Role of Variance in Manufacturing System Design: The Impact on Capacity Planning and Capital Investment Decisions

by

Joel W. Cain

B.S. Mechanical Engineering
Stanford University, 1995

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Master of Science in Mechanical Engineering
at the
Massachusetts Institute of Technology

August 1997
[September 1997]

Author...
Department of Mechanical Engineering
August 9, 1997

Certified by...
David S. Cochran
Thesis Supervisor

Accepted by...
Ain A. Sonin
Chairman, Graduate Committee

Role of Variance in Manufacturing System Design: The Impact on Capacity Planning and Capital Investment Decisions

by

Joel W. Cain

Submitted to the Department of Mechanical Engineering
on August 8, 1997 in partial fulfillment of the requirements for
the degree of Master of Science in Mechanical Engineering

Abstract

Variation in manufacturing systems is seldom seen as the principal performance driver. Instead, the anticipated behavior of most systems is obtained from models of expectation. However, as the system scope increases, having more dynamic variables, the compound effect of variation may cause models of expectation to lose relevance; in fact, under conditions of constant or non-trending change and adaptation, models of expectation cannot hold.

For systems where change is the rule rather than the exception, several well-defined methodologies have been advanced. System Dynamics with causal loop representation, Axiomatic Design with the recognition of coupling interdependencies, and Object-Orientation with the notion of encapsulation, are all approaches which capitalize on the nature of change for system optimization and improvement. Drawing upon the talents offered by each approach, this thesis develops a methodology called Object-Oriented Axiomatic Design (OOAD) which specifically addresses the role of variation in multi-objective distributed systems.

OOAD is applied to manufacturing capacity planning in an organizational context. Two case studies are presented, demonstrating the application of OOAD and illuminating a number of commonplace characteristics of variance that are functions not only of capacity planning, but also the manufacturing capital investment process. Specific recommendations are made to improve manufacturing system performance where variation is a principal driver.

Thesis Supervisor: David S. Cochran
Title: Assistant Professor of Mechanical Engineering

Acknowledgements

I would like to thank Professor David Cochran of MIT, and Mike Ledford and Larry Granger of Ford Motor Company. Without their sponsorship and continued support of this work, I would not have gained the insight, knowledge, and appreciation of systems-related issues.

Special thanks also goes to Ed Adams and Bob Gibson for their mentoring relationships during my stay at Ford. Debbie Christenson, Dick Duncan, Bob Goppold, Bob Jewell, John Kramer, Marilyn Maddox, and Mel Sillmon all gave freely of themselves to support this research. Each person has contributed in an unique way for which I am grateful.

My family and friends have been an incredible source of strength and encouragement during the past two years. In particular, I would like to thank Carol for her undying commitment and generosity in time and spirit that has carried me through.

This thesis is dedicated to my father, David, who has taught me to seek excellence in all things and to hold back at nothing.

In memory of Phil Pezzaglia.

Contents

List of Figures

Chapter 1

Introduction

1.1 Background and Motivation

Manufacturing capital investing is among the most difficult decisions made by commodity industries and industries otherwise characterized by heavy capital investment requirements. The predictive nature of capital requirements, coupled with long investment lead times, long asset lifetimes, and the insoluble nature of most investments, makes capital investing a risky endeavor. If this is misjudged or misguided, it can quickly lead to disastrous results. A thrust into a stagnant market, under- or over-capacitization, or an investment in process technologies rapidly eclipsed by new developments can bankrupt or cripple a company for years. Compounding the magnitude of this investment risk is the sheer complexity of the planning process. Not only are a diverse set of variables involved, such as forecast demand, competitive market assessment, future product offerings, and capital costs (each itself a function of current and future expected capacities), but the process itself is simultaneously centralized and decentralized. This paradox creates a tension that makes coordination difficult. On the one hand, the entire supply chain (both internal and external) must be capacitized to the same level for each product line. Whichever process step has the least capacity acts as the bottleneck. This renders useless excess investments elsewhere in the chain, causing costly lost or delayed sales when demand exceeds bottleneck capacity. In this manner, it is essential that the entire company and its supply chain march to the same tune. Decision making must be centralized, for if investment decisions were made in an independent manner, it is inevitable among the lot of decisions that some capacities will be excessive while others inadequate. The net result is wasted investment and lost sales. On the other hand, the capital investment process is decentralized; the type and cost of capital investment must be determined on an individual basis. Each manufacturing process step has its own cost and technology issues. Given this, how can overall capacity levels be determined centrally, when the costs of investment are understood only at a local level? This dichotomy requires the establishment of a well-defined process to manage manufacturing capital investing.

The bulk of academic literature on manufacturing capital investing falls into one of two categories. A number of authors focus on descriptive analyses of the nature and range of problems facing companies (see for example, Bakke [1], Hammosfahr [2], and Hayes[3]). Others provide specific optimization models that can be applied as decision making support to better help companies make educated and objective decisions (see for example, Manne [4], [5], Kalotay [6], and Luss [7]). While the different studies have focused on varied subjects in manufacturing capital investing, such as optimal timing of expansions under specific demand profiles (Luss [8]) or the determination of the optimal amount of manufacturing flexibility (Fine [9]), they all have one common thread: *variation.* If it weren't for volatility in the market and uncertainty in future requirements, capital investment decisions would be trivial. As it is, however, variation is assumed (either implicitly or explicitly) to play a critical role in investment performance, and most published works either illustrate the critical effect variation plays on organizations or provide mathematical models that specifically accommodate variance.

Despite the high level of development and validation of the body of published work, there are two key oversights, in the author's opinion, that have left much of this field of study unaddressed. Firstly, the role of variation has been principally treated as an external (to the company) driver of performance through exogenous demand variation or as operative in a single functional area rather than as an endogenous source resulting from independent decision making in an interdependent system. Secondly, the role of organizational process design and manufacturing system design have been largely ignored as potential solution areas. Instead, specific optimization tools that determine the optimum capacity volume level or best technology choice have been prescribed. In the author's opinion these oversights have ignored many of the causes of inferior performance as well as the range of possible solutions that can be employed.

In contrast to much of the contemporary work that addresses the role of variance in manufacturing capital investing, this work takes a different tack. The thesis of this work is that 1) the key driver of performance in a complicated, distributed process such as manufacturing capital investing is variation at an organizational level and 2) that if methods can be developed to characterize the role of variation in an interdependent system, proven optimization techniques such as those espoused in academic and business literature can be readily employed to eliminate adverse behavior, resulting in global performance improvement.

This work explores this role of variance in manufacturing capital investing in the following manner. First, a general methodology is developed to characterize, in both an analytic and descriptive manner, the role of variance in interdependent distributed systems. The lack of an adequate methodology has necessitated the development of an approach that draws upon three well-established methods: Axiomatic Design, System Dynamics, and Object-Oriented methodologies. Keeping with the development of the general methodology, generic solutions are provided. Some or all of these solutions will find specific application in particular situations. Then the developed methodology is applied in two case studies. As an integral element of the manufacturing capital investment process, the purpose of the capacity case studies is three-fold. First, the case studies demonstrate that performance of capital investments is as much a result of organizational behavior as it is a function of exogenous variation. Second, application of the methodology alongside proven techniques demonstrate how global performance improvements can be achieved. Third, the specific subject detail in the case studies highlight common problems alongside salable solutions that can be successfully applied to other business situations.

The author assumes that the reader comes to this work with a basic understanding of differential calculus and elementary statistics. Furthermore, this work draws upon Axiomatic Design, System Dynamics, and Object-Oriented methods. With respect to the subject matter covered, the reader is assumed to have a background in manufacturing systems, with a working level knowledge of capacity planning.

It is the hope that this work will provide the tools and understanding necessary for the reader to analyze and optimize manufacturing capital investment processes subject to variation. The use of interdependence matrices, in conjunction with the eight developed techniques to reduce adverse coupling between systems, deepens the reader's understanding and ability to control the effects of variance. The framework presented here, while far from complete, is general in nature, and therefore widely applicable to many systems. Furthermore, this developed methodology provides a formal foundation that can be used as a baseline for more advanced work.

1.2 Thesis Objectives

The objectives of this work are diverse. Unlike other research works that are either only theoretical treatises or experimental examinations, this project has sought fundamental development, validated by direct application. Pursuit of a dualistic approach to this research has resulted in the following objectives:

1. Demonstrate that adverse behavior occurs in systems despite the simultaneous effort of different areas to provide optimal performance.
2. Develop a general methodology to characterize interdependent behavior in systems. Rather than a provincial approach that has effective but limited application, the objective of this work is to provide a general framework that has wide application despite the tradeoff of decreased specialization.
3. Provide a mathematical representation that lends objective support to the understanding of ambiguous situations.
4. Provide a diagrammatical representation for abstraction of key concepts and communication of ideas.

5. Develop techniques for the optimization of multi-objective systems where control is decentralized.
6. Present the developed methodology in a cogent manner that enables it to serve as the basis for further development in this area.
7. Address specific problems in manufacturing capital investing and capacity planning due to variance using the developed methodology.
8. Provide a concise set of recommendations to resolve the problems identified in the case studies.

The remainder of this thesis strives to meet these objectives.

1.3 Thesis Outline

This work is divided into two major areas: theory and application. Chapter two presents Object-Oriented Axiomatic Design, the theoretical framework for characterizing and optimizing the behavior of interdependent distributed systems subject to variance. Chapters three and four apply this methodology to actual problems confronting two companies with unique operating environments. In both cases, the studies specifically address the role of variance in manufacturing capacity planning.

Object-Oriented Axiomatic Design is presented in four sections. Section 2.1 broaches issues related to variance in systems. Section 2.2 argues for the development of a formalized methodology. Section 2.3 develops this methodology, providing the structural representation, techniques for issue resolution, as well as a simplified example demonstrating application of the methodology. Finally, Section 2.4 summarizes the concepts presented in Object-Oriented Axiomatic Design, essential for the application of the methodology to practical situations.

Chapter 3 establishes the role of variance in manufacturing, by applying Object-Oriented Axiomatic Design to a company that manufactures temperature and pressure sensors. The first two sections, 3.1 and 3.2, provide background, while sections 3.3 and 3.4 characterize and optimize the behavior of the system, respectively. Concluding remarks are presented in 3.5.

The application of Object-Oriented Axiomatic Design to the role of variance in manufacturing capital investing involves capacity planning at Ford Motor Company, the case study presented in Chapter 4. Like the first case study, the work performed at Ford has been broken down into three general areas. The first sections, 4.1 – 4.4, provide an overview of the business and present an executive summary of case findings. Section 4.5 characterizes the system from the Object-Oriented Axiomatic Design perspective. Section 4.6 applies techniques developed in Section 2.3 to the system characterization of Section 4.5, with the goal of improving overall system performance by eliminating unwanted behavior. The final recommendations to Ford are presented in Section 4.7, and concluding remarks are made in Section 4.8.

The thesis work developed and applied in the previous sections is summarized in Chapter 5. This section concludes by identifying areas for further research and development.

Chapter 2

Object-Oriented Axiomatic Design: A Synthesis of Axiomatic Design, System Dynamics, and Object-Based Methodologies

2.1 Overview

Recently, systems-based approaches to problem solving have received increased attention by the engineering and business communities. This has been spurred by an increasing number of product and process related problems that require a systems rather than functional perspective for solution.

Object-Oriented Axiomatic Design is a general methodology that has been developed to address the role of variation in systems. While this methodology has been advanced for the purpose of characterizing and optimizing the role of variance in manufacturing capital investing, it is believed that the concepts presented here are widely applicable. OOAD draws upon the well-established work of System Dynamics, Axiomatic Design, and Object-Oriented methodologies widely used in software development. These concepts were pioneered by such luminaries as Jay

Forrester (System Dynamics), Nam P. Suh (Axiomatic Design), Grady Booch, Peter Coad, J.Rumbaugh, and Ed Yourdon (Object Oriented methodologies).

This methodology is the cornerstone of two case studies which focus on systems-related issues in manufacturing capacity planning. OOAD provided the analytical framework, the diagrammatical representation, and the applicable techniques for system optimization in each of these studies. The studies can be found in Chapters 3 and 4 of this work.

2.1.1 Outline of OOAD Development

This chapter presents Object-Oriented Axiomatic Design over the span of five sections. Section 2.1 highlights seven issues in systems characterization related to variance. These issues are common to many systems, but are specifically relevant to organizational systems. Each of the issues addressed in 2.1 played an integral role in the development of OOAD. Section 2.2 provides the motivation for the development of a formalized methodology, while Section 2.3 gives the objectives, structural representation, and multi-objective optimization methods necessary for application of OOAD to actual situations. To put closure on the section, a simple case example is presented to demonstrate application of the methodology. Because the expository nature of the OOAD method presented in Section 2.3 does not lend itself to quick reference, a summary of the main concepts and techniques are given in Section 2.4. Finally, a brief review of Axiomatic Design, System Dynamics, and Object-Oriented methods can be found in the Appendix, Section 2.5.

2.1.2 Issues In Systems Characterization Related to Variance

There are a number of issues that must be addressed if the role of variance in systems is to be effectively characterized. These are listed below and discussed in detail in the sections that follow. Each one plays a specific role in the behavior of systems.

- Exogenous and Endogenous Sources of Variation
- Dynamic Complexity
- Differential Delays
- Decentralization of Control
- Coupling Interdependencies
- Stochastic Variance
- Propagation of Variance

2.1.2.1 Exogenous/Endogenous Sources of Variation

Principally, there are two distinct types of variation in systems. Endogenous sources of variation are generated from within the system of interest. Exogenous sources of variation can be considered as external influences that affect the given system.

In practice it is sometimes difficult to assign variation as exogenous or endogenous. The difficulty lies in the fact that systems, by definition, are comprised of an assemblage of objects united by some form of regular interaction or interdependence.[1] Thus, changes made by one object in the system may affect its environment, which, in turn, may react to affect the original

---

[1] Webster's New Collegiate Dictionary, 1961, Cambridge, MA: The Riverside Press.

object. In this situation, the variation itself is endogenous despite the fact that the change to the system came from the environment. For example, if a company chooses to add capacity to a manufacturing line, it may be able to lower the price of the product, thereby stimulating demand and taking market share from a competitor. This change in the product demand is clearly from an endogenous change – the change in price. But, what happens if the competitor of this company learns of this capacity action and decides to counter by adding capacity of its own? If the other company does this, the original company's demand may fall despite its price reduction. Is this an exogenous form of variation, caused by an independent decision on the part of the other company or is it an endogenous form of variation, resulting directly from the original changes the first company made?

Causation is often difficult to assign. However, to the extent that endogenous sources of variation can be identified, they can be modified since they can be considered "owned" by the object or group that causes them. Likewise, exogenous sources of variation can be treated. Just as air conditioning systems in buildings are used to insulate people from environmental temperature variation, companies or organizations within a firm can employ techniques to protect themselves from unwanted but uncontrollable variation. These techniques will be discussed in detail in Section 2.3.3.

2.1.2.2 Dynamic Complexity

Often the interactive behavior in systems is highly nonlinear. This nonlinearity deprives people of their intuition of systems. Typically people assume systems to be linear. People will take the observations of a cause-effect relationship and project this onto future changes. Inventory control in manufacturing systems is a good example. When the system has high levels of inventory, a reduction of ten percent will linearly decrease the system throughput time by ten percent (Little's Law). However, as the system begins to get stretched thin, the buffering effect of inventory becomes increasingly insufficient. Variances in the manufacturing process cause stockouts and delays. Thus, when inventory is stretched thin, a further reduction in inventory of ten percent will actually cause average throughput times to increase rather than decrease. Unfortunately, the existence of this nonlinear behavior cannot be observed when inventory levels are high. This means that a person will continue to project the linear effects of inventory reduction for each subsequent reduction even though the underlying system contains a significant nonlinear structure.

2.1.2.3 Differential Delays

Differential delays substantially increase the complexity of systems behavior. Learning is difficult when the response to a change is separated from the change by a large period of time. This results in two effects. Either the person or group causing the change will overshoot the objective by not recognizing the existence of the delay or causal links will be inaccurately assigned. When searching for an explanation for a change, the natural tendency is to assign causation to things that are close in both space and time. When different systems have different time constants and multiple changes are made simultaneously, it becomes quite difficult to properly understand the dynamics of the system.

2.1.2.4 Decentralization of Control: Local Optimization

Increasing complexity in systems drives an age old problem: what to do when the complexity of a situation becomes greater than what any single one individual can handle. In all cases, the work must be distributed by bringing in more people. However, this does not mean that control is distributed. In fact historically, and to some degree today, control remains centralized in large organizations. Centralization of control ensures that all of the right decisions are made by those who have the "complete picture". By aggregating information upward through the control structure, leaders reduce the effective complexity of the system to something that can be handled by single individuals.

However, there is a drawback to centralized control: *time*. Companies having many levels of review structure are slow to respond to market movements. It takes time to send condensed information to those who make the decisions, time for those individuals to make the decisions, and time for the action steps to be passed down to subordinates. This processing time is uncompetitive in today's rapidly evolving market. As a consequence many companies have had to decentralize their control structures, allowing people at the "front lines" with expertise in only certain functional areas to make key decisions [10]. While expedient from the perspective of time-to-market and closeness to the customer, these decisions often sub-optimize corporate performance due to local, rather than global control structures. In Axiomatic Design parlance, these side effects are said to result from *coupled interdependencies* [11].

2.1.2.5 Coupling Interdependencies

Consider two independently controlled entities that interoperate as part of a larger system. These two subsystems (whether they are people, machines, software modules, or companies) are considered interdependent if changes made to one affect the performance of the other. Of course if this were not the case then they would not be part of the same system since a system is defined by the *interaction* or *interdependence* of two or more objects[2]. Moreover, these entities are said to be *coupled* if, in the pursuit of an objective, one entity affects the performance of one of the other entity's objectives [11]. This coupling can be either advantageous or adverse depending upon the circumstance. Advantageous coupling occurs when the improvement of state of one entity also benefits the other. Adverse coupling is the converse. This will be discussed further in Section 2.3.2.8.

2.1.2.6 Stochastic Variance

Stochastic variance is the variance associated with a stochastic process. A stochastic process is a particular function of time where each point in time along the path of the function is a random variable subject to a mean and variance. Another way of describing a stochastic process is the combination of a signal and noise. The signal is the time value of the stochastic process when the probabilistic variance is suppressed. The noise, on the other hand is the statistical variation of the stochastic process from the underlying signal. For example, product demand may be modeled as a stochastic process where the underlying demand signal is a particular cyclical function of time based on the economy. Deviation from this basic behavior trend is the random

---

[2] Webster's New Collegiate Dictionary, 1961, Cambridge, MA: The Riverside Press.

or probabilistic component. Therefore changes in demand are due to 1) fundamental changes in the demand structure over time and 2) noise which is the random deviation from the basic demand structure.

While in practice few truly random events ever occur, the confluence of innumerable changes may make the complete assignment of causality a hopeless undertaking. In this manner, causality is only assigned to the underlying signal. Changes to this underlying signal denote coupling, either exogenous or endogenous, depending on the source. Furthermore, as previously described, this coupling can be either advantageous or adverse. By contrast, the "noise" is unsigned (or random) variance. This is what has been termed *stochastic variance* [12]. The expected value of the stochastic variance is the underlying signal at any particular time, t. Thus, as a component of the stochastic process, stochastic variance has an expected value of zero but positive statistical variance (second moment).

Taken together, coupling interdependencies and stochastic variance form a spanning set [13], i.e. there exist no other identifiable sources of variation. Coupling interdependencies are the signed (either positive or negative) sources of variation while stochastic variance is the unsigned, or random fluctuation affecting a system in a non-trending manner.

Given that stochastic variance has been adequately defined it is worth considering its effect upon a system. Despite the fact that stochastic variance is unsigned, i.e. there is no long-term average expected change in either the positive or negative direction, its effect on the system is not benign. Stochastic variation affects system performance in three ways:

1. Variance increases the dynamics in the system, increasing the complexity and obfuscating the true drivers of change.
2. Variance creates waste; change from one state to another and then back again yields no net benefit to the system yet expends time, energy, and resources.
3. Variance reduces system reliability and dependability, increasing risk.

Increased variance makes system behavior more chaotic. Ideally, people either in business or in their personal lives want to only react to fundamental changes. When these fundamentals are obscured by noise, people either react to the wrong thing or are unsure what to react to at all. Increased waste is the second side effect of higher variances. This is self-explanatory considering that any expenditure of effort that does not create value is waste. Finally, the third adverse effect of stochastic variance is that stochastic variance increases risk. Increased variance increases risk for two reasons. First, increased variance makes the system behavior less deterministic. This decreases the certainty of the value of the system at any given point in time. Secondly, higher variance increases risk by virtue of the time dependency of other interacting systems. Since systems themselves are dynamic, having to respond to change, the failure of a key variable not being at its expected value at a particular point in time may critically affect the performance of the system. Therefore if a system is dependent on either of two variables having the same expected behavior but one subject to higher stochastic variance than the other, the higher variance variable will put the system at greater risk.

Increased Risk Through Decreased Reliability: An Example in Statistical Process Control

The first case of increased risk caused by decreased reliability is best illustrated by an example of statistical process control on a manufacturing line. Consider two different sheet metal stamping machines that can be used to produce a particular part with critical dimension, ***a***, that has a design tolerance of ±.003 inches. The first machine is highly repeatable with a 0.1% probability of dimension ***a*** being greater or lesser than the .003" tolerance. The second machine is not as repeatable, yet it too can meet the .003" tolerance 99.9% of the time. The less repeatable machine is expected to produce the same number of defective parts as the first machine because it has a different distribution profile. This is shown in Figure 2-1, below.

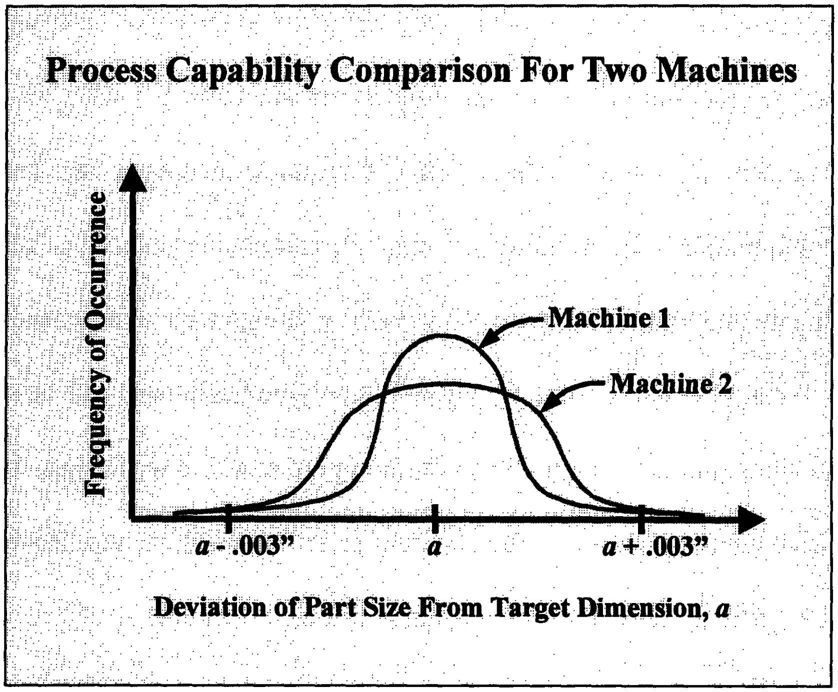


Figure 2-1: Process Capability Comparison For Two Stamping Machines

Since both machines have the same probability of making defects (parts outside the allowable tolerance) it would seem that the company should be indifferent regarding which machine should be used. However, when the sample dimensions, ***a***, are plotted on control charts, the second machine will demonstrate much greater variance (evidenced by a greater standard deviation of size) between the control limits than the first machine. Since the first machine has a much smaller range of variation, shifts of the mean or variance can be much more readily identified, giving the operator more time to correct the problem before defective parts begin to get produced. For the other machine, it may not be *until* defective parts begin to be produced that the problem is identified. The comparison of these two stamping machines provides an example of how, in this case, higher variance of a critical dimension makes it more difficult to determine the state of the machine. Because the condition of the machine is less fully understood, confidence in its reliability is decreased and therefore it cannot be considered as reliable or trustworthy for use.

Increased Risk Through Decreased Dependability: The Security Market Line

The second source of risk through increased variance is due to decreased dependability. This adverse effect of variance occurs in many situations, but is most formally treated in financial markets. The Security Market Line (SML) is widely accepted as a model of investors' risk aversion. The Security Market Line plots the relationship between the magnitude of fundamental variance and expected financial returns. This relationship is shown in Figure 2-2.

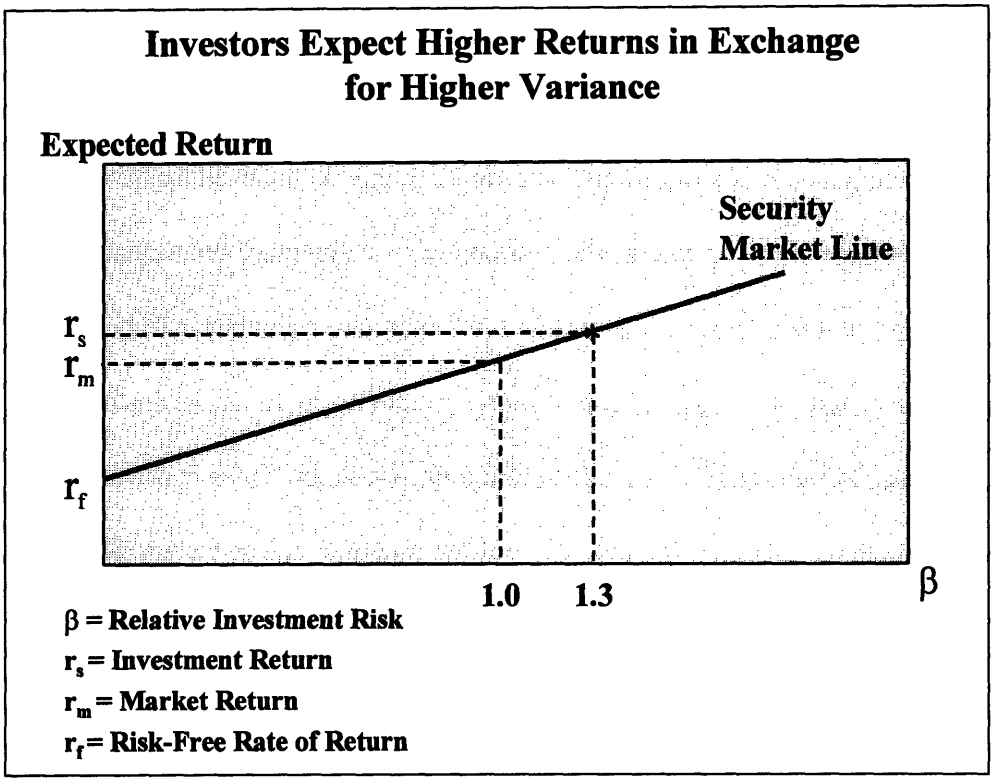


Figure 2-2: Relationship Between Risk and Return: SML

When the asset value of an investor's security is subject to a much higher degree of non-diversifiable variation, the asset is considered more risky and the investor requires a higher expected financial return for the higher risk borne. This risk is measured against the aggregate market risk as a reference. The risk of an individual security is determined from the following relationship,

$$\beta_S = \rho_{S,M} \cdot \frac{\sigma_S}{\sigma_M} \qquad (2\text{-}1)$$

where the variance (standard deviation) of the security, σ_S, is compared with the market volatility, σ_M. Then, only the undiversified component is kept by multiplying the resulting ratio by the correlation of the two movements, $\rho_{S,M}$. The uncorrelated component of variance is dropped since investors can hedge by investing in other securities with other variances.

The principal reason investors require higher returns in exchange for higher variance is due to the time-based dependency they have with these assets. Since the higher variance security will deviate further from its expected value (high or low) than the low variance security, its particular value at any point in time, ***t***, is known with less certainty. This adversely conflicts with the investor's need to have funds available at any particular point in time for the purchase of a car or home, for example. Since the value of the security with the higher variance is not known with as much certainty at any point in time (despite the consistency of its long-term average value), it is a less dependable source of funds. Thus, the decreased dependability due to increased variance makes this type of investment less attractive unless some sort of compensation is offered.

Despite the definitions for stochastic variance, it remains a more abstract concept than that of coupling interdependencies. The fact that causality cannot be specifically assigned makes its presence all the more insidious in organizations. Since variance is ubiquitous among organizations, it is worthwhile to illustrate its adverse effects through a specific example.

Example: Stochastic Variance Due to Cash Flow Volatility

Consider a volatile revenue cash flow structure from the perspective of a corporate controller's office. The company under consideration is a toy company that has been a solid performer in its market for more than ten years with a consistent annual growth rate of close to 10%. Despite the steady growth, quarterly sales have been highly variable – regularly deviating by 50% or more. New product introductions tend to be hit or miss and expansion into global markets has brought mixed results. Both of these effects have contributed to the volatility of the revenue stream. While some of this variation was induced by particular strategies employed by the company, the bulk can be attributed to the fickle nature of the toy market and therefore can be considered an exogenous source of variation. Since this variation is not attributed to a specific cause and low demand in one quarter is expected to be offset by high demand in a subsequent quarter, the interdependency between changes in demand and the company's profit objectives cannot be considered a coupling.

So why then, should the company care whether its revenue stream deviates by ten or fifty percent on a quarterly basis if its annual growth remains at 10%? The reason is threefold:

- The company must retain a higher cash reserve if the variance is 50%
- Changes in sales revenue will make revenue a poorer indicator of corporate performance if the revenue variance is 50%; it cannot be discerned whether a drop in sales is due to something gone awry internally or just a vagary of the market
- The company must have a lower debt/equity ratio if its revenue variance is 50% instead of ten percent

The company must maintain a higher cash reserve if the revenue variance is 50% because a decrease in sales will necessarily be offset by cash in order to cover the relatively constant business costs. This effect is sub-optimal since a smaller portion of the company's finances can be put to productive use by the firm. The second reason why the company would prefer a smaller revenue variance is that it makes revenues a better predictor of corporate performance. In fact, this is a chief reason why companies with more steady revenue streams (such as utilities) can be more highly leveraged. If the company has a high degree of revenue volatility, it must maintain a lower debt/equity ratio if it is to maintain solvency throughout downturns (alternatively it can retain a high debt/equity ratio, but at the cost of increased risk). The drawback to lower debt/equity ratios is that it becomes more difficult for the company to finance growth and capitalize on its success.

As a result of these three reasons, the toy company is better off if it can achieve similar earnings without the same degree of revenue variance. Unfortunately for this company, the revenue variance is stochastic. There is little it can do short of changing its target market segments to reduce the magnitude of this variance. The best thing for the company to do is to ensure that this revenue variance is not propagated throughout the organization. Employee morale, project budgets, forecast sales, etc. would be served no favors if their values marched in tune with demand. As the following section discusses, propagation of variance should be minimized wherever possible.

2.1.2.7 Propagation of Variance

Propagation of variance occurs any time stochastic variance is propagated through a system. If one process reacts to unnecessary change and as a consequence of its actions, causes another process to go through unnecessary change as well, variance has been propagated. Propagation of variance should be distinguished from a series of change events resulting from coupling interdependencies. An action in one part of the system may set off a number of other changes as the rest of the system reacts to the initial change. If the couplings are adverse, it is unfortunate that the rest of the system should have to respond to such a change, but such is the nature of coupling. If the system is to be improved, the change should not be restricted, but rather the coupling needs to be eliminated. By contrast, the propagation of variance is in itself change that is unnecessary and unwanted. To improve the system, this variation itself must be reduced or eliminated.

2.2 Motivation for a Formalized Methodology

The pervasive yet abstract nature of variation in systems makes it necessary but difficult to develop tools that characterize and optimize the dynamic behavior of systems. Variation can never be observed by direct point-in-time observation. Furthermore, variance in systems adheres to no functional boundaries. The role of variation in manufacturing systems involves numerous factors, each stemming from a different discipline. Multiple functional organizations within a company such as marketing and product development play roles. So do other areas such as personnel, competitive actions in the market place, and the economy also contribute to the role of variance in manufacturing systems. This expansive nature of variation and lack of adherence to a particular functional discipline increases the complexity of the analysis and precludes most existing methodologies due to lack of generality.

Because most systems in business or society are subject to complex, uncertain dynamics that strongly affect system performance, it is essential that they can be characterized in such a way that key relationships can be easily and objectively represented. Most systems require the cooperative effort of multiple individuals to enforce change. The need to promote universal understanding of the role of variance in systems provides the impetus for the development of a formalized methodology that can characterize these relationships.

2.2.1 Complexity of Systems

For the purposes of this work, a complex system is defined along two dimensions. A system is said to be complex if 1) more than one person is necessary to fully understand the system behavior, i.e. no single individual can know all the key operatives associated with the particular system and 2) multiple individuals can and do act autonomously in a manner that affects system performance. Both of these conditions are necessary for a system to be considered complex. In most cases though, if one condition is met the other will be satisfied as well. If more than one person is necessary to understand the system, it is usually the case that more than one person is free to act on the system.

Complex systems require explicit characterization. Process or system improvement is achieved only if, at a certain level of abstraction, each relevant individual understands how his or her role in the system affects the greater whole. Unless some level of global understanding is achieved, it will be impossible to improve the system. Each individual will have his own mental model of the system from his own frame of reference. A methodology providing explicit characterization enables the dynamics of a system to be understood on a global level in such a way that individual biases can be filtered out and consensus buy-in can be achieved. This allows steps to be taken that improve overall system performance.

2.2.2 Existing Methodologies Inadequate

Numerous methodologies have been developed for the design and analysis of products and systems. Such examples include: Structured Analysis and Design Technique (SADT), Axiomatic Design, System Dynamics, and Object-Oriented Design and Analysis. Each methodology has strength in its own right, but none have been applied to address the dynamic behavior resulting from the relationship between objectives and control in complex systems. Rather than attempt to develop an entirely new methodology, a synthesis of the most relevant concepts and techniques from each of these techniques has provided the basis for what has been termed *Object-Oriented Axiomatic Design.* Object-Oriented Axiomatic Design (OOAD) draws from System Dynamics, Object-Oriented Methods, and Axiomatic Design to provide an approach that can be used to improve the performance of complex systems.

2.3 Object-Oriented Axiomatic Design

Object-Oriented Axiomatic Design is an analysis and optimization approach that specifically addresses the nonlinear dynamics and disequilibrium found in distributed, interdependent systems. OOAD draws liberally from Axiomatic Design, Systems Dynamics, and Object-Oriented methodologies. A high-level review of these approaches can be found in Section 2.5.

2.3.1 OOAD Objectives

There are five key objectives to OOAD development. These are to:

- Explicitly recognize the boundary interfaces present in distributed systems
- Formally capture the dynamic behavior between interoperating groups within a system
- Clearly map the relationship of decision variables to objective states, identifying coupling interdependencies
- Succinctly abstract and communicate key concepts by means of diagrammatical representation
- Significantly improve system performance and organizational alignment through the application of well-defined techniques

The developed OOAD methodology presented in the following sections attempts to meet all five objectives in a logical and cogent manner.

2.3.2 Structural Representation

This section provides the structural representation of Object-Oriented Axiomatic Design. Presentation of the approach is broken down into ten sections. While each section builds upon the previous sections, each section discusses an unique attribute of the OOAD methodology. Development of OOAD begins with the representation of the most basic feedback loops in systems: balancing and reinforcing loops. This section, 2.3.2.1, is founded upon the System Dynamics methodology for systems characterization. The next section, 2.3.2.2, establishes the relationship between interdependent loops in systems. The mathematical representation allows the causal relationships between control variables and objective states to be explicitly defined. This work leads naturally to Section 2.3.2.3, which introduces the Axiomatic approach that enables the concise representation of larger systems. Section 2.3.2.4 is a critical link in OOAD. This section establishes a direct mathematical mapping between causal loops and design matrices. An example of the characterization of a system and the mapping between causal loops and design matrices is given in Section 2.3.2.5.

Sections 2.3.2.1 through 2.3.2.5 have developed the basic building blocks of the Object-Oriented Axiomatic Design approach. These basic building blocks form the basis of the next large step in OOAD development: *encapsulation*. Seldom do groups in systems operate in isolation. Often, the greatest problems resulting from the dynamic behavior of systems are not self-contained in autonomously operating groups. Many problems can be easily optimized within the control space of a group. Rather, the most difficult problems occur between different groups within a system. The degree of encapsulation is a measure of the independence between distinct control structures. The description of encapsulation, extension to the notion of boundaries, with a development of the Bounded Interdependence Matrix, and subsequent coupling interdependencies across boundaries, are presented in Sections 2.3.2.6, 2.3.2.7, and 2.3.2.8, respectively. Finally, straightforward extension of the basic approach provides differing levels of detail (Section 2.3.2.9) and abstraction (Section 2.3.2.10).

Together, these sections provide a comprehensive overview of the OOAD methodology. While simple examples are included in the sections to illustrate various concepts, the reader is referred to the case example, Section 2.3.4, for a unified application of the approach.

2.3.2.1 Objective Functions: Reinforcing & Balancing

Systems consist principally of two types of feedback loops: balancing and reinforcing [14]. These loops are described and graphically represented in the first part of the section. Then they will be represented mathematically. Finally, at the end of the section, an example is given.

Feedback loops, particularly the schematic representation of stock and flow structures, was pioneered by Jay Forrester [14] and popularized by Peter Senge [15]. The following sections are based on their original work.

Characteristics of Reinforcing and Balancing Feedback Loops

Balancing and reinforcing loops, as defined below, are called causal loops. Balancing (or goal seeking) behavior has several defining characteristics:

- Conservative behavior. The system seeks to maintain a steady state
- Exogenous objective state
- Delay between change in controller and change in state
- Controller governed by gap between desired and actual state
- Capable of operating in isolation

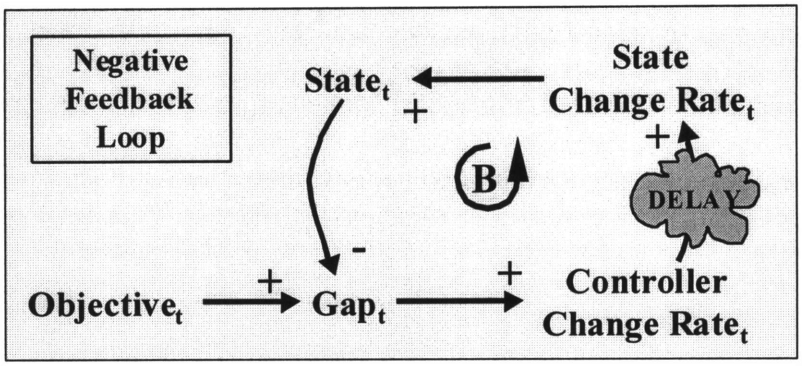


Figure 2-3: Graphical Representation of Goal Seeking Behavior

Balancing systems work to achieve and maintain an objective state that matches the objective of the system. The objective of the system is any objective, such as maintaining a certain number of days worth of inventory on hand. This objective is set exogenously to the system; that is, it is not a function of the system itself. The third characteristic of balancing systems is that there always exists a delay between the instance in which change is made and the time in which a change in the objective state is observed. This occurs because of the physical nature of all systems. Any change from one state to another necessarily has a rate of change in a period of time. Depending upon the nature of a given system, this delay can be short or long, playing a role which can be negligible in some cases and critical in others.

To maintain a system at the objective state, the system must have a controller and a way to measure the objective state. Given the perceived gap between the objective state and the objective, the controller state is changed so as to close the gap between the objective state and the system objective. Since both the objective state of the system and the controller of that state are contained within the same system, a balancing system will be self-regulating, i.e. it will manage itself in an autonomous manner.

Reinforcing loops are the second type of loop found in a system. Compared to balancing loops, reinforcing loops are:

- Non-conservative. System seeks to continuously amplify either positively or negatively
- Incapable of operating in isolation. Every observed reinforcing loop is interdependent with other reinforcing and/or balancing loops. This is necessary otherwise the system would exhibit non-conservative behavior
- Characterized by either maximum or minimum objective functions

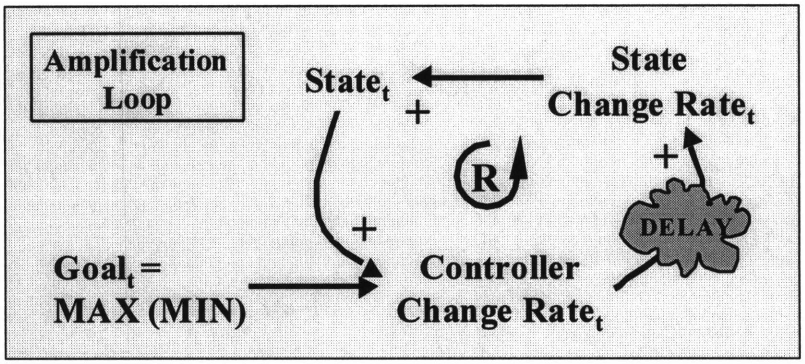


Figure 2-4: Graphical Representation of Reinforcing Loop

Maximum and minimum functions occur in virtually all systems. However, the amplifying behavior intrinsic to the nature of reinforcing loops is seldom seen; reinforcing behavior of an isolated system is non-conservative. Perpetual growth or decline does not occur because other loops begin to play increasingly strong roles, limiting amplification. For example, a company may want to maximize the machine rate for a particular manufacturing operation. Furthermore, if they decide that spindle speed for the cutting tool will be the control variable through which the machining rate will be increased. Even though the spindle speed may be able to be increased continuously, the material removal rate will not, since at higher spindle speeds thermal effects begin to play a dominant role. As the work piece temperature increases, (due to increased spindle speed) the ability of the tool blade to remove material decreases. This secondary, or side, effect creates the balancing behavior that prevents a runaway system. Therefore, systems containing reinforcing loops always include either balancing loops or reinforcing loops in the opposite direction, counteracting the principal loop's progress toward its given objective.

Mathematical Representation:

Both balancing and reinforcing loops can be characterized mathematically. Figure 2-5 and Figure 2-6 show the basic mathematical relationships for balancing and reinforcing loops.

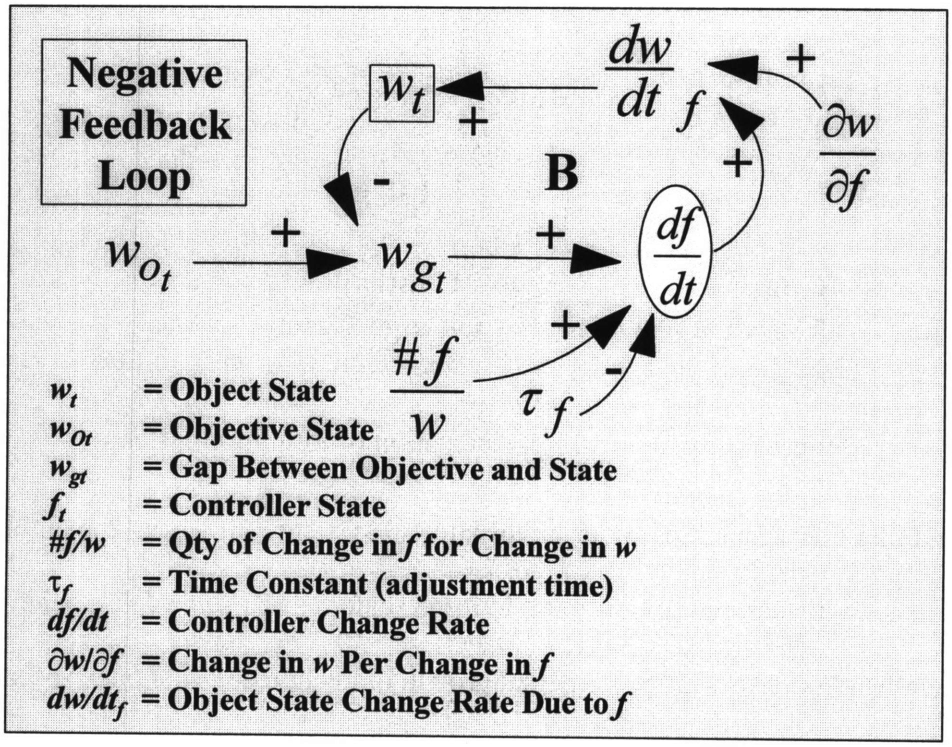


Figure 2-5: Mathematical Representation of Balancing Loop

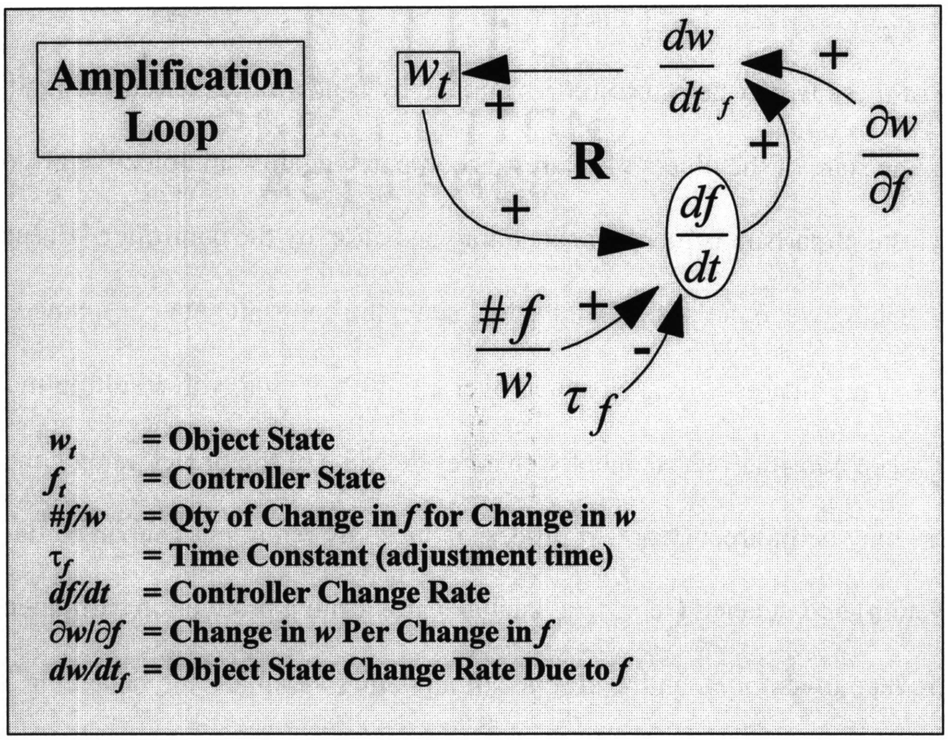


Figure 2-6: Mathematical Representation of Reinforcing Loop

The figures show the states (or levels) of different variables at different points in time, denoted by the subscript t. Within each loop, variables are either rates or levels. A level is the time-based integral of a rate and may, for example, represent work-in-process inventory or the defect rate of an assembly. This is illustrated in the figures by w_t, where w_t is an objective state variable, designated by a square. By contrast, the control variable, in this case df/dt, is designated by an oval. Comparison of the two different types of loops shows that their structure is nearly identical. The only difference is how information regarding the objective state of the system, w_t, is fed to the control variable. In a balancing loop, Figure 2-5, the gap between the desired and objective state, w_{gt}, is used to determine how much the control variable should be changed. On the other hand, for a reinforcing loop, Figure 2-6, there is no notion of a gap since the objective function of the system is to maximally increase or decrease the state of the system. In this manner, the state of the system is only used to determine the desired direction of change.

The behavior of these two systems can be determined by following the circuit around the loop. In either loop, the state of the object of interest is compared against an objective or goal state of that object. This comparison is then translated into an action by the controller. Based on the object state or gap between the desired and actual state, the controller state is changed by a certain amount. This change is then executed over a finite time interval, resulting in a rate of change, ***df/dt***. The controller subsequently changes the state of the object being controlled. The change process may be indirect (work-in-process reduced by removing kanban cards in a production system) and delayed (removing kanban cards from a production card queue does not immediately reduce inventory). In the absence of any other variables, the balancing loop will exhibit goal seeking behavior while the reinforcing loop will exhibit exponential growth behavior.

Example Involving Feedback Loops

Consider the feedback system associated with filling a glass ¾ full with water from a tap. This is a classic negative feedback system. The system, shown in Figure 2-7, is not unlike the basic feedback loop shown in Figure 2-3.

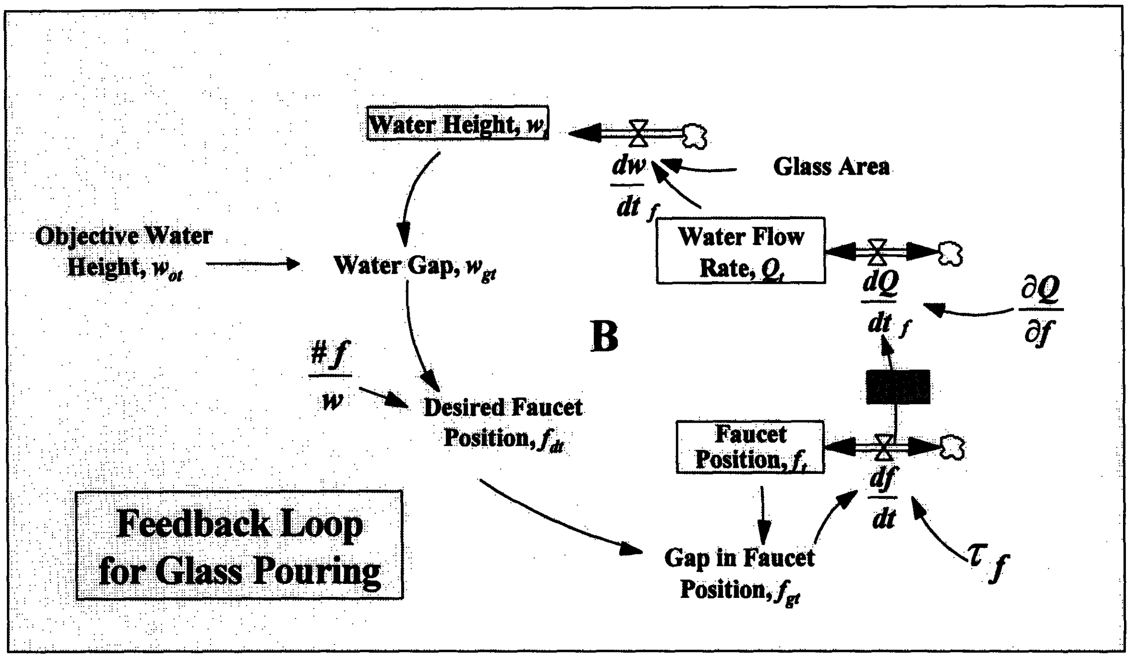


Figure 2-7: Feedback Loop Example

The objective state is the height of water in the glass and the gap is the difference between the objective (the glass ¾ full) and the current object state (height of water in the glass). The controller in this case is a water faucet knob (via a person). As the glass is filled, the faucet is adjusted to ensure the desired fill rate is achieved. Then, as the height of the water approaches its desired state, the flow rate is decreased to ensure that that the glass is not filled beyond ¾. In other words, the objective state change rate ***(dw/dt)*** is dictated by the faucet position. Like the basic negative feedback architecture, the knob's effect on the water flow rate is indirect: a rotation of the knob changes the position of a check valve, closing or opening the water passageway. Likewise, there is a delay between the change in position of the controller and the system response. The most significant delay is due to the time required for the water to flow from the valve to the glass. These delays and indirect relationships play significant roles in multivariate systems.

2.3.2.2 Interdependence Between Co-Existing Control Structures

Often reinforcing and balancing loops affect one another in an interdependent manner. These interdependencies can affect the system in nonlinear, and therefore non-intuitive, ways [16]. To demonstrate the interactive behavior of multiple feedback loops, it is worthwhile to consider an example.

Example: Productivity Improvement [17]

Consider a productivity improvement process in a manufacturing organization. While this process may have a number of objectives, the percent increase in task productivity is a good measure of improvement. Sometimes increased computer support or automation are chosen as methods for improvement, but for this example assume that the interest is to increase labor productivity through improved technique.

Like most improvement programs, productivity improvement requires a substantial investment in time in such things as training classes or kaizen sessions [18]. If these classes or sessions are effectively designed and run, the company will realize productivity improvements. However the benefit will not be realized immediately. Often the delay for productivity improvement will be anywhere from a few weeks to many months depending on the program [19]. The basic structure of this loop is shown in Figure 2-8.

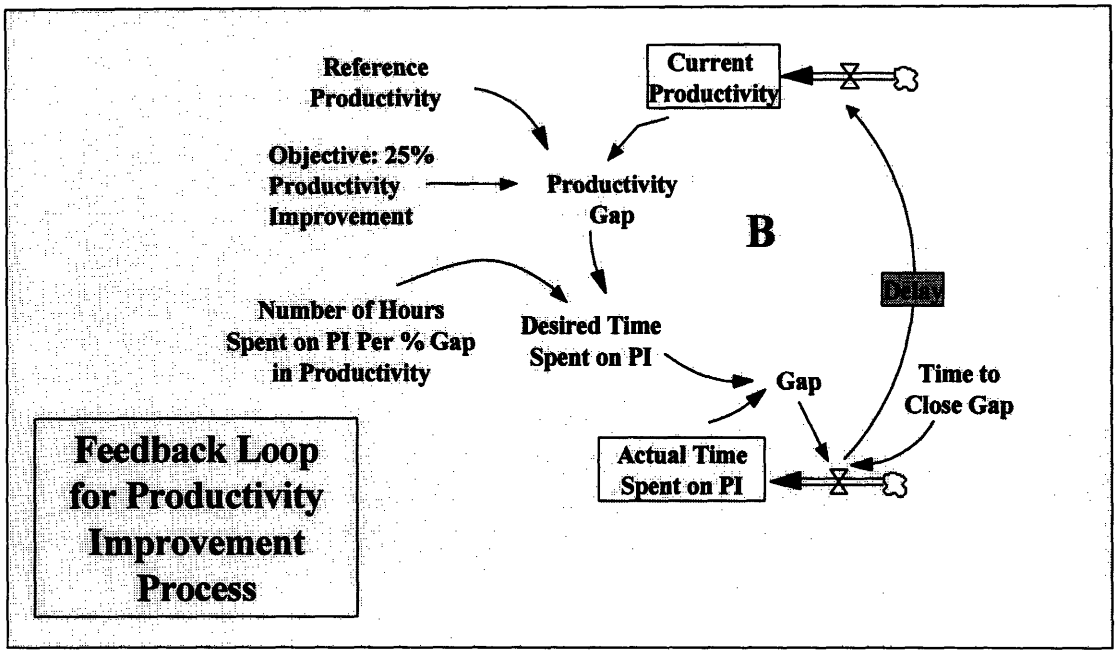


Figure 2-8: Productivity Improvement Loop

As shown in the figure, an increase in time spent on productivity improvement does not immediately increase the productivity of the system. This is due to the delay between the time changes are made in the time spent on productivity improvement, and the point in time the actual productivity increases.

The delay between the initiation of the productivity improvement process and the resultant boost in productivity causes an interdependence to be introduced. An *interdependence* is defined as the condition whereby a change in one variable in a system changes the value of an objective other than the principal objective given all other variables held constant [11]. The productivity improvement process takes time away from what would otherwise be time spent completing tasks. Since less time is available to complete tasks and productivity has not improved, it becomes more difficult to complete tasks in a timely manner. The interdependence is this: the productivity improvement process affects the organization's second objective of completing one hundred percent of its tasks on time. The control variable allowing the organization to meet this objective is the amount of time spent working per day. While the company may employ some long-term controls such as hiring or reducing work content through product redesign, overtime is often the only short term solution. This second loop is shown in Figure 2-9.

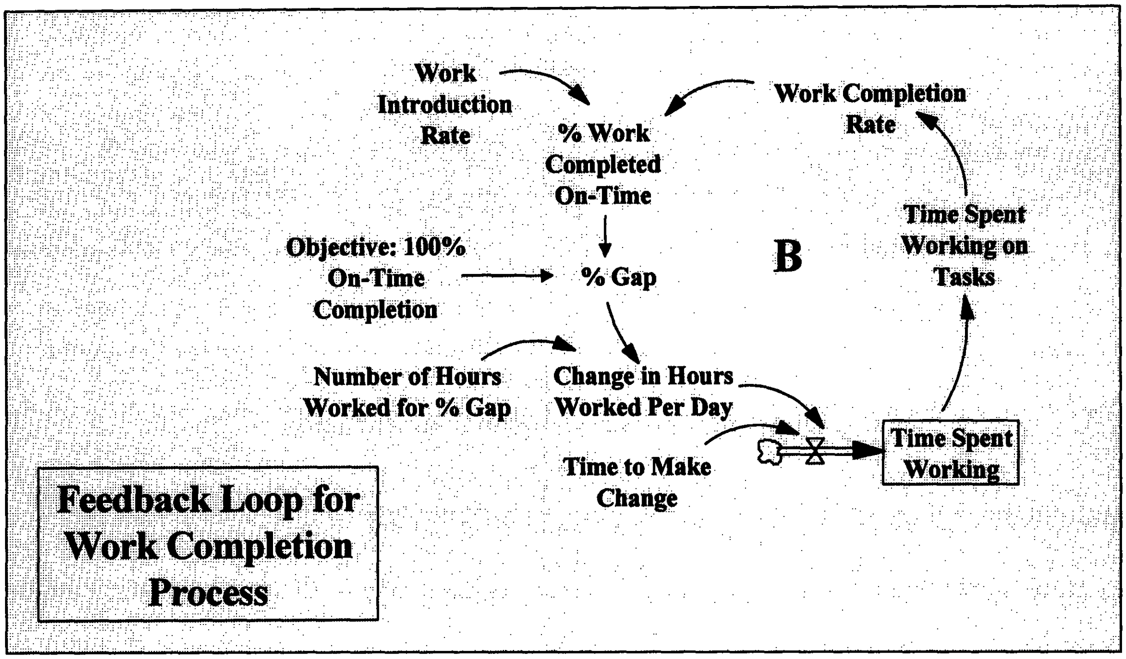


Figure 2-9: Work Completion Loop

The interdependence between these loops occurs in two places. First, improvements in productivity increase the work completion rate thereby increasing the percentage of work completed on time (a desirable effect). Second, time spent on process improvement immediately takes away from time spent working on tasks, reducing the amount of work completed on time (the side effect). These interdependencies are shown in Figure 2-10.

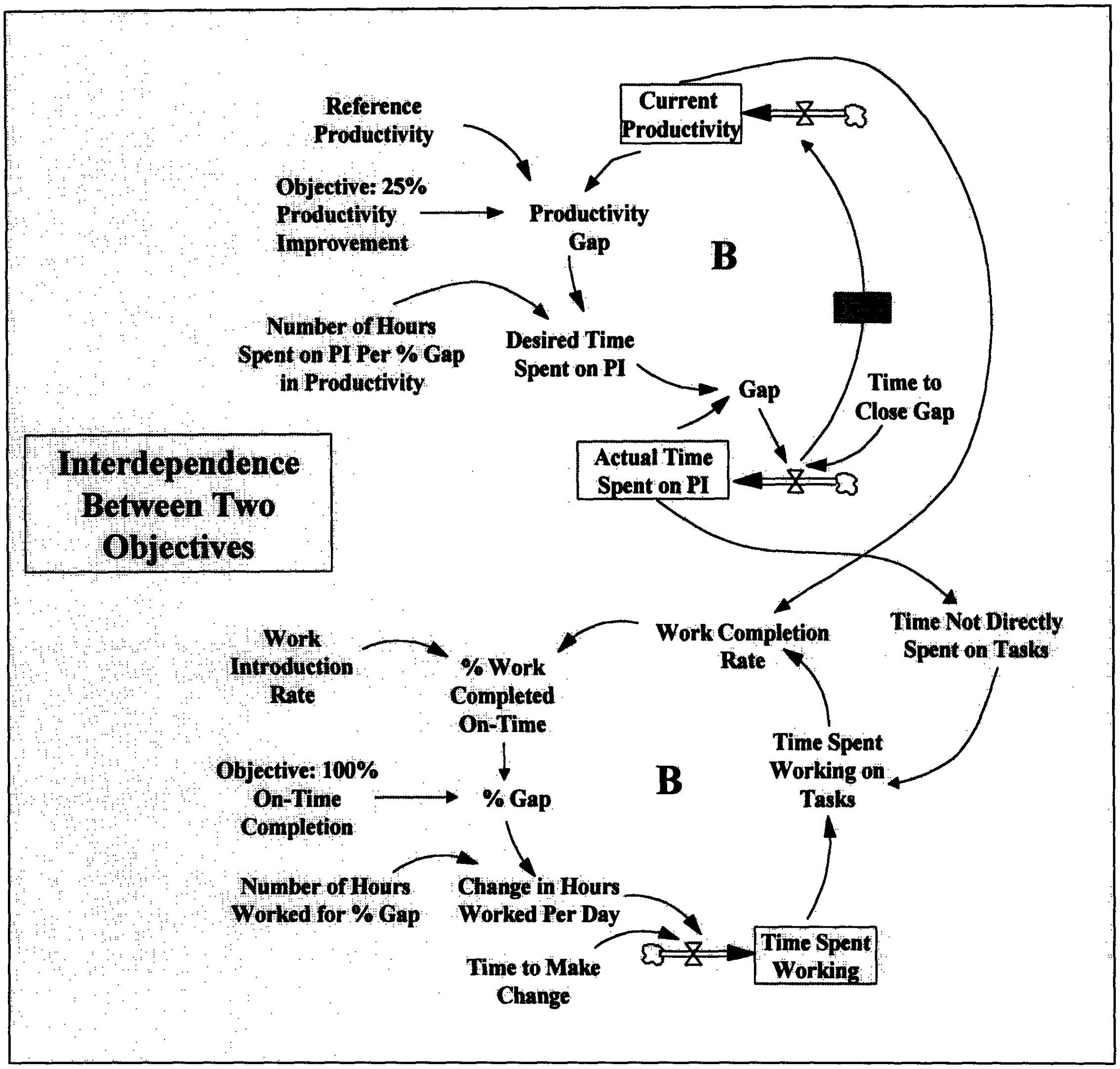


Figure 2-10: Interdependence Between Two Objectives

Since changes in the amount of time spent on productivity improvement affect the ability to complete work on time, the two feedback loops are said to be interdependent. Alternatively stated, for an interdependent system, a change in one control variable requires a change in another control variable for their respective objectives to be met. For the productivity improvement example, a change in the time spent on productivity improvement will require a change to be made in the time spent working if both objectives (productivity improvement and on-time completion) are to be simultaneously met.

The above example has shown two important concepts in OOAD. First, it has shown that the dynamic behavior of systems, from the control variable to the objective state (including intermediate variables), can be represented by causal loop diagrams. Second, it shows how these loops combine to form interdependencies in actual systems. This foundation will be revisited in Section 2.3.2.4, when the mapping between Causal Loop Diagrams and Axiomatic Design matrices is established.

2.3.2.3 Representing Larger Systems: Multiplicity of Objectives and Controls

As the previous section suggests, the behavior of real-world systems are often driven by a number of loops where the relationship between control variables and objective states may be separated by a number of intermediate variables. As illustrated in Figure 2-10, these diagrams can be quite complex, making the relationships between control and objective (not to mention interdependencies!) quite difficult to discern. While the causal loop representation is absolutely essential for objective system characterization, it does not lend itself well to concise representation of key relationships. Fortunately, the Axiomatic Design methodology has a compact structure that allows the key relationships of the underlying interdependencies to be portrayed in a straightforward manner. This is achieved through the use of an Axiomatic Design matrix alongside the System Dynamics structural representation.

Seldom do systems consist of a single objective function. In practice, managers, engineers, and scientists must simultaneously manage multiple objectives. A useful way of abstracting the systems' principal elements, namely, the objective function, the control function, and their relationship to each other, is through the use of an Axiomatic Design matrix. Axiomatic Design matrices were first proposed by Nam P. Suh [11]. The graphical and mathematical representations in the following sections are adaptations of his work. A simple design matrix is shown in Figure 2-11.

Design Matrix	Control 1	Control 2	Control 3
Objective 1	X	O	O
Objective 2	O	X	X

Figure 2-11: Graphical Description of Simple Design Matrix

As the figure shows, *Control 1* only affects *Objective 1.* Likewise, *Control 2* and *Control 3* only affect *Objective 2.* These loops are independent of each other as represented by the "**O**" demarcation on the matrix off-axis diagonal, indicating the absence of a relationship. The use of the letter "**X**" indicates that a change in state of the control will change the state of the objective. In most systems, there are a multitude of controls at a person's disposal, which can be used to affect the state of a system. Nevertheless, only a few controls are ever used. It is only the controls that are actively used to optimize the outcome which are included in the design matrix.

The design matrix is useful when an abstraction of the underlying system behavior is desired. However, the use of the design matrix does not contain valuable information necessary for evaluating system performance in the "**X**" and "**O**" format. First, it does not indicate the sign of the relationship between the control and the objective variable. The sign may be positive, negative, or dependent on the temporal state of the system. Secondly, the design matrix does not represent any of the intermediate variables that might exist between the control and the objective state. This is significant because the strength of the relationship between the control and objective may be useful to know, but only the intermediate variable relationships can be measured or determined.

Interdependencies such as the ones in the productivity improvement example, introduced in Section 2.3.2.2, can be shown by adding an "X" to the off-axis diagonal as shown in Figure 2-12.

Productivity Improvement Design Matrix	**Change in Time Spent on Productivity Improvement**	**Change in Time Spent Working**
Change in Productivity	X	O
Change in % of Tasks Completed On Time	X	X

Figure 2-12: Productivity Improvement Graphical Design Matrix

The off-axis "X" indicates that an interdependency exists between *Control 1* (Time Spent on Productivity Improvement) and the two objectives. What the off-axis mark doesn't show, however, is that the interdependency is a really a function of two distinct effects – separate effects that affect the system in different ways over time. Initially, an increase in time spent on productivity improvement reduces the time spent completing tasks, decreasing the percentage of tasks completed on time (a negative relationship). But, after a delay, an increase in time spent on productivity improvement will increase the percentage of tasks completed on time (a positive relationship).

As a result of the complex and nonlinear nature of systems, neither the design matrix nor the causal loop are sufficient by themselves to enable effective characterization, description, and optimization of system behavior. The System Dynamics modeling of system behavior is both comprehensive and general, but it lacks the abstraction necessary to be able to clearly convey key interdependencies when numerous variables (or loops) are involved. In contrast, Axiomatic Design provides high levels of abstraction, but ignores the intermediate variables between the control variables and the objective variables in the system. This loss of information makes it difficult to demonstrate the "path" between the control variable and the objective state and therefore makes it difficult to present interdependencies in a manner that can be easily understood and accepted as an accurate representation of the system.

It is therefore essential that causal loops and design matrices are used together and that the mapping between them be explicit. In the following section, the relationship between causal loops and design matrices is developed.

2.3.2.4 Mapping Between Causal Loops and Design Matrices

Establishing a mapping between causal loops and design matrices requires the development of a sound mathematical relationship between the two methods. The mathematical representation of a simple design matrix is shown below:

$$\begin{Bmatrix} FR_1 \\ FR_2 \end{Bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{Bmatrix} DP_1 \\ DP_2 \end{Bmatrix} \tag{2-2}$$

Taking the time-based derivative of both sides yields the following relationship:

$$\begin{Bmatrix} \dfrac{dFR_1}{dt} \\ \dfrac{dFR_2}{dt} \end{Bmatrix} = \begin{bmatrix} \dfrac{\partial FR_1}{\partial DP_1} & \dfrac{\partial FR_1}{\partial DP_2} \\ \dfrac{\partial FR_2}{\partial DP_1} & \dfrac{\partial FR_2}{\partial DP_2} \end{bmatrix} \begin{Bmatrix} \dfrac{dDP_1}{dt} \\ \dfrac{dDP_2}{dt} \end{Bmatrix} \tag{2-3}$$

Allowing variables ***w*** and ***x*** to represent FRs 1 and 2, and variables ***f*** and ***g***, to represent DPs 1 and 2, respectively, results in the design matrix, below:

$$\begin{Bmatrix} \dfrac{dw}{dt} \\ \dfrac{dx}{dt} \end{Bmatrix} = \begin{bmatrix} \dfrac{\partial w}{\partial f} & \dfrac{\partial w}{\partial g} \\ \dfrac{\partial x}{\partial f} & \dfrac{\partial x}{\partial g} \end{bmatrix} \begin{Bmatrix} \dfrac{df}{dt} \\ \dfrac{dg}{dt} \end{Bmatrix} \tag{2-4}$$

This relationship can then be compared to two generic feedback loops. As shown in Figure 2-13, the representations are the same except that in the feedback loop diagram the rates of change for ***w*** and ***x*** are shown to be separately affected by the control variables, ***f*** and ***g***.

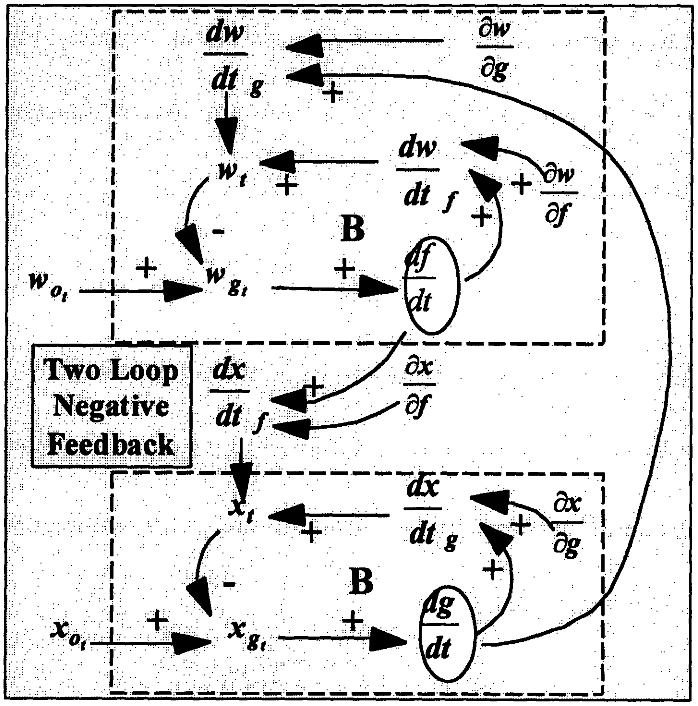


Figure 2-13: Mathematical Representation of Two Loop Negative Feedback

The on-axis relationships are shown in the main feedback loops (signified by the "B", for balancing), while the off-axis elements of the design matrix are shown crossing from one loop to another. In both forms (causal loop diagram or design matrix notation) the off-axis couplings are represented by the partial of the off-axis objective state divided by the partial of the control variable. In both the equation and the figure, these are shown by $\partial x/\partial f$ and $\partial w/\partial g$.

Simultaneous representation of design matrices and causal loops enables the critical interdependencies to be distilled from a set of complex relationships without losing fundamental representation. Taken separately, the design matrix is an insufficient system descriptor and the causal loop representation (though comprehensively describing the system) does not highlight the key relationships in the system.

The next section demonstrates this technique by revisiting the productivity improvement process example described earlier.

2.3.2.5 Example: Characterization of Productivity Improvement Process Through Design Matrix – Causal Loop Combination

This example develops the design matrix representation of the productivity improvement process through the use of a comprehensive causal loop model. This model, a detailed version of the two loop representation (Figure 2-10), is shown in Figure 2-14.

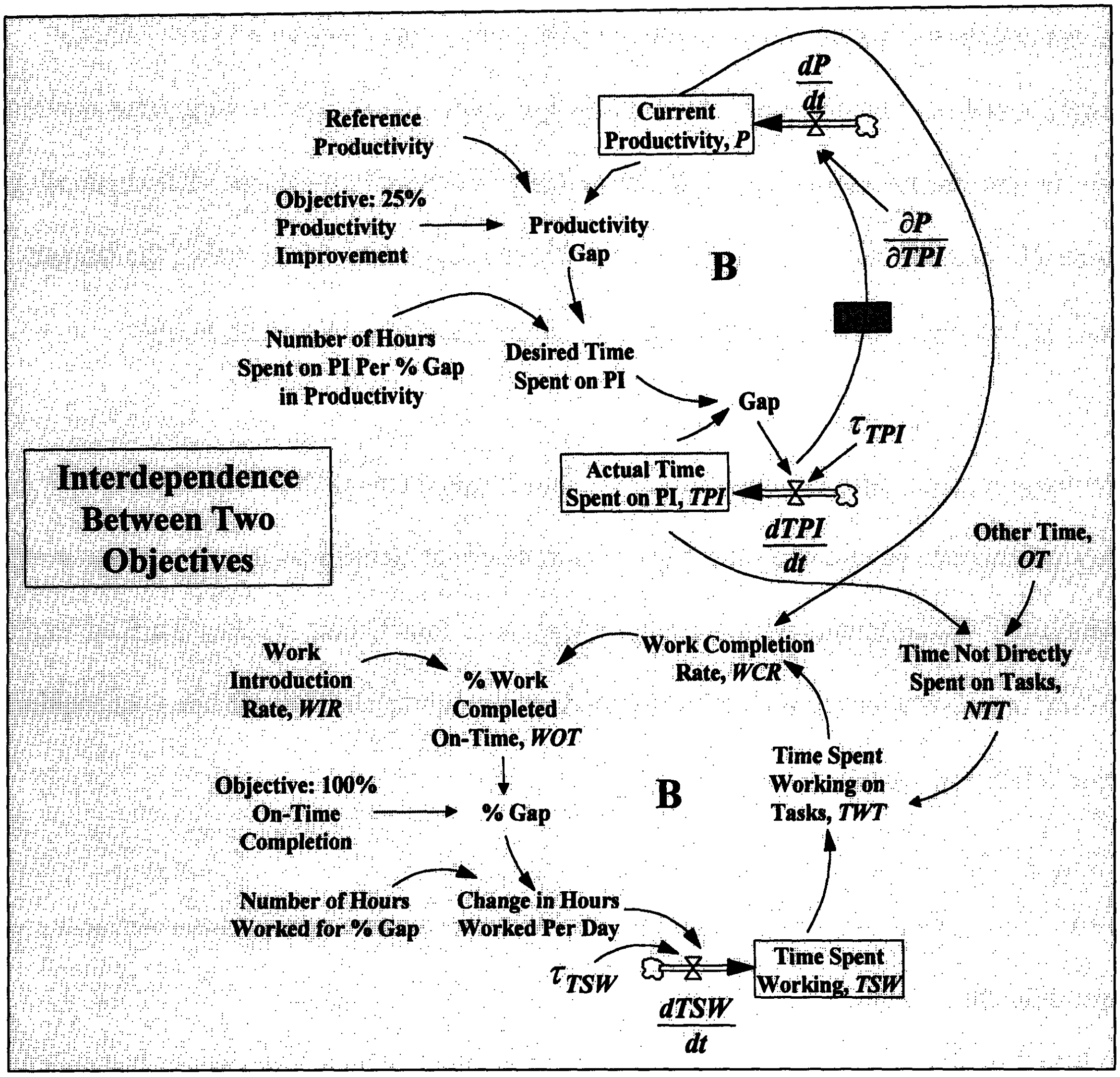


Figure 2-14: Detailed Causal Loop Diagram of Productivity Improvement Process

The following notation is used to describe the system:

P	Productivity
dP/dt	Change in ***P*** per change in time
TPI	Time spent on Productivity Improvement
dTPI/dt	Change in ***TPI*** per time
∂P/∂TPI	Partial differential of ***P*** with respect to ***TPI***
NTT	Non Task Time

OT	Other Time (not task time)
TWT	Time spent Working on Tasks
TSW	Time Spent Working
dTSW/dt	Change in ***TSW*** per change in time
WIR	Work Introduction Rate
WCR	Work Completion Rate
WOT	Percentage of Work completed On Time
$\partial WOT/\partial P$	Partial differential of ***WOT*** with respect to ***P***
$\partial WOT/\partial TSW$	Partial differential of ***WOT*** with respect to ***TSW***
$\partial WOT/\partial TPI$	Partial differential of ***WOT*** with respect to ***TPI***

To obtain the design matrix representation for this system, the four partial differentials must be determined. Since the two objectives are functions of ***P*** and ***WOT*** and the control variables are ***TPI*** and ***TSW***, respectively, the four partial differentials are:

$$\frac{\partial WOT}{\partial TSW},\ \frac{\partial P}{\partial TPI},\ \frac{\partial WOT}{\partial TPI},\ \frac{\partial P}{\partial TSW}$$

The following derivation determines these partial differentials.

First, ***WOT*** is determined in terms of control variables:

$$WOT=\frac{WCR}{WIR}=\frac{P*TWT}{WIR}=\frac{P*(TSW-NTT)}{WIR}=\frac{P*(TSW-TPI-OT)}{WIR} \tag{2-5}$$

Second, the differentials of the objective functions are written in terms of the immediately influencing intermediate variables:

$$\frac{dP}{dt} = \frac{dP}{dt}_{TPI}$$

$$\frac{dWOT}{dt} = \frac{DWOT}{dt}_{P} + \frac{DWOT}{dt}_{TSW} + \frac{DWOT}{dt}_{TPI} \tag{2-6}$$

Expanding,

$$\frac{dP}{dt}_{TPI} = \frac{\partial P}{\partial TPI} \cdot \frac{dTPI}{dt}$$

$$\frac{dWOT}{dt}_{P} = \frac{\partial WOT}{\partial P} \cdot \frac{dP}{dt}$$

$$\frac{dWOT}{dt}_{TSW} = \frac{\partial WOT}{\partial TSW} \cdot \frac{dTSW}{dt}$$

$$\frac{dWOT}{dt}_{TPI} = \frac{\partial WOT}{\partial TPI} \cdot \frac{dTPI}{dt}$$

Noting that $\frac{dP}{dt} = \frac{\partial P}{\partial TPI} \cdot \frac{dTPI}{dt}$,

$$\frac{dWOT}{dt}_{P} = \frac{\partial WOT}{\partial P} \cdot \frac{\partial P}{\partial TPI} \cdot \frac{dTPI}{dt}$$

Taking partial derivatives of (2-5),

$$\frac{\partial WOT}{\partial P} = \frac{TSW - TPI - OT}{WIR}$$

$$\frac{\partial WOT}{\partial TSW} = \frac{P}{WIR}$$

$$\frac{\partial WOT}{\partial TPI} = \frac{-P}{WIR}$$

Substituting into (2-6),

$$\frac{dWOT}{dt}=\left(\frac{TSW-TPI-OT}{WIR}\cdot\frac{\partial P}{\partial TPI}+\frac{-P}{WIR}\right)\frac{dTPI}{dt}+\frac{P}{WIR}\cdot\frac{dTSW}{dt}$$

Since ***P*** is not a function of indirect variables,

$$\frac{dP}{dt}=\frac{\partial P}{\partial TPI}\cdot\frac{dTPI}{dt}$$

where $\frac{\partial P}{\partial TPI}$ is not developed but known to be generally positive and independent of the other control variables and where $\frac{\partial P}{\partial TSW}=0$. Since $\frac{\partial P}{\partial TPI}$ is the change in productivity per time spent on productivity improvement, its specific value will be different for different manufacturing processes. For example, $\frac{\partial P}{\partial TPI}$ for semiconductor manufacturing may be extremely low since fabs are updated every several years. By contrast, a new startup contract manufacturer may find it extremely easy to achieve productivity improvements with minimal effort. In both cases though, the relationship will be positive.

Taken together, these partial differentials completely define the design matrix for this system.

$$\begin{Bmatrix}\frac{dP}{dt}\\ \frac{dWOT}{dt}\end{Bmatrix}=\begin{bmatrix}\frac{\partial P}{\partial TPI} & \frac{\partial P}{\partial TSW}\\ \frac{\partial WOT}{\partial TPI} & \frac{\partial WOT}{\partial TSW}\end{bmatrix}\begin{Bmatrix}\frac{dTPI}{dt}\\ \frac{dTSW}{dt}\end{Bmatrix} \tag{2-7}$$

The matrix above is the design matrix for the productivity improvement example. Substituting the derivation from above yields:

$$\begin{Bmatrix}\frac{dP}{dt}\\ \frac{dWOT}{dt}\end{Bmatrix}=\begin{bmatrix}+ & 0\\ \frac{TSW-TPI-OT}{WIR}\cdot\frac{\partial P}{\partial TPI}+\frac{-P}{WIR} & \frac{P}{WIR}\end{bmatrix}\begin{Bmatrix}\frac{dTPI}{dt}\\ \frac{dTSW}{dt}\end{Bmatrix} \tag{2-8}$$

This relationship shows the coupling relationships in detail. The "+" is used since the productivity improvement per time invested will be unique for each manufacturing productivity improvement process.

The relationship in (2-8) can be simplified further by substituting "X" for matrix sub-element that has coupling. This simplification is shown below:

$$\begin{Bmatrix} \frac{dP}{dt} \\ \frac{dWOT}{dt} \end{Bmatrix} = \begin{bmatrix} X & 0 \\ X & X \end{bmatrix} \begin{Bmatrix} \frac{dTPI}{dt} \\ \frac{dTSW}{dt} \end{Bmatrix} \tag{2-9}$$

As seen in this example, changes in time spent working have no effect on the productivity of the people performing the tasks. However, as previously described, there is an interdependence between the time spent on productivity improvement and the percent of work completed on time. While the design matrix shows an effect, it does not show the sign of the interdependency which in this case is neither positive nor negative. Instead it changes from negative to positive as the effect of time spent on productivity improvement begins to boost productivity.

2.3.2.6 Encapsulation

Encapsulation is an essential characteristic within the Object-Oriented Axiomatic Design paradigm. Strictly speaking, when part of a system is encapsulated, it has no common shared state with other parts of the system. This means that the encapsulated part of the system is not operated on by any other part of the system other than through pre-defined interfaces. In a business context, this means that a group, be it a department, business unit, or company, has the freedom to act in an independent manner. Though the objectives of the group may be prescribed, the group has the freedom to employ any method to meet its objectives. The group should be free to optimize about its objectives without affecting the capability or performance of any other group. If this is not the case and a change made by one group compromises another, they cannot each be considered encapsulated as Figure 2-15 illustrates.

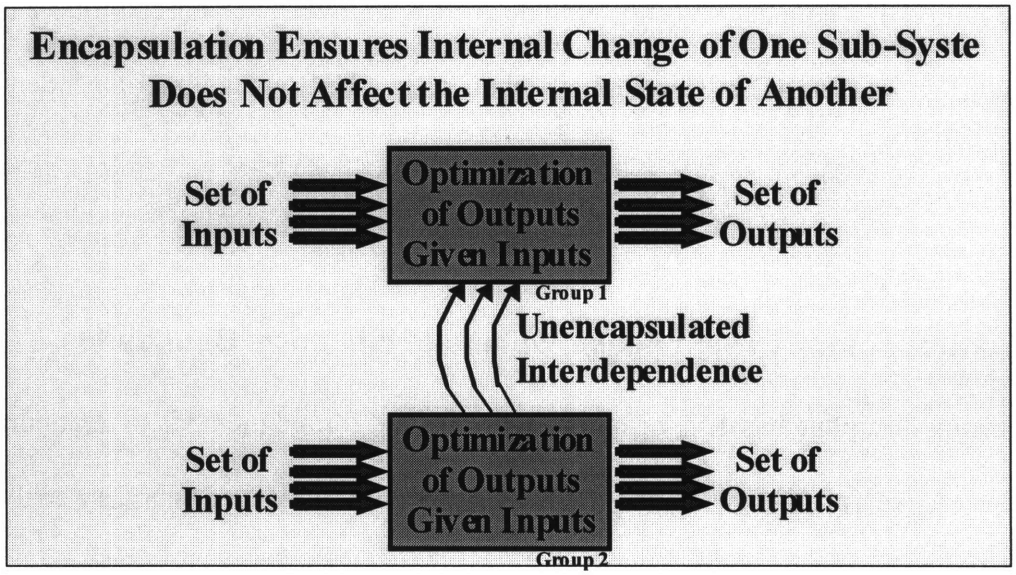


Figure 2-15: Unencapsulated System

Encapsulation, then, is a measure of the degree to which objects in a system are free to employ any number of methods available to them to best meet their objectives, without adversely affecting any other group. An encapsulated approach to system design and control is extremely powerful and effective in dynamic, changing environments, as Object-Oriented software development approaches have demonstrated. The rapid and virtually universal adoption of object-based approaches is testimony to how large, distributed, and evolutionary systems can undergo continuous improvement and change without fail. Encapsulation has been the key enabler, allowing rapid improvement to objects within a sub-system without compromise to the performance of objects elsewhere in the same system. While groups within organizational systems cannot be encapsulated to the same degree as software objects (due to physical laws, shared resources, etc. that do not exist in the world of virtual objects), the same principles apply, making an encapsulated state, as it is defined here, a goal worth seeking.

To measure the degree of encapsulation (or the converse – interdependence) between objects in a system, the notion of a boundary must be embraced. The following section introduces the boundary concept into OOAD by developing what is called the Bounded Interdependence Matrix.

2.3.2.7 Control Points, Design Space, Sphere of Influence, Boundaries

The engineering discipline has developed many elegant and formal methods to treat boundary conditions in engineering design and analysis. In many areas, the use of boundary analysis has proved a powerful technique for designing systems whose expected behavior is representative of actual operating behavior. The notion of a control volume in fluid mechanics and the free-body-diagram in statics are both methods of setting boundaries that separate the external environment from the localized area of interest. In electrical engineering, minimizing impedance mismatch is a way to minimize boundary disruptions. Additionally, the idea of a Thevenin equivalent where entire systems can be modeled as a series resistor and a source is a method of selectively designing parts of a much larger system. In all cases, the proper treatment of boundaries is critical for understanding and properly developing a design. In Finite Element Analysis for example, the treatment of boundary conditions between surfaces or between different types of elements is arguably the single most important area contributing to the overall success of the technique.

Many systems must functionally interoperate with other systems which can significantly affect the behavior of the design. As evidenced by the rapid acceptance of object-oriented methods in software development, the ability to provide sustained interoperability in dynamic environments has been vital to the long-term viability of software systems. The same is true for organizational systems.

While standard Axiomatic Design works well to solve isolated systems or devices the designer has full control over, it does not provide a formal treatment of boundaries or boundary conditions. Fortunately, the extension of Axiomatic Design matrices to provide formal recognition of boundaries within systems is straightforward. As shown in Figure 2-16, the boundary of a system can be represented in the following way:

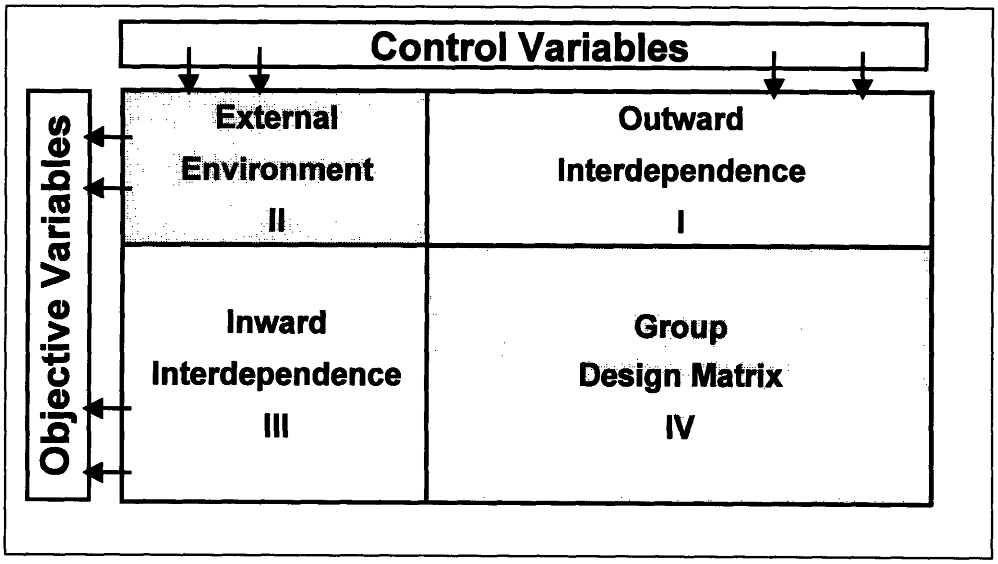


Figure 2-16: Bounded Interdependence Matrix (Simplified)

Figure 2-16 shows four quadrants of interdependence. Each quadrant represents interdependencies from the perspective of a single group, where a group is defined as the range over which behavior can be controlled in a centralized and non-autonomous manner. This group has a set of objectives (objective variables) and a set of controls (control variables) which are used to meet the group's objectives. The Group Design Matrix, Quadrant IV, is the design matrix that spans the sphere of influence of this group. Quadrant IV relates how the group's control variables (CVs) relate to its objectives. If the group's objectives and controls are considered in isolation, the Bounded Interdependence Matrix reduces to a standard Axiomatic Design matrix where the objective variables can be equated with FRs and the control variables can be equated with DPs [11].

However, no system exists alone; relative to a particular group with its "sphere of influence", everything else can be considered the group's environment. This can be treated either as the environment as a whole (as represented in the figure) or as a collection of other external groups. Not only do multiple autonomous groups exist, but there is often interaction between the external environment and the group of interest. This interaction can result from changes made by the particular (or principal) group which affect its environment as represented by Quadrant I, or from changes by some other group (in the environment) which affect the principal group's objectives. This second form of interaction is represented by Quadrant III. In both cases, the interaction across the boundary (shown by the black lines in Figure 2-16) occurs due to the dynamic interdependent behavior of multiple groups each trying to optimize its own set of objectives. Quadrants I and III are considered to be the off-axis interdependencies between groups. The existence of any relationship in either quadrant indicates that an interdependency exists which crosses the boundary of control of the principal group. These off-axis interdependencies can critically affect system performance, and if they are not considered, the characterization of the system may be significantly different from what is realized in practice.

The final area of the Bounded Interdependence Matrix, shown in Figure 2-16, is Quadrant II. Quadrant II represents the design matrix of the external environment, i.e. the matrix representing the relationship between the control variables in the external environment and the external environment's objective variables. This matrix of the external environment serves two purposes. First, it maps the relationship between control variables and objective variables for groups in the environment. This makes the larger system easier to understand, and therefore makes it easier to identify interdependencies which cross boundaries. Secondly, a mapping of Quadrant II provides valuable information to the principal group of how the environment might react to changes made by the principal group. For example, a company may want to know how a competitor will respond to a change in the company's price of a product or service. This is significant since a price war may reduce the earnings of all companies involved.

Since the Bounded Interdependence Matrix is used extensively in OOAD, it is worthwhile to more fully develop its structure. A detailed Bounded Interdependence Matrix is shown below:

	ECV_1	ECV_2	ECV_3	CV_1	CV_2	CV_3	CV_4	CV_5
EFR_1	X							
EFR_2		X						
EFR_3			X		X			
FR_1				X				
FR_2		X			X			
FR_3			X			X		
FR_4	X						X	
FR_5								X

Figure 2-17: Bounded Interdependence Matrix (Detailed)

This matrix is identical in structure to the Bounded Interdependence Matrix of Figure 2-16. In the figure above, each of the control variables and objective variables have been broken down into two groups. ECV and CV represent the control variables of the environment and principal group, respectively. Likewise, EFR and FR represent the objective variables (Functional Requirements in Axiomatic Design terminology) of the environment and principal group, respectively.

Figure 2-17 also shows how changes by the principal group can affect the environment, which in turn, can react and create a feedback response which loops back to the original group, affecting its performance. For example, a change made by **CV_3** to better meet its FR, can cause an imbalance in the external environment's **FR_3**. This is of no concern for the principal group unless the environment decides to respond to this action by changing its **ECV_3**, which, in turn, upsets the state of **FR_3**, a responsibility of the principal group. In this manner, systems with multiple, independent sources of control can exhibit closed-loop feedback behavior which is difficult to understand unless the key relationships governing this behavior are identified.

The Bounded Interdependence Matrix representation of Figure 2-17 can be generalized. The general form for analyzing and characterizing the different types of interdependencies is shown in Figure 2-18. It can be seen, comparing this mathematical representation with Equation (2-4), that the mathematical relationships between the Design Matrix and Bounded Interdependence Matrix are identical.

$$\begin{array}{cccc}
\textbf{Obj.} & \textbf{Gap} & \textbf{State} & \hfill \textbf{Control} \\
\begin{matrix} w_{1_\alpha} \\ \vdots \\ w_{n_\alpha} \\ x_{1_\alpha} \\ \vdots \\ x_{n_\alpha} \end{matrix} &
\begin{matrix} \Delta w_1 \\ \vdots \\ \Delta w_n \\ \Delta x_1 \\ \vdots \\ \Delta x_n \end{matrix} &
\begin{matrix} \int_t dw_1/dt \\ \vdots \\ \int_t dw_n/dt \\ \int_t dx_1/dt \\ \vdots \\ \int_t dx_n/dt \end{matrix} &
\left\{ \left[\begin{array}{c} dw_1 \\ \vdots \\ dw_n \\ \hline dx_1 \\ \vdots \\ dx_n \end{array}\right]
\left[\begin{array}{ccc|ccc}
\partial w_1/\partial f_1 & \cdots & \partial w_1/\partial f_n & \partial w_1/\partial h_1 & \cdots & \partial w_1/\partial h_n \\
\vdots & DM_1 & \vdots & \vdots & I_{2\rightarrow 1} & \vdots \\
\partial w_n/\partial f_1 & \cdots & \partial w_n/\partial f_n & \partial w_n/\partial h_1 & \cdots & \partial w_n/\partial h_n \\
\hline
\partial x_1/\partial f_1 & \cdots & \partial x_1/\partial f_n & \partial x_1/\partial h_1 & \cdots & \partial x_1/\partial h_n \\
\vdots & I_{1\rightarrow 2} & \vdots & \vdots & DM_2 & \vdots \\
\partial x_n/\partial f_1 & \cdots & \partial x_n/\partial f_n & \partial x_n/\partial h_1 & \cdots & \partial x_n/\partial h_n
\end{array}\right]
\left[\begin{array}{c} df_1 \\ \vdots \\ df_n \\ \hline dh_1 \\ \vdots \\ dh_n \end{array}\right] \right\}
\end{array}$$

Figure 2-18: General Bounded Interdependence Matrix

Figure 2-18 shows the four quadrants of the Bounded Interdependence Matrix previously described. DM_1 and DM_2 represent the external and principal design matrices, respectively. $I_{1\rightarrow 2}$ is the inward interdependence or the interdependence from DM_1 to DM_2. Similarly, $I_{2\rightarrow 1}$ is the outward interdependence or the interdependence from DM_2 to DM_1.

While this general interdependence matrix only shows the boundary between two groups, it is general and can be expanded to represent ***n*** groups. As well, the objectives of each of these groups can independently be maximizing, minimizing, or goal seeking, to represent the reinforcing and balancing feedback behavior that characterizes any given system.

2.3.2.8 Coupling Interdependencies

This section develops a notion of coupling interdependencies that is different from the definition used in Axiomatic Design. Unlike Axiomatic Design which focuses on the interdependencies contained within the sphere of influence of a single design team or organizational group, OOAD exclusively focuses on the interdependencies which bridge multiple teams or groups within a system. There are two principal reasons for the shift of focus from the self-contained interdependencies within a group's control to the interdependencies which cross control boundaries. First, most people are fairly smart. This means that one way or another, people will optimize the areas over which they have control and for which they are held responsible. If interdependencies exist which adversely affect another objective within the group, it is unfortunate, but not tragic. As Suh states in [11], the drawback to the presence of coupling interdependencies is that iteration may be required. However, the dynamics *between* different groups may not be nearly as benign. Since no single group has the hegemony to control the overall outcome or state of a system, interdependencies caused by one group's actions to improve its own performance may seriously compromise the performance of another group. This dichotomy is the second reason why OOAD emphasizes the interdependencies between groups. As this section will show, coupling interdependencies between groups can lead to some very undesirable dynamic behavior that could not occur within the context of a single group or individual. The development of this component of OOAD is made possible by the notion of encapsulation and boundaries. It is detailed in the pages which follow.

Several distinct types of interdependencies have been identified. Each yields distinctly different behavior; some interdependencies are beneficial and therefore desirable while others cause deficiencies in system behavior and should be eliminated. Although the term *interdependence* has been previously used to characterize any situation where changes made to one objective affect another, the following sections use a more specific definition of interdependence. Recognizing the role of boundaries in systems gives rise to two distinct types of interdependencies. Encapsulated interdependencies can be fully optimized from within the grouping. The second type of interdependency occurs when a group's objective state is affected in a manner that it cannot directly control. Under this circumstance, a group affected by an external influence cannot independently optimize its performance. In this circumstance, the two groups having interdependencies between them are said to be *coupled.* Subsequently, when the term *coupling* is used, it refers to two groups, having independent sources of control, where a change made by one group affects the other's performance.

In many systems, multiple types of coupling exist. The presence of each type of coupling causes a fundamentally different dynamic response. To develop the basic types of coupling, a subset of the matrix representation developed in Section 2.3.2.8 will be used in the sections which follow.

A total of six different types of couplings have been identified. These are: *Adverse Stable Coupling, Adverse Unstable Coupling, Advantageous Stable Coupling, Advantageous Unstable Coupling, Exogenous Coupling,* and *Temporal Coupling.* Each of these types of couplings are detailed in the following sections. For the purposes of explanation, the sections will make use of a generic bounded interdependence matrix. The matrix uses the following structural representation and element notation:

Bounded Interdependence Matrix

$$\begin{array}{c|cc|cc}
\text{Gap} & & & & \\
G_1 & A_{11} & A_{12} & A_{13} & A_{14} \\
G_2 & A_{21} & A_{22} & A_{23} & A_{24} \\
\hline
G_3 & A_{31} & A_{32} & A_{33} & A_{34} \\
G_4 & A_{41} & A_{42} & A_{43} & A_{44}
\end{array}$$

Figure 2-19: Generic Bounded Interdependence Matrix

The gap, or difference between the objective and the current objective state is designated by ***G***G_i. This gap, if it is positive, indicates that the state of objective variable should be increased to close the gap. It follows that if the gap is negative, the state of the objective variable is higher than its objective and should therefore be lowered to eliminate the gap. The elements of the Bounded Interdependence Matrix are designated by A_{ij}. The on-axis elements have identical subscripts, i.e. $i = j$. All other elements can be considered as off-axis elements.

Adverse Coupling

Adverse coupling occurs when an interdependence exists between two groups where the change of state of the objective in one group adversely affects the state of another with respect to its objectives. This is seen in Figure 2-20.

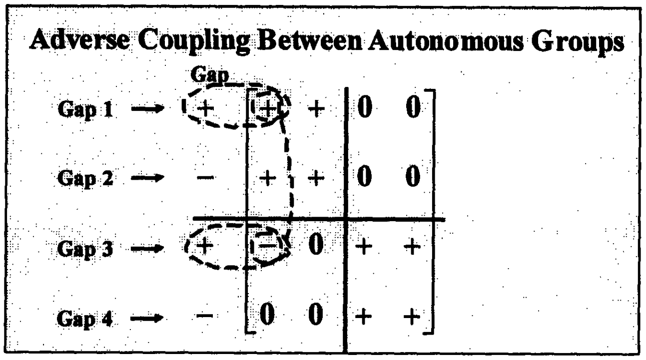


Figure 2-20: Adverse Coupling Between Autonomous Groups

In the figure, the first group (top left quadrant) will want to decrease the positive gap existing between its objective and current objective state. However, doing so will create a larger gap between the objective and objective state in the second group. Since the off-axis coupling does not affect the first group's objectives, the first group will optimize its performance at the expense of the second group, causing adverse coupling behavior. The second group can do little to counter this behavior since an increase in one of its control variables will adversely affect its second objective where a negative gap exists.

Advantageous Coupling

Advantageous coupling occurs when the actions taken to improve the state of one group have an unintended secondary benefit to another group. In this manner, advantageous coupling is desirable since two groups benefit through the efforts of a single group. This effect is shown in Figure 2-21, below.

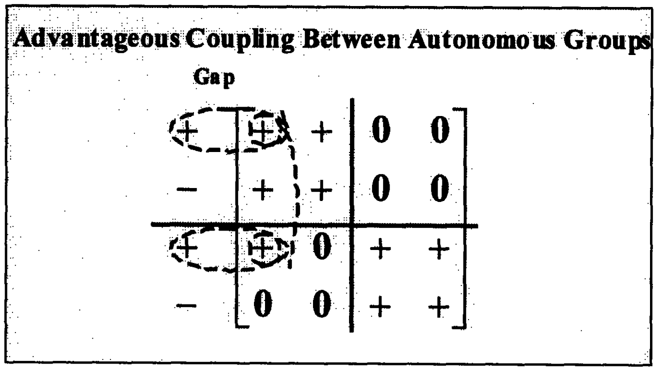


Figure 2-21: Advantageous Coupling Between Autonomous Groups

The previous two sections have dealt with adverse and advantageous coupling where open loops exist. As seen in Figure 2-20 and Figure 2-21, no closed-circuit loops exist and therefore feedback behavior between autonomous groups is impossible. The following section addresses the specific condition where closed-loop feedback behavior does exist.

System Stability: Amplification Index Criterion

Whenever feedback loops exist, the possibility of unstable or amplifying behavior exists. This behavior is acutely manifest when a group's good intentions to improve performance actually cause the system state to worsen due to the feedback response by other parts of the system. To assess whether or not a system is intrinsically stable, a criterion, called the Amplification Index, will be used. The Amplification Index is the negative ratio of the magnitude of the expected output given an input where the input is the change to a certain state variable and the output is the delayed system feedback to that same state variable that is the consequence of any interdependencies.

The Amplification Index, **AI**, is defined as

$$\mathrm{AI} = \frac{\text{Product of Off-Axis Partial Differentials}}{\text{Product of On-Axis Partial Differentials}}$$

The criteria for stability are as follows:

AI<0	Amplified Convergent
AI=0	Unamplified
0<AI<1	Attenuated Convergent
AI=1	Oscillatory
AI>1	Divergent

Every system containing an interdependence loop, such as the one shown in Figure 2-22 has the ability to demonstrate any one of the above behaviors. Amplified Convergent is a type of feedback behavior where the initial change causes a positive feedback reaction by another group. Unamplified behavior occurs when no complete loop exists; one of the terms in the matrix have a value of zero. Attenuated convergent behavior occurs when the Amplification Index is between zero and one. Negative feedback characterizes this condition where the magnitude of the original change is mitigated by the reaction from another group. Oscillatory behavior occurs when the Amplification Index is precisely one. In this situation, any action by the principal group creates an equal and opposite reaction by another group. The result is oscillatory behavior with no net change to the system state over time. Divergent behavior is the last of the Amplification Index criteria. When the Amplification Index is greater than one, the system will exhibit divergent behavior where any action made by the principal group will cause a reaction that is greater in magnitude than the original change. Such a system is intrinsically unstable and control is extremely difficult to manage.

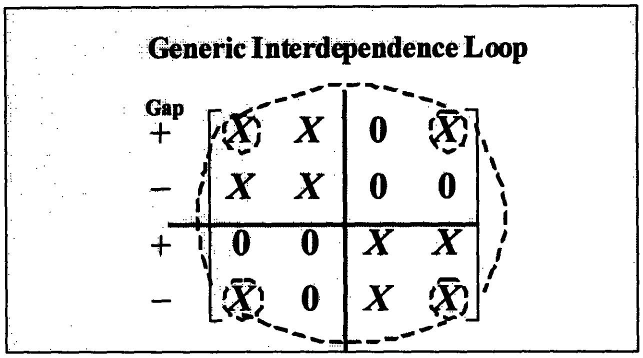


Figure 2-22: Generic Interdependence Loop

A simple mathematical model was developed to characterize the system response for each of the aforementioned conditions. The result is shown in Figure 2-23.

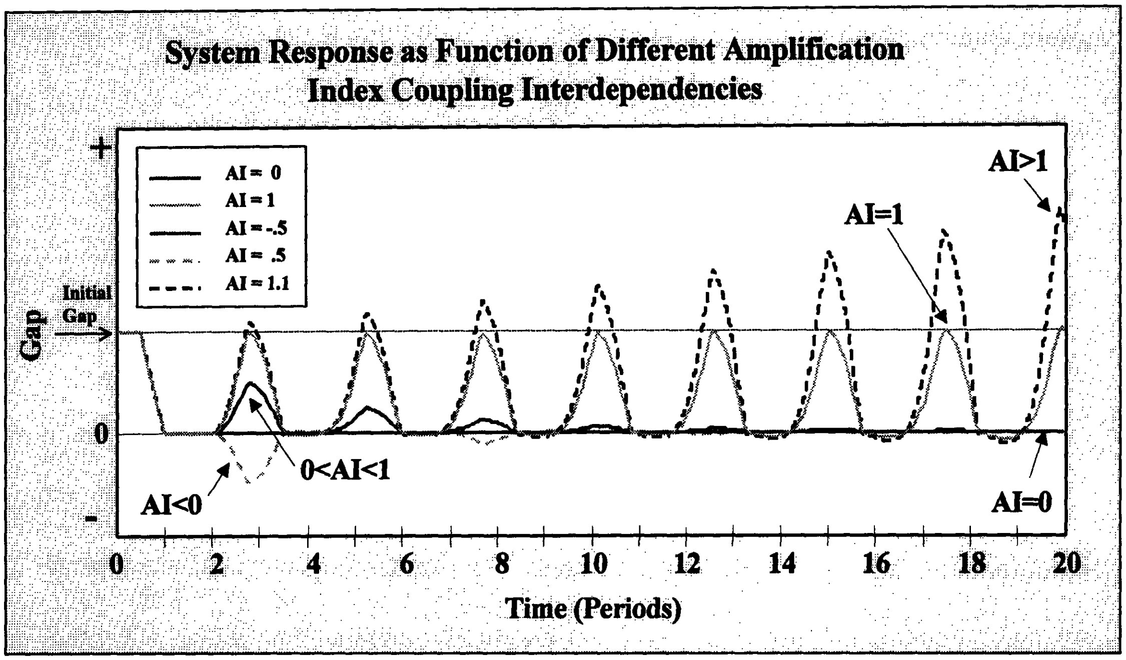


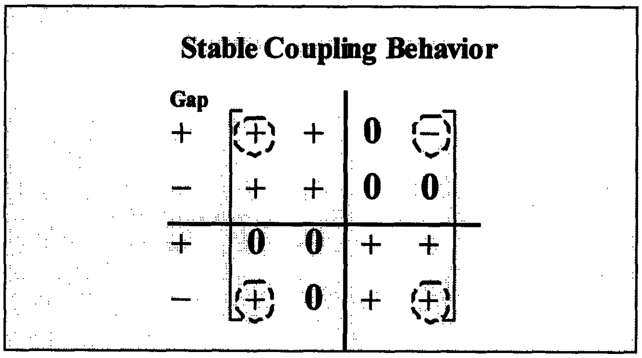
Figure 2-23: System Response Due to Different Coupling Interdependencies

While the specific behavior of any feedback system is contingent upon the delay period between the initial action and the reaction, the general behavior will be similar to those illustrated in Figure 2-23. As seen, depending on the value of the Amplification Index, the dynamic system behavior will be either stable or unstable. Each of these conditions will be discussed in the following sections.

Stable Systems

Stable systems can be considered convergent – that is, they will always goal seek toward their objectives. All systems having unilateral interdependence with another group will be inherently stable. If the first group makes a change to optimize its own objectives, it will affect another group, but it is not possible for this interdependence to feedback to the original group in any form.

It is also possible for two groups coupling with each other (therefore possessing closed interdependence loops) to be stable. The situation encountered in Figure 2-24 will always result in stable behavior.

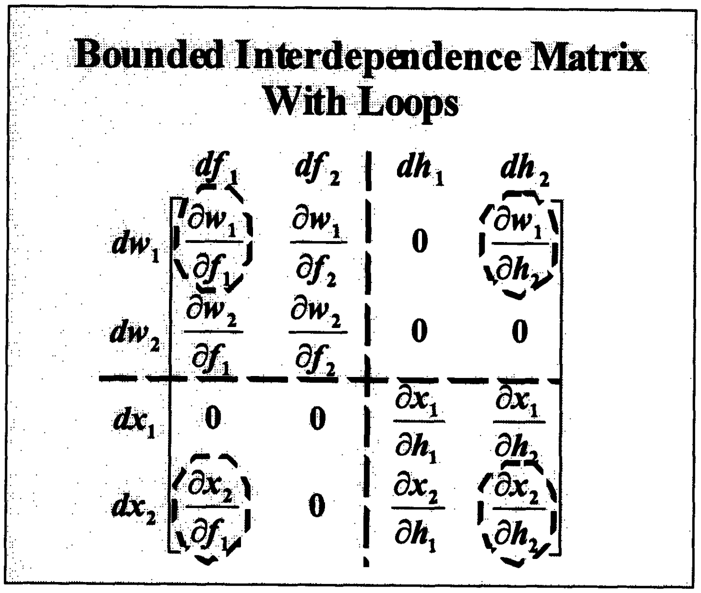


$$\begin{array}{c|cc|cc} \text{Gap} & & & & \\ + & \boxed{+} & + & 0 & \boxed{-} \\ - & + & + & 0 & 0 \\ \hline + & 0 & 0 & + & + \\ - & \boxed{+} & 0 & + & \boxed{+} \end{array}$$

Figure 2-24: Stable Coupling Behavior

Whenever a loop such as the one comprised by A_{11}, A_{41}, A_{44}, A_{14}, exists, the Amplification Index, **AI**, should be computed to determine if system stability criteria have been met.

Thus, if the Amplification Index is less than 1, the coupling will be stable irrespective of whether the coupling is advantageous or adverse. For the design matrix shown in Figure 2-25,



$$\begin{array}{c|cc|cc} & df_1 & df_2 & dh_1 & dh_2 \\ dw_1 & \boxed{\frac{\partial w_1}{\partial f_1}} & \frac{\partial w_1}{\partial f_2} & 0 & \boxed{\frac{\partial w_1}{\partial h_2}} \\ dw_2 & \frac{\partial w_2}{\partial f_1} & \frac{\partial w_2}{\partial f_2} & 0 & 0 \\ \hline dx_1 & 0 & 0 & \frac{\partial x_1}{\partial h_1} & \frac{\partial x_1}{\partial h_2} \\ dx_2 & \boxed{\frac{\partial x_2}{\partial f_1}} & 0 & \frac{\partial x_2}{\partial h_1} & \boxed{\frac{\partial x_2}{\partial h_2}} \end{array}$$

Figure 2-25: Bounded Interdependence Matrix With Loops the Amplification Index is given by

$$\mathbf{AI} = \frac{\dfrac{\partial x_2}{\partial f_1} \cdot \dfrac{\partial w_1}{\partial h_2}}{\dfrac{\partial w_1}{\partial f_1} \cdot \dfrac{\partial x_2}{\partial h_2}} \qquad (2\text{-}10)$$

Since $\partial w_1/\partial h_2$ is the only partial differential having a negative value, the overall sign of the Amplification Index must be negative and the system is therefore stable. The system will eventually converge.

Unstable Coupling

Instability in a system occurs when actions intended to improve performance (i.e. come closer to the objectives) actually worsen the state of the system. Frequently, the dynamic complexity of systems or large time constants between action and reaction make system instability difficult to recognize. Despite this, unstable coupling behavior is not uncommon. A *death spiral* is a coined term referring to companies which shortchange their long term fundamentals for short term results. The short term improvement comes at the cost of a longer term reaction that makes the state of the system even worse, further encouraging this behavior if the company is to maintain an image of robust earnings. Slowly, the company will be run into the ground.

Unstable behavior resulting in divergence occurs when the Amplification Index, **AI**, is greater than one. This condition arises when 1) the product of the off-axis coupling elements is greater than the product of the diagonal elements in the loop and 2) the sign of product of all the elements in the loop is positive. This is illustrated in Figure 2-26.

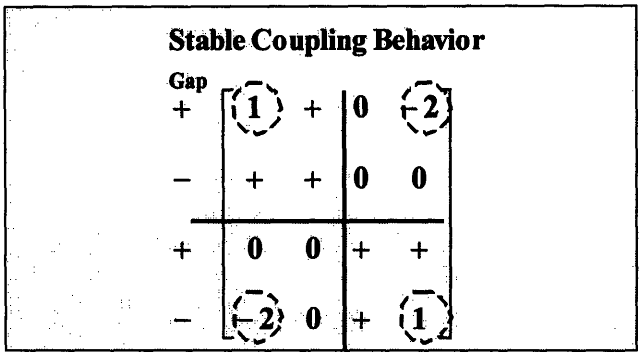


Figure 2-26: Unstable Coupling Behavior

As Figure 2-26 shows, the Amplification Index is equal to four. Any action by the first group (upper left quadrant) to improve its performance will cause a reaction by the second group that is greater in magnitude, but opposite in sign from the original change. The net result, after the delay between action and reaction, is that the overall state of the system is worse.

The above descriptions of advantageous and adverse coupling along with stable and unstable systems fully describe the first four types of coupling: adverse stable, adverse unstable, advantageous stable, and advantageous unstable. The following sections describe the remaining two types of coupling: external and temporal.

External Coupling

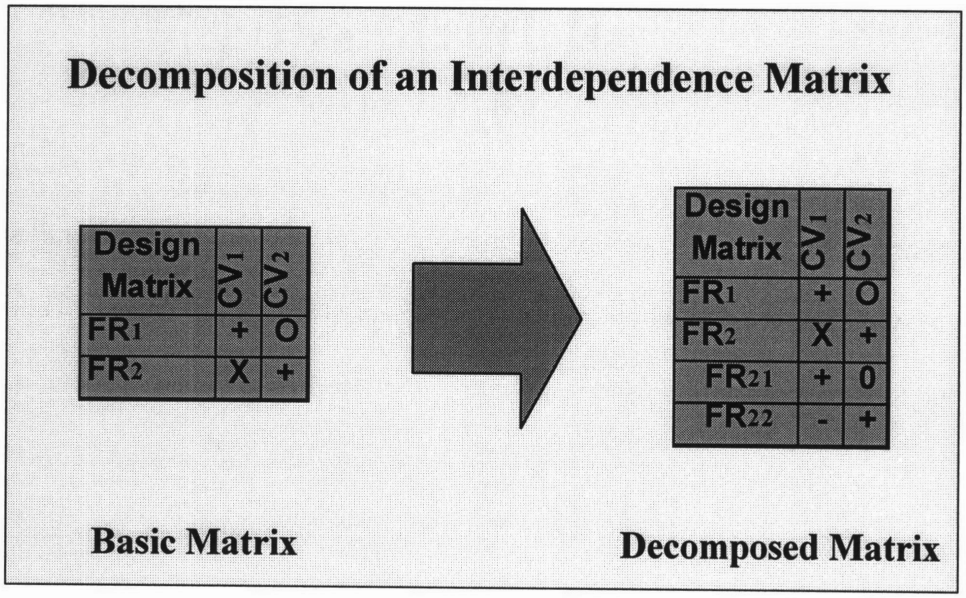
External coupling is caused by an exogenous variables outside the system. Depending on the way this variable reacts with the system of interest, the coupling may be advantageous or adverse.

Temporal Coupling

Temporal coupling occurs as a result of delays in a system. When changes are made to the state of a system, there may be transitory couplings which temporarily affect the state objectives (either positively or negatively) of some other part of the system. Since there is no long term net change, i.e. an initial deviation from an objective will be reversed after a delay without the intervention of any other part of the system, it is often not clear how to characterize the interaction. Should the temporal coupling even be considered? In general, a temporal coupling should be treated as any other coupling if its temporary effect is strong enough to cause reaction in any other part of the system. Since, for temporal coupling, the nature of the effect upon the system will depend upon the elapsed time from the original change, neither a "+" or "-" demarcation can be used. Instead, the temporal coupling should be identified by an "**X**".

2.3.2.9 Decomposition

Decomposition is an essential element of OOAD. As the example in Section 2.3.2.5 showed, situations will frequently arise where the coupling between the control and objective variable is not explicitly positive or negative. In this circumstance it is necessary to decompose the interdependence matrix to a level where meaningful relationships can be obtained. With decomposition of the objective variables, it is necessary to decompose along lines that will make the resultant decomposed matrix elements either explicitly positive or negative. This is shown in Figure 2-27.



Decomposition of an Interdependence Matrix

Design Matrix	CV_1	CV_2
FR_1	+	O
FR_2	X	+

Basic Matrix

Design Matrix	CV_1	CV_2
FR_1	+	O
FR_2	X	+
FR_{21}	+	0
FR_{22}	-	+

Decomposed Matrix

Figure 2-27: Decomposition of an Interdependence Matrix

Figure 2-27 shows FR_1 decomposed into two sub-objectives, FR_{11} and FR_{12}. Decomposition of objective variables is best explained by way of example. The productivity improvement example of 2.3.2.2 furnishes a good illustration. In Equation (2-8), the following design matrix representation was developed.

$$\begin{Bmatrix} \dfrac{dP}{dt} \\ \dfrac{dWOT}{dt} \end{Bmatrix} = \begin{bmatrix} + & 0 \\ \dfrac{TSW-TPI-OT}{WIR}\cdot\dfrac{\partial P}{\partial TPI}+\dfrac{-P}{WIR} & \dfrac{P}{WIR} \end{bmatrix} \begin{Bmatrix} \dfrac{dTPI}{dt} \\ \dfrac{dTSW}{dt} \end{Bmatrix}$$

As shown, the off-axis element, A_{21}, is not explicitly positive or negative. This can be explained in the following manner. The numerator of the left term in the element, $TSW - TPI - OT$, will always be positive since the time spent working, TSW, cannot drop below zero no matter how much the other time, OT, and the time on productivity improvement, TPI, is increased. Likewise, the work introduction rate, WIR, and the change in productivity per change in time spent on productivity improvement, $\partial P/\partial TPI$, will always be positive. This means that the left term will always be positive. By contrast, the right term, $-P/WIR$ will always be negative. This must be the case since both the productivity, ***P***, and the work introduction rate, ***WIR***, must be positive at all times.

Since the right and left terms of the off-axis element have opposite signs, the overall sign of the matrix element will depend upon whichever term is greater at any point in time. This observation led to the "**X**" notation used in Equation (2-9), repeated below.

$$\begin{Bmatrix} \dfrac{dP}{dt} \\ \dfrac{dWOT}{dt} \end{Bmatrix} = \begin{bmatrix} X & 0 \\ X & X \end{bmatrix} \begin{Bmatrix} \dfrac{dTPI}{dt} \\ \dfrac{dTSW}{dt} \end{Bmatrix}$$

As the above interdependence matrix shows, the "**X**" notation poorly conveys the way in which changes to the control variable (time spent on productivity improvement, ***TPI***) affect the work completed on-time, ***WOT***, objective. In this case, decomposition of the second objective variable in the interdependence matrix will clarify the nature of the interdependency. By studying Equation (2-8), it can be seen that there are two components which affect how a change in the time spent on productivity improvement, ***TPI***, will influence the amount of work completed on time. The first component is the net time available to produce product. This is the left term in matrix element A_{21}. The second component is the current level of productivity, ***P***. By decomposing, FR_{21} becomes the work completed on time due to available time (the left component) and FR_{22} becomes the work completed on time due to productivity (the right component). As shown in Figure 2-27, this decomposition eliminates the previous ambiguity and shows how increases to the time spent on productivity improvement, ***TPI***, will decrease the work completed on time when the productivity is high compared to the available time to produce products. In the same way, it shows the converse; when the productivity is low compared to the time available to produce products, increases in the time spent on productivity improvement will increase the work completed on time.

While the previous discussion focused on the decomposition of objective variables, it is also sometimes possible to decompose control variables. For example, the capacity of equipment for a manufacturing line can either be set on an aggregate basis or the capacity of each piece of equipment can be set individually. Sometimes it may be useful to refer to the total capacity control variable (with each element of capacity moved the same amount whenever a change is made) while at other times it may be advantageous to refer to the capacity of the individual machine bottleneck. Decomposition allows both levels of control to be represented simultaneously.

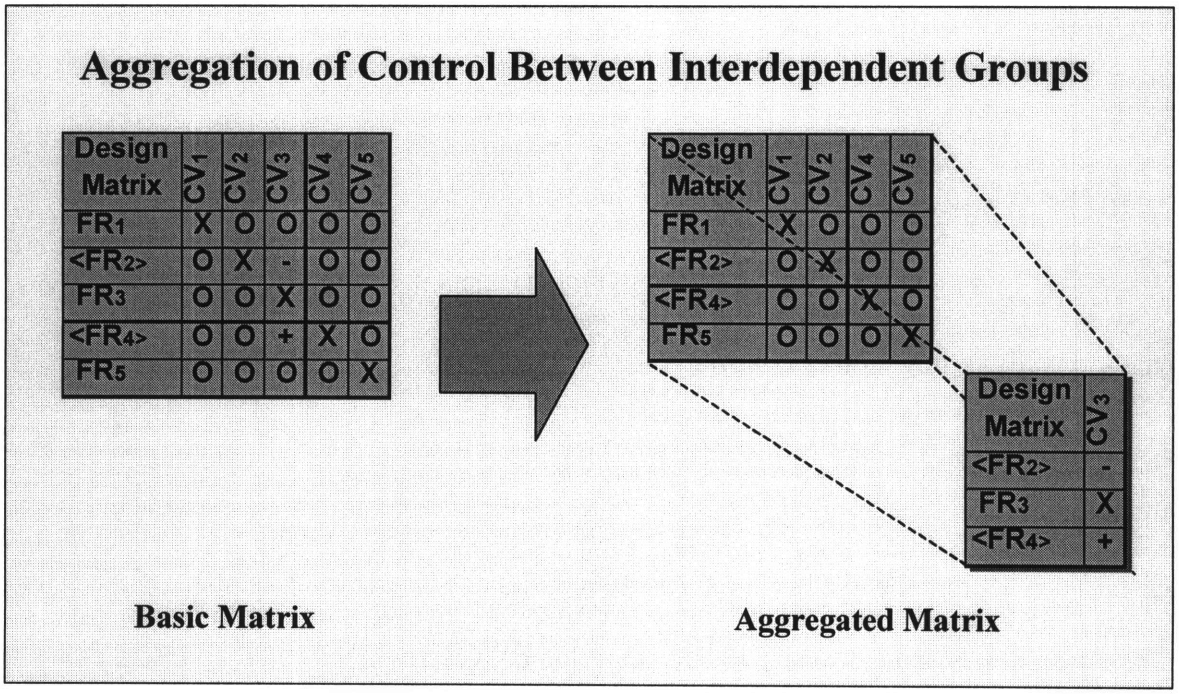
2.3.2.10 Aggregation

There are two principal types of aggregation in OOAD. The first is aggregation of objectives and the second is aggregation of control.

Aggregation of objectives is commonplace. For example, a company may have several divisions, each selling different product. On an individual basis, each division may have its own objective for sales revenue. These sales revenues objectives are often rolled up into the overall sales revenue for the company. By aggregating the objectives, the strength of each division's control variable on the overall corporate sales revenue can be determined. As shown in this example, linear superposition is the most straightforward form of aggregation. It is also possible, yet less frequent to aggregate by means of the product or quotient of different objectives.

Combining the defect rate objective with the sales rate objective to determine the quality objective for the manufacturing process in parts per million (ppm) is one such example.

The second type of aggregation is aggregation of control. Aggregation of control is not so much where control variables are combined into a single control variable (pointlessly reducing the granularity of control), but where control previously residing within one of two interdependent groups is relinquished and given to an aggregate or meta-object. This aggregate object is a senior object in the control hierarchy and makes decisions with the relinquished control variable that optimizes the objective performance of the different objective variables residing among the different groups. This form of aggregation is illustrated in Figure 2-28, below:



Aggregation of Control Between Interdependent Groups

Design Matrix	CV_1	CV_2	CV_3	CV_4	CV_5
FR_1	X	O	O	O	O
$<FR_2>$	O	X	-	O	O
FR_3	O	O	X	O	O
$<FR_4>$	O	O	+	X	O
FR_5	O	O	O	O	X

Basic Matrix

Design Matrix	CV_1	CV_2	CV_4	CV_5
FR_1	X	O	O	O
$<FR_2>$	O	X	O	O
$<FR_4>$	O	O	X	O
FR_5	O	O	O	X

Design Matrix	CV_3
$<FR_2>$	-
FR_3	X
$<FR_4>$	+

Aggregated Matrix

Figure 2-28: Aggregation of Control

As Figure 2-28 shows, aggregation of control enables a control hierarchy to be established. This control hierarchy allows previously distributed control to be centralized, eliminating the sources of adverse behavior.

2.3.3 Multi-Objective Optimization: Methods to Improve System Performance

2.3.3.1 Survey of Techniques to Reduce Adverse Couplings

A number of techniques can be used to improve system performance. These techniques strive to improve the performance of each organization on a global basis, as well as local basis. Compared with traditional process improvement techniques which merely achieve local, but not necessarily global, refinement, OOAD provides two principal benefits through boundary analysis. Firstly, boundary analysis enables methods to be applied which achieve local improvement that is also global. Secondly, boundary analysis enables global performance improvement through the reduction of unwanted coupling behavior between the disparate organizations.

In general there are three methods of achieving global performance improvement. To each of these, a variety of techniques can be employed. These are enumerated below.

A) Improve the system locally where it achieves global improvement, i.e. prevent "improvements" with costs to other parts of the system that meet or exceed the local benefit

   1. Select or add different controls
   2. Improve the methods by which objectives are measured

B) Attain global improvement by reducing or eliminating unwanted coupling interdependencies

   1. Select or add different controls

3. Select different objectives
4. Expand sphere of influence or reallocate objectives and controls to form different groupings
5. Passively buffer to reduce coupling strength

C) Attain global improvement by preventing coupling behavior without necessarily eliminating existing interdependencies

6. Aggregate control
7. Actively buffer
8. Establish Parity-Based Transactional Interdependence (Efficient Marketplace Model)

Each technique has a unique effect on the system. Selecting different controls changes the control variables employed by a group to attain its objectives.

Select or Add Different Controls

Different controls affect the objective state or output of the system through different mechanisms. This fundamentally changes the nature of the system and can result in either improved local performance or reduced coupling interdependencies. For example, it may be found that on-hand inventory costs can be reduced more effectively by decreasing batch sizes in manufacturing operations rather than decreasing finished goods inventory (FGI) levels. Compared to a reduction of FGI which increases the frequency of stockouts (adversely affecting another objective), smaller batch sizes may actually improve manufacturing lead times and quality (advantageously affecting other objectives) without affecting the frequency of stockouts (an interdependence existing with the other control variable). Thus, the selection of batch size as a control variable reduces undesirable interdependencies as well as improving local performance.

Improve the Methods by Which Objectives are Measured

The second technique, improving the methods by which objectives are measured, recognizes that frequently the objectives used to influence behavior and measure performance are really only estimators or predictors for actual objectives which cannot be measured directly. For example, at Ford Motor Company, actual return on sales can only be measured post-facto, at which time little can be done to change the outcome. In light of this, a predictive or calculated return on sales (ROS) is used instead. Since this calculated ROS does not perfectly correlate with the actual ROS, it is possible that in pursuit of optimizing the calculated ROS to meet its objective, decisions are made that actually serve to *reduce* the actual return on sales that will eventually be observed. In this sense, if the calculated ROS is improved to be a more accurate predictor of actual ROS, local performance will be improved.

Select Different Objectives

Selecting different objectives will fundamentally change the coupling interdependencies between the new and pre-existing objectives. Clearly this can result in the reduction or elimination of coupling interdependencies.

Reallocate Objectives and Controls to Form Different Groupings

Reallocating the objectives and controls to form different groupings is the fourth technique that can be employed. Sometimes the current grouping of objectives doesn't make sense. It may be possible to completely eliminate coupling interdependencies by merely reassigning the responsibility of different objectives to different organizations. In practice this technique may be very difficult to apply. Regrouping requires fundamental change to the organizational structure and this may encounter either physical or personal barriers.

Passively Buffer

Passive Buffering is a technique that can be used when a coupling interdependency is a function of other state variables. If these variables can be changed, the coupling strength can be reduced.

Consider the situation where the objective state, $\boldsymbol{F}$, is a function of two variables, $\boldsymbol{A}$, $\boldsymbol{B}$, where $\boldsymbol{A}$, is either an exogenous variable or a control variable possessed by another group. $\boldsymbol{B}$ is controlled by the group with objective $\boldsymbol{F}$. If $\boldsymbol{F=A+B}$, $\boldsymbol{A}$ will couple with $\boldsymbol{F}$ to a constant degree, independent of changes by $\boldsymbol{B}$. This is observed by noting that $\boldsymbol{\partial F/\partial A = 1}\ \Re(\boldsymbol{B})$. In this case, there is no way to reduce the coupling strength of $\boldsymbol{A}$. However, if the function is different, such that $\boldsymbol{F=AB}$, $\boldsymbol{\partial F/\partial A = B}$. In this case, the coupling strength of $\boldsymbol{A}$ is a direct function of the *state* of $\boldsymbol{B}$. If $\boldsymbol{B}$'s state is changed to a lower level, the off-axis coupling due to $\boldsymbol{A}$ will be reduced, achieving the desired result. This is called *passive buffering*. The reason this is called passive buffering (as opposed to active buffering) is because the coupling behavior of $\boldsymbol{F}$ due to $\boldsymbol{A}$ is a function of the state of $\boldsymbol{B}$. In other words, the state of $\boldsymbol{B}$ passively dampens or amplifies the change signal of $\boldsymbol{A}$ through its normal path of interaction.

Passive buffering also works for many different forms of interaction as well. If $\boldsymbol{F=AB+B}$ instead of $\boldsymbol{F=AB}$, the magnitude of $\boldsymbol{B}$ can be decreased to passively buffer the coupling caused by $\boldsymbol{A}$

without reducing the strength of ***B*** as a control variable. This can be observed by noting that $\partial F/\partial B = A+1$ and therefore is not a function of the state of ***B***. Of course this assumes that when the magnitude of ***B*** is reduced to passively buffer the system, there exists some other auxiliary control variable that can be used to re-equilibrate the system.

The final method for improving global performance involves more of a management of interdependencies than a reduction or elimination of them. Three different techniques can be used to reduce unwanted coupling behavior without requiring the interdependencies to be eliminated. First, control can be aggregated.

Aggregate Control

Aggregating control reduces the independence of decision-making. No longer does a group have the autonomous ability to change the state of its control variable, optimizing its performance at the expense of another group elsewhere in the system. Instead a meta-grouping (i.e. an executive committee or senior management forum) will make changes to the aggregated control variable that result in global improvement.

Actively Buffer

The second technique to reduce coupling behavior is to actively buffer the system. Active buffering explicitly adds an additional component that counteracts the adverse coupling effect. Typically this technique manifests itself in the form of hedging. For example, in unionized plants, the local union effectively controls the number of hours union members work. If the union wants to improve the state of one of its objectives such as the average wage per worker, it may call a strike (reduce the number of hours worked per union member to zero). The intent and effect of this change is to adversely couple with one of the employers objectives, namely net revenue. To eliminate this coupling behavior there are several things the employer can do. On one hand it can attempt to break the union and replace the workforce with a non-unionized labor force. This eliminates the coupling interdependence by eliminating union control. Alternatively the company can elect to take out an insurance policy or financial hedge against lost revenues in the event of a strike. By adding the hedge the company does not change the control of number of hours worked per union member nor does it change the effect that a strike has on sales revenue. Rather, the hedge actively buffers the system from the effect of the coupling behavior by compensating for the lost sales. The financial return from the hedge offsets the revenue lost, leaving the net return unaffected.

Establish Parity-Based Transactional Interdependence

The final technique available to improve global performance by reducing coupling behavior is the use of Parity-Based Transactional Interdependence. This technique basically involves a voluntary contractual agreement where one group will make a change that adversely affects its own performance while benefiting the state of the other group and the other group makes a reciprocal change. The net result is that both groups benefit and global system improvement takes place. This occurs despite the fact that the nature of the coupling interdependencies hasn't changed. However, for this technique to be saleable, it is necessary that 1) both groups adversely couple with each other or one group adversely couples with another and a transaction method exists to exchange value between the two groups, 2) a common method of valuation exists to perform a cost – benefit analysis between different objectives, and 3) the benefit to the second group exceeds the value of the cost to the first group when a change is made by the first group.

Each of the techniques described above are useful in increasing global system performance. Despite this, often only a small subset will be used. It is essential that the points of leverage of the system be identified and high leverage solutions are found. Additionally, the anticipated benefit of a change will have to be weighed against the cost of executing the change. Adherence to this process will ensure that only superior solutions become implemented.

2.3.4 Case Example: Barilla SpA

Consider the case of Barilla, the Italian pasta producer [20]. In the early 1990s, the company became increasingly plagued with product demand volatility. These demand fluctuations imposed heavy burdens on the company's manufacturing and distribution systems. The company's distribution centers typically stocked two and one half weeks of inventory with periodic weekly reviews. Order lead times were typically one and a half weeks. Despite the large amount of inventory, the demand fluctuation was significant enough to cause substantial stockouts. This demand fluctuation not only adversely affected Barilla's profitability, but hurt its reputation as a reliable source for dry pasta.

The sources of demand fluctuation were unclear. Aside from fundamental demand fluctuation due to the vagaries of retail customer purchases, there was no obvious cause of demand volatility. After reviewing historical data, it was believed that increased demand fluctuation potentially came from three sources. The first source was the sales and marketing organization within Barilla. An integral component of Barilla's sales strategy was the use of trade promotions. These promotions stimulated demand for short periods of time by allowing customers to order unlimited amounts of pasta at a certain percentage discount. Similarly, Barilla also offered volume discounts by covering all shipping costs (2-3% of order cost) for orders placed in full truck-load quantities. The trade promotions exacerbated demand fluctuation by encouraging buyers to purchase quantities in excess of current demand and to let inventories fall below reorder points on the expectation of an upcoming promotion. Thus, these price policies, with the intent of increasing net revenue to Barilla, taxed the operational side of the organization, making it difficult to maintain service levels.

The second source of demand variation was believed to come from the downstream distributors reduced inventory levels. As Barilla and other manufacturers increased the variety of their product offerings, distributors and retailers were forced to decrease their inventory holding if total shelf space was to remain constant. This, in turn, reduced their ability to damp out demand fluctuation; instead they passed it upstream where it was acutely felt at Barilla.

The third and final source of variation was believed to come from the distributors' or retailers' order policies. Most did not use a systematic reorder process. At best they used a periodic review process that lacked any forecasting packages or other analytic tools for determining order quantities. Any orders coming from the distributors that were not lock-step with their customers were seen, by Barilla, as additional sources of variation that the company would be forced to absorb.

To address this adverse coupling between demand and inventory stockouts, several proposals were evaluated by the company. The first was to push the distributors and retailers to increase the amount of inventory carried. While this change would increase service fill levels, it would adversely couple with the companies' objectives to maintain or reduce inventory holding costs. Therefore, it was expected to meet substantial resistance by those who would have to bear the additional costs. The second proposal was to increase the manufacturing flexibility within Barilla's factories, but this idea was quickly dismissed due to strict temperature and humidity requirements required in the production of the pasta. All but limited changes to the production process would be expected to adversely affect the quality of the pasta. The third proposal was to limit the frequency and magnitude of trade promotions my the marketing and sales organization within Barilla. While this would reduce demand volatility by making customers' orders more steady, it was strongly believed to adversely affect sales revenues. Thus, it too was rejected. A derivative approach that addresses the same problem would be to build up inventory prior to Barilla's promotions to its customers. This approach enables promotions to be continued and would increase inventories only momentarily, but would require coordinated planning between manufacturing and marketing and sales.

The final approach proposed was to install what was called the Just-in-Time Distribution (JITD) system. This system would in effect entirely bypass the order process currently used by the distributors. Instead, all sell-through data as well as current inventory levels would be relayed back to Barilla daily. Based on this data, Barilla would determine distribution quantities as well as develop forecasts for its production schedules. This change was expected to reduce the distributor-induced variation, increasing service levels and allowing overall inventory levels to be reduced. The main drawback to this process was that it required the distributors to cede control of their order process to Barilla.

After much consternation and the necessity for Barilla to perform proof-of-concept internal pilot tests, the latter two approaches were implemented. Inventory was stockpiled in anticipation of promotions so that production would not get any “surprises”. Distributors eventually agreed to the JITD system and the resultant reduction in required inventory and increase in customer service rates were dramatic. Inventory levels dropped by over 50%, customer stockout rates decreased by 80%, and demand volatility decreased by well over 50%.

Analyzing the changes Barilla considered from an OOAD perspective, it can be seen that the different proposals employ different techniques to reduce the coupling effect of demand. The relationship between the proposals and a select set of Barilla corporate objectives are shown below:

Barilla SpA Interdependence Matrix		**Exogenous Variables**	Demand EV_1	**Marketing & Sales CVs**	Frequency of Sales Promotions CV_1	Discount Level in Sales Promotions CV_2	**Logistics CVs**	Inventory Levels CV_3	**Manufacturing CVs**	Manufacturing Flexibility CV_4	**Aggregated CVs**	Promotional Inventory Build-Up CV_5	Just-in-Time Distribution (JITD) CV_6
Marketing & Sales Objectives													
FR_1	Max Sales Revenue		+		+	-		0		0		0	0
Logistics Objectives													
FR_2	Min # Stockouts		+		+	+		-		-		-	-
FR_3	Min Inventory Holding Costs		0		0	0		+		0		0	-
Manufacturing Objectives													
FR_4	Produce Highest Quality Product		0		0	0		0		-		0	0

Figure 2-29: Barilla SpA Interdependence Matrix

The first proposal to reduce the adverse coupling effect of demand fluctuation was to increase inventory carried. This technique *passively buffers* the system against stockouts since stockouts are a direct function of inventory levels. Similarly, adding manufacturing flexibility *passively buffers* the system since the flexibility is available whether or not its use is required. The third approach, reducing the number of promotions, were it enacted would require some form of *aggregation of control* since it is against Marketing & Sales wishes to make such a change. Increasing inventory levels prior to promotions is an *active buffering* technique. It actively buffers the system since the inventory is only added when the demand is going to increase due to a promotion. With this approach the additional inventory will offset the change in demand caused by the promotion and the original inventory level will be left unaffected. The final technique considered was the JITD system. This system makes use of the *parity-based transactional interdependence.* The distributors must agree to give up decision-making power about their orders on the basis that Barilla will reduce their stockout frequency and required on-hand inventory. Likewise, Barilla will have the additional responsibility and workload of making the distributors' orders for them on the assumption that the increased workload will result in greater demand stability.

Because the demonstrated advantages of the JITD system improved the performance of both Barilla and its distributors, the JITD approach was selected over the other proposals.

2.4 Summary of Concepts in OOAD

1) There exist two types of objectives: goal seeking and maximum/minimum objectives

2) There exist two types of feedback loops: balancing and reinforcing

3) Mathematical representation of feedback loops maps the relationship between the input (control variable) and the output (objective variable)

4) An interdependence exists when a change in one control variable (CV) in a system changes the state of one or more objective variables (FR) other than the principle objective given all else held constant

5) An Axiomatic Design Matrix is a useful tool for abstracting the relationship between the control variable and the objective variable

6) Four representations exist for characterizing the nature of interdependencies in a design matrix: O, X, +, -

7) The performance of real-world systems is critically dependent upon the span of authority that exists to affect the outcome. This span of authority or sphere of influence defines a boundary within the system, i.e. an object or group

8) Encapsulation does not define where boundaries exist, but rather the degree of independence that one object has from another

9) Ideal systems have fully encapsulated objects, each fully meeting their objectives

10) The Bounded Interdependence Matrix captures the location of boundaries in a system and the off-axis quadrants in the matrix capture the couplings across the boundaries

11) Coupling is the presence of an interdependence between two groups

12) Six different types of interdependencies exist: adverse stable coupling, advantageous stable coupling, adverse unstable coupling, advantageous unstable coupling, external interdependence, and temporal interdependence

13) Adverse stable coupling is characterized by improvement to one objective that adversely affects another but in such a way that the feedback effect, if any, has lower magnitude than the original change

14) Advantageous stable coupling is characterized by improvement to one objective that also improves another but in such a way that the feedback effect, if any, has lower magnitude than the original change

15) Adverse unstable coupling is characterized by improvement to one objective that adversely affects another but in such a way that the feedback effect makes the state of the system worse than before the change had been made

16) Advantageous unstable coupling is characterized by improvement to one objective that also improves another but in such a way that the feedback effect has greater magnitude than the original change. This effect accelerates the system toward its goals. However it will also cause one or both of these goals to be overshot, changing the state of coupling in the system to adverse unstable

17) External coupling is characterized by a variable whose change in state is not a direct consequence of changes to other control variables within the system. This external variable (EV) couples with the system if changes in its state change the objective state of one or more objective state variables (FRs) in the system

18) Temporal coupling is characterized by transitory coupling behavior which causes no net change to the off-axis objective variable (after an extended period), despite permanent change by the affecting control variable

19) The Amplification Index is a criterion for system stability

20) Aggregation is a method of encapsulating control of multiple objects within a larger object. This meta-object will manage couplings between the aggregated objects as if they were intra-object interdependencies

21) The degree of coupling can be measured in two ways: $\frac{\partial FR}{\partial CV}$ or $\frac{\sigma_{FR} / FR}{\sigma_{CV} / CV}$

22) Coupling can be can be reduced by 1) decreasing the strength of off-axis elements relative to on-axis elements or 2) aggregating (or centralizing) control

23) Decomposition of the interdependence matrix should be used when the sign of the interdependence changes with the state of the independent control variable. This is not an effective approach when the sign of the interdependency is a function of other variables in the system

2.5 Appendix: Review of Key Concepts Leveraged in OOAD Development

The development of Object-Oriented Axiomatic Design has drawn from three disciplines in three different fields. Axiomatic Design from Mechanical Engineering, System Dynamics from Management Science, and Object-Orientation from Computer Science are each well-developed. The following sections are intended to provide only the highest-level review of the subjects as pertinent to this work. For a more fundamental treatment of any subject, the reader is directed elsewhere[3].

2.5.1 Review of Axiomatic Design

Axiomatic Design is a formal design methodology developed by Nam P. Suh. For further review of topics covering Axiomatic Design refer to: [11], [21], [22], [23], [24], [25], [26], and [27].

2.5.1.1 Objectives

Axiomatic Design has one principle objective: to provide a formal design methodology for multi-objective problems. Suh has identified two design axioms: the Independence and Information Axioms [26]. These axioms form the basis of the Axiomatic Design methodology.

---

[3] Seminal works and other relevant readings will be referenced at the beginning of each section.

Axiomatic Design strives to provide the designer or design team the ability to simultaneously meet multiple design objectives with the minimum compromise and design complexity. To achieve this, it is necessary to, 1) be able to critique the quality of the synthesized design, and, 2) be able to provide the framework necessary to insure an optimal solution. Axiomatic Design pursues both of these goals.

2.5.1.2 Key Concepts

Four key concepts are integral to the Axiomatic Design Methodology [27]:

1. Independence / Information Axiom
2. Matrix Representation
3. Domains – separating the objective / requirement from the means (control / embodiment)
4. Decomposition (precedence relationships)

Independence and Information Axioms

Together, the independence and information axioms form the foundation of Axiomatic Design. The first axiom, the Independence Axiom, addresses the ability of a design to effectively meet two or more objectives. The second axiom, the Information Axiom, addresses the complexity of a particular design solution. This complexity is measured as both the information content embodied within the design, and the necessary information that must be externally applied to insure the design can meet its functional requirements.

The definitions of the two axioms are given below:

1. Maintain the independence of functional requirements
2. Minimize the information content of the design

These two axioms fully encompass Axiomatic Design. The intent behind the Independence Axiom is to make the design process straightforward and non-iterative. If independence of the functional requirements is maintained, there will be a single parameter of the design for each and every design requirement such that a change to any parameter will only affect one requirement of the design. The Information Axiom derives from the basic tenet that simpler designs, which sufficiently and adequately meet all functional requirements of the design, will be better than more complex solutions.

Matrix Representation

The mapping between design parameters and functional requirements is given by a design matrix. This mathematical relationship between FRs and DPs is shown below:

$$\begin{Bmatrix} FR_1 \\ FR_2 \end{Bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{Bmatrix} DP_1 \\ DP_2 \end{Bmatrix}$$

In an ideal design, the off-axis elements in the design matrix, A_{12} and A_{21}, will be zero. In the situation where they are not, coupling is said to exist. The strength of this coupling can be measured by taking the partials of the FRs with respect to DP, as shown below:

$$\begin{Bmatrix} \frac{dFR_1}{dt} \\ \frac{dFR_2}{dt} \end{Bmatrix} = \begin{bmatrix} \frac{\partial FR_1}{\partial DP_1} & \frac{\partial FR_1}{\partial DP_2} \\ \frac{\partial FR_2}{\partial DP_1} & \frac{\partial FR_2}{\partial DP_2} \end{bmatrix} \begin{Bmatrix} \frac{dDP_1}{dt} \\ \frac{dDP_2}{dt} \end{Bmatrix}$$

Coupling is reduced by either increasing the strength of the on-axis relationships or reducing the strength of the off-axis relationships. This can be achieved by selecting different DPs.

Domains

In Axiomatic Design there are four domains. These are: the customer domain, the functional domain, the physical domain, and the process domain. The previous sections on Axiomatic Design have focused on the functional domain (FRs) and the physical domain (DPs). A more formal approach recognizes that the functional requirements must be determined through some process. By establishing a one-to-one mapping from a customer want to a function of the design, i.e. a FR, each customer want is precisely met by a single element of the design solution. Likewise, the mapping between the process domain and the physical domain determines how each physical design parameter will be met through some sort of "manufacturing" process. A design matrix similar to the FR-DP matrix is used for the mapping between the DP-PV domains. Coupling can also exist.

Decomposition

The final significant concept in Axiomatic Design is decomposition. Decomposition is an integral part of Axiomatic Design. Unlike other design approaches which emphasize design synthesis, or OOAD which uses decomposition for the "mining" of embodied solutions, Axiomatic Design uses decomposition to drive the design process. Starting with the highest-level conceptual requirements, Axiomatic Design "zig-zags" from requirement (FR) to solution (DP) and then back to requirement (FR) in ever increasing layers of specialization. This process is repeated until the entire design is mapped out from the highest concepts to the minutest details.

At each level of the design hierarchy, the design matrices are evaluated for coupling. If the design is found to be coupled, then the designer must select new DPs and "back up" the design tree as much as necessary for an uncoupled solution to be found. If this prescribed process is strictly followed, the final design solution will be uncoupled; each customer want will be met by precisely one design parameter which can be easily realized through a single process variable.

2.5.2 Review of System Dynamics

For a review of System Dynamics refer to [14], [15], [16], [28], [29], [30], [31], [32], and [33].

2.5.2.1 Objectives

System Dynamics is a very general field with broad application. Depending on the specific situation, the objectives may be quite different. However, most applications of System Dynamics share the following objectives:

- Be able to model dynamic, nonlinear systems demonstrating transient behavior
- Obtain a more fundamental understanding of a system by reconciling and representing different mental models through the use of causal loop diagrams
- Simulate system behavior under varying policy structures

2.5.2.2 Key Concepts

There are four key concepts comprising the System Dynamics methodology. These are:

- Closed Boundary
- Feedback Loops as Building Blocks

- Level and Rate Variables
- Policy Structure

Jay Forrester, professor emeritus at MIT and founder of System Dynamics eloquently summarizes these four concepts in his paper, *Market Growth as Influenced by Capital Investment* [33]. In his own words, each concept will be discussed.

Closed Boundary

In defining a system, [it is best to] start at the broadest perspective with the concept of the closed boundary. The boundary encloses the system of interest. It states that the modes of behavior under study are created by the interaction of the system components within the boundary. The boundary implies that no influences from outside of the boundary are necessary for generating the particular behavior being investigated. So saying, it follows that the behavior of interest must be identified before the boundary behavior can be determined. From this it follows that one starts not with the construction of a model of a system but rather one starts by identifying a problem, a set of symptoms, and a behavior mode which is the subject of study. Without a purpose, there can be no answer to the question of what system components are important. Without a purpose, it is impossible to define the system boundary.

But given a purpose, one should then define the boundary which encloses the smallest permissible number of components. One asks not if a component is merely present in the system. Instead, one asks if the behavior of interest will disappear or be improperly represented if the component is omitted. If the component can be omitted without defeating the purpose of the system study, the component should be excluded and the boundary thereby made smaller.

An essential basis for identifying and organizing a system structure is to have a sharply and properly defined purpose.

Feedback Loops as Building Blocks

Inside the closed boundary one finds a structure of interacting feedback loops. The feedback loop is the structural setting within which all decisions are made. The feedback loop is a closed path. A decision is based on the observed state of the system. The decision produces action which alters the state of the system and the new state gives rise to new information as the input to further decisions. The feedback loop implies the circularity of cause and effect, where the system produces the decision which produces the action which produces change in the system. One has not properly identified the structure surrounding a decision point until the loops are closed between the consequences of the decision and the influence of those consequences on future decisions.

Level and Rate Variables

Within the feedback loop we find the next lower hierarchy of structure. To represent the activity within a feedback loop requires two and only two distinctly different kinds of variables – the levels and the rates. The levels represent the system condition at any point in time. In engineering, the level variables are often referred to as the system state variables. In economics, the system levels are often spoken of as stocks. The levels are the accumulations within the system. Mathematically they are integrations.

The rate variables represent the system activity. The rate equations are the policy statements in the system which define how the existing conditions of the system produce a decision stream controlling action.

The clear separation of system concepts into the two classes of variables – levels and rates – has interesting and useful consequences. The level variables are the integrations of those rates of flow which cause the particular level to change. It follows that a level variable depends only on the associated rates and never depends on any other level variable. Furthermore, in any system, be it mechanical, physical, or social, rates of flow are not instantaneously observable. No rate of flow can depend on the simultaneous value of any other rate. Rates depend only on the values of the level variables. If levels depend only on rates and rates depend only on levels, it follows that any path through the structure of system will encounter alternating level and rate variables.

Policy Structure

An important substructure exists within the equation that defines a rate variable. A rate equation defining a rate variable is a statement of system policy. Such a policy statement describes how and why decisions are made. A policy statement incorporates four components – the goal of the decision point, the observed conditions as a basis for decision, the discrepancy between goal and observed conditions, and the desired action based on the discrepancy.

A decision is made for a purpose. The purpose implies a goal that the decision process is trying to achieve. The policy statement that determines a rate variable does so in an attempt to bring the system toward the goal. The goal is sometimes adequately represented as a constant objective; more often the goal is itself a result of the past history of the system that has established traditions to guide present action. Whether or not the goal is actually achieved depends on how the system as a whole responds to the particular decision point. Usually, the competition for resource allocation results in the system falling short of most of the goals. The goal at the particular decision point is compared with the observed system condition as a guide to action.

One must distinguish observed conditions from the actual conditions of a system. A system model must incorporate both actual and apparent system levels (the levels describe the condition or state of the system. Where an important difference can exist between what the system is and what it is thought to be (and these differences are especially prevalent in the marketing sector of a company), one represents both, and explicitly shows how the apparent states arise out of the true states. A decision can be based only on the observed conditions, that is, the available information. Very often, substantial deviations exist between the true conditions of a system and the observed conditions. The discrepancy can arise from delay in recognizing changes in the system, random error, bias in not wanting to believe what is visible, distortion, insensitivity, and misinterpretation of meaning.

The policy statement makes a comparison of the goal and apparent condition to detect a discrepancy. The discrepancy may be in the form of a difference, a ratio, or some other indicator of lack of agreement. On the basis of the discrepancy, the policy describes the action to be taken.

2.5.3 Review of Object-Oriented Methodology

A large body of work has been written on Object-Oriented Analysis (OOA), Object-Oriented Modeling (OOM), and Object-Oriented Design (OOD). Seminal works can be found in [34], [35], and [36].

2.5.3.1 Description and Objectives

Object-oriented methods have gained widespread acceptance in software development due to several advantages over traditional structured techniques. Compare Figure 2-30 with Figure 2-31. The primary advantages are:

- Scalability
- Ease of Modification and Refinement
- Stability and Quality Control
- Reusability
- Modularity

Each advantage will be discussed in the sections which follow.

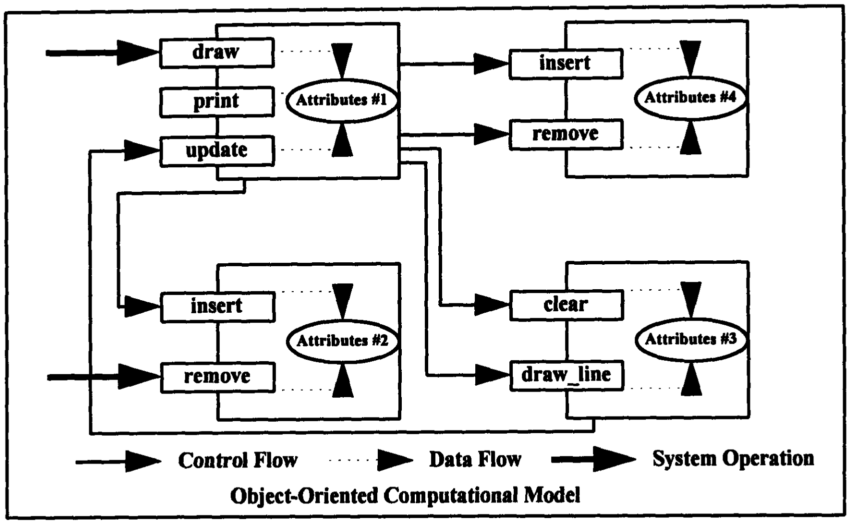


Figure 2-30: Object-Oriented Computational Model

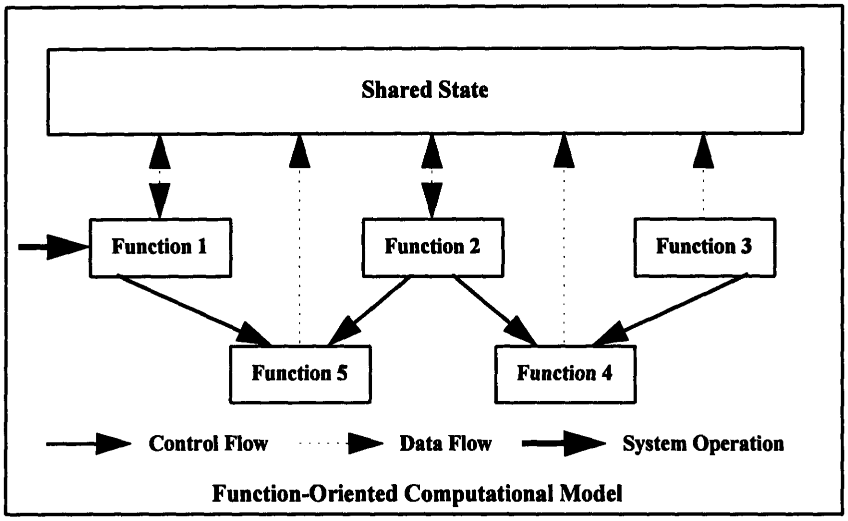


Figure 2-31: Function-Oriented Computational Model

Object-oriented systems are scalable. Scalability is achieved through the use of abstraction and object definition. Abstraction is the quality of labeling and describing the essential characteristics of something independent of the mechanism that will eventually be used to realize it. Object definition is merely the physical substantiation of the abstraction. Because objects are distinct from one another as well as Lego-like in nature, systems defined by interoperating objects can easily be expanded by adding more objects to the system. By contrast, it may be quite difficult to extend the capabilities of a function-oriented system unless the system was designed to accommodate specific future changes.

Modification and refinement is also easier to achieve in object-oriented systems. This is achieved because objects are encapsulated, i.e. the mechanisms employed by the object to perform its actions are self-contained. This means that the internal structure of objects can be changed without adversely affecting any other part of the system. This localization is not possible with unencapsulated systems that have heavy shared states. In these systems, any change to one part of the system to either extend or improve upon its offerings may adversely affect another part of the system. This arises because there is no control over what parts of the system can and cannot be employed by other parts of the system to achieve their end objectives. Furthermore, the dependencies may not be explicitly defined and when a change is made to one part of the system it will not be known whether or not another part of the system will be affected.

Object-oriented methods also yield better system stability and quality control. Since the relation between all interacting objects are explicit, bugs are less common and can be traced much more easily. Encapsulation not only helps scalability and refinement, but it also serves to isolate certain parts of the system from other parts of the system, localizing any problems that arise.

Reusability is enhanced in object-orientation through the concept of inheritance. All objects are derived from a class structure. Class structures are basically specialization trees. In this sense classes define "kind of" relationships among different attributes. The class structure enables an endless array of objects to be created from different combinations of class attributes. This facilitates the development of class libraries enabling reuse not only within one project but between projects as well.

As systems have become increasingly complex, it has become important to perform development in a distributed manner. Encapsulation enables systems to be subdivided. This allows multiple teams to work on the same system concurrently. The independence resulting from encapsulation ensures that development by one group will not unknowingly affect the performance of another group's work.

2.5.3.2 Key Concepts

Several fundamental concepts in object-oriented analysis have been incorporated in the Object-Oriented Axiomatic Design Approach. These are:

- Object Definition
- Encapsulation
- Aggregation

Each of these concepts add substantial strength to Object-Oriented Axiomatic Design. By virtue of the recognition of objects, boundaries are established. These boundaries establish independent sources of control. Encapsulation is much more difficult to achieve in real-world systems than in software development. However, the benefits are the same. Encapsulated systems can change the mechanisms employed to achieve a particular result without affecting any other parts of a system. Furthermore, encapsulation establishes a protocol between objects, ensuring that the object interfaces are well understood. Finally, aggregation establishes ownership among multiple objects. From a perspective of control, this ownership enables the structure of the objects themselves to be changed. This means that adverse coupling behavior can be eliminated.

Chapter 3

Case Study: Capacity Planning at United Electric Controls

3.1 Outline of Case Study

The case study of capacity planning at United Electric Controls is broken down into five major sections. Section 3.1 provides an overview of the company and the nature of its business. Section 3.2 provides an overview of the purpose and structure of the case study performed at United Electric. The system is characterized according to Object-Oriented Axiomatic Design in Section 3.3. This section develops the influence diagrams and interdependence matrices necessary for the elimination of adverse coupling behavior due to variation. Techniques applied to eliminate unwanted coupling behavior are presented in Section 3.4. This section includes the development of an interdependent capacity model and its predictive results. The case conclusion is subsequently found in Section 3.5.

3.2 Overview of Business

United Electric Controls is a small, privately held company with 200 employees and $40 Million in sales. It has fifteen product lines which are predominantly temperature and pressure controls used in both military and civilian applications. A defining characteristic of the business is the number of small orders it must handle each year. Most orders are for less than ten products. Furthermore, most orders are made for semi-custom products which are from the fifteen product lines but have specified one of several thousand possible option combinations. For this reason, the company is principally make-to-order.

The case study at United Electric (UE) focused on the production process referred to as the "120 Line". This line produced two significant revenue products for UE: the 105 Series and the 120 Series. Any product stemming from either of these two series was a temperature or pressure controller and came in as many as five hundred different option configurations. At the time of the study the company was in the process of refining a new type of production system that had been in place less than six months. In contrast to the old system which was a traditional MRP (Material Requirements Planning) batch process, the new process was a single-piece-flow kanban system which borrowed heavily from the Toyota production system. Despite the fact that the manufacturing process wasn't operating smoothly and the production workers were reluctant to change at the time of the case study, the motivation for change by management remained strong. Management believed that switching to the single-piece-flow pull system would enable the company to reduce inventory and decrease production lead times which, in turn, would allow UE to expand its business.

The 105 and 120 Series assembly process involved a total of seven "cells" or production areas. The fist area was called "Kitting". Kitting was where the components of the 105 and 120 products were collected from inventory and put in a tray. The next step was assembly. Because of differing assembly cycle times, the 105 and 120 products were manually assembled in separate assembly cells. From each of these cells, products would be transferred to either a temperature test cell or a pressure test cell depending upon the type of product. Since both series contained either temperature or pressure products, the separate assembly flows would become mixed at the different test cells. After successful testing, each assembly would move to a single packaging area. At this juncture, the separate temperature and pressure flows would be combined and all product would travel through the same packaging area. The last area the products passed through was shipping where orders were completed and shipping labels attached.

3.3 Overview of Study

3.3.1 Primary Objectives Encountered

United Electric has three chief objectives which were pervasive through the operations arm of the company. UE's primary objectives are to (1) ensure customer satisfaction through on-time delivery of quality products, (2) maximize earnings, and (3) maintain employee satisfaction. Each of the company's improvement efforts strive to enhance the company's position along at least one of these dimensions.

3.3.2 Problem Statement

At any given time, United Electric has a number of improvement efforts in process. At the time of this case study, UE was striving to simultaneously reduce inventory, capacity, and order lead times. The company had made substantial progress in both inventory and capacity reduction, but had had little success with its efforts in lead time reduction. United Electric believed that its arbitrary scheduling policies were the predominant source of costly delays. Toward this end, the company wished to study and develop heuristics for the production scheduling process.

The production scheduling process comprised both intra-day scheduling, i.e. the sequence of individual orders within the day, and inter-day scheduling, the quantity of orders produced each successive day. At the time of the study, the company had just institutionalized a policy that required the scheduling department to sequence orders on a first-in, first-out basis, with adherence to average daily totals for product mix and overall production quantity. Prior to this, orders were nominally scheduled on a first-in, first-out basis, but almost constantly overridden through expedited orders. The assumption was that if average daily totals could be strictly followed, variation to the manufacturing line in terms of mix and quantity would be minimized, thereby stabilizing production and decreasing the order backlog.

Despite the recent implementation of the aforementioned decision rules, the company had noticed little improvement in order lead time. Members of the operations management group at UE hoped that an in-depth study of the scheduling process would explain the lack of improvement and provide superior scheduling rules that would attain desired lead time performance.

Embarking upon the study of this process, the author quickly realized that the corporate objective to simultaneously reduce inventory, capacity, and lead time, had ignored two critical factors. First, the company did not address the interdependence between inventory capacity and order lead time. Second, it had ignored the role of demand variance. Given this circumstance, it was hypothesized that the company was trying to optimize a coupled system without understanding the nature of the couplings. This opened the door for the characterization of system behavior based on Object-Oriented Axiomatic Design. The expected outcome from this characterization would be an understanding of the couplings of the system, thereby illuminating the true sources of order delays. Furthermore, the OOAD characterization could then be used as a basis for the reduction of coupling behavior in the system.

3.4 System Characterization

To characterize the system in a meaningful manner, it was necessary to relate the variables used by the company in decision making back to the company's principal objectives, as described in Section 3.3.1. Since order lead time is the key operative in this study, it was first necessary to map the relationship between OLT and customer satisfaction, and the relationship between OLT and corporate profitability. Ultimately, the company needed to be in a position to understand how its control policies affect the bottom line. A mapping of the relationship between its control variables (decision variables) and its principal objectives could be attained once the effect of changes to the control variables upon OLT was established. The following sections detail this development.

3.4.1 Relationship Between Customer Satisfaction and Lead Time

Because UE serves a variety of customers with different requirements, not all customers expect the same product delivery lead time. Some customers at UE need sensors within one day to replace a broken unit, while others order at regular intervals. These different customer profiles are shown below:

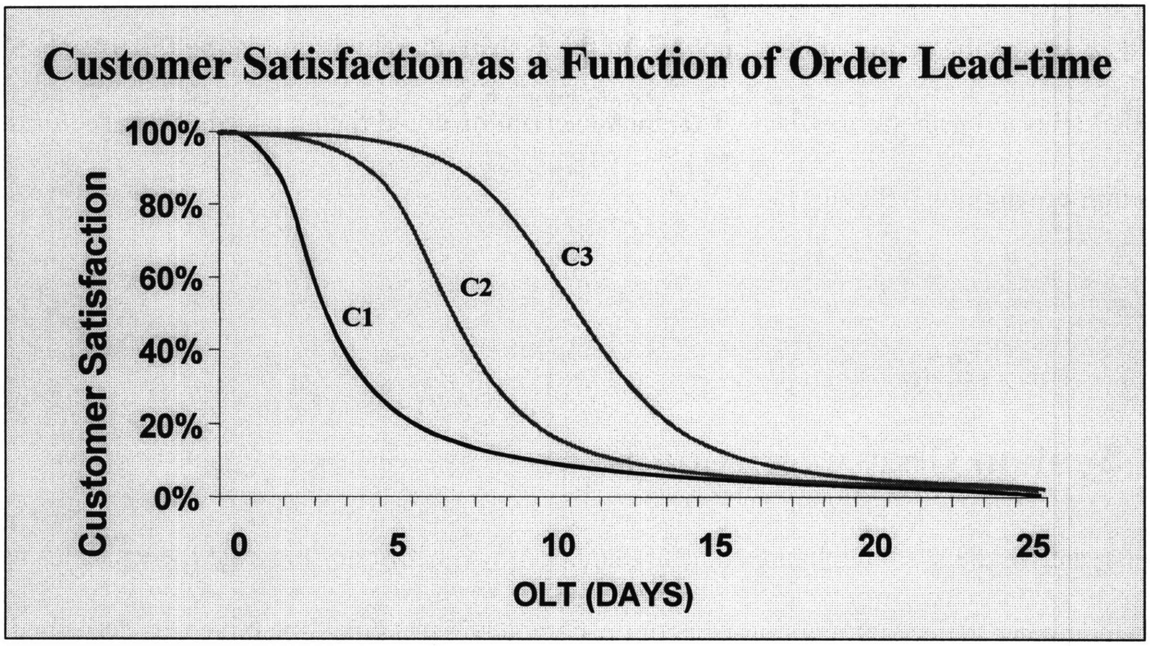


Figure 3-1: Customer Satisfaction as a Function of Order Lead Time.

This figure contains three different customer profiles (C_1, C_2, and C_3) each of which displays "s-shaped" behavior. There will be a short enough OLT for which the customer is indifferent (i.e. a customer may be 100% satisfied if the OLT five hours or less). Likewise, there is a long-enough OLT after which the customer's satisfaction no longer decreases (i.e. if a customer is displeased at an OLT of 50 days, he/she will be equally displeased if the OLT is 51 days). Thus this curve is expected to be flat at very short OLTs and at very long OLTs, leading to the s-shape.

3.4.2 Relationship Between Demand and Lead Time

A second critical relationship is the demand/OLT elasticity curve, shown in Figure 3-2. Clearly, as the OLT is shortened, demand for the product will grow; shorter lead times and therefore higher service levels increase the product's attractiveness. There is a small enough OLT, however, for which the product gains no attractiveness and demand does not increase. Similarly, demand for a product will drop as OLT increases, but after a long OLT, demand will approach zero. For this reason, the demand/OLT elasticity curve also has s-shaped behavior.

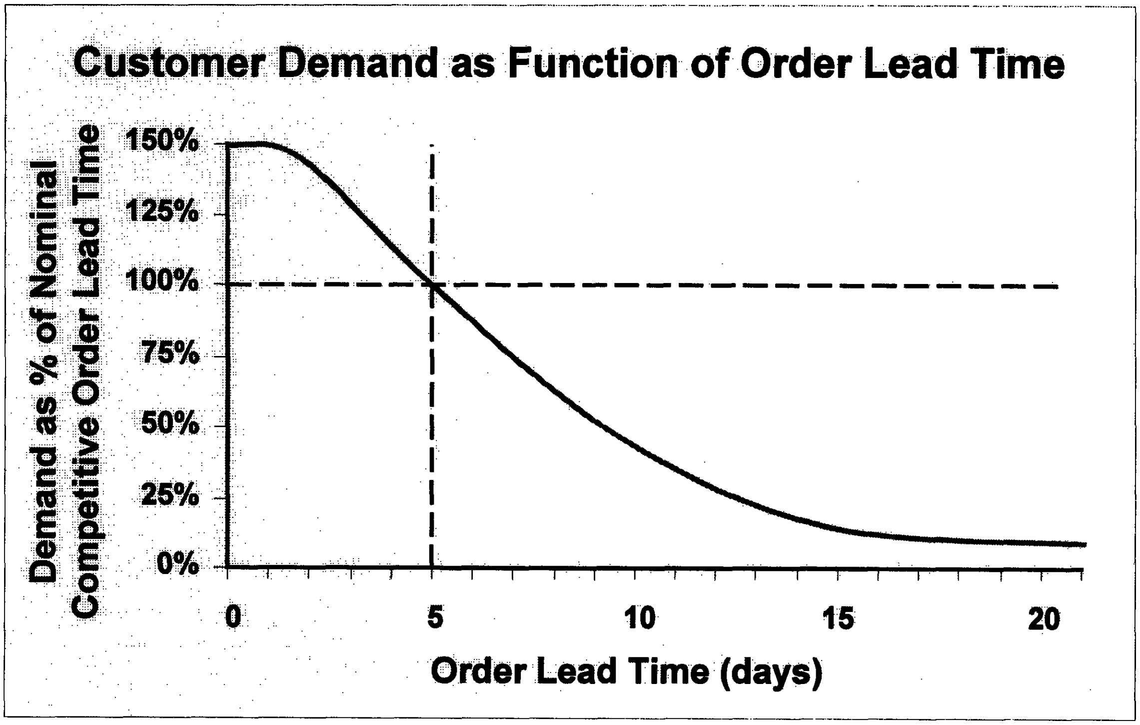


Figure 3-2: Customer Demand as Function of Order Lead Time

3.4.3 Relationship Between Capacity and Lead Time

Lastly, there is an important relationship between capacity and lead time. This relationship is particularly important in make-to-order environments, since, by definition, there can be no finished goods inventory (FGI) under such a production policy. Eliminating FGI (not to be confused with work-in-process inventory which is a necessary and strategic element in make-to-order systems) dramatically simplifies the relationship between lead time and capacity.

3.4.3.1 Common Understanding of Relationship Between Capacity and Lead Time

At a first order level this relationship is understood by virtually all people. There are two parts to this understanding:

- On a local basis, lead time is inversely proportional to capacity
- On an average basis, demand in excess of capacity will result in perpetually increasing lead times and demand less than capacity will be stable with lead times equal to the throughput time of the manufacturing system

<u>Local Relationship Between Capacity and Demand</u>

For example, if the production capacity is one hundred units per day and a day's demand comes in at one hundred units, the expected lead time is one day. On the other day, if demand comes in at two hundred units, the manufacturing will require twice the time to produce the order and therefore the expected lead time is two days. Likewise, if the daily orders come in at fifty units, it should take only half a day to produce the orders. This relationship is shown below.

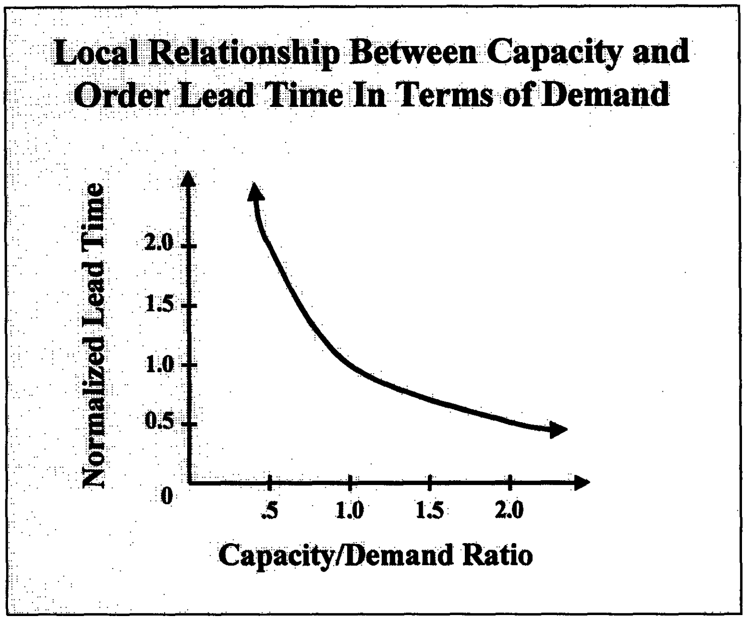


Figure 3-3: Local Relationship Between Capacity and Demand

Average Relationship Between Capacity and Demand

By comparison, on an average basis, the common understanding is that the order lead time will be the manufacturing throughput time when the average demand is at or less than the capacity of the system. The belief is that if the average demand equals the capacity of the system, sufficient capacity will exist to meet all the demand. Some days demand will be higher than others, but if average demand capacity exists, then the higher and lower demands will "balance out". In more concrete terms, people use the following relationship as their mental model:

$$OLT = (D + B) * Takt + Flow\,Time + Other \tag{3-1}$$

where

OLT = Order Lead Time

D = Daily Demand (units)

B = Backlog

$Takt$ = Takt Time (Time Per Unit Produced)

$Flow\ Time$ = Manufacturing Throughput Time or Elapsed Time From Beginning to End of Manufacturing Process

$Other$ = Order Processing Time + Shipping Time + Scheduling Inefficiencies + Manufacturing Losses

Since the takt time, the flow time, and other organizational losses can be considered constant relative to demand fluctuation, the only concern due to demand fluctuation is the backlog. At UE, the fluctuation in backlog was assumed to be proportional to demand, i.e. if demand increased 10% the backlog would increase 10% and if demand fell 10%, the backlog would fall 10%. The average backlog was assumed to be driven by scheduling inefficiencies and manufacturing losses. Given this, the average lead time will be the lead time at the average demand rate. In this manner, variation in demand will "balance out". This relationship is shown in (3–2).

$$\overline{OLT} = \left(\overline{D} + \overline{B}\right) * Takt + Flow\,Time + Other \tag{3–2}$$

Furthermore, according to Little's Law,

$$Flow\,Time = WIP * Takt$$

where

WIP = Work-In-Process

substitution into (3–2) yields

$$\overline{OLT} = \left(\overline{D} + \overline{B} + WIP\right) * Takt + Other \tag{3–3}$$

Finally, it is noted that

$$C = \frac{A}{Takt} \qquad (3\text{–}4)$$

where,

C = Manufacturing Capacity (units per day)

A = Availability (number of minutes available per day for production)

Therefore

$$C \propto \frac{1}{Takt}$$

Thus from (3–3) and (3–4), it can be seen that the expected order lead time is inversely proportional to the ratio of demand to capacity. As long as average demand is less than or equal to capacity, this relationship is expected to hold. If average demand equals capacity, then according to (3–3), the order lead time will be the sum of the time to work off the backlog, the time to work through the day's orders, and the flow time through the manufacturing process, among other time losses. If, on the other hand, demand is greater than capacity, the backlog will increase until customers begin to balk at the length of the lead time. Under this condition, the lead time will reach a constant value despite relative differences in the ratio demand to capacity. This assumed relationship between capacity and demand is shown in Figure 3-4, below:

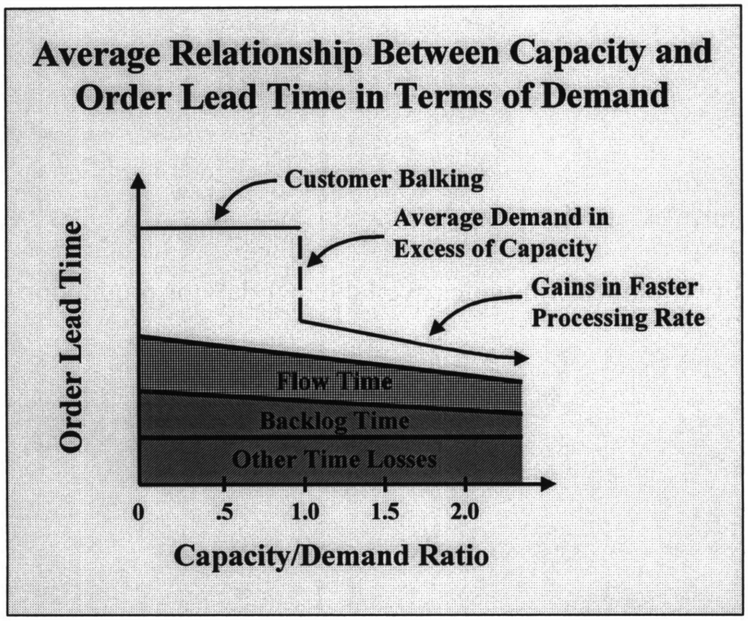


Figure 3-4: Average Relationship Between Capacity and Demand

3.4.3.2 UE Decision Process Based on Common Understanding of Capacity – Lead Time Relationship

At United Electric during the time of the study, the average lead time was two weeks. To increase customer satisfaction, the company wanted to bring this down substantially, to less than two days, if possible. From Equation (3–3), it can be seen that there are four ways to decrease the order lead time. The company can either decrease the takt time (the inverse of available capacity), the drivers of backlog, the WIP, or other losses such as order processing time, scheduling inefficiencies, manufacturing losses, or shipping time.

After discussing the alternatives and seeking the advice of a Toyota Motor Company consultant, United Electric decided it would set the takt time to the available hours divided by the average daily demand, work overtime to completely work off the backlog, and then use kaizen and strict scheduling approaches to reduce the other sources of lost time. The rationale for this approach was twofold.

1. Reducing the takt time increases the processing rate at the factory. This not only costs money since it is the equivalent to increasing capacity, but if it is lowered below the average daily demand rate, then the factory will be effectively overproducing, i.e. it will produce at a faster rate than demand. Additionally, increasing the production rate beyond the average demand rate only disguises other sources of time loss that affect the system. Therefore, any reduction in takt time below the average demand rate was considered undesirable.
2. If the takt time is set to the available time divided by the average daily demand and the backlog is worked off through overtime, the backlog should stay low since the average production rate will be the average order rate. Because the company would be producing at the average demand rate under this policy, the manufacturing process would be stable. The same quantity of product would be produced each day at the same production rate. This would enable the manufacturing process to be more efficiently and effectively run, reducing the likelihood of errors. Furthermore, with the takt time set at this value, it will take only one day to produce a day's worth of orders. Since the flow time at the factory was only half a day, this approach should enable the company to complete orders within a day and a half, meeting the two day objective.

This strategy of producing to average daily demand, reducing inventory, and following strict scheduling rules was widely believed to be able to reduce the order lead time dramatically.

3.4.3.3 Demand Variance: The Missing Link Between Capacity and Lead Time

Unfortunately though, this very policy caused the order lead time to increase at UE. The reason is that, unlike the widespread belief that average backlog was driven by periods of production below the average demand rate or scheduling inefficiencies, the average backlog size (and therefore OLT) is fundamentally driven by demand variance. It is the relationship between demand variation and capacity that principally determines backlog size. The syllogistic error was that backlog size was not considered to be a function of takt time; only the rate which manufacturing worked through the backlog was expected to be affected by changes in the takt time.

The reason why the takt time (or capacity) of the manufacturing system was not expected to affect the size of the backlog is that the role that variation in demand played was not understood.

The Three Components of Demand Variation

There are three components of demand variation that affect lead time performance. These are:

- Expected Demand
- Variance of Demand Relative to Expectation
- Autocorrelation of Demand Relative to Expectation

These three dimensions fully and independently characterize the demand profile that the company must support. Each of these will be described in detail in the following sections.

Expected Demand

The product forecast is the future expected demand. This expected future demand is the company's best guess of future demand. Typically companies can do no better than forecast a general trend for growth or decline of the product, seasonality of the product if there is any, or a combination of the two. The deviation of actual demand from the forecast, which is known to occur, is assumed random. Growth or seasonal variation strongly affect the system capacity requirements which in turn affect order lead times. In a growth scenario, setting the takt time to the average daily demand for any historical period will result in a perpetually increasing backlog. Demand will always be in excess of capacity. Similarly, when demand is seasonal, setting the takt time to the mean will also result in excessive backlogs whenever demand cycles above average. This effect is shown in Figure 3-5, below.

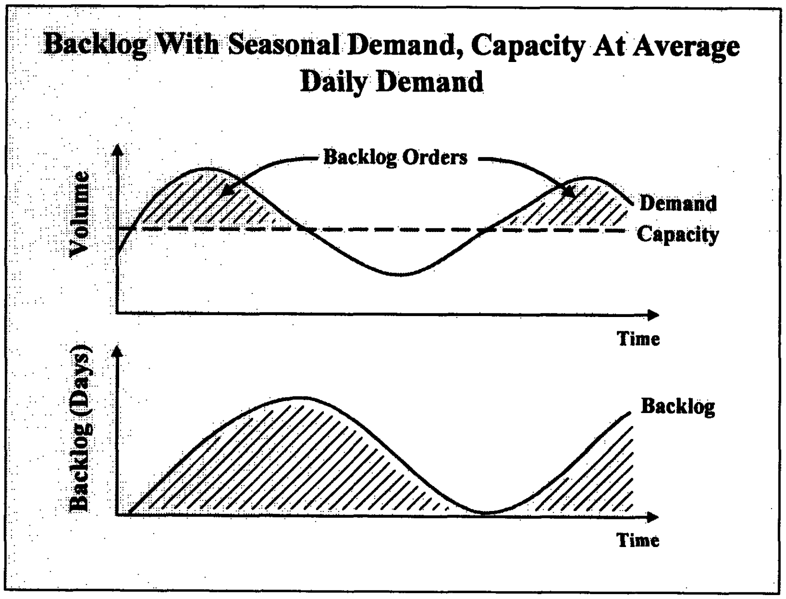


Figure 3-5: Backlog With Seasonal Demand, Capacity At Average Daily Demand

Because of the effect of expected demand variation on backlog, the first change the company should make is to set the base takt time to the forecast value of daily demand. Ignoring the effects of variation of demand from the forecast, this improvement will completely eliminate backlog build-up due to expected demand variation. Of course, how the company provides the physical capacity to attain the takt time is another matter. For seasonal demand about a constant value, the company can either install sufficient capacity to capture the cycle peak demand or it can find ways to bring in capacity as it is necessary, e.g. temporary labor or overtime.

Fortunately for United Electric, the product demand is neither seasonal nor on a growth trend. This can be observed from the fourth-order polynomial regression of the daily demand in Figure 3-6.

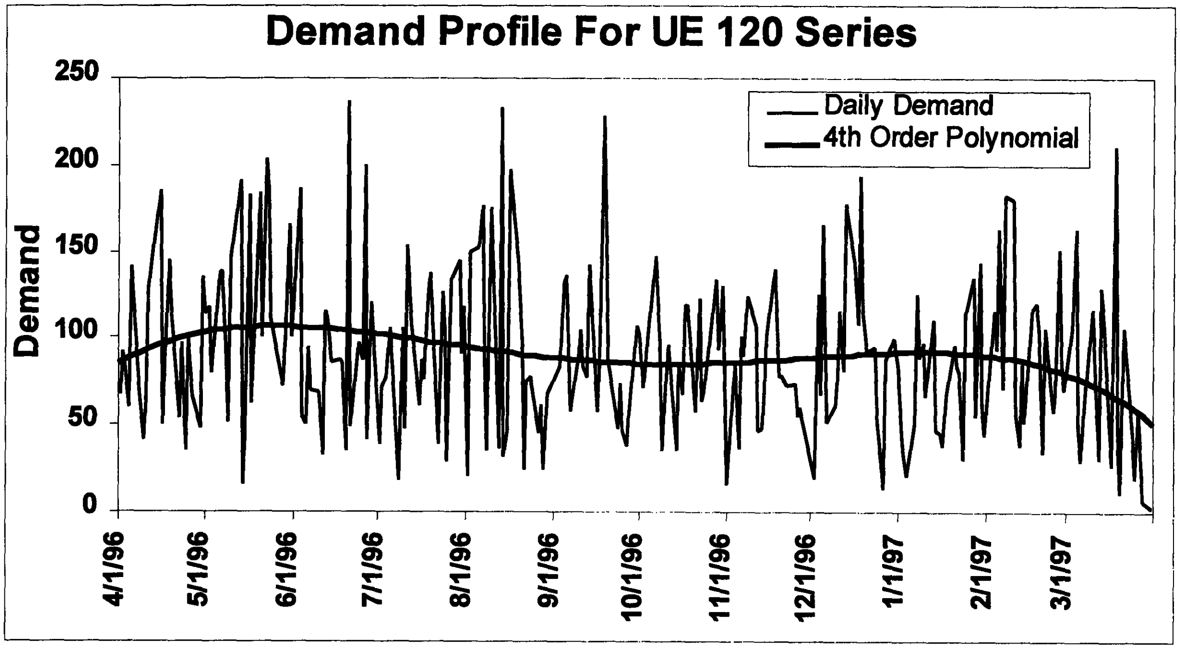


Figure 3-6: Demand Profile for UE 120 Series

The fourth-order fit has sufficient terms to capture any form of seasonal behavior that might exist. If there is any seasonality, it is only a few percentage points as the plot illustrates. Because of the steady, if not sloping downward slightly, trend in demand, United Electric is not in a position of having to respond to changes in the expected demand. This enables the base takt time to be set at the average daily demand, without incurring any increases in backlog due to variation of the expected demand.

Variance of Demand Relative to Expectation

The second dimension of demand variation affecting order lead time is the variance of demand relative to the expected or forecast demand. As the variance increases, so too will the backlog if the capacity in the system is left the same. Therefore, if two different demand profiles have the same expected value but different variances from the expected value (see Figure 3-7), the takt time based on the expected daily demand will be the same for each, but the backlog will be worse for the demand profile with the higher variance (Figure 3-8).

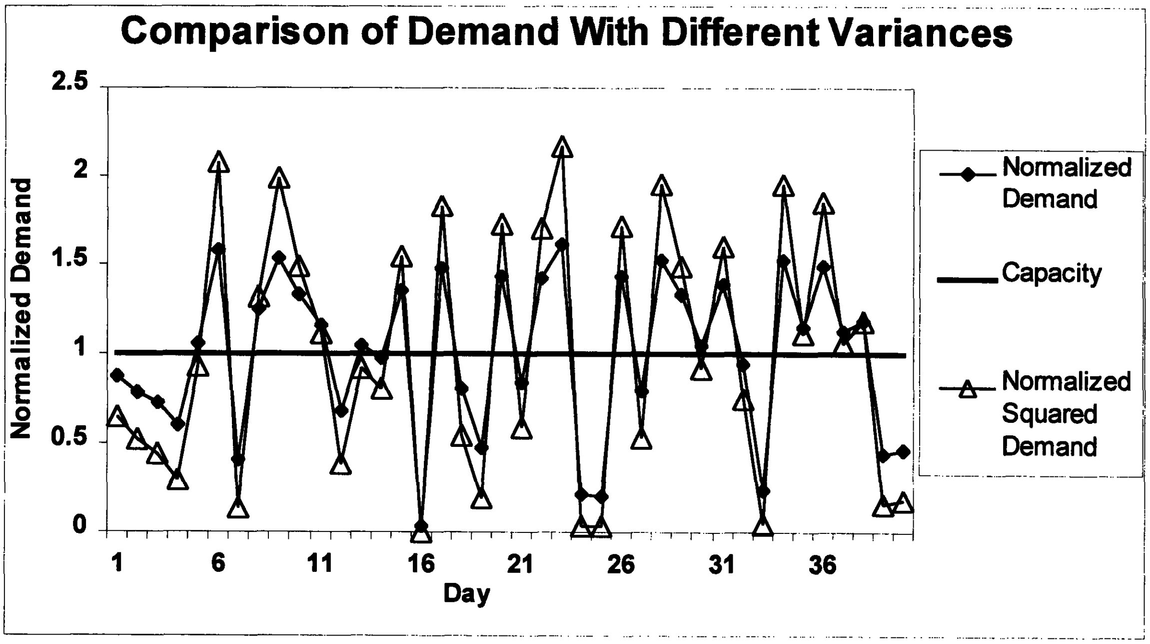


Figure 3-7: Comparison of Demand Profiles With Different Variances

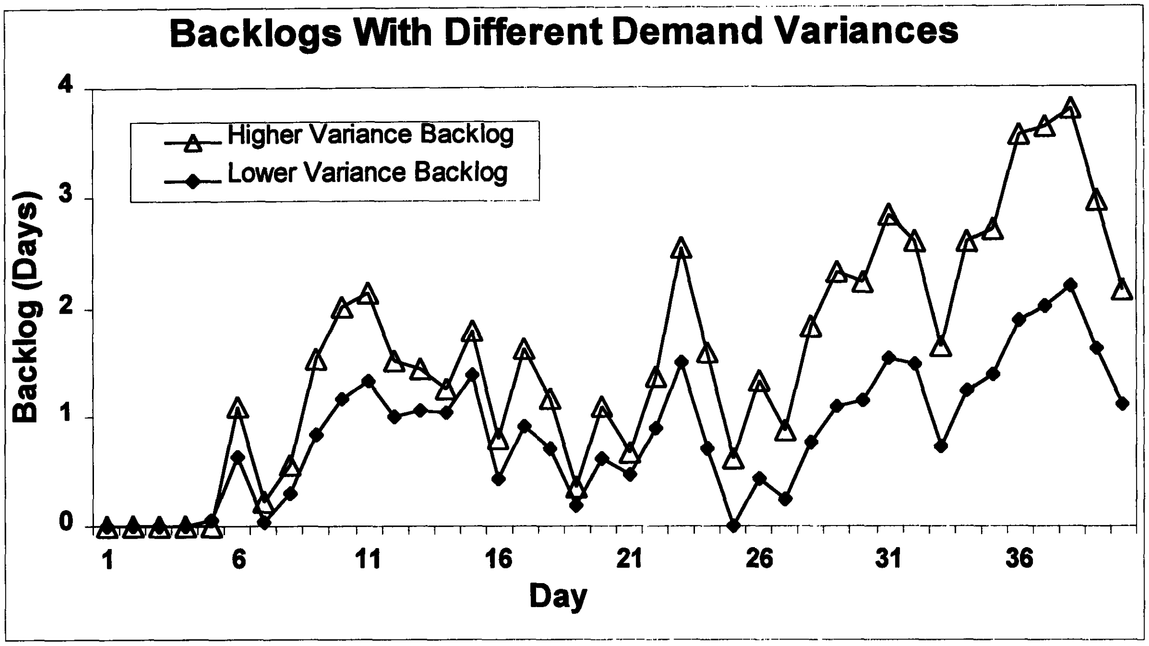


Figure 3-8: Comparison of Backlogs for Different Demand Variances

This effect of demand variance was not considered by the production planners at UE. If they had understood that increases in variance given a takt time would increase the order lead time, they would have been able to deduce the inverse – that increases in takt time given a particular variance would decrease the order lead time. Since United Electric interfaces directly with the final customer, there is no way to reduce the demand variance (as opposed to other supply-distribution systems where supply chain management and channel partnering can eliminate sources of variation from suppliers and partners). This means that in a make-to-order environment, the only way to decrease the backlog is to either decrease the takt time or work longer hours. Since this relationship was not understood, it was believed that if the backlog was worked off, it would not return as long as strict internal processes were followed.

Autocorrelation of Demand Relative to Expectation

Autocorrelation is the third dimension of demand that affects order lead time. Unlike variance, which measures the deviation of demand from the expected value, autocorrelation measures the serial correlation of demand. In other words, if demand deviates high from the expected value, the autocorrelation is the likelihood that the demand at the next time interval will also be high. In this manner, autocorrelation is fully independent of variance. The equation for autocorrelation is given in (3–5) [37].

$$\rho_{x_i,x_{i-1}} = \frac{Cov(x_i, x_{i-1})}{\sigma_x^2} = \frac{\dfrac{\sum_{i=1}^{n}(x_i - \bar{x})(x_{i-1} - \bar{x})}{n-1}}{\sigma_x^2} \tag{3–5}$$

High levels of autocorrelation will increase order lead times substantially. With a high autocorrelation, a high deviation of demand from the expected value will be followed by another high value of demand. When high demands stack up, the lead times will increase substantially due to the inability of the manufacturing system to accommodate the higher demand. For comparison, lead times will never stack up if the demand has a perfectly negative autocorrelation. In this situation, demand high on one day would be immediately followed by low demand on the subsequent day. This effect insures that lead times would never increase beyond one day.

At United Electric, the autocorrelation was found to be almost exactly zero. This means that if demand on one day is high, the following day's demand is as likely as to remain high as it is to go low. This situation can be compared to circumstances where UE's demand might be negatively correlated, as shown in Figure 3-9, or positively correlated, as shown in Figure 3-10.

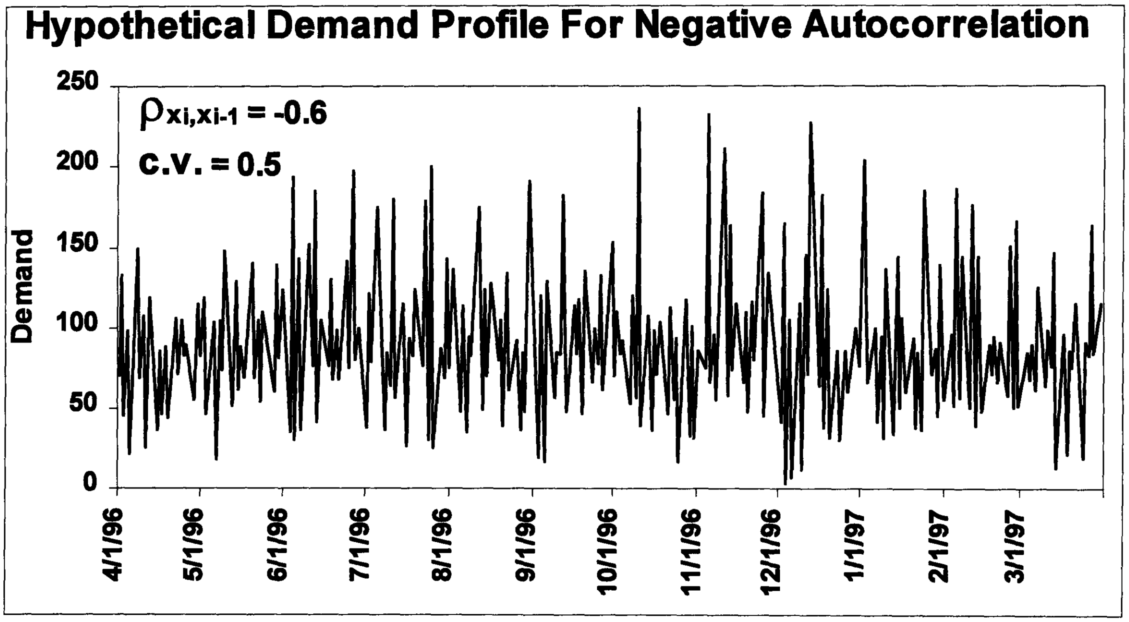


Figure 3-9: Hypothetical Demand Profile for Negative Autocorrelation

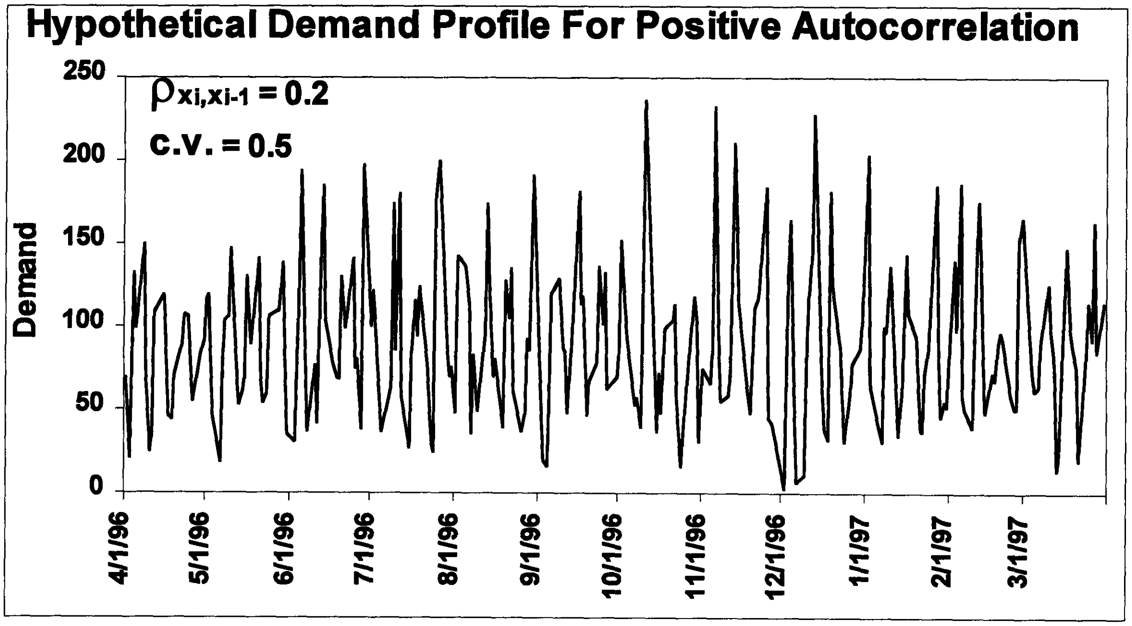


Figure 3-10: Hypothetical Demand Profile for Positive Autocorrelation

The differences between the two profiles are subtle, yet the effect on the manufacturing system in terms of backlog will be substantially different. While conditions of negative autocorrelation are desirable, they occur infrequently. On the contrary, there can be numerous situations where a high degree of positive autocorrelation occurs. For example, companies that are in industries where there are heavy promotional programs, either by the company itself or its competitors, will create high levels of autocorrelation. In these circumstances, autocorrelation will be caused by a promotional program that creates a run of high product demand for a period of time. United Electric's autocorrelation of zero indicates that demand follows a truly random pattern.

Because United Electric's autocorrelation is zero, there will be times when the company has a run of demand that is higher than the expected value. This effect means that autocorrelation plays a role in order lead time and therefore cannot be ignored. Since the autocorrelation cannot be easily changed for the better, the only way the company can treat this problem is to change its takt time to account for this effect. Just as increases in takt time can decrease order lead time given a demand variance, so will increases in takt time decrease order lead time given a particular autocorrelation. The arguments for both variance and autocorrelation are identical.

3.4.3.4 Actual Relationship Between Capacity and OLT Given Demand

The three measures of demand – expected value, variance, and autocorrelation – fully characterize the demand profile. While these three measures are independent, they do not independently affect lead times in the manufacturing system. Given a particular demand profile, changes in capacity will affect order lead times in a nonlinear manner.

3.4.4 System Characterization Through Use of OOAD Techniques

The previous sections have established the relationship between OLT and customer satisfaction, OLT and demand, and capacity and OLT given demand. Together, these relationships provide the basis for a mapping between the company's control variables and its key objectives.

However, the one-dimensional functions previously illustrated in Figure 3-1, Figure 3-2, and Figure 3-4 are inadequate. The system is coupled. Changes in both demand and capacity have a coupling effect on system behavior. Since changes in demand affect OLTs and demand itself cannot be explicitly controlled, it can be considered a coupling variable. Capacity, on the other hand, is clearly controlled by the company, but couples with corporate objectives because increases in capacity increase investment and carrying costs while increasing sales revenues. These couplings necessitate the development of an influence diagram and interdependence matrix if system performance is to be improved.

Influence Diagram

The influence diagram developed for UE (Figure 3-11) demonstrates how three controls, quoted lead time, inventory, and capacity, affect the company's two principle objectives, net revenue and customer satisfaction.

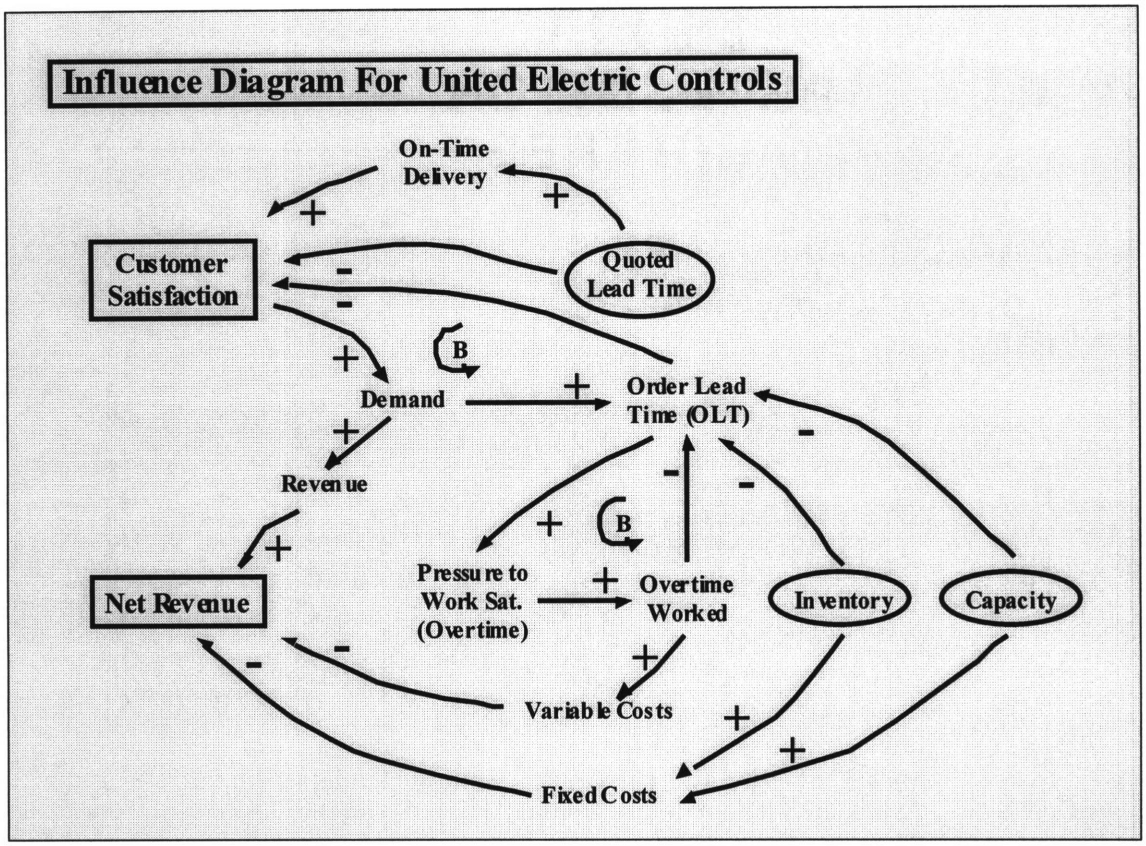
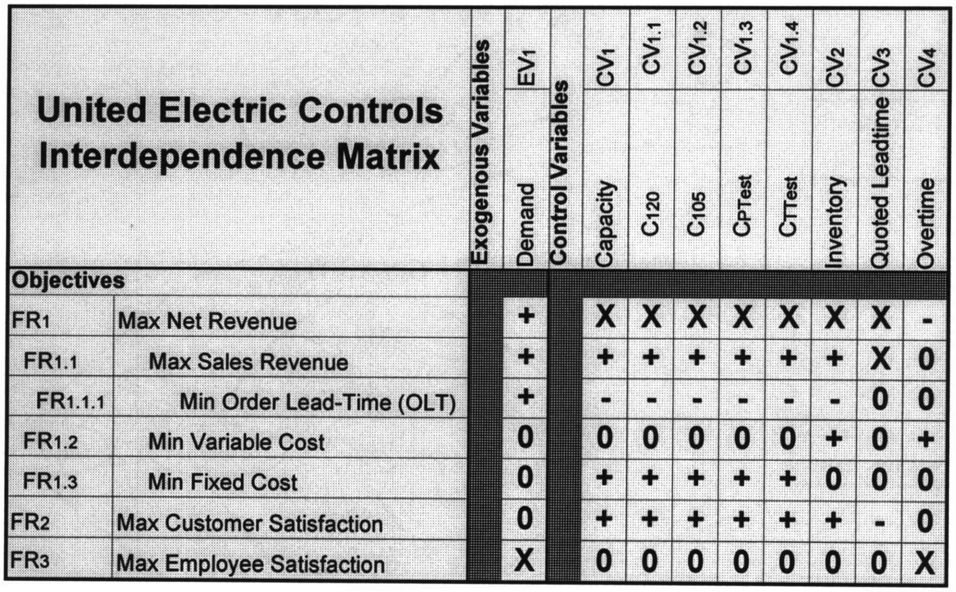


Figure 3-11: Influence Diagram for United Electric Controls

The first control variable, *Quoted Lead Time*, increases customer satisfaction as the lead time is decreased. However, this also decreases the likelihood of on-time deliveries, since less time is available. The second control variable, *Inventory*, was not a function of the scheduling process; it was taken as "given." *Capacity*, on the other hand, was a direct function of scheduling. The daily production quantity established the takt time (or capacity) of the system. As the influence diagram shows, increased capacity increases OLT, which in turn increases customer satisfaction and therefore demand. Increased demand, of course, results in higher revenues for the company. Additionally however, increases in capacity increase fixed costs, which in turn detract from net revenues. Thus, the influence diagram shows that the system is highly interdependent. These complex relationships can be easily abstracted through an interdependence matrix.

Interdependence Matrix

Figure 3-12 shows the key interdependencies and their couplings to UE's objectives.



United Electric Controls Interdependence Matrix		Exogenous Variables	EV_1 Demand	Control Variables	CV_1 Capacity	$CV_{1.1}$ C_{120}	$CV_{1.2}$ C_{105}	$CV_{1.3}$ C_{PTest}	$CV_{1.4}$ C_{TTest}	CV_2 Inventory	CV_3 Quoted Leadtime	CV_4 Overtime
Objectives												
FR_1	Max Net Revenue		+		X	X	X	X	X	X	X	-
$FR_{1.1}$	Max Sales Revenue		+		+	+	+	+	+	+	X	0
$FR_{1.1.1}$	Min Order Lead-Time (OLT)		+		-	-	-	-	-	-	0	0
$FR_{1.2}$	Min Variable Cost		0		0	0	0	0	0	+	0	+
$FR_{1.3}$	Min Fixed Cost		0		+	+	+	+	+	0	0	0
FR_2	Max Customer Satisfaction		0		+	+	+	+	+	+	-	0
FR_3	Max Employee Satisfaction		X		0	0	0	0	0	0	0	X

Figure 3-12: United Electric Controls Interdependence Matrix

The interdependence matrix shows that increases in capacity increase revenues, decrease lead times, increase fixed costs, and increase customer satisfaction. Per the discussion surrounding Figure 3-1 and Figure 3-2, it is understandable how increases in capacity, by changing order lead time, increase the volume of product demand and customer satisfaction. Unfortunately, any benefit in revenue by increased capacity is offset by increased investment costs. This effect reflects adverse couplings present in the system, whereby an improvement to one area detracts from the performance of another. Since these relationships are nonlinear, it is worthwhile to construct a schematic relationship of their effect on net revenue, the principle objective. This is shown below:

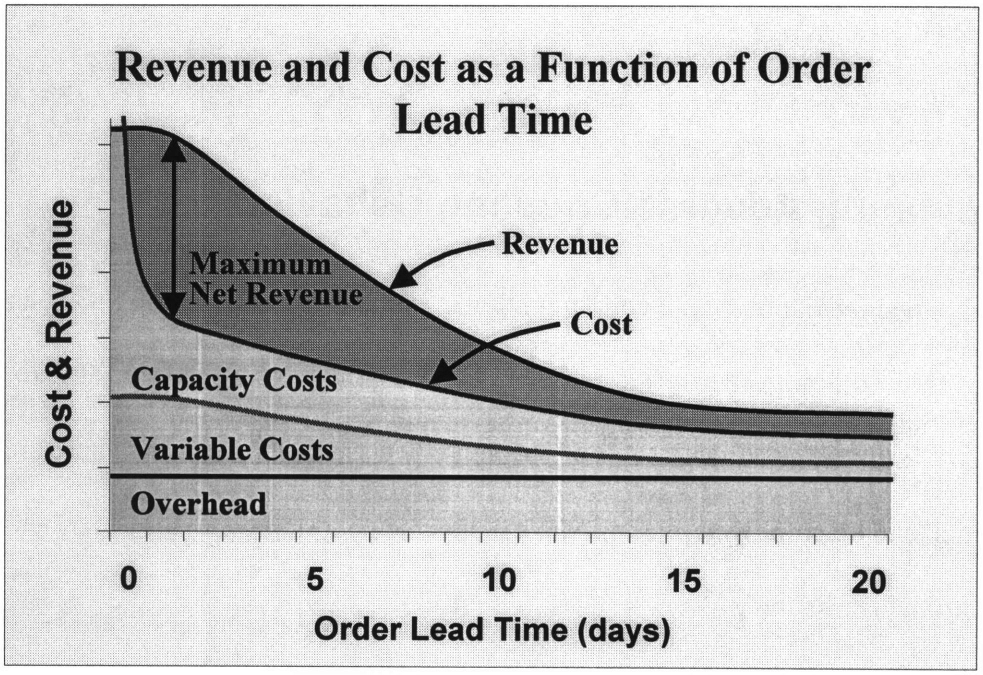


Figure 3-13: Revenue and Cost as Function of Order Lead Time

Figure 3-13 shows the net revenue as a function of OLT, with curves for overhead, variable costs, and capacity costs. As the OLT decreases, net income increases, but flattens out as the OLT approaches the "indifference" point. Capacity costs, nevertheless, approach infinity as the OLT goes to zero. These coupling relationships indicate that there is an optimal point in terms of maximized revenue. This optimum is obtained at the maximum distance between the total costs and total income.

Despite the fact that there exists a global optimum, in the current state of the system, the system does not naturally goal seek to that point. This occurs for two reasons. First, the relationships shown in Figure 3-13 are not well understood by the company. Second, the responsibilities for improved objective performance do not lie within the domain of the group that wields control. In this manner, it is necessary for the company to employ tactics that reduce adverse coupling behavior and therefore encourage the company to achieve optimal performance.

3.5 Eliminating Adverse Coupling Behavior

Having established the nature of the coupling interdependencies at UE and the reasons why it had been unsuccessful in its attempts to reduce lead times, the next logical step was to develop techniques to eliminate the adverse coupling behavior present in the system. There are three methods which can be employed to eliminate the adverse coupling behavior described in the previous section. These are:

- Aggregate control, understand couplings and optimize
- Reduce or eliminate couplings between objectives
- Reduce sensitivity to changes in demand

Due to the fundamental nature of these couplings, there was no method readily available to either reduce sensitivity to changes in demand or to reduce or eliminate these couplings. If the company were to fully understand the nature of the couplings, it would be possible to aggregate control in such a way that changes to capacity would only be made if it served to increase the company's overall performance. This aggregation of control, alongside a clear understanding of the coupling relationships present, would eliminate local decisions that adversely affected the overall corporate performance (e.g. produce to the average takt time). Toward this end, it was necessary to more fully develop the relationship between the principal control variable, capacity, and the primary objective, net revenue. This was essential for a clear understanding of the system structure.

UE employees felt that the relationship between OLT and sales revenue was understood. Similarly, the relationship between capacity and fixed cost was also understood. However, the company did not have a grasp of the relationship between capacity and OLT. The next section addresses this relationship.

Coupling Strength Between Capacity and Order Lead Time

Changes to the level of capacity at United Electric were not necessarily done in an aggregate manner. In actual practice, there were four independent capacities in the manufacturing line. These were: the capacity of the 120 cell, the 105 cell, the temperature test cell, and the pressure test cell. The different capacities reflect two different product families and two different product types, all of which shared the same production system.

Each of the four capacities affect the performance of the system in an unique manner. The capacity of the 105 cell determines the maximum amount of 105 products that can be passed on to the temperature and pressure cells. This is analogous for the 120 cell. Both the 105 and 120 product lines produce either temperature or pressure sensors that must pass through the respective test cells. This way, changes to any of these capacities are expected to affect order lead times in some manner. Because each of these decisions can be made independently, the interdependence matrix, Figure 3-12, has been decomposed one level to show each of these control variables individually.

As shown in the interdependence matrix, coupling to OLT by changes in capacity is given by:

$$\frac{\partial C_{120}}{\partial OLT}, \frac{\partial C_{105}}{\partial OLT}, \frac{\partial C_{PTest}}{\partial OLT}, \frac{\partial C_{TTest}}{\partial OLT}$$

These couplings were known to be not only nonlinear, but also functions of each other. The effect of a change in capacity of any one cell was dependent on the states of the other cells. For example, if the assembly cells (120 or 105) had extremely low levels of capacity, any amount of capacity increase in either of the test cells would have no effect on order lead times. Of course, seldom is the manufacturing system in this condition. Often, capacities are relatively balanced, making it particularly difficult to develop an intuition about the effect of any one change on order lead time. Therefore, to determine the coupling strength between capacity and OLT, it was necessary to build a computer model.

3.5.1 Capacity Model Development

The capacity model development has been broken down into three sections. The first section discusses three components of capacity. These are: the expected minimum daily demand, the known volatility, and the expected investment lead time. The second section describes the calculation of standing capacity, i.e. the amount of capacity available on a daily basis. The third section describes the features of the capacity model and the final section describes its formulation. Results from the capacity model are presented in Section 3.5.2.

3.5.1.1 The Three Components of Capacity

Capacitization is critical to the system performance of any company. Capacity can be subdivided into three categories. First, is the *expected minimum daily demand* that will be experienced over the life of the capital assets. This is the minimum daily demand that will be experienced day-in and day-out by the company. This component of the total daily volume can be considered extremely stable and dedicated; inflexible facilities can be installed toward this end. For UE this is approximately 20% of the total volume for its products.

The second category of capacity is *known volatility*. Known volatility is the total known variation. This is the expected variation (forecast variation) plus the pattern of variance from the forecast (variance and autocorrelation). At UE there is no appreciable seasonal component of demand; therefore the forecast is stable. However, the known volatility on a daily basis is high. Daily bookings show virtually no temporal correlation, i.e. an autocorrelation of zero, and typically have a c.v. (coefficient of variation) of between 0.5 and 1.0 over a one year horizon. This means that on a short (day-to-day) basis, demand will be highly volatile while on a monthly or yearly basis the demand will be relatively stable. These patterns of demand are seen in Figure 3-6.

The third component of capacity is the requirement to support customer demand in excess of the forecast over the lead time to install new capacity. This lead time is a function of the manufacturing system design such as the nature of local bottlenecks and the increment size of capacity. At UE, mean demand and volatility of demand is both stable and accurate when an exponential smoothing forecast technique is used. This means that monthly changes show no trend behavior and experience period-to-period changes of less than 10%. As a consequence, it is not necessary to carry any of this third capacity component given that volume changes of 10% can be accommodated within the lead time to increase capacity.

In summary, UE must capacitize its manufacturing process to optimally support a known highly volatile but stable and unseasonal customer demand process. This requires UE to set a desired standing capacity in such a manner as to achieve corporate strategic and operational objectives, where the standing capacity is defined as the amount of capacity available on a day-to-day basis. Distinguishing between flexible and inflexible capacity requirements as well as providing for short-lead capacity additions is not relevant to UE based on this demand profile. Therefore the selection of the standing capacity requirement is the most critical capacity parameter affecting the success of the manufacturing line. The following section is dedicated toward this end.

3.5.1.2 Calculation of Standing Capacity

The standing capacity requirement (the required manufacturing capacity carried on a day-to-day basis) is typically characterized by the mean demand (over a sufficient interval) plus *uplift*. Uplift is given as a percentage of the mean demand and represents the amount of capacity dedicated toward supporting demand variance (both certain and uncertain). To determine desired uplift, the company must determine what portion of demand variance will be accommodated by capacity, FGI, or order lead time. To the extent that the company adds capacity or inventory, the order lead time will be reduced. The specific mix of inventory, capacity, and FGI, will be dependent on the nature of the business. If customer tolerance for lead time is within the production lead time, the company has no alternative but to keep FGI. However, for UE the customer has an order lead time preference which enables "make-to-order" to be possible. Furthermore, the broad range of SKU per manufacturing line makes inventory carrying undesirable relative to manufacturing on demand. Only where the cost of carried capacity can be reduced by an amount in excess of the cost of carrying inventory should inventory be used to buffer variance instead of capacity. For UE this may be only for its highest volume runners. Once the desired FGI stocking levels have been established, the remaining volume must be captured in a capacity model which trades off capacity and order lead time.

3.5.1.3 Description of Capacity Model

The capacity model has been built to assess the tradeoff between capacity and order lead time for both the 120 and 105 product lines at UE. To calculate the specific relationship between capacity and order lead time, several factors must be considered.

1. Scheduling and production lead time must be adequately determined. At UE the scheduling lead time is approximately four hours (heijunka loading takes place in four hour intervals). Production lead time is approximately four hours once the heijunka card has been pulled. In total the scheduling and production lead time is about one day. The remainder of order lead time is either delay due to stockout, transport (shipping) time, or order queue time. In this model, effects of stockouts, in-process order sequencing changes, and transportation times were neglected.
2. Order sequence is important. The capacity model assumes that all orders are processed on a (by day) FIFO basis. If, for example, some of Wednesdays orders are processed a day or more before Tuesdays orders, the model will not be representative of actual practice.
3. Perfect yield and no production initial backlog is assumed. In practice, production variance both in volume and yield will affect order lead times. However, United Electric has a policy of producing all the orders slated for a particular day, even if

overtime is required. The developed capacity model assumes that all orders slated for production on a given day will be produced on that day.

4. Capacity is constraining. No more product can be produced on any given day that exceeds the capacity of the system. At UE, system capacity will be a function of mix since some products share parts of the production process while being separate on others. If the demand for product through a stage in the manufacturing process exceeds its daily capacity, the additional product will not be built, irrespective of capacity elsewhere in the system. The developed capacity model captures this effect.
5. Daily production mix is important. When a system has an interdependent capacity structure (such as UE with separate 105, 120, P, and T cells), the choice of which products to produce on a given day and which to queue to the next day when demand exceeds capacity can markedly affect order lead time. The developed capacity model captures this effect by prioritizing 105 demand before 120 demand (due to its higher c.v.) and demand for pressure product before temperature on the assumption that capacity in temperature test will be higher than pressure test due to physical differences in the current process.

3.5.1.4 Capacity Model Formulation

The United Electric Capacity Model consists of four modules. The first module is a data sheet of daily bookings. This booking information categorizes the daily demand by major manufacturing process. For the system studied this was the 105P, 105T, 120P, and 120T product lines. Data from one year of bookings was used to ensure statistical integrity of the model.

The second module was the capacity module. First, a forecast of mean daily demand for the 120, 105, Pressure, and Temperature products was developed. This forecast was based on the exponential smooth forecast technique where the alpha time constant was optimized by performing a least squared error fit to the historical data. Each forecast was projected forward a specified amount of time (two weeks) and held constant for a fixed period of time (one month) before being updated to a new value. This method of forecasting demand provides better performance than merely a historical mean and reflects the periodic update process used by UE. Once this mean value was determined, a daily capacity requirement is attained by adding the input uplift. Note that this capacity requirement is in units per day and does not recognize any variance in product cycle time that may be caused by a change in mix. The assumption is that UE will invariantly produce the required demand. Finally a required takt time is calculated based on an input available time. This available time is somewhat discretionary, subject to the number of shifts worked, and if product can be scheduled into overtime. Fortunately the capacity model is insensitive to this effect since it is a model output.

Independent Capacity Model Formulation

The third module is an independent OLT engine. This model takes bookings data, "schedules and produces" it on a FIFO basis, assuming independence between each cell. By means of analogy, this production process can be viewed similar to water being poured daily into a funnel. Since the funnel has a fixed outflow diameter (capacity) any water (demand) exceeding the maximum volume that can flow through the orifice will queue in the mouth of the funnel. This water will pass through the orifice before the next day's pouring does (assuming no mixing, i.e. FIFO). However, if there is a day where there is no backlog of water and the daily quantity is less than the capacity, only the daily demand will be produced. Though the excess capacity is a waste of capacity, the alternative is the waste of over (early) production. Using this approach, OLT can be calculated as the elapsed time between the day the order arrives (poured into the funnel) and the day the last piece in the order passes through the orifice. Again, if demand exceeds capacity, the capacity quantity of product will be produced. This method was applied to each of the four cells at UE: the 105, 120, Pressure and Temperature test cells. Based on the chosen uplift and the way orders stacked, order lead times were tracked for each part of the production process over the course of one year. Except for queue delays, the only other assumed delay is half a day for scheduling and half a day for production. Typical results for a 10% uplift on the 120 Cell are shown in Figure 3-14.

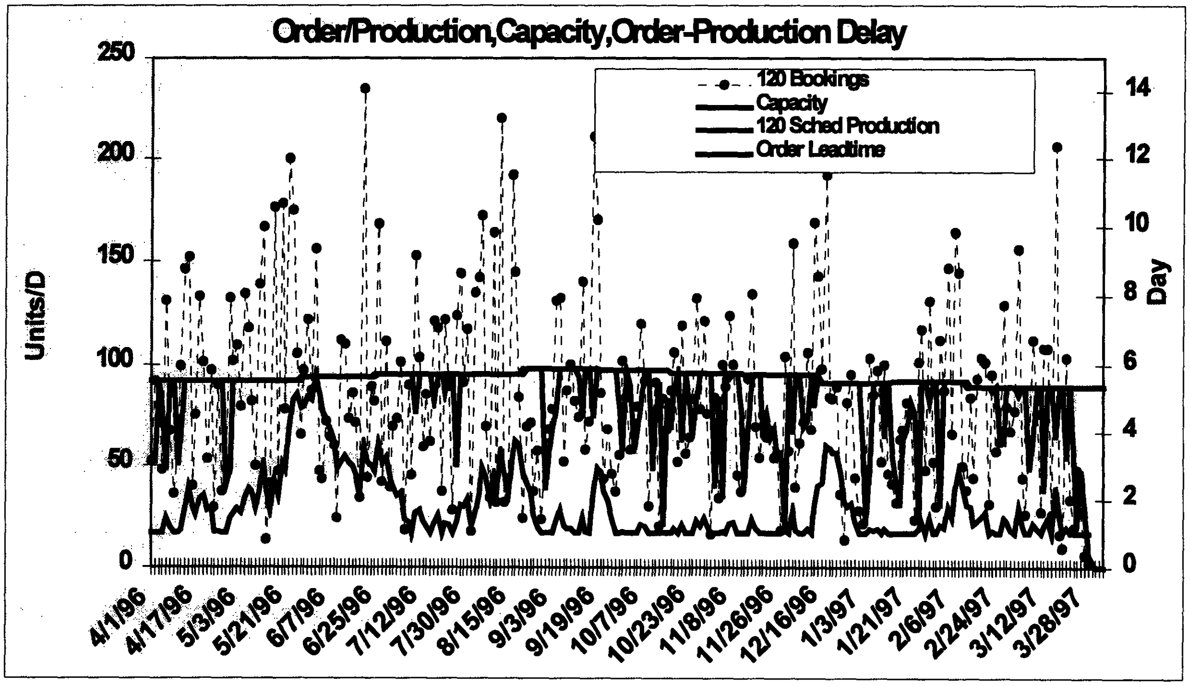


Figure 3-14: Lead Time For 120 Cell With 10% Uplift

As can be seen, whenever demand exceeds capacity, a delay is introduced. Furthermore, production cannot be held to a steady rate without introducing even further delays. Once performed for the 120 cell, this process is repeated for each of the other cells with their characteristic demand and assigned uplift. The independent assumption between process stages means that the order lead time (OLT) is the maximum order lead time for each of the four stages. This is intuitively obvious since the "order lead time clock" keeps ticking until the last product from a day's order has exited the production process.

Capacity Interdependence Formulation

The fourth module developed was the interdependent capacity module. This module builds on each of the three previous modules but recognizes that each stage of the capacity at UE is not independent but instead *highly coupled.* This observation meant that product could only be scheduled into the process if it could pass through the entire process during the day it was scheduled. To illustrate this point, imagine that two machining cells supply a final assembly cell that produces any range of mix of the two similar but different products. This is shown in Figure 3-15, below:

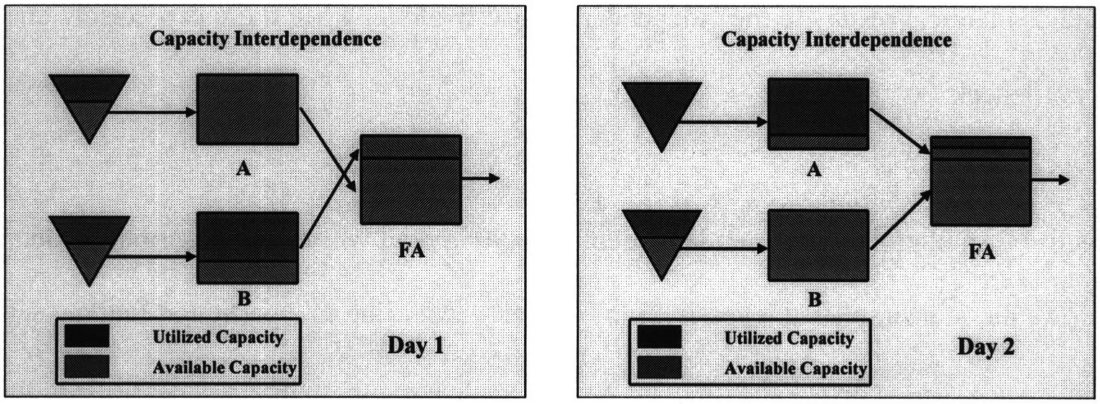


Figure 3-15: Illustration of Capacity Interdependence

The example shows that the demand for A fully utilizes Cell A. This, in turn, only partially utilizes the final assembly cell. The remaining capacity can be used to produce product B.

However, since little capacity is available, a queue is forced upstream of the assembly cell even though cell B has substantial excess capacity. On the second day, it may be that a large order (the size of B capacity) arrives, while the quantity ordered of type A becomes low. Because a portion of the B parts were delayed on the previous day, *though sufficient B capacity existed,* the second days orders become delayed as well. This happens despite the fact that sufficient capacity exists at the downstream final assembly to assemble more type B parts were they to get through their machining process. As a consequence of this interaction, actual system capacity is reduced and the order lead times are increased relative to the independent base case. Due to this effect, the order lead time results of independent capacity model only hold true when there is only one bottleneck in the process throughout time. This means that for relatively balanced lines the independent capacity model won't hold. Since the fourth module was developed, it stands that it should be used for all OLT calculations. United Electric has a demand characterization that results in swings similar to those illustrated in the example thereby necessitating its application. Use of the interdependent capacity model has been partially validated by confirming that it yields identical results to the independent capacity model when there is precisely one system bottleneck.

The specific formulation of the interdependent capacity model will not be described in detail here. However, at a simplified level, it "fills" capacity according to a particular loading scheme. There are a variety of schemes that may be used, some achieving higher average utilization than others. Despite the variety of load policies, if a daily FIFO is to be maintained, the backlog orders must be given first preference over the more recent orders. With this knowledge, the model fills the backlog capacity and analytically calculates the time until all orders have been completed, recognizing that different types or product with different backlog quantities will free up capacity at different times in the future.

3.5.2 Results

The capacity model gives strong support for the use of capacity uplift as a control variable for meeting OLT objectives. More specifically, the model provided the following insights:

- Small amounts of capacity uplift substantially reduce order lead times
- Diminishing returns for order lead time improvement exist as capacity is increased, i.e. each additional decrease in lead time requires an ever greater investment in capacity
- The interdependence between the different components of capacity, e.g. C_{120}, was shown to be particularly strong at small uplift levels

The results from the capacity model are shown below:

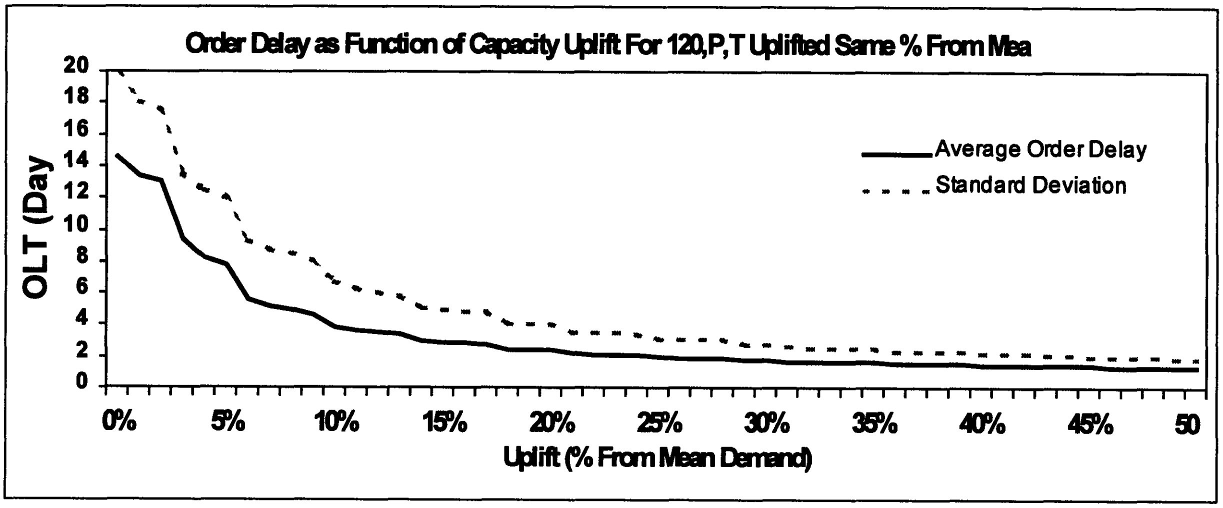


Figure 3-16: Order Lead Time as Function of Capacity Uplift – Same Uplift Throughout

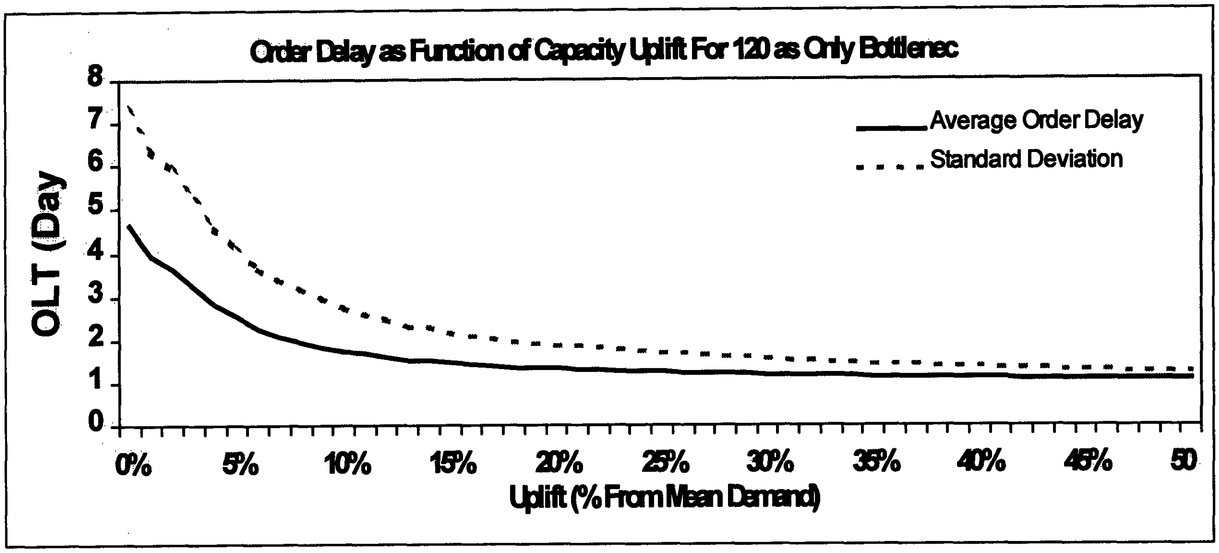


Figure 3-17: Order Lead Time as Function of Uplift – Assembly Cell Only Bottleneck

The plots above show the relationship between capacity and order lead time (OLT) for the most recent year of bookings at UE. The first plot shows the relationship between uplift and OLT if the 120 cell is considered the only process constraint (the result of the independent model) This assumes that no 105 orders are delayed and any range of pressure/temperature mix can be readily accommodated. The second plot shows the tradeoff between OLT and capacity if each stage in the assembly process were uplifted by the same amount, i.e. the 120 cell is uplifted 20% from its average daily demand, and the pressure and temperature tests are each uplifted 20% from their mean daily demands. Since each segment of capacity has been uplifted the same amount from their respective mean demands, the lines are relatively balanced (assuming similar demand profiles, c.v., etc.) As the graph shows, the OLT for this scenario is substantially worse for any given uplift than the situation when there is only one bottleneck. This finding strongly indicates that the company should not "balance" capacity in the system if it continues to use a pull system and can afford the additional capacity addition. The second key finding in the results is that the uplift has continuously less leverage in its ability to reduce lead time as uplift is increased. This indicates that a small amount of uplift makes a substantial improvement in system performance.

As discussed in the introduction, there is a coupling relationship between the cost of capacity and the change in revenues. This is a nonlinear relationship and UE must determine where its point of maximum net earnings lies. This will enable UE to choose the appropriate OLT based on the mean and standard deviation of order lead time. This desired OLT will then determine the required uplift that UE should carry for each of its production stages.

3.6 Concluding Remarks

This case study has demonstrated how Object-Oriented Axiomatic Design can be used to characterize the dynamic nature of variation in capacity planning through the use of influence diagrams and interdependence matrices. Application of the principles to reduce unwanted coupling behavior outlined in Section 2.3.3, in conjunction with the computer model, have given United Electric the understanding necessary to substantially improve corporate performance. The diagrammatical methods employed provide a clear mapping of existing interdependencies, as well as methods to control behavior. As a result, the company now understands why it cannot simultaneously reduce capacity, inventory, and order lead times. Moreover, the mapping of the coupling strength between capacity and order lead time gives the company the ability to optimize corporate performance.

In summary the following recommendations were made to UE:

- Establish uplift levels in the range of 10% - 15% for the main assembly line
- Set uplift levels for other processes, such as temperature or pressure tests, to greater than twenty-five percent

Setting capacity to "excess" levels on sub-processes reduces the interdependence strength of different capacity levels on order lead time. Finally, setting the standing capacity of the assembly cell to 10% - 15% will reduce lead times to approximately five days. This improvement in lead time will only be achieved if this extra amount of capacity is built into the system on a daily basis. Either by scheduled overtime or higher staffing, the company must be able to produce up to the uplift quantity. If the drive to increase worker and machine utilization pushes production managers to reduce capacity from the system, the lead times will revert back to the previous state.

Chapter 4

Case Study: Capacity Planning at Ford Motor Company

4.1 Outline of Case Study

The case study at Ford Motor Company is broken down into seven major sections. The first section, 4.2, *Page 158*, describes the study itself. Both the project plan and expected deliverables are discussed. The second section, 4.3, *Page 161*, provides an overview of the market and description of capacity planning process at Ford. The third section, 4.4, *Page 172*, presents the key findings resulting from the work. Section 4.5, *Page 173*, characterizes the system at Ford from an Object-Oriented Axiomatic Design perspective, specifically addressing the interface between two internal organizations, Capacity Planning and Product Development. This provides the basis for a detailed discussion of the results stemming from the OOAD interdependence work. This portion of the case study, found in Section 4.6, *Page 194*, describes how the propagation of variance can be reduced within the Capacity Planning function (Section 4.6.1), how changes to the Capacity Planning process can result in global performance improvement (Section 4.6.2), possible improvements to the Product Development organization (Section 4.6.3), reduced coupling between the Capacity Planning and Product Development organizations (Section 4.6.4), as well as methods that can be used as a hedge against variation (Section 4.6.5). Section 4.6.6 discusses other OOAD techniques evaluated, but not employed as part of the case work. Section 4.7 provides a summary of recommendations presented to Ford management that directly result from the OOAD Interdependence Study. The case study conclusion is given in Section 4.8. The final section of the case study, Section 4.9, *Page255*, is an appendix containing detailed mathematical derivations of key functions used in the OOAD analysis.

4.2 Description of Study

In 1994 Ford Motor Company launched a global reengineering effort called Ford 2000. The mission of this effort was twofold. First it would merge two independently operating arms of the company, Ford of Europe and North American Operations (NAO) into one single global organization, Ford Automotive Operations (FAO). The second component of the reengineering mission was to establish a center for process leadership (called Process Leadership imaginatively enough) that would lead the company in organizational learning and process improvement. The following year, five distinct reengineering efforts were simultaneously deployed. Each of these efforts focused on a separate process within the company.

Manufacturing capacity planning was one such process. A fragmented, poorly understood, and costly (a large portion of the company's $950M annual lost sales were attributed to the capacity planning process) process characterized the current state of Capacity Planning at Ford. In 1995, a reengineering effort was launched to resolve these issues. That same year, MIT was invited to work with Process Leadership and develop a formal design methodology that could be used for design and analysis by the different reengineering teams. It was felt that Capacity Planning Reengineering (CPR) would serve as a good test-bed or "lab" for this work since the capacity planning process interoperated closely with all other major groups within Ford and had to satisfy a varied and sometimes competing set of customer wants.

The ensuing study, over the course of a sixteen month period, developed and applied the Object-Oriented Axiomatic Design approach described in Chapter 2. The original project plan broke the study into four steps. Each step is detailed below:

1. Identify couplings.

The first step of the study was to develop and apply a methodology that would identify major couplings that cross the boundary between Capacity Planning and the other major components of Ford. Working together with members of the CPR team, the perimeter of the new Capacity Planning Process would be explored to determine how interactions with other organizations would promote or inhibit the performance of Capacity Planning. Comprehensive interviews would be undertaken with key individuals who were considered to be heavily impacted by the new process.

2. Classify couplings.

As couplings were identified, they would be classified according to a taxonomy that was to be concurrently developed. Having found instances of coupling, the relevant question to ask is: are these coupling relationships desirable in that they promote cross-functional integration and alignment, giving Ford competitive advantage or do they reduce inter-departmental flexibility and interfere with efficient operations? Developing this taxonomy would add formalism and provide a vehicle for communication of these issues.

3. *Resolve couplings.*

   Providing recommendations for resolving couplings would be integral to the success of the Object-Oriented Axiomatic Design effort. Analysis was to be performed on each type of macro coupling that occurred at the Capacity Planning interface to determine the type and variety of solutions which would be viable options for Ford. These coupling effects would then be ranked according to a grading system incorporating both the impact to Ford as a whole and the ease by which solutions may be implemented. A final report/presentation would summarize key findings and progress to date on the resolution of adverse couplings.

4. *Compile coupling methodology.*

   Finally, the methodology employed was to be captured in a well-documented form so that it can be reused in other areas of process engineering. This would provide Ford a fully validated approach for dealing with high-level couplings within the corporation. Workshops and training programs could then be conducted to transfer the methodology for use in other parts of the organization.

These four steps were followed closely. Not only was a formal methodology developed and applied, satisfying parts 1, 2, and 4 of the MIT-CPR interdependence study, but significant contribution was made in the elucidation and resolution of high-level couplings. Final recommendations were reviewed with senior management at Ford and expected to bring substantial benefit not just to the capacity planning process itself, but to Ford as a whole.

The purpose of this written case is to meet three objectives. First, it demonstrates the application of Object-Oriented Axiomatic Design in an actual business context. Second, it highlights several common problems caused by variance that confront capacity planning and capital investing functions within companies. Third, the details of this case are provided to a level of depth that will allow Ford to continue developmental work in this area.

4.3 Overview of Industry and Capacity Planning Process

The overview of the automotive industry and the capacity planning process at Ford is broken down into three parts. The first section, 4.3.1, provides an overview of the U.S. auto industry, the predominant source of Ford's sales revenue. The second section, 4.3.2, provides an overview of the current state capacity planning process used at Ford. The term *current state* is used to refer to the capacity planning process used by Ford prior to the reengineering effort as opposed to the *future state* process which will be the new process resulting from the Ford 2000 reengineering effort. The third and final overview section, 4.3.3, details the current process assumptions. Both the process and assumptions are essential for understanding the sources of problems with the current process as well as being essential for establishing a framework for the proposed axiomatic-based resolution of adverse systemic behavior. Section 4.4 will follow the overview by presenting the major findings of the work.

4.3.1 Nature of Industry

Five key attributes characterize the U.S. auto industry:

- Cyclical Market Movements
- Low Domestic Growth
- Unpredictable Market Segment Shifts
- Highly Integrated Product Offerings
- High Capital Investment Costs

These attributes make capacity planning a central element within automotive companies. Investment decisions can be risky and involve incredible sums; mistakes or sub-optimal decisions can cost the company millions of dollars. As a consequence, the capacity planning process is a highly visible part of Ford operations.

To provide a more comprehensive background to the case study, each attribute will be discussed in the following sections.

Cyclical Market Movements

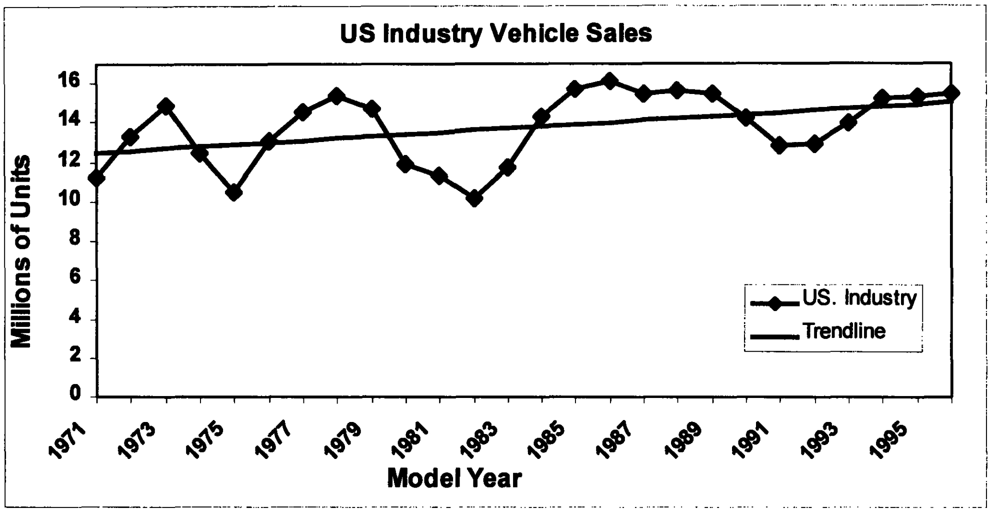


Figure 4-1: U.S. Auto Industry Vehicle Sales

Understandably, as America's largest industry, the U.S. auto industry is fundamentally tied to the performance of the economy. Thus, the economic cycle strongly affects annual auto sales as shown in Figure 4-1. This cyclicality, in recent years has been about 15% when measured as a peak deviation from the historical trend (regression in Figure 4-1).

From the perspective of capacity planning, this cyclical industry dictates that excess capacity is invested to ensure that the company can capture the high demand occurring during cycle peaks. The tradeoff is that the company must invest and carry capacity that will be used only a fraction of the time. A key role of capacity planning, then, is to establish a manufacturing capital investment structure that minimizes the cost of lost sales and cost of unused capacity.

Low Domestic Growth

The U.S. market is nearly flat. Over the past 25 years, automobile sales have increased an average of less than 1% per year. The slope of the trend-line in Figure 4-1 indicates that the market is essentially saturated. Since the U.S. market is the largest market in the world, low growth makes competitors compete predominately on cost and market share. Misjudgments of required capacity result in either lost share or a burgeoning cost structure. Excess capacity is essentially waste; since growth is so limited, it is unlikely to ever be used.

Unpredictable Market Shifts

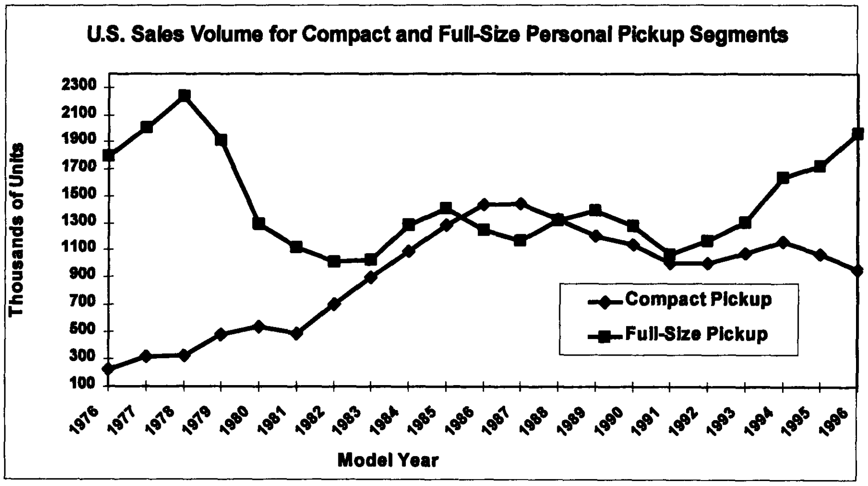


Figure 4-2: U.S. Industry Segment Trend for Compact and Full-Size Pickup

What may be even more difficult to manage than cyclicality, from the perspective of capacity planning, is the volatility in the individual vehicle segments. These segments, dividing the total industry volume into fifteen categories such as compact utility, luxury, and basic small, experience sudden and unpredictable changes in direction in addition to prolonged periods of growth or decline. While changes in segment volume generally show positive correlation with each other due to the cyclical movements of the entire industry, these segments move negatively with respect to each other when measured as a percent share of industry.

As Figure 4-2 shows, changes in industry volume coupled with changes in segment percent share of industry result in substantial segment volume changes over time. During the period 1989 through 1996, for example, production of full-size pickups increased by nearly 50% while during the same period, production of compact pickups fell by 25%. For comparison, the movement of the entire auto industry during this time was quite limited.

Segment movements occur for a variety of reasons, few of which can be forecast well. Customers change preferences due to changes in the economy, gas prices, personal image, age, personal utility, and the relative attractiveness of product offerings in competing segments among others. This segment volatility is not limited to a few segments; rather all segments are affected as shown in Figure 4-3, below.

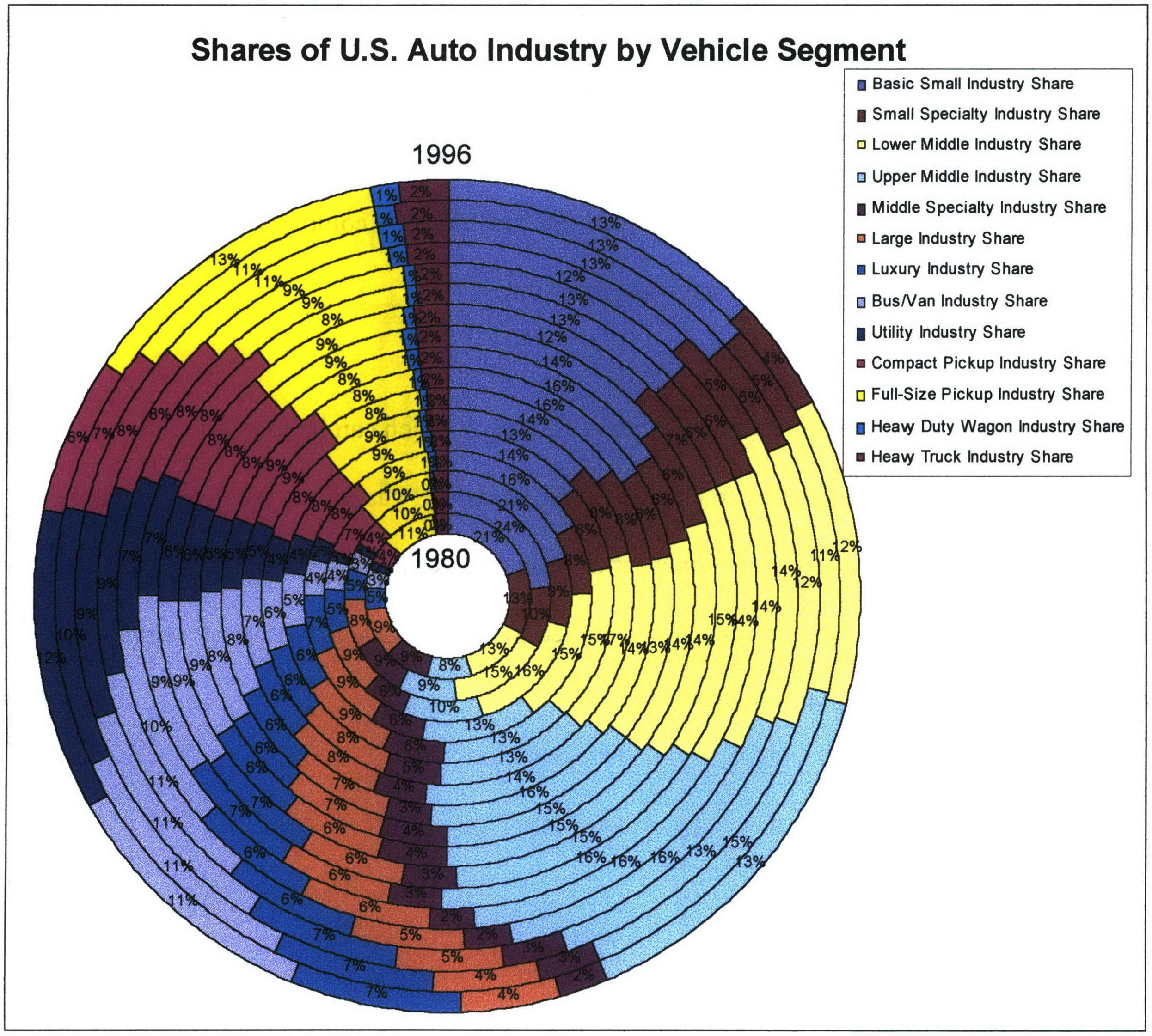


Figure 4-3: Segment Shares of U.S. Auto Industry

Even movements of share size by 1% are significant. This amounts to an increase or decrease of approximately 150,000 vehicle sales per year for that segment alone.

This segment volatility is also reflected at a corporate level although clouded slightly by changes in the corporate strategic direction, and by constrained demand fluctuation due to capacity constraints. This volatility, measured as segement share as a percent of total company sales is shown for Ford Motor Company in Figure 4-4.

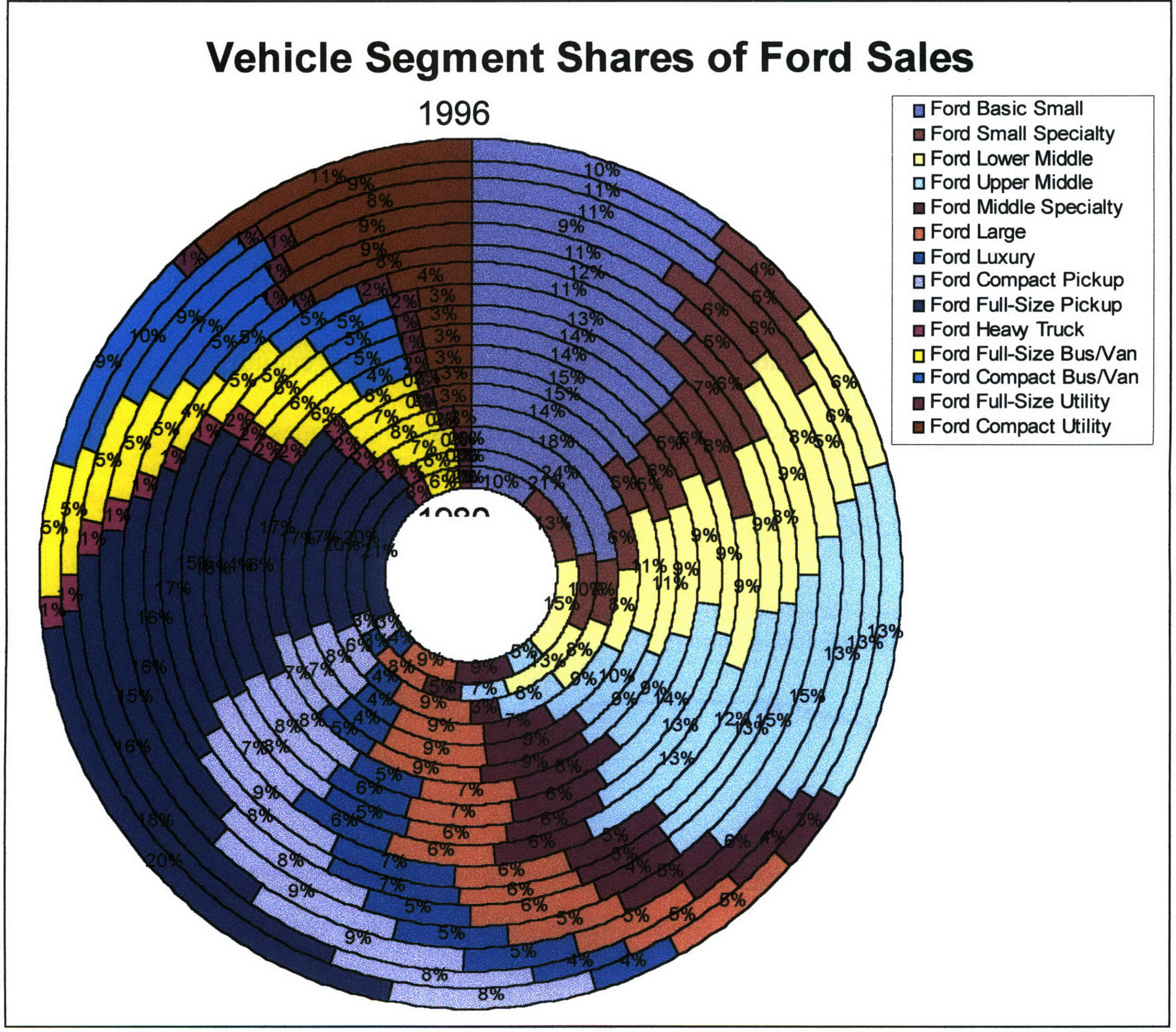


Figure 4-4: Ford Segment Volatility as Percent of Company Sales

Highly Integrated Product Offerings

The auto industry is further complicated by highly integrated product offerings. Each vehicle is the assemblage of more than three thousand parts. Each of these parts travels a different path through multiple manufacturing and assembly processes before it is incorporated into the final product. Some of these products are dedicated to single vehicle lines, while other components are sourced to multiple lines. Aside from optional vehicle components, if insufficient capacity exists for any standard component, an inadequate number of cars will be produced, and sales will be lost. For this reason, capacity planning is an essential element of each and every manufacturing and assembly operation.

High Capital Investment Costs

The auto industry is heavily capital intensive. As shown in Figure 4-5, capital expenditures in facilities and tooling consume a notable portion of the company's revenues.

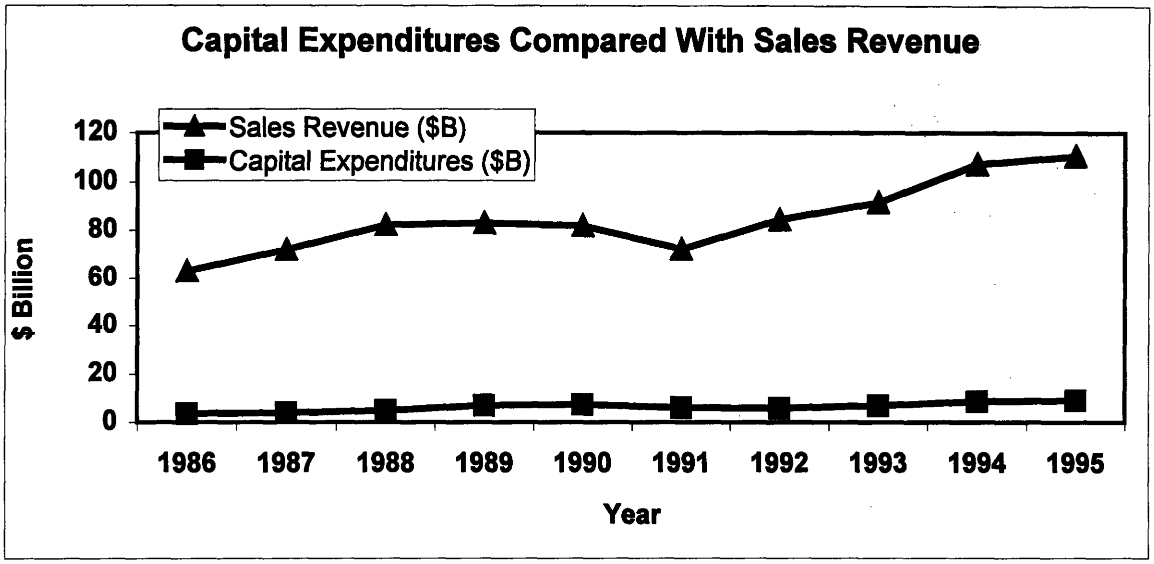


Figure 4-5: Ford Capital Expenditures vs. Sales Revenue

On average capital expenditures have been seven percent of Ford's sales revenues. With the company only earning 2.5% return on sales, (ROS), it is extremely important to keep capital costs under control. A change in capital costs as small as a couple percentage points will cost the company hundreds of millions of dollars annually.

4.3.2 Overview of Current State Capacity Planning Process

The capacity planning process at Ford is multifaceted, interfacing with numerous organizations and business processes within the company and its suppliers. Instead of providing a comprehensive overview, only select attributes of the process will be highlighted.

Three major elements comprise the capacity planning process:

- Establish Capacity Protection Volumes (CPVs)
- Use 5-Box Process for capacity issue identification and resolution
- Perform Capital Investment Analysis

CPVs are the required capacity volume levels for every component and sub-component made by Ford or its direct suppliers world-wide. Capacity Protection Volumes ensure the entire manufacturing base is capacitized to the same level. Because automobiles contain so many parts, it is absolutely essential that there are no shortages that prevent the vehicles from being built. Manufacturing capacity levels are an instrumental piece of this equation.

An eight-step process is used to determine part-level CPVs:

1. Interface with Cycle Planning (a strategy and finance group) to establish new product life-cycle strategies that make optimal use of existing facilities
2. Establish uplift parameters for each vehicle segment to determine most profitable level of protection
3. Add to forecast projections for all forward planning years
4. Constrain to assembly/engine/transmission capacity where appropriate

5. Establish world-wide take rates for vehicle options
6. Establish protection levels for vehicle options
7. Establish protection levels for combinatorial options (i.e. power windows but not power door locks requires use of a unique wiring harness)
8. Explode volume through the Bill of Material to establish capacity protection volumes by part type

The result of this process is a capacity protection volume by part type and forecast year. This number is the required capacity level that must be installed.

The "5-Box Process" is a systematic approach used throughout the company to ensure that capacity is always aligned with customer demand. This proactive approach ensures that any capacity discrepancies (shortfalls or excesses) are identified and resolved in a timely manner. The 5-Box process contains five steps similar to Edwards Deming's PDCA loop. These are:

- Determine capacity requirements
- Identify gaps between requirements and capacity
- Develop recommendations
- Approve recommendations
- Implement resolutions

This process is iterative as shown below.

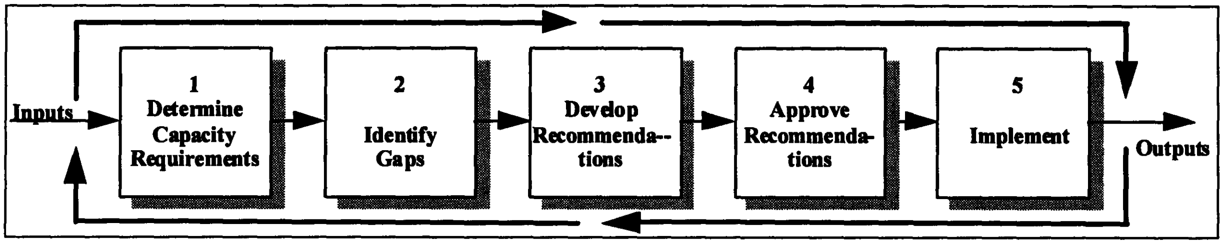


Figure 4-6: "5-Box Process" – Ford Capacity Planning

The final component to Ford's Capacity Planning is capital investment analysis. The Capacity Planning group analyzes, proposes, and secures funding for capital intensive capacity investments. This is particularly true for expansion proposals. Most frequently, these in-depth analyses are performed when capacity volumes are constrained.

4.3.3 Current Process Assumptions

As with any large process, innumerable assumptions exist regarding the capacity planning process. The ones related to this work are presented here. The capacity planning process has made the following assumptions:

- Type of capacity has no effect on the amount of capacity required for protection, i.e., capacity is capacity. One hundred thousand units per year of capacity is assumed the same whether it is flexible, dedicated, cellular, etc.
- 100% forecast accuracy
- All investments incur the same amount of risk, and therefore use the company cost of capital in financial evaluations
- Lost sales caused by capacity constraints where demand is greater than CPV are acceptable. The assumption is that the company cannot profitably afford to support capacity levels higher than CPV
- Lost sales caused by capacity constraints where demand is less than CPV are unacceptable and the fault of the manufacturing organization. The assumption is that

CPVs never change and insufficient capacity is the result of negligent investing rather than changes to CPV requirements within investment lead times.

These assumptions are only process assumptions; the individuals involved in and customers of the process frequently understand all-too-well how far the process deviates from reality. Nonetheless, it is this process that is a principal driver of corporate performance.

4.4 Findings

In this case study several findings were made. It was found that:

- Substantial room for improvement in capacity planning process alone
- Propagation of variance by Capacity Planning can be reduced
- Adverse coupling exists between the Product Development organization and Capacity Planning organization
- Coupling can be reduced and system performance can be improved by changing the way objective measures, namely ROS, within Product Development are calculated
- Observed couplings can be managed by aggregating control and establishing a joint manufacturing and product development strategy

These findings resulted in a series of specific recommendations to improve system performance. The summary of recommendations resulting from this work can be found in Section 4.7, beginning on *Page 248*.

4.5 Characterization of System

4.5.1 Definition of Variance

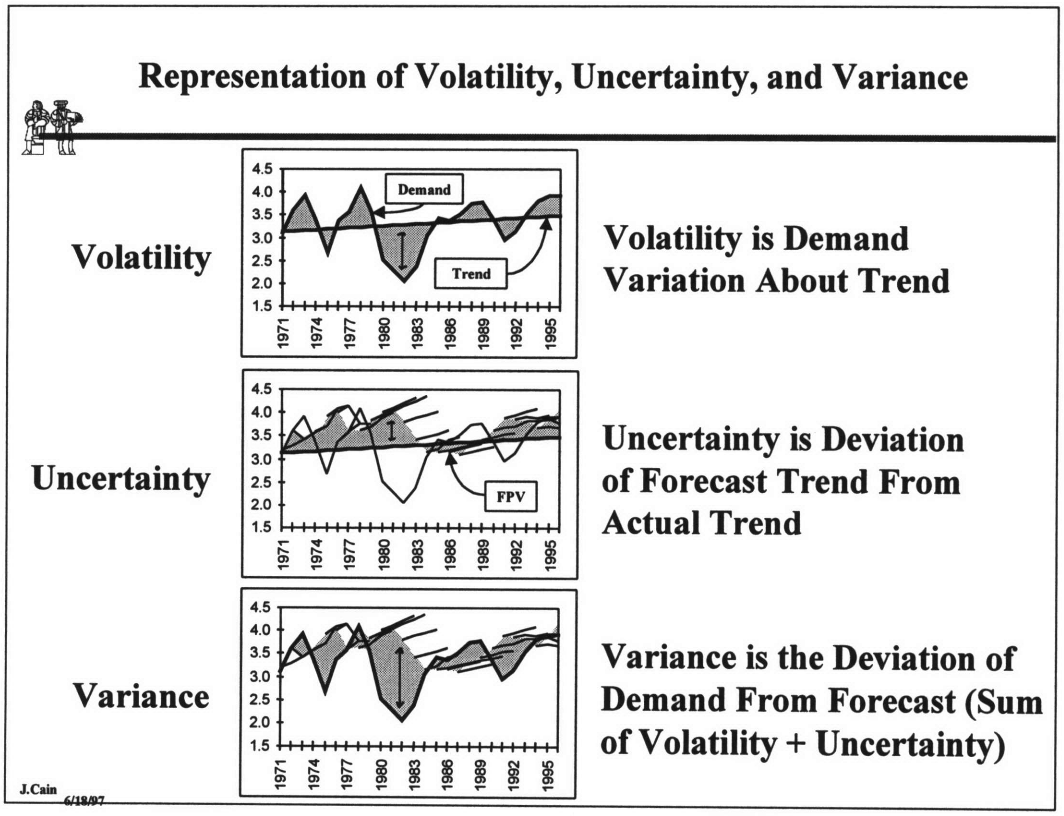
First and foremost, capacity planning is an information process. The quality of this process is governed by its ability to perform along three dimensions:

- The ability to properly select and process the most relevant information pertaining to the objectives of capacity planning
- The ability to efficiently process and disseminate this information in order to respond to market dynamics in an appropriate and timely manner
- The ability to effectively discern and filter out the irrelevant information among the input information gathered. This minimizes the error in the output data (capacity planning volume statements, recommendations, and actions) due to noise or aberrations in the input data

Since capacity planning is inherently tied to the forecast process, it is valuable to distinguish between the useful information (actual demand movements) and unwanted information (forecast uncertainty or error). The component of the forecast due to actual demand movements will be called *demand volatility* and the component of the forecast due to forecast error will be called *demand uncertainty*. The sum of demand volatility and uncertainty will be called *demand variance*. Since demand volatility and uncertainty are correlated, i.e. they cannot safely be assumed independent, demand variance is the *non-additive* sum of the two components.

Fortunately, demand variance is not difficult to measure. Demand variance is measured as the deviation of actual demand from the forecast demand.

For illustration, the differences between demand volatility, uncertainty and variance are shown below.



Representation of Volatility, Uncertainty, and Variance

Volatility

Volatility is Demand Variation About Trend

Uncertainty

Uncertainty is Deviation of Forecast Trend From Actual Trend

Variance

Variance is the Deviation of Demand From Forecast (Sum of Volatility + Uncertainty)

J.Cain 6/18/97

Figure 4-7: Representation of Volatility, Uncertainty, and Variance

These definitions of *volatility*, *uncertainty*, and *variance* will be strictly used throughout the rest of the case.

4.5.2 Propagation of Variance

Propagation of variance occurs when one process changes in such a way as to deviate from its expected value and then passes this change on to another process, causing it to be affected as well. In a manufacturing setting an example of this behavior would be when a variation of cycle time of an upstream process affects the performance of a downstream process. In the context of capacity planning, propagation of variance would be measured by how much changes in forecast values causes downstream customers of the capacity planning process to be affected.

Propagation of variance is a substantial issue at Ford where capacity numbers constantly change as the forecasts change. These changes generate numerous capacity studies, creating substantial make-work and undermining customers' confidence in the accuracy of the capacity protection volumes. Any changes that can be made to reduce this propagation of variance will reduce cost and improve the operational effectiveness of the organization.

4.5.3 Coupling Interdependencies

Early in the study with Ford, it was speculated that there was a high likelihood for coupling interdependencies to exist across the boundary between Capacity Planning and Product Development. It was proposed that the operational behavior within the two communities caused adverse effects reflected in the objective measures of the two groups. More specifically, it has historically been assumed that Capacity Planning volumes were considered fully sufficient for adequately protecting the company against the risk of capacity related stock-outs (CSOs) and preventing the installation of unnecessary capacity. However, this study hypothesized that the type of capacity invested is not independent of its effect on the avoidance of CSOs and excess capacity.

A more detailed study of this effect revealed that the Product Development organization's efforts to increase its return on sales (ROS) adversely couples with the Capacity Planning organization's two objectives, to minimize CSOs while simultaneously maximizing capacity utilization. Thus this warranted a full mapping of the interdependencies, in order to characterize the couplings present. This then would enable a set of actions to be proposed that would effectively manage or eliminate the adverse effects.

When the analysis of this specific coupling was presented to Ford management, it was concurred that these initial findings were in agreement with the general sentiment that the company's incentive structure was driving the company away from necessary manufacturing flexibility. It was believed that the automotive market was becoming more volatile, reducing overall predictability and thereby necessitating manufacturing practices that increased manufacturing flexibility rather than ones that decreased it. This sentiment is reflected in Figure 4-8.

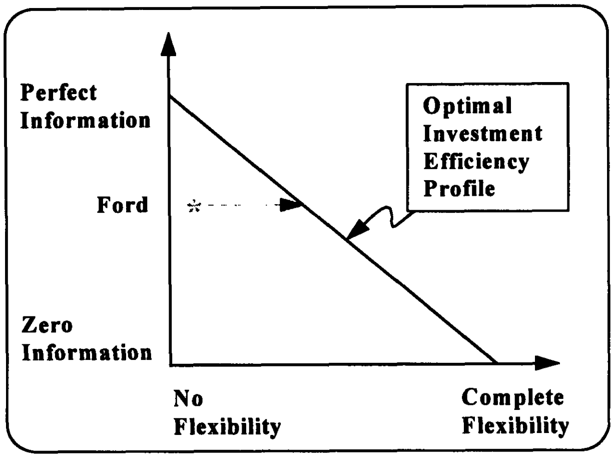


Figure 4-8: Investment Efficiency Profile Given Quality of Information

While senior management held this view, it was unclear to them the specific mechanism which induced this behavior, and even less clear to what degree this adversely affected the company's performance. The expectation was that the application of Object-Oriented Axiomatic Design (OOAD) would elucidate the specific drivers and mechanisms that give rise to this behavior. The hope, then, would be that this methodology would provide the basis for a set of recommendations to improve the alignment between these two organizations.

4.5.3.1 Mapping of the Boundary Interface Between Capacity Planning and Product Development

The mapping between Capacity Planning (CP) and Product Development (PD) is comprised of two components. The first is the mapping of the relationship between the principal control variables and objectives for PD. A similar approach is then used for Capacity Planning. The second step is to analyze the boundary interface between the two groups.

Mapping of the Product Development Domain

One of the strongest incentives within Product Development is the Return on Sales (ROS) metric. ROS is defined as the profit before tax divided by total revenues for a given product (4-1).

$$ROS = \frac{PBT}{R} \tag{4-1}$$

Currently, Ford's ROS averages between two and three percent, while competitors achieve ROS figures as high as four to five percent. The corporate objective is to consistently meet or exceed a five percent ROS. To increase ROS, there are several things which the company can do. Before looking at what controls exist to increase ROS, it is worthwhile to look at the variables which actually comprise the calculated ROS figure. Principally, ROS can be improved either through increasing revenues or through decreasing costs incurred. These variables are shown in detail in Figure 4-9, below.

ROS Influence Matrix	Salesi	PPUi	VCPUi	VMPUi	FTLE	OHi	Lifecycle	Other
ROS	+	+	-	-	-	-	+	-

Figure 4-9: Return on Sales Influence Matrix

Other than the Sales and Price Per Unit (PPU), the Variable Cost Per Unit (VPPU), the Variable Marketing Per Unit (VMPU), Facilities, Tooling, Launch, and Engineering (FTLE), and Overhead (OH) can be considered costs that in each year, *i*, over the life of the product offering, reduce the ROS. In the product development process, a calculated ROS is used to predict the performance of the product offering in the marketplace. During the design and development of a new product offering, the ROS is determined by using the expected (forecast or estimated) values for each year over the expected product lifecycle. To build the influence matrix, each of these variables are changed (given all other variables held constant) and the change in the output is observed. If the change in the dependent variable is of the same sign as the variable changed, a "+" is used. If the outcome is in the opposite direction, a "-" is used. If the equation is a nonlinear function, dependent on the state of other variables to determine the direction of outcome, an "**X**" is used. The full equation used in computation of ROS is not shown here but can be found in appendix Section 4.9.1.

While the influence matrix is useful in demonstrating the relationships the key variables have on the objective metric, it has two weaknesses that make it necessary to develop influence diagrams for ROS. The first reason is that the influence matrix does not show the control variables that the product development organization can use to improve the Return on Sales. For example, while it is possible to increase the price per unit and increase ROS (all else constant), it is not possible to change the overhead cost directly. The second reason is that the influence matrix does not show the interdependencies between the intermediate variables. An example of this effect is the commonly recognized relationship between expected sales and price per unit, (PPU). The demand elasticity curve widely used dictates that for typical commodities demand will fall as price increases. Depending where one is on the curve, an increase in price per unit (PPU) may or may not result in a net increase in ROS.

To more fully capture these relationships and provide a basis for demonstrating how Product Development's drive to increase ROS couples with Capacity Planning's objectives, a ROS influence diagram was developed. This diagram is shown in Figure 4-10.

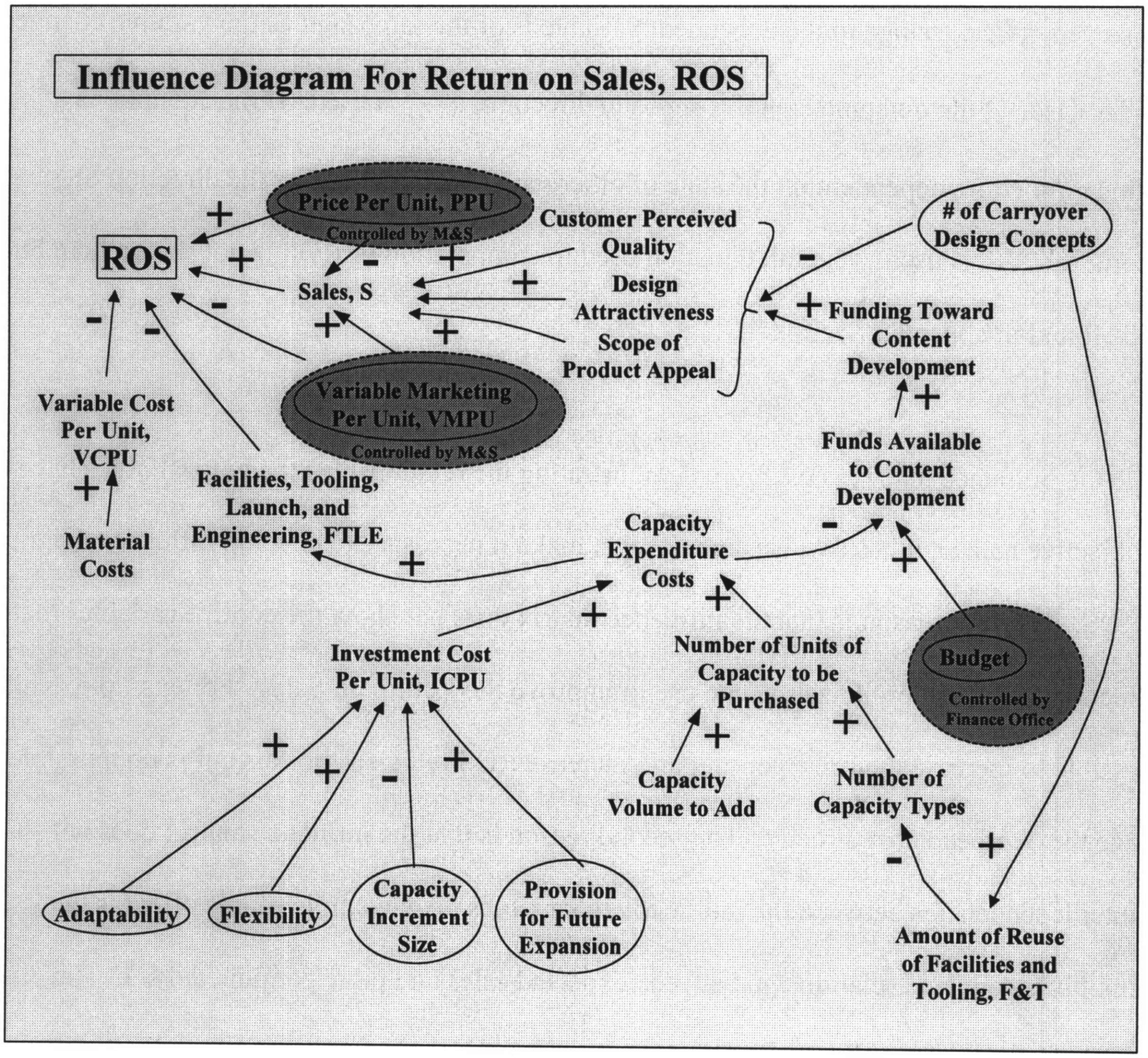


Figure 4-10: Influence Diagram for Return on Sales, ROS

As shown in the figure, there are a number of controls "owned" by Product Development. In the influence diagram the objective, ROS, is indicated by a square and the control variables are indicated by ovals. The principal controls presented here are: Adaptability, Flexibility, Capacity

Increment Size, Provision for Future Expansion, Program Budget, and Number of Carryover Design Concepts.

Adaptability refers to the ability to convert manufacturing capacity dedicated toward the production of one product to another product. Flexibility refers to the flexibility to manufacture two or more different product types on the same line with minimal changeover time. Designing products that support flexible and adaptable manufacturing typically increase engineering and capital investment costs. Similarly, provision for future expansion requires additional up front investment on the premise that a future capacity expansion is likely.

Capacity increment size is proportional to 1/MCT (Machine Cycle Time) and is typically viewed to reduce the capital investment costs if increment size is increased. This is due to scale economies that are achieved with larger manufacturing lines. While the advantages of economies of scale are seldom contested in terms of investment cost, this enduring quest for larger and faster manufacturing lines has actually added scale inefficiencies in many cases. Investment Cost Per Unit, (ICPU) actually begins to increase with increasing scale due to material processing rate limits and complexity effects that begin to dominate once the system exceeds a certain size. This behavior is shown in Figure 4-11.

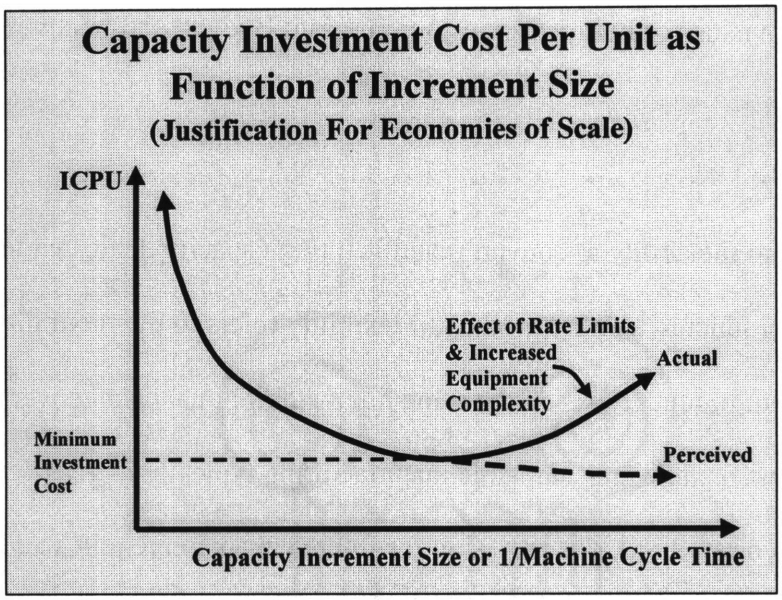


Figure 4-11: Economies of Scale Justification for Larger Capacity Increments

From discussions with members of the product development and manufacturing communities it is widely believed that the company is either below the point of decreasing returns on the whole or that these inefficiencies never become dominant terms and that economies of scale always exist. For this reason, the Return On Sales Influence Diagram, Figure 4-10, will assume that the relationship between Capacity Increment Size and ICPU is always positive.

Incentive Behavior to Increase ROS

Because ROS is a principal metric in Product Development, there is substantial pressure to increase its value. Based on the relationships between the product development control variables and the objective metric, ROS, it was found that there is incentive to:

- Increase reuse of previous design concepts on future model development. Increasing reuse will lower capital investment costs but at some point begin to adversely affect sales due to design compromise. Reuse will be increased to the point where the marginal contribution maximizes ROS

- Decrease the use of flexible manufacturing. There is no benefit to flexible manufacturing in the computed ROS
- Decrease the use of flexible manufacturing. Because adaptable manufacturing techniques principally benefit other or future vehicle programs there is no benefit to increasing adaptability in the computed ROS
- Eliminate provisions for future capacity expansion capability. There is no benefit to provision for expansion in the computed ROS
- Increase capacity increment size. Economies of scale encourage ever larger manufacturing lines

These incentives are expected to give rise to the behavior trend observed at the company.

Mapping the Capacity Planning Domain

The second step in mapping the boundary interface between Product Development and Capacity Planning is to develop an influence diagram for Capacity Planning. This influence diagram is shown in Figure 4-12.

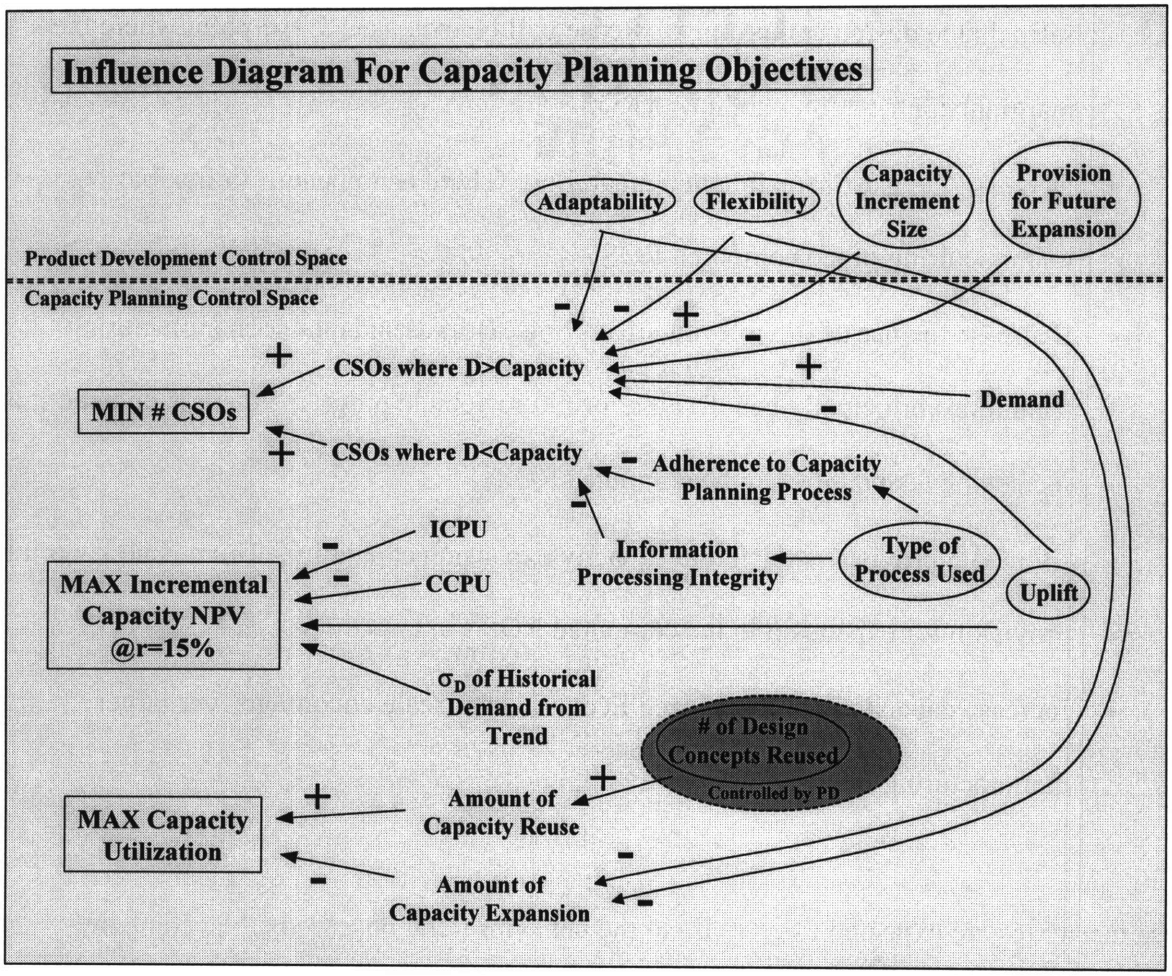


Figure 4-12: Influence Diagram for Capacity Planning Process

As shown, Capacity Planning has three key objectives:

- Minimize Capacity Related Stockouts, CSOs
- Maximize the Incremental Capacity NPV
- Maximize the company-wide average Capacity Utilization

To affect these objectives, Capacity Planning possesses two principal control variables. The first is *capacity uplift* and the second is the *type of capacity planning process* used throughout the company. Capacity uplift is principally used to maximize the incremental capacity NPV. This control variable seeks to balance the tradeoff between increased capacity investment and carry costs at higher uplift levels against the opportunity cost of expected lost sales when demand exceeds capacity. No sign is provided in the influence diagram since an increase of uplift may or may not increase the incremental NPV depending on the state of the system. The secondary or indirect effect of uplift is to decrease capacity related stockouts as uplift is increased. This effect is understood, but not formally recognized by the uplift process and therefore is not a factor considered by Capacity Planning when it seeks to maximize its performance.

The second control employed by Capacity Planning is the type of capacity planning process used in the development of Capacity Protection Volumes for each part number. Errors in the capacitization process increase the quantity of capacity related stockouts. This situation arises when a stockout occurs when demand is less than the capacity protection volume, indicating that the manufacturing facility was incorrectly facilitized. Since Capacity Planning owns the process used and is responsible for the effectiveness of its execution, any process improvement will result in a decreased number of capacity related stockouts.

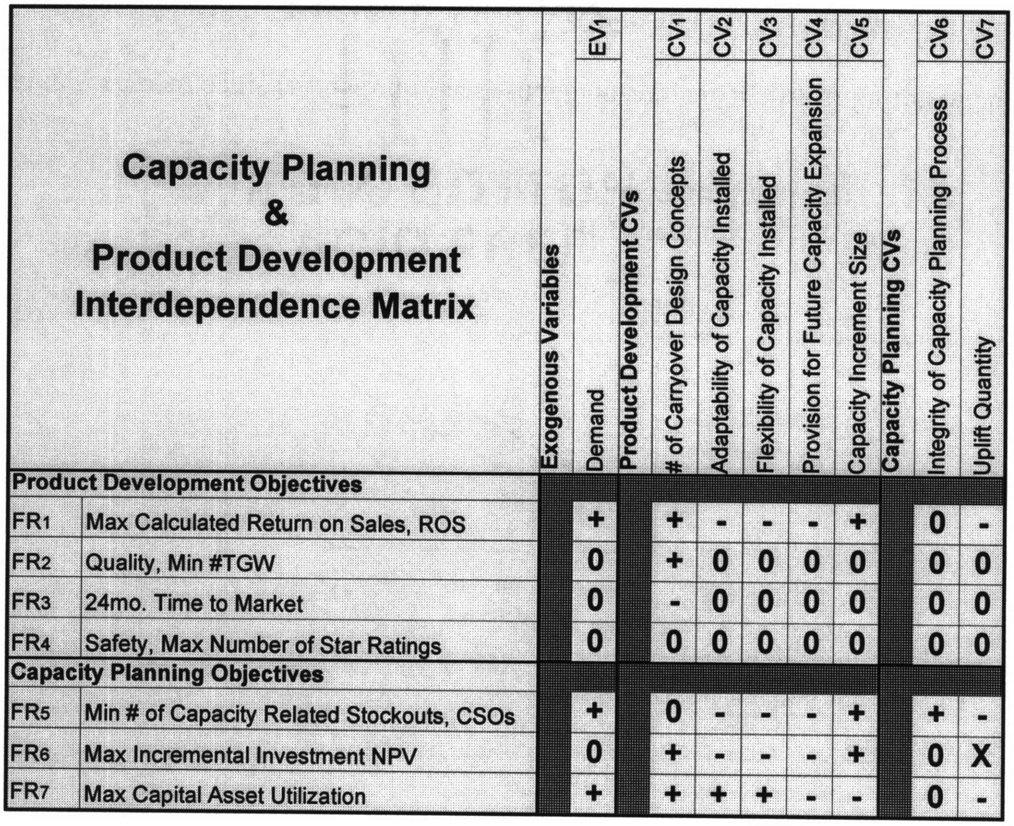
The Capacity Planning – Product Development Interface

Aside from the Capacity Planning control variables, the five Product Development control variables depicted in Figure 4-12 markedly affect the performance of the number of capacity related stock outs and the overall capacity utilization of the company. Reductions in flexible manufacturing, manufacturing adaptability, and provisions for future expansion limit the company's ability to handle situations when demand deviates from forecast and therefore will increase the Capacity Related Stockouts over time. As well, increases in the capacity increment size and amount of design reuse will increase the CSOs and facility utilizations, respectively.

Since interdependencies cross the boundary interface between Capacity Planning and Product Development, coupling exists. From the two influence diagrams, the nature of the coupling behavior can be inferred. However, an interdependence matrix will explicitly recognize these couplings and facilitate explanation and the development of recommendations to reduce or eliminate instances of adverse coupling behavior. This was the next step in the process of characterizing the interdependent system between Capacity Planning and Product Development. It's development is discussed in the following section.

4.5.3.2 Development of an Interdependence Matrix to capture the couplings between Capacity Planning and Product Development

To capture the couplings between Product Development and Capacity Planning at a high level, an bounded interdependence matrix was built. This bounded interdependence matrix, shown in Figure 4-13, highlights some of the most significant objectives and controls for both Capacity Planning and Product Development. The interdependencies shown in the matrix are abstracted from the influence diagrams presented previously. These interdependencies represent the current state interaction between the two groups, i.e. the couplings that were found to exist at the time the analysis was performed.



Capacity Planning & Product Development Interdependence Matrix		Exogenous Variables	Product Development CVs					Capacity Planning CVs	
		EV$_1$ Demand	CV$_1$ # of Carryover Design Concepts	CV$_2$ Adaptability of Capacity Installed	CV$_3$ Flexibility of Capacity Installed	CV$_4$ Provision for Future Capacity Expansion	CV$_5$ Capacity Increment Size	CV$_6$ Integrity of Capacity Planning Process	CV$_7$ Uplift Quantity
Product Development Objectives									
FR$_1$	Max Calculated Return on Sales, ROS	+	+	-	-	-	+	0	-
FR$_2$	Quality, Min #TGW	0	+	0	0	0	0	0	0
FR$_3$	24mo. Time to Market	0	-	0	0	0	0	0	0
FR$_4$	Safety, Max Number of Star Ratings	0	0	0	0	0	0	0	0
Capacity Planning Objectives									
FR$_5$	Min # of Capacity Related Stockouts, CSOs	+	0	-	-	-	+	+	-
FR$_6$	Max Incremental Investment NPV	0	+	-	-	-	+	0	X
FR$_7$	Max Capital Asset Utilization	+	+	+	+	-	-	0	-

Figure 4-13: Current State Interdependence Matrix for PD and CP

As shown in the figure, advantageous coupling exists for CV$_1$, the number of carryover design concepts, coupling (either advantageous or adverse depending on the sign of FR$_6$-CV$_7$) exists for CV$_7$, the uplift quantity, and a combination of advantageous and adverse coupling exists for control variables 1-5. In addition to these couplings, product demand can be considered an exogenous variable, which, depending on its direction of change, can advantageously or adversely couple with FR 5 and FR 7, the number of CSOs and capacity utilization, respectively.

Couplings resulting from changes in Control Variables (CVs)

The coupling resulting from CV_1 is advantageous (see Figure 4-14) since increases in reuse to improve ROS also reduce the average investment cost per unit. This results because a greater percentage of previously existing capacity will be reused if a capacity change is made. Additionally, increased reuse results in higher capacity utilization since a greater portion of manufacturing capacity will be used in new product offerings that would otherwise be idled.

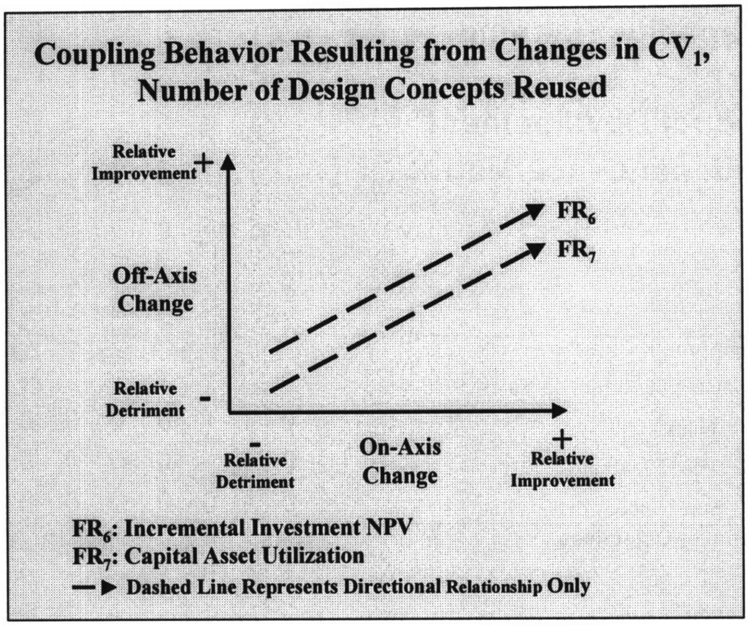


Figure 4-14: Current State Coupling Behavior Due to CV_1

The coupling resulting from CV_7 can be either positive or negative, as shown in Figure 4-15.

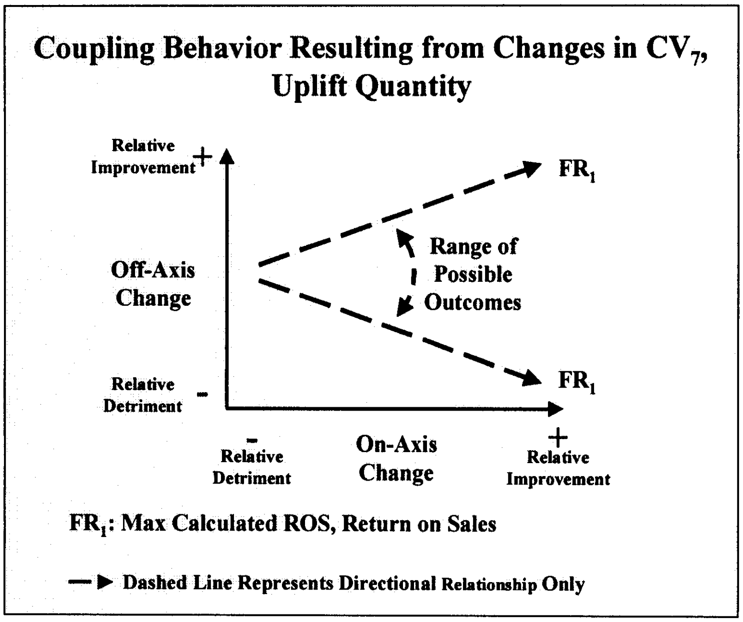


Figure 4-15: Current State Coupling Behavior Due to CV_7

If increasing the capacity uplift increases the incremental NPV, (FR_1), ROS will be reduced. However, if increasing the capacity uplift results in capacity investment and carry costs that do not generate the expected additional revenue (through additional sales) to offset the incremental investment and carry costs, the uplift will be reduced. Reducing the uplift decreases the total amount of capacity that needs to be invested by Product Development thereby increasing the calculated ROS.

Coupling caused by changes in the adaptability of capacity installed has both advantageous and adverse components. See Figure 4-16.

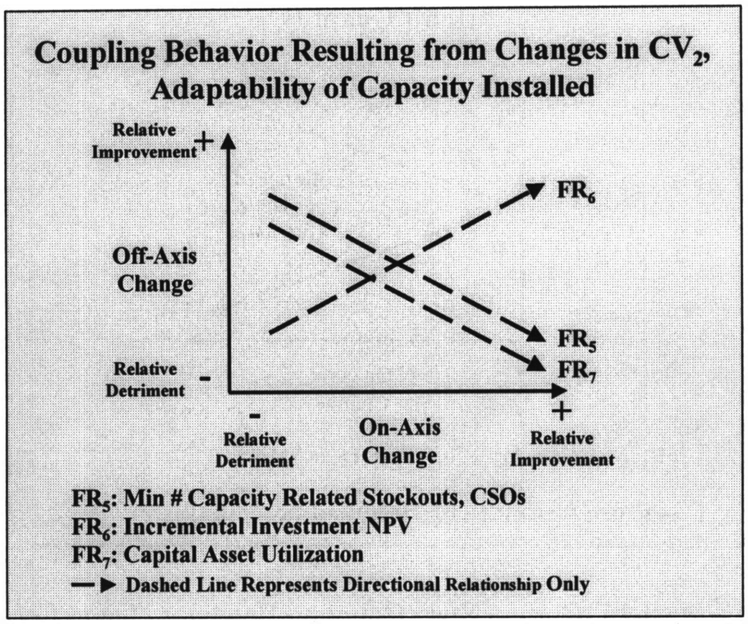


Figure 4-16: Current State Coupling Behavior Due to CV_2

Reducing the adaptability of the manufacturing capacity may lower the investment cost due to a simpler or more dedicated design. On the other hand, reduced adaptability leaves manufacturing capacity less responsive to changes in demand. Capacity related stockouts will increase as capacity adaptability is reduced, since the lead time for converting capacity (assuming adaptable capacity is available) is less than the investment lead time for new capacity. Likewise, reduced adaptability causes capital asset utilizations to fall. This situation arises frequently when demand in a vehicle segment experiences a prolonged downturn and capacity is idled. Dedicated capacity that cannot be adapted toward another use is a waste of a valuable asset.

Reduced flexibility has the same effect on corporate performance as reduced adaptability.

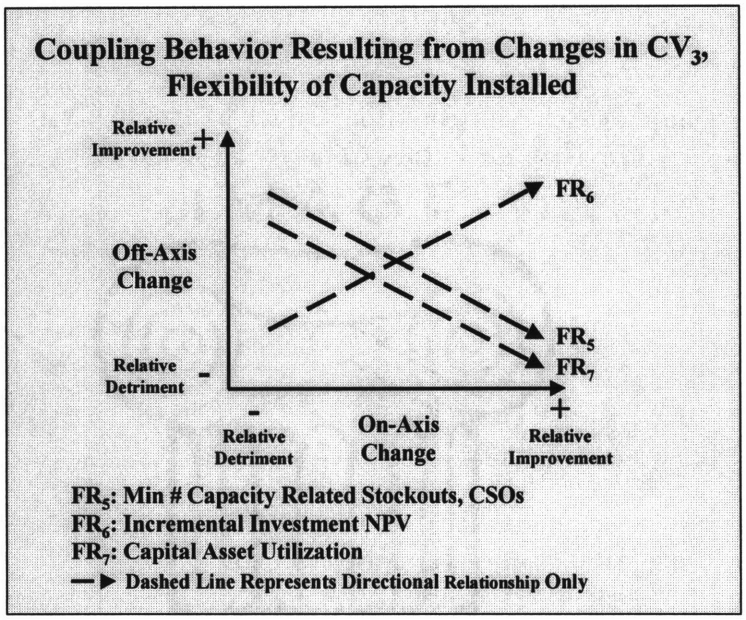


Figure 4-17: Current State Coupling Behavior Due to CV_3

Lost sales occur when production is at a maximum capacity on one line, while another has idle time. Were the two lines flexible, the second line could handle the excess demand. This would not only reduce the probability of lost sales, but would increase asset utilization.

The coupling behavior caused by the provision for future expansion has a higher upside than downside when provisions for options are increased. This is shown in Figure 4-18, below.

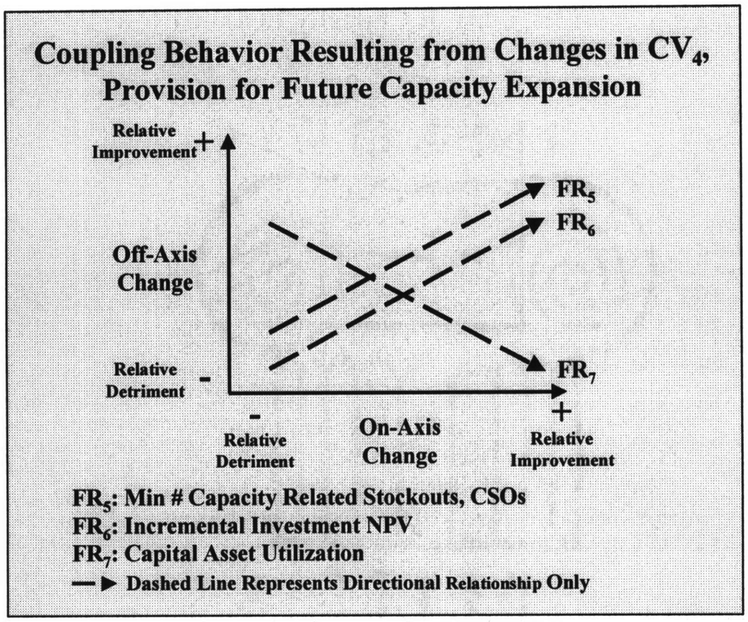


Figure 4-18: Current State Coupling Behavior Due to CV_4

Option provision increases investment returns and reduces lost sales, with the only down side of decreased asset utilizations. Decreased utilizations result when provisions for expansion decrease expansion lead times and make such expansions easier to cost-justify.

Since the current state trend of product development was to decrease options provision with the intent of increasing ROS, the overall nature of the coupling to Capacity Planning's objectives can be considered adverse.

The final coupling identified in the PD-CP interdependence matrix is the effect of changes in capacity increment size upon Capacity Planning's objectives. As shown in Figure 4-19,

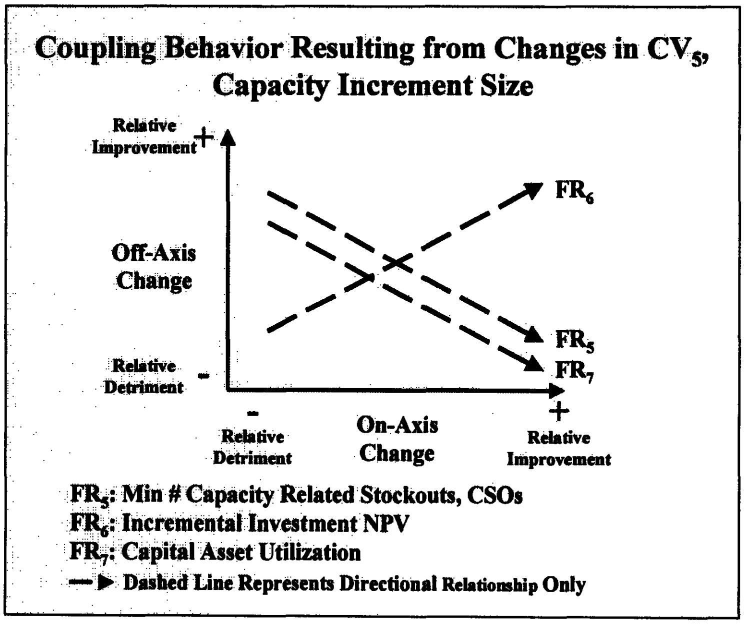


Figure 4-19: Current State Coupling Behavior Due to CV_5

increases in capacity increment size advantageously couple with incremental investment NPVs since economies of scale achieve lower investment costs per unit as the increment size is increased. However, increased capacity increments cause more stockouts and decrease utilizations. Stockouts are increased since increased increments have accordingly longer investment lead times, leaving the manufacturing system less responsive to changes in demand. Since increasing increment size makes capacity less likely to match demand, lower capacity utilizations are to be expected.

Couplings Resulting from Exogenous Variation

The coupling caused by changes in demand is stochastic. To the extent that demand increases, the overall coupling effect on the Capacity Planning organization will be advantageous. This is shown by EV_1, in the PD-CP interdependence matrix, Figure 4-13. However, since variance in demand is stochastic, i.e. principally driven by market movements that cannot be explicitly controlled, the stochastic variance component of demand adversely affects system performance. The nature of this effect is described in 2.1.1.6, and should be minimized if at all possible.

4.6 Results of OOAD Interdependence Study

The mapping of coupling interdependencies between Ford's product development and capacity planning organizations highlighted significant issues between the two organizations. This work provided the basis for the elimination of the coupling behavior between the two systems. The results of this OOAD Interdependence Study both in terms of findings and recommendations are detailed in the following sections, 4.6.1 through 4.6.6.

Section 4.6.1 addresses the propagation of variance. Specifically this section outlines techniques that can be used to damp the propagation of variance from upstream processes to downstream dependent customers.

Section 4.6.2 provides recommendations that can be used to improve the performance of capacity planning without adversely affecting other parts of the organization. This ensures that a local improvement results in global improvement.

Section 4.6.3 is similar to 4.6.2 except that it provides recommendations for improvement to the product development process.

Section 4.6.4 specifically addresses the adverse coupling behavior that was found to exist between Product Development and Capacity Planning. Recommendations are made to reduce the unwanted couplings.

Section 4.6.5 focuses on additional methods that can be used to hedge against variation. These methods assume demand to be an exogenous variable. Considering how demand affects different system objectives such as utilization, capacity stockouts, return on sales, etc, the methods outlined in this section can be used to reduce the off-axis coupling strength between demand and the objectives. These techniques are not in themselves a set of recommended changes such as those outlined in 4.6.4; instead they are a set of tools that can be used as warranted to achieve optimal results in an environment of perpetual change.

Section 4.6.6 provides a brief discussion of other techniques integral to OOAD but not part of the final set of recommendations to Ford. The purpose of the discussion is to provide insight into how different techniques might apply to a study such as this one performed at Ford.

4.6.1 Reduced Propagation of Variance

Propagation of variance is undesirable. Change fundamentally necessary for improvement should be welcomed, but change that occurs only because other things have changed is undesirable. As discussed in 2.1.1.7, propagation of variance increases system complexity, making it more difficult to comprehend as well as causing make-work that wastes resources.

Propagation of variance was acutely felt by downstream customers of the capacity planning process at Ford. In the course of studying the current state process and interviewing customers of the process, it became clear that propagation of variance was a substantial issue. The following comments made by Ford employees illustrate this point:

- “Capacity is mismatched: some plants work max overtime while others sit idle.” – Sr. Manager, Product Strategy Office
- “Capacity volumes can’t be trusted – they are either wrong or ‘all over the map’.” – Manufacturing Engineer
- “My process is only as good as the numbers coming into it.” – Capacity Planner
- “If Capacity Planning could just figure out what it wanted we wouldn’t have all these frivolous [capacity] studies.” – Plant Industrial Engineer
- “When subsequent Blue Book [reports containing required capacity volumes] issues contain different numbers for the same forecast period it is clear that someone in forecasting doesn’t know what’s going on.” – Manager in Vehicle Operations
- “Forecast projections by the Wall Street Journal are more stable and consistent than the ones issued by your company [Ford]. Improve your process and I’ll take it more seriously.” - Supplier

While none of the above comments make the differentiation, there are two components to propagation of variance. The first is variance due to fundamental market movements. The second is due to unnecessary change from within the firm. This second form of variance, called stochastic variance, is what the customers of capacity planning rightly consider waste. It is the responsibility of the capacity planning organization to prevent stochastic variance from being passed downstream even if the variance did not originate from within Capacity Planning.

Three sections are devoted to the reduction of variance propagation. Section 4.6.1.1 addresses the observation that much of the variance downstream customers experience is due to uncertainty in the forecast process. Application of some simple forecast techniques illustrate that substantial room for improvement exists. The second section, 4.6.1.2, uses knowledge about forecast uncertainty to buffer downstream processes from known uncertainties in the forecast process during forward planning years. The proposed technique does not reduce the stochastic variance input into the capacity planning process, but it does reduce the amount that is propagated to capacity planning customers. The third section, 4.6.1.3, outlines a risk-based capital investment strategy that can be used to reduce the variance passed on to the shareholders and stakeholders who have provided the funds necessary for the company's capital expenditures.

4.6.1.1 Improved Forecast Technique

Improvements to the forecast process for "free demand" sales volumes will 1) reduce the error between the expected and actual demand in any given year and 2) reduce the magnitude of changes to the forecast values over the forecast horizon. While the effects of forecast changes often only change spreadsheet models of projected earnings, cash flow, etc. in financial planning parts of the organization, the effect of forecast changes to the manufacturing capital investing part of the organization are costlier. Since all capital investments are made to a forecast, changes to the forecast often render previous investments unnecessary or require additional capital to be spent in an area where the forecast has increased. Reducing forecast variance (improving forecast accuracy) therefore reduces the variance propagated to the capital investing community.

Forecasting demand for each of Ford's vehicle lines involves three separate forecasts, industry volume, segment share, and share of segment. Industry volume is the aggregate volume for all cars and trucks in the U.S. auto industry. The segment share is the percent share that each vehicle segment commands, e.g. the percent share that Compact Utility vehicles command when all manufacturers are combined. Finally, the vehicle share of segment is the market share that each vehicle line (Ford Taurus for example) is expected to control. This share is measured as a percentage of the segment size. Multiplying all three of the forecasts together yields the projected demand for each product offering:

$$FPV_{i,i-j} = \overline{IV_{i,i-j}} * \overline{IS_{i,i-j}} * \overline{VSS_{i,i-j}} \tag{4-2}$$

where

FPV = Forecast Planning Volume

IV = Industry Volume

IS = Segment Share of Industry Volume

VSS = Vehicle Share of Segment Volume

i = Projection Year

i-j = Year Forecast Was Made

j = Forecast Horizon in Years

A variety of techniques are used to produce the different forecasts. The Economics Office at Ford uses a modified econometric model to forecast industry volumes while Marketing & Sales uses Delphi techniques to arrive at the segment and vehicle shares.

Although the purpose of the OOAD Interdependence Study was not to improve the forecast process at Ford, the availability of historical data made it easy to develop some simple models that could be used to identify areas of leverage in forecast improvement and therefore reduction in variance propagation.

From evaluation of historic data, it was found that the greatest source of error was in the forecast of segment shares. Volatility of industry volumes was relatively small (13%) and therefore forecast accuracy was high. Similarly, forecasts of vehicle segment shares were also good. This was either due to a quality process used in the projections of the market share Ford would capture in new product offerings and the expected product lifecycle or it was due to the "push" incentive structure that pushes the sales organization to meet share targets (current forecasts). Either way, vehicle share of segment was typically not far from the projections. On a relative basis, then, segment movements as a share of industry volume were the most difficult to predict. Improving forecast accuracy of segment share was expected to have the greatest impact in minimizing the propagation of variance.

Using twenty-five years of historical data for every vehicle segment, both in terms of actual and forecast demand, a simple forecast model was built and compared to actual Ford forecasts. Though any number of different forecast techniques could be used, the model developed was an equally weighted projection of a moving average and linear regression of historical segment shares. To optimize the accuracy of this model, two techniques were applied. First, the forecasts were adjusted such that the sum of the independently forecast shares for each segment would equal 100%. This improves forecast accuracy since, by definition, the sum of shares will always be 100%. Secondly, an Excel-based linear program was written to minimize the residual sum of squares between forecast and actual by varying the number of years history to be included in the moving average and linear regression. This optimization was performed independently for each year on the forecast horizon.

The result of the model show a significant improvement over the current forecast process used at Ford. This can be seen in Figure 4-20, below.

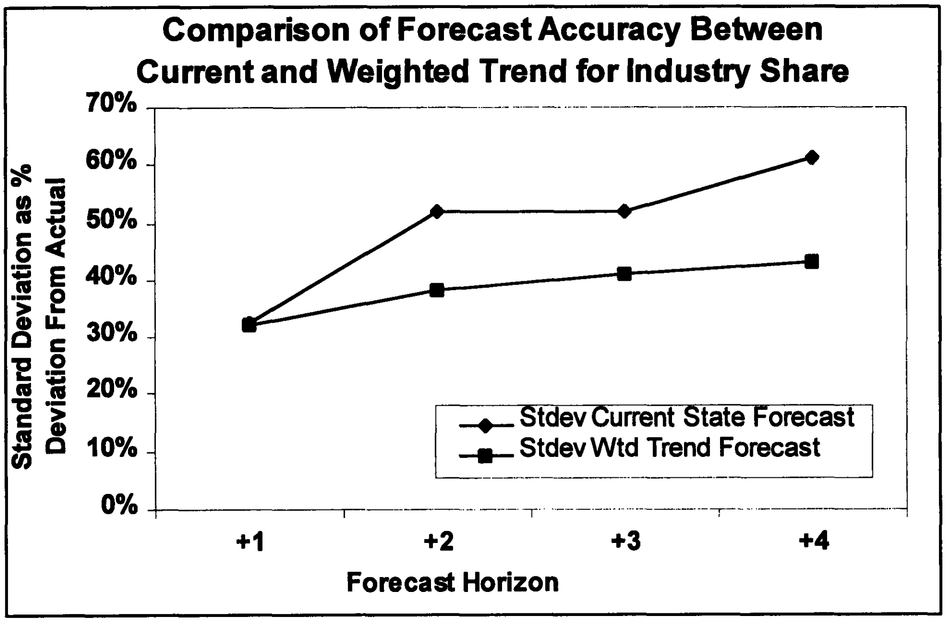


Figure 4-20: Comparison of Accuracy for Different Forecasts

As seen in the figure, the weighted trend forecast technique is superior to the current process by more than a year of forecast accuracy for all forecasts more than one year into the future. Since capacity planning is dependent on long-term forecasts, this is a significant improvement.

While the forecast model was not tested against new data to validate its accuracy and an analysis of variance was not performed to assess the appropriateness of the model as a forecast tool, it does illustrate that there is substantial room for improvement in the current process used at Ford. Any improvement will reduce the magnitude of forecast variance. This, in turn, will decrease the magnitude of change that the capacity planning organization passes on to its constituents.

4.6.1.2 Inclusion of Future Uncertainty in Uplift Process

The second way to reduce the propagation of variation is to include future uncertainty into the uplift process. Currently the uplift process, which determines the required level of vehicle volume protection, is based only on historical demand volatility. The volatility, which is measured as the distribution of demand relative to a historical linear regression, is assumed to be the distribution of demand relative to the forecast for all forward planning years. This is simply not true. It would be the case only if forecasts were 100% accurate. In practice, future demand deviates from the current forecast according to a distribution (the deviation of actual demand from the forecast – see Figure 4-21 for example) that widens the further the projection goes into the future. This effect is illustrated in Figure 4-22.

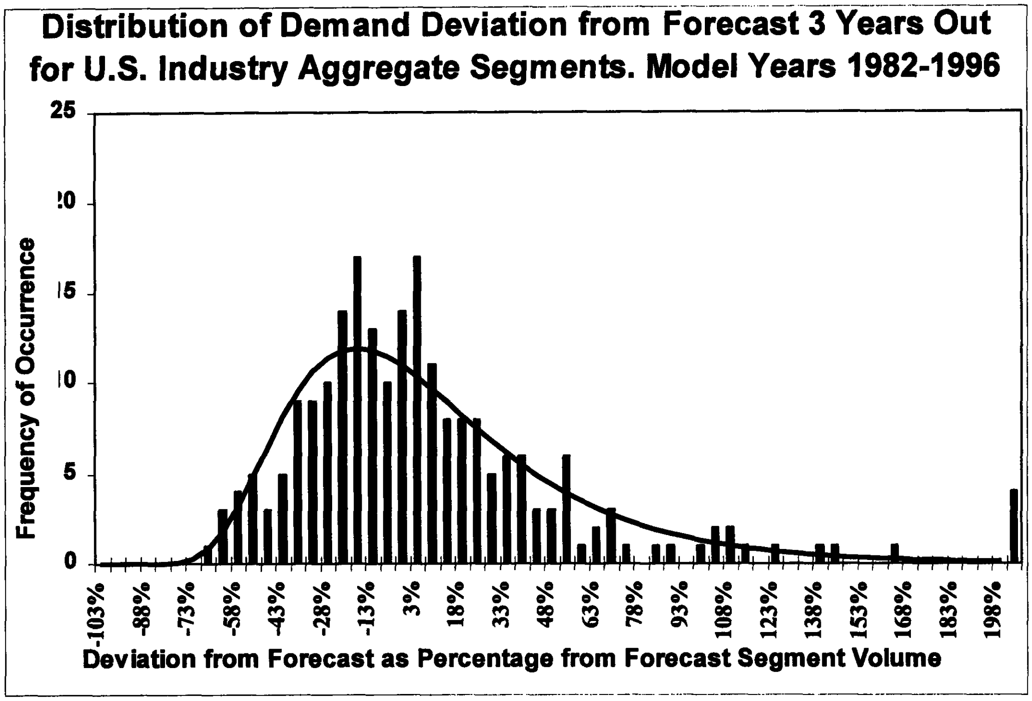


Figure 4-21: Forecast-Demand Variance Greater than Demand Volatility

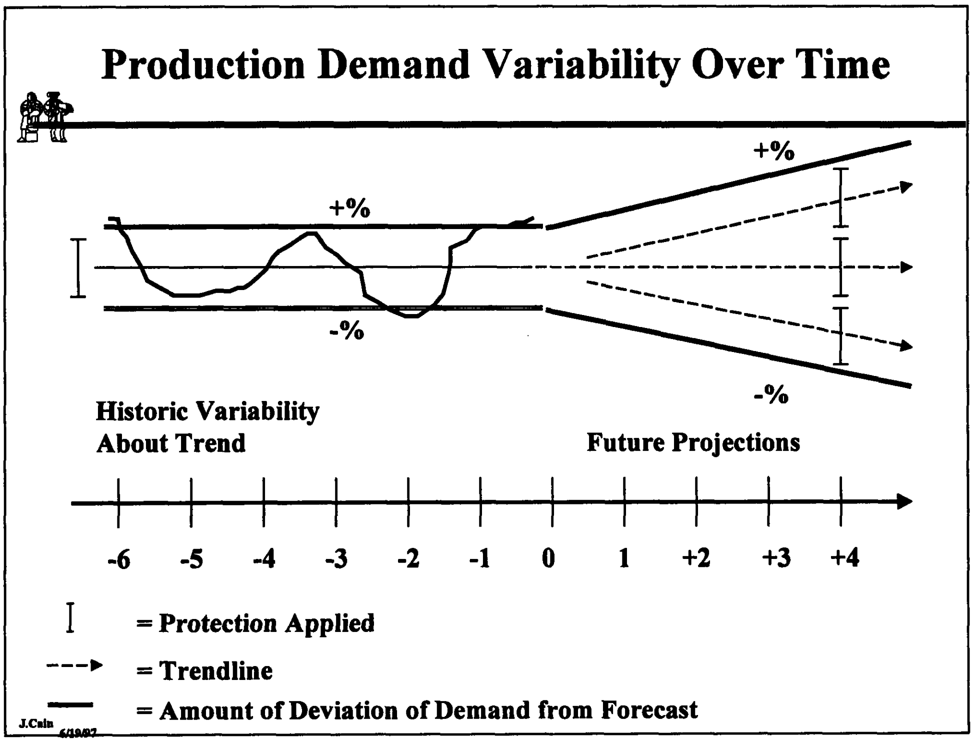


Figure 4-22: Uncertainty Increases Over the Forecast Horizon

If this uncertainty were to be included in the uplift process, uplifts for the same vehicle would increase for each forecast year into the future. The uplift will increase since the probability distribution of demand has a wider variance in future forecast years and therefore a higher probability of demand exceeding capacity for a given uplift. This is a fundamental change compared to the current process that has the same uplift percentage for all forward planning years.

Uncertainty-based uplifts will look similar to those illustrated in Figure 4-23.

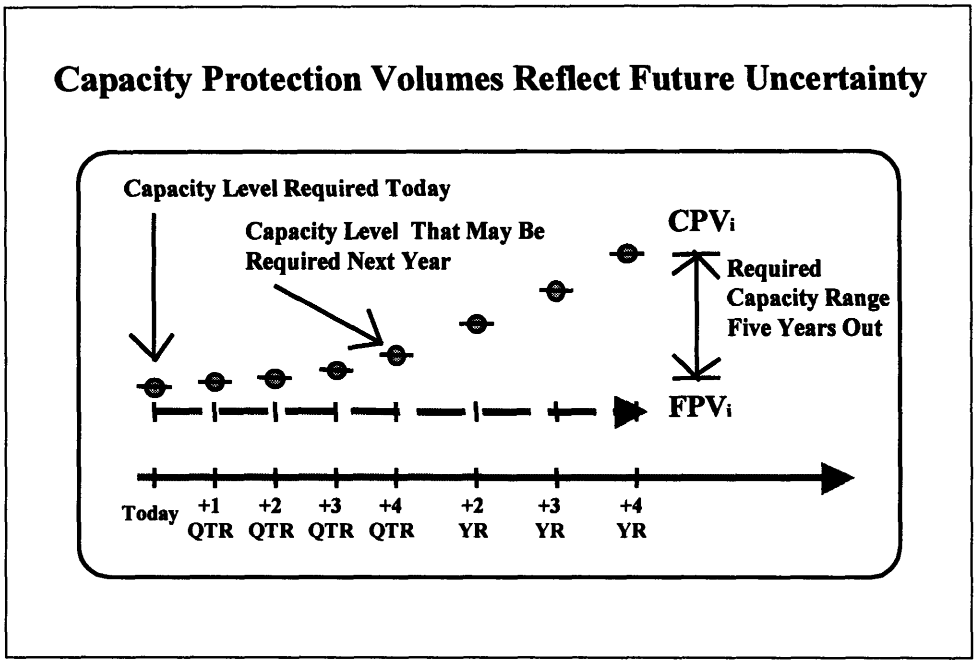


Figure 4-23: Forecast Uncertainty Applied to Capacity Protection Volumes

As shown, uplifts increase *relative* to the current forecast in forward planning years. This is a fundamental change from the current constant uplift approach. The current process assigns the same uplift to each year of the forecast to generate the capacity protection volume, CPV. The resultant CPV is then to be interpreted as the capacity the manufacturing organization will be required (and therefore expected) to support in each coming year. However, since forecasts are uncertain and change, fifty percent of the time, they will increase, and fifty percent of the time they will decrease. In practice, each time the CPVs increase, a new study must be performed to determine whether or not the existing capacity will be sufficient and if not, how much it would cost to increase capacity to the new level. Unfortunately these studies are often frivolous. By the time the future required capacity becomes the current required capacity, the volumes have changed. Half the time the volumes will have decreased, rendering the previous studies unnecessary.

In contrast, uncertainty-based uplifts are higher in future years to account for the amount the required capacity volumes may *increase* during the period of time where they move from a future requirement to the current requirement. In other words, the uncertainty-based uplift is the level the manufacturing organization must be able to support in each future year were that volume level to become reality when that year arrives. Thus, the current (one year forecast) requirement is the maximum affordable capacity the company has decided to support and is the amount required to be physically available on the factory floor. In accordance with this, all forward years' capacity protection volumes are *not* required to be currently supportable on the factory floor, but must *be able* to be met during the lead time that the future volume becomes the current required volume.

The advantages of this change in uplift strategy are threefold:

1. The difference between the current required capacity protection volume and the future, higher volumes, creates a tension that stresses the *capability* to support future changes that are not currently justified investment levels.

2. The inclusion of uncertainty in forward planning years reduces the number of "surprises" where an investment previously capacitized to a supposedly sufficient level suddenly becomes inadequate because capacity protection volumes have increased.
3. The decreasing uplift quantity that occurs as the forecast horizon becomes shorter reduces the frequency and magnitude of circumstances where the subsequent forecast requirement is higher than the previous forecast requirement. This reduces the propagation of variance.

The tension created by the difference between current and future requirements encourages a fundamentally different investment strategy. The "official" interpretation of capacity protection volumes is that investments made to CPV will be sufficient for the life of the asset. This encourages dedicated, unalterable investments with minimum investment costs. On the other hand, inclusion of uncertainty into the uplift process encourages flexible investments with future expansion options and short investment lead times.

Propagation of variance occurs when the forecast increases above previous levels. When this occurs, capacity protection volumes will increase relative to previous levels and customers of the capacity planning process will have to perform studies to assess the supportability of the new capacity volumes. In the event of a forecast downturn, propagation of variance will typically not be an issue since the previously supportable level will still be sufficient under the new, lower capacity requirement.

Inclusion of uncertainty into the uplift process reduces the propagation of variance by decreasing the number of times capacity requirements increase for the same forecast year. The reason for this is that for each movement upward in the subsequent forecast issue, the uncertainty will have decreased by the equivalent of one year of forecast uncertainty. For example, if the original forecast for a future year is low, i.e. lower than it ought to due to uncertainty, then in the following year it will be expected to increase due to a one year more accurate forecast. This would cause the capacity protection volume to increase except for the fact that the uplift has decreased by approximately the amount of the increase because the forecast is now one year more accurate. As a result, the new capacity protection volume is expected to be the same as the previous volume. In other words, the inclusion of uncertainty has already taken into account the typical changes in forecasts that occur from year to year and therefore changes to the forecast are less likely to cause change to the capacity planning customers. This effect is illustrated in Figure 4-24.

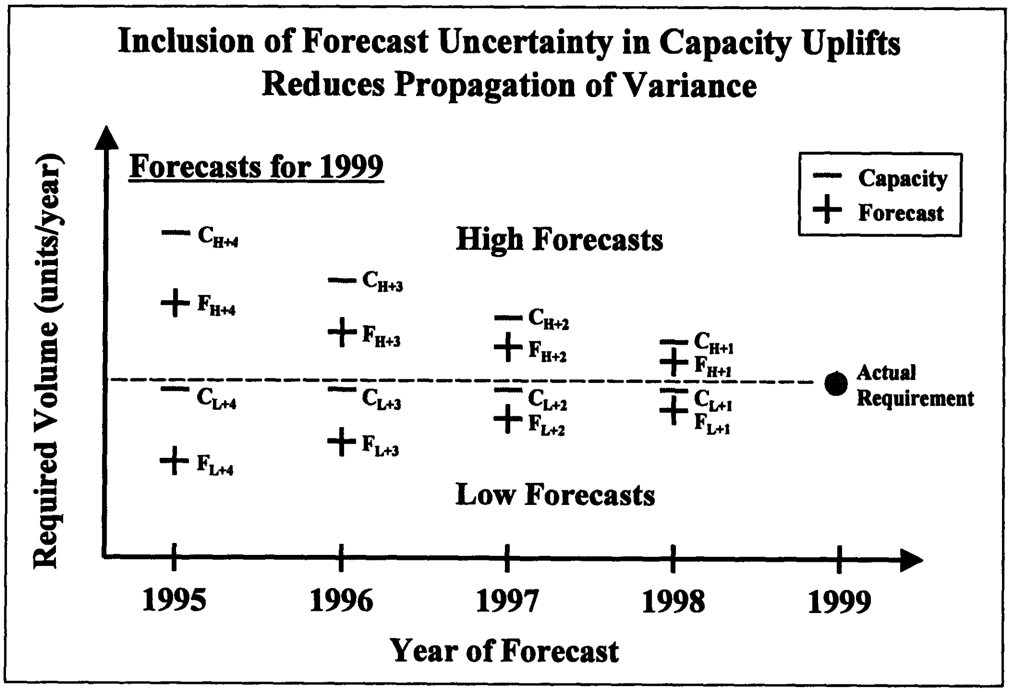


Figure 4-24: Reduced Propagation of Variance in Capacity Uplifts

For this reason, propagation of variance will be reduced if forecast uncertainty is incorporated into the uplift process.

4.6.1.3 Development of Risk-Based Capital Investment Strategy

Propagation of variance is also a significant issue to investors in the company. Each company typically has a different risk class. This risk class is based on an industry Beta which relates variation in investment return to fundamental variation in market return. For example, utilities have historically been quite stable, with a Beta of less than one. The automotive industry, on the other hand, has a Beta greater than one, because it produces durable goods which are more sensitive to industry movements.

From a particular company's perspective, all investments made by the company must meet or exceed the security market line discussed in Section 2.1.1.6 of Chapter 2. In practice, it is extremely difficult to assess the risk associated with any particular investment. Typically, the company uses the calculated cost of capital as a hurdle rate for all investments within the firm. This hurdle, or discount, rate is calculated in the following manner [38]:

$$\text{cost of capital} = r_{assets} = r_f + \beta_{assets} r_{p_m} = r_f + \beta_{assets}\left(r_m - r_f\right) \tag{4–3}$$

where

$$\beta_{assets} = \frac{D}{V}\beta_{Debt} + \frac{E}{V}\beta_{Equity} \tag{4–4}$$

r_f is the risk-free rate

r_{pm} is the market risk premium

r_m is the expected market return

D is the outstanding debt of the firm

V is the book value of the firm

E is the firm's equity

Finally, both B_{Debt} and B_{Equity} are the company's respective risk classes. These are calculated using the Capital Asset Pricing Model, given in Equation (2-1) of Chapter 2, Section 2.1.2.6. The calculated cost of capital is used to filter out investments that do not exceed the return of the security market line given the corporate risk. The assumption inherent in the cost of capital is that all investments by the company incur the same amount of risk. This assumption holds in general, but over the course of this work it was found that Ford's investment policies actually serve to increase the risk to investors without a compensating additional return. This means that a portion of the company's investments falls below the security market line. The specific finding was that investments made at higher levels relative to the forecast had much higher risk exposure than those that were made at lower levels. These respective investments are illustrated in Figure 4-25.

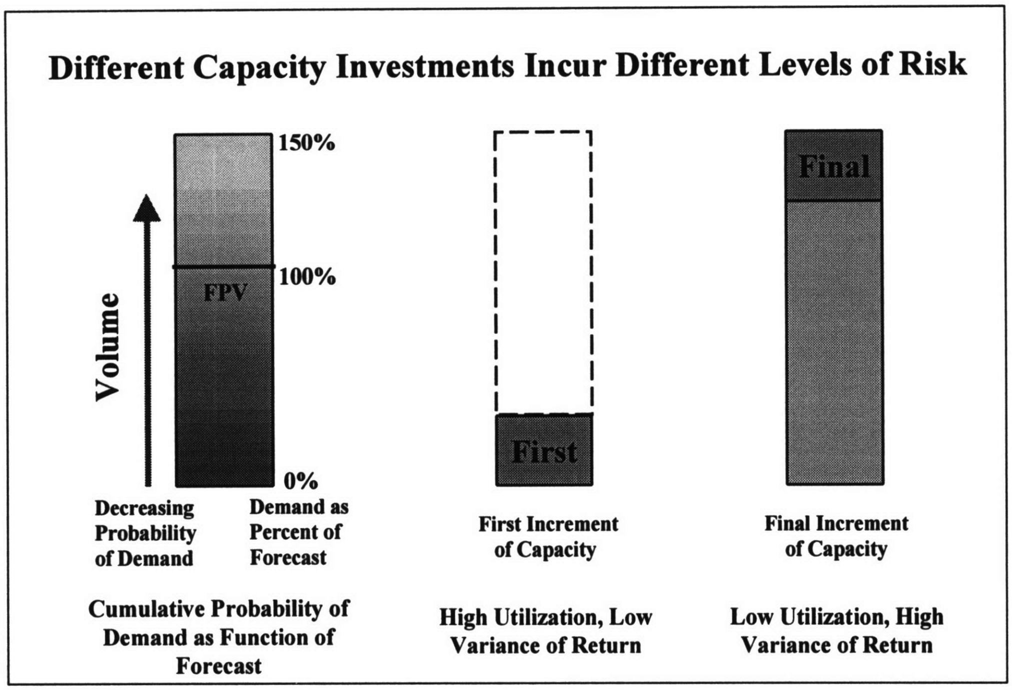


Figure 4-25: Relative Risk Exposure for Different Capacity Investments

Risk increases at higher levels of capacity because the actual capacity usage will substantially vary from the expected level of usage. For example, the final increment of capacity may have the expected usage of 25% although this may result in 100% usage in one in four years. On the contrary, the first increment of capacity will have an expected usage of nearly 100% and in practice actual usage will rarely deviate from this value. Since both investments are subject to the same demand and therefore the same risk correlation, the investment with the higher variance of return incurs the higher risk.

The Capacity Planning process used to calculate Capacity Protection Volumes maximizes the investment NPV to the company by increasing the capacity uplift until the incremental return is less than the corporate cost of capital. This process assumes the corporate level of risk for all investments even though optimal uplift levels will be different for different vehicle lines due to different investment costs and per unit profits per line. This process is shown below:

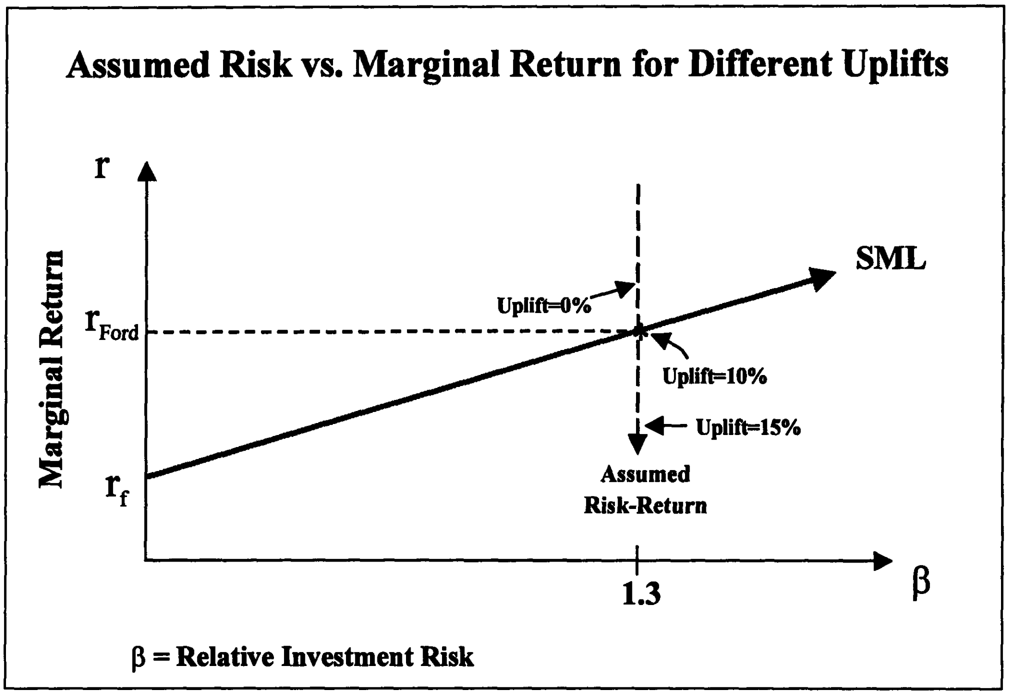


Figure 4-26: Assumed Risk vs. Marginal Return for Different Uplifts

While the assumption of corporate risk may be adequate for assessing the average return for a large increment discrete investment such as a new engine line, it is not adequate for calculating the capacity protection level for an entire vehicle line. While the correlation between variation in return for an incremental investment (due to demand fluctuation) and market variation is the same for all investment levels, the variance of investment return is clearly not the same. In fact, variance of marginal investment return was found to grow exponentially as uplift increased. This finding was obtained by developing a statistical model that calculated the expectation and variance of return by using the historical density distribution of demand relative to the expected value for an average-size vehicle segment. The model used numerical integration techniques over a range of possible uplifts to determine these values. Results are shown in Figure 4-27, below.

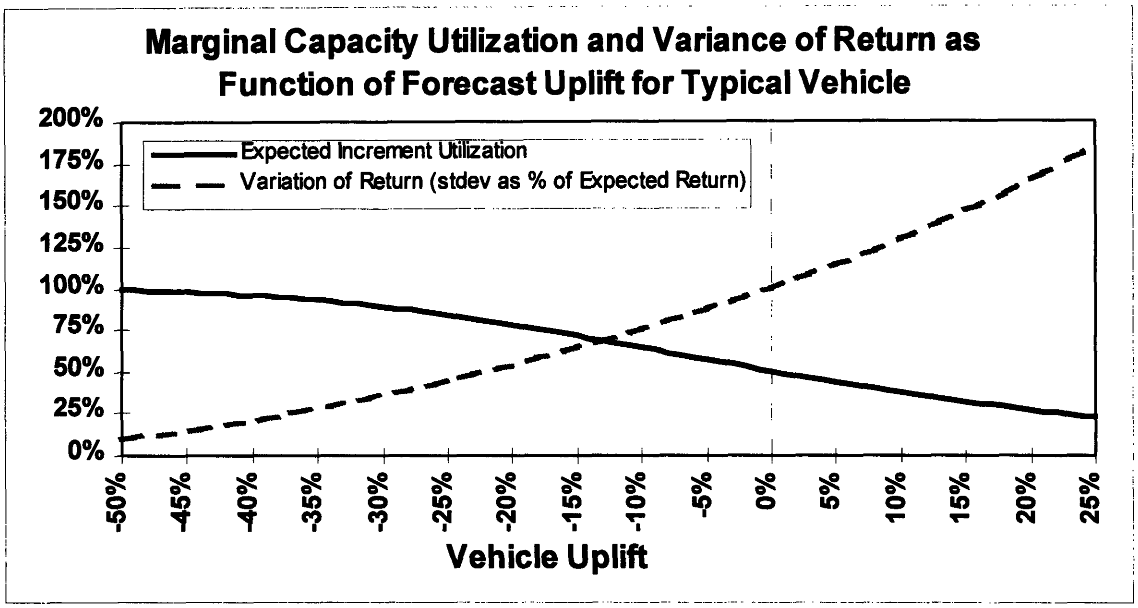


Figure 4-27: Marginal Capacity Utilization and Variance of Return as a Function of Uplift

As shown, the variance of return increases substantially as the uplift is increased. Given a constant correlation between changes in demand and market movements, increases in the variance of return will be directly proportional to increase in risk. Applying the findings of Figure 4-27 to Figure 4-26 yields the following result:

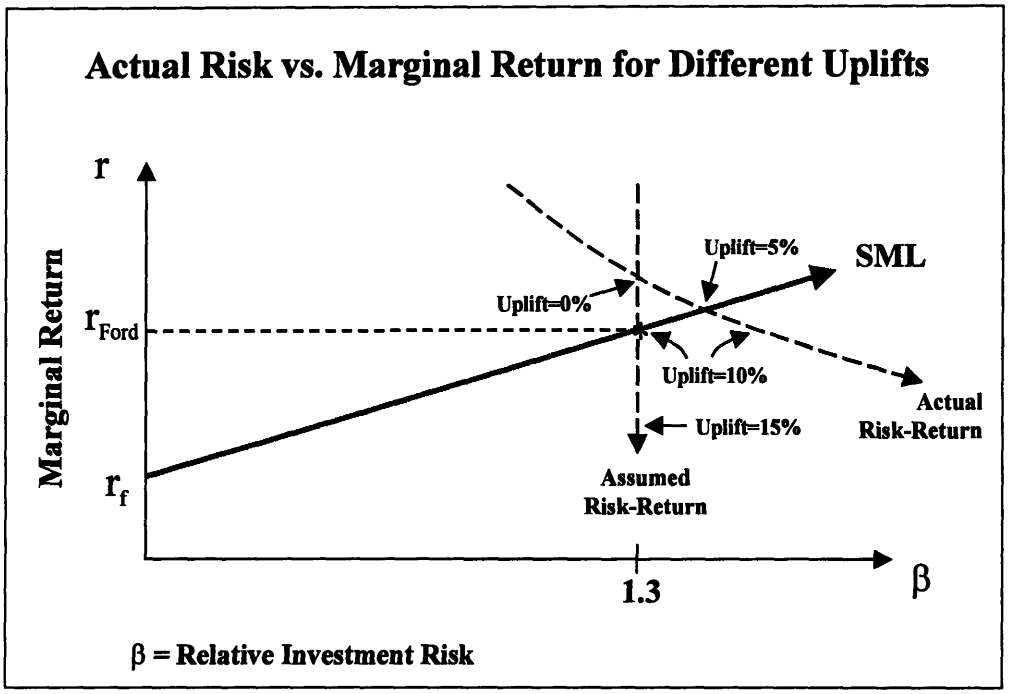
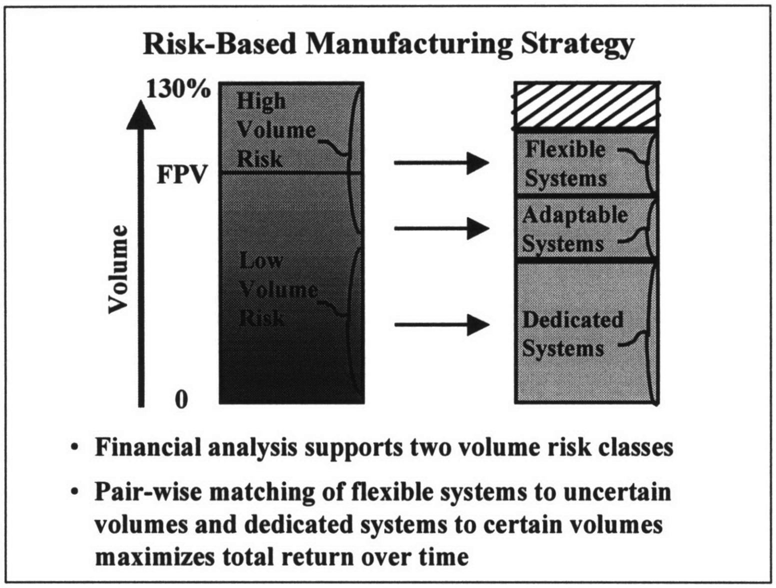


Figure 4-28: Actual Risk vs. Marginal Return for Different Uplifts

By ignoring the relationship between risk and marginal return as a function of volume, capacity protection volumes are higher than they ought to be, passing a disproportionate share of demand variation on to investors. Even though this variance does not affect the expected value of the investment return, it does increase the uncertainty and variance of the cash flow.

Two strategies stand out for decreasing this propagation of variance. The first is to decrease the vehicle uplift by accounting for the relative increase in risk as the uplift is increased. The second strategy is to adopt a risk-based manufacturing strategy that utilizes flexible and adaptable systems for the upper increments of investment volume. Manufacturing flexibility effectively pools the variance between products, reducing the sensitivity to changes in demand. This strategy is pictorially described in Figure 4-29.



Risk-Based Manufacturing Strategy

- **Financial analysis supports two volume risk classes**
- **Pair-wise matching of flexible systems to uncertain volumes and dedicated systems to certain volumes maximizes total return over time**

Figure 4-29: Risk-Based Manufacturing Strategy

This investment strategy is not actually the same for all vehicle lines. Differentiation should be based on the following categories:

- Correlation between forecast-demand deviation
- Relative life-cycle phase between products
- Industry segment size

The correlation between forecast-demand deviation is important since all capacity investments are made to a forecast and it is the demand movements relative to this forecast that cause shortages and under utilization. Having flexibility between products shifted in phase in terms of their life-cycle enables lower capacity investment levels without an increase in expected lost sales since the decrease in sales over time of the older product will be offset by the higher sales of the newer product.

Finally the industry segment size should be a factor since it was found that for the same level of uplift, smaller segments have higher demand variances. This observation was made by 1) taking the squared deviation of the natural log of the percent deviation of actual demand from forecast demand and plotting it against vehicle segment volume (industry level) for that forecast and 2) taking a nonlinear regression through the resulting set of squared deviations. A sample of this regression is shown in Figure 4-30. The analysis was performed separately for each forecast year (one year into the future, two years into the future, ...etc.), across all vehicle segments, and over the most recent fifteen years of forecasts within Ford.

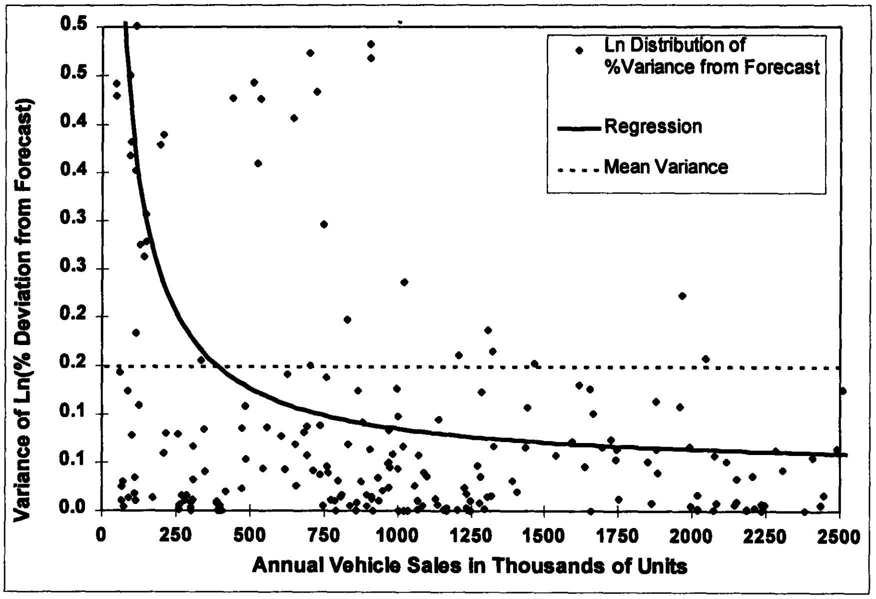


Figure 4-30: Regression of Demand-Forecast Deviation as Function of Segment Volume

Unfortunately, due to the complexity of the analysis (regression through a sample variance distribution) it was difficult to attain confidence intervals for the regression. However, the independent results for each forecast year were consistent with each other, giving credence to the evidence. See Figure 4-31. Furthermore, in total nearly 1000 years of forecast results were used in the analysis (fifteen segments * fifteen years of forecasts * four years of forecast horizon for each forecast), allaying concerns about sample size typically encountered when assessing long-range forecast accuracy.

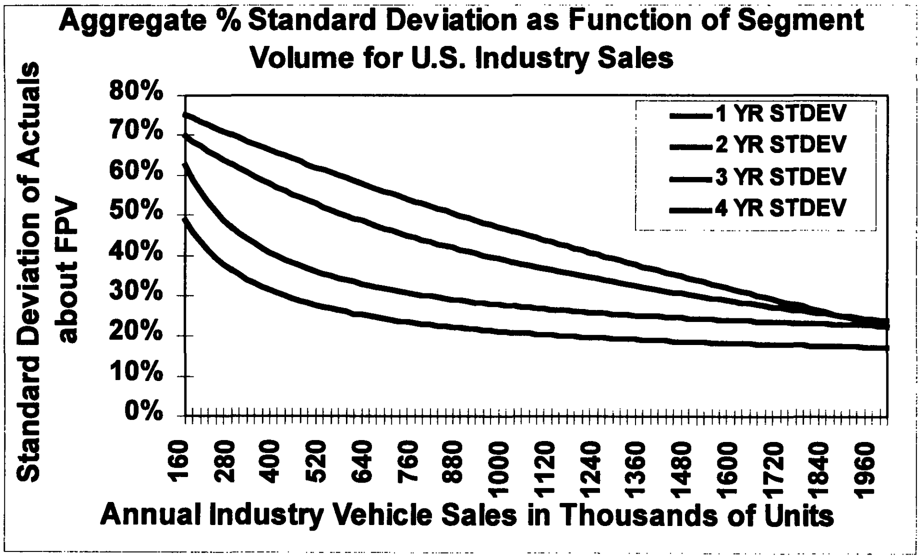


Figure 4-31: Forecast-Demand Variance by Year and Volume

The results of the analysis were plotted for each vehicle segment in the U.S. auto industry. This is shown in Table 4-1.

Forecast - Demand Variance as a Function of Vehicle Segment Size and Forecast Horizon						
Vehicle Segment	Segment Volume (units)	Assumed Standard Deviation	1YR Standard Deviation	2YR Standard Deviation	3YR Standard Deviation	4YR Standard Deviation
Luxury	1,000,000	20%	21%	28%	39%	46%
Full-Size Utility	200,000	23%	43%	56%	67%	73%
Compact Utility	1,600,000	20%	18%	24%	29%	31%
Full-Size Pickup	2,000,000	20%	17%	23%	24%	22%
Large	700,000	20%	24%	31%	46%	55%
Small Specialty	700,000	13%	24%	31%	46%	55%
Heavy Truck	300,000	23%	35%	45%	62%	69%
Compact Bus/Van	1,200,000	20%	20%	26%	35%	41%
Upper Middle	2,200,000	20%	17%	22%	22%	17%
Middle Specialty	300,000	20%	35%	45%	62%	69%
Full-Size Bus/Van	400,000	25%	31%	40%	57%	65%
Compact Pickup	1,000,000	15%	21%	28%	39%	46%
Lower Middle	1,900,000	13%	18%	23%	25%	25%
Basic Small	1,900,000	13%	18%	23%	25%	25%
Heavy Duty Wagon	140,000	–	53%	67%	71%	76%

Table 4-1: Forecast-Demand Variance as Function of Vehicle Segment Size

As the table shows, small vehicle segments tend to be quite volatile in terms of percent change. A number of reasons have been speculated for this volatility. One possibility is that movements in segment share of larger segments force greater change among smaller segments, e.g. a five percent change in volume of a large segment may result in a fifteen percent change to a smaller segment. Another possible reason is that smaller segments are more sensitive to competitive actions such as pricing incentives and new product offerings.

Irrespective of the causes of volatility, the data shows that product offerings among smaller segments will be subject to much greater volume risk. This means that investment in dedicated facilities may be a particularly risky proposition. The effect is compounded further when it is noted that smaller segments often yield smaller returns since the overall market is smaller, but the product development costs are nearly the same. As a result it is essential that the risk-based investment strategy described in Figure 4-29 be function of segment volume as well. The resulting relationship between industry segment size and risk-based investment strategy is shown below.

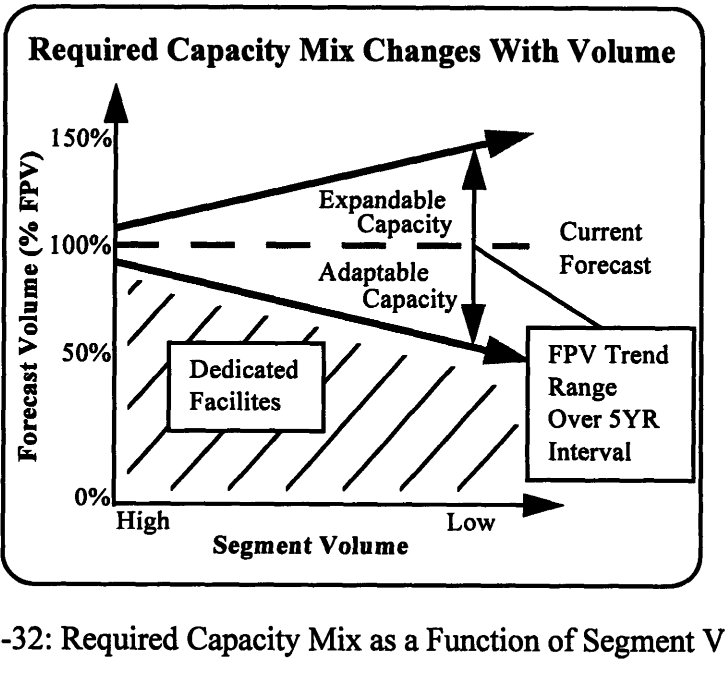


Figure 4-32: Required Capacity Mix as a Function of Segment Volume

If this risk-based investment strategy is pursued, variance to shareholders will be reduced without having to cut back on vehicle uplifts which would otherwise cut back on the company's market share in times of high demand and restrict its freedom in operational strategy.

4.6.2 Improved Capacity Planning Performance

Part of the results of the OOAD Interdependence study at Ford was the discovery of ways to improve the performance of the capacity planning function without adversely affecting other parts of the organization. These findings are expected to markedly improve the capacity planning process at Ford – a process that hasn't changed for nearly twenty years.

Assumption of Unalterable Investments

One of the most significant findings when this study was conducted was that the current state process used for calculating the capacity protection volumes (CPV) assumed that any investment decision was unalterable for the life of that investment. This would mean that no changes to capacity could be made once the original investment had been completed. The principal reason guiding this assumption was:

- Forecasts would never change, i.e. a forecast five years into the future would be identical to the forecast for that same year made only one year into the future
- Demand would not probabilistically deviate from the forecast any differently for far term forecast years as for near term forecast years

Given these assumptions, there would be no need to alter the capacity from values known a priori and differing levels of flexibility, adaptability, and expansion capability would have no effect on the expected revenues portion of the calculated incremental NPV equation. This situation is shown in Figure 4-33.

Capacity Planning & Product Development Interdependence Matrix		**Exogenous Variables**	EV_1 Demand	**Product Development CVs**	CV_1 # of Carryover Design Concepts	CV_2 Adaptability of Capacity Installed	CV_3 Flexibility of Capacity Installed	CV_4 Provision for Future Capacity Expansion	CV_5 Capacity Increment Size	**Capacity Planning CVs**	CV_6 Integrity of Capacity Planning Process	CV_7 Uplift Quantity
Product Development Objectives												
FR_1	Max Calculated Return on Sales, ROS		+		+	-	-	-	+		0	-
Capacity Planning Objectives												
FR_5	Min # of Capacity Related Stockouts, CSOs		+		0	-	-	-	+		+	-
FR_6	Max Incremental Investment NPV		0		+	-	-	-	+		0	X
FR_{61}	Investment Cost of NPV Calculation		0		-	+	+	+	-		0	+
FR_{62}	Return Portion of NPV Calculation		0		0	0	0	0	0		0	+
FR_7	Max Capital Asset Utilization		+		+	+	+	-	-		0	-

Figure 4-33: NPV Decomposition of Current State Interdependence Matrix

As shown, provisions for expansion only affect the investment cost side of the calculation. As long as there never was a need to alter the capacity invested, this method would be sound. This would only be true, however, if forecasts never changed and the deviation of demand from the forecast was constant over the forecast horizon.

<u>Why the Assumption of Perfect Forecasting Leads to the Build-up of Excess Capacity</u>

Unfortunately, neither of these assumptions were valid and the incremental NPV calculation was flawed, leading the manufacturing base to install a greater than optimal amount of capacity over the long run. This effect is illustrated below:

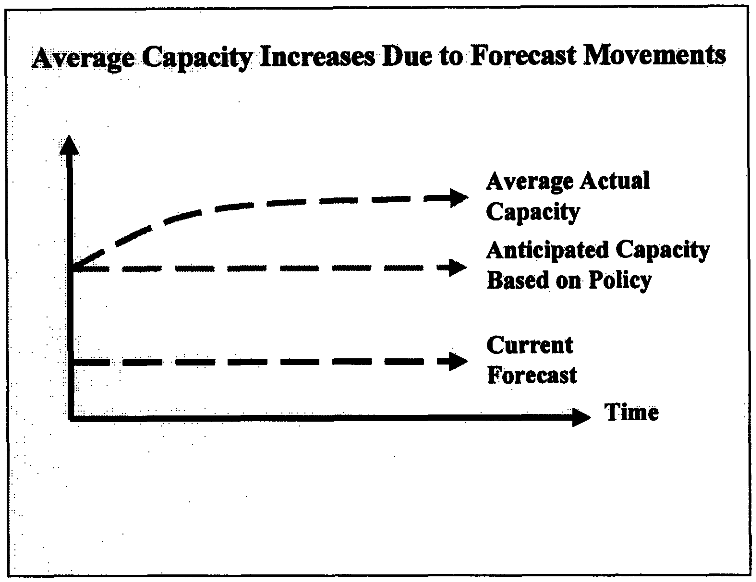


Figure 4-34: Average Capacity Increases Due to Forecast Movements

Of course, this finding raises the question of how it is possible for the company to invest more capacity than it intends. So why does average capacity increase over time?

- Forecasts move over time. Since capacity is a direct function of forecast, it must move as well.
- Capacity must move up when CPV moves up relative to where it was before – each part of the company must adhere to the capacity protection requirements at all times.
- Capacity seldom decreases when CPV moves down relative to where it was previously. This is due to the illiquidity and sunk cost of capital assets.

What is the consequence of the above behavior?

- On average the last unit of capacity in system is utilized less frequently than the CPV model predicts. This occurs simply because there is more capacity resulting from the model's behavior than the model assumes. Only if capacity were to move up and down with forecast movements would this not be the case.

Strategies to Improve Capacity Planning by Eliminating Excess Capital Investing

The combination of process behavior and above assumptions result in greater than desired capacity investment levels over time. Fortunately there are some straightforward ways that this sub-optimal behavior can be eliminated. The problem can be resolved by making the following changes:

- Assume actual investment decisions are made to a single forecast
- Assume utilization of the last unit of capacity originally invested to forecast will be expected to be utilized based on the statistical deviation of *demand from forecast*
- Assume options to expand capacity exist (with typical lead times) and are driven by changes in the forecast
- Assume options to abandon do not exist, i.e. all investments are sunk (valid assumption based on large capacity increments used – bottlenecks can be broken with larger units of capacity, but larger machines aren't replaced with smaller ones if demand falls)
- Use options theory to establish the initial capacity investment levels given the above assumptions

Resultant Changes to Ford by the Adoption of an Option-Based Capacity Uplift Strategy

Earlier statistical work has enabled forecast-demand variances to be readily applied (as seen in Figure 4-31) to the current process. Instead of demand volatility, forecast-demand variance would be substituted in the uplift calculation. The only other change is to incorporate the existence and value of options into the capacity planning process. This change will make the present value of actual investment decisions higher and total capacity invested lower.

Furthermore, this change will not adversely affect any of the objectives elsewhere at Ford. The interdependence matrix resulting from this change is shown below:

Capacity Planning & Product Development Interdependence Matrix		Exogenous Variables	EV_1 Demand	Product Development CVs	CV_1 # of Carryover Design Concepts	CV_2 Adaptability of Capacity Installed	CV_3 Flexibility of Capacity Installed	CV_4 Provision for Future Capacity Expansion	CV_5 Capacity Increment Size	Capacity Planning CVs	CV_6 Integrity of Capacity Planning Process	CV_7 Uplift Quantity
Product Development Objectives												
FR_1	Max Calculated Return on Sales, ROS		+		+	-	-	-	+		0	-
Capacity Planning Objectives												
FR_5	Min # of Capacity Related Stockouts, CSOs		+		0	-	-	-	+		+	-
FR_6	Max Incremental Investment NPV		0		+	-	-	-	+		0	X
FR_{61}	Investment Cost of NPV Calculation		0		-	+	+	+	-		0	+
FR_{62}	Return Portion of NPV Calculation		0		0	0	0	+	0		0	+
FR_7	Max Capital Asset Utilization		+		+	+	+	-	-		0	-

Figure 4-35: Resulting Interdependence Matrix: Recommended Changes to NPV

Implementing these changes will markedly affect the cost structure at Ford. Ford typically spends seven to eight billion dollars per year on capital investment as shown in Figure 4-5. Assuming that cost of capacity scales with size, even a reduction as small as one percent will yield an annual savings of as much as eighty million dollars. Since the CPVs are used for Ford world-wide and all of its Tier-One suppliers, changes to capacity uplift will have extremely high leverage.

4.6.2.1 Development of the Options Approach to Capital Investing

<u>The Black-Scholes Option Pricing Formula</u>

The basis for this change in the capacity planning process draws from the Black-Scholes [39] work on financial options. Black-Scholes developed an options pricing model that determines the present value of an option based on the future likelihood and value of exercising that option. This equation is shown below:

$$C = SN(x) - Ke^{-rt}N\left(x - \sigma\sqrt{t}\right) \tag{4–5}$$

where

$$x = \frac{\ln\left(S/K\right) + \left(r + .5\sigma^2\right)\cdot t}{\sigma\sqrt{t}} \tag{4–6}$$

In the option pricing formula,

C = Present Option Value

S = Present Asset Value of Option

$N()$ = Cumulative Standard Normal Distribution

K = Exercise Price of Option

r = Risk Free Discount Rate

t = Number of Years Until Option Expiration

σ = Standard Deviation of Asset Value Per Year (annualized standard deviation of the natural logarithms of the price relatives)

The Black-Scholes formula can be used for pricing any options-based asset provided

- Changes in the asset value over time are lognormally distributed

- The cumulative variance over time is the sum of the variance of independent random variables, i.e. $t\sigma^2$. Alternatively, it is assumed that there is no autocorrelation of price movements, that movements follow a random walk
- For every price movement, there are a continuum of outcomes

An Example Using Options Theory in Capital Investing

To put the notion of options in the context of capacity planning it is worthwhile to consider a simple example. This example applies options theory to a manufacturing capital investment.

The Scenario:

Today you must make an investment decision. So far, you have already decided it is worth investing in 100k units of capacity with a one year lead time and nine year life. What you must also consider, however, is a proposal that your chief manufacturing engineer presented you. She said that if you spend an additional $1M today, you will have the option of investing $50M next year, and gaining additional profits (through increased sales) that have a present expected value of $40M. The resultant expansion will increase total capacity to 125k units. If you do not spend the $1M now, she says that the same capacity increase will cost $80M next year due to a partial tear-down of the line and replacement of certain facilities which cannot support the increased line rate. This tear down is not assumed to affect current production or the installation lead time of new capacity compared to the optional expansion opportunity.

The Dilemma:

What should you do? Is it worth spending a certain $1M on an option that expects to net a loss of $10M? The only thing you know for sure is that the future is highly uncertain and as a result of historical forecast fluctuation, the standard deviation of net expected profits is 25%.

The Answer:

You should spend the $1M. The opportunity cost of not investing is $1.5M.

Justification:

While the current option asset value is $10M less than the investment cost, there is a chance that next year's forecast will increase, raising the expected revenues to an amount that creates a positive NPV for the option. If, on the other hand, the forecast remains the same or declines, the investment will be unfavorable and the $50M will not be invested. To determine the value of investing the $1M, the difference in the value of two mutually exclusive options must be evaluated. The first is the option that costs $1M with an exercise price of $50M and the second is one that costs nothing with and exercise price of $80M. Both have the same asset value. Using the Black-Scholes method, it is found that the $1M option has a present value of $1.57M and the free option has a present value of $22k. Thus, by spending the $1M, an additional $550k in profits can be expected over not doing so.

By realizing that the future option to expand was similar to a financial option, the up front investment could be financially justified. Without options theory, only certainty-equivalent investment and return values would be used. These would suggest a loss and the investment to provide future expansion capability would never be made.

Assumptions Necessary for the Adoption of Options in Capital Investing

The implementation of Black-Scholes option pricing formula into the capacity uplift process will be similar to the example just discussed. The options model in the capacity uplift process makes the following assumptions:

1. The strike (or exercise) price, *K*, is the present value of the total investment and carry cost of the optional expansion at the point in time the decision to expand is made.
2. The stock (or asset) value is the present value of the certainty equivalent revenues for the option were it exercised.
3. The one-period standard deviation is the annualized standard deviation of the natural logarithms of the forecast relatives.
4. If the option to expand will be exercised, it will be exercised within the first year of the initial investment decision.

In addition to these assumptions integral to the option pricing formula, the following additional assumptions are made:

5. The one year variance of forecast can be observed directly so that the requirement of serial independence of future forecast movements can be relaxed.
6. The standard deviation of the natural logarithm of relative forecast movements is normally distributed.

7. Expected revenue varies directly with forecast movements and therefore is identically distributed, i.e. it is a scalar multiple of expected demand (forecast volume).
8. The capacity investment and carry costs are identical to the initial investment on a per unit basis.
9. There is no additional cost to provide the option to expand.
10. The option lead time (the time between when the decision is made to exercise the option and the time it takes until it is available) is assumed to be identical to the initial investment lead time.

Discussion of Assumptions Necessary for the Application of Options to Capital Investing

The first assumption, the exercise price requires the system-wide average increment size to be known a priori. Given this, the exercise price is the certainty-equivalent present value of exercising the option at the time the option is exercised. For capacity planning, this exercise price is the present value of all investment and fixed costs for the life of the option asset. Since investment and fixed costs are relatively certain and insensitive to market movements, the risk free discount rate can be used.

The current asset value is the present value of the expected additional revenues that would be received each year for the life of the asset were the option exercised. The present value of this revenue stream is determined by discounting the expected cash flow for each year at the risk-based discount rate.

To determine the expected revenue each year from the exercised option, it is necessary to calculate the expected usage of the expansion. This is achieved by solving the following equation,

$$R_{O_i} = PPU \cdot \left(\Pr(D_i \leq C_1)\cdot 0 + \Pr(C_1 < D_i < C_2)\cdot\left(E\left[D_i|_{C_1<D_i<C_2}\right] - C_1\right) + \Pr(D_i \geq C_2)\cdot(C_2 - C_1)\right) \tag{4-7}$$

where R_{Oi} is the expected revenue from the option in year i and PPU is the profit per unit. D_i is the demand in year i and C_1 is the capacity protection volume and C_2 is the upper capacity limit on the expansion option. These variables are illustrated in Figure 4-36. Qualitatively, (4-7) is zero usage when demand is less than the base capacity protection, 100% usage when demand is greater than C_2, and a certain expected usage based on the probability density distribution over the range of option volume when the demand is within the volume limits of the expansion option. This comprises the three terms in (4-7).

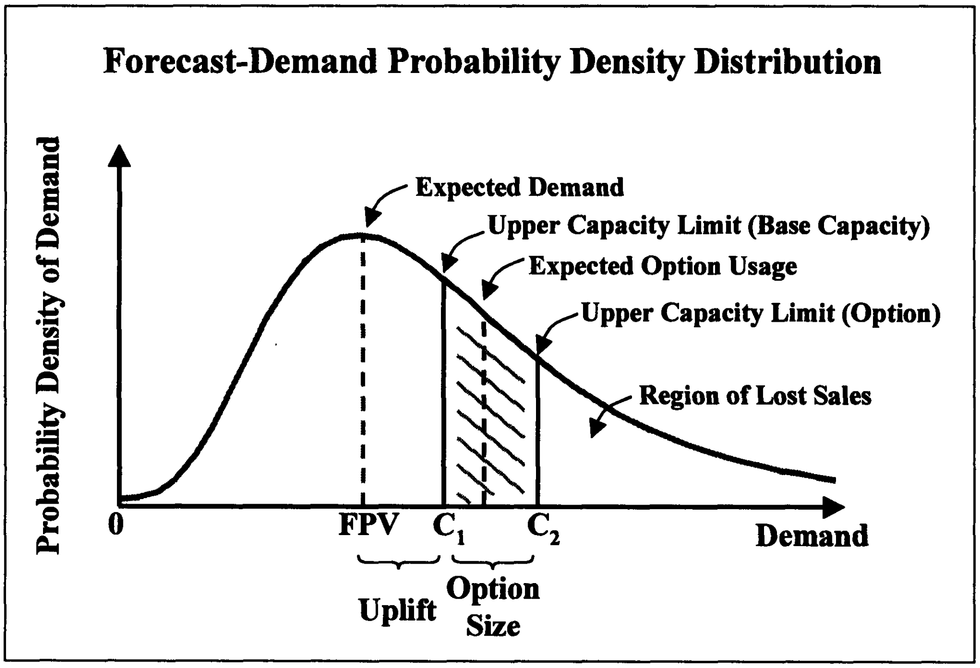


Figure 4-36: Single Year Forecast-Demand Probability Distribution

To determine the expected demand given $C_1<D_t<C_2$, it is necessary to integrate over the range from C_1 to C_2. This must performed numerically since there is no closed form solution to any part of the integral of the lognormal density function.

Assumption 3 is satisfied by a statistical analysis similar to the one performed for the forecast-demand variance in Section 4.6.1.3 as exemplified in Figure 4-31. The forecast movements are a function of both time and segment volume. This is shown in Figure 4-37, below.

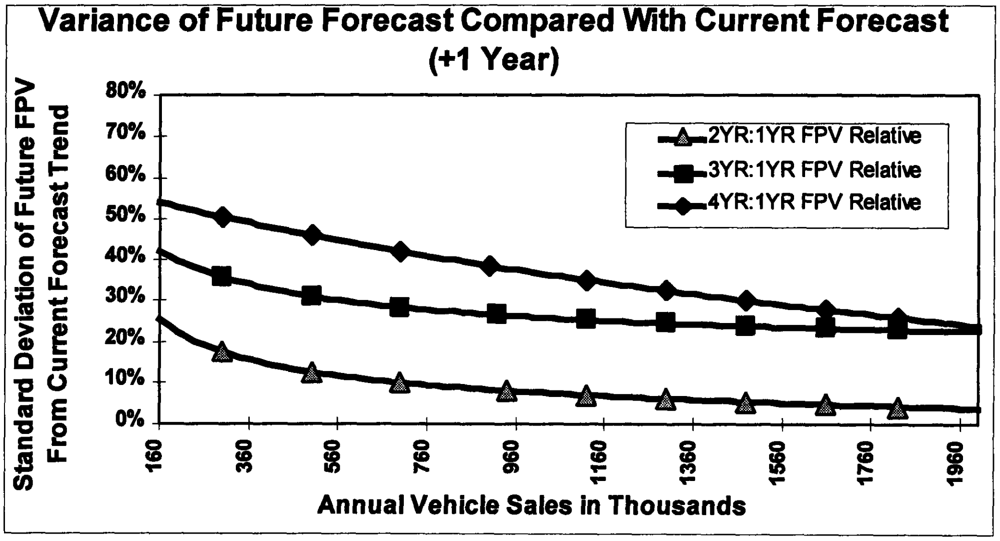


Figure 4-37: Forecast Movements as Function of Time and Vehicle Segment

The fourth assumption, that the option will be exercised within one year, is conservative. Due to the sales behavior of new product introductions and the long runs of growth or decline in vehicle segments, it is unlikely the option will be exercised in future years if the payoff in the first year is poor. Nonetheless, any option will increase in value if more time is allotted for its exercise. Because of this, the one year exercise assumption will be slightly conservative.

The fifth assumption eliminates the need for forecast movements to be serially independent. Since the decision whether or not to exercise the option is assumed to take place in one year, it is not necessary to use any other value than the one year forecast variance.

Assumption 6 requires that the relative forecast movements be lognormally distributed. This was found to be the case. See Figure 4-38, for example.

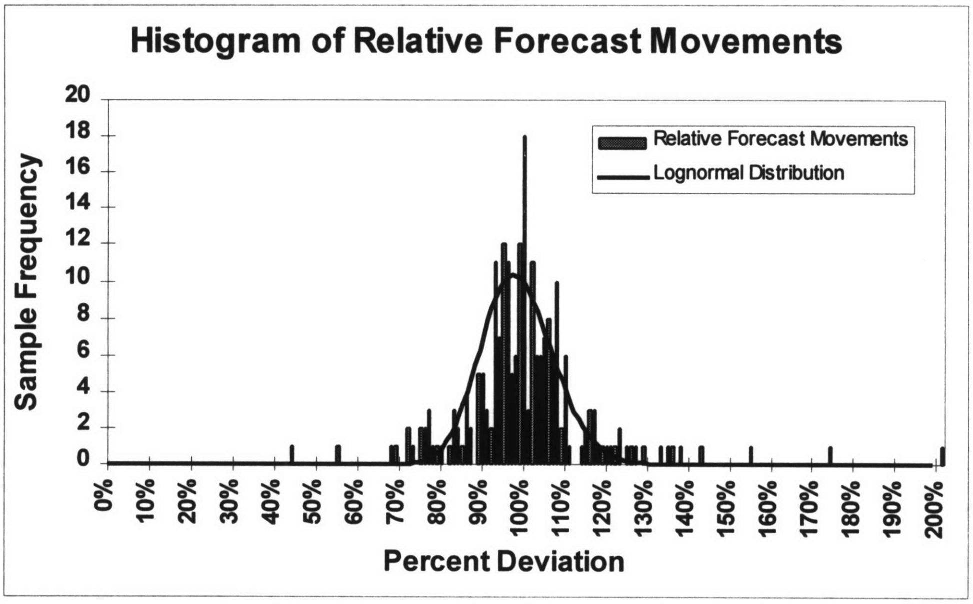


Figure 4-38: Basis for a Lognormal Distribution of Relative Forecast Movements

Assumptions 6 through 10 are only constants. The values are typical, but the company can easily adjust them if more representative values are found.

Results of Options-Based Incremental NPV Capacity Uplift Model

The use of options theory was fully integrated into a modified incremental NPV capacity uplift model. Applying the options theory substantially reduced the corporate uplifts across all vehicle lines. In fact, the results indicate that for all but the most profitable models, the vehicle uplifts can and should be completely eliminated.

These results indicate Ford could gain substantial savings by including the effects of options in the capacity planning process. However, because the model was not validated nor pilot tested in an actual investment scenario, no specific vehicle-by-vehicle recommendations will be made. What this work does demonstrate though, is that the use of options theory can substantially improve the capacity planning process at Ford without adversely affecting any other part of the company – in fact, objective performance for other parts of the company actually increase since total capital investment costs will be reduced by these changes.

4.6.3 Improved Return on Sales for Product Development

A significant issue in the continuing trend of decreasing flexibility in manufacturing is the fact that the calculated return on sales does not recognize the existence of any interdependence between provisions enabling manufacturing flexibility and the expected sales. This "independence" can be seen in the Current State Interdependence Matrix, Figure 4-39, when Return on Sales is decomposed one level.

Capacity Planning & Product Development Interdependence Matrix		Exogenous Variables	Product Development CVs					Capacity Planning CVs	
		EV_1 Demand	CV_1 # of Carryover Design Concepts	CV_2 Adaptability of Capacity Installed	CV_3 Flexibility of Capacity Installed	CV_4 Provision for Future Capacity Expansion	CV_5 Capacity Increment Size	CV_6 Integrity of Capacity Planning Process	CV_7 Uplift Quantity
Product Development Objectives									
FR_1	Max Calculated Return on Sales, ROS	+	+	-	-	-	+	0	-
FR_{11}	Sales Revenue	+	0	0	0	0	0	0	0
Capacity Planning Objectives									
FR_5	Min # of Capacity Related Stockouts, CSOs	+	0	-	-	-	+	+	-
FR_6	Max Incremental Investment NPV	0	+	-	-	-	+	0	X
FR_7	Max Capital Asset Utilization	+	+	+	+	-	-	0	-

Figure 4-39: ROS Decomposition of Current State Interdependence Matrix

As shown along FR_{11}, changes in either the type or volume of capacity are not expected to have any effect on the calculated ROS. The reason for this is that a "free demand" forecast value is used when the return on sales is calculated. However, a free demand sales forecast is clearly not the expected sales forecast. The very process used to calculate the uplift (percent capacity protection applied on top of the free demand forecast volume) expects to lose a certain percentage of sales due to insufficient capacity. By design this means that on average the actual sales will always be lower than the forecast sales due to imposed capacity constraints. Furthermore, as discussed in 4.5.3.2, changes to manufacturing flexibility such as reduced adaptability, provision for future expansion, etc. clearly do affect the quantity of lost sales.

The consequence of this current state model is to reduce flexible manufacturing practices below the value that would optimize the actual return on sales, yielding a higher calculated return on sales than what is actually realized. Noting this effect of flexibility, a modified sales forecast can be used in the ROS calculation. Equation (4-12) is one such example.

$$\overline{Sales_{ij}} = \sum_{k=1,k\neq j}^{p} \overline{Sales_{D_{ij>C}}} + \overline{Sales_{D_{ij<C},D_{ik<C}}} + \overline{Sales_{D_{ij<C},D_{ik>C}}} \tag{4-12}$$

This equation, detailed in Appendix Section 4.9.4, adjusts expected sales based on flexible capacity. While not shown in (4-12), the amount of flexible capacity is an integral part of the equation (a detailed derivation can be found on *Page 269*). Since the capacity volume is given to Product Development, the only variable it explicitly controls in this equation is flexibility. Although the equation only shows sales volume as a function of flexibility, extension to include asset adaptability, provision for expansion, and capacity increment size is expected to be straightforward. This is left as an exercise for the reader. Finally, implementation is not difficult, since all parameters have either been calculated or can be easily estimated.

Making this change to product development results in the following interdependence matrix:

Capacity Planning & Product Development Interdependence Matrix		Exogenous Variables	EV_1 Demand	Product Development CVs	CV_1 # of Carryover Design Concepts	CV_2 Adaptability of Capacity Installed	CV_3 Flexibility of Capacity Installed	CV_4 Provision for Future Capacity Expansion	CV_5 Capacity Increment Size	Capacity Planning CVs	CV_6 Integrity of Capacity Planning Process	CV_7 Uplift Quantity
Product Development Objectives												
FR_1	Max Calculated Return on Sales, ROS		+		+	-	-	-	+		0	X
FR_1	Sales Revenue		+		0	+	+	+	-		0	+
Capacity Planning Objectives												
FR_5	Min # of Capacity Related Stockouts, CSOs		+		0	-	-	-	+		+	-
FR_6	Max Incremental Investment NPV		0		+	-	-	-	+		0	X
FR_7	Max Capital Asset Utilization		+		+	+	+	-	-		0	-

Figure 4-40: Resulting Interdependence Matrix: Recommended Changes to ROS

This change achieves the following results:

- Makes calculated ROS a better predictor for actual ROS
- Puts a lower bound on the minimum desired flexible manufacturing practices by Product Development
- Increases the amount of manufacturing flexibility from current levels thus advantageously lowering CSOs and increasing utilizations
- Leads to an increase in actual ROS by causing decisions to be made that improve manufacturing performance

While not explicitly reducing the adverse coupling relationships to Capacity Planning, these changes will not introduce any additional couplings into the system. The net result is local improvement in Product Development that is also global improvement for the company.

4.6.4 Reduced Adverse Coupling Behavior Between Organizational Units

A core purpose of this interdependence study at Ford was to eliminate the coupling between two functional organizations within the company. While the previous two sections, 4.6.2 and 4.6.3, have presented approaches that achieve global system improvement through local change, this section proposes how global improvement can be achieved through the elimination of adverse coupling behavior between local groups.

The coupling interdependencies described in Section 4.5.3 are quite fundamental. Adverse couplings resulting from changes within Product Development to increase the calculated ROS are difficult to eliminate. A substantial delay will exist between the time a change is made and the time the coupling effect is felt. This delay will be several years, in most cases. Having such a long delay between cause and (side) effect creates two problems. First, the managers who originally made the decision, the effect of which has just become apparent, have long since moved to another position within the company. Secondly, the Capacity Planning organization is forced to take a reactive stance, since the effect of reduced response capability only manifests itself when demand threatens to exceed installed capability.

Both problems can be eliminated if the key decisions regarding the type of capacity installed (flexibility, adaptability, provision for expansion, capacity increment size) can be moved upstream of both the Capacity Planning and Product Development processes. Such a change moves these decisions, typically made by individual product program teams, into the domain of the product strategy group (PSO) within Ford. At this point in the development process, both product and capacity issues are considered concurrently. This upstream process is considered to be a window of strategic development in capacity planning. It is shown schematically in Figure 4-41 below.

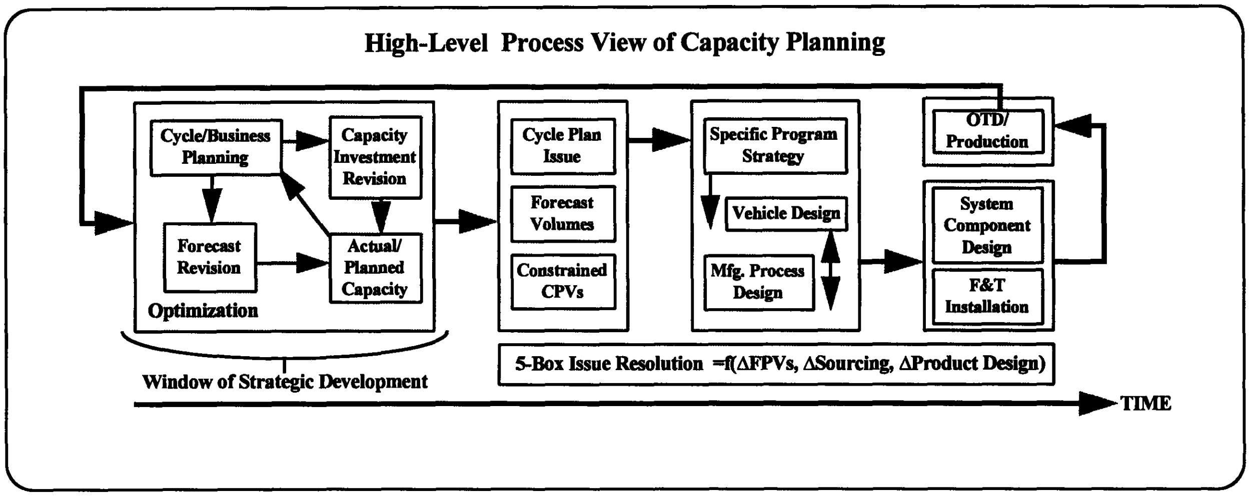
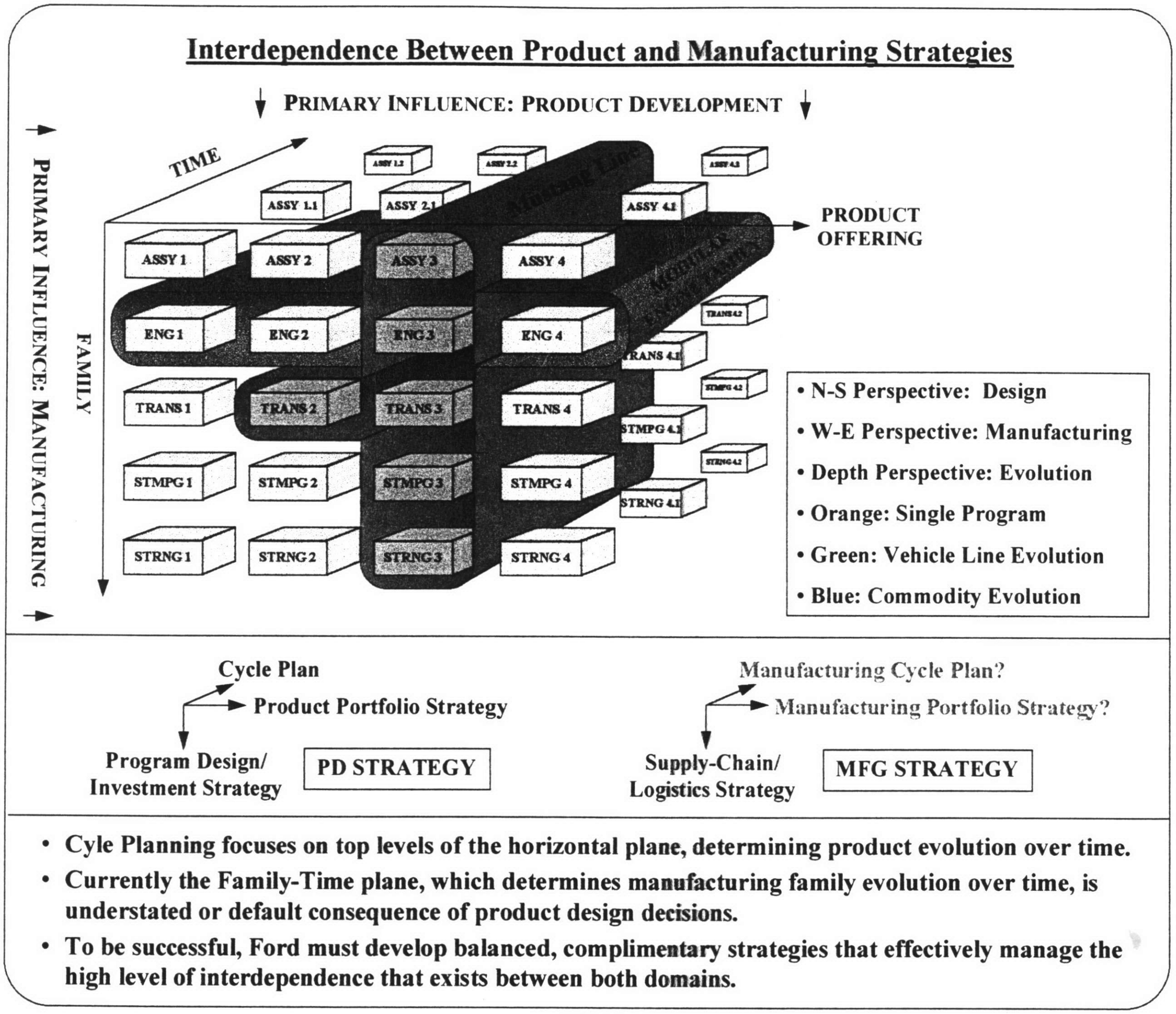


Figure 4-41: High Level Process View of Capacity Planning

By jointly considering product and capacity strategies, the previous coupling behavior will no longer occur. While the couplings will still exist, i.e. changing the amount of manufacturing flexibility will still affect Capacity Planning's objectives, there will no longer be the incentive for one group to optimize its performance at the expense of another.

This recommended change to the decision-making process effectively aggregates control, as described in Chapter 2, Section 2.3.3. As a technique, aggregation of control is effective, but requires greater coordination than other approaches. Nevertheless, compared to other methods considered (see Section 4.6.6 for a review of the other techniques evaluated), aggregation of control was the best alternative.

Figure 4-42 provides a conceptual framework for a joint manufacturing and product strategy.



Interdependence Between Product and Manufacturing Strategies

- N-S Perspective: Design
- W-E Perspective: Manufacturing
- Depth Perspective: Evolution
- Orange: Single Program
- Green: Vehicle Line Evolution
- Blue: Commodity Evolution

- Cyle Planning focuses on top levels of the horizontal plane, determining product evolution over time.
- Currently the Family-Time plane, which determines manufacturing family evolution over time, is understated or default consequence of product design decisions.
- To be successful, Ford must develop balanced, complimentary strategies that effectively manage the high level of interdependence that exists between both domains.

Figure 4-42: Joint PD-MFG Cycle Plan Strategy

Currently, the Product Strategy Office (PSO) issues a cycle plan that only considers the evolution of product design over time. While expected capacities (in terms of volume) were considered in the forecast plan's development, the *type* of capacity invested is considered a default consequence. The proposed structure results in a cycle plan that is the intersection of product (Mustang) and manufacturing (engine facilities) families. This can be seen in the figure. Producing a joint plan requires alignment between product development and manufacturing.

This makes capital investment decisions an integral part of the future forecast strategy, rather than a passive consequence of product development decisions.

4.6.5 Application of Additional Methods that Hedge Against Variation

The development of a joint PD-MFG cycle plan enables portfolio management of the manufacturing investment base. If a particular commodity, such as transmission assembly facilities, are managed on a portfolio basis, specific policies such as chained manufacturing flexibility and replicable cellular manufacturing begin to make sense. The key attribute of these approaches is that they provide a hedge against variation, either real (as in actual market movements) or illusive (as in forecast uncertainty). The joint cycle plan may call for a specific percentage of the manufacturing install base to be comprised of facilities providing this capability.

The following sections summarize each hedging method, as realized in actual manufacturing facilities. These methods may either be applied on a specific basis, or incorporated into an overall corporate manufacturing capital investment policy.

4.6.5.1 Unbalanced Capacity Investment Strategy

The unbalanced capacity investment strategy provides the capability to support higher future volume with minimal excess initial investment. Specifically, this dimension supports an unbalanced capacity investment strategy with surplus capacity designed into the system in cases where expansion reuse is limited, modular expansion capability does not exist, or where long-lead investments are required.

This strategy recognizes what will become the "unbreakable" bottlenecks in the event of a proposed capacity increase and incorporates a solution into the system during the manufacturing system design phase (Figure 4-43). These select investments will be initially capacitized to an amount in excess of the current rated capacity based on the provision for future possible expansion even though current forecast demand cannot justify such a high capacity. The incremental investment cost incurred by establishing an unbalanced line will be more than offset by the value of the option to expand in the future at substantially reduced expense. Furthermore, since future expansion is a viable option, risk of lost sales is substantially reduced over the "non-option" investment strategy. If implemented on a total supply-chain scale, this strategy has the direct consequence of reducing protection-level volume requirements for all initial investments, thus resulting in an across-the-board cut in initial capacity expenditures that exceeds the cost premium of line imbalance. The result: lower investment cash outlay with no increase in sales risk.

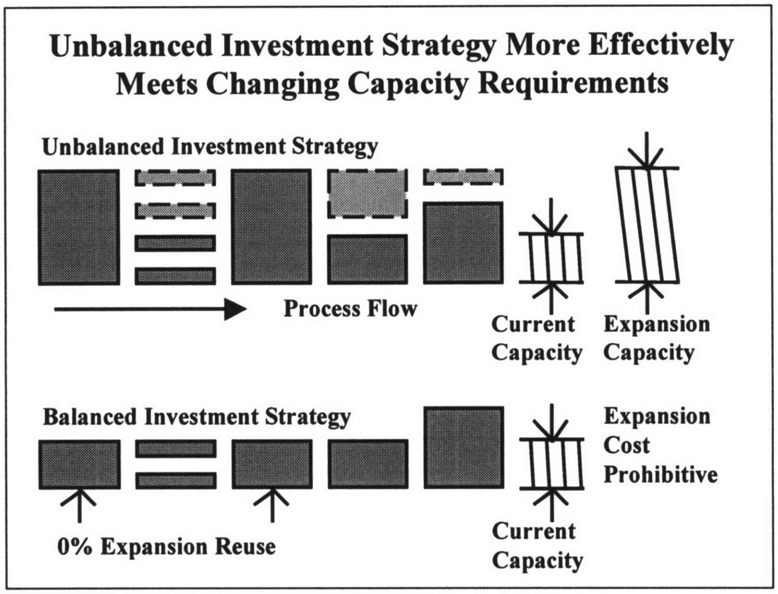


Figure 4-43: Unbalanced Capacity Investment Strategy

4.6.5.2 Modular Investment Strategy

The modular investment strategy enables rapid expansion capability by allowing small-increment capacity additions that are replications of the current system design. Since the original capacity is retained upon expansion, one hundred percent reuse is achieved. This is illustrated in Figure 4-44.

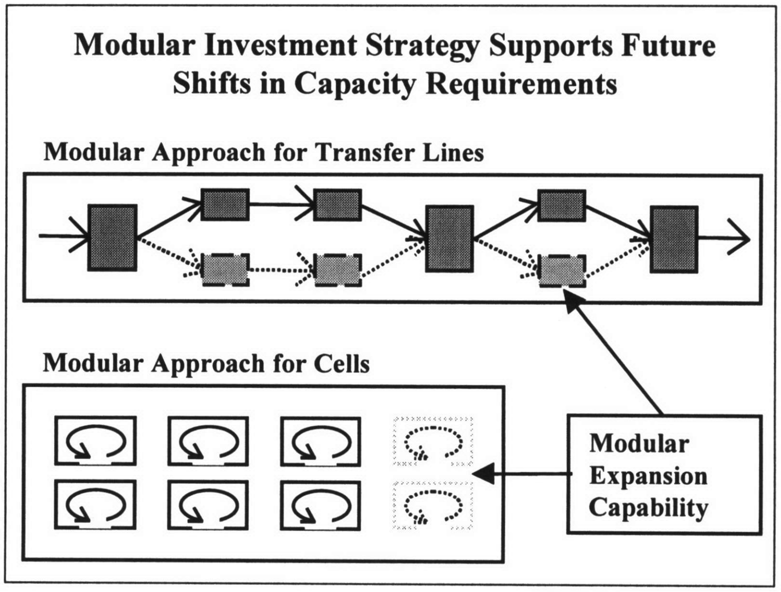


Figure 4-44: Modular Investment Strategy

Like the unbalanced capacity investment strategy, this approach reduces initial volume investment requirements in an amount exceeding the cost premium of the provision for optional expansion. Small-increment modular approaches, particularly cellular manufacturing, serve as key enablers for adaptable manufacturing systems designs.

4.6.5.3 Adaptable Manufacturing System Design

The adaptable manufacturing system design allows partial conversion of a manufacturing system from the production of one product to the production of another product. The approach is as much a function of machine design as it is a function of systems design. To be adaptable, a particular capital investment must be capable of being changed from the intended purpose of producing one type of product to another type of product in such a manner that an acceptable ROA is achieved. Commonly this approach is heralded under a reusability strategy wherein a given capital asset is adapted (or used without modification) and integrated into the design of a replacement product. Often however, a reusability strategy is reduced to a reuse effort since it is difficult to assess the future needs of a replacement product a priori, at the time the original capital investment is made. As a result, future design "degrees of freedom" may be restricted in order to accommodate the reuse strategy.

This difficulty does not exist when an adaptable manufacturing system is considered. To design a manufacturing system to be adaptable, the system must be made capable of being converted from the production of one product type to another while both products concurrently remain in production (though not by the same piece of equipment). Since both product types are known before the initial investment is made, it is far easier to develop a facilities and tooling design that can accommodate either product, than it would be if the second product had not yet been designed.

4.6.5.4 Flexible Manufacturing System Design

The flexible manufacturing system design establishes a cross-loadable or chained manufacturing system whereby the sum of demand for two or more products is known to be relatively stable yet the individual demand of a single product may vary unpredictably. In this system, the product changeover must support small-lot production.

Flexible manufacturing has numerous benefits such as production smoothing and high volume economies of scale for low volume products. Furthermore, when capacity becomes the production constraint in a flexible system, the producer can hedge losses by only “shorting” its least profitable product. Another benefit often cited is the higher utilization and fewer lost sales are achieved when negatively correlated demand streams are pooled together via a single manufacturing process. Unfortunately for the durable goods producers such as the auto industry and other seasonal industries, product demand is often positively correlated, thus making the previous claim hard to justify. However, what is not often cited is the effect of forecast uncertainty. While production forecasts are fairly predictable at a total industry or total corporate level, they become highly uncertain at a segment or vehicle line level, particularly so beyond the first couple years. Auto industry vehicle line forecasts are almost entirely negatively correlated to each other since any forecast error in one product must produce the opposite error in one or more products if the total forecast is to remain relatively accurate (see Sections 4.6.1.1 and 4.6.1.2 for a related discussion). Since capacity investments are made on a forecast basis and the goal of capacity planning is to simultaneously minimize investment costs and the costs of lost sales, substantial reductions in capacity volume can be made without increasing the number of lost sales when flexible manufacturing is used.

4.6.5.5 Variable Volume Fixed Asset Investment Strategy

The variable volume fixed asset investment strategy encourages efficient operations at different production rates given sunk investment costs. In this method, production volume and variable costs are designed to vary linearly.

The basis of this strategy stems from the fact that in an uncertain and volatile environment, the manufacturing plant will seldom run at its design point or point of maximum efficiency. In such a situation it may not be possible to change the value of the underlying assets, but it may be possible to change the variable costs. This strategy seeks to achieve operating costs that are 150% of the expected cost when production volume is at 150% of its expected value. Likewise, if production is down 50%, operating costs are down 50%. This strategy is particularly important for long-lead dedicated manufacturing systems such as engines and transmissions.

Since engine lead times are on the order of 36 months or more, the actual production volume may be substantially different from what was originally forecast. With labor lead times of only a few months and demand often being relatively stable once production begins, labor contracts are generally not a problem. However, how efficient is the manufacturing facility, if after rebalancing, still must employ 80% of the workforce to produce 50% of the volume? System designs with fewer individual process steps and team-based crews or cells where production manning can be easily changed support this strategy.

4.6.6 Other Techniques Evaluated

This section describes the application of techniques developed in Section 2.3.3 of Chapter 2, which were not selected as recommended approaches for Ford to reduce coupling interdependencies. The purpose of this section is twofold. First, it explains why these techniques were not recommended to Ford, and second, how they might have worked in a different context. It is hoped that these explanations promote the understanding of Object-Oriented Axiomatic Design to different situations.

Technique 1: Select or Add Different Controls to Achieve Local Improvement and Reduced Coupling

It is quite difficult to add or select different controls for this system. Uplift, flexibility, adaptability, and provision for expansion are all quite fundamental.

It may be possible to reduce coupling by simply eliminating flexibility, adaptability, and provisions for future expansion as control variables to improve ROS. Instead, other controls such as the styling or design concept selection could be used to meet the objective. Since these controls are available to the product planning community, they are less likely to couple with objectives of other groups. The downside is that these controls may exert less leverage on the system and be more difficult to implement. Additionally, the decisions about control variables affecting manufacturing flexibility must still be made, and whether they are made by Product Development or some other organization, they will continue to affect the calculated ROS. For these reasons, this is not the most viable solution.

Technique 3: Select Different Objectives to Reduce Coupling Strength

Return on sales, the number of capacity-related stockouts, asset utilizations, and investment return are all fundamental. This makes it difficult to select different objectives. It would be interesting, however, to consider changing the CSO objective to a capacity fill-rate measure, instead. This fill-rate objective would determine the fraction of orders lost, rather than just the intervals during which free demand could not be met. Unfortunately, implementation of this objective would be difficult, and the amount of coupling reduced remains unclear.

Technique 4: Expand the Sphere of Influence or Reallocate Objectives and Controls to Form Different Groupings

Expanding the sphere of influence of either Capacity Planning or Product Development will cause functional strength to be lost. The current partitioning seems appropriate as each group is focused on a mission critical aspect of the company. With respect to reallocation of objectives and controls, the greatest improvement will be achieved by allocating the controls that determine the type of capacity to the Capacity Planning organization. This change orients capital investing to support volume fluctuation at the possible expense of product design. Furthermore, the current organizational infrastructure provides no support for such a change.

Technique 5: Passively Buffer to Reduce Coupling Strength

Changes to flexibility, adaptability, provision for expansion can be used to passively buffer against changes in demand. Passive buffering can be used when a coupling interdependency is a function of other state variables. The reader is referred to the discussion of passive buffering in Section 2.3.3 in Chapter 2 for a more detailed explanation.

The coupling interdependencies present at the interface between Product Design and Capacity Planning offers two opportunities where passive buffering can be used to reduce the coupling strength of off-axis terms. The first opportunity is to reduce the coupling effect caused by changes in Product Development's control variables. The degree to which changes in flexibility, adaptability, provision for future expansion, capacity increment size affect lost sales and utilizations can be reduced by increasing capacity uplift. Unfortunately, capacity uplift is not a free variable that can be used to buffer Capacity Planning's objectives from Product Development's changes. Changes in uplift critically affect calculated investment NPVs.

The second opportunity for passive buffering lies in the ability to reduce the effect of swings in demand upon CSOs and capital asset utilizations. Currently, increases in demand will adversely affect the number of CSOs and decreases in demand will adversely affect capital asset utilizations. However, the extent to which swings in demand affect these objectives is a function of the flexibility and adaptability of the manufacturing system. This is apparent by referring to Equations (4-10) and (4-11) in the Appendix, Section 4.9. If the partial derivative of these equations is taken with respect to demand, the resulting term (the coupling strength due to changes in demand) will be a function of flexibility and adaptability. Therefore, if flexibility and adaptability are increased, the manufacturing system will be less sensitive to swings in demand. This intuitive result supports the use of passive buffering to improve system performance.

The reason why these controls are not used to passively buffer the system is that, like changes to capacity uplift, these changes are not free. For this reason, the current Product Development process uses these variables to cut costs. Therefore, since these control variables couple with other parts of the system focused on cost, it is not recommended that they be used apart from a cost-optimization approach.

Technique 8: Establish Parity-Based Transactional Interdependence to Prevent Coupling Behavior

A parity-based transactional interdependence model cannot be developed to prevent coupling behavior between Product Development and Capacity Planning. The reason for this is simple: Capacity Planning has everything to gain from changes to the way Product Development makes decisions, but has nothing to offer Product Development in return. In other words, Capacity Planning has no "money" by which to enter this transaction, and therefore no buying power with Product Development.

4.7 Summary of Recommendations

The Object-Oriented Axiomatic Design interdependence study conducted at Ford has resulted in numerous insights and opportunities for improvement. From this study, a series of four recommendations were presented to Ford. These steps would enhance established business practices by rationalizing the interdependence between the quantity and type of manufacturing capacity. Taken in concert, these actions improve the match between capacity and demand by creating a more robust and responsive manufacturing system.

The recommendations are as follows:

- Develop manufacturing investment strategy policy
- Issue forecasts in form of volume ranges
- Update CPV uplift process to reflect forecast uncertainty, future options
- Modify investment financial analysis: account for risk, revise labor and overhead

A detailed explanation of each is given in the following sections, giving a description of the proposal as well as the rationale behind the recommendation.

<u>Develop Manufacturing Investment Strategy Policy</u>

Description:

Enterprise-level manufacturing strategy as integrated part of Cycle Plan

- Creates manufacturing portfolio strategy paralleling the product portfolio strategy

- Establishes system-wide commodity requirements by vehicle segment for flexibility/adaptability, e.g. 70% dedicated, 15% adaptable, 15% flexible

Program-level manufacturing investment strategy policy

- Supports manufacturing cycle plan requirements
- Details the five key dimensions of manufacturing system design that enable enterprise-level requirements to be met (see Section 4.6.5)
- Establishes rigid requirements for commonization of design variables along key manufacturing dimensions, e.g. common locators, standardized process sequence, modular engine architecture

Rationale:

- No current strategy
- Need for manufacturing strategy development concurrent with product strategy
- Current cycle planning is primarily product driven with manufacturing as a tactical consequence or passive input constraint
- Proliferation of locally-optimized, near-sighted investment decisions
- Current funding allocation process enables independent decision making on interdependent subjects
- Need for more proactive stance on capacity planning. Reward manufacturing system designs designed to respond to changes in future demand requirements
- Flexible/Adaptable manufacturing systems, as evidenced by JIT and agile systems, require a total system or portfolio perspective to be successful

Issue Forecasts in Form of Volume Ranges

Description:

- Volume range is statistical variance by vehicle segment applied to forecasts as upper and lower "control limits"
- Volume ranges provide internal and external capacity planners confidence intervals for short and long-term capacity and production requirements
- Volume range provides wider intervals for far-term forecast years and vehicle segments subject to greater forecast-demand variance

Rationale:

- Single point forecasts with varying levels of accuracy convey limited useful information
- Ford forecast variance at product level is substantial; even more so for vehicle options
- Current business process assumes constant, accurate information that in practice is highly uncertain and variable
- Flexible/Adaptable manufacturing systems cannot be justified with a business process that a) assumes all forecast projections to be deterministic and b) has no knowledge of the extent of variation that must be supported over the life of the manufacturing asset

Update CPV Uplift Process to Reflect Forecast Uncertainty, Future Options

Description:

- Updated CPV is vehicle-level uplift which includes forecast uncertainty, differential investment lead times, and the effect of investment options in addition to vehicle profitability, invest & carry costs, and historic demand volatility

- Updated CPV provides a capacity protection level that increases in forward planning years. This recognizes that capacity requirements for those future years *may* become higher in the future. A constituent part of this approach is the requirement that higher future CPVs must be supportable. However, current (updated) CPVs do not financially justify today's capacitization to the higher future requirements. To satisfy these two seemingly divergent requirements, the manufacturing system must be adaptable

Rationale:

- Relaxes assumption that all forecasts have 100% accuracy for all forecast years
- Creates tension between current maximum affordable protection levels and higher future protection requirements. This promotes the design of flexible/expandable/adaptable manufacturing systems
- Provides higher protection requirements for long-lead investments exposed to greater volume variation
- Reduces system-wide uplift levels by relaxing assumption that any future changes to demand during the life of an asset will result in irrecoverable lost sales or low utilizations

Modify Investment Financial Analysis: Account for Risk, Revise Labor & Overhead

Description:

- Method recognizes and differentiates between low risk and high risk capital investments

- Method recognizes and differentiates between fixed and variable manufacturing assets[4]
- Method considers labor and overhead as function of total life-cycle cost
- Evaluates flexible/adaptable investments subject to volume risk using validated financial methods for investment options valuation and diversification through portfolio asset management
- Evaluates investment proposals under a range of probable outcomes rather than by the proposal yielding the highest expected return at the expected production volume

Rationale:

- Current financial models, generating maximum returns when manufacturing investment cost per unit is minimized, drive the manufacturing system to install only high-volume, fully dedicated, unalterable capital assets. In light of change and uncertainty, this undermines investment efficiency
- Financial system must recognize risk and be capable of valuing options if Ford is to financially justify flexible/adaptable systems

Development and Deployment of Proposed Actions

While much of the background work has been performed on each of the aforementioned topics, they are neither complete nor have been independently validated. To ensure broad acceptance and success of these efforts, the following approach is requested:

---

[4] Fixed assets are those dedicated to a particular use and have little abandonment value. Variable assets are assets having greater liquidity such as flexible or adaptable systems which can be easily converted to another use if demand requirements change.

- Secure sponsorship from senior management
- Charter cross-functional team with representation from Manufacturing, MBO, PSO, PD, Finance, Supply Base, and Process Leadership for facilitization
- Develop business case for financial justification and project plan for timing
- Identify and secure pilot program(s) for prove-out
- Codify best practices for uniform enterprise-wide deployment

This approach minimizes risk, facilitates development, and creates equity share for all impacted organizations.

4.8 Concluding Remarks

This case study with Ford Motor Company has demonstrated how Object-Oriented Axiomatic Design can be used to characterize the coupling interdependencies present at the interface between two internal, autonomous organizations. Coupling interdependencies identified through the use of influence diagrams and bounded interdependence matrices have provided the basis for a set of recommendations which improve overall corporate performance. Furthermore, the diagrammatical techniques employed can be used as a tool to galvanize change within the firm. Despite the depth of analysis performed at Ford, the nature of problems confronting senior management are complex. As the work progressed, it was found that even seasoned employees had very different mental models of the underlying system structure within the company. It would be worthwhile to extend the application of OOAD to these situations, in order to promote organizational alignment, and also to further develop this methodology as a valuable tool for process improvement.

4.9 Appendix: Derivation of Mathematical Relationships

4.9.1 Derivation of Computed Return on Sales

Return on Sales, ROS is the quotient of Profit Before Tax, PBT and Revenues, R.

$$ROS = \frac{PBT}{R} \tag{4-8}$$

Within product programs, an estimated ROS is used as a performance measure. This computed ROS can be expanded as follows:

$$ROS_C = \frac{\sum_{i=1}^{\overline{L}} \overline{S_i}\left(\overline{PPU_i} - \overline{VCPU_i} - \overline{VMPU_i}\right) - I}{\sum_{i=1}^{\overline{L}} \overline{S_i} \cdot \overline{PPU_i}} = 1 - \frac{\sum_{i=1}^{\overline{L}} \overline{S_i}\left(\overline{VCPU_i} + \overline{VMPU_i}\right)}{\sum_{i=1}^{\overline{L}} \overline{S_i} \cdot \overline{PPU_i}} - \frac{FTLE + \sum_{i=1}^{\overline{L}} \overline{OH_i} + Other}{\sum_{i=1}^{\overline{L}} \overline{S_i} \cdot \overline{PPU_i}}$$

(4-9): Computed Return On Sales

where

$\overline{S_i}$ = expected number of Sales in year ***i***

$\overline{L}$ = expected number of years of life of product

$\overline{PPU_i}$ = expected Price Per Unit in year ***i***

$\overline{VCPU_i}$ = expected Variable Cost Per Unit in year ***i***

$\overline{VMPU_i}$ = expected Variable Marketing cost Per Unit in year ***i***

I = total Investment cost

$FTLE$ = investment cost of Facilities, Tooling, Launch, and Engineering

$\overline{OH_i}$ = expected Overhead costs in year ***i***

$Other$ = Other non-recurring costs

4.9.2 Derivation of Capacity Related Stock Outs, CSOs:

$$\overline{CSO} = \sum_{j=1}^{p} \frac{\sum_{i=1}^{\overline{L}} P\left(D_{ij} > \overline{Effective\ Capacity_{ij}}\right)}{\overline{L}} + \sum_{j=1}^{p} P\left(Data\ Integrity\ Error_j\right) \tag{4-10}$$

$$P\left(D_{ij} > \overline{Effective\ Capacity_{ij}}\right) = P\left(D_{ij} > \overline{EC_{i,j}}\right) = 1 - \int_{0}^{\overline{EC_{i,j}}} f_{(D_j)} dD_j \tag{4-10a}$$

where

$$f_{(D_j)} = \frac{1}{\sqrt{2\pi}\sigma_{D_j} D_j} e^{-\frac{\left(\ln(D_j) - \overline{D_j}\right)^2}{2\sigma_{D_j}^2}} \tag{4-10b}$$

$$\sigma_D^2 = \frac{\sum_{t=1}^{n} \left(\ln(D_t) - \overline{D}\right)^2}{n-1} \tag{4-10c}$$

$$\overline{D} = \frac{\sum_{t=1}^{n} \ln(D_t)}{n} \tag{4-10d}$$

where

$$D_t = \frac{Demand_t}{FPV_{t-\overline{LT},t}} \tag{4-10e}$$

$$\overline{EC_{i,j}} = \frac{\overline{C_{i,j}} + \sum_{k=1, i \neq j,k}^{p} \overline{A_{i,jk}} \cdot \overline{FC_{i,jk}}}{FPV_{i-\overline{LT},i,j}} \tag{4-10f}$$

where

$$\overline{C_{i,j}} = \overline{C_{i-1,j}} + P\left(CPV_{i-\overline{LT},i,j} > CPV_{i-1-\overline{LT},i,j} \ldots CPV_{1-\overline{LT},1,j}\right) \cdot \max\left(\begin{array}{l} \left(\overline{CPV_{i-\overline{LT},i,j}} - \overline{CPV_{i-1-\overline{LT},i-1,j}}\right)_{CPV_{i-\overline{LT},i,j} > CPV_{i-1-\overline{LT},i-1,j} \ldots CPV_{1-\overline{LT},1,j}} \\ - \sum_{k=1,k\neq i,j}^{p} \left(\overline{A_{i,jk}} - \overline{A_{i-1,jk}}\right) \cdot \overline{FC_{i,jk}}, 0 \end{array} \right) \tag{4-10g}$$

where

$$\left.\overline{CPV_{i-\overline{LT},i,j}}\right|_{CPV_{i-\overline{LT},i,j} > CPV_{i-1-\overline{LT},i-1,j} \ldots CPV_{1-\overline{LT},1,j}} = E\left[\left.\frac{CPV_{i-\overline{LT},i,j}}{CPV_{i-1-\overline{LT},i-1,j}}\right|_{CPV_{i-\overline{LT},i,j} > CPV_{i-1-\overline{LT},i-1,j} \ldots CPV_{1-\overline{LT},1,j}} \right] \cdot \left.\overline{CPV_{i-1-\overline{LT},i-1,j}}\right|_{CPV_{i-1-\overline{LT},i-1,j} > CPV_{i-2-\overline{LT},i-2,j} \ldots CPV_{1-\overline{LT},1,j}} \tag{4-10h}$$

and

$$\overline{A_{i,jk}} = \left.\frac{\overline{A_{i-1,jk}} \cdot \overline{FC_{jk}} - \left(\overline{CPV_{i-\overline{LT},i,k}} - \overline{CPV_{i-1-\overline{LT},i-1,k}}\right)_{CPV_{i-1-\overline{LT},i-1,j} > CPV_{i-2-\overline{LT},i-2,j} \ldots CPV_{1-\overline{LT},1,j}}}{\overline{FC_{i,jk}}}\right|_{0<\overline{A_{i,jk}}<1} \tag{4-10i}$$

where

$$\left.\overline{CPV_{i-\overline{LT},i,k}}\right|_{CPV_{i-\overline{LT},i,j} > CPV_{i-1-\overline{LT},i-1,j} \ldots CPV_{1-\overline{LT},1,j}} = E\left[\left.\frac{CPV_{i-\overline{LT},i,k}}{CPV_{i-1-\overline{LT},i-1,k}}\right|_{CPV_{i-\overline{LT},i,j} > CPV_{i-1-\overline{LT},i-1,j} \ldots CPV_{1-\overline{LT},1,j}} \right] \cdot \left.\overline{CPV_{i-1-\overline{LT},i-1,k}}\right|_{CPV_{i-1-\overline{LT},i-1,j} > CPV_{i-2-\overline{LT},i-2,j} \ldots CPV_{1-\overline{LT},1,j}} \tag{4-10j}$$

and

$$\overline{FC_{i,jk}} = \overline{Percent\ Flexibility_{jk}} \cdot \overline{C_{i,j}} \tag{4-10k}$$

$$\begin{aligned} \overline{LT} = {} & \overline{LT_{Flexibility}} \cdot \overline{\%\ Flexibility} + \overline{LT_{Adaptability}} \cdot \overline{\%\ Adaptability} + \overline{LT_{Optional\ Expansion}} \cdot \overline{\%\ Optional\ Expansion} \\ & + \frac{\overline{Increment\ Size}}{Reference\ Increment\ Size} \cdot \overline{LT_{Reference\ Size\ New\ Installation}} \cdot \overline{\%\ New\ Installation} \end{aligned} \tag{4-10l}$$

<u>Key Assumptions:</u>

- The product life, L, is independent of the probability that demand, D, is greater than the effective capacity. In fact, these variables are positively correlated since the observed product life will increase if the demand for the product is strong, exceeding forecast sales and therefore likely exceeding the production capacity of the product

- A capacity stock out occurs due to demand only when the annual demand exceeds the annual capacity for the year. Thus, effects of seasonality and random demand fluctuation are ignored

Description of Equation:

The equation for Capacity-related Stock Outs (CSOs) is shown in (4-10). Stockouts due to capacity occur for three reasons. First, the capacity could be set to the required level yet demand still exceeds this level of capacity. In this situation, demand simply exceeds the amount the company is willing to support. The second situation occurs when the required capacity is in the process of being installed to the required level but due to installation lead time cannot support the current level of demand. This situation occurs when the forecast (and hence required capacity) changes within the capacity installation lead time. The third type of stock out occurs when, for any reason, the installed capacity is below the required level and the occurrence is not due to installation lead time. This situation is considered a data integrity error in (4-10) even though the cause may be due to a mis-communication in the capacity process or a lack of adherence to the required volumes, etc.

The expected number of CSOs is calculated by finding the average fraction of the time any particular product is expected to be in a stock out condition and then summing this fraction over all the product offerings. In this manner, the total number of capacity-related stock outs expected to occur in any given year can be calculated. To find the average fraction of time any particular product is expected to be in a stock out condition, the probability that its demand will exceed the expected capacity is calculated for each year of the expected product life cycle and then averaged over the number of years of product life. Therefore, to perform this calculation, the following variables and functions need to be known:

- The expected product life
- The expected capacity available each year to produce product
- The probability distribution of demand

Each of these terms and their constituent variables will be discussed in the sections that follow.

The expected product life cycle varies from product to product, but is typically about seven years for cars and eight to nine years for trucks.

The expected capacity available each year of the product life cycle is termed the *Effective Capacity* or *EC* and represents the sum of the dedicated and flexible capacity (in units per year) that can be allocated to the product in each year of the life cycle. This sum is then normalized by the forecast (FPV) used to set the capacity level, i.e. the current forecast of year **i** less the investment lead time, FPV_{i-LT}. The result is a capacity level that is measured in fractions of the forecast used at the time investment decisions were made, (4-10f).

The probability distribution of demand is based on the lognormal density distribution of demand (4-10b) where the standard deviation (4-10c).and the expected value (4-10d).are calculated by historical deviations of demand from the forecast of that demand made a period of time earlier equal to the investment lead time of capacity. These historical deviations are determined by the quotient of the actual demand and the forecast of that demand (4-10e). In the case of unbiased forecasts, the expected demand as a log fraction of forecast demand ($FPV_{i\text{-}LT}$) is will be equal to zero. The statistical parameters can be easily calculated as a function of lead time if the historical forecasts and respective demands are known for each product line.

The derivation of the expected dedicated capacity is shown in (4-10g). The expected dedicated capacity, ***C***, is measured in units per year and determined by the previous year's capacity plus the amount that can be expected to be added. The amount to be added each year is the probability that the current year's forecast is higher than all previous year's forecasts multiplied by the expected amount of dedicated capacity to be added. The expected amount of dedicated capacity that will be added is the difference between the current and previous year's capacity requirement (in units per year given the current and immediately previous capacity requirements are higher than each of their respective previous capacity requirements) less the amount of units of flexible capacity expected to become available. Increases in the availability of flexible capacity offset increases in capacity requirements as shown in the equation. For flexible capacity to be considered available, it cannot be counted toward any other capacity requirement. Thus, if additional flexible capacity is to become available, the complementary flexible product must have its capacity requirement decrease.

To determine the expected capacity requirement (CPV) in each year given that the current year's requirement is higher than all previous years (4-10h), it is necessary to know the expected percent change in capacity volume given that the most recent required capacity volume is higher than all previous volumes. This fraction is then multiplied by the previous year's expected volume to determine the current expected volume. The expected fractional increase in capacity volume can be calculated statistically through historical forecasts.

The final component of (4-10f) and (4-10g) requiring explanation is the calculation of available flexible capacity. The available flexible capacity is the number of units of flexible capacity, ***FC***, multiplied by the percentage of that flexible capacity available to the given product, ***j***. The number of units of flexible capacity can be measured in a number of ways with a simple and relatively accurate method being the average percent flexiblity (as a percent of dedicated capacity) times the amount of dedicated capacity for the given product (4-10k). The availablity of this flexible capacity (4-10i), is calculated for each year, ***i***, by taking the previous year's available capacity and subtracting the expected change in capacity requirement for the flexible product, ***k***. Hopefully products ***j*** and ***k*** are chosen to be flexible based on a historic negative correlation of their capacity requirements (i.e. demand movements) so that given an increase in CPV for product ***j***, the CPV for product ***k*** is expected to decrease, making more capacity available to product ***j***. The calculation for a relative change in CPV for ***k*** given an increase in the CPV for ***j*** is shown in (4-10j). This effect will increase the availability of flexible capacity to product ***j*** over time, up to a maximum of one hundred percent.

The last element of (4-10) to be explained is the expected capacity investment lead time, ***LT*** shown in (4-10l). For simplicity, the lead time is assumed to be a weighted average of different types of capital investments. The last component of this equation, the lead time for a new capacity installation is defined as a linear function of capacity increment size. The implicit assumption in this formulation is that lead time of new investments is a direct function of the size (units per year) of the investment made.

4.9.3 Derivation of Expected Utilization:

$$\overline{U}=\sum_{i=1}^{\overline{L}}\frac{\dfrac{\displaystyle\sum_{j=1}^{p}\sum_{k=1,k\neq j}^{p}\overline{Production_{D_{ij}>C}}+\overline{Production_{D_{ij}<C,D_{ik}<C}}+\overline{Production_{D_{ij}<C,D_{ik}>C}}}{\displaystyle\sum_{j=1}^{p}\sum_{k=1,k\neq j}^{p}\overline{C_{ij}}+\frac{\overline{FC_{i,jk}}}{2}}}{\overline{L}}$$

(4-11)

where

$$\overline{Production_{D_{ij}>C}}=P\left(D_{ij}>\frac{\overline{C_{i,j}}+\frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,j}}\right)\cdot\left(\overline{C_{i,j}}+\frac{\overline{FC_{i,jk}}}{2}\right)$$

(4-11a)

$$\overline{Production_{D_{ij}<C,D_{ik}<C}}=P\left(D_{ij}<\frac{\overline{C_{i,j}}+\frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,j}},D_{ik}<\frac{\overline{C_{i,k}}+\frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,k}}\right)\cdot E\left[D_{ij}\Big|_{D_{ij}<\frac{\overline{C_{i,j}}+\frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,j}},D_{ik}<\frac{\overline{C_{i,k}}+\frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,k}}}\right]\cdot FPV_{i-\overline{LT},i,j}$$

(4-11b)

$$\overline{Production_{D_{ij}<C,D_{ik}>C}}=P\left(D_{ij}<\frac{\overline{C_{i,j}}+\frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,j}},D_{ik}>\frac{\overline{C_{i,k}}+\frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,k}}\right)\cdot$$

$$Min\left(\begin{array}{l}E\left(D_{ij}\Big|_{D_{ij}<\frac{\overline{C_{i,j}}+\frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,j}},D_{ik}>\frac{\overline{C_{i,k}}+\frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,k}}}\right)\cdot FPV_{i-\overline{LT},i,j}+\\ Min\left(E\left(D_{ik}\Big|_{D_{ij}<\frac{\overline{C_{i,j}}+\frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,j}},D_{ik}>\frac{\overline{C_{i,k}}+\frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,k}}}\right)\cdot FPV_{i-\overline{LT},i,j}-\left(\overline{C_{i,k}}+\frac{\overline{FC_{i,jk}}}{2}\right),\frac{\overline{FC_{i,jk}}}{2}\right),\overline{C_{i,j}}+\frac{\overline{FC_{i,jk}}}{2}\end{array}\right)$$

(4-11c)

where

$$\overline{C_{i,j}} = \overline{C_{i-1,j}} + Max\left(P\left(CPV_{i-\overline{LT},i,j} > CPV_{i-1-\overline{LT},i,j} \ldots CPV_{1-\overline{LT},1,j} \right) \cdot \left(\begin{array}{l} \left(\overline{CPV_{i-\overline{LT},i,j}} - \overline{CPV_{i-1-\overline{LT},i-1,j}} \right)_{CPV_{i-\overline{LT},i,j} > CPV_{i-1-\overline{LT},i-1,j} \ldots CPV_{1-\overline{LT},1,j}} \\ - \sum_{k=1,k\neq i,j}^{p} \left(\overline{A_{i,jk}} - \overline{A_{i-1,jk}} \right) \cdot \overline{FC_{i,jk}} \end{array} \right), 0 \right) \tag{4-11d}$$

where

$$\left. \overline{CPV_{i-\overline{LT},i,j}} \right|_{CPV_{i-\overline{LT},i,j} > CPV_{i-1-\overline{LT},i-1,j} \ldots CPV_{1-\overline{LT},1,j}} = E\left[\left. \frac{CPV_{i-\overline{LT},i,j}}{CPV_{i-1-\overline{LT},i-1,j}} \right|_{CPV_{i-\overline{LT},i,j} > CPV_{i-1-\overline{LT},i-1,j} \ldots CPV_{1-\overline{LT},1,j}} \right] \cdot \left. \overline{CPV_{i-1-\overline{LT},i-1,j}} \right|_{CPV_{i-1-\overline{LT},i-1,j} > CPV_{i-2-\overline{LT},i-2,j} \ldots CPV_{1-\overline{LT},1,j}} \tag{4-11e}$$

and

$$\overline{A_{i,jk}} = \left. \frac{\overline{A_{i-1,jk}} \cdot \overline{FC_{jk}} - \left(\overline{CPV_{i-\overline{LT},i,k}} - \overline{CPV_{i-1-\overline{LT},i-1,k}} \right)_{CPV_{i-1-\overline{LT},i-1,j} > CPV_{i-2-\overline{LT},i-2,j} \ldots CPV_{1-\overline{LT},1,j}}}{\overline{FC_{i,jk}}} \right|_{0 < \overline{A_{i,jk}} < 1} \tag{4-11f}$$

where

$$\left. \overline{CPV_{i-\overline{LT},i,k}} \right|_{CPV_{i-\overline{LT},i,j} > CPV_{i-1-\overline{LT},i-1,j} \ldots CPV_{1-\overline{LT},1,j}} = E\left[\left. \frac{CPV_{i-\overline{LT},i,k}}{CPV_{i-1-\overline{LT},i-1,k}} \right|_{CPV_{i-\overline{LT},i,j} > CPV_{i-1-\overline{LT},i-1,j} \ldots CPV_{1-\overline{LT},1,j}} \right] \cdot \left. \overline{CPV_{i-1-\overline{LT},i-1,k}} \right|_{CPV_{i-1-\overline{LT},i-1,j} > CPV_{i-2-\overline{LT},i-2,j} \ldots CPV_{1-\overline{LT},1,j}} \tag{4-11g}$$

and

$$\overline{FC_{i,jk}} = \overline{Percent\ Flexibility_{jk}} \cdot \overline{C_{i,j}} \tag{4-11h}$$

$$\begin{aligned} \overline{LT} = {} & \overline{LT_{Flexibility}} \cdot \overline{\%\ Flexibility} + \overline{LT_{Adaptability}} \cdot \overline{\%\ Adaptability} + \overline{LT_{Optional\ Expansion}} \cdot \overline{\%\ Optional\ Expansion} \\ & + \frac{\overline{Increment\ Size}}{Reference\ Increment\ Size} \cdot \overline{LT_{Reference\ Size\ New\ Installation}} \cdot \overline{\%\ New\ Installation} \end{aligned} \tag{4-11i}$$

Description of Equation:

The notation used in (4-11) is similar to that of (4-10). Terms not defined herein can be found in 4.9.2.

The expected utilization, ***U***, is the average utilization across the company for each year of product life, averaged over the product life cycle, ***L***. The average utilization for a given year is the ratio of the total production to the total capacity. From the perspective of capacity, production can be classified in three ways. Production can occur where the demand for product ***j*** is greater than the assigned capacity (4-11a). The second situation arises when the demand for both products ***j*** and ***k*** is less than their respective capacities (4-11b). The third situation occurs when the demand for ***j*** is less than its capacity while the demand for ***k*** is greater than its respective capacity (4-11c). In this circumstance, a portion of the flexible capacity between ***j*** and ***k*** that is assigned to ***j*** will be used to produce product ***k***. The sum of the expected values of each of these mutually exclusive situations will be the total expected production through a given piece of capacity.

The expected production where demand of ***j*** is greater than its capacity is simply the probability that demand is greater than capacity multiplied by the capacity (4-11a). The capacity in this case is the dedicated capacity plus one half the flexible capacity, if any. By convention, half of any flexible capacity is assigned to product ***j*** and the other half is assigned to product ***k***.

The expected production where both ***j*** and ***k*** have demands less than their respective capacities is the joint probability of this occurrence multiplied by the expected value of demand given that demand for both ***j*** and ***k*** is less than capacity. Finally, this product is multiplied by the forecast volume, (FPV) to obtain an answer in units per year (4-11b). This step is necessary since the demand, D_{ij} is measured in deviations from the historical forecast.

The third component of (4-11) is the expected production when demand for ***j*** is less than its capacity and the demand for ***k*** is greater than its capacity (4-11c). This equation is similar to (4-11b) wherein the joint probability of occurrence is multiplied by the expected production given the joint probability. The difference between the two, however, lies in the expected production rate. To the extent that the demand for ***k*** is greater than its capacity it will use up to the amount of excess capacity expected to be available by ***j***.

The subsequent derivations, (4-11d) through (4-11i) are identical to the derivation and explanation in 4.9.2.

4.9.4 Derivation of Expected Sales Given Expected Capacity, Flexibility

$$\overline{Sales_{ij}} = \sum_{k=1,k\neq j}^{p} \overline{Sales_{D_{ij}>C}} + \overline{Sales_{D_{ij}<C,D_{ik}<C}} + \overline{Sales_{D_{ij}<C,D_{ik}>C}} \tag{4-12}$$

where

$$\overline{Sales_{D_{ij}>C}} = P\left(D_{ij} > \frac{\overline{C_{i,j}} + \frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,j}} \right) \cdot \left(\overline{C_{i,j}} + \frac{\overline{FC_{i,jk}}}{2} \right) \tag{4-12a}$$

$$\overline{Sales_{D_{ij}<C,D_{ik}<C}} = P\left(D_{ij} < \frac{\overline{C_{i,j}} + \frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,j}}, D_{ik} < \frac{\overline{C_{i,k}} + \frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,k}} \right) \cdot E\left[D_{ij} \Big|_{D_{ij} < \frac{\overline{C_{i,j}} + \frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,j}}, D_{ik} < \frac{\overline{C_{i,k}} + \frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,k}}} \right] \cdot FPV_{i-\overline{LT},i,j} \tag{4-12b}$$

$$\overline{Sales_{D_{ij}<C,D_{ik}>C}} = P\left(D_{ij} < \frac{\overline{C_{i,j}} + \frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,j}}, D_{ik} > \frac{\overline{C_{i,k}} + \frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,k}} \right) \cdot$$

$$Min\left(\begin{array}{l} E\left(D_{ij} \Big|_{D_{ij} < \frac{\overline{C_{i,j}} + \frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,j}}, D_{ik} > \frac{\overline{C_{i,k}} + \frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,k}}} \right) \cdot FPV_{i-\overline{LT},i,j} + \\ Min\left(E\left(D_{ik} \Big|_{D_{ij} < \frac{\overline{C_{i,j}} + \frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,j}}, D_{ik} > \frac{\overline{C_{i,k}} + \frac{\overline{FC_{i,jk}}}{2}}{FPV_{i-\overline{LT},i,k}}} \right) \cdot FPV_{i-\overline{LT},i,j} - \left(\overline{C_{i,k}} + \frac{\overline{FC_{i,jk}}}{2} \right), \frac{\overline{FC_{i,jk}}}{2} \right), \overline{C_{i,j}} + \frac{\overline{FC_{i,jk}}}{2} \end{array} \right) \tag{4-12c}$$

where

$$\overline{C_{i,j}} = \overline{C_{i-1,j}} + Max\left(P\left(CPV_{i-\overline{LT},i,j} > CPV_{i-1-\overline{LT},i,j} \ldots CPV_{1-\overline{LT},1,j} \right) \cdot \left(\begin{array}{l} \left(\overline{CPV_{i-\overline{LT},i,j}} - \overline{CPV_{i-1-\overline{LT},i-1,j}} \right)_{CPV_{i-\overline{LT},i,j} > CPV_{i-1-\overline{LT},i-1,j} \ldots CPV_{1-\overline{LT},1,j}} \\ - \sum_{k=1,k\neq i,j}^{p} \left(\overline{A_{i,jk}} - \overline{A_{i-1,jk}} \right) \cdot \overline{FC_{i,jk}} \end{array} \right), 0 \right) \tag{4-12d}$$

where

$$\left.\overline{CPV_{i-\overline{LT},i,j}}\right|_{CPV_{i-\overline{LT},i,j}>CPV_{i-1-\overline{LT},i-1,j}\ldots CPV_{1-\overline{LT},1,j}}=E\left[\left.\frac{CPV_{i-\overline{LT},i,j}}{CPV_{i-1-\overline{LT},i-1,j}}\right|_{CPV_{i-\overline{LT},i,j}>CPV_{i-1-\overline{LT},i-1,j}\ldots CPV_{1-\overline{LT},1,j}}\right]\cdot\left.\overline{CPV_{i-1-\overline{LT},i-1,j}}\right|_{CPV_{i-1-\overline{LT},i-1,j}>CPV_{i-2-\overline{LT},i-2,j}\ldots CPV_{1-\overline{LT},1,j}} \tag{4-12e}$$

and

$$\overline{A_{i,jk}}=\left.\frac{\overline{A_{i-1,jk}}\cdot\overline{FC_{jk}}-\left(\overline{CPV_{i-\overline{LT},i,k}}-\overline{CPV_{i-1-\overline{LT},i-1,k}}\right)_{CPV_{i-1-\overline{LT},i-1,j}>CPV_{i-2-\overline{LT},i-2,j}\ldots CPV_{1-\overline{LT},1,j}}}{\overline{FC_{i,jk}}}\right|_{0<\overline{A_{i,jk}}<1} \tag{4-12f}$$

where

$$\left.\overline{CPV_{i-\overline{LT},i,k}}\right|_{CPV_{i-\overline{LT},i,j}>CPV_{i-1-\overline{LT},i-1,j}\ldots CPV_{1-\overline{LT},1,j}}=E\left[\left.\frac{CPV_{i-\overline{LT},i,k}}{CPV_{i-1-\overline{LT},i-1,k}}\right|_{CPV_{i-\overline{LT},i,j}>CPV_{i-1-\overline{LT},i-1,j}\ldots CPV_{1-\overline{LT},1,j}}\right]\cdot\left.\overline{CPV_{i-1-\overline{LT},i-1,k}}\right|_{CPV_{i-1-\overline{LT},i-1,j}>CPV_{i-2-\overline{LT},i-2,j}\ldots CPV_{1-\overline{LT},1,j}} \tag{4-12g}$$

and

$$\overline{FC_{i,jk}}=\overline{Percent\ Flexibility_{jk}}\cdot\overline{C_{i,j}} \tag{4-12h}$$

$$\begin{aligned}\overline{LT}=\;&\overline{LT_{Flexibility}}\cdot\overline{\%\ Flexibility}+\overline{LT_{Adaptability}}\cdot\overline{\%\ Adaptability}+\overline{LT_{Optional\ Expansion}}\cdot\overline{\%\ Optional\ Expansion}\\&+\frac{\overline{Increment\ Size}}{Reference\ Increment\ Size}\cdot\overline{LT_{Reference\ Size\ New\ Installation}}\cdot\overline{\%\ New\ Installation}\end{aligned} \tag{4-12i}$$

Description of Equation:

The expected sales, (4-12) consists of three principal components: (4-12a), (4-12b), and (4-12c). Each of these are identical to the three components comprising the total expected production: (4-11a), (4-11b), and (4-11c). The reason that expected sales is not the expected demand is due to the fact that sales will always be constrained by capacity. Only what can be built will be sold. In certain industries, capacity constrained orders can be backlogged. However, for the auto industry, if the desired car is not available, the customer will balk and go to a competitor.

The subsequent derivations, (4-12d) through (4-12i) are identical to the derivation and explanation in 4.9.2.

Chapter 5

Conclusions and Recommendations

5.1 Summary of Contributions

This thesis has made several contributions to theory as well as practice. In accordance with the eight objectives outlined in Section 1.2, a total of eight significant contributions were made in this work. First, this thesis developed relationships characterizing the dynamic, interdependent behavior between autonomous groups in distributed systems. Concurrent with the presentation of these relationships, a mathematical mapping between Axiomatic Design and System Dynamics was established, capturing this interdependent behavior and providing an objective approach to distributed systems modeling. Then, in Section 2.3.2, the notion of a Bounded Interdependence Matrix was developed to provide a high-level diagrammatical abstraction of coupling interdependencies between groups. A total of six types of coupling interdependencies which drive system behavior were subsequently identified. In light of these interdependencies, eight techniques were established for the reduction or elimination of adverse coupling behavior, thereby enabling the optimization of multi-objective systems under distributed control.

Relative to the theoretical nature of the previous contributions, the case studies in this thesis contributed in specific ways. In Section 4.6.5, five manufacturing capital investment strategies were identified. These strategies provide a hedge against demand variation. The seventh contribution to this work was to extend the use of financial options to the capacity planning process. Finally, a risk-based manufacturing investment strategy was developed, explicitly recognizing the role of variance in manufacturing.

5.2 Recommendations For Further Work

While this thesis has primarily focused on the development and determination of an analytic approach to highlight the role of variance in manufacturing system design, neither the tools themselves nor the results are the ends of this work. At a basic level, this research has wrestled with two fundamental issues confronting modern manufacturing organizations. The first issue is to *understand* how local optimization or improvement creates coupling interdependencies which propagate through the established system, compromising overall performance. Secondly, given that the relationships between the control variables and performance objectives can be understood, the challenge becomes how to *represent* these relationships and behaviors in such a way that managers and practitioners alike can have confidence in the findings. Without both of these sufficiently addressed, it will not be possible to foster an environment of change for the overall improvement of the firm.

Toward this end, there are three areas recommended for further development. First, an effective method needs to be established for the representation of nonlinear couplings of two or more variables. It is not uncommon for the situation to arise where the coupling strength as well as sign of one control variable is a function of another control variable. This nonlinear coupling is dependent upon the states of other variables in the system at any given time and therefore does not lend itself to proper representation in the current Object-Oriented Axiomatic Design methodology.

Second, there is significant opportunity for improvement to the mathematical as well as diagrammatical representation of interdependencies caused by stochastic variance in systems. This work does little more than recognize the important role stochastic variance plays in distributed systems.

Third, an invaluable extension of this work would be the compilation of recurring behavior modes or system archetypes found in distributed, interdependent systems. Such a compilation would greatly facilitate the modeling and optimization of coupling interdependencies.

5.3 Conclusion

In recent years, a number of trends have become manifest in the business environment which underscore the need for a systems-perspective in the design and management of manufacturing systems. The lack of organizational and manufacturing slack resulting from companies' efforts to become lean, coupled with increased demand volatility due to market fragmentation has created an environment which is increasingly difficult to navigate. Stack on the increased system complexity and need for quick-response which has led to decentralized decision-making and it is evident that the role of variance will become an increasingly large determinant of overall manufacturing system performance. Thus, the onus is on companies to develop an active role in this area.

The intent of this work has been to increase awareness of the role of variance in manufacturing system design. The tools presented here are intended to be expedient toward this end.

References

1 Bakke, N.A. and Hellberg, R., 1993, The Challenges of Capacity Planning. *International Journal of Production Economics*, **30-31**, 243-264.

2 Hammosfahr, R.D. Jack, Pope, J.A., and Ardalan, A., 1993, Strategic Planning for Production Capacity. *International Journal of Operations and Production Management*, **13**, 41-53.

3 Hayes, R.H., Wheelwright, S.C., Clark, K.B., 1988, *Dynamic Manufacturing: Creating the Learning Organization* (New York, NY: The Free Press).

4 Manne, A.S., 1961, Capacity Expansion and Probabilistic Growth. *Econometrica*, **29**, 632-649.

5 Manne, A.S. (ed.), 1967, *Investments for Capacity Expansion: Size, Location, and Time-Phasing.* (Cambridge, MA: MIT Press)

6 Kalotay, A.J., 1973, Capacity Expansion and Specialization. *Management Science*, **20**, 56-64.

7 Luss, H., 1979, A Capacity Expansion Model for Two Facility Types. *Naval Research Logistics Quarterly*, **26**, 291-303.

8 Luss, H., 1982, Operations Research and Capacity Expansion Problems: A Survey, *Operations Research*, **30**, 907-947.

9 Fine, C.H., Freund, R.M., 1990, Optimal Investment in Product-Flexible Manufacturing Capacity, *Management Science*, **36**, 449-466.

10 Hammer, Michael, *Reengineering the Corporation*, 1994, New York, NY: Harper Business.

11 Suh, Nam P., The Principles of Design, 1990, New York, NY, Oxford University Press.

12 Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, 1965, New York, NY: McGraw-Hill Book Co.

13 Strang, Gilbert, *Linear Algebra and its Applications*, 1988, San Diego, CA: Harcourt Brace Jovanovich.

14 Forrester, Jay W., *Industrial Dynamics*, 1961, Portland, OR: Productivity Press.

15 Senge, Peter M., *The Fifth Discipline: The Art and Practice of the Learning Organization*, 1990, New York, NY: Doubleday.

16 Sterman, John D., "Learning in and about Complex Systems," *System Dynamics Review*, 1994, **10**, 291-330.

17 Adapted from Repenning, Nelson P., "The Improvement Paradox: Three Essays on Process Improvement Initiatives", *Ph.D. Thesis*, Alfred P. Sloan School of Management, MIT, 1996.

18 Imai, Masaki, *Kaizen: The Key to Japanese Competitive Success*, 1986, New York, NY: Random House Business Division.

19 Stata, Ray, "Organizational Learning – The Key to Management Innovation," *Sloan Management Review*, 1989, **Spring**, 63-74.

20 Adapted from Harvard Business School *Barilla SpA* case # 9-694-046

21 Albano, Leonard D., Suh, Nam P., "The Information Axiom and Its Implications," *Intelligent Concurrent Design: Fundamentals, Methodology, Modeling and Practice*, ASME 1993, **66**, 1-11.

22 Suh, Nam P., "Design Axioms and Quality Control," *Robotics and Computer-Integrated Manufacturing*, 1992, **9**, 367-378.

23 Kim, S-J., Suh, N.P., and Kim, S-G., "Design of Software Systems Based on Axiomatic Design," *Robotics and Computer-Integrated Manufacturing*, 1991, **Vol. 8 No. 4**, 243-255.

24 Black, J.T., "The Design of Manufacturing Systems: Axiomatic Approach," *Design, Analysis, and Control of Manufacturing Cells*, ASME-PED 1991, **53**, 1-14.

25 Suh, N.P., Bell, A.C., Gossard, D.C., "On an Axiomatic Approach to Manufacturing and Manufacturing Systems," *Journal of Engineering for Industry*, 1978, **100**, 127-130.

26 Suh, Nam P., "Development of the Science Base for the Manufacturing Field Through the Axiomatic Approach," *Robotics and Computer-Integrated Manufacturing*, 1984, **1**, 397-415.

27 Suh, Nam P., "Designing-in of Quality Through Axiomatic Design," *IEEE Transactions on Reliability*, 1995, **44**, 256-264.

28 Keough, M., Doman, A., "The CEO as Organization Designer," *The McKinsey Quarterly*, 1992, **2**, 3-30.

29 Augusto, L.A., Forrester, J.W., Lyneis, J.M., *System Dynamics*, 1980, New York: North Holland.

30 Forrester, J.W., *World Dynamics*, 1973, Cambridge, MA: Productivity Press.

31 Forrester, J.W., "A New Corporate Design," *Industrial Management Review*, **Vol. 7 No. 1**.

32 Forrester, J.W., *Urban Dynamics*, 1969, Cambridge, MA: Productivity Press.

33 Forrester, Jay W. "Market Growth as Influenced by Capital Investment," *Industrial Management Review*, Winter 1968, 9 no. 2, 83-105. (Sloan Management Review).

34 Booch, Grady, *Object-Oriented Analysis and Design With Applications*, 1993, Menlo Park, CA: Benjamin Cummings Press.

35 Coad, Peter, Yourdon, Ed, *Object-Oriented Analysis*, 1991, Englewood Cliffs, NJ: Prentice Hall.

36 Rumbaugh, J., et al, *Object-Oriented Modeling and Design*, 1991, Englewood Cliffs, NJ: Prentice Hall

37 Montgomery, Douglas C., et al, *Forecasting and Time Series Analysis*, 1990, New York, NY: McGraw-Hill, Inc.

38 Brealey, Richard A., Myers, Stewart C., *Principles of Corporate Finance*, 1996, New York, NY: The McGraw-Hill Companies, Inc.

39 Black, F. and Scholes, M., "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, May-June 1973, **81**, 637-654.

Bibliography

1. Arogyaswamy, B., Simmons, R.P., “Thriving on Interdependence: The Key to JIT Implementation,” *Production and Inventory Management Journal*, 1991, Third Quarter, 56-60.

2. Athaide, C., “Capacity Allocation and Safety Stocks in Manufacturing Systems,” *Ph.D., MIT Sloan School of Management*, 1992.

3. Bahrami, H., “The Emerging Flexible Organization: Perspectives from Silicon Valley,” *California Management Review*, Summer 1992, 33-52.

4. Bakke, H.A., Hellberg, R., “The Challenges of Capacity Planning,” *International Journal of Production Economics*, 1993, **30-31**, 243-264.

5. Beedle, M.A., “Object-Based Reengineering,” *Object Magazine*, March-April 1995, 53-58.

6. Black, T.A., Fine, C.H., Sachs, E.M., “A Method for Systems Design Using Precedence Relationships: An Application to Automotive Brake Systems,” Working Paper #3208-90-MS, MIT Sloan School of Management, October 1990.

7. Bowen, H.K., Clark, K.B., Holloway, C.A., Leonard-Barton, D., Wheelwright, S.C., “Regaining the Lead in Manufacturing,” *Harvard Business Review*, September-October 1994, 108-119.

8. Buss, A.H., Lawrence, S.R., Kropp, D.H., “Volume and Capacity Interaction in Facility Design,” *IIE Transactions*, July 1994, **Vol. 26 No. 4**, 36-49.

9. Chang, T.S., Ward, A.C., Lee, J., Jacox, E.H., “Distributed Design With Conceptual Robustness: A Procedure Based on Taguchi’s Parameter Design,” *Concurrent Product Design*, ASME 1994, **74**, 19-22.

10. Cohen, M.A., Halperin, R.M., “Optimal Technology Choice in a Dynamic-Stochastic Environment,” *Journal of Operations Management*, 1986, **Vol. 6 No. 3**, 317-331.

11. Colgan, J.A., “A Business Planning Model to Access the Tradeoff Between Inventory and Capacity for a Stage 1 Manufacturing Process,” *M.S.M.E. MIT Thesis*, 1994.

12. Crittenden, V.L., Crittenden, W.F., “Examining the Impact of Manufacturing and Marketing Capacity Decisions on Firm Profitability,” *International Journal of Production Economics*, 1995, **40**, 57-72.

13. Crowston, K., "A Taxonomy of Organizational Dependencies and Coordination Mechanisms," Working Paper, University of Michigan School of Business Administration, October 1994.

14. Cusumano, M.A., "The Limits of 'Lean'," *Sloan Management Review*, Summer 1994, 27-32.

15. Drucker, P.E., "The Emerging Theory of Manufacturing," *Harvard Business Review*, May-June 1990, 94-102.

16. Eisenhardt, K.M., "Speed and Strategic Choice: How Managers Accelerate Decision Making," *California Management Review*, Spring 1990, 39-53.

17. Fisher, M.L., Hammond, J.H., Obermeyer, W.R., Raman, A., "Making Supply Meet Demand in an Uncertain World," *Harvard Business Review*, May-June 1994, 83-93.

18. Freidenfelds, J., *Capacity Expansion*, 1981, New York, NY: Elsevier-North Holland.

19. Garvin, D.A., *Managing Quality*, 1988, New York, NY: Free Press.

20. Gebala, D.A., Eppinger, S.D., "Methods for Analyzing Design Procedures," *Design Theory and Methodology*, ASME 1991, **31**, 227-233.

21. Hayes, R.H., Pisano, G.P., "Beyond World-Class: The New Manufacturing Strategy," *Harvard Business Review*, January-February 1994, 77-86.

22. Hayes, R.H., Schmenner, R.W., "How Should You Organize Manufacturing?", *Harvard Business Review*, January-February 1978, 105-118.

23. Hayes, R.H., Wheelwright, S.C., Clark, K.B., *Dynamic Manufacturing*, 1988, New York, NY: The Free Press.

24. Hoover, S.P., Rinderle, J.R., "Abstractions, Design Views and Focusing," *Design Theory and Methodology*, ASME 1994, **68**, 115-129.

25. Hopp, W.J., Spearman, M.L., *Factory Physics: Foundations of Manufacturing Management*, 1996, Boston, MA: Irwin.

26. Jordan, W.C., Graves, S.C., "Principles of the Benefits of Manufacturing Process Flexibility," *Management Science*, 1995, **Vol. 41 No. 4**, 577-594.

27. Kamath, R.R., Liker, J.K., "A Second Look at Japanese Product Development," *Harvard Business Review*, November-December 1994, 154-170.

28. Kannapan, S.M., Bell, D.G., Taylor, D.L., "Structuring Information and Coordinating Teams in Product Development," *Design Theory and Methodology*, ASME 1993, **53**, 233-242.

29. Keough, M., Doman, A., "The CEO as Organization Designer," *The McKinsey Quarterly*, 1992, **2**, 3-30.

30. Krishnan, V., Eppinger, S.D., Whitney, D.E., "Iterative Overlapping: Accelerating Product Development By Preliminary Information Exchange," *Design Theory and Methodology*, ASME 1993, **53**, 223-231.

31. Krishnan, V., Eppinger, S.D., Whitney, D.E., "Towards a Cooperative Design Methodology: Analysis of Sequential Decision Strategies," *Design Theory and Methodology*, ASME 1991, **31**, 165-172.

32. Kusiah, A., Wang, J., "Qualitative Analysis of the Design Process," *Intelligent Concurrent Design: Fundamentals, Methodology, Modeling and Practice*, ASME 1993, **66**, 21-32.

33. Kusiak, A., Larson, T.N., Wang, J., "Reengineering of Design and Manufacturing Processes," *Computers And Industrial Engineering*, 1994, **Vol. 26 No. 3**, 521-536.

34. Leone, R.A., Meyer, J.R., "Capacity Strategies for the 1980s," *Harvard Business Review*, November-December 1980, 133-140.

35. Li, S., Tirupati, D., "Technology Choice and Capacity Expansion with Two Product Families Tradeoffs Between Scale and Scope," *International Journal of Production Research*, 1992, **Vol. 30 No. 4**, 887-907.

36. Li, S., Tirupati, D., "Technology Selection and Capacity Planning for Manufacturing Systems," from Kusiah, A., Ed., *Intelligent Design and Manufacturing*, 1992, New York, NY: John Wiley & Sons, Chapter 10, 257-287.

37. Luss, H., "A Heuristic for Capacity Expansion Planning with Multiple Facility Types," *Naval Research Logistics Quarterly*, 1986, **33**, 685-701.

38. Mair, A., "Honda's Global Flexifactory Network," *International Journal of Operations & Production Management*, 1994, **Vol. 14 No. 3**, 6-23.

39. Manne, A.S., *Investments for Capacity Expansion*, 1967, London: George Allen and Unwin.

40. McCutcheon, D.M., Raturi, A.S., Meredith, J.R., "The Customization – Reponsiveness Squeeze," *Sloan Management Review*, Winter 1994, 89-99.

41. Mintzberg, H., "The Fall and Rise of Strategic Planning," *Harvard Business Review*, January-February 1994, 107-114.

42. Morris, W.T., *A Capacity Decision System*, 1967, Homewood, IL: Richard D. Irwin.

43. Nelson, G., "Manufacturing Flexibility," *Strategic Directive A-123*, Internal Document: Ford Motor Company, June 1993.

44. Pimmler, T.U., Eppinger, S.D., "Integration Analysis of Product Decompositions," *Design Theory and Methodology*, ASME 1994, **68**, 343-351.

45. Pine, B.J., Victor, B., Boynton, A.C., "Making Mass Customization Work," *Harvard Business Review*, September-October 1993, 108-119.

46. Rohloff, M., "Decentralized Production Planning and Design of a Production Management System Based on Object-Oriented Architecture," *International Journal of Production Economics*, 1993, **30-31**, 365-383.

47. Schonberger, R.J., "Frugal Manufacturing," *Harvard Business Review*, September-October 1987, 95-100.

48. Sinha, D.S., Wei, J.C., "Stochastic Analysis of Flexible Process Choices," *European Journal of Operational Research*, 1992, **60**, 183-199.

49. Skinner, W., "Manufacturing – Missing Link in Corporate Strategy", *Harvard Business Review*, May-June 1969, 136-145.

50. Skinner, W., "The Focused Factory," *Harvard Business Review,* May-June 1974, 113-121.

51. Smith, R.P., Eppinger, S.D., Gopal, A., "Testing an Engineering Design Iteration Model in an Experimental Setting," *Design Theory and Methodology*, ASME 1992, **42**, 267-276.

52. Steward, D.V., "The Design Structure System: A Method for Managing the Design of Complex Systems," *IEEE Transactions on Engineering Management*, 1981, **Vol. EM-28 No. 3**, 71-74.

53. Suarez, F.F., Cusumano, M.A., Fine, C.H., "An Empirical Study of Flexibility in Manufacturing," *Sloan Management Review*, Fall 1995, 25-32.

54. Suarez, F.F., Cusumano, M.A., Fine, C.H., "Flexibility and Performance: A Literature Critique and Strategic Framework," Working Paper #50-91 International Center for Research on the Management of Technology, MIT, Cambridge, November 1991.

55. Upton, D.M., "Flexibility as Process Mobility: The Management of Plant Capabilities for Quick Response Manufacturing," *Journal of Operations Management*, 1995, **12**, 205-224.

56. Upton, D.M., "The Management of Manufacturing Flexibility," *California Management Review*, Winter 1994, 72-89.

57. Upton, D.M., "What Really Makes Factories Flexible?", *Harvard Business Review*, July-August 1995, 74-84.

58. Vercellis, C., "Multi-Criteria Models for Capacity Analysis and Aggregate Planning in Manufacturing Systems," *International Journal of Production Economics*, 1991, **23**, 261-272.

59. Verter, V., Dincer, M.C., "An Integrated Evaluation of Facility Location, Capacity Acquisition, and Technology Selection for Designing Global Manufacturing Strategies," *European Journal of Operational Research*, 1992, **60**, 1-18.

60. Ward, A., Liker, J.K., Cristiano, J.J., Sobek, D.K., "The Second Toyota Paradox: How Delaying Decisions Can Make Better Cars Faster," *Sloan Management Review*, Spring 1995, 43-61.

61. Wheelwright, S.C., *Capacity Planning and Facilities Choice,* 1979, Boston: Division of Research, Harvard Business School.

62. Womack, J.P., Jones, D.T., "From Lean Production to the Lean Enterprise," *Harvard Business Review*, March-April 1994, 93-103.

63. Zijm, W.H.M., Buitenhek, R., "Capacity Planning and Lead Time Management," *International Journal of Production Economics*, 1996, **46-47**, 165-179.