will be most likely to find profitable new activities in areas that are complementary to the newly increased activities. For example, suppose the y^i variables are nonnegative real numbers and that $y^i = 0$ at the initial optimum before the parameter change that increases the optimal value of x^i. Then, at the new optimum after the parameter change, y^i is still zero if $\partial g^i/\partial x_i \partial y^i \leq 0$,[5] but y^i can be positive if the reverse inequality holds. Even if the initial position was not an optimum, if the chosen level of x^i increases and the cross-partial with y^i is positive, then increasing y^i is now more attractive. Thus, the search for complementary new activities can help direct the activities of boundedly rational firms in a changing environment.

Sixth, the expected value of a supermodular function in which the choice variables are perturbed by random errors is higher when the perturbations are the same than when they are independent random variables. That is, if $\varepsilon_1, \ldots, \varepsilon_n$ are independent and identically distributed, then $\mathrm{E}[f(x_1 + \varepsilon_1, \ldots, x_n + \varepsilon_n)] \leq \mathrm{E}[f(x_1 + \varepsilon_1, \ldots, x_n + \varepsilon_1)]$. In this mathematical sense, when complementarities are present, 'fit' is important, that is, even mistaken variations from a plan are less costly when they are coordinated than when they are made independently.

[5]This holds because the objective function with x and $-y_i$ as arguments is supermodular, so the optimal value of y_i is a nonincreasing function of x. This is an instance of the 'sign reversal' re-ordering trick.

Seventh, an upward or downward movement of a whole system of complementary variables, once begun, tends to continue. This applies equally to the emergence and growth and to the decline and collapse of systems of complements. As one formalization of this idea, suppose that for each date t, x_t maximizes $f(x_t, x_{t-1})$ subject to $x_t \in S$, where x_{t-1} is fixed by history. If f is supermodular, S is a sublattice, and $x_t \geq x_{t-1}$ for some date t, then the conclusion is that $x_t \leq x_{t+1} \leq \cdots$. Similarly, if the values of x ever decrease, they will continue to do so ever after (until disturbed by some shock). The same implications remain true when the choices after some date t are made nonmyopically to maximize $\sum_{s \geq t} \delta^{s-t} f(x_s, x_{s-1})$ starting from any x_{t-1}.

Many of the popular growth models based on returns to scale can be fit into the foregoing framework, because returns to scale in those models is equivalent to complementarity of choices at different points in time. For example, suppose the payoff earned by a decision maker in period t is a convex function of the stock of capital at that time, which in turn depends on periodic investments. For example, the net benefit might be $B(\sum_{s \leq t} \rho^{t-s} I_s) - C(I_t)$, where B is convex. Then this objective is supermodular in the investment levels I_t: returns to scale in this sense imply that complementarity among investments at different points in time. Similarly, suppose the net capital stock at any date is $K_t = \rho^t K_0 + \sum_{s \leq t} \rho^{t-s} I_s$ and the net benefit at any date is $p_t K_t - I_t C_t(I_t/K_{t-1})$ where each C_t is increasing and convex and $K_0 > 0$. The functions C_t describe the average cost of investment; its argument is the rate of expansion of the capital stock. Then, investments at different points in time are mutually complementary, so higher early investments increase the pace of later investments. The benefits of nonmyopic investment planning in such models are much the same as the benefits from coordination in any other situation with extensive complementarities.

Eighth, a global form of the LeChatelier principle holds when the objective function is supermodular. Let the objective be $f(x_1, \ldots, x_n, \theta)$, let $x^*(\theta)$ denote the optimizing value of x as a function of θ, and let $x^*(\theta; S)$ denote the optimizing values of x when the components x_i, $i \in S$, of the x vector are constrained to be held fixed. Consider an increase in θ from θ' to θ''. Then $x^*(\theta') \leq x^*(\theta'', S') \leq x^*(\theta'', S'') \leq x^*(\theta'')$ for any $S'' \subseteq S'$. In particular, suppose the managers directing the different activities and functions in a firm each select their decision variables to maximize overall profits as a function of the environmental parameters. If they are not able to coordinate their choices, but rather each acts on the assumption that the others' choice variables are fixed at their current levels, then they will systematically under-respond to environmental changes.

This under-responsiveness includes the possibility that decentralized decision making fails to respond at all to the existence of higher common payoffs than are currently being realized. Such coordination failures inherently involve an element of supermodularity. Suppose that the current choices are $x' = (x'_1, \ldots, x'_n)$

yielding a payoff of $f(x')$. Suppose that $f(x_i, x'_{-i}) \leq f(x')$ for any $i = 1, \ldots, n$ and for any x_i, but that there exists $x'' > x'$ such that $f(x'') > f(x')$. (The assumption that $x'' > x'$ is without loss of generality, because we are free to reorder the components of x to make the inequality hold.) Then there is a supermodular function on the n-dimensional interval $[x', x''] = \{x \mid x'_i \leq x_i \leq x''_1\}$ that coincides with f on the relevant domain, $\{x', x'', (x''_i, x'_{-i})_{i=1}^{n}\}$.[6] In the 2×2 case, this domain is itself a sublattice of $\mathbb{R}^2$ and f is supermodular on this sublattice.

The next set of results concern the Nash equilibria of games with strategic complementarities. These are finite- or infinite-strategy games in which the strategy sets are compact sublattices and each player's payoff function is supermodular in the player's own strategy choice, and in which the player's marginal returns are nondecreasing functions of the competitors' strategy choices.[7] Such games have strategic complementarities in the terminology of Bulow, Geanakoplos, and Klemperer (1985): Best response functions are upward-sloping. In such games, there exist largest and smallest pure strategy Nash equilibria. Moreover, these coincide with the largest and smallest serially undominated strategy profiles,[8] the interval of strategy profiles between them contains all the strategies played in any common noncooperative solution concept, and every adaptive learning algorithm leads eventually to the exclusive play of strategies in that interval. Moreover, if the player's payoff functions are supermodular in their choice variables and a parameter, then the largest and smallest equilibrium profiles are nondecreasing vector functions of the parameter.

Tenth, returns to scale is a source of strategic complementarity in games. Diamond's (1982) macroeconomic search model and the network externality models of Farrell and Saloner (1986) and Katz and Shapiro (1986) provide good illustrations. In these models, the payoff of an individual player j has the form $f(\sum_i x_i) - C(x_j)$ where f is convex or $x_j f(\sum_{i \neq j} x_i) - C(x_j)$ where f is increasing. These conditions are usually interpreted as reflecting returns to scale in matching processes, telephone systems, shared technologies, and the like. The key implication is that the mixed partial derivative of the player j's payoff function with respect to x_j and any other x_i is positive. Moreover, complementarity – rather than general returns to scale – provides a better descriptive account in such applications. For example, the gains to personal

[6]The relevant supermodular function is given by $g(y) = f(x') + \sum_{i=1}^{n} [f(y_i, x'_{-1}) - f(x')]$ except that $g(x'') = f(x'')$.

[7]More generally, all the conclusions of the theory still apply when this nondecreasing marginal returns condition is replaced by the following weaker condition: an increase in a players strategy choice that is (weakly) profitable for that player given one specification of the other players' strategies is (weakly) profitable for any higher selection of the others' strategies.

[8]Serially undominated strategies are strategies that survive a process of iterated elimination of pure strategies that are strictly dominated by some other pure strategy.

computer users from focusing on just one or two standards is that it eases the development of complementary products including both software (operating systems, applications software) and hardware (fax boards, monitors, storage devices).

The eleventh point treats the problem of decentralized decision making by an n-member team with a supermodular objective function f defined on a product set $A_1 \times \cdots \times A_n$. Let the team's initial behavior be described by the point $y = (y_1, \ldots, y_n)$ from which no unilateral change by any single team member can increase the team's payoff. This is equivalent to saying that y is a Nash equilibrium of the n-player supermodular game in which each player's payoff function is given by f. In our third and sixth results, we identified reasons why the team might choose to restrict its search for a better point to one single (translated) orthant such as $\{x \mid x \geq y\}$. Suppose that it does so, effectively coordinates its search, and successfully locates a point x^* that maximizes f over that set. Then the result is that x^* is another Nash equilibrium.[9] Moreover, if the team members are initially instructed to play some specific point in the orthant $\{x \mid x \geq y\}$ and then are freed to pursue adaptive learning strategies in which they optimize their action choices given beliefs that are consistent with the others' past choices, their behavior will remain forever trapped in that orthant. (The same conclusions apply when the search is over the set $\{x \mid x \leq y\}$.) The significance of the first of these results is that no further improvement from the new strategy x^* is possible without further coordination among the team members, even when team members are free to search individually for improvements that violate the constraint $x \geq y$. The second result reinforces this message, holding that even if the coordinated move does not take the team members initially to a Nash equilibrium, individual adaptive learning strategies in the class most often considered cannot do better than to find that equilibrium.

Finally, introducing additional 'positive feedbacks' at any point inside the equilibrium interval of a supermodular game tends to increase the distance between the extremal equilibria and, in particular, to make the existence of multiple equilibria more likely. Formally, we start with an N-player supermodular game in which player i's payoff function is denoted by f_i. Let x_L and x_S denote the largest and smallest equilibria of the game and let $x \in [x_S, x_L]$. Suppose $g_1, \ldots, g_N$ are functions with the properties that the game

[9] *Proof*: Let B be the best response map for the game under consideration and let $B_y(x)$ identify the maximum best response to any x over the set $\{z \mid z \geq y\}$. Since y is a Nash equilibrium, $B(y) = y$. Since x^* is the optimum over the set $\{z \mid z \geq y\}$, $x^* = B_y(x^*)$. Finally, since the best response map of a supermodular game is nondecreasing and since $x^* \geq y$, $B(x^*) \geq B(y) = y$. Hence, the constraint that $x \geq y$ is not binding, so $B(x^*) = B_y(x^*) = x^*$. Notice that this argument does not require that each team member control just one decision variable, so the same conclusion applies without that restriction.

with payoffs $f_i + g_i$ is still supermodular and that, for all i, $g_i(y_i, y_{-i})$ is increasing in its first argument on the range $y \geq x$ and decreasing in its first argument on the range $y \leq x$. Let $\bar{x}_L$ and $\bar{x}_S$ be the largest and smallest equilibria of the game with payoff functions $f + g$. Then $\bar{x}_L \geq x_L$ and $\bar{x}_S \leq x_S$.[10]

3. Complementarities, strategy, and structure

Together, these results suggest a basis for thinking about coherence and fit among elements of strategy, structure, and process. They help us model how the elements of optimal firm strategy and structure are linked to one another and, using the comparative statics results, how they would change in a coherent fashion in a changing environment. We will provide examples of this in the next section. As well, they suggest how the strategy and structure of a boundedly rational firm might evolve over time with the adoption of new features that are complementary with existing practices and polices. This will be seen again in our analysis of Lincoln Electric.

These results also provide a basis for understanding why decentralized outcomes can be stable even if they are not optimal and despite experimentation by agents. To see this, consider again the 2×2 example from Table 1 in the preceding section and suppose different managers control x and y, but that both seek to maximize total profits as given by the entries in Table 1. Suppose θ increases from some value $\theta' < 1$ to a value $\theta'' > 1$. Initially, the choices are ('low', 'low'); the new optimum is ('high', 'high'). Yet no amount of individual, uncoordinated search will find an improvement, and the system can get stuck at the suboptimal original position. This example also illustrates how strong complementarities make it more likely that (i) individual adaptations will fail to converge upon optimal results, (ii) the distance from the team's equilibrium to its optimum can be large, and (iii) central strategic direction will be valuable. Note that the results on the efficacy of limited search mean that those providing the strategic direction need not have detailed knowledge of the payoff function in order to be able to help the individual units coordinate on an effective improvement: They literally need only identify the relevant complementarity structure in order to recommend a fruitful 'direction' for coordinated search.

Read more liberally, the results also suggest a reason why change in a system marked by strong and widespread complementarities may be difficult and why

---

[10] *Proof*: First, apply the theorem about the monotonicity of the greatest Nash equilibrium in a parameter to an artificial game where the strategy spaces are restricted to include only strategies greater than x. Then, by the argument of the preceding footnote, this Nash equilibrium is also the greatest equilibrium of the original (unrestricted) game. This proves that $\bar{x}_L \geq x_L$. A similar argument applies to the lowest equilibrium.

centrally directed change may be important for altering systems. Changing only a few of the system elements at a time to their optimal values may not come at all close to achieving all the benefits that are available through a fully coordinated move, and may even have negative payoffs. Of course, if those making the choices fail to recognize all the dimensions across which the complementarities operate, then they may fail to make the full range of necessary adaptations, with unfortunate results. At the same time, coordinating the general direction of a move may substantially ease the coordination problem while still retaining most of the potential benefits of change. Moreover, the systematic errors associated with centrally directed change are less costly than similarly large but uncoordinated errors of independently operating units.

4. Modern manufacturing vs. mass production

The first part of the twentieth century saw a paradigm shift in manufacturing as mass production replaced craft methods (Hounshell, 1984; Womack, Jones, and Roos, 1990). The mass production model spread from the U.S. automobile industry to become the dominant approach world-wide to manufacturing organization, bringing with it remarkable gains in production and wealth. The basic logic rested on interchangeable parts, the transfer line and economies of scale. As the model was refined and perfected, it also came to encompass characteristic features involving firms' product development, manufacturing and marketing strategies, their human resource practices, their internal information, control and decision systems, their relations with customers and suppliers, and their extent of vertical integration (see Table 2).

The last decades of the century are witnessing another such fundamental redefinition of the basic patterns of strategy, organization, and management in manufacturing firms. The changes began in the Japanese automobile industry in the 1950s, but they have now spread internationally and to other industries. In the new pattern that is emerging, the fundamental logic involves flexibility, speed, economies of scope, and exploitation of core competencies. In Milgrom and Roberts (1990) we called this new pattern 'modern manufacturing'; Womack, Jones, and Roos labelled it 'lean manufacturing'. As with mass production, the new pattern involves distinctive approaches to a whole range of policies and structures (see Table 3).

In our 1988 and 1990 papers we offered models involving some of the dimensions on which the two patterns differ. In each paper we asked why the features of the new pattern tended to be associated with one another and why it might be that they were being adopted now. In both papers, the modelled features are mutually complementary, fitting together and supporting one another, and the move towards adopting them is a profit-maximizing response

Table 2
Characteristic features of mass production

Logic: The transfer line, interchangeable parts, and economies of scale

Specialized machinery
Long production runs
Infrequent product changes
Mass marketing
Low worker skill requirements
Specialized skill jobs
Central expertise and coordination
Hierarchic planning and control
Vertical internal communication
Sequential product development
Static optimization
Accent on volume
High inventories
Supply management
Make to stock, Limited communication
Market dealings: Employees and suppliers
Vertical integration

Table 3
Characteristic features of modern manufacturing

Logic: Flexibility, speed, economies of scope, and core competencies

Flexible machines, Low set-up costs
Short production runs
Frequent product improvements
Targeted markets
Highly skilled, cross-trained workers
Worker initiative
Local information and self-regulation
Horizontal communication
Cross-functional development teams
Continuous improvement
Accent on cost and quality
Low inventories
Demand management
Make to order, Extensive communications
Long-term, trust-based relationships
Reliance on outside suppliers

to falling costs of flexible machines, data communications, and computation and to changes in demand that favor broader product lines or more frequent product improvements. (Such changes are plausibly associated with increasing income levels.) The models thus offered a possible explanation for the frequency with which they are seen together in successful manufacturing organizations and for the timing of their adoption.

Before we look at such models in more detail, an extremely simple version of the basic argument, restricted to just two of the relevant variables, provides a useful introduction. Focus on just two of the many decisions that must be made in developing a manufacturing strategy: the flexibility of the production equipment (as measured by the costs of changeovers) and the breadth of the product line. Increased flexibility makes increasing the breadth of the product line more attractive, because making more frequent changeovers and producing in smaller lot sizes allows the improved match with customer preferences to be achieved without having to incur high inventory costs. Simultaneously, a broadened product line increases the value of increased flexibility in the manufacturing process, because the lost economies of scale in inventory that accompany narrower markets for each product mean that it is advantageous to cut production runs and do more frequent changeovers. Manufacturing flexibility and product line breadth are complementary: Increasing either one makes increasing the other more attractive.

Thus, high levels of flexibility ought to be associated with broad product lines, and inflexible production technologies with limited product variety. Both constitute coherent patterns, and either can be successful and, indeed, optimal in the appropriate environment. Henry Ford's transfer line produced anything the customer wanted, as long as it was a black Model T. The entire factory had to be rebuilt when the product design was finally changed. The narrow product line and highly inflexible manufacturing fit one another. Moreover, they were arguably well adapted to the technological and market conditions of the time: They allowed Ford to dominate the market. At the other extreme, Toyota's manufacturing and product strategies represent another coherent pattern. On each of assembly lines, Toyota produces thousands of different variants of several basic designs, essentially to customers' individual orders, and it can rapidly switch these lines over to handle new models. Other aspects of the system are similar: One engine plant produces over three hundred and fifty variants of engine and transmission combinations on a daily, on-going basis. In the current environment, Toyota's approach, which is the archetypical example of lean or modern manufacturing, has been remarkably successful.

Of course, an attempt to achieve the manufacturing economies of the Ford system by narrowing the product line while using manufacturing equipment that is geared to flexibility would not work, nor would an attempt to gain the demand advantages of a broad line while using very inflexible equipment. More generally, mixing elements of two coherent patterns is unlikely to lead to another coherent pattern. As well, it is unreasonable to think that a move from one to the other pattern can be achieved without central coordination. As the 2×2 example suggests [and as DeGroote (1988) has shown in a more fully developed model], if different managers control a firm's two different choice variables in this problem, then – even though the managers share the objective of maximizing aggregate profits – each coherent pattern can represent a Nash equilibrium from which any unilateral change will strictly reduce profits.

As suggested in Tables 2 and 3, the actual range of variables involved in the shift from mass production to modern manufacturing is very large. However, so long as the problem exhibits the sorts of complementarities and nonconvexities that mark the 2 × 2 example, the point remains that changing only some of these from their mass production levels to those associated with lean production cannot generally be expected to yield an improvement, even if a full-scale move would be beneficial. This may account for some of the notable failures that have occurred in manufacturing firms that have attempted to adopt the new ways. For example, General Motors, once the most successful of mass producers, spent some $80 billion during the 1980s on robotics and other capital equipment normally associated with the new methods. It did not, however, make any serious adjustments in its human resource policies, its decision systems, its product development processes, or even in its basic manufacturing procedures. Either it failed to see the importance of making these complementary changes or else, for whatever reason, it was unable to make the changes that were required on these dimensions. The result was that those billions of dollars were largely wasted: GM in the early 1990s had assembly lines that should have been the most flexible in the world but that produced only a single model, while the corporation as a whole lost money at unprecedented rates.

A number of empirical studies and managerial articles have examined complementarities among various of the different aspects of manufacturing strategy and organization. Among the first was Jaikumar (1986, 1989), who noted a complementarity between the use of flexible machine tools and the breadth of the range of products being made, the length of production runs, and the level of work-in-process inventory. Interestingly, the Japanese firms he studied had realized this complementarity and had adapted their methods to take advantage of it, while the US firms in his sample on average had not done so. Instead, they were tending to use flexible equipment to mass-produce large volumes of a few items. See also Hayes and Jaikumar (1988), which accentuates the need for adopting a variety of organizational changes if the full benefits of flexible equipment are to be realized. Nemetz and Fry (1988), while not presenting any data, do draw a number of conclusions from other studies which support the complementarity of the elements in the modern manufacturing pattern. Brown, Reich, and Stern (1993), working from case studies, examine complementarities among different aspects of human resource policies. Helper and Levine (1994) and Kelley, Harrison, and McGrath (1994) examine the empirical evidence for interaction among internal, human resource practices and the nature of relations with suppliers. Brynjolfsson and Hitt (1993) find evidence in firm-level data for complementarities among aspects of investment in information technologies. MacDuffie and Krafcik (1992) find evidence for complementarities between aspects of human resource and manufacturing organization polices in affecting productivity and quality in automobile assembly. Ichniowski, Shaw, and Prennushi (1993) give evidence based on data from the US steel industry that a large number of human resource practices are complementary in affecting productivity. McMillan (1994) surveys research on changing supplier relations and finds support for existence of complementarities there. Finally, Parthasarthy and Sethi (1993) explicitly test for and find evidence of bilateral complementarities between flexible automation and a host of strategic and organizational variables in a multinational sample encompassing several different manufacturing industries.

Together these papers make a good case for the argument that the new pattern in manufacturing does reflect the existence of widespread complementarities. One task for theory is to capture some of these in formal models and to explicate their implications for strategy and structure.

Our 1988 paper attempted to do this, focusing on four of the elements of the system: the breadth of the product line, the extent of communication with customers, levels of finished goods inventories, and the choice of make-to-stock versus make-to-order. We found that make-to-stock and make-to-order were substitutes, and that (because of economies of scale in operating inventory systems) the firm's profit was a strictly convex function of the fraction of customers served on a make-to-order basis. Thus, profit-maximizing firms would tend to specialize, either making to stock or making to order for all customers. Meanwhile, high inventory levels are naturally complementary with producing to stock, while producing to order naturally involves higher levels of early communication with customers in order to plan production. Further, the breadth of the product line and the adoption of the make-to-order regime are complementary because of the economies of scale in inventory systems, which are foregone when the market demand is segmented more finely. Factors that increase the attractiveness of a broader product line (such as shifting tastes or a reduction in the cost of more flexible manufacturing equipment) or that reduce the costs of communication (such as improved telecommunications) tend to favor a shift to the make-to-order regime, lower inventories, and more communication with customers.

Our 1990 paper accentuated the choice of technology, capital investments, and operating systems. The choice variables were price, the production technology as represented by the marginal production cost, the number or frequency of product improvements, the design technology as represented by the marginal design cost of more product varieties, the order processing and delivery times, the number of set-ups per period, the costs of set-ups on new and existing products, and the probability of producing a defective batch requiring rework. In the model in that paper, there were complementarities among these variables that meant that technological changes that eased communication and computation and that lowered the costs of flexible machinery favored a systemic shift in all the variables. This shift involved lower prices,[11] more frequent product improvements, quicker order processing and delivery, more frequent set-ups, a lower chance of stock-outs, a lessened probability of defects, investments that reduce variable production costs and the cost of product redesigns, and the adoption of more flexible manufacturing methods with lower costs of shifting production among existing products and to newly redesigned ones. An extension of the model added reduced vertical integration to the pattern of changes.[12]

Together these models captured many of the aspects of the paradigm shift, but they did not address explicitly a range of human resource management policies that have also been identified as important aspects of the system (see the managerial and empirical work cited earlier). We here offer a model that focuses on some of these. A key element of the model will be the frequencies of product and process innovation. In the spirit of our 1990 paper, we will take falling costs of flexible manufacturing equipment and of product design as the changes in exogenous parameters in the model that lead to shifts in the other variables, but other approaches could have been taken.

The model will involve a dozen choice variables, which is an unusually large number for theoretical economics. Nevertheless, in some respects, the model is still too simple, and the analysis is insufficiently nuanced to be thoroughly satisfying as a treatment of the phenomena in question. Recognizing this, we offer the model as a first-pass attempt to capture some of the effects that have been noted in the managerial and empirical literatures. It also serves to illustrate how the methods surveyed earlier can be used to construct models with the desired comparative statics properties.

Consider a firm whose operating profits depend on its quantity, q, the frequency of new product introductions or product innovations, r, and the frequency or number of process improvements, i: $\pi = \pi(q, r, i)$. We want a model in which a parameter shift increases r and i, and it is convenient to have q increasing too. This leads us to assume that π is supermodular in these three variables. The content of this is that MR minus MC is increasing in each of r and i, while increasing the rate of product innovations increases the attractiveness of increasing the rate of process improvements.

A finer, but perhaps too simple, modelling might specify π as

$$\pi = qP(q,r) - C(q,i),$$

[11]The assumptions on demand originally presented in our 1990 paper actually need to be strengthened to obtain price as one of the elements of the system of complements. See Bushnell and Shepard (1994) and Topkis (1994) for alternative strengthenings.

[12]This analysis made extensive use of the sign-reversal technique, thus permitting the conclusion that some variables fall as others rise.

in which case the assumption that π is supermodular requires only that marginal revenue is increased by product improvements while marginal cost is reduced by process innovations. However, while it seems quite reasonable that revenues should be unaffected by process innovation, it is also plausible that variable costs might depend on r, so direct costs would be $C(q, r, i)$. If marginal costs are reduced by increases in r, reflecting product redesigns that not only are attractive to customers but also are cheaper at the margin to build, then for supermodularity of π the only additional assumption we need is that more frequent process innovations are more valuable the more often the product is being changed. This too seems quite natural. In the differentiable case, this assumption is $C_{ri} < 0$. If more frequent changes in the product tend to raise the marginal costs, however, perhaps because of lack of familiarity with the best way to build the new models, then supermodularity of π also requires that increasing r raises marginal revenue by more than it raises marginal cost.[13]

Undertaking product innovations involves costs, denoted R, for design and for any adjustments to the production system that are needed to produce the new model. We take R to depend on r and on three other variables, e, t, and m: $R = R(r, e, t, m)$. These new variables are, respectively, the efficiency of the design process, the level of training of the workforce, and the flexibility of the manufacturing equipment. Our assumption is that $(-R)$ is supermodular. This means first that increasing e, t, or m reduces the additional costs incurred in increasing the frequency of product innovation. This set of assumptions is almost definitional. The supermodularity assumption also requires that having better trained workers or a more flexible production system does not decrease the benefits of having a more efficient design process. Finally, it requires that having a more highly trained workforce does not reduce the benefit of having more flexible equipment in terms of carrying out product innovations. This latter assumption is perhaps somewhat problematic, because it might well be the case that flexibility of human and physical capital are substitutes. On this issue in the Japanese context, see Koike (1994).

The costs of achieving a particular level of design efficiency e are $E(e, \varepsilon)$, where ε is a parameter. We assume that increases in ε reduce the costs of increasing e, so that $(-E)$ is supermodular. With this, we might think of ε as representing the cost of computer-aided design (CAD) equipment. We might also interpret it as the development of cross-functional teams in product design. Similarly, the cost of achieving a given level of flexibility in the production system is $M(m, \mu)$, where increases in the parameter μ reduce the incremental costs incurred in increasing m so that $(-M)$ is supermodular. Thus, increases in μ might represent the falling costs of computer-numeric controlled machinery and sophisticated robotics.

---

[13]See Athey and Schmutzler (1994) for a much richer analysis of some of these issues.

The costs of process innovations depend on their frequency, i, and on a number of human resource and organizational design variables. Besides the level of training, t, these include the extent to which workers are given autonomy (denoted a) and are able to take actions on their own in light of their detailed knowledge of the production process; the extent of cross-training, denoted s, which helps workers better understand the production process and so facilitates their identifying potential process improvements; and the extent of horizontal communication, h, increases in which help ensure that process changes made at one point do not increase the workload of others. We denote the costs of process innovations by $I(i,t,a,s,h)$, and assume that $(-I)$ is supermodular. This means that increases in any of the other variables lower the costs of doing more process innovations, and that having more of any one of these does not lessen the benefits to having more of the others. For example, it requires that the benefits in lowering the cost of process improvements of having higher levels of worker autonomy are not reduced by also having more communication among workers or by their being better trained.[14]

We take the cost T of training to depend on its level, t, and on the ability level of the workers, b, with the assumption that higher ability levels make it cheaper to provide higher levels of training: $(-T)$ is supermodular. We let $B(b)$ denote the costs of obtaining a workforce of (average) ability b. These costs might be both the costs of more careful screening and any higher wage that is needed to attract such people. The costs of worker autonomy are $A(a)$; these might reflect moral hazard or the failure to adapt adequately to information that is available directly only to those at higher levels in the firm (Aoki, 1986). The costs of cross-training s are $S(s,g,w)$, where g indicates the use of worker groups or teams, organized at a cost $G(g)$, and w is the use of pay-for-skills programs in which workers are compensated not for the job to which they are assigned but for the set of skills they have acquired. We assume that increasing g or w does not increase the cost of additional cross-training and that increases in the use of teams do not make pay-for-skills less attractive. Thus, $(-S)$ is supermodular. If there is any extra cost to using pay-for-skills, it is denoted $W(w)$, and if increased horizontal communication is costly, this cost is denoted $H(h)$.

Our model of the firm's profits is thus

$$\begin{aligned}\Pi(q,r,i,e,t,m,a,s,h,b,g,w;\varepsilon,\mu) = \pi(q,r,i) - R(r,e,t,m) - E(e,\varepsilon) - M(m,\mu)\\ - I(i,t,a,s,h) - T(t,b) - B(b) - A(a)\\ - S(s,g,w) - G(g) - W(w) - H(h).\end{aligned}$$

[14]See Athey, Gans, Schaefer, and Stern (1994) for a richer model of the allocation of decision authority to workers and of some of the complementary decisions that go with this.

Assuming that the feasible values for choice variables lie in a sublattice in $\mathbb{R}^{12}$, then under the assumptions we have made, the objective function is supermodular. (Notice that this does not require divisibility of any of the variables, and we need make no restrictions of concavity of the functions or even on the signs of first derivatives!) Consequently, a fall in the costs of flexible manufacturing equipment (μ rises) or of computer-aided design equipment (ε rises) will lead to a systematic response:

- increased output,
- more frequent product innovatons,
- more frequent process improvements,
- higher levels of training,
- investment in more efficient product design procedures (CAD or cross-functional teams),
- investment in more flexible manufacturing equipment,
- greater autonomy for workers and better use of local information,
- more cross-training, use of teams and pay-for-skills,
- increased screening to identify more able prospective employees,
- increased horizontal communication.

This list captures a wide variety of the features of the new paradigm. What is perhaps most striking, however, is how simple it is to identify the assumptions needed to generate these results. The theory (particularly result #4) establishes the complementarity assumptions that are sufficient to imply the stated conclusions. Further, it establishes that these assumptions are in a certain sense the weakest ones that imply a *robust* comparative statics conclusion, that is, a conclusion that is quite independent of the specification of such functions as A, B, G, H, and W. This is one of the benefits thinking in terms of supermodularity. With this mode of analysis, attention is focussed squarely on the economic structure of the problem as represented by the complementarity assumptions, rather than on the technical issues of specifying tractable functional forms, ensuring the existence of interior optima, managing the case of multiple optima, characterizing the optimum by first-order conditions, and so on.

5. The Lincoln Electric Company

The methods of supermodular optimization and games are clearly useful for proving theorems about formal models, but they are also valuable in giving structure to informal analyses. The key is to use the notion of complementarity carefully, identifying two policies or inputs or activities as complementary precisely when doing (more of) one raises the return to doing (more of) the other. Once the reasonableness of the complementarity hypothesis is verified, one hardly needs to write down a fully specified mathematical model. As we have seen, certain kinds of conclusions follow directly from the complementary structure, without further technical assumptions.

To illustrate how such an informal analysis in conducted, consider the case of the Lincoln Electric Company. This is the most widely used business school teaching case:[15] for over 20 years it has been a staple of MBA classes, and thousands of prospective managers have attempted to understand the remarkable set of policies and procedures that Lincoln Electric employs. Taking a perspective based in the theory of supermodularity and complementarities gives a comprehensive and effective understanding.

Lincoln Electric is a highly successful manufacturer of arc welding equipment based in Cleveland, Ohio. Founded in 1895, it was profitable in every quarter from 1934 up through the beginning of the 1990s; it has never had a layoff; its productivity is far above the average in comparable manufacturing firms; its employees' average hourly earnings including bonuses are roughly double those of nearby manufacturing firms; it draws dozens of applicants for every job opening and suffers turnover of only about 0.5% per year (compared to 4%–5% in other electrical machinery manufacturers); and such giants as General Electric dropped out of the welding equipment business rather than continue to compete with Lincoln and its strategy of constantly lowering prices (and costs) in real terms while still providing superior service.

The firm is famous for its incentive systems that are the focus of the case and that center on widespread use of piece rates. However, the case description reveals a number of other distinctive features to the firm (see Table 4). A complementarity analysis helps us understand these and the relations among them.

The most prominent feature of Lincoln's particular practices is the extreme reliance on piece rates. Production workers are all paid on this basis, even typists were once paid by the keystroke, and (until safety problems arose) the crane operators were paid by the number of loads moved. These rates are set on the basis of time-and-motion studies. A standard output rate is established on the basis of the engineering analysis and from it the piece rate is determined so that a worker who produced at the standard would earn a competitive wage. The worker's actual pay is then the number of units produced times the piece rate (plus any bonus – see below). The firm's policy is to revise the standards only when new machinery or methods are introduced. Workers, are, however, always free to challenge the standards and to have new studies made, at which time the rate may be adjusted up or down.

Given that piece rates have been gradually fading from use elsewhere in American industry, the use of piece rates for manufacturing workers is of some

[15]HBS case #376-028, available from Harvard Business School, Boston, MA 02163.

Table 4
Distinctive features: The Lincoln Electric Company

Piece rates
Internal ownership
Worker–management communication mechanisms
Permanent employment
Bonuses as residual
Dividend payout target
High earnings, excess demand for jobs
Make, not buy
Promotion from within
Flexible work rules
Extensive (firm-specific) training
Old plant and equipment
High inventories
Occasional problems meeting demand
Strategy of being the low-cost producer

interest, and indeed, it has captured the bulk of the attention of many who have studied Lincoln (see, e.g., Wiley, 1990). However, other features that distinguish Lincoln from standard practice in manufacturing are also striking. The firm is largely owned by its employees and managers, and the company has long had both an open door policy for its top executives and institutionalized channels for direct communication between the two groups. There appears to be a target for dividends, with exceptionally high (respectively, substandard) returns accruing to (respectively, decreasing) employee bonuses. These bonuses are a very important part of employees' pay, normally equalling their direct compensation from piece-rate work on average. The individual bonuses are based on supervisors' evaluations on such factors as quality, cooperation, and ideas. There is also a permanent employment policy, with no history of layoffs even in severe recessions: Workers are guaranteed that they will be allowed to work (and earn the piece rates on what they produce or else a competitive wage if they are assigned to other tasks) at least 30 hours per week. At the same time, work rules are quite flexible by traditional U.S. manufacturing standards. Promotion from within, rather than external recruitment, is used whenever possible, and the firm also tends to make inputs internally rather than buying from the outside. The firm provides quite extensive training designed to produce firm-specific human capital: For example, salespeople learn how to make and use Lincoln's welding equipment. Even in the early 1970s, Lincoln was using cross-functional teams for product development while other American manufacturers still used sequential processes. It has relatively old plant and equipment and tends to have high inventories of both work-in-process and final goods. Finally, Lincoln sometimes has problems meeting demand: The case indicates that the only time

Lincoln loses a customer is when it cannot supply the customer in a timely manner.

Each of these features can be seen as part of a coherent pattern in which the pieces fit together in a complementary fashion, making the other pieces more valuable. It is simplest to see this by focusing on the complementarities between piece rates and each of the other features, picking up the complementarities between pairs of these other features in passing.

As is widely understood, paying piece rates encourages output-directed effort. The high employee earnings suggest both that the piece rates encourage them to work at more than the standard rate and that there is probably a selection effect as well, with highly motivated, able workers being differentially attracted to the firm. However, piece rates also give incentives to skimp on quality if quality is not easily monitored and if maintaining quality competes with generating volume. The bonus system helps counter this. In fact, each unit is stencilled with the initials of the people who worked on it, and if it fails after delivery because of a flaw in production, the responsible worker loses as much as 10% of his annual bonus. The bonus for cooperation also helps overcome the tendency for workers to resist helping one another or taking on temporary special tasks that need doing but cannot be paid on a piece rate (both of which would take away from the time when they could be producing and earning money). Thus the bonus and the piece-rate pay scheme are complementary: Using either one makes it more attractive to use the other.

Obviously, if piece rates are effective, different workers will work at different rates, making it necessary to shift workers around to balance the production line. This makes flexible work rules especially valuable and creates a need for work-in-process (WIP) inventories to allow individual workers to continue their production even when there is a temporary slowdown in the preceding or following production step. Thus, Lincoln's exceptionally high WIP inventory levels and flexible work rules are complementary with its piece-rate pay system.

A traditional problem with piece rates is the workers' fear that once they respond to the rates by working hard and thus reveal just how productive they can be, management will raise the output standard and/or lower the piece rate, thereby appropriating quasirents being received by the workers.[16] A host of Lincoln's features are responsive to this. First, roughly 80% of Lincoln's shares are owned by managers and workers, originally through direct stock holdings and, more recently, through an ESOP. This reduces the pressure for lowering piece rates compared to ownership by outside claimants. Moreover, although the employees do not own all the stock, they are essentially residual claimants through the dividend and bonus policies. This has a similar effect. The no-layoff

[16]Lazear (1986) has accentuated the importance of quasirents in this context.

policy supports the worker ownership policy, and, indirectly, piece rates: without it, management or other stockholders could jettison workers, forcing them sell their shares, and thereby gain control. Going the other way, having the workers in control makes the permanent employment promise more credible than it would be if the firm were controlled by outside investors.

The practice of changing rates only when there is a change in technology and methods is clearly aimed at overcoming the workers' fear of management's opportunistically lowering rates, and the communication system further helps develop the trust needed to make the system work. (It also supports the worker-ownership arrangement.) Still, any change in the rates will be an occasion for potential dispute, and especially so during a period of rapid learning-by-doing following a change in equipment. This may discourage making changes in capital as often as would otherwise be done and so may help explain the very low value of plant and equipment on Lincoln's balance sheet. (In the Harvard case, the total value of land, plant, and equipment is less than the value of inventory.)

Permanent employment makes it costly to respond to (possibly temporary) demand increases by adding employees. This accounts for the occasional delivery problems: despite the flexible work rules, Lincoln cannot easily expand production to meet peak demands. At the same time, with guaranteed work there will be some tendency for product to pile up when demand is slack, thus generating high finished-goods inventories in such times. Of course, the reluctance to add workers in the face of temporary demand surges and the need to keep workers occupied during downturns that a permanent employment policy engender increase the value of flexible rules.

The 'make, not buy' policy and the policy of internal promotion may both support other elements of the firm's approach. For example, if the firm uses former production workers in normal times to make inputs that can also be purchased externally, then in peak periods it can move these employees back to making welding equipment and purchase externally. This would give further flexibility that is complementary with the permanent employment policy. The evidence in the case on such matters is, however, not clear.

Together these various policies generate strong incentives for high and growing productivity and the means to achieve it. This is key to the success of Lincoln's chosen competitive strategy of being the low-cost producer. They also ensure that the staff is knowledgeable and that quality is maintained, and these support the provision of superior service.

An important puzzle is why Lincoln's successes have not been copied. What Lincoln does is no secret: the case is familiar to tens of thousands of MBAs; a constant stream of business and union leaders visit Lincoln every year to examine its pay practices; and the firm distributes a videotape about its incentive programs. Lincoln has even been featured on CBS Television's 'Sixty Minutes'. A common answer is that piece rates are unsuitable in situations where the work cannot be efficiently designed to be individually paced (for example, an assembly line) and to permit individual output to be readily identified (for example, team production or long lags between effort being exerted and performance being measured). Yet there are presumably many situations where piece rates are entirely feasible, and yet the documented trend is to move away from individual piece rates. Another potential explanation is that piece rates discourage co-operation and team work. Yet Lincoln's bonus system has apparently overcome this problem. Organized labor's traditional opposition to piece rates could also be a possible reason why Lincoln has not been copied, but this cannot easily explain the failure of nonunion firms to copy successfully. In any case, the high earnings achieved by Lincoln's employees and the manifest desirability of employment there ought to calm union concerns.

An alternative answer lies in a story of competition in the labor market. Matutes, Regibeau, and Rockett (1994) have argued that one firm's paying piece rates while competitors pay wages constitutes an equilibrium in a game of labor market competition for workers of differing abilities, and Lazear (1986) made the same point in a less formal, perfectly competitive model. In equilibrium the piece-rate firm attracts the most able and energetic workers, and their earnings are higher than those employed in the wage-paying firms. The firm paying piece rates is also more profitable, yet (because the situation is an equilibrium) none of its competitors can profitably copy its pay policy. This model accounts for some of the observed features of the Lincoln situation, but it does not seem to be quite adequate. In particular, it predicts that we ought to see similar patterns in other labor markets, and yet it does not seem that we actually do.

The complementarity perspective suggests a quite different answer. Other explanations focus on piece rates almost exclusively. Our discussion suggests that Lincoln's piece rates are a part of a system of mutually enhancing elements, and that one cannot simply pick out a single element, graft it onto a different system without the complementary features, and expect positive results. Analyses of Lincoln that focus on the piece rates and fail to appreciate that their value is dependent on their being supported by the bonus scheme, the ownership structure, the inventory policy, and so on, cannot explain the failures of other companies to mimic Lincoln's system successfully.

Further, even if those who might have copied Lincoln fully understood the significance of the complementarities, many of the elements are difficult to copy. It is easy to announce that the firm will pay piece rates. It is much harder to develop credibility for a no-layoff policy or the worker trust that Lincoln enjoys and has earned over the last sixty years, and harder still to do that while changing over the workforce from one that was self-selected to fit well in a more standard industrial environment to one that will thrive in the Lincoln system.

This latter interpretation is supported by Lincoln's own recent experience. Beginning in 1987, Lincoln expanded its overseas operations very rapidly: In 1987 it had two US plants and three abroad; by 1992 it had 23 plants in 15 countries. Many of these were obtained through acquisitions of existing operations. Lincoln's management's plan was to institute the full Lincoln system in each of these. But as Lincoln Chairman and CEO Donald Hastings acknowledged, the company had 'miscalculated the time it would take. The tenacity of foreign cultures to hang on to their unprofitable ways is startling to me. They have no sense of urgency to make profits ...'[17] In fact, in both 1990 and 1992, Lincoln overall lost money. Although its domestic operations remained profitable and some of the green-field sites overseas were successful, the losses in the acquired operations were more than enough to offset. Strikingly, Lincoln borrowed money to permit it to pay bonuses in its successful operations.

6. Conclusion

The formal notion of complementarities and the corresponding mathematics does seem to provide a promising way to give precision and analytical usefulness to the intuitive and often vague notions of 'fit' and 'synergies' among the elements of an organization's strategy and structure. An additional attraction of the mathematical approach described here is that it derives conclusions from complementarity assumptions alone, without any appeal to the kinds of assumptions that tend to proliferate in the alternative approaches. For complementarity analyses, one has no need for particular functional forms or for convexity, smoothness, or divisibility assumptions. At the same time, as the manufacturing model and the Lincoln case study illustrate, the complementarity perspective is useful both for proving theorems in formal models and for structuring less formal analyses.

The complementarity also raises some interesting research problems. One of these, mentioned earlier, is to estimate empirically the strength of the complementarities: Just how strongly are various elements of the systems linked? Also, which subcollections of activities can be broken off successfully and grafted onto another system? A second involves developing the analysis of overlapping systems of complements. We have noted, for example, that Lincoln uses product development teams, with members from both design engineering and manufacturing, and that it provides employment guarantees. These two characteristics are also common among leading Japanese manufacturing firms. However, Japanese firms differ enormously from Lincoln in their manufacturing practices (which emphasize teamwork and very low inventories), their incentive practices (e.g., paying for skills acquired, rather than output), and so on.

[17] As quoted by Chilton (1993).

A critical question for our theory is how the shared characteristics can be consistent with both systems, when they are so different in other respects. This presents a puzzle and a challenge for further work.

References

Aoki, Masahiko, 1986, Horizontal vs. vertical information structure of the firm, American Economic Review 76, 971–983.

Athey, Susan and Armin Schmutzler, 1994, Information, coordination and the organization of work, Draft (Stanford University, Stanford, CA).

Athey, Susan, Joshua Gans, Scott Schaefer, and Scott Stern, 1994, The allocation of decisions in organizations, Draft (Stanford University, Stanford, CA).

Bagwell, Kyle and Garey Ramey, 1994, Coordination economies, advertising, and search behavior in retail markets, American Economic Review 84, 498–517.

Brown, Clare, Michael Reich, and David Stern, 1993, Becoming a high-performance work organization: The role of security, employee involvement and training, International Journal of Human Resource Management 4, 247–275.

Brynjolfsson, Erik and Lorin Hitt, 1993, New evidence on the returns to information systems, Draft (Massachusetts Institute of Technology, Cambridge, MA).

Bulow, Jeremy I., John D. Geanakoplos, and Paul D. Klemperer, 1985, Multimarket oligopoly: Strategic substitutes and complements, Journal of Political Economy 93, 488–511.

Bushnell, P. Timothy and Allen D. Shepard, 1994, The economics of modern manufacturing: Technology, strategy and organization: Comment, Draft (Allegheny College, Meadville, PA).

Chandler, Alfred Dupont, 1962, Strategy and structure: Chapters in the history of the industrial enterprise (MIT Press, Cambridge MA).

Chilton, Kenneth W., 1993,The double-edged sword of administrative heritage: The case of Lincoln Electric (Washington University, St. Louis, MO).

DeGroote, Xavier, 1988, The strategic choice of production processes, Ph.D. dissertation (Stanford University, Stanford, CA).

Diamond, Peter, 1982, Aggregate demand management in search equilibrium, Journal of Political Economy 90, 881–894.

Farrell, Joseph and Garth Saloner, 1986, Installed base and compatibility: Innovation, product preannouncements, and predation, American Economic Review 76, 940–955.

Gates, Susan, Paul Milgrom, and John Roberts, 1994, Complementarities in the transition from socialism: A firm-level analysis, in: J. McMillan and B. Naughton, eds., Reforming Asian socialism: The growth of market institutions (University of Michigan Press, Ann Arbor, MI) forthcoming.

Hayes, Robert H. and Ramchandran Jaikumar, 1988, Manufacturing's crisis: New technologies, obsolete organizations, Harvard Business Review, Sept.–Oct., 77–85.

Helper, Susan and David I. Levine, 1994, Supplier/customer participation and worker participation: Is there a linkage?, Draft (Case Western Reserve University, Cleveland, OH; University of California, Berkeley, CA).

Holmstrom, Bengt and Paul Milgrom, 1994, The firm as an incentive system, American Economic Review 84, 972–991.

Hounshell, David A., 1984, From the American system to mass production: 1800–1932 (Johns Hopkins University Press, Baltimore, MD).

Ichniowski, Casey, Kathryn Shaw, and Giovanna Prennushi, 1993, The effects of human resource practices on productivity, Draft (Columbia University, New York, NY; Carnegie Mellon University, Pittsburgh, PA).

Jaikumar, Ramchandran, 1986, Post-industrial manufacturing, Harvard Business Review, Nov.–Dec., 69–76.

Jaikumar, Ramchandran, 1989, Japanese flexible manufacturing systems: Impact on the United States, Japan and the World Economy 1, 113–142.

Katz, Michael and Carl Shapiro, 1986, Technology adoption in the presence of network externalities, Journal of Political Economy 94, 822–841.

Kelley, Maryellen R., Bennett Harrison, and Cathleen McGrath, 1994, The congruence of external collaborative manufacturing with internal employee participation, Draft (Carnegie Mellon University, Pittsburgh, PA).

Koike, Kazuo, 1994, Learning and incentive systems in Japanese industry, in: M. Aoki and R. Dore, eds., The Japanese firm: Sources of competitive strength (Clarendon Press, Oxford) 41–65.

Lazear, Edward P., 1986, Salaries and piece rates, Journal of Business 59, 405–431.

MacDuffie, John Paul and John F. Krafcik, 1992, Integrating technology and human resources for high-performance manufacturing: Evidence from the international auto industry, in: Thomas A. Kochan and Michael Useem, eds., Transforming organizations (Oxford University Press, Oxford).

Matutes, Carmen, Pierre Regibeau, and Katharine Rockett, 1994, Compensation schemes and labor market competition: Piece rates versus wage rates, Journal of Economics and Management Strategy 3, 325–344.

Meyer, Margaret and Dilip Mookherjee, 1987, Incentives, compensation and social welfare, Review of Economic Studies 54, 209–226.

Meyer, Margaret, Paul Milgrom, and John Roberts, 1992, Organizational prospects, influence costs and ownership changes, Journal of Economics and Management Strategy 1, 9–35.

Milgrom, Paul and John Roberts, 1988, Communication and inventories as substitutes in Organizing production, Scandinavian Journal of Economics 90, 275–289.

Milgrom, Paul and John Roberts, 1990, The economics of modern manufacturing: Technology, strategy and organization, American Economic Review 80, 511–528.

Milgrom, Paul and John Roberts, 1990a, Rationalizability, learning and equilibrium in games with strategic complementarities, Econometrica 58, 1255–1278.

Milgrom, Paul and John Roberts, 1991, Adaptive and sophisticated learning in repeated normal form games, Games and Economic Behavior 3, 82–100.

Milgrom, Paul and John Roberts, 1992, Economics, organization and management (Prentice Hall, Englewood Cliffs, NJ).

Milgrom, Paul and John Roberts, 1994, Comparative statics, Draft (Stanford University, Stanford, CA).

Milgrom, Paul and John Roberts, 1994a, Comparing equilibria, American Economic Review 84, 441–459.

Milgrom, Paul and John Roberts, 1994b, Complementarities and systems: Understanding Japanese economic organization, Estudios Economicos 17, 3–42.

Milgrom, Paul and Christina Shannon, 1994, Monotone comparative statics, Econometrica 62, 157–180.

Milgrom, Paul, Yingyi Qian, and John Roberts, 1991, Complementarities, momentum, and the evolution of modern manufacturing, American Economic Review 81, 85–89.

Nemetz, Patricia L. and Louis W. Fry, 1988, Flexible manufacturing organizations: Implications for strategy formulation and organizational design, Academy of Management Review 13, 627–638.

Parthasarthy, Raghavan and S. Prakash Sethi, 1993, Relating strategy and structure to flexible automation: A test of fit and performance implications, Strategic Management Journal 14, 529–549.

Shannon, Christina, 1990, An ordinal theory of games with strategic complementarities, Working paper (Stanford University, Stanford, CA).

Shannon, Christina, 1992, Complementarities, comparative statics and nonconvexities in market economies, Ph.D. thesis (Stanford University, Stanford, CA).

Topkis, Donald M., 1976, The structure of sublattices of the product of *n* lattices, Pacific Journal of Mathematics 65, 525–532.

Topkis, Donald M., 1978, Minimizing a submodular function on a lattice, Operations Research 26, 305–321.

Topkis, Donald M., 1979, Equilibrium points in nonzero-sum *n*-person submodular games, Siam Journal of Control and Optimization 17, 773–787.

Topkis, Donald M., 1987, Activity optimization games with complementarity, European Journal of Operations Research 28, 358–368.

Topkis, Donald M., 1994, Manufacturing and market economics, Draft (University of California, Davis, CA).

Vives, Xavier, 1990, Nash equilibrium with strategic complementarities, Journal of Mathematical Economics 19, 305–321.

Wiley, Carolyn, 1990, Incentive plan pushes production, Personnel Journal, Aug., 86–91.

Womack, James, Daniel Jones, and Daniel Roos, 1990, The machine that changed the world (Rawson Associates, New York, NY).