Prototypes and Strategy:

Assigning Causal Credit Using Fuzzy Sets

by

Bruce Kogut*, John Paul MacDuffie**, and Charles Ragin***

October 2004
First working paper: August 1999

We would like to acknowledge the financial support of the Reginald H. Jones Center, and to thank the anonymous referees and editor, Mark Fitchman, John Lafkas, Anca Metiu, Richard Nelson, Michael Trick, Arie Schinnar, and participants in seminars at Carnegie Mellon, Columbia University, Ohio State University, Stockholm School of Economics, and the University of Chicago for their comments. Kriss Drass provided critical programming support and we thank him in memoriam for his work on our behalf.

---

*INSEAD, Fontainebleau.
**Department of Management, Wharton School, University of Pennsylvania
***Department of Sociology, University of Arizona.

Abstract

Strategies often are stylized on the basis of particular prototypes (e.g. differentiate or low cost) whose efficacy is uncertain often due to uncertainty of complex *interactions* among its elements. Because of the difficulty in assigning causal credit to a given element for an outcome, the adoption of better practices that constitute strategies is frequently characterized as lacking in causal validity. We apply Ragin's (2000) fuzzy logic methodology to identify high performance configurations in the 1989 data set of MacDuffie (1995). The results indicate that discrete prototypes of practices are associated with higher performance, but that the variety of outcomes points to experimentation and search. These results reflect the fundamental challenge of complex causality when there is limited diversity in observed experiments given the large number of choice variables. Fuzzy set methodology provides an approach to reduce this complexity by logical rules that permit an exploration of the simplifying assumptions. It is this interaction between prototypical understandings of strategy and exploration in the absence of data that is the most important contribution of this methodology.

Long-term strategy is the choice of capabilities that result in a bundle of attributes embodied in a product or service that allows a firm to position itself via other firms favorably in a market. This characterization suggests then two stages, the first involving the development of capabilities, and the second the exploitation of these capabilities to achieve a particular positioning in the market. The dynamic problem is then the development of capabilities that permits the firm to position competitively in markets for its products and services (Kogut and Kulatilaka, 2002).

The complicating feature of this choice is that these capabilities are embedded in human-machine relationships that are not additive in their effects. In the parlance of recent body of economics, these interactions define complementary practices whose efficacy depends upon the presence of the joint composition (Milgrom and Roberts, 1990). A classic example is the achievement of high performance work systems (MacDuffie, 1995). Such a work system consists of a bundle of practices that improves the productivity and quality of production. Candidate practices are work systems that use human resource policies that dictate incentives and training levels. Since the effective use of one practice is contingent upon the adoption of another practice, there are inherent interactions among these elements.

However, the lists of factors that can compose these systems are many, and the number of experiments is limited. Hence, the task of sorting out these interactions into configurations, or complements, of practices poses a problem of complex dimensionality. If we think of practices as taking on high or low values (e.g. present or absent), then the analysis of two practices suggests looking at a 2^k combinatorial problem. Because dimensionality enters as the exponent, the combinatorial space rapidly expands with the increase in practices. The inter-disciplinary interest in this problem is an indication, in fact, that dimensions are likely to be many. The choice of bundles is influenced by the economics of production, by the internal policies of a firm, and by institutional factors (e.g. unions or regulation). As a consequence, the statistical analysis to identify bundles and measure their effects is itself quite complex.

Recent attempts to sort out this problem have relied upon case descriptions and upon simulations. A case description cannot sort out complex causality and is incapable to determining bundles unless considerable controlled experimentation is permitted. We refer below to the problem of complexity as "assigning causal credit". Simulations can be useful. However, they often avoid the principal points of interest by stipulating a fixed technological landscape and dimensionality, assuming all combinations are visited, and being unable to confront empirical data.[1] In a more philosophical perspective, complexity poses not only the intriguing problem of the contingency of what is knowable, but also the human construction of what is believed to be contingent. Hence, we would like a method that searches for causality but in recognition of its contingent knowability and its human construction. The method we propose to identify 'bundles" under these conditions is fuzzy set methodology.

It is often missed in the literature on the transfer of best practices that there must first be agreement on what are the best practices. Because of the complexity of this assignment of causality, it is not surprising that we deal linguistically with such complexity by the use of fuzzy *prototype* categories that reduces multiple dimensions to discrete categories. For example, the strategy of divisionalization was often defined in reference to General Motors. The literature on high performance work practices in the auto industry has stressed the importance of the Toyota model of production as a point of emulation. The terms "Toyotaism" or "Ohnoism" (after an influential production engineer at Toyota) populate the academic discussion (Coriat, 1991), while the popular discussion has centered on "lean

---

[1] Activity analysis in operations research had long noted the problem of complements and the problem posed to optimization. A literature that addressed this type of questioning is "contingency theory"; see Miller, 1997, and

production" as the generic characterization of the Toyota model (Womack, Jones, and Roos, 1990). The Toyota Production System serves, in effect, as a prototype in the sense of Rosch (1978). Few firms, or plants, conform precisely to the typified Toyota operation, but approximate this idealized type through some degree of possession of the attributes that constitute membership in this category.

Evolving strategies often reflect this competition to migrate toward prototypical configurations that act as poles of attraction guiding the search for better practices. Behind this search is a set of recurring questions: Do these prototypic configurations lead to better outcomes? Did a firm that claims to have adopted "lean production" actually do it and if so, to what extent and to what effect? Is a firm that adopts only new work teams a better example of high performance work systems than a firm that adopts performance-based pay and extensive training? Or are they both examples of transitional systems, or variations of traditional work practice configurations? How do patterns of work practices interact with changes in production practices, such as the implementation of lean inventory buffers, and how much are the combined socio-technical innovations required to affect performance? In short, the inferential problem of assigning causal credit is easily overwhelmed by the *limited diversity* that the world offers as experiments, as well as the fundamental difficulty of categorizing these data into primary units of analysis.

Earlier work has sort to identify bundles by statistical analysis of data, often collected at the plant level.[2] For example, MacDuffie (1995) collected questionnaire observations from auto plants-- the data used in this article-- and developed constructs based on bundles of practices to test their interactive effects on performance, that is, to identify configurations. Similar efforts have been made by Ichniowski, Shaw, and Prennushi (1997) in their analysis

---

Ferguson and Ketchen, 1999, for recent statements. (We thank a referee for these suggestions.) Recent articles using case or simulations are Levinthal, 1997; Rivkin, 2000; and Siggelkow, 2001.

of steel plants. These efforts persistently face the difficulty of omitted influences and the risk of misspecification of the functional form. Comparative work, for example, has found that the adoption of work practices (e.g. mass production, or quality circles) is strongly contingent upon the institutional context of a country (Piore and Sabel, 1984; Cole, 1985). The interaction of contextual factors with work practices creates a high-level problem of dimensionality. As a result, it is very hard to sort out the influence of unobserved contextual factors from the proper specification and identification of the relationship among work practices. Because of the high order of dimensionality in the problem, research into complementarities among elements is often forced to apply *simplifying assumptions* about the interactions that are guided by these prototype understandings.

We seek to provide a grounded method for discovering configurations by applying an inductive fuzzy logic methodology.[3] Fuzzy logic is a classifier methodology that "assigns credit" to specific combinations of traits for achieving an outcome. The problem of credit assignment, to use Holland's (1992) phrasing, arises in the context of genetic algorithms that search for the string of genes responsible for particular phenotypic outcomes. Managerial practices are usually many elements strung together, with opaque clarity as to their causal implications. Unlike biological genes, practices are rarely crisply manifested, but rather are characterized by a fuzzy membership in prototype categories that are cognitively understood. Fuzzy logic, as developed in Ragin (2000), begins with the recognition that categorization is not unique and crisp and that diversity is limited. Based on a fuzzy categorization of membership, it identifies sufficient and necessary configurations, or complements, that explain a given outcome but in reference to simplifying assumptions. In this way, it assigns credit to the combination of elements that are causally responsible for the observed

[2] An excellent review is given in Pfeffer, 1998, and in Ichinowski and Shaw, 2003.
[3] This analysis was reported in the working paper Kogut, MacDuffie, and Ragin (1999).

outcomes, with the caveat that this credit is assigned in the context of limited diversity –the world cannot generate all experiments—and of explicit logical assumptions made by the analyst.

After explaining the methodology, we analyze MacDuffie's (1995) data on high performance systems in the world auto industry. MacDuffie collected data on 70 auto assembly plants throughout the world. He formed three constructs from multiple questionnaire items to measure lean buffers, new work systems, and human resource practices. While controlling for other factors, he found that each of these constructs positively influenced productivity and quality in separate regressions. He also tested for their two-way and three-way interactions, using both multiplicative and log-additive specifications. The results showed that the interactions also were correlated with better performance, suggesting that there were complementarities in their joint interaction. Not all the interactions were positive, and there was modest indication of a lack of robustness in the analysis of quality.[4]

Strategies consist, of course, of more than just the choice of production elements and include such positioning factors as pricing or market choice. In Figure 1, we depict the formulation of strategy as consisting of the state variables that describe a firm's resources and hence its capabilities (one the left hand side) and the choice of markets, prices, and other positioning choice variables (on the right hand side). In our analysis, we hold positioning variables constant by focusing on auto plants that are producing cars for a similar mass market with considerable cross-country shipment of product. By this choice, we analyze for a cross-section in time a sample of plants to determine the configuration of practices and technologies –what can be called production strategies-- that are complements for achieving

[4] Hunter and Lafkas (1998) also show a link to wages from the adoption of high performance systems. See Pil and MacDuffie (1996) for a more recent discussion of bundles and diffusion in the auto industry.

high performance. We define high performance as the joint achievement of high productivity and quality. Through iteration between the fuzzy configurations and the qualitative data (see also MacDuffie 1996, 1997), we seek to provide a rich analysis of high performance cases that lends itself to generalization.

Motivation

The vast debate over the definition of Japanese production methods reveals a history of a discursive search for better practices amid a time of heightened competition and yet create uncertainty over the complex causality in regard to performance. Many academics played important roles in defining and diffusing understandings regarding Japanese practices. For example, Ouchi's (1981) Theory Z analysis pointed to the importance of managerial techniques as the source of competitive gain for Japanese enterprises. In a strikingly precocious study, Schonberger (1982) discussed the combination of practices required to achieve Japanese high quality and high performance in manufacturing plants. Studies were made that rebutted the claim that the source of cost advantage is lower capital costs (see for example Flaherty, 1984). By 1985, a major study on the world automobile industry concluded that the Japanese approach to production organization established a new standard of best practices (Altschuler *et al*, 1984: 161). At the same time, some union studies took a skeptical attitude towards such initiatives as quality circles (Parker, 1985). In addition, there was considerable skepticism over lean production techniques that unions saw as methods to "speed up" the line.[5]

In the studies focused on a single sector, such as automobiles, the growing body of field observations and data suggested a number of practices that might explain a perceived Japanese cost advantage. Yet, there was disagreement over how to categorize these practices and over the variation in Japan that posed the question of what exemplified "Japanese" manufacturing. This debate continues in more recent studies, such as the overview offered by Liker, Fruin, and Adler (1999) that concludes that the Japanese Management System, in their terminology, cannot be reduced to a prototypical configuration exemplified by Toyota.

This debate around best examples, or the ideal type, suggests that the discourse at this time was around category formation (what constitutes new practices) and around prototypes by which to anchor these understandings. (See, for example, Rosch (1978) and the early statement by Lakoff (1973).) In Lakoff's (1987) analysis of prototype categorization, people hold category concepts that are characterized by central members, or objects. Members more distant from these central prototypes are peripheral; hence categories are radial, with central and peripheral membership. A classic example of a prototype illustration is the category of birds (Lakoff, 1987: 44-45). Though most people would agree that a robin is an excellent member of the category of birds, an ostrich or penguin are more distant members. Scientifically, their membership may be satisfied by a definition of the required genetic makeup of a bird. However, cognitively, people hold a prototypical image of a bird, and membership to this class is characterized by a radial property in which some members are attributed a higher degree of membership than others. In fact, members to the same category may hold no feature in common, and yet the implicit categorization may link them through a "category chaining." For example, a penguin and ostrich may have no common defining characteristic of "birdness," and yet belong to the same category due to their sharing different traits in common with the central trait.

Fuzzy sets are, as discussed below, exactly these polythetic categories that classify membership by a type of chaining rule. The methodology classifies cases by membership,

---

[5] Adler (1993) provides an incisive examination of this debate by looking at the General Motors-Toyota joint venture; Sengerburger (1992) reviews some of the reactions of unions in several countries.

treating them as characterized by configurations of attributes. It infers causality by testing all combinations against their membership value in the set of outcomes (e.g. productivity) and, thereby, assigns credit to the individual factors that are logically identified as explanatory, either separately or as discrete combinations. It then returns to the field observations by analyzing the prototypical cases. It is this iteration between formal classification and qualitative assessment that distinguishes fuzzy set methodology from more statistical approaches.

Yet all of these studies collect data on somewhat different variables, propose different bundles or clusters of practices, and suffer from the problems of unobserved effects and the difficulty of estimating the full set of interactions among practices, as noted above. In the language of an inductive analysis, these results diverge because of a disagreement about the size of the dimensional space, the variables that define this space, and the specification of the complexity of these variables. Logical analysis resolves these issues by conceding them. The determination of a configuration of variables that are causally related to a given outcome (e.g. high performance) is sensitive to dimensionality and limited diversity. This problem is not eliminated by complex distributional assumptions regarding unobserved effects. To the contrary, the problem (which manifests itself in the Boolean logic as contradictions, or as unexplored diversity) is an invitation to return to the cases, informed by an inductive empiricism combined with explicit theoretical suppositions.

In the academic discussion, the eventual evidence pointed to the claim that best practices could be represented by a prototype drawn from the Japanese examples that consisted of advanced automation and three *sets of practices*: work, inventory management, and human resources. Ichinowski et al. (1997) determined that these factors were the complements that were suitable for steel plants producing for an environment marked by an increasing combination of cost and quality considerations. Similarly, MacDuffie (1995) argued that these three practices, while controlling for technology and scale, produced jointly high performance, as measured by quality and productivity. In the work below, we propose this prototype as the working theory: plants that are characterized by all three of these practices dominate those that characterized by two or, even more so, by one or none. It is possible, in fact, that in the absence of one or two of these practices, the best choice would be *not* to choose the third practice. Thus, we would like to have a method that relates polythetic categories to performance outcomes. We propose a fuzzy set methodology for this purpose.

Ideal type profile analysis, as proposed by a reviewer of this paper, assumes that all elements of the ideal type be considering when examining the fit of each case to this type. In the fuzzy set analysis, the goal is to examine the different configurations of features derived from a prototype (or ideal type if preferred) that are linked to specific outcomes. In effect, fuzzy set analysis disassembles the ideal type and then reassembles them systematically through testing their causal relation to an outcome. This method is not atheoretical; it starts with a prototype and then provides a more exhaustive inferential engine to identify multiple conjunctural causation. If, by ideal type analysis, it should be meant the testing all possible configurations for their causal claims, it then indeed converges to the Boolean (or fuzzy set) methodology. However, ideal type or contingency theory has not produced any adequate alternative methodology, because of a failure to understand the conceptual challenges, and opportunities, to exploring causal complexity.

Boolean crisp sets

Given this complexity, a natural approach is to turn to non-parametric methodologies that rely upon rankings and that engage the researcher in trying to identify the causality. One approach is to identify logically the possible interactions as bundles of complements that define a configuration. The analysis of configurations confronts the difficulty of trying to understand "configurations" whose elements share an unspecified and unknown relationship among themselves in reference to an observed outcome. In crisp Boolean logic, these elements are coded 0 or 1, and their observed effect is also coded as 0 or 1. Each configuration indicates, consequently, a truth statement that pairs a particular configuration of elements to a binary outcome.

Qualitative comparative analysis uses Boolean logic to identify the minimal list of configurations that determine the truth condition of the observed cases (Ragin, 1987). It proceeds by inductively coding the configuration and truth condition of each case, and then applying a "logic" algorithm developed for electronic circuit design to find robust causal (or functional) relationships that reduce the observed truth table to a minimal number of logical statements.

To return to the example of the auto industry, it is often posited that new work practices (e.g. work teams plus job rotation plus off-line problem-solving groups) and certain human resource practices (e.g. extensive training, performance-based pay) are required to achieve a high performance system. We would code the two causal factors as 0 if absent in a given factory, and as 1 if present; similarly, we code high performance as absent, 0, or present, 1. Since any causal element can take 2 values, there are then 2^k, or 4, possible configurations: {0,0}, {0,1}, {1,0}, {1,1}. Let's make the critical assumption-- to which we will return later-- that we empirically observe each of these configurations, and each configuration has a corresponding truth value of low performance (0) or high performance (1).

We want to pose the question what is the minimal "covering" logic to which we can reduce the 4 possible configurations. This reduction is both an empirical and logical question, that is, we need to know the empirical truth values in order to make the logical reduction. Consider, for example, two configurations where the first two columns refer to work teams and training, respectively, and the third column gives the truth value for high performance.

Case 1: 1 0: 1

Case 2: 1 1: 1

In this case, the second factor is clearly redundant and the presence of work teams is *sufficient* to cause high performance. Our two-dimensional box collapses to a line whose end points [0,1] sufficiently determine the truth condition. By sufficiency, we mean the logical inference that an effect is present whenever a given cause is also present. We can also say that a configuration is sufficient if, whenever the member factors are jointly present, they always generate a given effect.

To illustrate necessity, consider an effect that has 3 potential causes. To continue our example, we can add, to work teams and training, the third causal condition (column 3 below) of whether a factory is lean (1 for low inventory buffers) or not lean (0 for high buffers). Three factories have the following configuration and associated truth values:

Case 1: 101: 1

Case 2: 111: 1

Case 3: 100: 0

For these configurations, we no longer can claim that work teams (column 1 entries) are sufficient, for they are present in case 3 and yet the effect of high performance was not observed. A comparison of cases 1 and 2 eliminates training as a causal factor and implies that high performance is caused by the joint presence of work teams and low buffers. Case 3 indicates, though, that work teams are not sufficient to cause high performance in the absence of low buffers; such practices are necessary but not sufficient. Work teams were present in every configuration associated with high performance. Thus we can infer that they are a necessary condition; if they are not present, high performance is not observed.

The logic of necessary and sufficiency conditions is essentially, then, a statement about the set-theoretic relationships between cause (X) and effect (Y). A necessary condition always subsumes the set of outcomes. There are cases in which a necessary cause is present but there is no effect, but there is never a case in which the effect is present but the necessary cause is not. In other words, there is no case in which Y but not X. (We relax this statement below to hold true statistically, but not absolutely.) Sufficiency implies that the outcome also includes the set of sufficient causes. There may be cases where high performance exists but a sufficient cause is missing, but a sufficient cause cannot be present without the presence of high performance. In other words, there is no case in which X but not Y.

Thus a cause (X) that is sufficient or necessary for a given effect (Y) implies the following relationships:

$$\begin{aligned}&\text{X is a necessary condition}: Y \subseteq X \text{ if } Y \Rightarrow X\\&\text{X is a sufficient condition}: Y \supseteq X \text{ if } Y \Leftarrow X\end{aligned}$$

In the case that Y and X are subsets of each other, then we can infer that X is a necessary and sufficient cause of Y.

Of course, causes need not be individually sufficient or necessary and the logical reduction of cases may result in a complex array of causal configurations. Boolean comparative analysis essentially is an inductive logic to find the minimal set of configurations that explains the truth condition. A configuration is itself the intersection of factors whose conjunction causes an outcome. To say that the combination of lean buffers and new work practices cause high performance through their joint presence is logically equivalent to stating that their intersection is causally associated with a particular truth condition. By intersection, we mean that lean buffers “AND” new work practices causes high performance.

These simple definitions formalize some of the discussion on universality, contingency, and configuration. A sufficient condition is universal; a necessary condition – when not also sufficient—is contingent, or perhaps better said, all causal combinations are contingent on its presence (see Delery and Doty, 1996). For social science, it is common to find that a given effect is associated with multiple configurations. Multiple conjunctural causation is characterized by the condition of an effect being produced by different combinations of factors. A listing of these causal combinations is expressed logically as the union of the configurations. Union means, for example, that lean buffers "OR" new work practices causes high performance. (In this example, we would conclude that either condition is sufficient.)

Boolean minimization relies upon two principal operations:

Absorption: A+ AB= A

Reduction: AB + Ab= A(B+b) =A(1)= A

The second operation is derived directly from the distributive and complement laws of Boolean algebra.[6] The first operation derives from the laws of subset. If AB is the intersection of the sets A and B, then this intersection must be equal to, or be a subset of, A.

How many possible logical configurations are there? In the degenerate case of no variance in the truth condition, each configuration is causally associated with the outcome and, consequently, there is no possible reduction in the configurations. With variance in truth conditions, the application of Boolean logic reduces configurations to simpler causal statements.

In the earlier example, we skipped by an important point that a factor might be causal in its presence or absence, or be redundant. The 2^k calculation, illustrated above, assumes

---

[6] In Boolean (and fuzzy) algebra, union (logical OR) is indicated with a plus sign (e.g. A+B), while intersection (logical AND) is inducated through multiplication (e.g. AB).

that each factor is causal. As we saw, the application of Boolean logic seeks to reduce these configurations to more robust and general relationships, and some factors might drop out as redundant. Lean buffers, for example, might cause high performance; not lean buffers might also cause high performance (perhaps in conjunction with high volume); lean or not lean buffers may have no effect at all. Let's demarcate the presence of "lean buffers" by a big B, "not lean buffers" by lower case b, and its absence of any effect by eliminating it from the causal configuration, denoted by "-". We have then 3 possible states that lean buffers might take-- present (B), absent (b), no causal effect (-). Similarly, we use "T", "t", and "-" to denote teams, not teams, and no causal effect of teams. Consequently, if n (the number of possible causal factors) is 2, we have $3^n - 1$, or 8, possible causal combinations: {b,t}, {B,t},{b,T},{B,T},{B,-},{b,-},{-,T},{-,t}.

If the number of cases is large, the probabilistic significance of each observed configuration can be tested against a benchmark proportion, called p*, that represents an analogue to the researcher's prior of the mean success of a "very good" theoretical prediction. The realized success of a configuration in correctly predicting a truth value can be compared against this benchmark, and this deviation-- along with the sample size and estimate of the sample variance-- can be used to calculate a Z-score as a measure of probabilistic significance:[7]

$$\frac{(P - P^{*}) - \frac{1}{2N}}{\sqrt{\frac{pq}{N}}} \geq z$$

Obviously, if the number of cases is small, it will be difficult to reach significance.

This latter observation raises the important issue that some configurations will not be observed. This problem of *limited diversity* is distinct from the issue of specification error

[7] For n < 30, a binomial probability test can be used.

through omitted variables. Of the possible interpretations, two are particularly important. The first is that limited diversity reflects a weakness in the research design to sample cases for all experimental combinations. An analogue would be a study of the effects of smoking on mortality of men and women that failed to include any observations on smoking women. But another possibility is that nature does not run all experiments. This possibility raises the question of what should be the inference from missing configurations. The Boolean approach forces the researcher to analyze the implications of unobserved logical combinations. This contrasts sharply with conventional statistical analysis, where regions of the vector space that lack cases are included in the results by implication, with no thoughtful consideration of these regions. Through an examination of limited diversity directly, the researcher is invited to explore existing and *possible* worlds.

Fuzziness:

It is an obvious objection that the world rarely conforms to a binary, or crisp, characterization. A rich person is different than very rich. Sexual membership as male or female is, biologically, relatively crisp in some respects, but less so in others. It is clearly not crisp if the question is sexual preference or sexual identification. It is common in social science research to rely on categories to offer discrete approximations of a continuum. For example, rich countries have per capita income in excess of $15000, middle income is less than $15,000 but more than $5,000, and the income of poor countries is less than $5000. It is possible to code each of these discrete categories as three binary variables. The logical complexity increases dramatically through this method, since the number of configurations increases exponentially by 2^n.

However, there is a more fundamental issue than logical complexity concerning the way people categorize and describe phenomena. It was noted early that individual often classify on the basis of prototypes. Prototypes are best examples of members belonging to the same category. The usage of prototypes implies, therefore, that the degree of membership is a gradient, with more distant members holding lower degrees of membership.[8] Using this concept, we define membership in a fuzzy set of a given member x in the fuzzy set of A as

$$m_A(x) = Degree(x \in A)$$

Degree of membership can be geometrically portrayed by a hypercube in which a set is no longer constrained to be located at one of the "crisp" vertices. The simple case is a straight line:

0____________________.5____________________1

The two end points are the crisp values of 1 or 0, in or out of the set. Values in between identify fuzzy membership, e.g. fairly rich countries or not very rich countries (Klir and Yuan, 1995). The mid-point, .5, is of interest, for it defines maximal fuzziness (or what Kosko (1993) refers to as maximal entropy) and it represents a natural cognitive anchor.

A prevailing practice in statistical work is to combine like-items into a scale by imposing a functional transformation. For example, the data can be factor analyzed, or transformed into z-scores while testing for their inter-item discrimination. Membership values in a fuzzy set can also be subjected to scaling. The caveat to scaling is that since the causal analysis (as described below) relies upon greater than, or less than, relations (rather than correlations), the results are very sensitive to the data values.

Partially as a consequence of this sensitivity, the assignment of membership can be strongly influenced by linguistic hedges (Klir and Yuan, 1995: 230-231). Zadeh (1972) proposed that such a hedge as "very" signifies that membership values should be squared (what he called concentration). The hedge "fairly" is naturally captured by taking the square

[8] We flag that there is a debate regarding prototype theory and fuzzy logic. For example, Lakoff (1973) sees fuzzy logic as insufficient for fully accounting for observed categorization heuristics.

root of membership (or what he referred to as "dilation"). These transformations have a common sense property. Clearly, an apple that has a membership value of .5 in the set of red apples should have a lower membership value in the set of very red apples.

The above example relies intuitively upon a notion of subsets. An important property upon which we rely heavily in the analysis below is that membership of x in a subset of A is less than or equal to membership in the set of A:

$m_{B\leq A}(x) \leq m_A(x)$

Figure 2 provides a graphical illustration that membership of X in the subset of A, defined by a 2-dimensional space, lies in the domain of the set of A.

Fuzzy set logic:

The categorization of entities by their degree of membership means that categories are not exclusive. This property has the attractive feature of conforming to commonsense notions of categories: people can be somewhat religious or somewhat moral. Manufacturing plants similarly have high membership in new work practices, but low membership in team organization. This property of membership, however, poses the question of how should we define the intersection and union of fuzzy sets. What is the membership value of a plant in the intersection of new work practices and work organization?

Because membership values are binary, logical operations on fuzzy sets are more complicated than crisp operations, though fairly simple. The key difference is that membership values in a fuzzy set lies in the interval of [0,1]. As a result, the operations of negation, union, and intersection must heed the membership values.

Negation: In crisp logic, the set of A has the complement of the set of not-A. (See Klir and Yuan, 1995:50). This operation applies also to fuzzy sets. Consider the set A whose element X has a fuzzy membership denoted by a point along the unit interval. Then, negation is simply

$m_{\bar{A}}(x) = 1 - m_A(x)$

This definition is technically intuitive, and yet deserves a note of caution. For while the complement of rich is not rich, we would not want to say that the complement of rich is poor. We may view Portugal as holding a membership value of .4 in the set of rich countries, and hence the value of .6 in the set of not rich countries. Yet, we may assess its membership in the set of poor countries as considerably less than .6. Language matters in understanding fuzzy sets, and the use of a predicate logic does not eradicate the ambiguity in linguistic terms and quantifiers.

Union: The union of two sets is logically denoted as an "or" operation. The union of A and B implies that x belongs to A or B. However, this denotation is complicated in the context of fuzzy logic, because the membership of x in A or B can take on any value between, and including, 0 and 1. Fuzzy logic applies the union operator by taking the maximum of the membership value of X in each of the two sets:

$m_{A \cup B}(x) = \max(m_A(x), m_B(x))$

If X is short and smart with membership values of .5 and .8 respectively, in these two sets, X has then a membership value of .8 in the set of people who are short or smart . This definition corresponds intuitively with the implication of an "or" operation. That is, x is a member of set A or set B with degree of membership equal to its maximum membership in each set.

Intersection: Fuzzy logic defines the intersection operator as the minimum of the membership degree of X in each of the two sets:

$m_{A \cap B}(x) = \min(m_A(x), m_B(x))$

The intersection of two sets is logically denoted as an "and" operation. To belong to two sets means that X is member of both set A and set B. If X is not jointly a member, then it does not belong to the intersection. Again, we see a complication that X is likely to have different membership degrees in the two sets. It is unappealing that X's membership in the intersection should be greater than its membership in either of the individual sets.

The application of the minimum operator makes intuitive sense and is consistent with a prototype theory of membership. Consider the adjectives of big and furry to describe dogs. A given dog can be furry and very small, and it has membership values of .9 and .10 in the respective sets of furry and big. To average these membership values would give the misleading impression that furry can linearly compensate for being small. It might be surprising, having purchased a dog by the internet without a photo and who bore only the characterization as "a more or less" member in the set of big and furry dogs, to open a big box containing a Pekinese. To most, a Pekinese has a low degree membership in the club of dogs who are both furry and big.[9] The minimum operator also makes formal sense. Recall the earlier definition of complementarities as supermodular. Since the value of doing two things together is higher than when they are apart, it makes sense to guarantee that the arguments to the function are all increasing. Taking the maximum would neglect the inferior argument. The minimum indexes increases in the joint presence of two variables by the least value. This permits a direct test of whether the minimum of doing two (or three or more) is associated with increases in performance.

Fuzzy causal inference:

Assigning membership values to all possible combinations constitutes the first step in the analysis. The second step is to derive those combinations, or complements, that explain the causality of observed outcomes. Causality in fuzzy logic shares some of the intuitive properties commonly confronted in statistical work. In linear specifications, we ask how does y vary with more of x. Fuzzy causal inference relies upon the set-theoretic definitions of necessity and sufficiency to identify factors that satisfy the sub-set axioms (Ragin, 2000). For necessity, the outcome is a subset of the causal factor. Necessity implies, then, that the membership degree of a case in a causal factor should be associated with a *smaller* membership value in an outcome. For sufficiency, the causal factor is a subset of the outcome. Sufficiency implies, then, that the membership degree of a case in the causal factor should be associated with a *larger* membership value in an outcome.

A graphical illustration of determining necessary and sufficient conditions can be given by graphing the degree of membership in a hypercube in which a set is no longer constrained to be located at one of the "crisp" vertices. Figure 3 shows a hypothetical relationship between lean buffers and the causal outcome of high performance. Lean buffers satisfies the axiomatic definition of a necessary condition, because all cases have larger membership degrees in it than in the causal outcome.

Figure 4a portrays the analysis of sufficiency. Since the membership value in work teams uniformly is less than the membership degree in the causal outcome for all cases, we conclude that lean buffers is sufficient. Figure 4b illustrates the same analysis for a configuration of two factors (lean buffers and work teams). Since we are looking at their joint effect (or intersection), we take the minimum of each case's membership value in these two factors. The minimum effectively moves the distribution of dots to the left, except for the unlikely case that the membership values in the two causal factors are the same.

It is obvious that a given factor cannot be both sufficient and necessary, except for the cases when the causal factor and causal outcome share the same membership values. Empirically, we expect that a causal factor or configuration will not be found only above or

---

[9] Hampton (1997) summarizes some of the objections from cognitive psychology to fuzzy set definitions of prototypes. Part of these objections consist of problems of taking intersections among nested sets, a classic

below the diagonal. The statistical formula to calculate the z-score, as given above, permits an assessment of the statistical significance of necessity and sufficiency. Moreover, since, for fuzzy set logic, every case has a membership value in a configuration, the problems of small sample size are much less severe than for crisp logic.

The calculation of the z-score requires the researcher to state a benchmark. Here the linguistic hedge suggests the choice of the benchmark proportionality. To ask, for example, if the observed proportion is significantly greater than "usually necessary" indicates a benchmark of .65. A benchmark of "very necessary" implies a value slightly greater than .7 benchmark. (The linguistic hedge of "very" is mathematically equivalent to squaring the membership value, as discussed earlier; the square of .71 is approximately .5, the cognitive anchor where a member is maximally more or less a member of the set of "very necessary" causes. We use the value of .65 in the following analysis.) Whereas these benchmarks may seem arbitrary (but no more arbitrary than the conventions governing questionnaire scaling such as a Cronbach alpha or significance tests), sensitivity analysis around the benchmark easily provides a way to assess robustness. In addition, sensitivity of measurement error can be examined by adjusting the diagonal to accept errors that differ by a stated percentage off the diagonal.

The determination of fuzzy sets proceeds, then, by statistically identifying necessary causes. Cases that reveal zero membership in the necessary causes are eliminated (by definition, they cannot satisfy the logical condition of necessity). Sufficient causes are then found by identifying causal configurations that statistically satisfy the requirement that their membership values are less than the causal outcome.[10] This analysis generates then a listing, or union, of sufficient configurations, conditioned on the initial identification of necessary

---

paradox in set theory. We empirically avoid these operations below.

causes. To achieve a global assessment of the statistical strength of the analysis, a membership score in the sufficient configurations for each case can be calculated. The comparison of this membership degree against the observed membership in the causal outcome serves to generate a test statistic to determine the significance of the classification success of the method.

Any cause that is individually sufficient is also sufficient jointly. (Proof is available on request.) Necessity of one cause does not mean, however, that two necessary causes are jointly necessary. However, any jointly necessary conditions are also individually necessary. (A proof is available on request.) It is thus justified to apply rules of Boolean absorption to fuzzy sets. Since the configuration Ab is a subset of the configuration of A (i.e. Ab is an intersection and hence a subset of A), the union of two configurations Ab and A logically implies that x will have a membership value equal to its membership value in A. Thus, Ab+A logically reduces to A.

For example, the statement that tall men must shave can be absorbed into the statement men must shave. To a great extent, this rule captures the meaning of a radial category. Peripheral members are absorbed into more basic representations of the category.

However, the rule of Boolean reduction does not apply. Since (B+b) equals max(B,b) and not 1—as in crisp logic, the crisp law of complements does not hold and Ab + AB does not reduce further. Fuzzy set analysis consequently loses some of the logical sharpness of the crisp method, since configurations do not easily reduce to more general and simpler causal factors.

This loss of sharpness is compensated partly by the statistical analysis that tests each configuration for significance. Since all cases (e.g. auto plants) are members *to some degree*

[10] Theoretically, if enough cases lie exactly on the diagonal, a cause can be found to be both sufficient and necessary.

in each configuration, each configuration has a sample size equal to the number of all plants in the sample.[11] This property greatly facilitates the application of statistical methods, as described above. The configurations that pass significance can then be minimized by the absorption rule that applies to both crisp and fuzzy sets.

The final step of the analysis then assigns cases to configurations by choosing the maximum membership value of that case in the minimized configurations. For example, an analysis of auto plant productivity might find that technology and human resource management constitute one configuration and technology and high scale form another. A given plant has a membership score of .4 in the first and .7 in the second (each score is derived by taking the intersection, or minimum, of the two practices constituting that configuration). The assignment rule would then assign this plant to the second configuration.

This reduction can obviously assign plants that are bad examples of a particular configuration. It makes little sense, for example, to claim that a given plant is characterized by high performance work practices when it belongs weakly to every attribute set that defines this configuration. This possibility conforms with a prototype theory of classifications whereby an ostrich is bad example of a bird. It also reflects a methodological weakness in fuzzy sets insofar that operations of intersections can assign members to classes that are not commonsensical. Lazarfeld (1937) offers, as noted before, a proposed solution to this type of problem by ruling out implausible combinations. (This intervention is broadly standard in statistical methodologies, such as in confirmatory factor analysis or model specification.) In a similar fashion, we propose to allow for the use of commonsense and theoretical intervention in two forms. First, in the interpretation of the configurations, we look at the "better" prototypical examples, that is, those cases that score .5 or more in a configuration.

---

[11] For the analysis of necessity, we lose cases whose outcome values are 0.

Secondly, to reduce the overall solution space, we check the *simplifying assumptions* that eliminate configurations that grossly violate theoretical and commonsensical relationships. As in the case of Boolean comparative analysis, the fuzzy set methodology faces the problem of limited diversity. Consider figure 2 that provides a two dimensional representation of operations on fuzzy sets. Imagine that the graph is divided into four quadrants from each of the midpoints at .5. The corners represent the crisp sets, and in this way, each quadrant is associated with a given crisp configuration. Limited diversity arises when there is no case in a quadrant. For Boolean analysis, limited diversity is obvious, as no case will show the configuration.

For fuzzy sets, since all cases have membership in all configurations, it is necessary to be especially careful to check that a causal configuration is not derived from an assumption that is not strongly justified by the empirical data. This verification is conducted by enumerating all the crisp sets and identifying those that have no cases with membership values greater than .5. This list can be used to isolate the combinations of factors for which there is little empirical evidence. This then poses the question if these combinations, that Ragin (2000) labels simplifying assumptions, are justified to play a role in deriving the minimized configurations. To check robustness, the researcher can check if these assumptions have been incorporated into the results of the sufficiency analysis. If this is the case, the researcher can either eliminate the simplifying assumption, which may change the results, or decide for theoretical reasons that the assumption should be retained. Both of these strategies have analogues in other methods. Econometrics often infers from the absence of a condition that decision makers did not choose this configuration because it was not profit maximizing. This provides information and can be used in the estimations (see Athey and Stern, 1999, for an example). The second strategy is more common and arises in multiple variable regressions when one factor is not significant, but contributes to the overall estimation. An advantage with the Boolean and fuzzy set methodologies is that the researcher can explicitly identify the simplifying assumptions used in the minimization and decide, based on theory or field knowledge, if they should be eliminated or retained.

Sample and Variables:

We apply the technique of fuzzy sets to identify bundles, or complementary practices, among technical and organizational factors affecting manufacturing performance in the world auto industry. The International Assembly Plant Study was sponsored by the International Motor Vehicle Program (IMVP) at M.I.T. Ninety assembly plants were contacted, representing 24 producers in 16 countries, and approximately 60% of total assembly plant capacity worldwide. Survey responses were received from 70 plants during 1989 and early 1990. These plants were divided into "volume" and "luxury" categories (the latter defined as plants producing automobiles with a 1989 U.S. base price of over $23,000), on the assumption that the production systems for these product types might differ substantially. This paper includes data from the 62 volume plants, whose surveys were more complete; because of missing data, only 57 plants are used for productivity and 45 for quality. The actual samples used in the logical analysis are 56 and 44, respectively, as the analysis assigns one plant in each sample a zero membership in the outcome and consequently eliminates it from the analysis.

Table 1 lists the distribution of the 62 volume plants by regional category. The proportion of plants in different regions corresponds closely to the proportion of worldwide production volume associated with those regions, with a slight underrepresentation of Japanese plants in Japan and overrepresentation of New Entrant and Australian plants, whose volume is low. Plants were chosen to achieve a balanced distribution across regions and companies, and to reflect a range of performance within each participating company, minimizing the potential for selectivity bias.

Questionnaire Administration and Data Collection

Questionnaires were sent to a contact person, often the plant manager, who distributed different sections to the appropriate departmental manager or staff group. Plants and companies were guaranteed complete confidentiality and, in return for their participation, received a feedback report comparing their responses with mean scores for different regions. All 90 plants that were contacted were visited by one of the researchers between 1987 and 1990. Early visits provided the field observations that became the foundation of the assembly plant questionnaire. Some of these plants were used to pilot the questionnaire as well. For the 70 plants that returned a questionnaire, the visit often followed receipt of the questionnaire, providing an opportunity to fill in missing data, clarify responses that were unclear or not internally consistent, and carry out interviews to aid the later interpretation of data analyses. When the visit preceded receipt of a questionnaire, this same follow-up process to improve data accuracy was carried out via phone and fax. We calculate membership degrees for both productivity and quality measures from the sample of plants for which there are usable outcome data. Some cases eliminated later due to missing data for the independent variables anchored the performance scaling at these extreme values; thus the ultimate membership scores for performance do not necessarily vary from 0 to 1.

As the measures are described in detailed in MacDuffie (1995), we supply only brief descriptions here.

Measures - Dependent Variables

Productivity. Productivity is defined as the hours of actual working effort required to build a vehicle at a given assembly plant, adjusted for comparability across plants by a methodology developed by Krafcik (1988). The productivity methodology focuses on a set of standard activities that are common across all plants in the survey, to control for differences in vertical integration. Since a large vehicle requires more effort to assemble than a small vehicle, adjustments are made to standardize for vehicle size. Adjustments are also made to standardize for the number of welds, which differs across designs and therefore affects headcount in the body shop.

This scale was fit to a [0,1] interval. Then, because high labor hours per vehicle indicates low productivity, we took the complement (i.e. subtracted the membership degree from 1) to create a reverse scale that indicates monotonic increases in productivity.

Quality. The quality measure is derived from the 1989 survey of new car buyers in the U.S., carried out by J.D. Power. The variable measures the number of defects per 100 vehicles. It is adjusted to reflect only those defects that an assembly plant can affect, i.e. omitting defects related to the engine or transmission, while emphasizing defects related to the fit and finish of body panels, paint quality, and the integrity of electrical connections (Krafcik, 1988). As with productivity, by taking the complement, we reverse scaled this measure.

Measures - Independent Variables[12]

Production Organization Measures. To measure the organizational logic of lean vs. mass production systems, three component indices were constructed -- Use of Buffers, Work Systems, and HRM Policies. The variables included in these indices reflect choices, based on fieldwork, about what items to include in the assembly plant questionnaire as well as statistical tests aimed at boosting the internal reliability of each index. Reliability tests are reported in MacDuffie (1995).

[12] We choose to work with scaled measures rather than each item; obviously, dimensionality would explode otherwise. It is possible to work out fuzzy ways to reduce these items; we relied upon our case knowledge to evaluate the scales.

Each of the three component indices is composed of multiple variables, described below. All variables are standardized by conversion to z-scores before being additively combined to form indices. Each variable in an index receives equal weight, because there was no clear conceptual basis for assigning differential weights. For ease of interpretation, a linear transformation is applied to the summed z-scores for each component index, such that 0 is the plant with the lowest score in the sample and 100 is the plant with the highest score. The validation of these indices is described in the next section.

i) Use of Buffers: This index measures a set of production practices that are indicative of overall production philosophy with respect to buffers (e.g. incoming and work-in-process inventory). A high score on this index signifies a minimal buffer "lean production" approach, and a low score, a large buffer "mass production" approach. It consists of three items:

ii) Work Systems: This index captures how work is organized, in terms of both formal work structures and the allocation of work responsibilities, and the participation of employees in production-related problem-solving activity. A low score for this variable indicates a work system with a narrow division of labor that is "specializing" in orientation, and a high score indicates a "multiskilling" orientation.

iii) HRM Policies: This index measures a set of policies that affects the "psychological contract" between the employee and the organization, and hence employee motivation and commitment. A low score for this variable indicates a "low commitment" set of HRM policies and a high score indicates "high commitment" policies.

Control Variables

The idea of control variables is standard in regression analysis to eliminate potential influences in non-experimental settings. For Boolean or fuzzy set analysis, they pose added dimensionality that can quickly complicate the logical inferences, especially for small data sets. We chose, therefore, to work with three control variables to capture technology, scale, and model age; to explore robustness, we also added part complexity.

Technology (Automation) The main technology variable, the automated percentage of direct production steps, captures the level of both flexible and fixed automation. For each functional area, a proxy measure for direct production activities was developed; see MacDuffie (1995) for details. Then a weighted average level of automation for the plant was calculated, based on the amount of direct labor each functional area requires in an average unautomated plant.

Scale. This is defined as the average number of vehicles built during a standard, non-overtime day, adjusted for capacity utilization. Overtime is not included in either production levels or hours worked, which adjusts for overcapacity situations.

Model Design Age. This is defined as the weighted average number of years since a major model change introduction for each of the products currently being built at each plant. This measure is a partial proxy for manufacturability in the assembly area, under the assumption that products designed more recently are more likely than older products to have been conceived with ease of assembly in mind. While older designs, by moving down the learning curve, could be associated with fewer hours per car, most evidence suggests that the benefits of more manufacturable designs outweigh learning curve gains (Womack, Jones, and Roos, 1990).

Parts Complexity. This measure is compiled from two subgroups of variables: parts or component variation and factors influencing the logistics of material and parts flow and the administrative/coordination requirements for dealing with suppliers. All these variables are scored on a 1-6 scale, where 1 is the lowest and 6 the highest complexity level. They are additively combined and the resulting index is rescaled from 0 to 100, as above.

Table 2 contains descriptive statistics for the variables used here. Means are based on the rescaling of each variable from 0 to 1, as required by fuzzy set analysis. The mean for productivity as transformed is roughly centered in the middle of this distribution. The control variable means reflect the fact that the predominance of plants have relatively high levels of automation and relatively young product designs. The mean for scale is relatively low because the largest plant, scored as 1, is an extreme outlier in terms of size; we discuss the effects of this outlier on the analysis below. Finally, means for the indices linked to lean production reveal that the use of lean buffers is most common in this sample, with a mean near .5, while the means for both the HRM and Work Systems indices are considerably lower.

As Table 2 also shows, both the variables capturing lean production (WORK, BUFF, HRM) and the control variables (SCALE, AGE, TECH) are significantly correlated with productivity. Indeed, the weakest correlation is with scale, suggesting that economies of scale are not such a dominant influence on labor productivity in this setting as it commonly supposed. Correlations among the three indices of lean production are also quite high, as the conceptualization of this overall production system would suggest. While plants with high scores for lean practices also tend to be highly automated and have younger products, they are not necessarily large; the correlations between the three production organization indices and scale are not significant. Scale and technology are strongly correlated, however, as both capture different aspects of capital investment at a given plant.

Analysis of cases

Fuzzy set methodology is a classifier technique that combines logic with the researchers' knowledge of the terrain. The search for the fuzzy sets of complementary activities involves first an analysis of necessity, then of sufficiency. If the analysis reveals any necessary conditions, this condition then appears in all configurations that pass the sufficiency test. We first calculate all 3^n-1 combinations for the variables. These variables include the controls (i.e. scale, technology, model age) and the organizational factors (i.e. new work practices, advanced human resource management practices, and lean buffers.) There are consequently 728 causal combinations to test. The test statistic for sufficiency compares the proportion of the times that the minimal value of a configuration (defined by the intersection operator) is less the value of the outcome (productivity or quality) against some benchmark. We use .65 as the threshold for sufficiency, as this hurdle resulted in the most parsimonious results. The causal combinations that pass this test are then submitted to an "absorption" algorithm to derive the minimal configuration.

We made two decisions to arrive at robust solutions. First, we squared the measures for productivity and quality. Squaring serves to accentuate the hedge "very", as noted earlier, and served to dissipate the bunching of outcome variables. A plant with a high productivity score is "very" productive. Second, we were sensitive to the potential that the inferential engine by which all permutations are taken and then tested for necessity and sufficiency might lead to outcomes that have low empirical and theoretical support. Thus, for a configuration evaluated as sufficient, we would like to verify that the conclusion was not reached by an inference from assuming a configuration to be empirically valid when the actual support is low. This error arises from the problem of limited diversity discussed earlier. We made then the following decision rule: for all simplifying assumptions (configurations for which the empirical support is weak), if two out of three production organization indices (WORK, HRM, BUFFERS,) were in a not-condition, we rejected this simplifying assumption and did not allow it to contribute to logical absorption. This decision rule resulted in a more parsimonious and robust set of solutions. We discuss the applications of this rule below.

Productivity Analysis:

Table 3 provides the baseline test for productivity-squared that includes the indices of Buffers, Work Practices, and HRM Practices as well as controls for scale of production, level of automation, and average age of the models being assembled. Recall that intersection is represented by multiplication (AB), whereas union is represented by (A+B). The necessary cause analysis indicates three necessary conditions ($p < .01$): not-scale, a low (young) product age, and a high level of automation. (The statistical test is one-tail, as we do not care about cases that fall below the benchmark.) While this result was as expected for product age and automation, it seemed unusual to find not-scale, i.e. a relatively low level of daily production, to be associated with higher labor productivity, i.e. fewer hours per vehicle. After all, the auto industry is generally regarded as the prototypical example of economies of scale.

Upon investigation, we found that the division of the sample into scale and not-scale categories was heavily influenced by the presence of a single outlier case. This plant, the largest in the world at that point in time, had a daily level of production more than four times the sample mean and 30% more than the second highest volume plant. This plant was also relatively inefficient, particularly in relation to its supposed scale advantage; it can in many ways be viewed as a prime example of the diseconomies of scale. Because of this outlier, the classification procedure is assigning membership in the set of "extremely large" and "not extremely large" plants. "Not-scale" as a necessary condition contains nearly 90% of the sample, all plants with scores above .5 in this set of "not extremely large" plants. Besides the outlier plant, five other plants have scores of above .5 in the "extremely large plant" category and hence don't meet the necessary condition of "not scale". It is worth noting that many plants in the "not-scale" subset are well above any threshold of minimum efficient scale, and operate with a production volume well above the world average; these are not low-volume plants, they are simply not "extremely large".

Exploring Complexity

We have emphasized that a primary advantage of Boolean or fuzzy analysis is the exploration of the effects of missing combinations, or combinations of low probability. We examined the simplifying assumptions involved in the sufficiency analysis. One such assumption included "not" conditions for two of the three indices of production organization, specifically not-buffers and not-HRM. According to this assumption, highly productive plants were associated neither with low levels of buffers (or inventory, repair space, utility workers) nor with high levels of commitment-inducing human resource management practices. Based on prior analyses of this data set and extensive fieldwork at these plants, we concluded that this particular assumption (and following our decision rule, any assumption that negated two or more of the production organization indices) was implausible, and we excluded it. After this exclusion, the sufficiency analysis for productivity generates three causal combinations.

The second configuration (not-scale, WORK, BUFF, AGE, TECHNOLOGY) contains 6 plants, five of which surpass the threshold value of .5. These plants are all located in Japan and most closely resemble the lean production ideal type. While their highest sufficiency score is in this configuration, four of the five plants also have a sufficiency score greater than .5 in the previous configuration. This suggests that all three of the production indices (HRM, WORK, and BUFF) are identified as sufficiency conditions for being a high productive plant in this grouping, beyond the necessary conditions of high automation levels and low product age. These results are very supportive of the consensual understanding of Japanese high performance work systems.

The third configuration (not-scale, BUFF, hrm, AGE, TECHNOLOGY) contains 33 plants and is the most geographically diverse group, ranging from the U.S. and Europe to Australia to Brazil, Taiwan, and Korea; it includes no Japanese plants or transplants. Only six of these plants surpass the .5 threshold. What characterizes these six plants is that they have pursued productivity through a different adaptation of the lean production model, namely a heavy emphasis on the reduction of buffers and a minimal emphasis on "high commitment" HRM practices or new work practices. The other plants in this category have low scores on various of these variables. Some have very low levels of automation, others build very old product designs, and many have very large buffers of inventory (which generates a low score on BUFF). Any of these could be the primary reason that these 27 cases are not identified in the set of "very productive" plants. These non-productive characteristics also frequently overlap; many of the plants in New Entrant countries have low automation, old product designs, **and** a production system reliant on large buffers.

The national diversity of these grouping also suggests two interpretations. The first is that the historical point in time when these surveys were collected reflected an incomplete diffusion. This interpretation is in line with the finding of the predominance of Japanese plants in the first and second configurations that satisfy the .5 hurdle. The second, and related interpretation, is that plants in other countries were still experimenting in the context of different national environments. Practices such as those related to teams were anathema to nations, as they challenged both union and firm control over the workplace. It is not surprising in this light that the third group shows a groping for new combinations that did not lead, however, to high productivity.

We undertook one sensitivity analysis to test the effect of choosing a .5 threshold for membership in a causal configuration. Changing the threshold to .4 adds two plants to the first configuration, no plants to the second configuration, and nine plants to the third configuration. These plants did not alter the substantive interpretation of the categories. The difference between applying a membership threshold of .4 versus .5 appears to be a matter of degree and not of kind. Plants with sufficiency scores above .5 are simply stronger members of the set of very productive plants. Therefore, we continue, in subsequent analyses, to apply .5 as the threshold for membership in a configuration.

Quality Analysis:

In order to identify high performance plants (defined as plants that are highly productive and have high quality), we turn next to the analysis of quality. Because we have only data on quality for 43 plants, we report in table 4 the productivity analysis for this smaller subset to test for robustness. The necessary and sufficient conditions are unchanged; indeed, the fit measure is identical. This smaller sample is used for the remaining analyses.

The results for the necessary conditions –which do not report here-- are the same as for productivity, although the significance level for the technology variable is somewhat weaker ($p < .05$ rather than $p < .01$). This is consistent with earlier analyses (MacDuffie, 1995) which found automation level was not strongly correlated with quality-- even though most high quality plants were highly automated, many high-automation plants had quite poor quality. We thus treat scale as a necessary condition, and let technology be determined by the sufficiency tests.

Exploring Complexity

For the sufficiency analysis given in Table 5, we excluded three simplifying assumptions, following our decision rule regarding the infeasibility of any such assumption in which two out of three production organization indices (WORK, HRM, BUFFERS,) were in a not-condition. Five causal configurations result from this analysis. The first configuration consists of a combination of lean buffers and work-related practices, such as problem solving. The next two configurations each contain two organization indices combined in different ways (WORK BUFFERS; WORK HRM; and BUFFERS HRM) along with the necessary conditions. The final two configurations both contain HRM; "not-work" is also included in the fourth configuration and "not buffers" in the fifth configuration. As with productivity, this analysis reveals differences in the extent to which plants with strong membership in the category of high-performing plants have implemented certain of the production organization policies of lean production. Whereas for productivity, plants with minimal buffers but more traditional HRM policies achieved respectable performance, the pattern for quality differs. Here it is high-commitment HRM policies that are most consistently associated with high level of quality performance; HRM appears in four of the five configurations. It is not a necessary condition because one configuration exists for plants for which WORK and BUFFERS are sufficient to predict quality without HRM being causally relevant.

High Performance Analysis:

We defined earlier high performance plants as those producing quality autos at high levels of efficiency. We took therefore the intersection (i.e. the minimum) of productivity and quality to form a single outcome called high performance. In Table 6, we examine plants that achieve high performance in both productivity and quality, that is, the "high-performance system" plants. The same three necessary conditions hold, with a significance level of $p < .01$ for TECHNOLOGY once again. We exclude only one simplifying assumption here, the same assumption identified in the productivity analysis.

This analysis identifies four causal configurations in the sample of 43 plants for which we have both productivity and quality data. Using the threshold of 0.5, we find only 12 plants are strong members of this category of high-performance plants. There are two plants in the first configuration, six plants in the second configuration, four plants in the third configuration, and none in the fourth configuration. This reduction in the number of plants is not surprising. Many plants are able to maximize either productivity or quality by trading off against the other outcome, but only the highest-performing plants are able to achieve both productivity and quality simultaneously.

There is not a high level of differentiation in performance among the configurations in this analysis. Some plants have identical membership scores in two of the three configurations; we treat these plants are members of both configurations in the performance analyses below. Still other plants have their highest membership score in one configuration, but have a membership score above the 0.5 threshold in another configuration, indicating a strong overlap in the influence of the sufficient conditions across these configurations.

The first configuration (not-scale WORK BUFFERS AGE TECHNOLOGY) contains six plants located in Japan. Four of these plants have identical scores for the third configuration (not-scale WORK HRM AGE TECHNOLOGY), and the other two also have strong membership (score > 0.5) in the third configuration. The first four plants confirm quite closely to the lean production ideal type. Their identical scores across these configurations reinforces the conceptual argument about mutual interdependence across the three aspects of production organization measured here, and the positive consequences of this interdependence for simultaneous achievement of high productivity and high quality. In contrast, the latter two plants are distinguished by a somewhat lower adherence to commitment-inducing HRM policies in relation to plants in the other two configurations.

The second configuration (not-scale work HRM AGE TECHNOLOGY) contains six plants that were all included in the first configuration of the productivity analysis (see Table 3). Four of these plants are Japanese transplants located in North America, and the other two are located in Mexico and Korea; the latter two also manufacture autos of Japanese design. In relation to the other two configurations, these plants have very high scores on HRM but lower scores on WORK and BUFFERS because they had only partially implemented on-line/off-line work team activities and Just-in-Time inventory policies at this point in time. For the Japanese transplants, these scores reflect not only the relatively young age of these plants but also the decision to make small group activities more voluntary than in Japan, and the necessity of stocking higher levels of inventory given the much greater geographical dispersion of the supply chain in the U.S.

The presence of plants in Mexico and Korea in this category of "high performance" plants suggests that product design may play some role in a plant's performance level, since superior design-for-manufacturability can make assembly both more efficient and less vulnerable to defects. But it also suggests that many of the production organization policies can be transferred successfully to settings in emerging economies, where automation levels are typically quite low. In such plants, high levels of worker training and high levels of selectivity for jobs viewed as quite desirable, in terms of pay, benefits, and job security, helps compensate for the generally lower level of education among the workforce.

The third configuration (not-scale WORK HRM AGE TECHNOLOGY), as mentioned above, contains four Japanese plants that are also members (with identical scores) of the first configuration. In contrast, the fourth configuration (not-scale not-buffers HRM AGE TECHNOLOGY) contains no plants with scores above 0.5, suggesting that plants with relatively large buffer stocks and only modest adoption of flexible work practices are not capable of achieving membership in the category of "high performance system" plants, even if their use of commitment-inducing HRM policies is extremely high.

In Figure 5, we graph the relationship between the observed (actual) high performance of a plant and the maximum value the plant takes in any of the four configurations. Given the classification system that seeks to align configurations and performance, it is not surprising the scores lie along the diagonal. The interesting aspect of the figure is the identification of how few plants and their associated best configuration are prototypes of high performance.

Tables 7 and 8 examine the performance means for the causal configurations identified in the productivity (Table 3) and "high performance system" (Table 6) analyses.

For productivity, the second configuration (not-scale WORK BUFFERS AGE TECHNOLOGY) has the best average labor hours per vehicle (17.5); plants in the other configurations require 51% and 80% more hours per vehicle, on average. The combined analysis of "high performance systems" given in Table 8 (which corresponds to Table 6 and to Figure 5), there is much less difference across the configurations. The four Japanese plants that possess membership in configurations one and three have the best combined performance, at an average of 19.1 hours per vehicle and 44 defects per 100 vehicles; by virtue of this combined membership, we know that they have high scores on WORK HRM and BUFFERS. Consistent with the earlier analyses, the configuration with the best quality performance (#2, at 43.2 defects per 100 vehicles) features high scores on HRM, while the configuration with the best productivity performance (#1, at 18.9 hours per vehicle) features high scores on WORK and BUFFERS.

Thus while there is no one single configuration of production characteristics associated with "high performance systems", lean production achieves performance advantages through the complementary interactions across two of three key areas of production organization: the management of buffers, the organization of work, and the human resource. These policies yield high levels of skill and flexibility in the workforce and induce high levels of performance. However, these results do not confirm the three-way interaction associated with the prototype of complementarities among all three dimensions of a production organization.

Discussion

The above results present a cross-section in the diffusion of practices that began in Japan.[13] High performance systems are generally associated with Japanese plants located in

[13] One of the referees asked for the use of firm dummy variables. Treating firm membership would, obviously, explode the dimensionality that we treat. More importantly, as all our data are the plant level and we are

Japan or outside. We did not find that all three work practices were complements associated with high performance, but we did find that two configurations of two of these three practices were complementary. The diffusion interpretation is further suggested by the plants outside of Japan that evidenced a greater variability in the degree to which they implemented these practices. Generally, higher performance plants were those that more successfully emulated "Toyotaism", that is the complementary implementation of these practices.

Comparing these results to MacDuffie (1995), we can identify a few important differences in methodological treatments and conclusions. Like the MacDuffie analysis, the fuzzy set methodology rejects a three-way interaction (though the latter approach induces this result simultaneously for productivity and quality). The set theoretic treatment of the cases allows configurations to be identified rather than a sub-set tested for the statistical significant of multiplicative interactions; see Ichinowski and Shaw (1997). Thus, we can see more clearly why, for example, that MacDuffie's tests of complementarities to achieve high quality were more problematic; clearly the interactions among practices are highly complex. We are also easily able to define high performance as the intersection of high productivity and high quality, and avoid separate tests for each. Finally, the analysis allows for an exploration of assumptions and the exploration of combinations (even if membership may be weak).

That a few combinations of practices can be assigned causality for the achievement of high performance systems across many countries suggests a transition period of experimentation, whereby diversity in configurations –whether planned or not—permitted an exploration of practices, to decouple old practices and recouple new ones. It is an important question, which these single cross-section data cannot answer, whether this transition lead to a convergence in a single set of best practice or in competing prototypes. We did identify one

---

holding constant the product market, we prefer to look at firm effects by looking at the membership of the plants in each configuration and then identifying firm, or nationality, effects.

"universalistic" element of small scale as a necessary condition (see earlier discussion regarding the topology of Delery and Doty, 1996); it is possible that in times of transition, smaller factories provide better experimental conditions. For this cross-section, we did not however find a single configuration, but several associated with high performance. In large part, these findings of multiple paths to a similar outcome restate the idea of "equifinality" proposed by Miles and Snow (1978). It will take a time series to sort out whether this multiple conjunctural causation is the product of multiple equilibria, multiple environments, or a snapshot in a historical process yet to converge to a best configuration.

Conclusion

The methodological treatment of complexity by fuzzy inference permits a cautious assignment of causal credit. In our application, we analyze an example where performance itself is two-dimensional (productivity and quality). We provide a method –the intersection of the two solutions—to show how causal assignment to configurations is still possible. This analysis is directly primarily at the understanding of the choice of capabilities, as we held the product market constant across the plants. Obviously, the full combination of capabilities and product market positioning requires a fuller treatment of a firm's strategic decisions.

The world does not generate enough experimentation to sort through all causal claims; the attribution of strategies or any entity to particular categories can only be made with fuzzy membership claims. Fuzzy set logic expresses this fundamental limitation on possible inferences. For even if we had full substantive understandings of the correct choice of strategy in particular environments, the complex interactions observed in practice poses two related problems of assigning membership and causality. The membership problem is, as we have seen, how to identify correctly the match between strategy and noisy environments. The causal credit problem is how do we know causality when observed or unobserved factors outside the model influence the strategy choice.

Our proposal is to recognize the inherent complexity facing researchers and decision makers and to develop inferential methods of exploration that render explicit the challenge of assigning membership and assigning causal credit. Rather than control for unobserved sources of variation, or lack of variation itself, we propose a systematic treatment of, one, how people (researchers and managers) think about the world through prototypes and, two, how causal relationships can be inferred through reduction and exploration of assumptions. The conclusions to this exercise reflect informed thought-experiments about possible worlds through exploratory data analysis. It is this avenue of analyzing worlds that may exist that is the most intriguing aspect of the application of logic to empirical cases. This perspective broadens the analysis from induction for the purpose of asserting general claims towards the disciplined examination of worlds logically possible but empirically and historically unobserved.

References

Adler, Paul, 1993, "The Learning Bureaucracy: New United Motors Manufacturing, Inc.", in Barry M. Staw and Larry Cummings (eds.), *Research in Organizational Behavior*, 15: 111-194, Greenwich, CT: JAI Press.

Altschuler, Alan, Martin Anderson, Daniel Jones, Daniel Roos, James Womack, 1984, *The Future of the Automobile*, Cambridge, MA: MIT Press.

Arora, Ashish, 1996, "Testing for Complementarities in Reduced Form Regressions: A Note," *Economics Letters*, 50:51-55.

Athey, Susan and Scott Stern, 1996, "An Empirical Framework for Testing Theories About Complementarity in Organizational Design," Cambridge, MA: NBER Working Paper 6600.

Becker, Howard, 1992, "Cases, Causes, Conjunctures, Stories, and Imagery," in Charles Ragin and Howard Becker (eds.), *What is a Case? Exploring the Foundations of Social Inquiry*, New York: Cambridge University Press.

Clark, Kim B. and Takahiro Fujimoto, 1991, *Product Development Performance : Strategy, Organization, and Management in the World Auto Industry*, Boston: Harvard Business School Press.

Cole, Robert E., 1985, "The Macropolitics of Organizational Change: A Comparative Analysis of the Spread of Small-Group Activities," *Administrative Science Quarterly*, 30:560-585.

Coriat, Benjamin, 1991, *Penser a l'envers. Travail et organisation dans l'Enterprise Japonaise*, Paris: Christian Bourgois Editeur.

Delery, John and D. Harold Doty, 1996, "Modes of Theorizing in Strategic Human Resource Management: Tests of Universalistic, Contingency, and Configurational Performance Predictions," (in Special Research Forum: Human Resource Management and Organizational Performance), *Academy of Management Journal*, 39: 802-835.

Ferguson, David and Ketchen, D. 1999, "Organizational configurations and performance: The role of statistical power in extant research," *Strategic Management Journal* 20(4): 385-395.

Flaherty, M. Therese, 1984, "Finance," in Daniel I. Okimoto, Takuo Sugano and Franklin B. Weinstein (eds.), *Competitive Edge: The Semiconductor Industry in the U.S. and Japan*, pp. 134-176, Stanford: Stanford University Press.

Hampton, James, 1997, "Conceptual Combination, " in Koen Lamberts and David Shanks (eds.), pp. 133-160, *Knowledge, Concepts, and Categories*, , Cambridge: MIT Press.

Holland, John, 1992, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, Cambridge, MA: MIT Press.

Hunter, Larry W. and John J. Lafkas, 1998, "Information Technology, Work Practices, and Wages," Wharton School, Philadelphia: Financial Institutions Center Working Paper 98-02.

Ichniowski, Casey, Thomas Kochan, David Levine, Craig Olson, and George Strauss, 1996, "What Works at Work: Overview and Assessment," *Industrial Relations*, 35:325-333.

Ichniowski, Casey and Kathryn Shaw, 2003, "Beyond Incentive Pay: Insiders' Estimates of the Value of Complementariy Human Resource Management Practices," *Journal of Economic Perspectives*, 17: 155-180.

Ichniowski, C., K. Shaw and G. Prennushi, 1997, "The Effects of Human Resource Management Practices on Productivity," *American Economic Review*, 87:291-313.

Klir, George J. and Bo Yuan, 1995, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Upper Saddle River, NJ: Prentice Hall PTR.

Kogut, Bruce and Nalin Kulatilaka, 2001, "Capabilities as Real Options", *Organization Science, Vol. 12, No. 6, pp. 744-758*.

Kogut, Bruce, John Paul MacDuffie, and Charles Ragin, 1999, "Prototypes and Fuzzy Work Practices: Assigning Causal Credit for Performance," Reginald H. Jones Center Working Paper.

Kosko, Bart, 1993, *Fuzzy Thinking: The New Science of Fuzzy Logic*, New York: Hyperion.

Krafcik, John, 1988, "Triumph of the Lean Production System," *Sloan Management Review*, 30:41-52.

Lakoff, George, 1973, "Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts," *Journal of Philosophical Logic*, 2:458-508.

Lakoff, George, 1987, *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*, Chicago: University of Chicago Press.

Levinthal, Dan, 1997, "Adaptation on rugged landscapes," *Management Science*, 43: .

Liker, Jeffrey, W. Mark Fruin, and Paul Adler, 1999, *Remade in America: Transplanting and Transforming Japanese Management Systems*, New York: Oxford University Press.

MacDuffie, John Paul, 1995, "Human Resource Bundles and Manufacturing Performance: Organizational Logic and Flexible Production Systems in the World Auto Industry," *Industrial and Labor Relations Review*, 48:192-221.

MacDuffie, John Paul, 1996, "International Trends in Work Organization in the Auto Industry: National-Level vs. Company-Level Perspectives," in Kirsten Wever and Lowell Turner (eds.), *The Comparative Political Economy of Industrial Relations*, pp. 71-113, Madison, WI: Industrial Relations Research Association.

MacDuffie, John Paul, 1997, "The Road to 'Root Cause': Shop-floor Problem-Solving at Three Auto Assembly Plants," *Management Science*, 43:479-502.

Miles, Raymond and Charles Snow, 1978, *Organizational Strategy, Structure and Process*, New York: McGraw-Hill.

Milgrom, Paul and John Roberts, 1990, "The Economics of Modern Manufacturing: Technology, Strategy and Organization," *American Economic Review*, 80:511-528.

Miller, Danny, 1996, "Configurations revisited," *Strategic Management Journal*, 17: 505-512.

Ouchi, William G., 1981, *Theory Z*, Reading, MA: Addison-Wesley Publishing Co.

Parker, Mike, 1985, *Inside the Circle: A Union Guide to QWL*, Boston: South End Press.

Pfeffer, Jeffrey, 1998, *The Human Equation: Building Profits by Putting People First*, Boston: HBS Press.

Pil, Frits, and John Paul MacDuffie, 1996, "The Adoption of High-Involvement Work Practices," *Industrial Relations*, 35:423-455.

Piore, Michael J. and Charles F. Sabel, 1984, *The Second Industrial Divide: Possibilities for Prosperity*, New York: Basic Books.

Ragin, Charles, 1987, *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*, Berkeley: University of California Press.

Ragin, Charles, 2000, *Fuzzy-Set Social Science*, Chicago: University of Chicago Press.

Rivkin, Jan, 2000, "Imitation of Complex Strategies," *Management Science*, 46, 6: 824-845.

Rosch, Eleanor, 1978, "Principles of Categorization," in Eleanor Rosch and BB Lloyd (eds.), *Cognition and Categorization*, Hillsdale, N.J.: Lawrence Erlbaum Associates.

Schonberger, Richard J., 1982, *Japanese Manufacturing Techniques: Nine Hidden Lessons in Simplicity*, New York: The Free Press.

Sengenberger, Werner, 1992, “Lean Production- The Way of Working and Producing in the Future?,” forum “Labour in a Changing World Economy”, International Institute for Labour Studies, International Labour Organization.

Siggelkow, Nikolai, 2001, "Change in the Presence of Fit: The Rise, the Fall, and the Renaissance of Liz Claiborne." *Academy of Management Journal* 44: .

Womack, James, Daniel T. Jones and Daniel Roos, 1990, *The Machine That Changed the World*, New York: Rawson Associates.

Zadeh, L.A., 1972, “A Fuzzy Set Interpretation of Linguistics Hedges,” *Journal of Cybernetics*, 2:4-34.

Table 1
Composition of Volume Assembly Plant Sample

Regional Category	*n*
Japan (J/J)	8
Japanese-parent plants in North America (J/NA)	4
U.S.-parent plants in North America (US/NA)	14
Europe (All/E)	19
New Entrants, including East Asia, Mexico Brazil (All/NE)	11
Australia (All/Aus)	6
Total	**62**

Source: International Assembly Plant Study

Table 2
Descriptive Statistics[14]

Variable	Mean	Standard Deviation	Pearson Correlation PROD	SCALE	WORK	BUFF	HRM	AGE	TECH
PROD	.5512	.2208	1.000	.306*	.587**	.502**	.529**	.558**	.685**
SCALE	.2289	.1841	.306*	1.000	.222	.238	.188	.137	.514**
WORK	.2202	.2712	.587**	.222	1.000	.651**	.652**	.304*	.292*
BUFF	.4698	.2655	.502**	.238	.651**	1.000	.586**	.542**	.382**
HRM	.3388	.3192	.529**	.188	.652**	.586**	1.000	.350**	.461**
AGE	.7293	.2108	.558**	.137	.304*	.542**	.350**	1.000	.525**
TECH	.6626	.2008	.685**	.514**	.292*	.382**	.461**	.525**	1.000

* Correlation is significant at the 0.05 level ($p < .05$)
** Correlation is significant at the 0.01 level ($p < .01$)

[14] As Quality reduces the size of the data set, we do not include the descriptive statistics for it here. They are available on request from the authors.

Table 3
Fuzzy-Set Analysis of Complements: Results for Productivity
(Number of cases: 56)

A. NECESSARY CAUSE ANALYSIS

Variable	N Cause Outcome	Observed Proportion	z	p
scale	48	0.86	3.11	0.001*
SCALE	6	0.11		
work	43	0.77	1.71	0.004
WORK	7	0.13		
buffers	31	0.55		
BUFFERS	31	0.55		
hrm	40	0.71	0.87	0.193
HRM	18	0.32		
age	12	.021		
AGE	50	0.89	3.67	0.000*
technology	20	0.36		
TECHNOLOGY	49	0.88	3.39	0.000*

b: SUFFICIENT CAUSE ANALYSIS*

scale HRM AGE TECHNOLOGY +
scale WORK BUFFERS AGE TECHNOLOGY +
scale BUFFERS hrm AGE TECHNOLOGY

*(Exclusion of Simplifying Assumptions:
scale WORK buffers hrm AGE TECH)

Test Proportion: 0.65
Significance level: < 0.01
Fuzzy Adjustment: 0.05

Table 4
Robustness Test for Productivity by Varying N
(Number of Cases: 43)

Sufficient Cause Analysis Shown Only:*

scale HRM AGE TECHNOLOGY +
scale WORK BUFFERS AGE TECHNOLOGY +
scale BUFFERS hrm AGE TECHNOLOGY

*(Exclusion of Simplifying Assumptions:
scale WORK buffers hrm AGE TECHNOLOGY)

Test Proportion: 0.65
Significance level: < 0.01
Fuzzy Adjustment: 0.05

Table 5
Results for High Performance Systems: Fuzzy-Set Analysis of Quality and Performance
(Number of cases: 43)

Sufficient Cause Analysis Shown Only:

scale WORK BUFFERS AGE +
scale WORK HRM AGE +
scale BUFFERS HRM AGE +
scale work HRM AGE TECHNOLOGY +
scale buffers HRM AGE TECHNOLOGY

*(Exclusion of Simplifying Assumptions:
scale WORK buffers hrm AGE TECHNOLOGY
scale work buffers HRM AGE TECHNOLOGY
scale WORK buffers hrm AGE TECHNOLOGY)

Test Proportion: 0.65
Significance level: < 0.01
Fuzzy Adjustment: 0.05

Table 6
Robustness Results for High Performance Systems
By Varying Excluding Assumptions
(Number of cases: 43)

Sufficient Cause Analysis Shown Only:*

scale WORK BUFFERS AGE TECHNOLOGY +
scale work HRM AGE TECHNOLOGY +
scale WORK HRM AGE TECHNOLOGY +
scale buffers HRM AGE TECHNOLOGY

*(Exclusion of Simplifying Assumptions:
scale WORK buffers hrm AGE TECHNOLOGY)

Test Proportion: 0.65
Significance level: < 0.01
Fuzzy Adjustment: 0.05

Table 7
Performance Means for Productivity Configurations

Configuration	**Productivity (hours per vehicle)**
Group 1 (scale HRM AGE TECH) threshold = .5	26.5
Group 2 (scale WORK BUFF AGE TECH) threshold = .5	17.5
Group 3 (scale BUFF hrm AGE TECH) threshold = .5	31.4

Table 8
Performance Means for "High Performance" (Productivity <u>and</u> Quality) Configurations

Configuration	**Productivity (hours per vehicle)**	**Quality (defects per 100 vehicles)**
Group 1 (scale WORK BUFF AGE TECH) threshold = .5	18.9	53.5
Group 2 (scale work HRM AGE TECH) threshold = .5	24.4	43.2
Group 3 (scale WORK HRM AGE TECH) threshold = .5	19.1	44.1

Figure 1
Strategic Choice in the Long- and Short-run:

Resource, Capability, and Markets

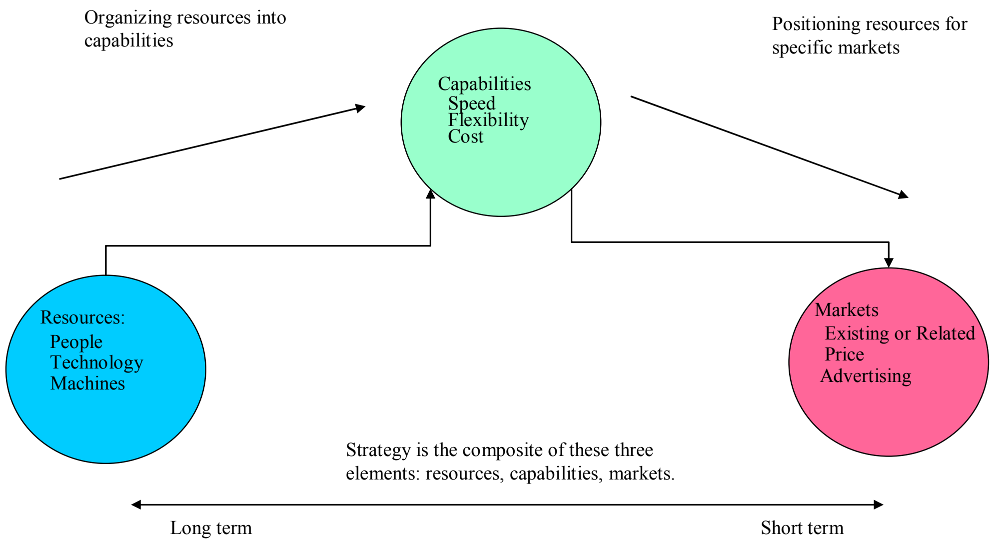


Figure 2
Subsets

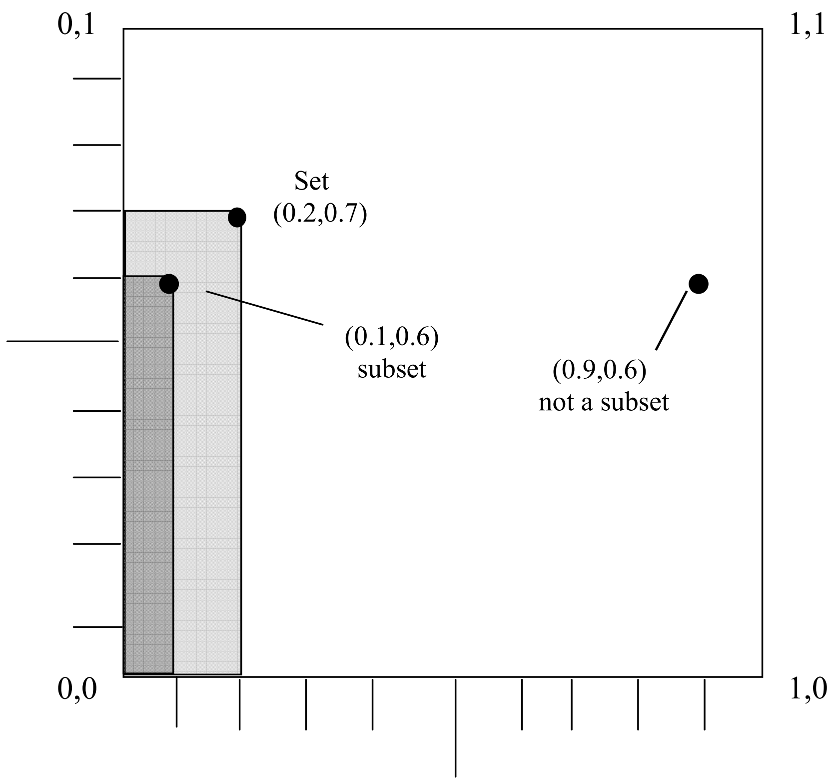


Figure 3
Plot of Fuzzy Relationship of Necessary Condition and Causal Effect

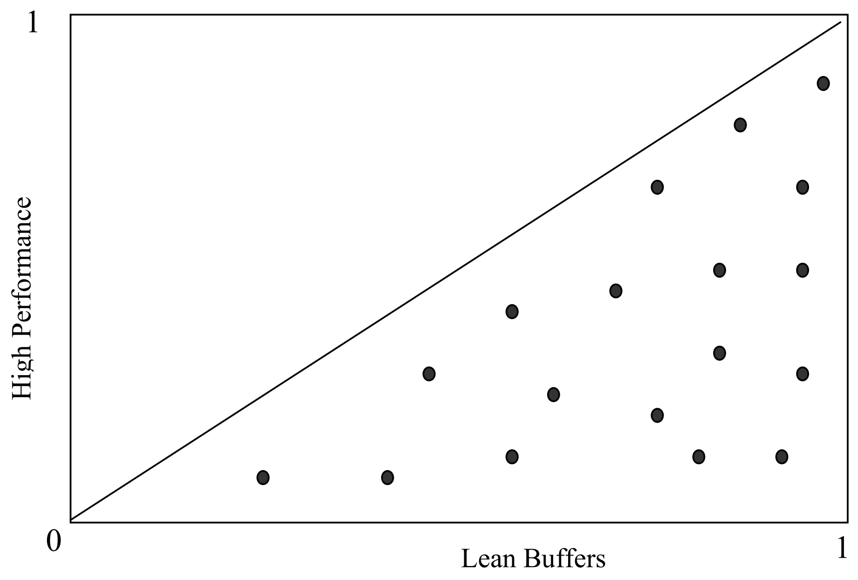
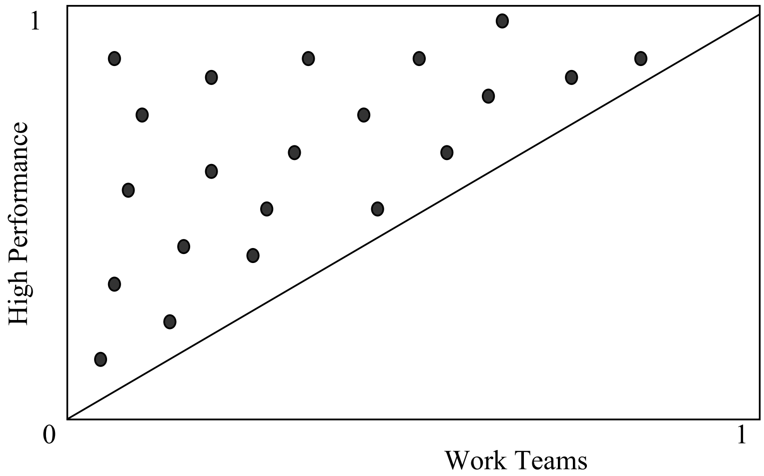


Figure 4
Plot of Fuzzy Relationship of Sufficient Condition and Causal Effect

a. Single sufficient cause



b. Two sufficient causes

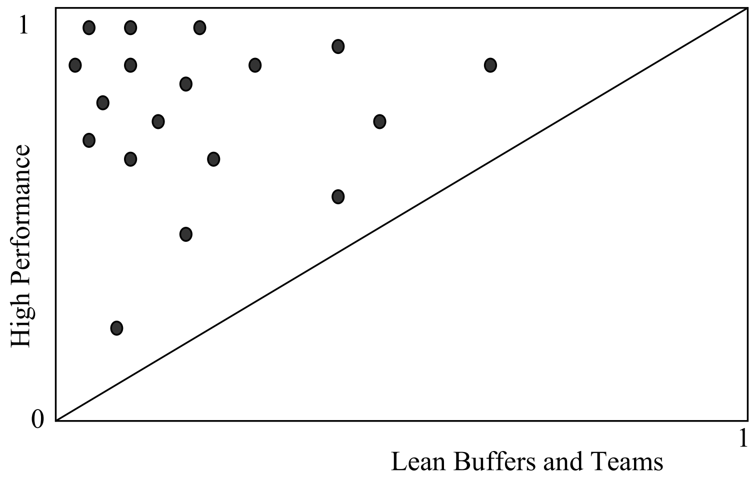


Figure 5

Scatter Plot of Actual and Predicted Maximum Membership Value for High Performance