1993

A Restoration-Based Model for Materials Management in a Global Manufacturing Environment.

Gary Wayne Clendenen
Louisiana State University and Agricultural & Mechanical College

Order Number 9419873

A restoration-based model for materials management in a global manufacturing environment

Clendenen, Gary Wayne, Ph.D.

The Louisiana State University and Agricultural and Mechanical Col., 1993

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

A RESTORATION BASED MODEL FOR MATERIALS MANAGEMENT IN A GLOBAL MANUFACTURING ENVIRONMENT

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Interdepartmental Program in Business Administration

by
Gary Wayne Clendenen
B.S., New Mexico State University, 1972
M.A., University of New Mexico, 1973
M.S., University of Washington, 1975
December, 1993

ACKNOWLEDGEMENTS

First, I would like to thank Dr. Dan B. Rinks, the chairman of my committee. His help and suggestions during this formative stage of my research career was of immeasurable value.

The members of my doctoral committee, Dr. Kwei Tang, Dr. Peter Kelle, Dr. Jim Pruett, Dr. Ishwar Murthy, and Dr. Linguo Gong gave freely of their time and offered helpful guidance and support throughout the entire process. A special thanks is also due to Dr. Ralph Pike for his suggestions regarding the solving of the nonlinear models.

I would also like to thank the Board of Regents for supporting me with a fellowship during my tenure at Louisiana State University. Finally, I wish to thank my family for their support and patience.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF TABLES (continued)

LIST OF FIGURES

LIST OF FIGURES (continued)

ABSTRACT

Globalization of manufacturing along with increased competition has made effective planning and control more important than ever. At the same time, it is more difficult than ever to achieve effective planning and control due to larger leadtimes and shorter product life cycles. The objective of this research is to explore the importance of control strategy on materials management in global manufacturing networks.

Control strategies in common use and others that have recently been proposed in the literature are reviewed and classified along a push/pull gradient. It is shown that one of them, the restoration control strategy, can be used to represent a wide range of pull systems as well as certain elements of push systems.

Using concepts underlying the restoration strategy, two models are developed for aggregate planning in a global manufacturing network. One model requires that all demands be met whereas the other allows some sales to be lost. Application of either of the models to a specific network results in values for decision variables, including target inventories and restoration coefficients. Target inventories are aggregate values that can be disaggregated to finer levels of detail. Values for restoration coefficients help identify the best control strategy.

Both models apply to multi-echelon networks of any design and under known demand. Both formulations are nonlinear, mixed-integer programming models that have proven to be difficult to solve for the general case. Relaxing the integrality constraints allows the models to be solved using commercially available software although optimality cannot be guaranteed due to nonconvexity of constraints.

The models were applied to a specific network. The restoration model with no lost sales was found to have severe limitations; however, the restoration model that allows lost sales provided results that were stable. The relationships between the decision variables and holding costs, labor costs, and demand variation were explored using the simulation technique of batch means. Among other things, results indicated that a control strategy very similar to base stock was most appropriate for the specific network studied.

CHAPTER 1 INTRODUCTION

1.1 Overview

At the end of World War II, manufacturing productivity of the United States far surpassed that of any other nation for a number of reasons. Foremost, American industry and society had not been torn apart during the war as badly as had those of Europe and Japan. Yet, there were more fundamental reasons for America's postwar success, including superior technology, availability of capital, and a highly educated work force. Furthermore, the United States enjoyed a uniquely large and affluent home market (e.g., Porter, 1990). In summary, the 1950's were a period of high demand and little competition for U.S. industries; a period during which American industry flourished.

During this period, American firms grew complacent about production issues. Production came to be viewed as a cost center, rather than as a potential strategic weapon. Production jobs were perceived by management as dead ends and did not attract the best people. Instead, the focus of corporate America shifted to marketing as new mediums of mass communication such as television, became commonly available (e.g., Dertouzos et al., 1989).

In the late 1960's and early 1970's broad segments of American industry began to lose competitive advantage. For example, America's 1971 balance in merchandise trade was a

deficit for the first time in the twentieth century. The erosion in competitive advantage is also illustrated by the decline of the share of the American car market held by American owned companies that fell from almost 100% in 1955 to 80% by the mid seventies. Underlying the fall in competitive advantage was the fact that America's annual percent change in manufacturing productivity was significantly below those of Japan, Italy, France, Germany, and Canada throughout the 1960's and 1970's (Dertouzos et al., 1989). The overall decline in competitive advantage across a broad spectrum of industries has continued into the 1990's as illustrated by the fact that only 65% of the American car market is currently held by American owned companies, in 1992.

In contrast, Japan emerged from the war with limited capital, few natural resources and a new form of government. In addition, the zaibatsu (giant holding companies) that had fueled prewar industrial growth had been disbanded. In spite of these obstacles, Japanese government, industry, and workers have jointly propelled their economic system to levels rivaling the United States in less than fifty years.

A number of different reasons have been given for Japan's rapid emergence, including the role of government, cultural aspects, and the workforce. It is true that Japanese government has and still does actively support their industries. An example of this is the keiretsu, a

loose connection between numerous firms usually with a financial institution at the center. This type of organization is illegal in the United States. It is also true that cultural homogeneity and a motivated and educated workforce have been contributing factors. Indeed, Japan's principle factor advantage has been, perhaps, their workforce (Porter, 1990).

However, the factor that may have contributed most to Japan's emergence is the management system that they have evolved during the past few decades. Supporting evidence is given by the significant improvements recorded by the numerous U.S. companies that have adopted portions of the Japanese manufacturing practices (e.g., Voss and Robinson, 1987). Although it has proven itself in Japan, the Japanese management system cannot be brought as is into the Western World. Laws are different, the workforce is less educated and more heterogeneous, supplier relationships have historically been adversarial, and geographic distances are much greater. Yet, it is imperative that Western firms identify and incorporate the best features of the Japanese techniques into their own practices. The problem for Western firms is to find adaptations of the Japanese management style that are both efficient and suitable given the environment in which they operate.

1.2 Motivation for the Study

The nature of competition in repetitive manufacturing is radically different than it was fifty years ago. First, there is a clear trend towards globalization, both in regards to manufacturing as well as distribution. An example of this is a U.S. automaker with plants in Mexico and the U.S. competing for European market share against a Japanese automaker with plants in Japan and the U.S. Second, competition is now stronger than ever before. This factor is forcing manufacturers to simultaneously improve quality and service even as there is downward pressure on prices. Finally, product life cycles are becoming shorter and forecasts of demand are becoming less reliable as time based competition is becoming a reality (Stalk and Hout, 1990).

Logistics is one area that has become particularly important with the trend towards globalization (DeRoulet, 1991; Willersdorf, 1991). Indeed, some believe that logistics is a key element for success given recent trends towards free trade in both Europe and North America (Lieb, 1991). Others argue that logistics in itself is not enough; effective planning and control across the entire manufacturing and distribution network is required. In view of this perception, new concepts such as logistics strategy (Perry, 1991), integration of marketing and logistics

(Lambert and Cook, 1990), and supply chain management (Battaglia and Tyndall, 1991) have emerged.

Materials management (MM) is used in this study for the broad picture of planning and control described above. Materials management refers to the coordinated production and transportation of products across a network of plants and distribution outlets working towards the same finished product(s). In the context of a global network, materials management extends to the coordination of plants/outlets located hundreds or even thousands of miles apart. The global characteristic increases the time element, since shipments between plants/outlets may require substantial transit times. As a result, leadtimes are longer and plans cannot be changed as easily. In summary, effective materials management in global networks promises challenging planning and control issues.

Figure 1-1 depicts a generalized view of a global manufacturing network. The small circles within each rectangle represent a manufacturing or production process in which value is added. The idea of value added at a distribution outlet can be interpreted as being closer to the customer. The small squares following the circles represent buffers in which inventory is stored. Henceforth, the word "node" is used to represent both a manufacturing process (a circle) plus its subsequent buffer (a square).

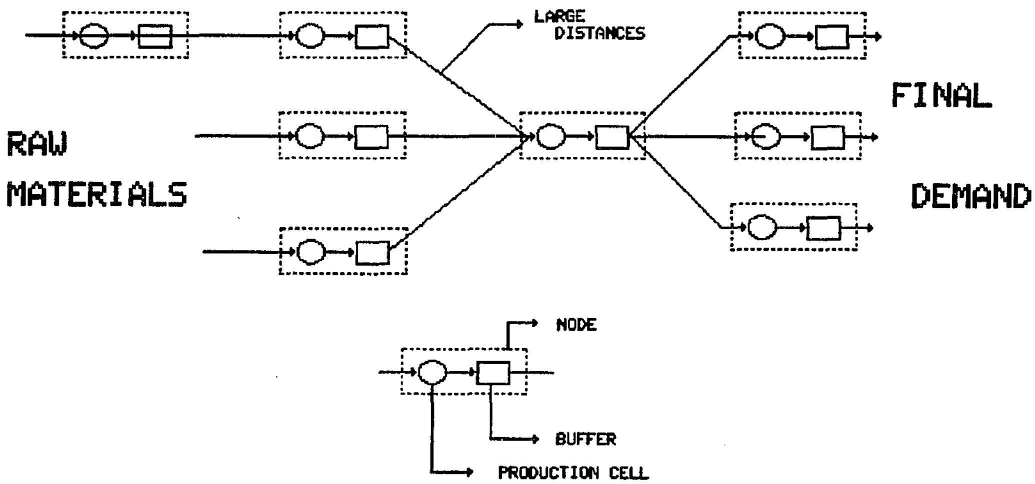


FIGURE 1-1
GENERAL VIEW OF A
GLOBAL MANUFACTURING NETWORK

In terms of a global network, a node can be thought of as a manufacturing process and product store at a single facility. Global networks will tend to have large distances between nodes so that the arcs between the nodes in Figure 1-1 potentially represent large distances and large delivery leadtimes. The function of materials management in a global network is to effectively coordinate production schedules and inventories at all nodes as well as shipment schedules between nodes.

As manufacturing and distribution transcends individual countries, managers must cope with additional uncertainties due to cultural differences and large geographical separations. No longer is the goal to optimize productivity or efficiency at one node, if that was ever the objective. Rather, the operative criteria is to maximize the productivity and efficiency of the entire network. Planning and control becomes complex in an environment of increased globalization and increased time based competition. It can be concluded that firms that excel at materials management have an advantage. In effect, materials management has become a strategic weapon in the bid for competitive advantage.

The new global realities call for new approaches to production and distribution. Manufacturers need planning and control systems that are reactive to day-to-day realities and to volatile shifts in consumer preferences.

They also need effective ways to coordinate production at facilities separated by large distances. Several new conceptual frameworks for planning and control have been proposed, including Lean Production (Womack et al., 1990) and Synchronous Manufacturing (Umble and Srikanth, 1990). Throughout the remainder of this study, the phrase "lean production" will be used to represent the minimal inventory, high product quality, and tightly coordinated manufacturing philosophy espoused by Womack et al. (1990).

It can be concluded that at this moment in time, materials management in global networks is more critical for success than ever before. It can also be concluded that both current and proposed systems are in states of flux as they grapple with the new order of things. This study is directed towards the identification and analysis of different planning and control strategies that may be used for material flows in a global network. The focus of this study is to help identify the "best" control system and to explore the effect of labor costs, holding costs, and demand variation on the values of associated policy parameters.

1.3 Relevant Issues

In order to be effective, a global manufacturer must resolve problems across a broad range of issues. This study will be limited to the consideration of those issues listed in Table 1-1. The issues have been categorized into inventory, production, and global issues solely for the

purposes of discussion. The implication of each issue upon a global manufacturing and distribution network will be discussed briefly in this section. Those factors thought to be most significant will be examined in more detail in Chapter 3.

TABLE 1-1
BASIC ISSUES RELEVANT TO GLOBAL MANUFACTURING

CATEGORY	ISSUE
Inventory	Order Quantity Target Inventory
Production	Yield Rate Delivery Leadtime Production Leadtime Capacity
Global	Network Architecture Coordination Characteristics of Demand

1.3.1 Inventory Issues

Inventory issues of concern to a cooperating global network are those same issues developed in classical multi-echelon inventory theory. After all, the problem under study is essentially a multi-echelon inventory and production problem. Thus, a control strategy for a global network must result in specific order quantities between nodes and specific target inventories at each node.

Target inventories are thought of in two different ways. First, inventories in an order-up-to system typically

remain below the order-up-to quantity, the target inventory. Under this concept, target inventories are more like an upper bound only achieved in the event that demands go to zero. On the other hand, target inventories are also thought of as the desired level of inventory at a node. In this case, inventories may fall below target values or exceed target values in the short term. However, in the long term it is expected that average inventories equal target values. Under both concepts of target inventory a production rule, sometimes called a smoothing rule, is used to adjust production with the goal of maintaining target inventories.

Actual inventories are compared to target values when creating production schedules. Actual inventories in a global manufacturing network include inventory in-transit to a node, in-wait at the node, in-process at the node, or in-buffer at the node. Accordingly, actual inventory at node i is defined to be:

$$Inventory_i = IT_i + IW_i + I_i + IB_i \qquad 1.1$$

where

$Inventory_i$ = Actual inventory at node i and in-transit to node i

IT_i = Work in-transit to node i from a predecessor node(s)

IW_i = Inventory in-wait or waiting to go into production at node i

I_i = Work in-process at node i

IB_i = In-buffer inventory at node i waiting to be shipped

Inventory as used in this study includes both safety stock as well as cycle stock.

1.3.2 Production Issues

Again, the important production issues are those of the classical multi-echelon case including yield rates, delivery leadtimes, production leadtimes, and capacity restrictions. Yield rate, or quality, has been a focal point for manufacturers over the past decade or more (e.g., Crosby, 1979). Numerous manufacturers have improved quality to the point that in many cases the differences between quality leaders and those with average quality is small (e.g., Womack et al., 1990). In effect, everyone must now offer quality in order to survive over the long run in a quality conscious world. For these reasons, and in the interests of model complexity, we will assume that yield rates are 100% for the remainder of this study.

The issues surrounding leadtime are more complex for a global network than in the classical multi-echelon setting. Leadtime now includes a significant element in terms of the time required for delivery. Geographical distances in a

global network can be quite large and shipments may be between different countries. Heavy or low-valued components will frequently need to be moved by rail or by ship. Uncertainty in delivery leadtimes is potentially compounded by uncertainty due to transportation carrier, distance, and customs. Weather may even influence shipping schedules for some carriers in certain regions. Large distances and slow transportation in a lean production environment suggests that a substantial proportion of the inventory in the network may be in-transit at any one moment in time.

Mismanaged materials management in a global network with large delivery leadtimes may result in substantial expediting. Expediting in a global network could result in large inefficiencies and be quite costly. On the other hand, expediting in a global network may simply not be possible. In that event, mismanagement could result in production shutdowns as nodes become starved for materials. One reaction against possible shutdowns is to carry higher inventories than otherwise normal, an expensive alternative. Clearly, large delivery leadtimes make effective planning and control more important than ever.

Production leadtimes and capacity restrictions are also important factors. Expediting at one node may result in increased demands at predecessor nodes or in increased workloads at downstream nodes, both of which may be problematic. For example, increased workloads at downstream

nodes may result in increased production leadtimes at those nodes and cause a production bottleneck. A bottleneck may disrupt the coordination of that portion of the network forcing additional expediting. Excess capacity may be built into the system to minimize expediting, but at high costs. In summary, the long delivery leadtimes associated with global networks make production leadtimes and capacity issues more important than ever.

1.3.3 Global Issues

One primary concern is the architecture or design of the manufacturing network. The possible components for building a network include serial, assembly, and distribution as shown in Figure 1-2. Of course, most global manufacturing networks will be composites made up of many of these blocks combined in various ways. Figure 1-3 shows some simple examples of common types of networks: serial, assembly, distribution, and conjoined. The design of the global network for a particular manufacturing/market objective is outside the realm of this work. In other words, it is assumed that a particular network already exists. Our goal is to find effective methods for planning and control, given a specific network architecture. The phrase "general network" is henceforth used to mean any of the network types shown in Figure 1-3.

Given a particular network, the problem becomes one of how best to coordinate production and transportation within

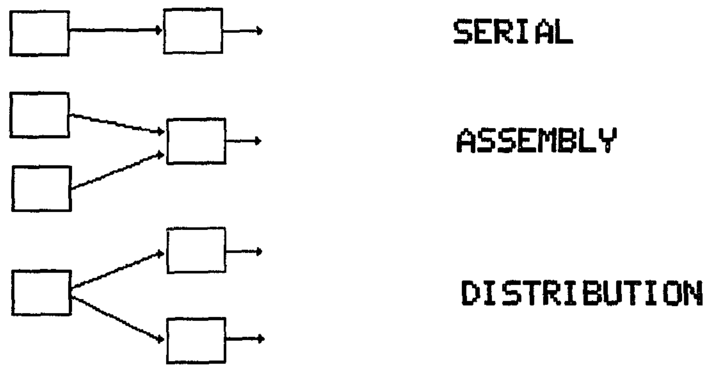


FIGURE 1-2
POSSIBLE COMPONENTS FOR BUILDING A
MANUFACTURING/DISTRIBUTION NETWORK

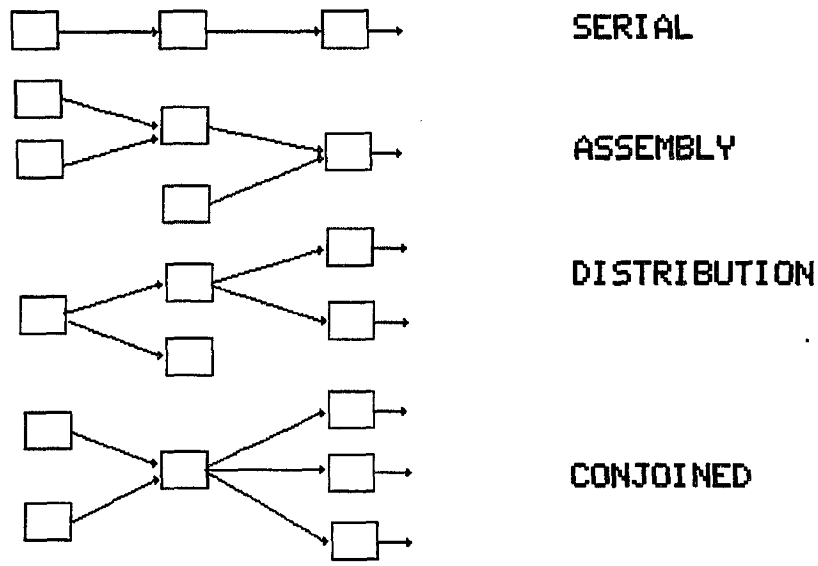


FIGURE 1-3
COMMON TYPES OF MANUFACTURING/DISTRIBUTION NETWORKS

the network. One possible alternative is a push strategy in which forecasts of demand, inventory status, and planned leadtimes are used to centrally schedule material flows.Push strategies have typically been used for manufacturing planning and control by Western firms. Another alternative is to use a pull strategy in which both production at a node and transportation between nodes is based on consumption at downstream nodes. Pull strategies have been used with great success by numerous Japanese manufacturers. Finally, a hybrid strategy with both pull and push components can be used for coordination. Push versus pull strategies will be discussed in depth in Chapter 2, the Literature Review.

Another important issue is in regards to the nature of demand for final product. Manufacturing today is highly competitive and consumer preferences are less predictable than previously. One example of the difficulties of forecasting consumer demand is given by Jordan and Graves (1991) who state that automobile sales forecasts (presumably at General Motors), 1 to 3 years in the future have historically differed from actual sales by 40%. In general, the assumption of a constant, predictable demand is not an appropriate assumption. Manufacturers today must learn to cope with volatile demand and aggressive competition. Effective planning and adaptable control seem to be basic requirements for global manufacturers and distributors in today's highly competitive marketplace.

1.4 Purpose of the Study

Clearly, managers of global manufacturing and distribution networks face complex inventory, production, coordination, and demand issues. Further, the complexity is compounded by the interaction between factors that managers control and those they do not control. For example, material managers may control the amount of overtime, but may have limited or no control over demand. The challenge is to find functional, adaptable strategies that allow a firm to at least compete, but preferably to excel at materials management. The purposes of this study are to:

1) identify the most important factors relevant to materials management;
2) examine different strategies for planning and control of materials management in global networks; and
3) construct a model that can be used to identify the best control rule for a particular network.

Particular emphasis will be placed on factors that a manager can control so that the study should have practical implications.

Chapter 2 summarizes the relevant literature. The literature review touches upon a variety of topics because the problem of interest encompasses a wide range of issues. First, select issues from multi-echelon production and inventory theory and aggregate planning will be reviewed.

Then, push and pull strategies for planning and control will be discussed and several existing and proposed control systems will be reviewed.

Chapter 3 presents two restoration models that subsume many control systems in common use and in the literature. The models have different assumptions regarding lost sales, but both can be applied to networks of any design. They incorporate parameters that can be used to help identify characteristics of the best control rule. Both model formulations are nonlinear and mixed-integer. Relaxing the integrality assumptions and removing setup costs from the model results in a nonlinear formulation that can be solved using commercially available software.

Chapter 4 presents the questions we wish to explore in regards to the nonlinear models presented and then develops the methodology for doing so. Specific hypotheses are generated and statistical tests for testing them are proposed. The methodology of Chapter 4 is applied to a specific global manufacturing network in Chapter 5. Finally, Chapter 6 provides a discussion of the results and identifies areas for further research.

CHAPTER 2 LITERATURE REVIEW

2.1 Introduction

As we have seen, materials management in a global network is a multifaceted problem. It is essentially a multi-echelon production and inventory planning and control problem with potentially large distances between nodes and therefore potentially large delivery leadtimes. Alternatively, the problem can also be viewed as a modification of the aggregate planning problem which explores the strategic issues of overtime/undertime, production schedules, and inventories. First, a review of select topics from the multi-echelon literature is given. Then, some fundamental concepts in hierarchical and aggregate planning are reviewed. The perspective is that it may be possible to extend or modify some of these concepts to apply to materials management in a global network.

Push and pull strategies for manufacturing planning and control are then reviewed and a framework for further discussion is adopted. Subsequently, manufacturing planning and control systems in common use are reviewed and classified along a push/pull gradient. Finally, several newly proposed systems for planning and control are explored. Each of these systems are classified along a push/pull gradient and analyzed for applicability to materials management in a global network.

2.2 Multi-Echelon Production and Inventory Theory

2.2.1 Background

The multi-echelon inventory problem received considerable attention beginning in the 1960's as evidenced in the review paper by Clark (1972). Generally, the early models were specific to one type of network, such as serial, assembly, or distribution. For example, the first model in this area was that of Clark and Scarf (1960), who assumed no setup costs in all but the lowest echelon of a serial network. They derived the optimal stocking policy under periodic review with stochastic demand and deterministic leadtimes. Clark and Scarf (1962) attempted to incorporate setup costs into their 1960 model, but only derived upper and lower bounds on the minimal cost. Numerous additional models have been presented for serial systems as illustrated in the review paper by Goyal and Gunasekaran (1990).

Results for assembly type networks have been more restricted. Schmidt and Nahmias (1985) derive a complicated optimal policy for a simple assembly structure under random demand. It is not apparent that their result can be extended to more complex networks. Schwarz and Schrage (1975) present a model for an assembly network with constant demand and leadtimes using the concept of echelon stock. They assume that the lot size at one stage is an integer multiple of the lot size at its immediate successor stage. They then suggest a myopic policy that examines two stages

at a time. Rosling (1989) proposes that under fairly restrictive assumptions including no setup costs, a more general assembly structure can be approximated by a series model. The application of this strategy to more complex problems remains to be demonstrated.

Distribution networks based upon one or more warehouses serving several retailers, have received substantial attention over the past two decades. Under the assumption that all leadtimes are zero, Bessler and Veinoit (1966) treat a simple arborescent structure and explore the near optimality of one-period policies. Schwarz (1973) derived an optimal policy for identical retailers under deterministic demand. Eppen and Schrage (1981) showed that expected holding and penalty costs were less using a centralized strategy in which (nearly) all inventory is held at the warehouse compared to a decentralized strategy in which no inventory is held at the warehouse.

Deuermeyer and Schwarz (1981) and Svoronous and Zipkin (1988) present models for estimating the service levels of distribution systems. Finally, Rogers and Tsubakitani (1991) recently presented a nonlinear optimization model for determining base stock levels under relatively general assumptions. Their objective was to minimize the total expected penalty costs of backorders subject to a budget constraint. The optimal solution was a newsboy style result

with an additional term: the Lagrangian multiplier for the budget constraint.

There are many models in the literature that incorporate either transportation issues or production issues with inventory policies. One of the earliest was by Baumol and Vinod (1970) who present an inventory theoretic model for determining the optimal choice of transport. They formulated the problem as an inventory problem using the transit time for the replenishment leadtime. Constable and Whybark (1978) present both exact and heuristic procedures for determining inventory reorder points, order quantities, and transportation alternatives for a two stage network. Their objective was to minimize the sum of transportation costs, carrying costs, ordering cost, and expected backorder costs. Burns et al. (1985) examine the coordination of distribution efforts to minimize inventory holding, production, and transportation costs associated with production and distribution. The deterministic nature of their model restricts its applicability.

Control theory represents another approach for modelling production and inventory systems. Control theory involves the use of feedback loops so that the system responds to change. A review of control theory applied to production problems can be found in Axsater (1982). Some recent models of production systems that use control theory include Popplewell and Bonney (1987) and O'Grady and Bonney

(1985). Recent work by Towill (1982;1992) combines the power of control theory with the flexibility of simulation. This work shows promise but is still at an early stage.

2.2.2 Lot Sizing in General Networks

Formulations of the basic multi-echelon lot sizing problem for general networks are given by McLaren (1976) and Heinrich and Schneeweiss (1986). Steinberg and Napier (1980) formulate the problem as a constrained, generalized network. Since it is intuitively easier to follow, a variation of the formulations of McLaren and Henrich and Schneeweiss is given after first listing the assumptions of the model:

1) Known, time-varying demand;
2) All leadtimes equal zero;
3) Periodic review;
4) Finite horizon;
5) No backorders, no lost sales;
6) Demand occurs only at nodes with no successor(s);
7) Uncapacitated;
8) Set up cost is independent of order quantity;
9) Order quantity may vary from period to period;
10) Products move at most one stage per time period; and
11) No more than 1 component from each predecessor is needed to produce a component at a node.

Note that assumption 11 can easily be relaxed; it is used to keep the notation as simple as possible.

The objective of the formulation is to minimize the sum of setup and holding costs across all nodes and for all time periods in the finite horizon. The constraints include inventory balance constraints at each node of the network. There is also a constraint that does not allow the decision variable X_i^t (production quantity) to be greater than zero unless the decision variable δ_i^t is equal to 1. The variable δ_i^t is effectively an on-off switch. It turns the setup cost on if production is scheduled at that node during that period. Otherwise, it turns the setup cost off. The formulation follows:

$$\min \sum_i \sum_t \{\delta_i^t S_i + h_i I_i^t\} \qquad 2.1$$

subject to:

$$I_i^t = I_i^{t-1} + X_i^t - D_i^t \qquad \forall t;\ \forall i \ni s(i)=\emptyset \qquad 2.2$$

$$I_i^t = I_i^{t-1} + X_i^t - \sum_{j \in s(i)} X_j^t \qquad \forall t;\ \forall i \ni s(i)\neq\emptyset \qquad 2.3$$

$$X_i^t \leq \delta_i^t M \qquad \forall t;\ \forall i \qquad 2.4$$

$$\delta_i^t \in \{0,1\} \qquad \forall t;\ \forall i \qquad 2.5$$

$$I_i^t,\ X_i^t \in \{0,1,2,3,\ldots\} \qquad \forall t;\ \forall i \qquad 2.6$$

where

D_i^t = Demand for finished product at node i during t

I_i^t = Inventory at node i at end of period t

X_i^t = Amount of product produced at node i during t

S_i = Setup cost at node i

h_i = Holding cost at node i per item per unit time

$$\delta_i^t = \begin{cases} 0 & \text{if no setup for component } i \text{ occurs during } t \\ 1 & \text{otherwise} \end{cases}$$

$s(i)$ = Set of immediate, downstream nodes to node i

M = A large number

The formulation above is an integer programming formulation that can be solved using standard procedures. However, as presented, its application to actual problems is quite limited based on the number and severity of assumptions. The formulation can readily be modified to incorporate constant leadtimes (McClain et al., 1982; Afentakis and Gavish, 1986) and capacity constraints at each node (Billington et al., 1983; Gavish and Johnson, 1990; Pochet and Wolsey, 1991). In fact, Hackman and Leachman (1989) introduce a general framework for the formulation of deterministic models for general networks.

A version of the formulation given by 2.1 - 2.6 that allows constant production leadtimes PL_i and restricts capacities C_i at each node i is given below. The model uses the idea that production scheduled ($t-PL_i$) time units ago at node i, namely X_i^{t-PLi}, will complete processing during period t. Importantly, an additional assumption of the formulation is that production at different nodes is perfectly coordinated by a control rule such as Materials Requirements Planning. Using this assumption, a node is never starved for materials so that the assumption of constant leadtimes is supported.

$$\min \sum_i \sum_t \{\delta_i^t S_i + h_i I_i^t\} \qquad 2.7$$

subject to:

$$I_i^t = I_i^{t-1} + X_i^{t-PL_i} - D_i^t \qquad \forall t;\ \forall i \ni s(i)=\emptyset \qquad 2.8$$

$$I_i^t = I_i^{t-1} + X_i^{t-PL_i} - \sum_{j \in s(i)} X_j^t \qquad \forall t;\ \forall i \ni s(i)\neq\emptyset \qquad 2.9$$

$$X_i^t \le \delta_i^t C_i \qquad \forall t;\ \forall i \qquad 2.10$$

$$\delta_i^t \in \{0,1\} \qquad \forall t;\ \forall i \qquad 2.11$$

$$I_i^t,\ X_i^t \in \{0,1,2,3,\ldots\} \qquad \forall t;\ \forall i \qquad 2.12$$

where

D_i^t = Demand for finished product at node i during t

I_i^t = Inventory at node i at end of period t

X_i^t = Production start at node i during period t

S_i = Setup cost at node i

h_i = Holding cost at node i per item per unit time

PL_i = Unavoidable delay between production and availability of an item at node i

C_i = Upper bound on capacity at node i

$$\delta_i^t = \begin{cases} 0 & \text{if no setup for component i occurs during t} \\ 1 & \text{otherwise} \end{cases}$$

s(i) = Set of immediate, downstream nodes to node i

This formulation is also an integer programming problem and can be solved using standard procedures.

Constraints 2.8 and 2.9 insure that production occurs PL_i time units in advance. Constraint 2.10 insures that scheduled production does not exceed capacity at the node. Jointly, the effect of constraints 2.8, 2.9, and 2.10 is to

shift production that exceeds capacity to earlier time periods. The net effect is an increase in leadtimes, inventories, and holding costs. Therefore, this model implicitly increases production leadtimes if needed to avoid capacity restrictions. This idea was pointed out by Billington et al. (1983) in the following model that also incorporates overtime and undertime considerations:

$$\min \sum_i \sum_t \{\delta_i^t S_i + h_i I_i^t\} + \sum_k \sum_t \{CO_k^t O_k^t + CU_k^t U_k^t\} \qquad 2.13$$

subject to:

$$I_i^t = I_i^{t-1} + y_i X_i^{t-PLi} - D_i^t \qquad \forall t;\ \forall i \ni s(i)=\emptyset \qquad 2.14$$

$$I_i^t = I_i^{t-1} + y_i X_i^{t-PLi} - \sum_{j \in s(i)} X_j^t \qquad \forall t;\ \forall i \ni s(i)\neq\emptyset \qquad 2.15$$

$$\sum_i (b_{ik} X_i^t + s_{ik}\delta_i^t) + U_k^t - O_k^t = R_k^t \qquad \forall t;\ \forall k \qquad 2.16$$

$$X_i^t \le \delta_i^t M \qquad \forall t;\ \forall i \qquad 2.17$$

$$\delta_i^t \in \{0,1\} \qquad \forall t;\ \forall i \qquad 2.18$$

$$I_i^t,\ X_i^t \in \{0,1,2,3,\ldots\} \qquad \forall t;\ \forall i \qquad 2.19$$

where

D_i^t = Demand for finished product at node i during t

I_i^t = Inventory at node i at end of period t

X_i^t = Production start at node i during period t

R_k^t = Capacity in units of time at k during t

S_i = Setup cost at node i

h_i = Holding cost at node i per item per unit time

$$\delta_i^t = \begin{cases} 0 & \text{if no setup for component i occurs during t} \\ 1 & \text{otherwise} \end{cases}$$

O_k^t = Amount of overtime at k during t

CO_k^t = Cost of overtime at k during t

U_k^t = Amount of undertime at k during t

CU_k^t = Cost of undertime at k during t

y_i = Average yield fraction

PL_i = Fixed minimum production leadtime

b_{ik} = Time required to produce one unit of i at k

s_{ik} = Time required to set up node k for item i

M = A large number

Billington et al. (1983) assume that PL_i represents a fixed, minimum value for leadtime at node i. The idea is that PL_i will be constant as long as scheduled production is less than capacity C_i^t. The term $y_i X_i^{t-PL_i}$ in constraints 2.14 and 2.15 corresponds to the percent of $X_i^{t-PL_i}$ that is, on average, of high enough quality. Constraint 2.16 states that the sum of production time plus setup time plus undertime minus overtime equals regular capacity. Note that the objective of minimizing costs will insure that only one of undertime or overtime will be nonzero for any period. Constraints 2.14, 2.15, and 2.16 work in tandem to shift production to earlier periods as needed to avoid actual production bottlenecks. Note that another constraint can easily be added if there is an absolute upper limit on overtime. The net effect of doing this is to increase the leadtime and work-in-process (WIP) inventories by shifting

work to earlier time periods as needed to avoid capacity constraints.

The model by Billington et al. assumes known demands, a general network, and perfect coordination by the control system. The model adjusts production schedules for capacity bottlenecks by shifting production to earlier periods as needed to minimize the sum of setup, holding, overtime, and undertime costs. This mechanism allows the model to implicitly change leadtimes if needed to avoid production bottlenecks. In summary, this model is an excellent one to begin to think in terms of materials management in global networks.

None of the above formulations make allowances for uncertainty in demand nor do they incorporate uncertainty in leadtimes. Yet, we know that uncertainty is a fact for many if not all global manufacturers. The question becomes one of whether or not the formulations given above are adequate, even for first order approximations given prevalent levels of uncertainty.

Numerous authors support the idea that uncertainty in demand can have a large effect on system performance (e.g., DeBodt et al., 1982; Grasso and Taylor, 1984). A formulation similar to that given by equations 2.7 - 2.12 but limited to two stages, is given by Beale et al. (1980). Their model allows demand to be stochastic. They present an approximate solution methodology, using techniques from

stochastic programming, under the assumption that demand is normally distributed with constant mean and variance.

Gong and Matsuo (1991) present a model for a serial system with stationary random yields and demands over infinite time periods. Their formulation seeks to minimize WIP rather than costs per se:

$$\min \lim_{t\to\infty} \sum_i E\{I_i^t\} \qquad 2.20$$

subject to:

$I_i^t = I_i^{t-1} + Z_i^t - D^t$	$\forall t;\ \forall i \ni s(i)=\emptyset$	2.21
$I_i^t = I_i^{t-1} + Z_i^t - X_j^t$	$\forall t;\ \forall i \ni s(i)\neq\emptyset;\ j\in s(i)$	2.22
$X_i^t \le I_{i-1}^{t-1}$	$\forall t; \forall i$	2.23
$X_i^t \le C_i$	$\forall t; \forall i$	2.24
$X_i^t \ge 0$	$\forall t; \forall i$	2.25
$P(I_i^t \ge D^t) \ge \delta$	$\forall t; \forall i$	2.26

where

D^t	= Demand at final node during t
I_i^t	= Inventory at node i at end of period t
X_i^t	= Production start at node i during t
C_i	= Capacity at node i
Z_i^t	= Yield at node i during t
s(i)	= Set of immediate, downstream nodes to node i

Gong and Matsuo's solution procedure is to first, ignore constraints 2.23 - 2.26 and derive steady state covariances of WIP and production quantities as functions of

parameters of control rules. Second, constraints 2.23 - 2.25 are converted into chance constraints. Note that constraint 2.26 is already a chance constraint. Finally, the feasibility of the derived production rule is restored when implemented. The model developed by Gong and Matsuo shows promise but is difficult to apply to general networks.

It has proven to be quite difficult to develop a model for a general network with both random demands and leadtimes. In particular, models incorporating stochastic leadtimes are generally limited to single stages (Kaplan, 1970; Nevison and Burstein, 1984; Anderson, 1989) or to the use of heuristics (e.g., Whybark and Williams, 1976; Nevison, 1985). For that reason, we now review hierarchical and aggregate planning.

2.2.3 Hierarchical Planning and Aggregate Planning

In general, the problem of manufacturing planning and control in a large firm is so complex that no single model incorporates all of the relevant factors. One response to the situation has been to schedule production at a node based on final demands, adjusted for leadtimes. Under this planning process, needs at downstream, intermediate nodes are effectively ignored. This strategy is commonly referred to as base stock (Silver, 1985).

Another mechanism for coping with the overall levels of complexity falls under the concept of hierarchical planning (e.g., Hax and Meal, 1975; Krajewski and Ritzman, 1977).

Hierarchical planning is the partitioning of a problem into procedures for making decisions at separate levels such as strategic, tactical, and operational. Hierarchical systems attempt to find solutions that:

1) perform well in regards to total costs;
2) are consistent in that lower level constraints fall within the bounds of higher level decisions; and
3) are implementable by the firm.

Hierarchical systems strive for good solutions but do not guarantee optimal solutions.

Some interesting work has occurred in the area of hierarchical planning recently. For example, Cohen and Lee (1988) develop a hierarchical model that incorporates uncertainty in both production and distribution. Several approximate, stochastic submodels are linked and a heuristic optimization procedure is introduced. This type of approach shows promise, but at this point is still in its infancy.

From within a hierarchical framework, aggregate planing can be viewed as a strategy for making higher level decisions in regards to production/shipment schedules, overtime/undertime, and hiring/firing policies (e.g., Nahmias, 1989; Vollman et al., 1988). The earliest work on aggregate planning appeared in Holt et al. (1955) and Holt et al. (1956). The basic concepts of the model presented in the book by Holt et al. (1960) are relevant to our work and will now be discussed.

Holt et al. (1960) present a model that helps to distinguish between the three possible coping mechanisms for varying demand:

1) maintain constant production by hiring and firing workers as needed (chase strategy);
2) maintain a constant work force and use overtime and undertime to vary production; and/or
3) maintain a constant workforce and production rate, but allow inventories to fluctuate.

The authors point out that the optimal strategy may be some combination of these three mechanisms.

Expressions for the costs used in the model, with slight modifications in notation, are:

$C_1W^t + C_{13}$	Regular Payroll Costs	2.27
$C_2(W^t - W^{t-1} - C_{11})^2$	Hiring & Firing Costs	2.28
$C_3(X^t - C_4W^t)^2 + C_5X^t - C_6W^t + C_{12}X^tW^t$	Overtime Costs	2.29
$C_7(TI^t - I^t)^2$	Inventory Related Costs	2.30

where:

C_i, C_{jk} are constants specific to the firm

W^t = Number of workers

X^t = Production scheduled

TI^t = Target inventory (or optimal net inventory)

I^t = Inventory position

Notice that regular payroll costs (2.27) are linear, whereas hiring and firing costs (2.28), overtime costs (2.29), and inventory related costs (2.30) each contain a quadratic (nonlinear) term. For each of these last three costs, the authors make the argument for a "U-shaped" curve. They propose that a quadratic form is a suitable approximation of reality yet is mathematically tractable. Their model follows:

$$\min \sum_{t} \left\{ C_1W^t + C_{13} + C_2(W^t - W^{t-1} - C_{11})^2 + C_3(X^t - C_4W^t)^2 + C_5X^t - C_6W^t + C_{12}X^tW^t + C_7(TI^t - I^t)^2 \right\} \qquad 2.31$$

subject to:

$$TI^t = C_8 + C_9S^t \qquad 2.32$$

$$I^{t-1} + X^t - S^t = I^t \qquad t = 1, 2, \ldots, T \qquad 2.33$$

where S^t is equal to the aggregate order rate.

In particular, notice the final term in the objective function: $C_7(TI^t - I^t)^2$. In effect, the model includes inventory costs only as actual inventories deviate from target inventories. Incidently, target inventories are derived externally to the model. Notice that the penalty associated with a deviation of actual inventories from target is severe due to the quadratic form of the term.

Other authors assume linear costs in models of aggregate planning. Hansmann and Hess (1960) present a linear programming formulation with the following decision variables: work force, production level, inventory level,

number of workers to be hired, number to be fired, overtime, undertime, and number of units to subcontract out. Their formulation is to minimize the sum of these costs subject to balance constraints on work force, production, and inventories. In contrast to Holt et al. (1960) this model specifically incorporates holding costs on inventories. Chung and Krajewski (1984) extend this model by including setup costs, resulting in a mixed-integer formulation.

More recent work on aggregate planning focuses on disaggregation (Bitran and Hax, 1981), or multiple products (Bergstrom and Smith, 1970), or include marketing/financial variables (Damon and Schramm, 1972; and Leitch, 1974). Before examining control systems in common use, we explore the concepts of push and pull strategies for planning and control.

2.3 Push versus Pull

Strategies for manufacturing planning and control are frequently classified as either push or pull. The push strategy has been the primary strategy in use in North America and Europe, although components of the pull system have been around since the days of Henry Ford (Womack et al., 1990). The Japanese, notably Toyota, were the first to develop the pull system into its current, highly polished form (e.g., Kimura and Terada, 1981; Monden, 1983). An overview of both push and pull strategies as they are commonly thought of, follows.

Push systems require a forecast of demand over a planning horizon. A centralized controller uses the forecast along with global data on inventory status and planned leadtimes to schedule material flows. Thus, production schedules and batch sizes are under the control of a central authority with access to global information. The central authority controls the release of orders to upstream stages near raw material input. Subsequent stages are responsible for processing any unfinished products that have come to them from upstream stages. Effectively, push systems schedule throughput and measure inventory. In turn, inventory is used in feedback loops to adjust subsequent production schedules.

In contrast, pull systems do not rely on long range forecasts and are generally viewed as being reactive with decentralized (local) control. Production at each stage is scheduled based upon consumption at the downstream stage(s). Pull systems are frequently associated with minimal inventories. In turn, minimal inventories require high product quality, short leadtimes, and tight coordination between adjacent work centers. Effectively, pull systems control inventory and measure throughput. In turn, throughput may be used to adjust inventory.

Variation in demand, machine reliability, leadtimes, and/or yield rates create a problem for both push and pull systems. Push systems tend to cope with uncertainty by

using safety stock and safety leadtimes. Nonetheless, push systems frequently result in inventory shortages at some stages and accumulations at others. Pull systems cope with uncertainty by reducing leadtimes and by freezing the production schedule for a short period of time such as 30 days. Further, pull systems do not respond well in general to lumpy demand. Incidently, both push and pull systems can be quite difficult for a manufacturer to implement.

The distinction between push and pull systems has been ascribed to order release (Karmarkar, 1986), to the use of global versus local information (Silver and Peterson, 1985), and to the degree of centralization (Takahaski et al., 1987). More recently, Pyke and Cohen (1990) state that it is misleading to try to classify a system as either push or pull. They argue that push and pull are characteristics of components of a system, rather than of the system as a whole. They propose a framework for classifying systems based on:

1) information used by the decision maker;
2) who has authority over the decision.

A summary of their work is given in Table 2-1.

The classification scheme provided by Pyke and Cohen supports the commonly held view that push systems tend to use global information to centrally plan material flows. The scheme also supports the commonly held view that localized data from downstream stages is both the

TABLE 2-1
DISTINGUISHING FEATURES BETWEEN PUSH AND PULL
TAKEN FROM PYKE AND COHEN (1990)

	PUSH	PULL
Authority	Upstream	Downstream
Information	Local to Upstream Global	Local to Downstream

information and authority required for material flows in pull systems. Cohen and Pyke suggest that each of the following components of a system be analyzed separately in terms of sources of authority and information:

1) determination of batch size;
2) timing of a production request;
3) setting of dispatch rules; and
4) interference mechanisms for handling emergency orders.

In summary, Pyke and Cohen suggest that each of these material control decisions can be classified along a push/pull gradient, whereas it may be difficult to classify the entire system along a push/pull gradient. This classification scheme is useful in analyzing systems and will be referred to again.

2.4 Systems in Common Use

2.4.1 Introduction

Manufacturers' urgent need for effective planning and control has resulted in the evolution of several systems

that are now in common use, in particular Materials Requirements Planning and kanban. Although neither of these promises optimality, both are highly functional in specific settings and both have received considerable attention in the literature. Both Materials Requirements Planning and kanban will be reviewed.

2.4.2 Materials Requirements Planning

Materials Requirements Planning (MRP) concepts have been around since the 1960's according to Anderson et al. (1982). The popular books by academicians New (1974), Wight (1974), and Orlicky (1975) legitimized and publicized the concepts. MRP and its successor (MRPII) represent a set of procedures for converting demand forecasts over a planning horizon into a formal schedule for each part. A centralized planning system uses demand forecasts, global data on inventory status, and planned leadtimes to schedule work at each stage. In terms of the classification scheme of Pyke and Cohen (1990), batch size and timing of production are clearly based on a push strategy. On the other hand, control for sequencing and interference for emergencies is frequently local-pull characteristics. Although MRP is commonly viewed as a push system, it has some characteristics of a pull system.

Numerous firms implemented MRP in the 1970's. Some were successful (Schroeder et al., 1981), but many were not (Woolsey, 1979; Kanet, 1990). At first, the failures were

attributed to insufficient education of the workforce and to the generally poor accuracy of the data (Cox and Clark, 1984). However, it has since been pointed out that some of the basic assumptions underlying MRP are not valid (e.g., Whybark and Williams, 1976; Karmarkar, 1989). For example, MRP logic assumes constant leadtimes, yet leadtimes tend to vary with the amount of production scheduled. Therefore, MRP does not always generate feasible plans. In addition, MRP has been criticized as being top heavy with paperwork and nonresponsive to changes at the shop floor (Cox and Clark, 1984; Baer, 1991). The slow response to events on the shop floor tend to result in unplanned inventory shortages and accumulations, a major criticism of MRP. In summary, although MRP has worked in specific circumstances, it remains far from a panacea.

The lure of MRP was that of near optimality using central planning and control. Yet, the manifestations of MRP were associated with inventory shortages and accumulations. Early on, researchers began working on methods of calculating order quantities (e.g., see Chapter 12 of Vollmann et al., 1988) in an attempt to improve coordination. The restrictive assumptions and complexity of calculations for optimal models have resulted in several heuristic lot sizing techniques (e.g., Groff, 1979; Gaither, 1981; Gaither, 1983). It is interesting to note the recent survey by Haddock and Hubicki (1989) that demonstrates that

the simplest lot sizing rules (lot for lot and fixed order quantity) are those currently used most often by MRP practitioners. Apparently, understandability and simplicity have superseded the more complex methods in terms of use.

Uncertainty in demand and in leadtimes is problematical for MRP practitioners. The usual process to compensate for uncertainty is through the use of safety stocks although safety leadtimes and excess capacity are also used (Schmitt, 1984). Safety stocks are the inventory remaining should actual demand equal forecast demand and actual leadtimes equal assumed leadtimes. For a specific situation, Carlson and Yano (1981) conclude that safety stocks are more appropriate for the final product than for intermediate components. Indeed, the use of safety stocks at the final stage seems to be a common practice for MRP practitioners in assembly networks (Lambrecht et al., 1981). Component commonality or distribution networks may call for a different strategy.

In summary, the final verdict regarding the use of MRP for manufacturing planning and control has not yet been heard. MRP has proven successful in certain applications and problematic in others. Many of those using MRP successfully have found it to require substantial modifications to conform to their particular manufacturing environment. Those components of MRP that seem most applicable to a global network include its ability to handle

lumpy demand and its potential to coordinate disparate, geographically separated nodes.

2.4.3 Kanban

The manufacturing philosophy that originated in Japan within the past few decades is commonly referred to as Just-In-Time (JIT). JIT has been widely publicized in numerous books including the early ones by Schonberger (1982), Hall (1983), and Hay (1988). JIT emphasizes the elimination of waste and continual improvement (kaizen). Defects and inventory are both viewed as wasteful, whereas reduced setup costs are mandatory. Successful JIT practitioners focus on continual improvements in quality and reductions in inventory. Among other things, JIT calls for total quality management, worker involvement, and small batch sizes.

Kanban is used in this study to represent the specific control system evolved by Japanese automakers such as Toyota, under the JIT philosophy. Under kanban, the size and number of inventory containers at a work center is centrally determined. The size of a container represents the order quantity and the size of the container times the number of containers equates to target inventory at the work center. In this case, target inventory at a work center is also the upper bound on inventory at the work center.

Under kanban, nothing is produced at a work center until triggered by inventory removal from a subsequent work center. Information flows from a work center to a

predecessor work center by cards; see Huang et al. (1983) for a description of the process. A card authorizes the predecessor work center to produce the number of products stated on the card. Overall, the system operates with a fixed number of cards, thus a fixed upper limit on inventory. The kanban objective of continually striving for lower inventories is achieved by continually reducing the number of cards, the number of products stated on each card, or both.

Adjacent work centers are tightly coupled under kanban. Information flow from a work center to the predecessor work center(s) is rapid due to low inventories and short leadtimes. In fact, low inventories and short leadtimes are required for kanban to work effectively (Hall, 1983; Monden, 1983). In a minimal inventory environment, product quality becomes a requirement. Under kanban, quality is deemed sufficiently important that each individual on the production line has the ability to shut down the entire line should quality fall (Monden, 1983).

Supplier relationships for kanban practitioners differ significantly from the typical adversarial supplier relationships in the Western world (e.g., Manoochehri, 1984). Kanban practitioners seek to establish long term relationships with a few good suppliers. Suppliers tend to be located nearby and make frequent, small deliveries of high quality products. Deliveries occur as frequently as

ten times per day at Toyota (Monden, 1983). The consistent high quality of products and stable production schedules tend to eliminate the need for functions such as receiving, inspection, and associated paper work. Requirements for kanban supplier relationships include reliability, consistency, small batch sizes, and short delivery leadtimes.

In terms of the classification scheme of Pyke and Cohen (1990), the timing of production under kanban is clearly a pull characteristic since the downstream stage is both the authority and information required to trigger production. Batch sizes (container sizes) are established centrally, a push characteristic, yet consumption at a downstream stage triggers the number of batches (containers) to produce, a pull characteristic. Both priorities and expediting procedures are centrally established in kanban. In summary, kanban is perceived as a pull system, yet it has some significant push components.

Recently, a number of studies have analyzed ways to establish the number of and/or size of inventory containers at each work center under a kanban system. Bitran and Chang (1987) present a math programming model for a kanban system in a deterministic multi-stage capacitated serial production system. Both the complexity of the model and the underlying assumptions of known demand and leadtimes limit its use in practice.

Moeeni and Chang (1990) present a simplified heuristic for determining a lower bound on the number of containers needed in a multistage, uncapacitated assembly tree structure under deterministic demand. Philipoom et al. (1987) look at the effect of variation and autocorrelation of processing times on the number of kanbans needed. Under the assumption that work centers can be decoupled and modelled separately, they propose a simulation approach for determining the number of containers at a work center. Although these models can be used in specific situations, underlying assumptions prevent application to a wide range of environments.

Deleersnyder et al. (1989) incorporate variability in demand and machine reliability into an analytical model of a serial production system. They show that production schedules are very sensitive to variation in both machine reliability and in demand, especially as inventory is reduced to the feasible minimum. This supports the common view that kanban systems operating at minimal inventory do not handle uncertainty well. In fact, manufacturers using kanban handle uncertainty by freezing the production schedules for short periods of time (Monden, 1983). Deleersnyder et al. show that the addition of small amounts of safety stock at all work centers can have a beneficial effect on the performance of the entire network.

The use of kanban requires that each work center must have excess capacity. For if a work center does not have excess capacity, it will never be able to catch up once it falls behind. Minimal inventories in the system would then guarantee long lasting shortages at downstream stages. So and Pinault (1988) use work center capacity in a queuing model that estimates the amount of safety stock needed at each stage in order to maintain a specific service level. The model applies only to a serial production line under kanban type control. It uses safety stock to protect against variation in processing times, machine breakdowns, and demand fluctuations.

The success of kanban as demonstrated by Japanese automakers has prompted the question of whether or not it can be implemented in North America and Europe. In particular, Huang et al. (1983) look at the effects of variable processing times, variable master production scheduling, and imbalances between production stages on kanban type control. They conclude that kanban cannot automatically be applied to American firms. Variability in processing times and demand rates have a definite impact on production. In summary, substantial changes in the production system must usually be made before implementing kanban. Sarker and Harris (1988) support this conclusion in their study on the effect of line imbalance on kanban.

Krajewski et al. (1987) compare kanban with MRP using a list of factors thought by managers to be important to manufacturing effectiveness. They conclude that a reorder point system performs fundamentally as well as kanban. Kanban is nothing more than a convenient way to implement a small lot size and high quality strategy. Rees et al. (1989) also support the idea that MRP can work well in their comparison of MRP lot-for-lot with kanban in an ill-structured production operation. Basically, these authors focus on the potential ineffectiveness of kanban in an environment different than that in which it has been applied. They suggest that the group technology and layout themes implemented by Japanese automakers are necessary for the full success of kanban. They also point out that MRP handles lumpy demand more readily than does kanban.

2.4.4 Summary

Both MRP and kanban work well in specific instances. MRP appears to be able to handle lumpy demand and should be able to coordinate disparate production facilities effectively. MRP also applies to a wider range of manufacturing facilities than does kanban. Yet, MRP seems to be associated with excess safety stocks and nonresponsiveness in many implementations.

On the other hand, kanban has proven to be very effective in certain high-volume, repetitive manufacturing environments. The drive towards implementation of kanban

results in reduced leadtimes, reduced inventories, and higher levels of product quality. Each of these are desirable goals in itself. Yet, kanban requires low set up costs, short leadtimes, and stable production schedules-features that do not normally exist for many manufacturing environments. We have seen that global networks tend to have long leadtimes. Therefore, without modification, kanban can not be expected to perform well in global networks.

In summary, it must be concluded that there are components of both MRP and kanban that are worthy. An objective with possible merit is to combine the best components of both systems. Table 2-2 shows the levels of push and pull for MRP and kanban as given by Pyke and Cohen (1990). This information will be used as a basis of comparison for several newly proposed control strategies that are discussed in the next section.

2.5 Newly Proposed Systems

One fundamental criticism of MRP systems are that they are not responsive to uncertainty so that inventory shortages and accumulations commonly occur. Significant amounts of safety stock tend to be required to smooth production systems under MRP control. In turn, kanban systems react poorly if there are significant set up costs or large fluctuations in demand. Further, classical kanban from Japan can not readily be applied in the Western world.

TABLE 2-2
LEVEL OF PUSH VERSUS PULL BY TYPE OF DECISION
FOR MRP AND KANBAN
TAKEN FROM PYKE AND COHEN (1990)

	MRP			
DECISION	AUTHORITY	INFORMATION	PUSH	PULL
Batch Size	Upstream	Global	▲	
Timing	Upstream	Global	▲	
Priorities	Downstream	Local		▲
Interference	Upstream	Local		▲

	KANBAN			
DECISION	AUTHORITY	INFORMATION	PUSH	PULL
Batch Size	Downstream Upstream	Local Global	 ▲	▲
Timing	Downstream	Local		▲
Priorities	Upstream	Global	▲	
Interference	Upstream	Global	▲	

Some believe the best planning and control system is dependent on the particular manufacturing environment of that firm (e.g., Ptak, 1991; Veatch and Wein, 1991). Still other authors have proposed new and different control strategies. These control strategies resemble MRP or kanban in certain respects but differ in others. Several of these recently proposed strategies will now be reviewed in the order in which they appear in the literature. Each will be analyzed in terms of its push and pull components and viewed in terms of its potential applicability to a global network.

Note that none of the systems promise optimality. A brief preview of these rules is provided in Table 2-3.

TABLE 2-3
SUMMARY OF PUSH VERSUS PULL LEVELS
FOR DIFFERENT CONTROL RULES
BATCH SIZE AND TIMING DECISIONS

RULE	DECISION	AUTHORITY	INFORMATION	PUSH-PULL				
MRP	Batch Size	Upstream	Global	▲				
	Timing	Upstream	Global	▲				
Kanban	Batch Size	Both	Both				▲	
	Timing	Downstream	Local					▲
Periodic Pull	Batch Size	Both	Both				▲	
	Timing	Downstream	Local					▲
Drum-Buffer-Rope	Batch Size	Upstream	Global	▲				
	Timing	Upstream	Global	▲				
Single Feedback	Batch Size	Both	Both				▲	
	Timing	Downstream	Local					▲
Multiple Feedback	Batch Size	Both	Both			*		
	Timing	Both	Both			*		
CONWIP	Batch Size	Both	Global		▲			
	Timing	Downstream	Global			▲		
Restoration	Batch Size	Both	Both			*		
	Timing	Both	Both			*		
Hodgson & Wang	Batch Size	Upstream	Global	▲				
	Timing	Both	Both			▲		
Production Authoriz. Cards	Batch Size	Both	Both			*		
	Timing	Both	Both			*		

* These values can range from push to pull depending on specific values for model parameters.

2.5.1 Periodic Pull System

As geographical separations become larger, it becomes impractical to move materials management data manually as currently occurs in the Japanese version of kanban. Kim

(1985) introduces the Periodic Pull System (PPS), as an operating variant of kanban particularly suited to larger geographical distances between stages. PPS is identical to kanban except that computers move information in contrast to the cards now used in the Japanese version of kanban. Information moves faster on a computer network than through manual exchanges of cards so that PPS should result in reduced leadtimes, an issue of utmost importance in a global network. Since PPS is kanban, it mirrors the push/pull classifications of kanban discussed earlier.

2.5.2 Drum-Buffer-Rope

Goldratt and Fox (1986) propose the drum-buffer-rope system in which a production line is effectively decomposed into two parts by a bottleneck. Their perspective is that the productivity of the entire line is a function of the productivity at the bottleneck. They therefore seek to maximize productivity at the bottleneck resource. The drum paces everything upstream of the bottleneck to the bottleneck itself. The rope symbolizes the connection from the bottleneck to the input at upstream stages. Production is pushed along stages upstream of the bottleneck according to a centrally determined schedule using global information. Production is also pushed along stages downstream of the bottleneck, in a manner similar to MRP. Therefore, each stage downstream of the bottleneck is responsible for completing any work that comes to it.

Drum-buffer-rope uses a centrally based push system to schedule the timing of and batch sizes of production. Presumably, priority setting and expediting in drum-buffer-rope are similar to those in MRP so that some local scheduling can occur within the centrally determined guidelines. Thus, priority setting and expediting each have both push and pull characteristics. It would appear as if drum-buffer-rope is primarily a push strategy.

It is interesting to note that Drum-Buffer-Rope can easily be altered in the direction of a more pull oriented system. After identifying the bottleneck resource, a target inventory can be set at the buffer in front of the bottleneck based on maximization of bottleneck productivity. Kanban type pull control rules can then be used above the bottleneck based on demand measured against target inventory at the bottleneck. Kanban type pull control could also be used below the bottleneck.

In terms of a global network, Drum-Buffer-Rope illustrates the importance of recognizing any permanent bottleneck. Drum-buffer-rope philosophy becomes muddled when applied to a global network in which several bottlenecks may exist simultaneously, or one in which the bottleneck can shift stochastically in time from one node to another. Drum-buffer-rope presents a strategy for coping with a network with a significant and long-lasting

imbalance; it does not appear to give prescriptions for a network in reasonable balance.

2.5.3 Single Feedback versus Multiple Feedback

Takahashi et al. (1987) explore two different "push" strategies for a three stage serial production line. One strategy uses a "single feedback method" in which production at a stage is controlled based upon inventory at the subsequent buffer. Actually, this "single feedback method" looks very similar to kanban; the authors do not provide sufficient information to logically separate the two. The second strategy is a "multi-feedback method" in which production at a stage is based upon inventory at multiple downstream buffers. This type of system is similar to the restoration strategy proposed by Tang (1990), which will be discussed in depth shortly.

Takahashi et al. focus on looking at the choice of a control system on the amplifications (highs and lows) in inventory at various stages. Their results suggest that the choice of a control system is intricately related to forecast errors, downtime, and loading ratio. They suggest that estimates of these quantities should be made before choosing a control strategy, but do not specify how to estimate these parameters. These authors conclude that uncertainty in demand and/or leadtime can have a significant effect on the effectiveness of a control strategy.

2.5.4 Constant Work-In-Process

Spearman et al. (1990) propose an interesting production system dubbed CONWIP for CONstant Work In Process. For a serial line, CONWIP is a generalization of kanban in which a production card for each finished product is assigned to a production line. Recall that kanban assigns a card for each container at a stage. Under CONWIP, the number of cards for the line is fixed so that there is a maximum number of items in-process at any moment in time. A fixed number of production cards implies a (nearly) constant WIP over the whole line. In turn, a constant WIP should result in (nearly) constant leadtimes and better predictability regarding throughput. The authors suggest that CONWIP retains many of the advantages of kanban since both adhere to the philosophy of "lean production", yet CONWIP applies to a broader range of manufacturing environments than does kanban.

Under CONWIP, jobs are released to upstream stages based primarily on consumption at the final stage(s), suggesting an overall pull strategy for both the size and timing of batches. Yet, once released, a job is pushed through stages according to a first come first served priority rule-a push trait. Another push trait of CONWIP is that it requires a centrally determined number of production cards, or amount of WIP, for the production line. Although CONWIP is introduced as a pull strategy, it has significant

push components. In several instances, Spearman et al. liken CONWIP to Drum-Buffer-Rope, a system we have already depicted as being based primarily on a push strategy.

Spearman et al. conceptually extend CONWIP to the control of an assembly network with parallel production lines using an MRP-like explosion process. Each serial piece of the production network is controlled using CONWIP. Production for each serial piece is "pulled" based on demand at the end of the serial piece as calculated using MRP logic. The authors point out that CONWIP supports MRP logic by stabilizing leadtimes, a critical assumption underlying MRP. They thus predict that the fusion of CONWIP and MRP will outperform MRP alone for assembly networks with parallel lines.

The concepts behind CONWIP are clear for a serial line, or for a network composed mainly of serial lines. In fact, it can be concluded that under CONWIP, a complex manufacturing network is essentially viewed as numerous serial lines coordinated by MRP logic. The question remains regarding the form CONWIP takes on for networks with a high number of short, parallel lines. Apparently, CONWIP approximates kanban for a serial line, but approaches MRP for a network composed of many short, parallel lines.

In terms of a global network, CONWIP appears to have several limitations. First, the strategy underlying CONWIP goes from mostly pull for serial lines to mostly push for

short, parallel lines. It is not apparent to us that the choice of push versus pull control strategy should depend solely on the architecture of the production network. Clearly the kanban system has worked well for Toyota, presumably a production network with many parallel lines. Second, it is possible under CONWIP that at one instant in time most of the WIP is near the end of the production line. This possibility in conjunction with large leadtimes in a global network could result in lumpy or cyclic output. Perhaps some additional mechanism is needed in CONWIP to pace throughput, but none is given.

Finally, CONWIP requires a centrally determined number for the number of production cards. Yet, no method for setting the number of cards is given. In fact, CONWIP may be quite sensitive to the amount of WIP inventory. Too much inventory may result in inventory accumulations and resulting inefficiencies whereas too little inventory may result in shortages. Shortages in a global network with long leadtimes may be catastrophic. In summary, the applicability of CONWIP to global networks may be problematic.

2.5.5 Restoration

Tang (1990) presents a control strategy in which production at one stage is scheduled based on downstream shortages of inventories from target inventories. This strategy is referred to as restoration since it attempts to

fully or partially restore inventory back to target levels. In concept, the restoration strategy applies to any network design. However, Tang's particular model restricts the application to serial production lines. His model includes two forms of uncertainty, uncertainty in yield rates and in demand. The model assumes unlimited capacity and backorders but no rework is allowed.

The approximate restoration rule proposed by Tang uses a linear production rule composed of two terms to calculate production at each node. The first term adjusts end item demand for yield rates, the percent of products without defects, at the current and subsequent nodes. In other words, production at node i is end item demand divided by the product of the yield rates at node i and all downstream nodes.

The second term adjusts production at node i for inventory shortages from prespecified target inventories at all nodes downstream to node i. In effect, production is pulled based on shortages of actual inventories from target inventories at downstream nodes. The second term uses prespecified restoration coefficients that provide the percent of inventory shortage to be restored. A restoration coefficient of 0.8 between node i and downstream node k means that 80% of the inventory shortage at node k is to be scheduled at node i. The production rule as presented by Tang (1990) is:

$$X_i{}^t = \frac{D^t}{\prod\limits_{a \in s'(i)} y_a} + \sum_{k \in s'(i)} \frac{[TI_k - I_i{}^t]\, r_{ik}}{\prod\limits_{a \in s'(i)} y_a} \qquad 2.34$$

$$0 \leq r_{ik} \leq 1$$

where

$X_i{}^t$ = Production start at node i during period t

D^t = Expected demand per period for final product during t

y_a = Yield rate at node a

TI_k = Target inventory at node k

$I_i{}^t$ = Buffer inventory at node k at the beginning of period t

r_{ik} = Restoration coefficient from node i to downstream node k-ranges from 0% to 100%

s'(i)= The set of node i and all nodes downstream to node i

Assuming perfect yields, the restoration rule can potentially schedule production at any of three extremes as shown in Figure 2-1. Figure 2-1 shows the concept using a serial line for illustration purposes only since the concept applies to any network architecture. At one extreme, the restoration rule schedules production at a stage based only on inventory shortages at the subsequent stage (r_{ik}=1 for k downstream and adjacent to i, else r_{ik} = 0). In this event, the restoration rule is similar to kanban and subsequently has the push/pull characteristics of kanban.

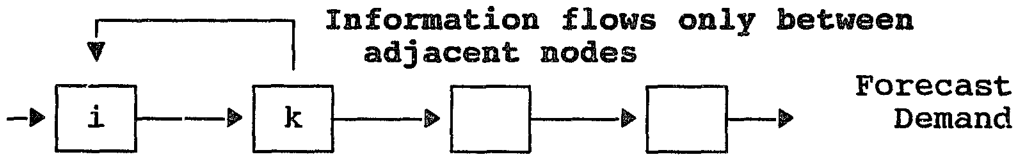
At a second extreme, the restoration rule schedules production based only on shortages at nodes that meet final demand: $r_{ik}=0$ if k does not satisfy final demand and $r_{ik}=1$ if k satisfies final demand. In this case, the restoration rule parallels the base stock control system in which production is scheduled based on end item demand. Shortages or excesses of inventory at intervening nodes are effectively ignored.

At the final extreme, the restoration rule schedules production at stage i based on inventory shortages at all stages downstream to i ($r_{ik}=1$ $k \in s'(i)$). At this extreme, batch size and timing are centrally derived using downstream information-indicative of both push and pull traits. Target inventories and restoration coefficients are centrally derived, a clear push trait. In fact, depending on values for restoration coefficients, the restoration model can vary from a predominately pull system to a system with clear push traits.

Tang proceeds to develop a heuristic by which both restoration coefficients and target inventories can be established for serial lines. Restoration coefficients are first found using a nonlinear math program. One constraint of the math program sets an upper limit on variance of production based on service level considerations. Target inventories are then set for expected final demand plus safety stock based on the variance of the quantity

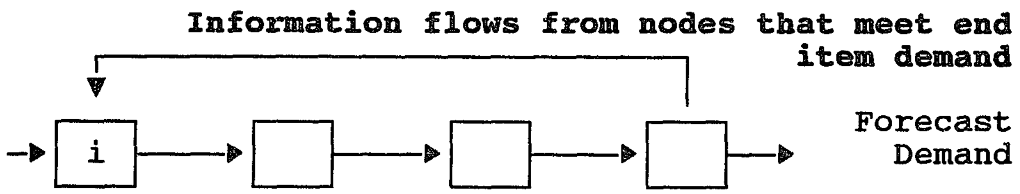
Approximates Kanban:

$$r_{ik} = \begin{cases} 1 & \text{for } k \text{ adjacent to and downstream of } i \\ 0 & \text{otherwise} \end{cases}$$



Approximates Base Stock:

$$r_{ik} = \begin{cases} 1 & \text{for } k \text{ a node that meets end item demand} \\ 0 & \text{otherwise} \end{cases}$$



Approximates a Push Strategy:

$$r_{ik} = \begin{cases} 1 & \text{for } k \text{ downstream to } i \\ 0 & \text{otherwise} \end{cases}$$

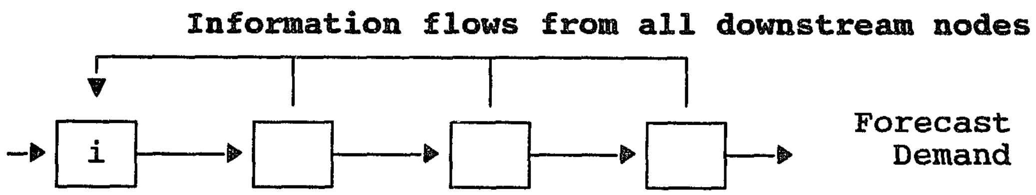


FIGURE 2-1
THE RESTORATION RULE AT ITS EXTREMES
ASSUMING PERFECT YIELDS

$(I_i^t - X_{i+1}^t)$ where I_i^t = buffer inventory at node i and X_{i+1}^t which equals the scheduled production at node (i+1). Overall, the restoration rule provides an interesting but not necessarily optimal rule for controlling production in a serial line. Conceptually, the extension of the rule to production lines other than serial is obvious. However, the extension of Tang's approximate analysis for finding restoration coefficients and target inventories is not.

In terms of a global network, the restoration concept is appealing for a number of reasons. First, it provides flexibility since parameter settings allow the rule to approximate kanban at one extreme or incorporate considerably more centralized, or push, components at the other extreme. Effectively, different rules can be constructed if required for different networks. Second, the restoration concept readily embodies the lean production concepts of minimal inventory, continual improvement, and quality. Presumably, this is a prerequisite in today's competitive, global arena.

Third, the restoration rule with $r_{ik} > 0$ for k downstream to i, works toward maintaining a stable amount of WIP between node i and the downstream node(s) of final demand. Since this is true for all nodes, restoration attempts to maintain stable WIP at each subset of nodes in the network. Equivalently, restoration attempts to maintain WIP inventories at their target level for each node.

Inventory shortages and accumulations should occur infrequently as long as restoration coefficients and target inventories have been set at appropriate values.

Fourth, constant WIP in every subset of nodes should result in nearly constant leadtimes within each subset. This has the advantage of helping to create predictability that makes planning more effective and control more realizable. In other words, the restoration rule may help offset some of the variation inherent in global networks. Finally, the restoration rule has the potential to tie several or all of the stages together so that information passes quickly throughout the network. This capability can be used to help overcome the large leadtimes in global networks.

In summary, the restoration rule appears to be applicable to global networks. It is interesting to note that the restoration rule has some similarities to the control rule used by NUMMI, a Toyota-General Motors joint venture in California (see both Tang, 1990 and Parker and Slaughter, 1988). Even so, the issue of how to optimally set restoration coefficients and target inventories for a general network remains. We will return to this issue in Chapter 3.

2.5.6 Hodgson and Wang

Hodgson and Wang (1991a; 1991b) look at several control strategies for two specific networks, both of which are

assembly type. These authors conclude that the best control strategy for the particular networks studied is to use a push strategy to release orders at all top, upstream stages. Subsequently, a pull strategy is used to control production at all downstream stages. Basically, they seem to be recommending a pull system, subject to a push strategy for order releases at those nodes closest to raw materials.

A question arises in terms of the effects on the system of a rapid increase or a lump in demand. A rapid increase in demand would cause the push system to release more orders at the top, upstream stages. Thinking in terms of a minimal inventory pull system, the system may not be able to adapt to the increased production requirements. As a result, inventory may accumulate and leadtimes grow. It would appear that some mechanism linking the push and pull components is needed.

2.5.7 Mean Weighted Variances

Gong and Matsuo (1991) propose a strategy for a serial multi-stage line based on minimizing WIP. Their formulation incorporates stationary random yield and demand over infinite time periods. Basically, their formulation is to minimize the expected value of WIP subject to:

1) inventory balance constraints;
2) feasibility constraints; and
3) service level constraints.

The last two constraints above are stated in terms of chance-constraints.

Gong and Matsuo propose a control rule to minimize weighted variances of WIP. Mean weighted variance is derived as an optimal solution to a system with a quadratic objective function. They show that mean weighted variance subsumes the restoration control rule, and that it is an improvement over the restoration control rule for serial systems. Limitations to the model include its complexity and its restriction to serial networks.

2.5.8 Production Authorization Cards

Finally, Buzacott and Shanthikumar (1992) state that the optimal control system depends on the particulars of each manufacturing environment. They present a control system called Production Authorization Cards that they claim generalizes MRP, OPT, kanban, CONWIP, in addition to several other control systems. Production Authorization Cards uses a set of tags, electronic signals, to control materials management throughout the network. The coordination logic does not include price information so that the system is limited to coordination within one firm or between cooperating firms. A brief description of Production Authorization Cards is provided in Figure 2-2 and in the discussion that follows.

Production Authorization Cards uses tags to control production. Each tag is associated with a single item so

that n separate tags are required for a batch of size n. A requisition tag is used by a manufacturing cell to notify the predecessor buffer to ship one item as soon as possible (see Figure 2-2). An order tag is used by a manufacturing cell to notify the predecessor buffer that a requisition tag for an item will be forthcoming in the near future. In effect, an order tag notifies a buffer in advance of demand. For every order tag, there will eventually be one requisition tag. The time lag between the order tag and the associated requisition tag can either be zero or a positive number.

Depending on availability of inventory at a buffer, an order tag may cause the buffer to send a production authorization tag to the predecessor production cell. The purpose of a production authorization tag is to authorize the production of one item. After production, and after the item has been moved to a buffer, the production authorization tag is released back to the buffer. As shown in Figure 2-2, a production authorization tag travels from a buffer to a predecessor production cell. It then travels along with the item back to the buffer, where the tag is released.

The number of production authorization tags at a buffer may be limited, by design. One effect of a limit on the number of tags is to place an upper bound on WIP. The limit on tags equals the maximum number of items that can be

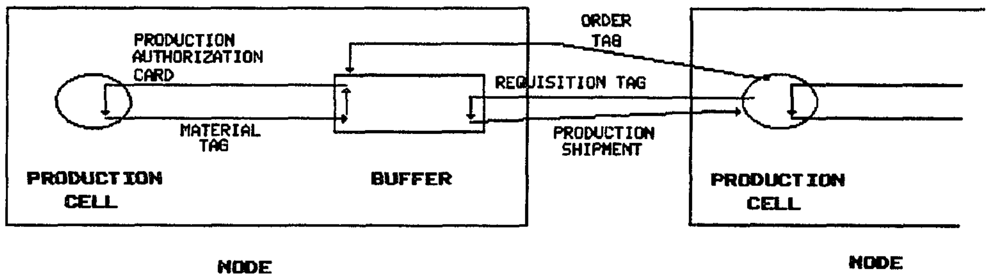


FIGURE 2-2
OVERVIEW OF PRODUCTION AUTHORIZATION CARDS

in-process at a cell at one time. A limit on the number of production authorization tags can also slow the flow of information through the network. This occurs as a buffer with no remaining production authorization tags receives additional order tags; the buffer has no production authorization tags remaining with which to request additional production. Furthermore, no tags will be available until the predecessor cell completes some WIP. Thus, no information regarding demand can move through the node until some WIP is completed and the associated production authorization tags are released. Effectively, this results in a delay in the flow of information on demand moving through the node.

Buzacott and Shanthikumar argue that under Production Authorization Cards, three parameters are needed to distinguish MRP from kanban from CONWIP from OPT, etc. The three parameters are:

z_i = Static inventory limit at node i
(Target Inventory)

k_{ij} = Limit on the number of active production authorization tags at node j for product i

τ_{ij} = The delay at cell j between the order tag for product i and the associated requisition tag

Within the framework of Production Authorization Cards, MRP is defined by:

$z_i \geq 0$ for all i $\rightarrow$ equates to safety stock

$k_{ij} = \infty$ for all i,j $\rightarrow$ no limit on WIP

$\tau_{ij} \geq 0$ for all i, j $\rightarrow$ delay determined by leadtimes

In other words, MRP has safety stock, no limit on WIP, and zero or positive times between the issuance between order tags and requisition tags.

Under Production Authorization Cards, the parameter values associated with kanban are:

$z_i > 0$ for all i

$k_{ij} = z_i$ for all i,j (assuming serial lines)

$\tau_{ij} = 0$ for all i, j

Target inventory for kanban is some number strictly greater than zero, albeit small for typical kanban installations. The limit on WIP equals target inventory, in contrast to MRP. The positive value for k_{ij} means that information about demands for final products does not automatically pass back to earlier nodes immediately-the flow of information may be limited by the finite number of production authorization tags. Similarly, other values of the same three parameters distinguish CONWIP and several other control rules. The reader is referred to the original paper for additional details.

A criticism of the model presented by Buzacott and Shanthikumar is that the parameter k_{ij} confounds the issue of limited WIP with a delay on the flow of information moving through the node. Thus, it is not possible under

Production Authorization Cards to model a system that limits WIP yet allows information to flow through all nodes without delay. For example, Production Authorization Cards does not subsume restoration in which work scheduled at a node is potentially based on inventory shortages at ALL downstream nodes with nonzero restoration coefficients. Even so, the model by Buzacott and Shanthikumar suggests the tantalizing notion that only a few parameters are needed to distinguish between control systems. We will extend this concept in Chapter 3 by presenting a modification of Production Authorization Cards.

2.6 Summary

Clearly, production scheduling and target inventories have been the focus of much of the manufacturing research over the past half century. Nonetheless, there are few results for complex, general networks and yet fewer results for those under random demands and/or leadtimes. In the meantime, a number of good, but not necessarily optimal, strategies have evolved for manufacturing planning and control. Historically, these systems have been somewhat loosely categorized as either push or pull. It was shown that it is more appropriate to categorize features of a system along a push-pull gradient than to characterize the entire system as either push or pull. Kanban, MRP, base stock, and several newly proposed control strategies were

reviewed and features examined in terms of push-pull gradients.

Production Authorization Cards, a system that subsumes kanban, MRP, OPT, and other control rules, was reviewed. It was shown that Production Authorization Cards does not subsume the restoration control strategy. Yet, the restoration control strategy is useful since it can be used to represent a continuum gradient from push to pull based on restoration coefficients. An extension of Production Authorization Cards that also subsumes restoration is needed.

In Chapter 3, a modified version of Production Authorization Cards is given. The modified version is used for recognition of the important parameters-those parameters whose values determine the characteristics of the best control rule. Model formulations for a general network under known demands are then developed and can be used to estimate these parameters. In particular, the concept of target inventory plays an integral role in all of the models developed.

CHAPTER 3 MODELS

3.1 Introduction

Table 1-1 in Chapter 1 listed nine factors important in regards to materials management in a global network. Based on earlier discussion, we choose to ignore yield rates by assuming zero defects. The issues that remain are the subjects of this study and are shown in Table 3-1.

Section 3.2 presents a generalized control system that is a modification of Production Authorization Cards. The modification requires a central controller of information and a slightly different set of parameters than those used by Buzacott and Shanthikumar (1992). The advantage of the modification is that the modified control system subsumes the restoration rule proposed by Tang (1990) in addition to kanban, CONWIP, made-to-order, MRP, and others (see Buzacott and Shanthikumar for details).

A math programming model is presented in Section 3.3.2 for the case of known demand in a general network. It is a mixed-integer, linear formulation that can be solved using standard techniques. More general math programming models that include a production rule based on the restoration concept are given in Sections 3.3.3 and 3.3.4. The model of Section 3.3.3 requires that every demand is met, regardless of cost whereas the model of Section 3.3.4 allows for lost sales. Decision variables of these two models include three

TABLE 3-1
ISSUES RELEVANT TO MATERIALS MANAGEMENT
IN A GLOBAL NETWORK

CATEGORY	ISSUE
Inventory	Order Quantity Target Inventory
Production	Delivery Leadtime Production Leadtime Capacity
Global	Network Architecture Coordination Characteristics of Demand

of the four parameters of the modified Production Authorization Cards system proposed in Section 3.2. The fourth parameter is not included for reasons discussed momentarily.

Limitations of each model are noted as is our inability to solve the mixed-integer, nonlinear formulation as given in Sections 3.3.3 and 3.3.4. However, a solution strategy for solving the models is developed and techniques for solving nonlinear math programming formulations are briefly discussed in Section 3.4. Questions to be explored in regards to the models of Sections 3.3.3 and 3.3.4 and the methodology for doing so are presented in Chapter 4.

3.2 Modified Control System

As described in Chapter 2, a major criticism of Production Authorization Cards is that it confounds the issue of constant WIP at a node with the flow of information

through the node. This prevents a system from passing information through a node in which all allowed inventory is in-process. Combining the concepts of Production Authorization Cards and restoration, a more general form of Production Authorization Cards is presented. The modification incorporates a central controller and a set of additional parameters, namely restoration coefficients. The modified version of Production Authorization Cards subsumes MRP, base stock, kanban, and restoration among other control systems.

Figure 3-1 shows a modified version of Production Authorization Cards suitable for our purposes. The figure shows a serial line for ease of exposition, although the concept applies to networks of any design. Requisition tags and production authorization tags work exactly as they did in Production Authorization Cards. The main difference lies in the addition of a central controller through which all order tags must now pass. Associated with the order tag from node i to each downstream node j is a restoration coefficient r_{ij} ($0 \le r_{ij} \le 1$). The restoration coefficient is identical in concept to the restoration coefficient concept in Tang's (1990) restoration model. Under this generalization of Production Authorization Cards, information can now pass through a node even though all allowable inventory is in-process.

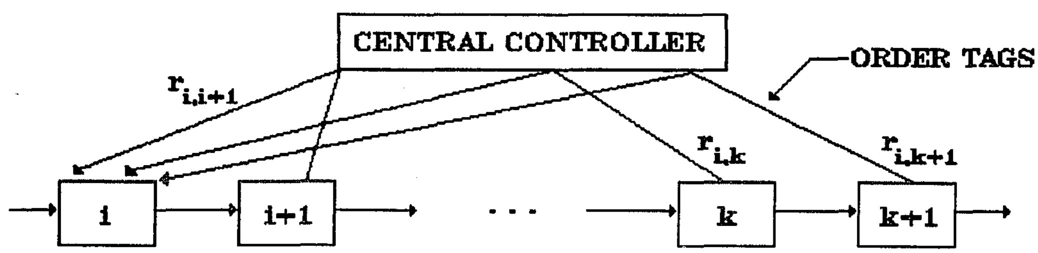


FIGURE 3-1
MODIFIED VERSION OF PRODUCTION AUTHORIZATION CARDS

No longer does the parameter k_{ij} both limit WIP at node j and delay passage of information through node j. The variable k_{ij} is relegated to represent the upper limit on WIP at node j. In order to clearly distinguish this difference, the symbol k_{ij} will no longer be used. Henceforth, the symbol UL_j will be used strictly for the upper limit of inventory at node j.

The four parameters needed to describe the modified system are:

TI_j = Target inventory at node j

UL_j = Upper limit on inventory at node j

r_{ij} = Restoration coefficient from node i to downstream note j

τ_{ij} = The delay at cell j between the order tag for component i and the associated requisition tag

Note that three of these conceptually follow from the work of Buzacott and Shanthikumar (1992), namely TI_j, UL_j, and τ_{ij}; and one follows after Tang (1990), namely r_{ij}.

The specific parameter values of the modified system that are associated with kanban, restoration, and MRP are shown in Table 3-2. Notice that target inventories TI_j are greater than zero for both kanban and restoration but in theory may be zero for MRP. In practice, MRP systems require safety stock, so that target inventories for actual MRP systems are also greater than zero. Notice also that kanban is the only one of the three systems with an upper

limit on WIP (UL_j) at a node. Neither restoration nor MRP set an upper limit on WIP.

TABLE 3-2
MODIFIED PRODUCTION AUTHORIZATION CARDS
PARAMETER VALUES FOR THREE CONTROL RULES

CONTROL SYSTEM	PARAMETER VALUES
Kanban	$TI_j > 0$ $UL_j = TI_j$ $\tau_{ij} = 0$ $r_{ij} = \begin{cases} 1 & \text{for k adjacent \& downstream to j} \\ 0 & \text{otherwise} \end{cases}$
Restoration	$TI_j > 0$ $UL_j > 0$ $\tau_{ij} = 0$ $0 \leq r_{ij} \leq 1$ for j downstream to i $r_{ij} = 0$ otherwise
MRP	$TI_j \geq 0$ $UL_j > 0$ $\tau_{ij} = \Sigma l_{ij} > 0$ for j downstream to i $r_{ij} = \begin{cases} 1 & \text{for j downstream to i*} \\ 0 & \text{otherwise} \end{cases}$

* The definition of r_{ij} under MRP applies only to those nodes i closest to raw materials.

MRP uses planned leadtimes (τ_{ij}) for estimating the delay between order tags and requisition tags. In contrast, kanban makes no allowances for leadtimes ($\tau_{ij} = 0$ for all i and j). A work center operating under kanban is to deliver product as soon as possible after being notified of a need at a subsequent work center. Restoration could probably be modified to explicitly include leadtime considerations but at the expense of model complexity.

In kanban, the only nonzero restoration coefficients (r_{ij}) are between a node and any adjacent, downstream nodes. In this case, the restoration coefficient equals 100%. Under base stock, the only nonzero restoration coefficients are between each node and any downstream node that satisfies final demands; these restoration coefficients equal 100%. Under MRP, order releases at those nodes furthest upstream are based on forecast demands, planned leadtimes, and existing inventories at all downstream nodes. Therefore, restoration coefficients between those nodes closest to raw materials and ALL downstream nodes equal 100%.

Under the restoration control rule, the restoration coefficient between a node and any downstream node can range from 0% to 100%. Thus, restoration can represent a gradient from something close to kanban on one extreme to something close to base stock, or perhaps MRP, on the other extreme. A primary goal of the final models of this chapter is to help identify the best control strategy using target inventories, restoration coefficients, and shortages of actual inventory from target.

One primary difference between kanban and restoration is that kanban has an upper limit on inventory at each node whereas restoration does not. Instead, restoration has a target inventory that the system tries to maintain, but may exceed on occasion. A second difference between kanban and restoration is that a node under kanban schedules production

based only on shortages at adjacent, downstream nodes. In contrast, nodes under restoration schedule production based on shortages at any or all downstream nodes depending on values of restoration coefficients.

The principal difference between kanban and restoration on one side and MRP on the other is that MRP includes planned leadtimes whereas the time lag between order tags and requisition tags equals zero for both kanban and restoration. Another distinction is that both kanban and restoration strive for constant WIP, in contrast to MRP. Recall that MRP schedules throughput then measures WIP. The end result is that WIP will typically vary more under MRP than under either kanban or restoration.

MRP differs from restoration in that under MRP, nodes furthest upstream automatically receive full information from all downstream nodes. In other words, all of the restoration coefficients to these nodes are 100%. Under MRP, the values of restoration coefficients to nodes that are not furthest upstream are not apparent. The restoration concept is a "pull" concept and cannot easily be modified to describe MRP-like "pushing" of inventory through intermediate nodes. Under restoration, each node receives information from some but not necessarily all downstream nodes. Further, the information from downstream nodes under restoration may be filtered by restoration coefficients less than 100%.

A modified version of Production Authorization Cards that requires four parameters has been identified. The modified version allows us to characterize and distinguish kanban from restoration from MRP. The modified version also allows us to characterize made-to-order, CONWIP, integral control, and OPT among other systems (see Buzacott and Shanthikumar, 1992).

In the next section, three models for a general network under known demand are presented. The first model does not contain a production smoothing rule and uses two of the four parameters of the modified control system. The second and third models do contain a production smoothing rule and use three of the four parameters of the modified control system. The perspective is that these final two models can be used to find optimal parameter values for particular networks. In turn, optimal parameter values can be used to identify characteristics of the "best" control rule for that particular network.

3.3 Math Programming Formulation

3.3.1 Introduction and Notation

A primary goal of this section is to formulate a fairly general math programming model for production scheduling and inventory control of a single product manufactured in a global network. The models of this section apply to general networks, networks of any design. Both production leadtimes and delivery leadtimes may be significant in global networks

so that both types of leadtimes are incorporated into the model. Large leadtimes suggest that it may be advisable to use overtime to avoid starving a node for input materials. On the other hand, costs may be such that it may, from time to time, be beneficial to plan undertime. Accordingly, allowances are also made in the model for both overtime and undertime.

A common assumption of previous models of complex networks is that a control rule perfectly coordinates production. Under this assumption, it is not possible for a node to have inventory on hand from an immediate predecessor without also having paired inventory on hand from all other immediate predecessors. For example, in order to mount a wheel onto an automobile, a work center must have a tire, a wheel mount, and several lug nuts. Perfect coordination assumes that a work center would never have, for example, tires and wheel mounts on hand but no lug nuts. In reality, the assumption of perfect coordination will not hold in general. Capacity constraints and leadtime variability will result in some shipments that are behind schedule and others that are ahead of schedule. The assumption of perfect coordination is relaxed in the models of this section.

Four inventory states are used in the models: in-transit, in-wait, in-process, and in-buffer. In-process and in-buffer states are commonly used in models of this type

and do note need further explanation. The in-transit state is needed since significant amounts of inventory may be in-transit at any instant in time in a global network with large distances between nodes. The in-wait state is required since the assumption of perfect coordination is relaxed and inventory may be at a node waiting to go into production. An overview of the notation is presented in Figure 3-2.

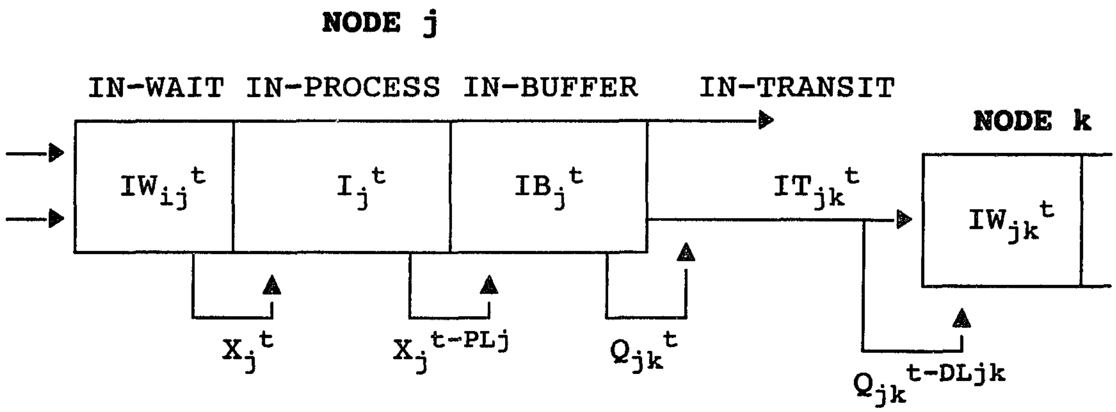
Work in-process inventory at the end of period t (I_j^t) equals work in-process inventory at the end of t-1 adjusted for product that begins processing during t (X_j^t) and product that finishes processing during t (X_j^{t-PLj}). This relationship is shown in the following equation:

$$I_j^t \;=\; I_j^{t-1} + X_j^t - X_j^{t-PLj} \qquad \forall j;\ t \geq 1 \qquad 3.1$$

In-buffer inventory at the end of period t (IB_j^t) equals ending in-buffer inventory at the end of t-1 adjusted for product that has completed processing during t (X_j^{t-PLj}) and any product that was removed from the buffer during t. Equation 3.2 is for nodes that meet final demand and equation 3.3 is for nodes that produce components for successor nodes:

$$IB_j^t = IB_j^{t-1} + X_j^{t-PLj} - D_j^t \qquad \forall j \ni s(j)=\emptyset;\ t \geq 1 \qquad 3.2$$

$$IB_j^t = IB_j^{t-1} + X_j^{t-PLj} - \sum_{k \in s(j)} Q_{jk}^t \qquad \forall j \ni s(j)\neq\emptyset;\ t \geq 1 \qquad 3.3$$



INVENTORY STATES:

IW_{ij}^{t} = In-wait inventory shipped from node i and on hand at node j at the end of period t
I_j^t = In-process inventory at node j at the end of t
IB_j^t = In-buffer inventory at node j at the end of t
IT_{jk}^{t} = In-transit inventory in shipment from node j to node k at the end of t

VARIABLES USED TO CHANGE INVENTORY STATES:

X_j^t = Amount of product scheduled to go into production at node j during period t
PL_j = Production leadtime at node j
Q_{jk}^{t} = Amount of product scheduled to be shipped from node j to node k during period t
DL_{jk} = Delivery leadtime from node j to node k

FIGURE 3-2
INVENTORY STATES AT A NODE

In-transit inventory between nodes j and k at the end of period t (IT_{jk}^{t}) equals in-transit inventory at the end of the prior period adjusted for any new shipments that are shipped from the predecessor (Q_{jk}^{t}) or received at the successor (Q_{jk}^{t-DLjk}). This relationship is expressed by the following equation:

$$IT_{jk}^{t} = IT_{jk}^{t-1} + Q_{jk}^{t} - Q_{jk}^{t-DLjk} \qquad \forall j \ni s(j) \neq \emptyset;\ k \in s(j);\ t \geq 1 \qquad 3.4$$

The final inventory balance constraint is found by realizing that in-wait inventory at node k that came from node j (IW_{jk}^{t}) is increased by product shipped from j DL_{jk} time units ago (Q_{jk}^{t-DLjk}) and decreased by product scheduled to go into production at node k during period t (X_{k}^{t}):

$$IW_{jk}^{t} = IW_{jk}^{t-1} + Q_{jk}^{t-DLjk} - X_{k}^{t} \qquad \forall j \ni s(j) \neq \emptyset;\ k \in s(j);\ t \geq 1 \qquad 3.5$$

The models of the next three sections are progressively more complex even though they are all based on the assumption of known demands. The model of Section 3.3.2 does not incorporate a production control rule and is included for the sake of completeness. The model of Section 3.3.3 includes a production control rule as does the model of Section 3.3.4. The production control rule is important

in that it provides the system with the ability to adjust to demands (or leadtimes) that differ from forecasts (or planned leadtimes). The model of Section 3.3.4 differs from those of Sections 3.3.2 and 3.3.3 in that it incorporates the possibility of lost sales.

3.3.2 Model without a Production Rule

The model introduced in this section applies to any network design, e.g., serial, assembly, conjoint, etc. Further, the assumption of perfect control is relaxed so that it is possible for a node to have product on hand from a predecessor while waiting on product from another predecessor. The purpose of this section is to introduce the basic assumptions and notation. Most of the assumptions listed here as well as the notation also apply to the models of Sections 3.3.3 and 3.3.4. The basic assumptions follow:

1) Networks can be of any design;
2) Finite horizon;
3) Demand is known and must be met;
4) External demands occur only at nodes with no successors;
5) A single product is produced at each node;
6) Production leadtimes and delivery leadtimes are constant and known;
7) There is a limit on the amount of overtime that may be scheduled at a node;
8) An unlimited supply of raw materials is available;

9) Time units are small enough so that a product does not move more than one inventory state per time period;

10) The production of 1 item at a node requires only 1 component from each predecessor; and

11) The only set up costs are those associated with the shipment of product from one node to another- there are no set up costs associated with production.

The notation that is used throughout the remainder of this study is presented here. Note that inventory, amount of work scheduled, overtime, and undertime are all given in number of items.

$IW_{jk}{}^t$ = In-wait inventory shipped from node j and on hand at node k at the end of period t

$IT_{jk}{}^t$ = In-transit inventory in shipment from node j to node k at the end of period t

$IB_j{}^t$ = In-buffer inventory at node j ready for shipment to a successor node at the end of period t

$I_j{}^t$ = In-process inventory that is in production at node j at the end of period t

$X_j{}^t$ = Amount of work scheduled to enter the in-process state at node j during period t

PL_j = Production leadtime at node j

X_j^{t-PLj} = Amount of work going from the in-process state to the in-buffer state at node j during t

Q_{jk}^t = Amount of product going from the in-buffer state at j to the in-transit from j to k state during t

DL_{jk} = Delivery leadtime for delivery from node j to node k

Q_{jk}^{t-DLjk} = Amount of product going from the in-transit state between j and k to the in-wait state at subsequent node k during period t

S_{jk} = Setup cost of initiating a shipment Q_{jk}^t from j to k

TI_{jk} = Target inventory for total inventory in-transit from j to k, in-wait at k from j, in-process at k, plus in-buffer at k; for $j \in p(k)$

INV_{ij}^t= Actual inventory in-transit from j to k, in-wait at k from j, in-process at k, plus in-buffer at k; for $j \in p(k)$

h_j = Holding cost for in-process inventory at j

h_j' = Holding cost for in-buffer inventory at j

h_{jk} = Holding cost for in-transit inventory from j to k

h_{jk}' = Holding cost for in-wait inventory at k, from j

O_j^t = Amount of overtime scheduled at node j during period t, in number of items

CO_j = Per item cost of overtime at node j

U_j^t = Amount of undertime scheduled at node j during period t, in number of items

CU_j = Per item cost of undertime at node j

$\delta_{jk}{}^t$ = 0 or 1 switch that turns shipment set up cost on or off

s(j) = Set of all immediate, successor nodes to node j

p(j) = Set of all immediate, predecessor nodes to node j

$D_j{}^t$ = Demand for final product at node j during t

R_j = Regular production capacity in number of items at node j

B_j = Upper bound on overtime at node j, in number of items

M = A large number

The basic theme underlying the model is to minimize the sum of setup costs, holding costs, undertime/overtime costs, and the cost of inventories falling either above or below target inventories. The first four categories of cost are those commonly found in models of this type and need no further explanation. In contrast, the cost associated with deviation of actual inventories from target inventories needs clarification.

First note that the concept of target inventories are important in this model even though demands and leadtimes are known with certainty. This is true since we are exploring higher level decisions in a hierarchical planning model. Presumably, results from this model, including

values for target inventories, will be disaggregated and subsequently used at lower levels.

In a global network, shortages will tend to result as inventories fall significantly below target levels. Shortages in an environment with long leadtimes will result in either substantial expediting costs or lost sales or both. On the other hand, inventories in excess of target result in large holding costs. More subtly, inventories in excess of target do not seem to be compatible with the theme of lean production now generally considered essential. Certainly, kanban practitioners dogmatically strive to identify and reduce excess inventories. The stated reason is that inventories hide flaws in the production or materials management processes. At any rate, it is clear that a V or U shaped cost function based on deviation from target inventory is appropriate.

Given that the cost function is of this general shape, the question becomes one of whether a linear cost function or a nonlinear function such as the quadratic one presented by Holt et al. (1960) is appropriate. We feel that significant deviations from target, in either direction, can be quite costly for global manufacturers practicing lean production so that the curve should be steep. In this section, we introduce this cost using linear terms so that the overall model remains linear. Two different terms are used so that the cost of a shortage does not necessarily

have to equal the cost of an excess of the same order of magnitude:

$C_1(TI_{jk} - INV_{jk}^t)^+$ is used when $TI_{jk} > INV_{jk}^t$ and

$C_2(TI_{jk} - INV_{jk}^t)^-$ is used when $TI_{jk} < INV_{jk}^t$.

These linear terms will be replaced by a single, quadratic term in the models of Sections 3.3.3 and 3.3.4.

Constraints include the inventory balance constraints introduced in Section 3.3.1, capacity constraints involving overtime and undertime (3.7-3.8), and restrictions on variables (3.9-3.12). The decision variables are the quantities Q_{jk}^t to be shipped, target inventories TI_{ij}, the associated δ_{jk}^t variable for setup cost, the amount to produce X_j^t, the amount of undertime U_j^t, and the amount of overtime O_j^t:

$$\min \sum_j \sum_t \Big\{ h_j I_j^t + h_j' IB_j^t + CO_j O_j^t + CU_j U_j^t + \sum_{k\epsilon s(j)} [\delta_{jk}^t S_{jk} + h_{jk} IT_{jk}^t + h_{jk}' IW_{jk}^t + C_1(TI_{jk}-INV_{jk}^t)^+ + C_2(TI_{jk}-INV_{jk}^t)^-] \Big\} \qquad 3.6$$

such that:

Inventory balance constraints 3.1 - 3.5, and

$$X_j^t + U_j^t - O_j^t = R_j \qquad \forall j,t \qquad 3.7$$

$$O_j^t \le B_j \qquad \forall j,t \qquad 3.8$$

$$Q_{jk}^t \le M\delta_{jk}^t \qquad \forall j,t;\ k\epsilon s(j) \qquad 3.9$$

$$\delta_{jk}^t \in \{0,1\} \qquad \forall j,k,t \qquad 3.10$$

$$I_j^t,\ IB_j^t,\ IT_{jk}^t,\ IW_{jk}^t,\ Q_{jk}^t,\ X_j^t,\ TI_{jk} \in \{0,1,2,\ldots\} \qquad \forall j,t;\ k\epsilon s(j) \qquad 3.11$$

$$O_j^t,\ U_j^t \ge 0 \qquad \forall j,t \qquad 3.12$$

Constraint 3.7 states that scheduled production X_j^t plus undertime minus overtime equals regular time. Again, note that X_j^t, U_j^t, O_j^t, and R_j are in terms of number of items and not in time units. Values for U_j^t and O_j^t will not both be greater than 0 for a particular j and t. For if so, a lower objective value can be obtained be decreasing both U_j^t and O_j^t until one or both are zero, while all other constraints remain satisfied. Constraint 3.8 places an upper bound on the amount of overtime that may be scheduled at node j during any t.

In tandem, constraints 3.7 - 3.8 in conjunction with the constraints given by 3.11 assure that scheduled production remains within the capacity of the facility. Production scheduled may range from a lower bound of zero to a finite upper bound. The maximum amount that can be scheduled is the lower of the limit set by:

1) the amount of inventory on hand at predecessor nodes; or
2) the sum of regular time plus overtime capacity.

Note that the model given by 3.1 - 3.12 is a mixed-integer, linear math program that can be solved using standard techniques. A solution for a specific network with specific costs would dictate production (X_j^t), shipment (Q_{jk}^t), overtime (O_j^t), and undertime (U_j^t) amounts as well as target inventories for the optimal control rule. In the

next section we introduce a restoration based control rule as a constraint.

3.3.3 Model with a Production Rule-No Lost Sales

In Section 3.2, four parameters (TI_j, UL_j, r_{ij}, and τ_{ij}) were shown to be associated with many control rules. If values of these four parameters are known, a control rule can potentially be inferred or at least characteristics can be identified. Originally, we looked for a model that could be used to estimate all four parameters for a particular manufacturing network. It became apparent that leadtime (τ_{ij}) made the model unduly complex for marginal gain. Accordingly, attention was focused on developing a model that can be used to estimate TI_{mn}, UL_j, and r_{jn}.

Based on the argument in the prior section, we choose to include a quadratic term in the objective function of the form:

$$K \bullet (\text{Target Inventory} - \text{Existing Inventory})^2$$

Recall that Holt et al. (1960) use a similar quadratic cost function in aggregate planning problems. Our model differs from the Holt et al. model in that target inventories are decision variables in this model. It is of interest to note that the quadratic loss function is also commonly used in quality control (e.g., Kackar, 1985).

The inclusion of this term in the objective function places a penalty on large deviations of inventories from target values. The severity of the penalty (quadratic)

should help to minimize both shortages and excesses in inventories, an important goal of lean production. The constant K in front of the quadratic term is a parameter set by management. Overall, the model chooses the "best" target inventories and a control rule is used to keep inventories near target values. This strategy attempts to maintain constant inventory at each node which has the added advantage of helping to maintain constant leadtimes. Thus, the strategy helps to maintain predictability for the purposes of production planning.

The upper limit on inventory at node j is designated by UL_j. This quantity is an upper limit on inventory at a node and does not include inventory in-transit to or from that node. Effectively, no upper limit is assumed for in-transit inventories since common carriers are typically used to move freight large distances. Inventory at a node is bounded, depending on whether or not it has a predecessor, as follows:

$$\left(\sum_{i \in p(j)} IW_{ij}^{t} \right) + I_j^t + IB_j^t \leq UL_j \qquad \forall t;\ \forall j \ni p(j) \neq \emptyset \qquad 3.13$$

$$I_j^t + IB_j^t \leq UL_j \qquad \forall t;\ \forall j \ni p(j) = \emptyset \qquad 3.14$$

UL_j may either be prespecified by management or it may be a decision variable of the model. If it is a decision variable, then its value will equal the maximum over all

periods of $\{I_j^t + IB_j^t\}$ for equation 3.14 and similarly for equation 3.13. Alternatively, UL_j may be fixed by the processing and storage capacities at existing facilities. In that case 3.13 and 3.14 become active constraints of the model.

Each node with one or more successors is assumed to have an information related coefficient called a restoration coefficient associated with each downstream node. The restoration coefficient r_{jn} is a number between 0 and 1 that applies to the flow of information between node j and downstream node n. Parallel to the restoration concept introduced in Chapter 2, the shipment quantity from node j to node k is a function of target inventories, existing inventories, and restoration coefficients:

$$Q_{jk}^t = \max\left\{0, \sum_{(mn,n)\in DS(jk)} (TI_{mn} - INV_{mn}^{t-1})r_{jn}\right\} \qquad \forall t;\ k \in s(j) \qquad 3.15$$

$$0 \leq r_{jn} \leq 1 \qquad \forall j,n \qquad 3.16$$

where

$INV_{mn}^t \;=\; IT_{mn}^t + IW_{mn}^t + I_n^t + IB_n^t$

$DS(jk)$ = (arc jk, node k) and the set of all arc-node pairs downstream to k, $j \in p(k)$

Equations 3.15 - 3.16 represent the production control rule, also referred to as the production smoothing rule. Note that order quantities cannot be negative. Note also

that order quantity Q_{jk}^{t} is a function of inventory shortages at the end of period (t-1) from target inventories. The quantity INV_{mn}^{t-1} is inventory at the end of period (t-1) along arc mn and at node n. Note that node m must be an immediate predecessor to node n. The notation $(mn,n) \in DS(jk)$ refers to arc mn, node n. The set DS(jk) is arc jk, node k and the set of all arc-nodes downstream to node k.

Thus the shipment quantity from node j to node k during period t is a function of the shortage of inventory at the end of (t-1) from target at arc-node jk, at node k and all arc-nodes downstream to node k. The amount of the shortage at a specific downstream arc-node (mn,n) that is to be shipped from j to k depends on r_{jn}. Kanban is represented by r_{jn} that equal 1 for nodes that are adjacent to one another and 0 for nodes separated by one or more intervening nodes. Base stock is represented by r_{jn} equal to 1 only for nodes n from which final demand is satisfied. Otherwise, r_{jn} equals zero.

The incorporation of the production control rule makes a significant contribution to the original model. First, the production rule helps smooth production requirements. Without it, the model might require lumpy production in order to satisfy lumpy demand. Second, the production control rule allows the model to react to errors in forecasts of demand. Finally, values of restoration

coefficients tell us which shortages and excess of inventories at downstream nodes are important to consider when scheduling production at a node. In effect, values for restoration coefficients along with shortages of actual inventories from target specify the best control strategy for the specific network.

It is important to realize that an additional assumption is required when incorporating the production control rule as a constraint in the model. Feasibility now requires that a value for target inventory exists for each node and that the value satisfies constraint 3.15. Not only that, our model assumes that the value for target inventory at a node must be level across all time periods in the planning horizon. In this section we present a model that requires that all demands must be met. In Section 3.3.4, a mechanism is incorporated into the model which allows for lost sales. The model for the no lost sales case is presented on the next page in its entirety.

$$\min \sum_{j}\sum_{t} \Big\{ h_j I_j^t + h_j' IB_j^t + CO_j O_j^t + CU_j U_j^t + \sum_{k \in s(j)} [\delta_{jk}^t S_{jk} + h_{jk} IT_{jk}^t + h_{jk}' IW_{jk}^t + K(TI_{jk} - INV_{jk}^t)^2] \Big\} \quad 3.17$$

such that:

$$IW_{jk}^t = IW_{jk}^{t-1} + Q_{jk}^{t-DLjk} - X_k^t \qquad \forall j \ni s(j) \neq \emptyset; k \in s(j); t \geq 1 \quad 3.18$$

$$IT_{jk}^t = IT_{jk}^{t-1} + Q_{jk}^t - Q_{jk}^{t-DLjk} \qquad \forall j \ni s(j) \neq \emptyset; k \in s(j); t \geq 1 \quad 3.19$$

$$IB_j^t = IB_j^{t-1} + X_j^{t-PLj} - D_j^t \qquad \forall j \ni s(j) = \emptyset;\ t \geq 1 \quad 3.20$$

$$IB_j^t = IB_j^{t-1} + X_j^{t-PLj} - \sum_{k \in s(j)} Q_{jk}^t \qquad \forall j \ni s(j) \neq \emptyset;\ t \geq 1 \quad 3.21$$

$$I_j^t = I_j^{t-1} + X_j^t - X_j^{t-PLj} \qquad \forall j;\ t \geq 1 \quad 3.22$$

$$X_j^t + U_j^t - O_j^t = R_j \qquad \forall j,t \quad 3.23$$

$$O_j^t \leq B_j \qquad \forall j,t \quad 3.24$$

$$Q_{jk}^t \leq M\delta_{jk}^t \qquad \forall j,t;\ k \in s(j) \quad 3.25$$

$$\Big(\sum_{i \in p(j)} IW_{ij}^t\Big) + I_j^t + IB_j^t \leq UL_j \qquad \forall t;\ \forall j \ni p(j) \neq \emptyset \quad 3.26$$

$$I_j^t + IB_j^t \leq UL_j \qquad \forall t;\ \forall j \ni p(j) = \emptyset \quad 3.27$$

$$Q_{jk}^t = \max\Big\{ 0, \sum_{(mn,n) \in DS(jk)} (TI_{mn} - INV_{mn}^{t-1}) r_{jn} \Big\} \qquad \forall t;\ k \in s(j) \quad 3.28$$

$$0 \leq r_{jn} \leq 1 \qquad \text{for n downstream to j} \quad 3.29$$

$$r_{jn} = 0 \qquad \text{otherwise} \quad 3.30$$

$$\delta_{jk}^t \in \{0,1\} \qquad \forall j,k,t \quad 3.31$$

$$I_j^t,\ IT_{jk}^t,\ IW_{jk}^t,\ Q_{jk}^t,\ X_j^t,\ TI_{jk} \in \{0,1,2,\ldots\} \qquad \forall j,t;\ k \in s(j) \quad 3.32$$

$$IB_j^t \in \{0,1,2,\ldots\} \qquad \forall j \ni s(j) \neq \emptyset;\ \forall t \quad 3.33$$

$$O_j^t,\ U_j^t \geq 0 \qquad \forall j,t \quad 3.34$$

The model given by 3.17 - 3.34 is a formulation that applies to a network of any design, under known leadtimes, and under known demands. The objective function is similar to that of section 3.3.2 except for the modified (TI_{jk}-INV_{jk}^t) cost term in the objective function. The additional constraints given by 3.26 and 3.27 represent upper bounds on inventories and the production control rule is given by constraints 3.28 - 3.30.

The parameters TI_{jk}, r_{jn}, and UL_j that are referred to in the modified version of Production Authorization Cards are decision variables in this model. When known, these decision variables in conjunction with actual inventories dictate a specific control rule. The decision variables Q_{jk}^t and δ_{jk}^t dictate shipment quantities and associated setup costs. The variable X_j^t represents the movement of inventory within a node, pulled by shipment quantities Q_{jk}^t at the end buffer for node j. The decision variables O_j^t and U_j^t dictate the amount of overtime and undertime (in number of items) to be scheduled at a node.

3.3.4 Model with a Production Rule-Lost Sales

The model of Section 3.3.3 requires that all demands be met and also specifies the existence of level target inventories and restoration coefficients for each planning horizon. That model is extended in this section to incorporate the possibility of lost sales. This requires an additional cost term in the objective function; the cost

associated with losing a sale and possibly a customer. It also requires that constraint 3.20 be modified to incorporate the possibility of a lost sale. In all other respects, the model is identical to that of Section 3.3.3.

$$\min \sum_{j}\sum_{t} \Big\{ h_j I_j{}^t + h_j{}' IB_j{}^t + CO_j O_j{}^t + CU_j U_j{}^t + \sum_{k \in s(j)} [\delta_{jk}{}^t S_{jk} + h_{jk} IT_{jk}{}^t + h_{jk}{}' IW_{jk}{}^t + K(TI_{jk} - INV_{jk}{}^t)^2] \Big\} + \sum_{\forall j \ni s(j)=\emptyset} \sum_{t} N_j LS_j{}^t \qquad 3.35$$

subject to:

Constraints 3.18, 3.19, and 3.21 - 3.34; and

$$IB_j{}^t = IB_j{}^{t-1} + X_j{}^{t-PLj} - D_j{}^t + LS_j{}^t \qquad \forall j \ni s(j)=\emptyset;\ t \geq 1 \qquad 3.36$$

where

$LS_j{}^t$ = Lost sales at node j during period t

N_j = Cost of a lost sale at node j

The model of this section, like that of Section 3.3.3, is a mixed integer, nonlinear optimization model. Again the prominent decision variables are target inventories and restoration coefficients. Some basic strategies for solving these mixed-integer, nonlinear models are given in the next section.

3.4 Solution Strategy

Mixed-integer models are typically solved using the technique of branch and bound. The models of Sections 3.3.3 and 3.3.4 are both mixed-integer and nonlinear requiring

numerous solves of the nonlinear model to complete each branch and bound procedure. As a consequence, it is computationally prohibitive to solve either model for networks with more than a few nodes. However, the models can be solved for larger problems using commercially available software if the integrality constraints given by 3.31 through 3.33 are relaxed. The relaxation of the integrality constraints requires that setup costs be removed from the model. In turn, this requires that the term $\delta_{jk}^{t}S_{jk}$ be removed from the objective function and also that constraint 3.31 be discarded. This relaxation is not unreasonable given that the model has been developed to apply to the manufacture of a single item (rather than multiple items).

There are two commonly used strategies for solving nonlinear optimization problems. First, methods based on linearized subproblems have received considerable attention over the last two decades (e.g., Schittkowsky, 1980; Drud, 1985). This technique requires the nonlinear constraints to be linearized and commonly uses an augmented Lagrangian function for the objective function. Feasibility may not be obtained until late in the solution process. This technique tends to work well for problems in which the only nonlinearity is in the objective function, but it may not work well for highly nonlinear models (Drud, 1985; 1992).

The second method is referred to as generalized reduced gradient (GRG). This technique first searches for a feasible solution, then follows along a feasible path using reduced gradients to direct the line search. GRG is more functional than the method of linearized subproblems for highly nonlinear models according to Drud (1985; 1992). GAMS/MINOS is an example of commercially available software based on linearized subproblems and GAMS/CONOPT is an example of commercially available software based on the method of GRG.

We cannot demonstrate the convexity of the constraint set, specifically the nonlinear constraints given by the production control rule (3.28). As a result, it is not possible to guarantee that any particular solution is globally optimal. Instead, we will seek to reassure ourselves that we have "good" solutions by exploring the stability of the models to different demand sequences during the simulation studies described in Chapters 4 and 5.

3.5 Summary

Production Authorization Cards was modified to include the restoration concept through the use of four parameters and a central controller. The modified Production Authorization Cards system was shown to subsume MRP, kanban, and restoration among other control strategies. For completeness, a math programming model without a production control rule was given in Section 3.3.2. A model with a

production control rule for the case of no lost sales was given in Section 3.3.3. Finally, a model with a production control rule for the case of lost sales was presented in Section 3.3.4. Both models with the production rule incorporate three of the four parameters used in the modified version of Production Authorization Cards and rely heavily on the concepts underlying the restoration control strategy.

If solvable, the models of Sections 3.3.3 and 3.3.4 would specify parameters that would determine the best control rule for shipment quantities between nodes as well as target inventories at each node. Unfortunately, the models have proved to be difficult to solve for the general case. A solution strategy is proposed in which setup costs are removed from the model and integrality constraints are relaxed. This relaxation allows the model to be solved using commercially available software, although global optimality cannot be guaranteed due to the nonconvexity of constraints. Finally, two different solution strategies for solving nonlinear optimization models are reviewed.

CHAPTER 4 METHODOLOGY

4.1 Introduction

The purpose of this chapter is to describe the basic questions we wish to explore in regards to the restoration models developed in Chapter 3. The methodology for answering these questions will also be developed. Incidently, the word "restoration" is henceforth used to refer to either the restoration model without lost sales or the restoration model with lost sales.

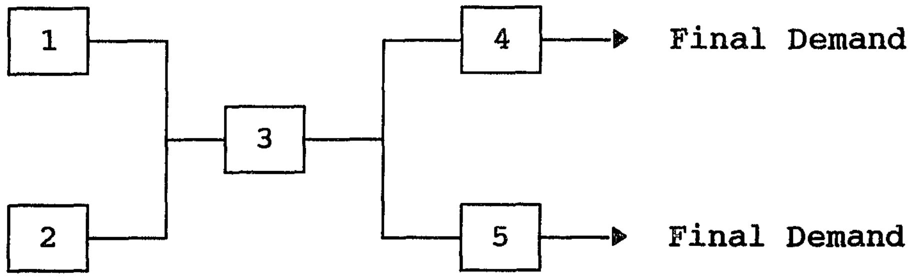
We have two fundamental goals. First, we wish to explore the effect of labor costs and holding costs on target inventories, actual inventories, and restoration coefficients. Second, we also wish to examine the effect of demand variability on target inventories, actual inventories, and restoration coefficients. Values for actual inventories say something about the amount and location of safety stock. Values for target inventories, actual inventories, and restoration coefficients together impute the "best" control strategy. Our underlying motivation is to identify general guidelines for relevant policy decisions implementable by management.

The issues introduced above will be studied in terms of a specific network. The network chosen is *loosely* patterned after the global manufacturing network Honda Motor Company uses to manufacture the Honda Accord automobiles. A

schematic of the network used is shown in Figure 4-1. As you can see, it is a five node, conjoined network with clear global connotations. Detailed information regarding costs, leadtimes, and demands for the base case can be found in Appendix A.

Two different simulation studies will be done. The first simulation holds demand variability constant and varies the levels of labor costs and holding costs. The levels for both of these factors are given in Section 4.2.1. The second simulation study holds labor costs and holding costs constant and varies the levels of demand variability. The levels of the factor demand variability are given in Section 4.2.2.

The idea of rolling horizons will be reviewed in Section 4.2.3. Rolling horizons represent a commonly used strategy for coping with forecast errors as well as deviations from planned production schedules. All test procedures proposed in this chapter and reported on in Chapter 5 will be done under rolling horizons. The method for calculating the length of the warmup period as well as the number and length of batches is given in Section 4.3. Finally, the experimental design is given in Section 4.4 along with the specific hypotheses that will be tested. The results of the simulation are presented in Chapter 5.



Node	Location
1	Japan
2	Mexico
3	Ohio
4	Europe
5	North America

FIGURE 4-1
SPECIFIC NETWORK USED FOR COMPARISONS

4.2 Overview

4.2.1 Labor Costs and Holding Costs as Factors

The primary decision variables of the restoration models are target inventories and restoration coefficients. In effect, order quantities are dependent variables that can be calculated given target inventories, actual inventories, and restoration coefficients. Our first objective is to determine the effect of labor costs and/or holding costs on the values of target inventories, actual inventories, and restoration coefficients. A 2x2 full factorial simulation experiment will be conducted using labor costs and holding costs as factors.

Comparisons will be made using the following general guidelines. The solving of the model for a specific network

with specific costs and specific forecasts of demands will result in associated values for target inventories and restoration coefficients. The solving of the model for the same network with an identical sequence of demands but different costs will result, possibly, in different values for target inventories and restoration coefficients. The design of the simulation study itself is given in Sections 4.3 and 4.4. Factor levels for labor costs and holding costs are identified in this section.

The restoration models were designed to apply to global networks involved in high volume, repetitive manufacturing. This type of manufacturing tends to have more automation thus lower labor costs, than do other types of manufacturing. For example, labor costs as a percent of total costs in the automotive industry vary from approximately 15% for Toyota to near 30% for General Motors (Economic Strategy Institute, 1992). At any rate, we somewhat arbitrarily set the low level of labor costs at 12.5% of value added and the high level at 25% of value added.

Holding costs include the time value of money, obsolescence, storage costs, etc. The principal component of holding costs, the time value of money, has varied widely as indicated by prime interest rates over 20% in 1980 versus less than 10% in 1992. Accordingly, the low level of holding cost is set at 15% and the high level is set at 30%.

The levels of both labor costs and holding costs are given in Table 4-1. We feel that the values for the levels of both types of costs are reasonable and that they encompass a wide range of manufacturing environments.

TABLE 4-1
COST FACTORS AND LEVELS

FACTOR	LEVEL	
	Low	High
Labor Costs	12.5%	25%
Holding Costs	15%	30%

4.2.2 Variability of Demand as a Factor

Our next objective is to determine the effect of demand variability on target inventories, actual inventories, and restoration coefficients. A single factor simulation experiment will be conducted with coefficient of variation as the sole factor. Again, the goal is to interpret results with regard to policy decisions implementable by managers and generalize results where possible.

The specific network under study has two nodes that meet final demand (see Figure 4-1). Average demands for these nodes differ since sales in Europe (node 4) average less than sales in North America (node 5). Since the means differ, factor levels for demand variation are given in terms of coefficients of variation rather than standard deviations. Recall that:

$$\text{Coefficient of Variation} = \frac{\text{Standard Deviation}}{\text{Mean}}$$

The use of coefficient of variation allows demand variation for nodes with different means to be standardized. For example, assume that nodes 4 and 5 have average demands of μ_4 and μ_5 respectively. If a coefficient of variation of 12% is used for both, then:

$$\frac{\sigma_4}{\mu_4} = 0.12 \quad \Rightarrow \quad \sigma_4 = 0.12\mu_4 \quad (= 12\% \text{ of } \mu_4)$$

$$\frac{\sigma_5}{\mu_5} = 0.12 \quad \Rightarrow \quad \sigma_5 = 0.12\mu_5 \quad (= 12\% \text{ of } \mu_5)$$

A constant coefficient of variation results in standard deviations that can be expressed as a percentage of their respective means. Likewise, confidence intervals for normally distributed random variables can also be expressed as a percent of their respective means.

The levels chosen for the factor coefficient of variation are given in Table 4-2. A factor level corresponding to a coefficient of variation of 0.01 is included for purposes of comparison only since few manufacturers face so low a variability in demands. Note that the nonlinear optimizers used to solve the model were unable to handle the case of level demand since it required the inverting of a singular matrix.

The asterisk beside the coefficient of variation of 0.12 means that this level of variation is roughly equivalent to that for the base case in Appendix A. Demands at the European and North American nodes are assumed to be

normally distributed with means of 1,000 cars per week and 1,500 cars per week, respectively. The following data may help put a perspective on the chosen levels for coefficients of variation:

Coefficient of Variation	95% Confidence Interval for Total Weekly Demands
0.01	2,500 ± 36
0.06	2,500 ± 216
0.12	2,500 ± 432
0.18	2,500 ± 649

We feel that this level of variation of demand is reasonable and includes the levels of variation faced by many manufacturers.

TABLE 4-2
LEVELS OF THE FACTOR COEFFICIENT OF VARIATION

FACTOR	LEVELS			
Coefficient of Variation	0.01	0.06	0.12*	0.18

* Roughly equivalent to the base network in Appendix A.

4.2.3 Rolling Horizons

All comparisons will be made using rolling horizons. Rolling horizons represent a commonly used technique that manufacturers use to cope with change and uncertainty. The concept is based on a planning horizon of finite length for which there exist forecasts of demand. A production plan is constructed for the planning horizon but only decisions relevant to the first few periods are implemented. The planning horizon is then rolled forward, initial inventories

and demand forecasts are adjusted, and a new production plan is constructed.

Forecasts over the planning horizon are required each time the model is solved. The planning horizon should be long enough to smooth short term fluctuations in demand, yet short enough to efficiently coordinate production with demands. The length of the planning horizon is outside the realm of this study. The length of the planning horizon tends to vary by industry; typical lengths for a particular industry can usually be found in the literature. A planning horizon of 24 weeks, or approximately 6 months, will be used for the network of Figure 4-1. The term "solve" is henceforth used to refer to one solve of the model associated with a specific planning horizon.

Only decisions pertaining to the first few periods are implemented under a rolling horizon strategy. For example, each solution of the restoration model provides values for target inventories and restoration coefficients that influence decisions for only a few periods. The period of time over which these decisions are implemented is henceforth referred to as the implementation period. The length of the implementation period also tends to be industry specific. Again, typical values for a specific industry can usually be found in the literature. An implementation period of 4 weeks, or approximately 1 month, will be used for the network of Figure 4-1. Each solve of

the model is associated with one planning horizon and one implementation period.

Each time a planning horizon is rolled forward, the first few forecasts of demand are dropped and an equivalent number of new demand forecasts are tacked onto the end. As a result, a somewhat different forecast of demand occurs each time the planning horizon is rolled forward. New and different sets of demands from those of the initial planning horizon will be encountered if the planning horizon is rolled forward a sufficient number of times.

The model is designed to minimize costs over a planning horizon. However, these are not the costs that are actually incurred by the firm. The relevant costs to a manufacturer are those incurred, in other words those costs associated with implementation periods. The actual costs over an extended time period are the sum of costs over adjacent implementation periods.

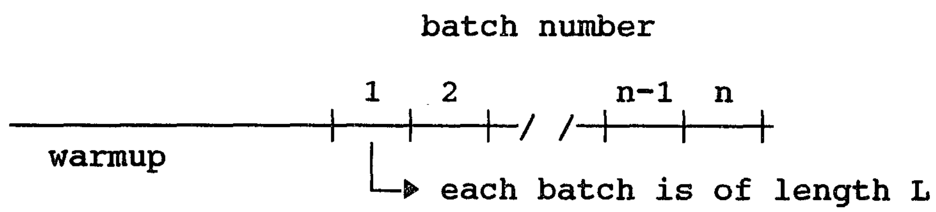
4.3 Preliminaries to the Experimental Design

4.3.1 Introduction

The technique to be used for forming replicates at each treatment level is batch means, a commonly used technique in simulation (e.g., Law and Kelton, 1982). This technique requires a single, long simulation run for each treatment level. Each simulation run requires numerous solves of the nonlinear model along adjacent planning horizons. Demands for different simulation runs will be generated from a

common random number stream. This is a variance reduction technique that effectively allows control strategies to be compared under similar experimental conditions (see Law and Kelton, 1982).

The length of the simulation run associated with each treatment level can be thought of in terms of (w + nL). The value w represents the length of the warmup period, a period of time at the beginning of the simulation during which no statistics are compiled. The purpose of the warmup is to reduce the influence of artificially chosen initial conditions on subsequent estimates of performance measures. Following the warmup, the simulation consists of n batches (replicates) each of length L:



Average values for target inventories, actual inventories, and restoration coefficients will be calculated from each batch. These values are assumed to be independent and identically distributed estimates of the true values for the specific set of costs and coefficient of variation used with that simulation run. These batch means, or replicates, will then be used in analysis of variance tests for significance.

Before proceeding, it is necessary to establish the length of the warmup period (w) as well as the length (L) of each batch and the number of batches (n). The length of the warmup period will be chosen using the average cost per period. The length used will be the greater of:

1) the number of periods required for the average cost per period to converge to a stable value; or
2) the length of the planning horizon.

The length of the warmup period will be established by applying the restoration model with no lost sales to the base network described in Appendix A. The same warmup period will then be used for all other factors and levels. Methods for fixing the length and number of batches will be discussed in the next two sections.

4.3.2 Batch Length

The analysis of variance tests we use require the assumption that batch means are independent and identically distributed according to a normal distribution. Of these assumptions, Law and Kelton (1982) present evidence that the correlation between batch means (the independence assumption) is potentially the most serious source of error for many simulation studies. Accordingly, a basic goal is to make the batches of significant length to minimize correlations between adjacent batch means.

On the other hand, there is a cost associated with large batch sizes, namely it may become computationally

infeasible to carry out the simulation. Each solve of the restoration model for the network of Figure 4-1 with a planning horizon of 24 periods involves in excess of 1,000 constraints. Forty batches of twenty implementation periods each would require the nonlinear model with over 1,000 constraints to be solved 800 times for each level of each factor-a prohibitive requirement. Clearly, it is necessary to find a balance. Batch lengths should be long enough to reduce correlation between batch means, but short enough to be computationally feasible.

The procedure we choose to follow for determining the batch lengths parallels the approach outlined by Law and Kelton (1982). First, the restoration model will be run for a long sequence of demands. Average costs per period will be computed assuming batch lengths of 2 periods, then of 3 periods, then of 4 periods, and so on. In each case, the autocorrelation of lag 1 for batch means will be calculated. Finally, a graph of autocorrelation of lag 1 of batch means versus length of batch will be constructed. The batch length is chosen so that its autocorrelation of lag 1 is less than 40% and also so that no longer batch length has an autocorrelation of lag 1 that exceeds 40%.

In no event will the batch length be shorter than the length of the planning horizon. As a result, each batch will be composed of at least two planning horizons that have no demands in common. The batch length will be established

using the restoration model with no lost sales as applied to the base network described in Appendix A. The same batch length will then be used for all levels of all other factors.

4.3.3 Number of Batches

The procedure we plan to use to determine the required number of batches follows that presented by Montgomery (1991). A set of preliminary runs using the restoration model without lost sales will be made for the base network of Appendix A. These runs will provide estimates of standard deviations which can then be used to find the required number of replicates.

The procedure begins with a prespecified probability (say 95%) of detecting differences in target inventories that exceed a prespecified value. A difference of 2 for target/actual inventories, which equates to about 5% of average values, was chosen based on preliminary runs. A difference of 0.1 or about 10% of average restoration coefficients was chosen based on the same preliminary runs. An iterative process is then used involving the following equation taken from Montgomery (1991):

$$\phi^2 = \frac{nD^2}{2\sigma^2} \qquad 4.1$$

where

n = Number of replicates

D = Prespecified difference we seek to delineate

σ^2 = Standard deviation of target inventories

If the null hypothesis of no treatment differences is false, the statistic

$$F = \frac{MS_{Treatments}}{MS_{Error}} \qquad 4.2$$

is distributed as a noncentral F random variable. The parameter ϕ^2 is related to the noncentrality of this distribution (see Montgomery for details).

The process is to solve 4.1 for ϕ using different values of the number of replicates (n). Operating curves are used for each ϕ to find the associated probability of failing to reject the null hypothesis given that it is false (Type II error). The number of replicates needed is the minimum that provides a satisfactory probability of a Type II error (say 5%).

The procedure leads to a required number of replicates for target inventory or actual inventory at a specific node, or for the restoration coefficient between two nodes. Our goal is to find the required number of replicates to handle the worst case of the four target inventories, four actual inventories, and eight restoration coefficients in the base case network of Appendix A. After determining the number of replicates, the same number will be applied to each level of each factor. However, it may be prohibitive to run the number of replicates suggested by this procedure. If so, we

will determine which specific target inventories, actual inventories, and/or restoration coefficients have abnormally high standard deviations. It may not be possible to test for significant differences for any decision variables with very high standard deviations simply due to computational constraints.

4.4 Experimental Design

4.4.1 2^2 Full Factorial-Labor Costs and Holding Costs

A 2^2 full factorial experiment will be done with labor costs and holding costs as factors. The coefficient of variation will be held constant, corresponding to the base network of Appendix A. Specific values for the levels of both types of costs were given in Section 4.2.1. A long simulation run consisting of a warmup period of length w, plus n batches of length L will be done at each factor level combination where w, n, and L are derived as indicated in Sections 4.3.2 and 4.3.3. The same values of w, n, and L will be used for each simulation run. As a result, there will be n replicates at low labor/low holding costs, n replicates at low labor/high holding costs, and so on. These replicates will form the basis for the analysis of variance that will now be described.

We assume that a linear relationship exists between target inventories and factors. The linear relationship is illustrated by the following equation:

$$TI_{ijk} = \mu + L_i + H_j + (LH)_{ij} + \epsilon_{ijk} \qquad 4.3$$

where

TI_{ijk} = target inventory with labor at ith level, holding at jth level, and kth replicate

μ = overall mean effect

L_i = effect of the ith level of labor costs

H_j = effect of the jth level of holding costs

$(LH)_{ij}$ = effect of the interaction between L_i and H_j

ϵ_{ijk} = a random error component

This same basic linear relationship is assumed to apply to each target inventory, actual inventory at each node, and each restoration coefficient.

We are interested in the main effects due to both labor costs and holding costs in addition to any interaction effect between labor costs and holding costs. The specific hypotheses we propose to test are:

H_0: $L_1 = L_2 = 0$ Main Effect of Labor Costs

H_1: at least one $L_i \neq 0$

and

H_0: $H_1 = H_2 = 0$ Main Effect of Holding Costs

H_1: at least one $H_i \neq 0$

and

H_0: $(LH)_{ij} = 0 \quad \forall\, i,j$ Interaction Effect

H_1: at least one $LH_{ij} \neq 0$

These three hypotheses will be tested independently for each target inventory, actual inventory at each node, and for each restoration coefficient. The analysis of variance itself is straight forward and can be found in any text on experimental design (e.g., Montgomery, 1991).

4.4.2 Single Factor-Variability of Demand

A separate, single factor experiment will be done using the low labor cost and high holding cost base network of Appendix A. The factor coefficient of variation of demand will have the four levels described in Section 4.2.2 (0.01, 0.06, 0.12, and 0.18). The simulation at each of the four levels will consist of a warmup of length w followed by n batches of length L, where w, n, and L are identical to those used in 4.4.1. Therefore, we will end up with n replicates at a coefficient of variation of 0.01, n replicates at a coefficient of variation of 0.06, and so on.

A linear relationship is assumed to exist between target inventory and the levels of the factor coefficient of variation. The linear relationship is expressed by the following equation:

$$TI_{ik} = \mu + V_i + \epsilon_{ik} \qquad 4.4$$

where

TI_{ik} = Target inventory with coefficient of variation at the ith level and kth replicate

μ = Overall mean effect

V_i = Effect of the ith level of coefficient of variation

ϵ_{ik} = A random error component

The same basic linear relationship is assumed to apply to each target inventory, actual inventory at each node, and each restoration coefficient.

In this case, we are interested in the main effect. The hypothesis we propose to test is:

H_0: $V_1 = V_2 = V_3 = V_4 = 0$

H_1: at least one $V_i \neq 0$

This hypothesis will be tested for each target inventory, actual inventory at each node, and each restoration coefficient. Again, the analysis of variance is straight forward and can be found in a text on experimental design.

4.5 Summary

Two distinct simulation experiments were proposed. The first one explores the effects of the two factors labor costs and holding costs on target inventories, actual inventories, and restoration coefficients. The second experiment looks at the effect of the factor variability of demand on target inventories, actual inventories, and restoration coefficients.

The simulation technique proposed for both experiments is batch means. This technique requires a long simulation run at each combination of factor levels. A method is

presented for determining the length of the warmup period (w) as well as the length (L) and number of batches (n). Once the values of w, n, and L are determined, these same values are used for each combination of factor levels.

Both experiments will be analyzed using analysis of variance. The first experiment will be a 2^2 full factorial experiment with the factors labor costs and holding costs. The second experiment will be a single factor experiment with the factor coefficient of variation. Specific tests of hypothesis are presented. In the next chapter, the design developed in this chapter is applied to a specific network.

CHAPTER 5 EXPERIMENTAL RESULTS

5.1 Introduction

This chapter contains the results of two separate simulation experiments. In both experiments, a network equivalent to the base network of Appendix A is used as one of the factor level combinations. The first simulation experiment is a 2^2 full factorial experiment with factors labor costs and holding costs. The second simulation experiment is a single factor experiment with the factor being coefficient of variation.

The network used in both experiments is identical to that shown in Figure 4-1 and is repeated in Figure 5-1 for convenience. The nodes are numbered as they are referred to throughout this chapter. The terminology arc 1-3 refers to an arc-node pair: the arc from node 1 to node 3 along with node 3. Therefore, actual inventory along arc 1-3 or equivalently actual inventory 1-3 refers to the sum of inventory:

in-transit from node 1 to node 3, plus

in-wait at node 3 (from node 1), plus

in-process at node 3, plus

in-buffer at node 3.

Similarly, target inventory 1-3 refers to target inventory along arc 1-3 but restoration coefficient 1-4 refers to the restoration coefficient linking node 1 to node 4.

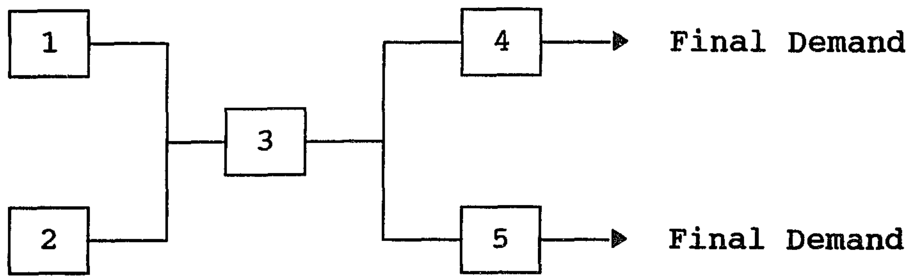


FIGURE 5-1
SPECIFIC NETWORK TO BE STUDIED

5.2 Warmup, Batch Length, and Number of Batches

The criteria outlined in Chapter 4 was used to find the appropriate values for length of warmup, batch length, and number of batches. Our approach was to first determine these values separately for the model without lost sales and the model with lost sales. The goal is to use the more conservative of the two values for length of warmup, batch length, and number of batches for all subsequent simulations.

Both models were applied to the base network of Appendix A. In each case, the model was solved (using GAMS/CONOPT) 200 times under a rolling horizon concept with a planning horizon of 24 weeks and an implementation period of 4 weeks. Two sets of 1,200 demands were generated from within GAMS (seed = 3945) based on N(10,σ=1) and N(15,σ=2) distributions. Note that demands were scaled down from N(1000,σ=100) to N(10,σ=1) and from N(1500,σ=200) to

N(15,σ=2) to aid in the convergence of the nonlinear algorithm. Identical demands were used for both models.

Based on the criteria of Section 4.2.1, a warmup of 28 periods is sufficient for either model. Supporting data for this conclusion and those that follow can be found in Appendix B. First, the criteria of sections 4.3.2 and 4.3.3 were applied to the model without lost sales. It was determined that over 100 replications of batch length 13 solves each are required to test the hypotheses at the desired levels for this form of the model. On this basis, the proposed simulations require over 12,000 solves of the nonlinear model, each containing over 1,000 constraints. Computational requirements are such that we are unable to do the simulations using this form of the model. Further problems related to the model without lost sales are discussed in the next section.

Next, the criteria of Sections 4.3.2 and 4.3.3 were applied to the model with lost sales. Again, supporting data is provided in Appendix B. It was found that 65 replications of batch length 7 solves each are required to test the hypotheses at the desired levels for this form of the model. Although the computational requirements are significantly less than those associated with the model without lost sales, the proposed simulations still require nearly 4,000 solves of the nonlinear model.

Table 5-1 shows the probability of a Type II error (β-level), corresponding to an α-level of 1%, based on 20 replications of the model with lost sales. Notice that the β-level is satisfactory for all parameters other than restoration coefficients 1-3 and 2-3. Clearly, variances associated with estimates of these two parameters are quite large. In Section 5.4, we show that these two restoration coefficients are largely irrelevant for other reasons. Accordingly, a decision was made to compute 20 replicates at each factor level combination for the restoration model with lost sales.

TABLE 5-1
PROBABILITY OF A TYPE II ERROR (β-LEVEL)
RESTORATION MODEL WITH LOST SALES
BASED ON 20 REPLICATES AND AN α-LEVEL OF 1%

Parameter	β-level
Target Inventory 1-3	0.02
Target Inventory 2-3	< 0.01
Target Inventory 3-4	< 0.01
Target Inventory 3-5	< 0.01
Restoration Coefficient 1-3	-
Restoration Coefficient 1-4	< 0.01
Restoration Coefficient 1-5	< 0.01
Restoration Coefficient 2-3	-
Restoration Coefficient 2-4	< 0.01
Restoration Coefficient 2-5	< 0.01
Restoration Coefficient 3-4	< 0.01
Restoration Coefficient 3-5	< 0.01

5.3 Results Based on the Restoration Model without Lost Sales

The solution process for the base network was stopped by GAMS/CONOPT on several occasions during the process of the 200 solves of the restoration model without lost sales. In order to restart the simulation, it was necessary to either:

1) adjust the starting guess for target inventories and restoration coefficients; and/or
2) back the process up as many as three or four solves from the point at which it stopped.

We were able to continue the solution process for this specific example until all 200 solves were calculated. However, there is no guarantee that this will always be the case. It is quite possible that the restoration model without lost sales may simply be infeasible from time to time.

Figure 5-2 shows the total 24 period cost versus the number of the solve for the restoration model without lost sales applied to the base network of Appendix A. As you can see, the total cost function contains numerous "spikes" composed of one or two solves in which the cost is abnormally high compared to most of the other solves. Table 5-2 decomposes the total costs into those due to labor costs, holding costs, and costs based on the deviation of inventory from target. Averages and standard deviations of

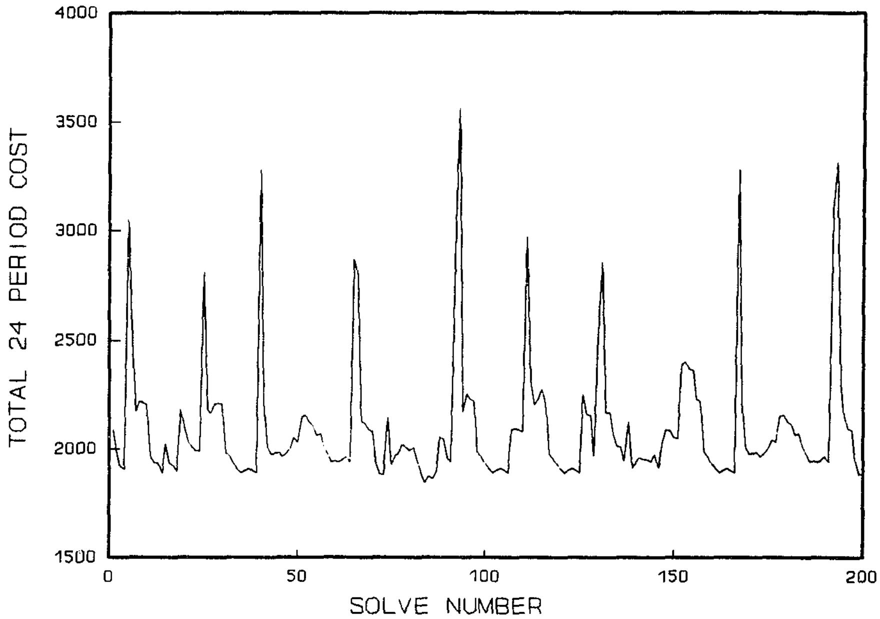


FIGURE 5-2
TOTAL 24 PERIOD COSTS
RESTORATION MODEL WITHOUT LOST SALES
LOW LABOR/HIGH HOLDING COSTS

these values are also shown. The relatively small standard deviations associated with labor costs and holding costs suggest that these values remain relatively stable. In contrast, the cost of inventories deviating from target varies widely as shown by the associated large standard deviation. In fact, the abnormally large costs apparent in Figure 5-2 are caused by abnormally large values of the term $K(TI_{jk}-INV_{jk}^{t})^2$ in the objective function.

TABLE 5-2
DECOMPOSITION OF TOTAL 24 PERIOD COSTS
BASED ON 200 SOLVES OF THE
RESTORATION MODEL WITHOUT LOST SALES
LOW LABOR/HIGH HOLDING COSTS

Cost	Average 24 Period Cost	Standard Deviation
Labor Costs	1,846.9	72.1
Holding Costs	78.3	38.5
Costs related to $K(TI-INV)^2$	170.5	255.7

Table 5-3 shows averages and standard deviations for target inventories, actual inventories, and restoration coefficients. As you can see, standard deviations are relatively small for all four actual inventories as well as target inventories along arcs 1-3 and 2-3. In contrast, target inventories along arcs 3-4 and particularly 3-5 have large standard deviations. Figure 5-3 shows total 24 period costs versus the sum of target inventory 3-4 and target inventory 3-5. The figure shows that, in every case, high

total costs occurred concurrently with high target inventories along arc 3-4 and/or arc 3-5.

TABLE 5-3
AVERAGES AND STANDARD DEVIATIONS OF DECISION VARIABLES
BASED ON 200 SOLVES OF THE
RESTORATION MODEL WITHOUT LOST SALES
LOW LABOR/HIGH HOLDING COSTS

Parameter	Average	Standard Deviation
Target Inventory 1-3	81.4	3.54
Target Inventory 2-3	55.6	2.79
Target Inventory 3-4	45.7	6.01
Target Inventory 3-5	54.3	12.08
Actual Inventory 1-3	80.4	2.81
Actual Inventory 2-3	55.2	2.43
Actual Inventory 3-4	33.2	1.51
Actual Inventory 3-5	35.8	2.82
Restoration Coefficient 1-3	0.56	0.424
Restoration Coefficient 1-4	0.88	0.239
Restoration Coefficient 1-5	0.85	0.224
Restoration Coefficient 2-3	0.64	0.414
Restoration Coefficient 2-4	0.88	0.255
Restoration Coefficient 2-5	0.88	0.207
Restoration Coefficient 3-4	0.89	0.197
Restoration Coefficient 3-5	0.90	0.205

In fact, occasions occur when using this form of the model when the best solution has an abnormally large target inventory along arc 3-4 and/or arc 3-5. We note that abnormally large target inventories were not observed along arc 1-3 nor along arc 2-3. Figure 5-4 shows target inventory 3-5 versus the associated restoration coefficient along arc 3-5. As you can see, abnormally large target inventories are consistently associated with abnormally

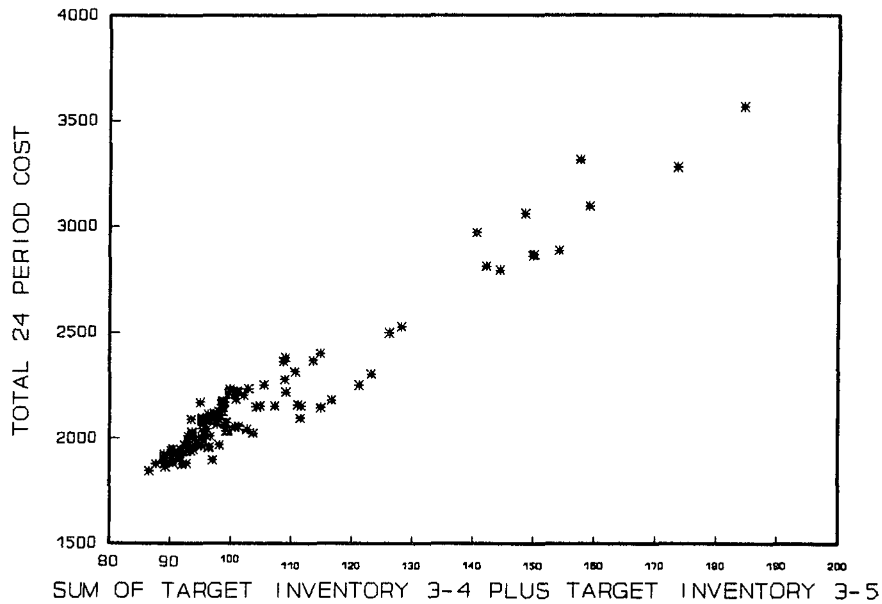


FIGURE 5-3
TOTAL COST VERSUS TARGET INVENTORY ALONG ARCS 3-4 AND 3-5
BASED ON 200 SOLVES OF THE
RESTORATION MODEL WITHOUT LOST SALES
LOW LABOR/HIGH HOLDING COSTS

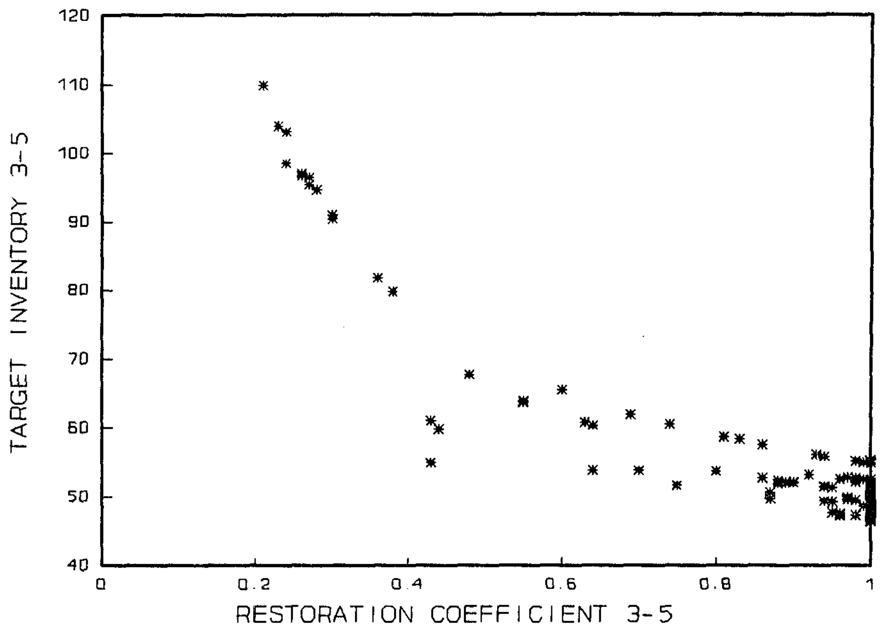


FIGURE 5-4
TARGET INVENTORY 3-5 VERSUS RESTORATION COEFFICIENT 3-5
BASED ON 200 SOLVES OF THE
RESTORATION MODEL WITHOUT LOST SALES
LOW LABOR/HIGH HOLDING COSTS

small restoration coefficients. This relationship also holds for target inventory 3-5 and restoration coefficients 1-5 and 2-5 as well as target inventory 3-4 and its associated restoration coefficients. In summary, abnormally large target inventories are offset by abnormally low restoration coefficients so that order quantities remain relatively stable.

At any rate, we are faced with the dilemma that the restoration model without lost sales:

1) may occasionally be infeasible; or
2) may occasionally converge to a solution that has abnormally high target inventories and abnormally low restoration coefficients.

Neither of these two occurrences are positive. In effect, occasional model infeasibility is a statement that it is not always possible to simultaneously satisfy the production control rule 3.28 and meet all demands-an unacceptable outcome. Occasional, abnormally high values for target inventories are also an unacceptable outcome since target inventories are aggregate values that are to be disaggregated to finer levels of detail. This level of variation in target inventories would place the shop floor into disarray.

We conclude that the restoration model without lost sales is based on a severely limiting assumption. The assumption is that level target inventories and restoration

coefficients that satisfy all constraints, including the constraint of meeting all demands, always exist. The limits of the restoration model without lost sales have been demonstrated; we now turn our attention to the form of the restoration model that allows lost sales.

5.4 Labor Costs and Holding Costs as Factors (Lost Sales)

5.4.1 General

A 2^2 full factorial experiment was done using the restoration model *with lost sales* and the factor levels of labor costs and holding costs described in Chapter 4. Each factor level combination (e.g., low labor/low holding, low labor/high holding, etc.) was simulated with 168 solves of the nonlinear model along a rolling horizon. For each factor level combination, results from the first 28 solves were discarded and 20 batches of 7 solves each were used in tests for significant differences.

GAMS/CONOPT did not stop the execution during any one of the four simulation runs in contrast to the restoration model without lost sales in which the execution was halted numerous times. Results from the model are also significantly more stable in the sense that the occasional, abnormally high costs associated with the model without lost sales no longer occur. This result is seen graphically by contrasting the total costs for the restoration model with lost sales in Figure 5-5 to the total costs from the restoration model without lost sales in Figure 5-2.

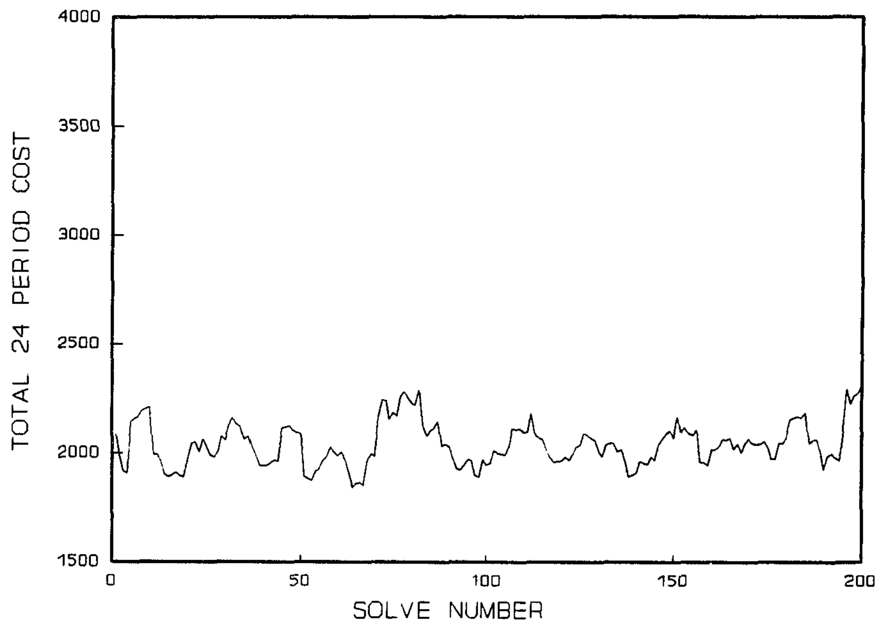


FIGURE 5-5
TOTAL 24 PERIOD COSTS
RESTORATION MODEL WITH LOST SALES
LOW LABOR/HIGH HOLDING COSTS

Table 5-4 shows averages and standard deviations of batch means of target inventories, actual inventories, and restoration coefficients for the base network. As you can see from the data, standard deviations are generally small compared to averages with the exception of those associated with restoration coefficients 1-3 and 2-3. We will return to these two restoration coefficients momentarily.

TABLE 5-4
AVERAGES AND STANDARD DEVIATIONS OF BATCH MEANS
BASED ON 20 REPLICATES OF THE
RESTORATION MODEL WITH LOST SALES
LOW LABOR/HIGH HOLDING COSTS

Parameter	Average	Standard Deviation
Target Inventory 1-3	80.8	1.71
Target Inventory 2-3	55.2	1.23
Target Inventory 3-4	43.0	0.95
Target Inventory 3-5	49.5	1.33
Actual Inventory 1-3	79.9	1.57
Actual Inventory 2-3	54.8	1.25
Actual Inventory 3-4	32.5	0.78
Actual Inventory 3-5	34.3	1.63
Restoration Coefficient 1-3	0.65	0.179
Restoration Coefficient 1-4	0.97	0.035
Restoration Coefficient 1-5	0.92	0.060
Restoration Coefficient 2-3	0.65	0.222
Restoration Coefficient 2-4	0.98	0.023
Restoration Coefficient 2-5	0.94	0.045
Restoration Coefficient 3-4	0.96	0.002
Restoration Coefficient 3-5	0.99	0.001

It can be concluded that results using the model with lost sales are computationally more tractable and are more

stable than those derived when using the model without lost sales. The question becomes one of at what cost. Table 5-5 shows average lost sales at nodes 4 and 5 for the base network as a percent of expected demand over the implementation period. It also shows the standard deviation of lost sales at both nodes. As you can see, lost sales average only 0.35% of expected demand at node 4 and 0.46% of expected demand at node 5; further, both standard deviations are relatively small. Clearly, the stable results of the model with lost sales come at a reasonable expense in terms of the amount of lost sales.

TABLE 5-5
FOUR PERIOD LOST SALES
AS A PERCENT OF EXPECTED DEMAND
RESTORATION MODEL WITH LOST SALES
LOW LABOR/HIGH HOLDING COSTS

Node	Average of Lost Sales*	Standard Deviation
4	0.35%	1.128
5	0.46%	1.669

* As a percent of expected demands over the implementation period

In Section 5.2, it was shown that it was computationally prohibitive to do the proposed simulations using the restoration model without lost sales. In Section 5.3, it was shown that this form of the model may occasionally: (1) be infeasible; or (2) result in abnormal

values for target inventories and restoration coefficients. In contrast, we have seen that the restoration model with lost sales is readily solvable, is computationally tractable with respect to the proposed simulations, and results in small amounts of lost sales. We conclude that the restoration model with lost sales is the only form of the model appropriate for use with the proposed simulation study.

Before proceeding to hypothesis testing, we will explore the control strategy suggested by the values for the restoration coefficients chosen by the model. Table 5-6 presents grand means based on 20 batches for target inventories, actual inventories at the end of implementation periods, and restoration coefficients for the base network. It also shows the average shortage of actual inventory from target. Notice that the average shortage along arcs 1-3 and 2-3 is very small compared to the average shortage along arcs 3-4 and 3-5. Notice also that most of the restoration coefficients are near 100%, but restoration coefficients 1-3 and 2-3 both average 65%.

The production control rule specifies that the order quantity at node 1 is a linear combination of shortages of actual inventories from target at downstream nodes. This relationship for the order quantity at node 1 is expressed algebraically as follows:

$$Q_{13}{}^{t} = (TI_{13}-INV_{13}{}^{t-1})r_{13} + (TI_{14}-INV_{14}{}^{t-1})r_{14} + (TI_{15}-INV_{15}{}^{t-1})r_{15} \qquad 5.1$$

where

TI_{13} = Target inventory 1-3

$INV_{13}{}^{t-1}$ = Actual inventory at t-1 along arc 1-3

r_{13} = Restoration coefficient 1-3

From Table 5-6, we see that the shortage $(TI_{13}-INV_{13}{}^{t-1})$ is on average quite small compared to shortages $(TI_{34}-INV_{34}{}^{t-1})$ and $(TI_{35}-INV_{35}{}^{t-1})$.

TABLE 5-6
GRAND MEANS FOR INVENTORIES & RESTORATION COEFFICIENTS
BASED ON 20 BATCHES
RESTORATION MODEL WITH LOST SALES
LOW LABOR/HIGH HOLDING COSTS

Arc	Target Inventory	Actual Inventory*	Shortage (TI - INV)	Restoration Coefficient	Average Value
1-3	80.78	79.85	0.92	1-3	0.65
2-3	55.20	54.81	0.38	1-4	0.97
3-4	42.96	32.49	10.47	1-5	0.92
3-5	49.46	34.26	15.20	2-3	0.65
				2-4	0.98
				2-5	0.94
				3-4	0.96
				3-5	0.99

* Actual inventory is measured at the end of the implementation period.

On average, only 2.5% of the production scheduled at node 1 is due to consumption at node 3. The remaining 97.5%

of the production scheduled at node 1 is due to shortages of inventories from target at the two nodes that meet final demand. The result is very similar at node 2, where an average of 99% of the production scheduled at node 2 is based on shortages from target at the two nodes that meet final demand. These results were very similar for the other factor level combinations of low labor/low holding, high labor/low holding, and high labor/high holding. In effect, results suggest that production at nodes 1 and 2 should be scheduled based almost entirely on shortages of actual inventory from target at nodes 4 and 5, the nodes that meet final demand.

Two conclusions can be drawn. First, the value of the restoration coefficient between nodes 1 and 3 as well as nodes 2 and 3 is largely immaterial. The shortage of actual inventory from target along arcs 1-3 and 2-3 is so small that it makes little difference if the associated restoration coefficient is 0% or 100%, the order quantities at nodes 1 and 2 are approximately the same. It appears as if shortages at nodes 4 and 5 cause production to be "pushed" from node 1 to node 3 and from node 2 to node 3. Sufficient inventory is pushed forward into node 3 so that shortages of actual inventories from target along arcs 1-3 and 2-3 remain small. Although we have not observed it, it is entirely possible that actual inventories along arcs 1-3

and 2-3 could occasionally exceed target due to this "pushing" effect.

The second conclusion is a result of the observation that all of the restoration coefficients to the nodes that meet final demand (nodes 4 and 5) are close to 100% as shown in Table 5-6. This fact suggests that production at nodes 1, 2, and 3 is scheduled based primarily on consumption at nodes 4 and 5. In other words, production is scheduled based primarily on consumption at nodes that meet final demand. Therefore, the model suggests that a strategy very close to base stock is best for the specific network under study. Further, this result holds for all factor level combinations of labor costs and holding costs used in the study.

5.4.2 Tests of Hypotheses

We wish to identify any significant main effects and interaction effects of labor costs and holding costs on target inventories, actual inventories, and restoration coefficients. Our procedure is to test hypotheses separately for each target inventory, actual inventory, and restoration coefficient. Fourteen separate analysis of variances are required, one each for each of the four target inventories, four actual inventories, and six of the eight restoration coefficients. No tests are done for restoration coefficients 1-3 and 2-3 for the reasons stated earlier. Hypotheses testing is done at the 1% level, corresponding to

a 1% probability of accepting the null hypotheses given that it is not true. The number of replicates was chosen so that the β-levels, the probability of accepting the null hypothesis given that it is false, is 2% or less.

Table 5-7 shows the analysis of variance for testing the following hypotheses relevant to target inventory 1-3:

H_0: $L_1 = L_2 = 0$ Main Effect of Labor Costs

H_1: at least one $L_i \neq 0$

H_0: $H_1 = H_2 = 0$ Main Effect of Holding Costs

H_1: at least one $H_i \neq 0$

H_0: $(LH)_{ij} = 0 \quad \forall i,j$ Interaction Effect

H_1: at least one $(LH)_{ij} \neq 0$

Based on data in the table, we cannot reject the hypothesis of no main effect due to labor costs on target inventory along arc 1-3. Nor can we reject the hypothesis of no interaction effect between labor costs and holding costs on the target inventory along arc 1-3. We do reject the hypothesis of no main effect of holding costs on target inventory 1-3. In other words, there may be a causal relationship between the level of holding costs and target inventory 1-3. Higher holding costs result in a lower target inventory along arc 1-3 and vice versa.

Similar analysis of variance tests are performed on the three remaining target inventories, actual inventories along all four arcs, and six of the eight restoration coefficients. The analysis of variance results for all 13

remaining tests are given in Appendix C. A summary of the overall results are given in Table 5-8.

TABLE 5-7
ANALYSIS OF VARIANCE FOR TARGET INVENTORY 1-3
2^2 FULL FACTORIAL EXPERIMENT USING LOST SALES MODEL
LABOR COSTS AND HOLDING COSTS ARE THE FACTORS

Source	DF	Mean-Square	F-Ratio	Prob > F
Labor Costs	1	2.4801	0.90	0.3447
Holding Costs	1	45.2146	16.48	0.0001*
Interaction	1	0.6882	0.25	0.6179
Error	76	2.7428		
Total Sum-Squares	79	256.8344		

* Significant at the 1% level.

In no case were there any significant main effects related to labor costs. In retrospect, this may be because only overtime/undertime costs were considered to be relevant costs and included in the model. Larger holding costs may have "swamped" any effects due to labor costs. Alternatively, labor costs may simply have little bearing on target inventories and restoration coefficients. Likewise, there were no significant interaction effects between labor costs and holding costs.

There were significant main effects of holding costs on target inventory 1-3 and target inventory 2-3. The data in Table 5-8 also show that there are significant effects on actual inventories along 1-3 and 2-3 due to holding costs.

Note the conspicuous absence of any significant differences in either target inventories or actual inventories along arcs 3-4 and 3-5.

TABLE 5-8
SIGNIFICANT RELATIONSHIPS BASED ON THE ANOVA TESTS
2^2 FULL FACTORIAL EXPERIMENT USING LOST SALES MODEL
LABOR COSTS AND HOLDING COSTS ARE THE FACTORS

Significant Effect[1]	Factor	Relationship[2]
TI 1-3 TI 2-3	Holding Costs Holding Costs	Negative Negative
AI 1-3 AI 2-3	Holding Costs Holding Costs	Negative Negative
RC 1-5 RC 2-5 RC 3-5	Holding Costs Holding Costs Holding Costs	Negative Negative Negative

[1] TI 1-3 = Target Inventory 1-3
AI 1-3 = Actual Inventory 1-3
RC 3-5 = Restoration Coefficient from Node 3 to Node 5

[2] A negative relationship implies that higher holding costs result in lower values of parameters. Therefore, higher holding costs results in lower values for all parameters listed in the table.

Apparently, target inventories and actual inventories at the nodes that meet final demand are not influenced by labor costs nor by holding costs, at least with respect to the levels of those two factors used in this study. Holding costs do appear to influence both target inventories and actual inventories at those nodes that do not meet final

demand. At these nodes, higher holding costs result in somewhat lower target inventories and actual inventories.

There appears to be some discretionary room regarding levels of target and actual inventories at nodes that do not satisfy final demands; these values fluctuate depending on holding costs. However, target and actual inventories at nodes that meet final demand do *not* seem to change with respect to changing holding costs. It appears that factors external to labor costs and holding costs are determinants of target inventories and actual inventories at nodes that meet final demand. At this point, it is not yet prudent to generalize this conclusion to other network architectures; additional research is needed.

The results in Table 5-8 also suggest that holding costs influence values of restoration coefficients from node 1 to node 5, from node 2 to node 5, and from node 3 to node 5. Holding costs have an inverse effect on these three restoration coefficients. All three restoration coefficients drop from near 100% at low holding costs to near 90% at high holding costs. In spite of this, shortages from target along arcs 1-3 and 2-3 never account for more than 6% of the order quantities from nodes 1 and 2. We conclude that the small decline in values of restoration coefficients does not detract from the earlier conclusion that the best control strategy for the network under study is fundamentally very similar to base stock.

We are not sure of the reasons that holding costs effect restoration coefficients to node 5 yet do not seem to influence restoration coefficients to node 4 which also meets final demand. Perhaps it is because the coefficient of variation at node 5 (0.13) is higher than that at node 4 (0.10). Perhaps it is because the mean demand at node 5 (15) is greater than the mean demand at node 4 (10). Perhaps it is because the leadtime from node 3 to node 4 (2 weeks) exceeds the leadtime from node 3 to node 5 (1 week). It may even be a combination of any or all of these factors.

5.5 Coefficient of Variation as Factor (Lost Sales)

5.5.1 General

A single factor experiment using four levels of coefficient of variation of demand was done. The levels used were 0.01, 0.06, 0.12, and 0.18. Assuming normally distributed demands at nodes 4 and 5, these coefficients of variation can be expressed in terms of 95% confidence intervals as follows: 2,500 ± 36, 2,500 ± 216, 2,500 ± 432, and 2,500 ± 649 respectively. The simulation technique used parallels that described in Section 5.4. One hundred sixty-eight solves along a rolling horizon were done at each of the four levels of variation in demand. In each case, data from the first 28 solves were discarded and analysis of variance tests were done using 20 replicates of 7 solves each. Again, we report that no difficulties were encountered when solving the model using GAMS/CONOPT.

Table 5-9 shows grand means from 20 batches for target inventories, actual inventories, and restoration coefficients for the base network which has a coefficient of variation = 0.12. The table also shows shortages of actual inventory from target. Since inventory shortages along arcs 1-3 and 2-3 are small, the conclusion can again be drawn that the restoration coefficients from nodes 1 to 3 and 2 to 3 are largely immaterial. Production at nodes 1 and 2 is scheduled based primarily on consumption at nodes 4 and 5, the nodes that meet final demand. Again, the model is suggesting that a strategy very similar to base stock is best for this particular network. This result is invariant to the levels for the coefficient of variation factor used in this study.

TABLE 5-9
GRAND MEANS FOR INVENTORIES & RESTORATION COEFFICIENTS
BASED ON 20 BATCHES
COEFFICIENTS SINGLE FACTOR EXPERIMENT
COEFFICIENT OF VARIATION = 0.12

Average Arc	Target Inventory	Actual Inventory	Shortage (TI - INV)	Restoration Coefficient	Value
1-3	80.56	79.76	0.80	1-3	0.62
2-3	55.05	54.66	0.40	1-4	0.95
3-4	43.59	32.94	10.65	1-5	0.94
3-5	49.06	33.94	15.12	2-3	0.62
				2-4	0.97
				2-5	0.95
				3-4	0.95
				3-5	0.99

5.5.2 Test of Hypotheses

Table 5-10 shows the analysis of variance for testing the following hypothesis relevant to target inventory 1-3:

H_0: $V_1 = V_2 = V_3 = V_4 = 0$ Effect of Variation in Demand

H_1: at least one $V_i \neq 0$

Based on the results in the table, we reject the hypothesis that the level of variation of demand does not influence target inventories along arc 1-3. Analysis of variance for the remaining three target inventories, all four sets of actual inventories, and six of the eight restoration coefficients are provided in Appendix D. The results are summarized in Table 5-11.

TABLE 5-10
ANALYSIS OF VARIANCE FOR TARGET INVENTORY 1-3
SINGLE FACTOR EXPERIMENT USING LOST SALES MODEL
COEFFICIENT OF VARIATION IS FACTOR

Source	DF	Mean-Square	F-Ratio	Prob > F
Coefficient of Variation	3	170.72	65.41	0.0000*
Error	76	2.61		
Total Sum-Squares	79	710.53		

* Significant at the 1% level.

Table 5-11 shows that there is a significant, positive relationship between coefficient of variation and all target and actual inventories in the network. This is shown graphically in Figures 5-6 and 5-7 where target inventories

and actual inventories appear as functions of coefficient of variation. Notice that there is a linear relationship between target inventories and coefficient of variation as well as between actual inventories and coefficients of variation. Larger actual inventories equate to increasing amounts of safety stock. It can be concluded that an appropriate response to increasing variation of demand is to increase safety stock.

TABLE 5-11
SIGNIFICANT RELATIONSHIPS BASED ON THE ANOVA TESTS
SINGLE FACTOR EXPERIMENT
COEFFICIENT OF VARIATION = 0.12

Significant Effect[1]	Factor	Relationship[2]
TI 1-3	Coefficient of Variation	Positive
TI 2-3	Coefficient of Variation	Positive
TI 3-4	Coefficient of Variation	Positive
TI 3-5	Coefficient of Variation	Positive
AI 1-3	Coefficient of Variation	Positive
AI 2-3	Coefficient of Variation	Positive
AI 3-4	Coefficient of Variation	Positive
AI 3-5	Coefficient of Variation	Positive
RC 1-4	Coefficient of Variation	Negative
RC 1-5	Coefficient of Variation	Negative
RC 2-4	Coefficient of Variation	Negative
RC 2-5	Coefficient of Variation	Negative
RC 3-4	Coefficient of Variation	Negative
RC 3-5	Coefficient of Variation	Negative

[1] TI 1-3 = Target Inventory 1-3
AI 1-3 = Actual Inventory 1-3
RC 3-5 = Restoration Coefficient from Node 3 to Node 5

[2] A positive relationship implies that a high coefficient of variation of demand results in a high value for the parameter. A negative relationship implies that a high coefficient of variation of demand results in a low value for the parameter.

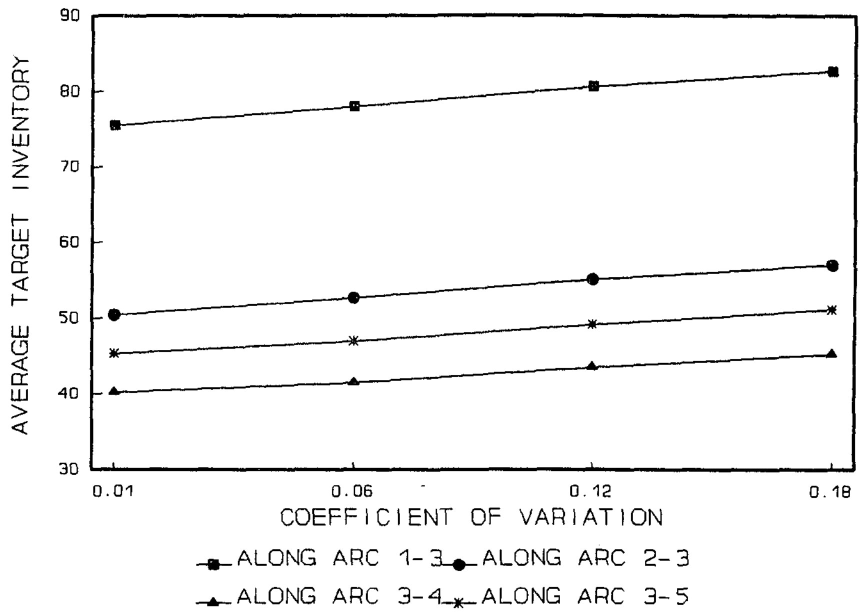


FIGURE 5-6
TARGET INVENTORIES VERSUS COEFFICIENT OF VARIATION
RESTORATION MODEL WITH LOST SALES

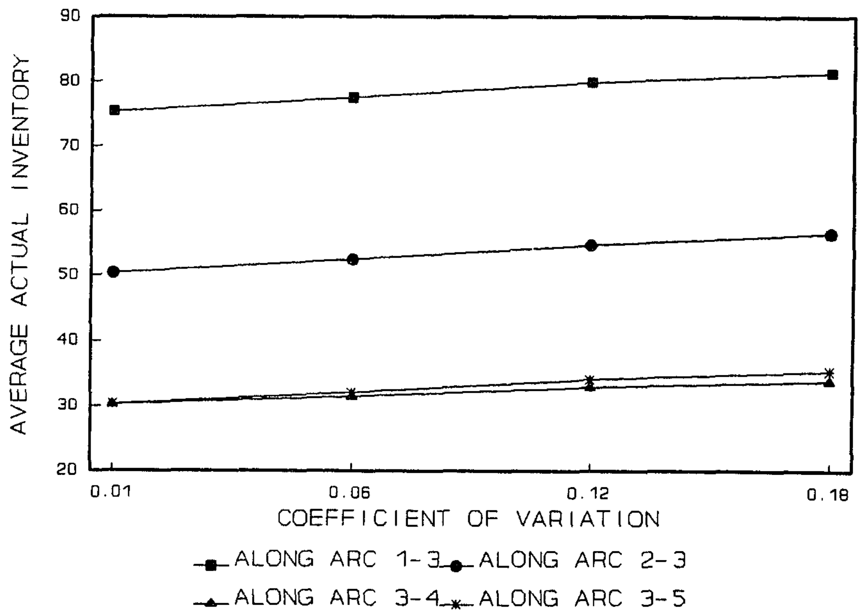


FIGURE 5-7
ACTUAL INVENTORIES VERSUS COEFFICIENT OF VARIATION
RESTORATION MODEL WITH LOST SALES

The idea of increasing safety stock to cope with increasing variation of demand is not new. The insight given by the model is in terms of where the safety stock should be placed. In practice, it is common to place extra safety stock at the final nodes, the nodes that meet final demand. This strategy was not chosen by the model. Rather, the restoration model chose to distribute safety stock more or less evenly throughout the network.

The restoration model reaffirms the use of a control strategy similar to base stock at all levels of variation of demand used in the study. Increasing variation of demand is best met by increasing target inventory throughout the network by more or less uniform amounts. In turn, this results in increased levels of actual inventories, or equivalently, increased levels of safety stock. The implications are that lean inventory systems operating under a pull control strategy cope best with increasing variation by increasing safety stocks uniformly throughout the network.

This result agrees with the common view that kanban systems do not handle variation in demands well. Rather, kanban practitioners (including Honda Motor Company) freeze the master production schedule for some period in time and attempt to level production. Results from the restoration model suggest that level production schedules are, in

fact, a requirement for operating a pull control system with minimal inventories.

Other, more subtle, effects are also demonstrated by the data in Table 5-11. The data shows that all restoration coefficients to nodes 4 and 5 decrease slightly with increasing variation in demand. In effect, increasing variation results in increased target inventories, actual inventories, and safety stock but decreased values for restoration coefficients to nodes that meet final demand. It would be interesting to observe this effect in a network, with more echelons.

Target inventories increase but restoration coefficients decrease with increasing variation of demands. The question is whether or not these two changes offset one another in regards to order quantities. Given that demand is stationary, these two quantities must offset each other. If not, inventories would either accumulate at final nodes (they do not) or inventories would fall to the point of causing infeasibilities, which also does not occur.

Increased variation in demand results in increased target inventories and actual inventories but decreased values for restoration coefficients. The relative slopes of the corresponding lines on Figures 5-6 and 5-7 show that target inventories increase more rapidly with increasing demand variability than do actual inventories. As a result, the average shortage of actual inventory from target grows

slightly with increasing demand variation. This relationship is shown graphically for arcs 3-4 and 3-5 in Figure 5-8. This nonlinear effect means that the cost of inventory deviating from target increases rapidly with increasing variation of demand. It is not clear what effects, if any, this would have on policy decisions.

5.6 Comparison of the Two Restoration Models

In retrospect, some similarities can be discerned for results from the restoration models *with* and *without* lost sales. Recall that total 24 period costs as well as target inventories 3-4 and 3-5 were occasionally abnormal when using the restoration model without lost sales. We decided to compare results from the two models using actual inventories. Table 5-12 shows average actual inventories and associated standard deviations based on 200 solves of the respective models for the factor levels: coefficients of variation 0.01 and coefficient of variation 0.12.

Notice that average values of actual inventories are almost identical for the two models at a coefficient of variation of 0.01. Notice also that average actual inventories at a coefficient of variation of 0.12 are slightly higher in the no lost sales model as compared to the lost sales model. Both models compensate for increasing demand variation with increased levels of actual inventory. The model without lost sales simply requires a higher level of actual inventories to cope with demand variability than

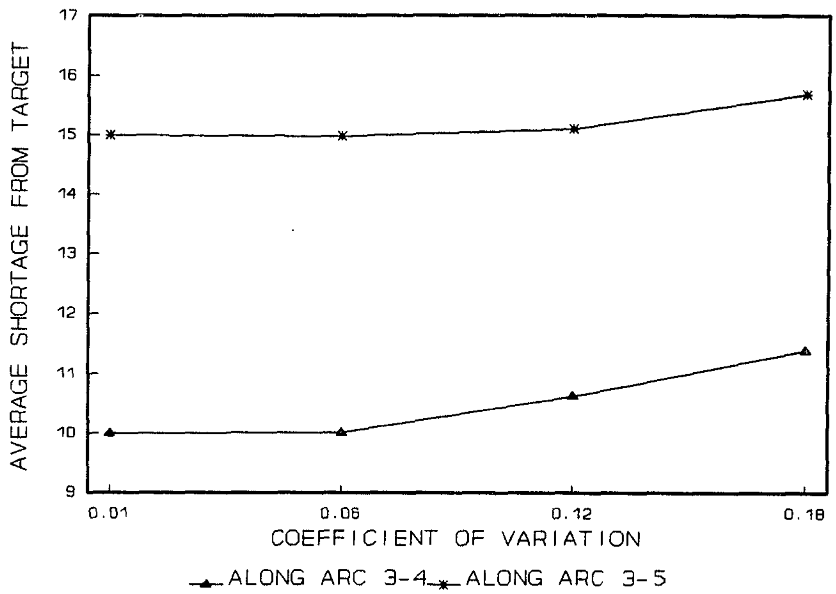


FIGURE 5-8
SHORTAGE OF INVENTORY FROM TARGET ALONG 3-4 AND 3-5
RESTORATION MODEL WITH LOST SALES

does the model with lost sales. Even so, the model without lost sales does not perform well as discussed previously.

TABLE 5-12
COMPARISON OF THE MODELS WITH AND WITHOUT LOST SALES
BASED ON AVERAGES FROM 200 SOLVES
LOW LABOR/HIGH HOLDING

	Actual Inventory With Lost Sales		Actual Inventory Without Lost Sales	
	Coef. of Variation		Coef. of Variation	
ARC	0.01	0.12	0.01	0.12
1-3	75.4	79.7	75.4	80.2
2-3	50.4	54.6	50.4	55.2
3-4	30.2	32.8	30.3	33.8
3-5	30.3	34.0	30.3	35.5

In summary, it appears as if the model without lost sales generally provides results quite similar to those from the model with lost sales. It is just that the model without lost sales occasionally results in abnormal target inventories, restoration coefficients, and total costs. Costs other than those related to the $K \cdot (TI_{jk} - INV_{jk}{}^{t})^2$ term in the objective function are very similar in both models. Neither model accumulates inventories at any point along the network. One model meets all demands whereas the other model typically meets over 99% of the demands. All in all, results from the two models appear similar.

5.7 Summary

An attempt was made to apply the restoration model without lost sales to the base network of Appendix A. Difficulties arose as it became apparent that this form of the model may occasionally: (1) be infeasible; or (2) result in abnormally large values for target inventories and costs. It was concluded that the restoration model without lost sales is too restrictive. The assumption that there always exist level values for target inventories and restoration coefficients that satisfy all constraints of this form of the model, including the meeting of all demands, is not valid. Accordingly, attention was turned to the restoration model with lost sales.

A 2^2 full factorial experiment was performed using labor costs and holding costs as factors. Neither labor costs nor the interaction of labor costs with holding costs were found to effect target inventories, actual inventories, or restoration coefficients. However, holding costs were found to have an effect on target inventories and actual inventories at nodes that do not meet final demands. Holding costs did not effect target inventories nor actual inventories at nodes that meet final demands. Apparently, target inventories and actual inventories at nodes that meet final demands are determined by factors other than labor costs and holding costs.

A single factor experiment was conducted with four levels of the factor coefficient of variation. Both target inventories and actual inventories were found to increase with increasing variation of demand. Increased actual inventories implies increased safety stock. Thus, increasing variation is countered by increasing safety stock. The restoration model suggests that the increased safety stock should not be concentrated at any particular node, rather the increased safety stock should be distributed more or less uniformly throughout the network.

The restoration model could have chosen the kanban type pull control strategy or any of several others as "best." It is interesting to note that the model implicitly chose a strategy very similar to pure base stock under every different factor level combination in which it was tested. Of course, this result is specific to the network of Appendix A which had no uncertainties in leadtimes or forecasts and no problems with quality. It would be interesting to extend the study by incorporating uncertainties in forecasts, defect levels, or leadtimes.

CHAPTER 6 SUMMARY AND CONCLUSION

6.1 Introduction

This research focused on materials management in a global manufacturing environment. Materials management as used here refers to the coordinated production and transportation of products across a network of plants and distribution outlets working towards the same finished product(s). In effect, materials management is an aggregate planning problem that encompasses issues related to production schedules and inventories in addition to control strategy.

An overview of the development and significance of the models is given in Section 6.2. Results based on the simulation studies and limitations of those results are summarized in Section 6.3 along with general guidelines for the use of the models. Finally, areas for future research are described in Section 6.4 and general conclusions are drawn in Section 6.5.

6.2 Overview

A review of the literature showed that few models exist for materials management in manufacturing networks of arbitrary design (e.g., serial, assembly, or conjoint). Those available assume perfect coordination between nodes by an externally defined control strategy such as MRP. We were unable to locate a model that can be used to specify the

best control strategy for networks of any design in addition to specifying aggregate values for inventories and production quantities. The fundamental goal of this research was to develop such a model.

A review of control strategies in common use showed that they are often described in terms of push or pull. At first glance, these two classifications appear dichotomous. However, further analysis revealed that it may be more appropriate to characterize components of an overall strategy using push or pull rather than using these words to describe the strategy as a whole. It turns out that many control strategies, including MRP and kanban, have both push and pull components. Control strategies in common use in addition to newly proposed ones found in the literature were reviewed and components of each were classified along a push versus pull gradient.

Production Authorization Cards, presented by Buzacott and Shanthikumar (1992), subsumes a wide range of control strategies. However, it does not subsume the restoration control strategy proposed by Tang (1990). The restoration strategy is based on the idea of restoring inventory at a node based on shortages and excesses of inventories from target values at down stream nodes. The restoration strategy subsumes a wide range of pull control strategies as well as certain aspects of push strategies. Further, the

restoration strategy is readily modified to apply to materials management in global manufacturing networks.

Production Authorization Cards was modified in this research so that it also subsumes the restoration strategy. The modified version resulted in the idea that many different control strategies can be distinguished based on values of only a few parameters such as target inventories and restoration coefficients. This concept was central to the development of several restoration based optimization models.

The restoration models developed in this research subsume a wide range of pull control strategies, including kanban and base stock. In addition, certain of the strategies subsumed by the restoration models exhibit push traits. Application of the restoration model to a specific network results in estimates for target inventories and restoration coefficients. These parameters in conjunction with actual inventories define the production control rule and implicitly identify the amount and location of safety stock. In summary, the models developed help define the best control strategy in addition to identifying aggregate values relating to inventory and production schedules.

6.3 Results, Limitations, and Guidelines for Use

A strategy very close to base stock was "best" for each of the simulation runs made. This result is specific to a five node, conjoint network with no forecast errors, no

uncertainties, and external demands only at the final nodes. Further research is needed to demonstrate whether or not a similar strategy is also best for networks that differ from the one studied.

Labor costs did not significantly influence values of target inventories, actual inventories, or restoration coefficients. It appears that values for these policy parameters are relatively invariant to misspecifications of or shifts in labor costs, at least within the range of labor costs used in this study. Holding costs did appear to influence target inventories at nodes that did not meet final demands. Higher holding costs resulted in lower target inventories at these nodes and vice versa. It is not yet clear how holding costs would influence target inventories in larger networks, those with more than three echelons of nodes. However, it is clear that it would behoove managers to accurately assess holding costs and their relationship to target inventories. There did not appear to be an interaction effect between labor costs and holding costs.

Not surprisingly, results from the simulation experiment suggest that an increase in demand variability is best countered with increased target inventories and actual inventories/safety stocks. Interestingly, our results suggest that an increase in demand variation is best met by increasing safety stock throughout all nodes in the network

more or less uniformly. The idea of handling increasing demand variation with additional safety stock at nodes that meet final demand was rejected as inferior. It would be interesting to further contrast these two strategies involving safety stock.

We postulate that increased demand variation in a minimal inventory, pull control system cannot be met solely by increasing safety stock at nodes that meet final demand. Increased safety stock at nodes that meet final demands simply allows the variation to move through those nodes to the next level of the network. Increased variation in component demand at this level requires increased safety stock and so on throughout the network. Therefore, increased demand variation under a minimal inventory, pull control strategy permeates the entire network-increased safety stock is required at every node.

It is clear that under a minimal inventory, pull control strategy increased demand variation results in increased holding costs throughout the network. There is a tradeoff between the costs of minimizing variation of demands and the cost of carrying extra safety stock. Management may choose to try to meet demands or to level demands using marketing incentives (for example).

This relationship may have contributed to the success of the Japanese automakers. Demand for Japanese imports has generally exceeded supply over the past 25 years. This

allowed Japanese automakers the luxury of leveling production schedules within planning horizons. Level production requirements under a pull control strategy allows for minimal inventories and minimal holding costs. It will be interesting to observe if the new global realities of more competition and higher demand variation force any of the Japanese firms to change their strategies.

6.4 Future Research

This research has enabled us to identify a number of areas worthy of further investigation. For purposes of discussion, future research is somewhat arbitrarily divided into factors internal to the manufacturing network and those external to the network.

6.4.1 Internal Factors

The simulation experiments focused on a specific five node, conjoint network. In particular, the conclusion that holding costs influence target inventories at nodes that do not meet final demand may be dependent on the specific network studied. Networks containing more than three echelons of nodes must be explored before results can be generalized.

It was also concluded that a strategy close to the base stock strategy was best for the specific network studied. It would be interesting to observe any changes to the control strategy based upon the introduction of 1) defects throughout the network and/or 2) final demands at

intermediate nodes. In both cases, we would anticipate that shortages from target inventory at intermediate nodes would become a more important factor in determining order quantities at the earliest nodes-but this remains to be demonstrated.

It is desirable to apply the model to networks substantially larger than the five node case studied in Chapter 5. A difficulty arises in that the number of decision variables rises rapidly with increasing numbers of nodes, particularly the number of restoration coefficients in a highly connected network. As it stands, it is probably not possible to solve the model even for moderate sized networks due to the nonlinearities. In applying the model to the five node network shown in Figure 5-1, we note that the target inventory and restoration coefficients along arc 1-3 behaved very similarly to those along arc 2-3 to changing costs and demand variations. If this relationship can be shown to hold in general along converging arcs, it may help us solve larger problems. Of course, the relationship must first be firmly established.

It would also be interesting to extend the assumption of single-item to multi-item. By doing so, we would anticipate that setup costs would become more important and should be reintroduced into the model. In this event, the model would again become mixed-integer and nonlinear. The difficulties we encountered in solving the model were in

regards to the without lost sales form of the model. It should be possible to solve the mixed-integer, nonlinear model with lost sales for small networks.

6.4.2 External Factors

A central, limiting assumption of the specific network used in this study was that demands were stationary. In practice this assumption almost never occurs since demands are usually cyclic or exhibit a trend. It would be of interest to explore the effectiveness of the restoration control strategy under nonstationary demands. Given that target inventories are highly sensitive to variation in demand, we project that trends will be difficult to handle in a minimal inventory, pull control system-but this remains to be demonstrated.

The only control strategies explored in the simulation experiments of Chapter 5 are those subsumed by the restoration model (e.g., kanban, base stock, CONWIP, etc.). No comparison has been made with a control strategy that is not subsumed by the restoration model, such as MRP. It would be interesting to compare the restoration model to MRP for a network containing uncertainties such as forecast errors. Indeed, it would be quite interesting to compare the restoration strategy to the commonly used strategies of kanban and MRP in a realistic setting. An appropriate methodology for a study of this type would be Monte Carlo simulation.

6.5 Conclusion

This research resulted in a general model for materials management in a global manufacturing network that subsumes a wide range of pull control strategies. The model can be used to determine aggregate values of target inventories and implicitly, actual inventories and safety stock. Results from the model appear to be stable and are consistent with observations. The model can be used to identify the "best" control strategy for a specific network along with aggregate values of target inventories and order quantities. In turn, these values can be disaggregated using concepts from hierarchical planning.

BIBLIOGRAPHY

Afentakis, P., and B. Gavish, "Optimal Lot-Sizing Algorithms for Complex Product Structures," Operations Research, 34 (1986), 237-249.

Anderson, E.J., "A Note on the Dynamic Lot-Size Model with Uncertainty in Demand and Supply Processes," Management Science, 35 (1989), 635-640.

Anderson, J.C., R.G. Schroeder, S.E. Tupe, and E.M. White, "Material Requirements Planning systems: The state of the Art," Production and Inventory Management, 23 (1982), 51-66.

Axsater, S., "Control Theory Concepts in Production and Inventory Control," Int. Journal of Systems Science, 16 (1982), 161-169.

Baer, T., "How to Schedule in Unexpected Events," Computerworld, February 11, 1991, 61.

Battaglia, A.J., and G.R. Tyndall, "Working on the Supply Chain," Chief Executive, 66 (1991), 42-45.

Baumol, W.J., and H.D. Vinod, "An Inventory theoretic Model of Freight Transport Demand," Management Science, 16 (1970), 413-421.

Beale, E.M.L., J.J.H. Forrest, and C.J. Taylor, "Multi-Time-Period Stochastic Programming," in M.A.H. Dempster (ed.), Stochastic Programming, New York: Academic Press, 387-402, 1980.

Bergstrom, G.L., and B.E. Smith, "Multi-item Production Planning-An Extension of the HMMS Rules," Management Science, 16 (1970), 614-29.

Bessler, S.A., and A.F. Veinoit, "Optimal Policy for a Dynamic Multi-Echelon Inventory Model," Naval Res. Logistics Qtr., 13 (1966), 355-389.

Billington, P.J., J.O. McClain, and L.J. Thomas, "Mathematical Programming Approaches to Capacity-Constrained MRP Systems: Review, Formulation and Problem Reduction," Management Science, 29 (1983), 1126-1141.

Bitran, G.R., and L. Chang, "A Mathematical Programming Approach to a Deterministic Kanban System," Management Science, 33 (1987), 427-441.

——— and A. Hax, "Disaggregation and Resource Allocation Using Convex Knapsack Problems with Bounded Variables," Management Science, 27 (1981), 431-441.

Burns, L.D., R.W. Hall, D.E. Blumenfeld, and C.F. Daganzo, "Distribution Strategies that Minimize Transportation and Inventory Costs," Operations Research, 33 (1985), 469-490.

Buzacott, J.A., and J.G. Shanthikumar, "A General Approach for Coordinating Production in Multiple-Cell Manufacturing Systems," Production and Operations Management, 1 (1992), 34-52.

Carlson, R., and C.A. Yano, "Buffering Against Demand Uncertainty in Material Requirements Planning Systems-Systems with no Emergency Setups," Technical Report #81-2, Dept. of I.E. and Eng. Mgmt., Stanford University, 29, 1981.

Chung, C., and L.J. Krajewski, "Planning Horizons for Master Production Scheduling," Journal of Operations Management, August (1984), 389-406.

Clark, A.J., "An Informal 'Survey of Multi-Echelon Inventory Theory," Naval Research Log. Qtr., 19 (1972), 621-650.

——— and H.E. Scarf, "Optimal Policies for a Multi-echelon Inventory Problem," Management Science, 6 (1960), 475-490.

——— and H. Scarf, "Approximate Solutions to a Simple Multi-Echelon Inventory Problem," in Studies in Applied Probability and Management Science, 88-110, K.J. arrow, S. Karlin and H. Scarf (eds), Stanford, Stanford University Press, 1962.

Cohen, M.A., and H.L. Lee, "Strategic Analysis of Integrated Production-Distribution Systems: Models and Methods," Operations Research, 36 (1988), 216-228.

Constable, G.K. and D.C. Whybark, "The Interaction of Transportation and Inventory Decisions," Decision Sciences, 9 (1978), 688-699.

Cox, J.F., and S.J. Clark, "Problems in Implementing and Operating a Manufacturing Resource Planning Information System," J. of Management Information Systems, 1 (1984), 81-101.

Crosby, P.B., Quality is Free: The Art of Making Quality Certain, New York: McGraw-Hill, 1979.

Damon, W.W., and R. Schramm, "A Simultaneous Decision Model for Production, Marketing, and Finance," Management Science, 19 (1972), 161-172.

DeBodt, M.A., L.N. Van Wassenhove, and L.F. Gelders, "Lot Sizing and Safety Stock Decisions in an MRP System with Demand Uncertainty," Eng. Costs and Prod. Econ, 6 (1982), 67-75.

Deleersnyder, J., T.J. Hodgson, H. Muller, and P.J. O'Grady, "Kanban Controlled Pull Systems: An Analytic Approach," Management Science, 35 (1989), 1079-1091.

Deuermeyer, B.L. and Schwarz, L.B., "A Model for the Analysis of System Service Level in a Warehouse-Retailer Distribution System," In Multi-Level Production/Inventory Control System: Theory and Practice, L.B. Schwarz, Ed., Vol 16, 1981, Studies in the Management Sciences, North-Holland, Amsterdam, 163-194.

DeRoulet, D.G. "Define a Global Logistics Strategy," Transportation & Distribution, 32 (1991), 38-40.

Dertouzos, M.L., R.K. Lester, and R.M. Solow, Made in America, New York: Harper Perennial, 1989.

Drud, A., "A GRG Code for Large Sparse Dynamic Nonlinear Optimization Problems," Mathematical Programming, 31 (1985), 153-191.

———, "A Large-Scale GRG Code," ORSA Journal on Computing (1993), forthcoming.

Eppen, G., and L. Schrage, "Centralized Ordering Policies in a Multi-Warehouse System with Lead Times and Random Demand," in L. Schwarz (ed.), Multi-Level Production Inventory Control Systems: Theory and Practice, New York: North Holland, 51-67, 1981.

Gaither, N., "A Near Optimal Lot-Sizing Model for Material Requirements Planning Systems," Prod and Inv Mgmt., fourth Qtr (1981), 74-89.

———, "An Improved Lot-Sizing Model for MRP Systems," Prod and Inv Mgmt., (1983), 10-19.

Gavish, B., and R.E. Johnson, "A Fully Polynomial Approximation Scheme for Single Product Scheduling in a Finite Capacity Facility," Operations Research, 38 (1990), 70-83.

Goldratt, E.M., and R.E. Fox, The Race, New York, North River Press, Croton-On-Hudson, 1986.

Gong, L., and H. Matsuo, "Stabilizing Work-In-Process and Smoothing Production in a Production System with Random Yield," Working Paper, U. Texas, 1991.

Goyal, S.K., and A. Gunasekaran, "Multi-Stage Production-Inventory Systems," European J. of Operational Research, 46 (1990), 1-20.

Grasso, E.T., and B.W. Taylor, III, "A Simulation-Based Experimental Investigation of Supply/Timing Uncertainty in MRP Systems, Int. J. Production Research, 22 (1984), 485-497.

Groff, G.K., "A Lot Sizing Rule for Time-Phased Component Demand," Prod and Inv. Mgmt., First Qtr (1979), 46-53.

Hackman, S.T., and R.C. Leachman, "A General Framework for Modeling Production," Management Science, 35 (1989), 478-495.

Haddock, J., and D.E. Hubicki, "Which Lot-Sizing Techniques are Used in Materials Requirements Planning?, "Prod. and Inv. Mgmt. Journal, Third Qtr (1989), 53-56.

Hall, R., Zero Inventories, Homewood Ill, Dow Jones-Irwin, 1983.

Hansmann, F., and S.W. Hess, "A Linear Programming Approach to Production and Employment Scheduling," Management Technology, 1 (1960), 46-51.

Hax, A.C., and H.C. Meal, "Hierarchical Integration of Production Planning and Scheduling," In Logistics (M.A. Geesler, Ed), Studies in Management Science, Vol. 1, New York, North-Holland, 53-69, 1975.

Hay, E.J., The Just-In-time Breakthrough, Implementing the New Manufacturing Basics, New York: John Wiley & Sons, 1988.

Heinrich, C.E., and C. Schneeweiss, "Multi-Stage Lot-Sizing for General Production Systems," in S. Axsater, C. Schneeweiss, and E. Silver (eds.), Multi-Stage Production Planning and Inventory, Berlin: Springer-Verlag, 150-181, 1986.

Hodgson, T.J., and D. Wang, "Optimal Hybrid Push/Pull Control Strategies for a Parallel Multistage System: Part I," Int. J. Production Research, 29 (1991a), 1279-1287.

———, and D. Wang, "Optimal Hybrid Push/Pull Control Strategies for a Parallel Multistage System: Part II," Int. J. Production Research, 29 (1991b), 1453-1460.

Holt, C.C., F. Modigliani, and H.A. Simon, "A Linear Decision Rule for Employment and Production Scheduling," Management Science, 2 (1955), 1-30.

———, F. Modigliani, and J.F. Muth, "Derivation of a Linear Decision Rule for Production and Employment," Management Science, 2 (1956), 159-177.

———, F. Modigliani, J.F. Muth, and H.A. Simon, Planning Production Inventories, and Workforce, Englewood Cliffs, N.J., Prentice-Hall, 1960.

Huang, P.Y., L.P. Rees, and B.W. Taylor III, "A Simulation Analysis of Japanese Just-in-Time Technique (With Kanban) for a Multiline, Multistage Production System," Decision Sciences, 14 (1983), 326-343.

Jordan, W.C., and S.C. Graves, "Principles on the Benefits of Manufacturing Process Flexibility," Working Paper #3296-91-MSA, Alfred P. Sloan School of Management, M.I.T., May, 1991.

Kackar, R.N., "Off-Line Quality Control, Parameter Design, and the Taguchi Method," Journal of Quality Technology, 17 (1985), 176-188.

Kanet, J.J., "MRP96: Time to Rethink Manufacturing Logistics," Prod. and Inventory Management, 29 (2nd Qtr, 1990), 57-61.

Kaplan, R.S., "A Dynamic Inventory Model with Stochastic Leadtimes," Management Science, 16 (1970), 491-507.

Karmarkar, U., "Push, Pull and Hybrid Control Schemes," Working Paper No. QM8614, University of Rochester, Rochester, NY, 1986.

———, "Getting Control of Just-In-time," Harvard Business Review, Sept-Oct (1989), 122-131.

Kim, T., "Just-In-Time Manufacturing System: A Periodic Pull System," Int. J. of Production Research, 23 (1985), 553-562.

Kimura, O., and H. Terada, "Design and Analysis of Pull System, A Method of Multi-Stage Production Control," Int. J. of Production Research, 19 (1981), 241-253.

Krajewski, L.J., B.E. King, L.P. Ritzman, and D.S. Wong, "Kanban, MRP, and Shaping the Manufacturing Environment," Management Science, 33 (1987), 39-57.

———, L.P. Ritzman, "Disaggregation in Manufacturing and Service Organizations: Survey of Problems and Research," Decision Sciences, 8 (1977), 1-18.

Lambert, D.M. and R.L. Cook, "Integrating Marketing and Logistics for Increased Profit," Business, Jul-Sep (1990), 22-29.

Lambrecht, M.R., J.V. Ecken, and H. Vanderveken, "Review of Optimal and Heuristic Methods for a Class of Facilities in Series Dynamic Lot-Size Problems," In Multi-Level Production/Inventory Control Systems: Theory and Practice, L.B. Schwarz, Ed., Vol 16, 1981, Studies in the Management Sciences, North-Holland, Amsterdam, 69-94.

———, J.A. Muckstadt, and R. Luyten, "Protective Stocks in Multi-Stage Production Systems," Int. J. Production Research, 22 (1984), 1001-1025.

Law, A.M., and W.D. Kelton, Simulation Modeling and Analysis, New York, McGraw-Hill Book Company, 1982.

Leitch, R.A., "Marketing Strategy and Optimal Production Schedule," Management Science, 20 (1974), 903-911.

Lieb, R.C., "Logistics as a Strategic Variable in Post-1992 Europe," Transportation Journal, Summer (1991), 56-61.

Manoochehri, G.H., "Suppliers and the Just-In-Time Concept," J. of Purchasing and Materials Management, Winter (1984), 16-21.

McLaren, B.J., "A Study of Multiple Level Lot Sizing Techniques for Materials Requirements Planning Systems," Unpublished Ph.D. Dissertation, Purdue Univ., 1976.

McClain, J.O., W.L. Maxwell, J.A. Muckstadt, L.J. Thomas, and E.N. Weiss, "On MRP Lot Sizing," Management Science, 28 (1982), 582-584.

Moeeni, F., and Y. Chang, "An Approximate Solution to Deterministic Kanban Systems," Decision Sciences, 21 (1990), 596-607.

Monden, Y., Toyota Production System, Practical Approach to Management, Norcross, Ga., I.E. and Management Press, 1983.

Montgomery, D.C., Design and Analysis of Experiments, 3rd Edition, New York, John Wiley & Sons, 1991.

Nahmias, S., "Managing Reparable Item Inventory Systems: A Review," In Multi-Level Production Inventory Control Systems: Theory and Practice, L.B. Schwarz, Ed., Vol 16, 1981, Studies in the Management Sciences, North-Holland, Amsterdam, 253-277.

Nevison, C.H., "A Cost Adjustment Heuristic for Dynamic Lot-Sizing with Uncertain Demand Timing," Operations Research, 33 (1985), 1342-1353.

———, and M. Burstein, "The Dynamic Lot-Size Model with Stochastic Lead Times," Management Science, 30 (1984), 100-109.

New, C., Requirements Planning, Essex, England: Gower Press, 1974.

O'Grady, P.J. and M.C. Bonney, "The Optimum Behavior of a Multi-Echelon Production and Inventory System," Engineering Costs and Production Economics, 9 (1985), 23-32.

Orlicky, J., Materials Requirements Planning, New York: McGraw-Hill, 1975.

Parker, M., and J. Slaughter, "Management by Stress," Technology Review, October (1988), MIT, 37-44.

Perry, J.H., "Emerging Economic and Technological Futures: Implications for Design and Management of Logistics Systems in the 1990s," J. of Business Logistics, 12 (1991), 1-16.

Philipoom, P.R., L.P. Rees, B.W. Taylor, and P.Y. Huang, "An Investigation of the Factors Influencing the Number of Kanbans Required in the Implementation of the JIT technique with Kanbans," International Journal of Production Research, 25 (1987), 457-472.

Pochet, Y., and L.A. Wolsey, "Solving Multi-Item Lot-Sizing Problems Using Strong Cutting Planes," Management Science, 37 (1991), 53-67.

Popplewell, K., and M.C. Bonney, "The Application of Discrete Linear Control Theory to the Analysis and Simulation of Multi-Product, Multi-Level Production Control Systems," Int. Journal of Production Research, 25 (1987), 45-56.

Porter, M.E., The Competitive Advantage of Nations, New York, Free Press, 1990.

Ptak, C.A., ""Mrp, MRP II, OPT, JIT, and CIM-Succession, Evolution, or Necessary Combination," Production and Inventory Management Journal, 2nd Qtr. (1991), 7-11.

Pyke, D.F., and M.A. Cohen, "Push and Pull in Manufacturing and Distribution Systems," J. of Operations Management, 9 (1990), 24-42.

Rees, L.P., P.Y. Huang, and B.W. Taylor III, "A Comparative Analysis of an MRP Lot-for-Lot System and a Kanban System for a Multistage Production Operation," Int. J. Production Research 27 (1989), 1427-1443.

Rogers, D.F., and S. Tsubakitani, "Inventory Positioning/Partitioning for Backorders Optimization for a Class of Multi-Echelon Inventory Problems," Decision Sciences, 22 (1991), 536-558.

Rosling, K., "Optimal Inventory Policies for Assembly Systems under Random Demands," Operations Research, 32 (1989), 565-579.

Sarker, B.R., and R.D. Harris, "The Effect of Imbalance in a Just-In-time Production System: A Simulation Study," Int. J. Production Research, 26 (1988), 1-18.

Schittkowsky, J.K., <u>Nonlinear Programming Codes, Lecture Notes in Economics and Mathematical Systems</u>, Vol 183, Heidelberg: Springer Verlag, 1980.

Schmidt, C.P., and S. Nahmias, "Optimal Policy for a Two-Stage Assembly System under Random Demand," Operations Research, 33 (1985), 1130-1145.

Schmitt, T.G., "Resolving Uncertainty in Manufacturing Systems, "J. of Operations Mgmt., 4 (1984), 331-345.

Schonberger, R., <u>Japanese Manufacturing Techniques</u>, New York, Free Press, 1982.

Schroeder, R.G., J.C. Anderson, S.E. Tudy, and E.M. White, "A Study of MRP Benefits and Costs," J. of Operations Management, 2 (1981), 1-9.

Schwarz, L.B., "A Simple Continuous review Deterministic One-Warehouse n-Retailer Inventory Problem," Management Science, 19 (1973), 555-566.

———, and L. Schrage, "Optimal and System Myopic Policies for Multi-Echelon Production/Inventory Assembly Systems," Management Science, 21 (1975), 1285-1294.

Silver, E.A., and R. Peterson, Decision Systems for Inventory Management and Production Planning, 2nd Ed., New York: John Wiley & Son, 1985.

So, K.C., and S.C. Pinault, "Allocating Buffer Storage in a Pull System," International Journal of Production Research, 26 (1988), 1959-1980.

Spearman, M.L., D.L. Woodruff, and W.J. Hopp, "CONWIP: A Pull alternative to Kanban," Int. J. of Production Research, 28 (1990), 879-894.

Stalk, G., and T.M. Hout, Competing Against Time, New York: Free Press, 1990.

Steinberg, E., and H.A. Napier, "Optimal Multi-Level Lot Sizing for Requirements Planning Systems," Management Science, 26 (1980), 1258-1271.

Svoronous, A. and P. Zipkin, "Estimating the Performance of Multi-Level Inventory Systems," Operations Research, 25 (1988), 57-72.

Takahashi, K., R. Muramatsu, and K. Ishii, "Feedback Method of Production Ordering System in Multi-Stage Production and Inventory Systems," Int. J. Production Research, 25 (1987), 925-941.

Tang, C.S., "The Impact of Uncertainty on a Production Line," Management Science, 36 (1990), 1518-1531.

Towill, D., "Dynamic Analysis of an Inventory and Order-Based Production Control System," Int. Journal of Production Research, 20 (1982).

———, "The Practical Use of Control Theory in Manufacturing Systems Design and Manufacturing Management," Journal of Systems Engineering, 2 (1992), 2-14.

Umble, M.M., and M.L. Srikanth, Synchronous Manufacturing: Principles for World Class Excellence, Dallas, Tx., South-Western Publishing Co., 1990.

Veatch, M., and L. Wein, "Optimal Control of a Multistage Production/Inventory System," Paper Presented at 1991 Conference on Multi-Echelon Inventory Systems Held at Berkeley, Ca.

Vollman, T.E., W.L. Berry, and D.C. Whybark, Manufacturing Planning and Control Systems, 2nd Ed., Homewood, Ill., Irwin, 1988.

Voss, C., and S. Robinson, "The Application of Just-In-time Techniques," Int. J. of Operations and Production Management, 7 (1987), 46-52.

Whybark, D.C., and J.G. Williams, "Material Requirements Planning Under Uncertainty," Decision Sciences, 7 (1976), 595-606.

Wight, O., Production and Inventory Management in the Computer Age, Boston, Cahners Books, 1974.

Willersdorf, R.G. "Adding Value through Logistics Management," Int. Journal of Physical Distribution & Logistics Management, 21 (1991), 6-8.

Woolsey, G., "Ten Ways to Go Down with MRP," Interfaces, 5 (1979), 77-80.

Womack, J.P., P.T. Jones, and D. Roos, The Machine that Changed the World, New York, Harper Collins, 1990.

APPENDIX A

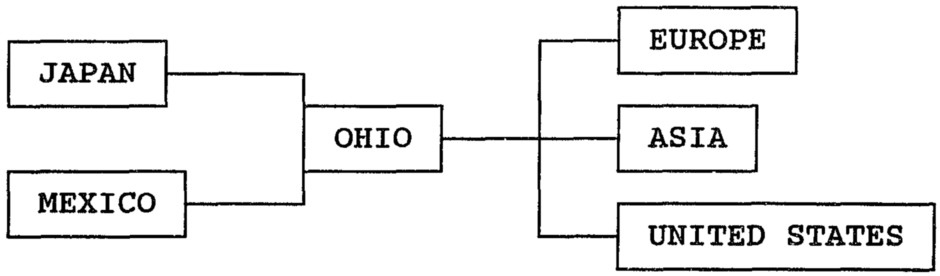
COST DERIVATION FOR A SPECIFIC NETWORK

The models of Chapter 3 apply to a broad range of manufacturing networks. However, the specific network used for the numerical studies of Chapter 5 is patterned after the organization that Honda Motor Company, LTD. uses to manufacture automobiles. Certain figures requested from Honda were considered proprietary and were not released to us. Accordingly, the specific network and costs that we use are only loosely patterned after the Honda case.

Honda builds various parts for automobile engines and transmissions in Japan. These products are moved by ship and by truck to the Ohio region. Wiring harnesses and various other components are constructed in Mexico and moved by rail or by truck to Ohio. Numerous suppliers in the Ohio and Canada region also supply parts to the Marysville, Ohio plant. Honda claims that 75% of the finished product is made up of local content, where local content refers to products made in the Ohio/Canada region.

Final assembly is done in Marysville, Ohio after which the finished automobiles are shipped to distribution points in the United States. Automobiles are also transported to Fort Lauderdale, Florida for shipment to Europe and to Portland, Oregon for shipment to Asia. Honda's average sales per model were about 125,000 per year in 1991, or approximately 2,500 per week.

The global manufacturing network used by Honda follows:



It is important to note that the Ohio node represents a number of suppliers in the Ohio/Canada area in addition to the final assembly plant in Ohio.

The asking price for a new Honda Accord is currently around $18,000. We assume their total cost for this vehicle at the point of final sale to be $15,000. Further, we assume their cost just prior to shipment from Ohio to be $13,000. A 75% local content translates to a value added of $9,750 in the Ohio area, leaving a value of $3,250 for products produced in Japan and Mexico. It is assumed that about 70% of the $3,250, or $2,250 in value, is from parts made in Japan and that the remaining $1,000 in value is from parts made in Mexico.

Holding costs including the costs of capital, obsolescence, and storage and are assumed to be 30% per year. It is of interest to note that the aversion JIT practitioners have for inventory, results implicitly in a value of significantly more than 30% on holding costs. Certainly the theme of lean production requires that a high cost be placed on inventories.

Each time interval in the network studied is assumed to be one week, so that:

$$\text{Holding Costs for One Period} = \text{Value} \cdot \frac{30\%}{52}$$

Shipments within North America, from Mexico to Ohio and from Ohio to any destination in the United States, are assumed to require one week. Shipments between continents, from Japan to Ohio and from Ohio to Asia or Europe, are assumed to require twice as long: two weeks.

Labor costs can be separated into overhead related costs and direct labor costs. We assume that overhead costs are fixed so that only direct labor costs are relevant to our model. Further, we also assume that regular time, direct labor costs are not relevant costs. Therefore, only overtime/undertime costs associated with direct labor are relevant to the model. Based on a report put out by the Economic Strategy Institute (1992), direct labor costs are assumed to be 12.5% of the value added at each node. Assuming overtime costs are one and one-half times regular time costs, the relevant cost of overtime equates to 50% of the regular labor cost. Overtime costs at a node are calculated as follows:

$$\text{Overtime Costs} = \overbrace{\text{Value Added} \cdot 12.5\%}^{\text{Regular time costs}} \cdot 50\%$$

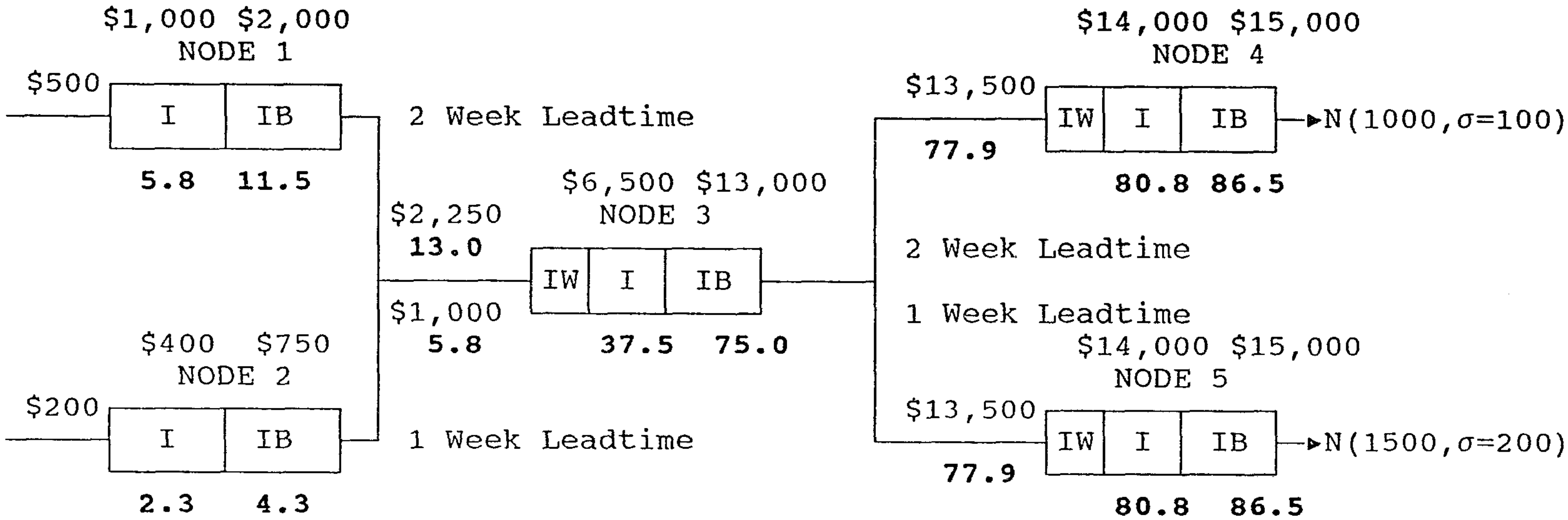
Significant and prolonged undertime would, of course, be quite costly to a firm. By assumption, we choose to

model the case in which employment is relatively well balanced with production requirements. Note that plants running with "lean" numbers of workers will not tend to have much undertime. In that event, the small amounts of undertime that occur are used for maintenance or other activities of value to the firm. Accordingly, we assume that the relevant cost associated with undertime is small, equal to 20% of regular time costs. Therefore, undertime costs at a node are calculated as follows:

$$\text{Undertime Costs} = \overbrace{\text{Value Added} \bullet 12.5\%}^{\text{Regular time costs}} \bullet 20\%$$

The specific network to be studied is the five node conjoint network shown in the following figure. In drawing the parallel with the Honda manufacturing example, nodes 1 and 2 represent Japan and Mexico respectively. Node 3 represents subcontractors and final assembly located in the Ohio/Canada region. Node 4 represents distribution to Europe and Asia combined and node 5 represents distribution to the United States. Please note that the network is loosely patterned after the Honda network and is not intended to be an exact replica.

It is desired that final demands at nodes 4 and 5 be stationary with a combined mean of 2,500 automobiles per week. This mean corresponds roughly to Honda's average volume per model. Specific sequences of demands for nodes 4 and 5 will be derived from a Normal (1000,σ=100) and a



<u>Capacities/Limits</u>:

		NODES 1	2	3	4	5	
Regular Capacity	=	2,500	2,500	2,500	1,000	1,500	Number of items
Limit on Overtime	=	1,000	1,000	1,000	1,000	1,000	

<u>On a Per Item Basis</u>:

		1	2	3	4	5	
Regular Labor Costs	=	$188	$70	$1,200	$188	$188	= 12.5% of value added
Overtime Costs	=	$ 94	$35	$ 600	$ 94	$ 94	= 50% of Regular Time
Undertime Costs	=	$ 38	$14	$ 240	$ 38	$ 38	= 20% of Regular Time

* Figures in **bold** are weekly holding costs.

FIGURE A-1
SPECIFIC NETWORK AND COSTS USED FOR EXAMPLE IN CHAPTER 5
LOOSELY PATTERNED AFTER THE HONDA MOTOR COMPANY NETWORK

Normal (1500,σ=200) distribution, respectively. The effect of this strategy is a combined final demand that is distributed according to a Normal (2,500,σ=224) distribution. By assumption, final demand excludes both trends in sales and any cycles in sales. In addition, total final demand will be within 448 (2 standard deviations) of the mean of 2,500, 95% of the time.

In the table, per item costs are listed above each node and holding costs per week are listed below each node. Leadtimes required to complete shipments between nodes are noted between the nodes in the figure. All production leadtimes are assumed to be one. Amounts of regular time capacity for each node and limits on the amount of overtime possible at each node are listed at the bottom of the figure. The network is assumed to be balanced with respect to production capacity versus demand. In other words, production capacity is defined to equal the expected demands from the distributions from which demands are derived.

The cost per item for regular time, overtime, and undertime is also shown at the bottom of the figure. Again, regular time costs are not considered to be costs relevant to the model. Overtime costs that exceed regular time costs are relevant as are undertime costs that result in no productivity to the firm. Overtime costs per unit are defined to be 50% of regular time costs, and undertime costs per unit are defined to be 20% of regular time costs.

APPENDIX B

WARMUP, BATCH LENGTH, AND NUMBER OF BATCHES

Further details related to the selection of the warmup period, batch length, and number of batches are included in this appendix. Figures B-1 and B-2 and Table B-1 were derived by applying the restoration model without lost sales to the low labor/high holding cost network of Appendix A. Figures B-3 and B-4 and Table B-2 were derived using the restoration model with lost sales.

The average total 24 period cost per period based on the restoration model without lost sales is shown in Figure B-1. As you can see, average costs stabilize quickly. A warmup of 28 periods satisfies the criteria listed in Chapter 4. The data in Figure B-2 suggest that a batch length of at least 13 solves of the model is required to justify the assumption of uncorrelated batch means. Table B-1 shows the β-levels associated with an α-level of 5% and 10, 20, and 30 replicates. As you can see 30 replicates does not result in satisfactory β-levels for three of the parameters: target inventory 3-5, restoration coefficient 1-3, and restoration coefficient 2-3. Since the computational requirements associated with over 30 batches of 13 solves each are prohibitive, attention was turned to the restoration model with lost sales.

Figure B-3 shows the average total 24 period cost per period using the restoration model with lost sales. A

warmup of 28 periods remains sufficient. Figure B-4 shows that a batch length of 7 solves results in autocorrelations below 40% and is therefore satisfactory for our purposes. Table B-2 shows β-levels associated with an α-level of 1% and 10, 20, and 30 replicates. Twenty replicates result in satisfactory β-levels ($\leq$2%) for all parameters other than restoration coefficients 1-3 and 2-3. In Chapter 5, these two restoration coefficients are shown to be largely immaterial for other reasons. Accordingly, 20 replicates were done at each treatment level.

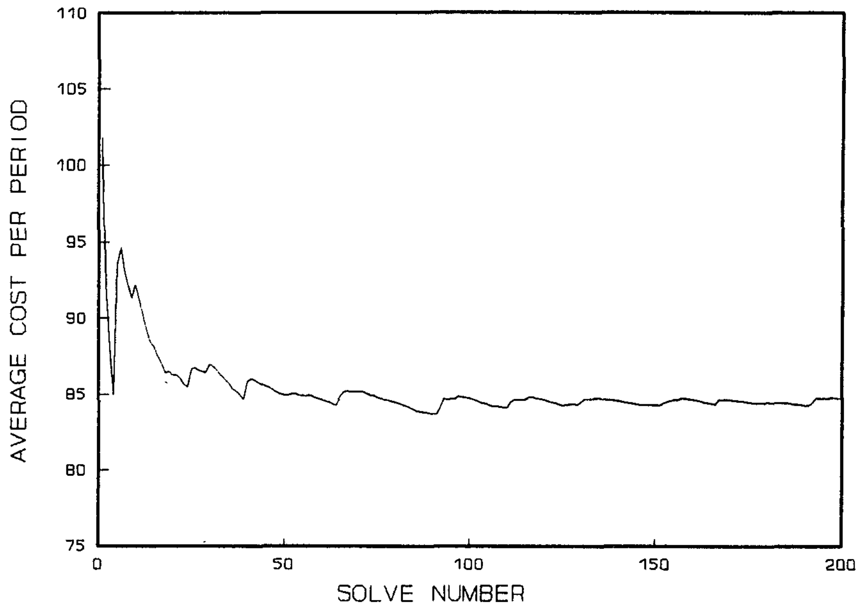


FIGURE B-1
AVERAGE 24 PERIOD COST PER PERIOD
RESTORATION MODEL WITHOUT LOST SALES
LOW LABOR/HIGH HOLDING COSTS

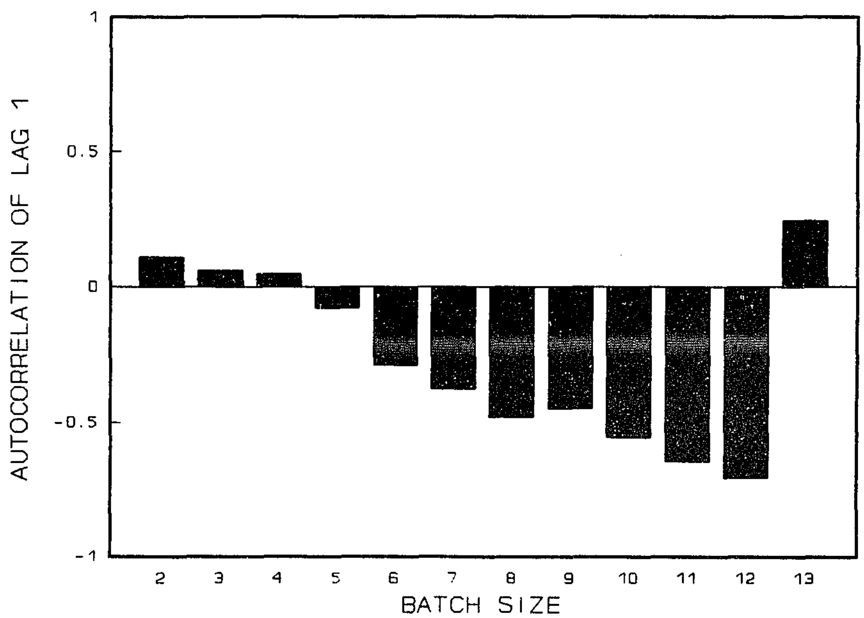


FIGURE B-2
LAG 1 AUTOCORRELATION OF TOTAL COSTS VERSUS BATCH SIZE
RESTORATION MODEL WITHOUT LOST SALES
LOW LABOR/HIGH HOLDING COSTS

TABLE B-1
MEANS, VARIANCES, AND β-LEVELS[1] (α-LEVEL = 5%)
RESTORATION MODEL WITHOUT LOST SALES
APPLIED TO THE BASE NETWORK OF APPENDIX A
BASED ON 20 BATCHES

Target Inventory[2]	Mean	Variance	ϕ^2;[3]	β-Level		
				n=10	n=20	n=30
1-3	81.31	2.84	0.704n	0.09	<0.01	<0.01
2-3	55.54	2.09	0.957n	0.03	<0.01	<0.01
3-4	45.78	6.43	0.311n	0.38	0.08	0.02
3-5	53.81	32.26	0.062n	-	-	0.60

Restoration Coefficient[2]	Mean	Variance	ϕ^2;[3]	β-Level		
				n=10	n=20	n=30
1-3	0.56	0.053	0.094n	-	0.60	0.38
1-4	0.87	0.008	0.625n	0.10	<0.01	<0.01
1-5	0.87	0.009	0.556n	0.16	<0.01	<0.01
2-3	0.68	0.063	0.079n	-	-	0.43
2-4	0.86	0.014	0.357n	0.30	0.06	<0.01
2-5	0.89	0.008	0.625n	0.10	<0.01	<0.01
3-4	0.89	0.007	0.714n	0.08	<0.01	<0.01
3-5	0.91	0.008	0.625n	0.10	<0.01	<0.01

1 β-Levels = Probability of rejecting the null hypothesis given that it is true. Values for the β-levels are taken from operating characteristic charts in Montgomery (1991).

2 Target Inventory 1-3 refers to the sum of inventory:
in-transit from node 1 to node 3, plus
in-wait at node 3 (from node 1), plus
in-process at node 3, plus
in-buffer at node 3.

3 $\phi^2 = \frac{nD^2}{2\sigma^2}$ where n = Number of replicates
D = Prespecified difference to delineate
σ^2 = Variance of target inventories

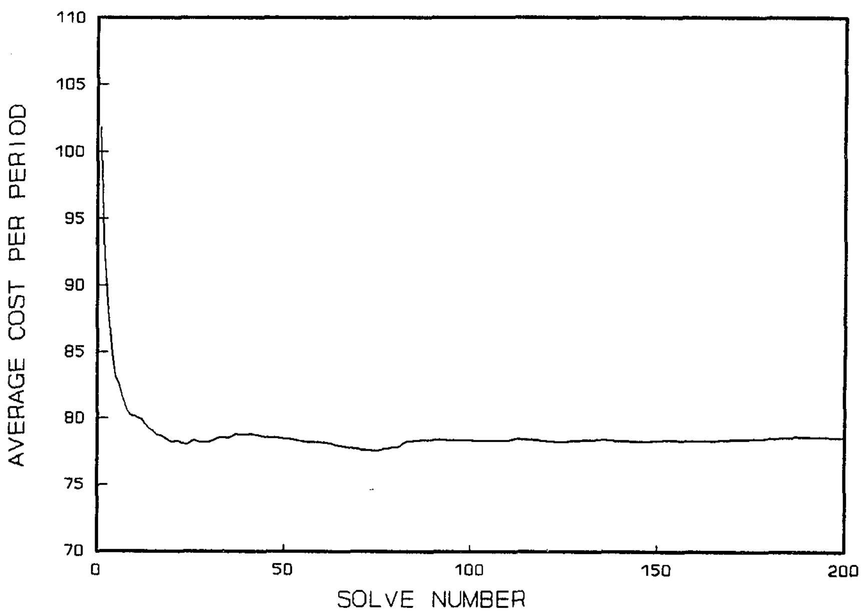


FIGURE B-3
AVERAGE 24 PERIOD COST PER PERIOD
RESTORATION MODEL WITH LOST SALES
LOW LABOR/HIGH HOLDING COSTS

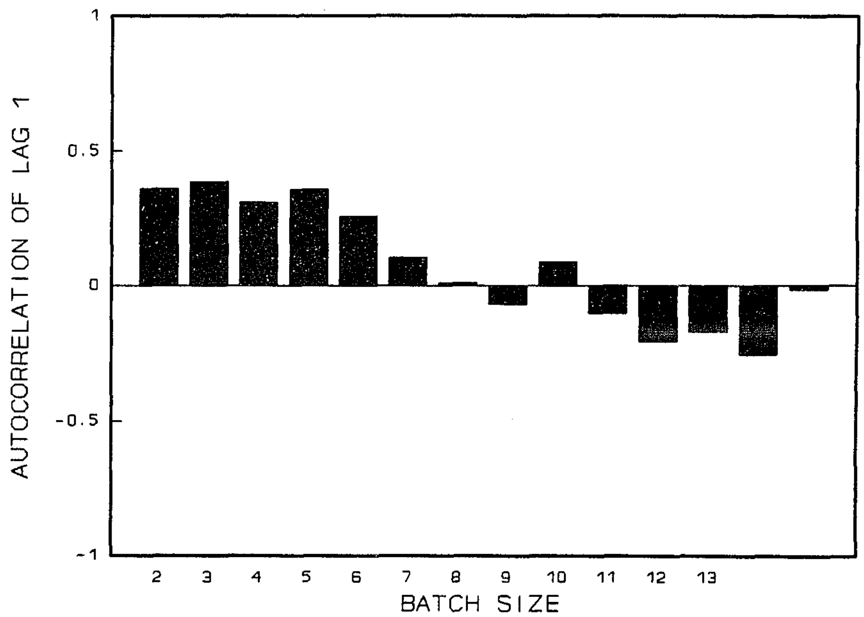


FIGURE B-4
LAG 1 AUTOCORRELATION OF TOTAL COSTS VERSUS BATCH SIZE
RESTORATION MODEL WITH LOST SALES
LOW LABOR/HIGH HOLDING COSTS

TABLE B-2
MEANS, VARIANCES, AND β-LEVELS[1] (α-LEVEL = 1%)
RESTORATION MODEL WITH LOST SALES
APPLIED TO THE BASE NETWORK OF APPENDIX A
BASED ON 20 BATCHES

Target Inventory[2]	Mean	Variance	ϕ^2[3]	β-Level n=10	β-Level n=20	β-Level n=30
1-3	80.78	2.93	0.682n	0.30	0.02	<0.01
2-3	55.20	1.52	1.316n	<0.05	<0.01	<0.01
3-4	42.96	0.90	2.223n	<0.01	<0.01	<0.01
3-5	49.46	1.77	1.130n	<0.07	<0.01	<0.01

Restoration Coefficient[2]	Mean	Variance	ϕ^2[3]	β-Level n=10	β-Level n=20	β-Level n=30
1-3	0.65	0.032	0.156n	-	-	0.40
1-4	0.97	0.001	5.000n	<0.01	<0.01	<0.01
1-5	0.92	0.004	1.250n	0.06	<0.01	<0.01
2-3	0.65	0.049	0.102n	-	-	-
2-4	0.98	0.001	5.000n	<0.01	<0.01	<0.01
2-5	0.94	0.002	2.500n	<0.01	<0.01	<0.01
3-4	0.96	0.002	2.500n	<0.01	<0.01	<0.01
3-5	0.99	0.001	5.000n	<0.01	<0.01	<0.01

[1] β-Levels = Probability of rejecting the null hypothesis given that it is true. Values for the β-levels are taken from operating characteristic charts in Montgomery (1991).

[2] Target Inventory 1-3 refers to the sum of inventory:
in-transit from node 1 to node 3, plus
in-wait at node 3 (from node 1), plus
in-process at node 3, plus
in-buffer at node 3.

[3] $\phi^2 = \frac{nD^2}{2\sigma^2}$ where n = Number of replicates
D = Prespecified difference to delineate
σ^2= Variance of target inventories

APPENDIX C

ANALYSIS OF VARIANCES FOR TARGET INVENTORIES, ACTUAL INVENTORIES, AND RESTORATION COEFFICIENTS 2X2 FULL FACTORIAL EXPERIMENT FACTORS LABOR COSTS AND HOLDING COSTS

TABLE C-1
ANALYSIS OF VARIANCE FOR TARGET INVENTORY 2-3
2^2 FULL FACTORIAL EXPERIMENT USING LOST SALES MODEL
FACTORS LABOR COSTS AND HOLDING COSTS

Source	DF	Mean-Square	F-Ratio	Prob > F
Labor Costs	1	7.9281	4.33	0.0408
Holding Costs	1	46.9777	25.66	0.0000*
Interaction	1	0.7409	0.40	0.5266
Error	76	1.8307		
Total Sum-Squares	79	194.7798		

TABLE C-2
ANALYSIS OF VARIANCE FOR TARGET INVENTORY 3-4
2^2 FULL FACTORIAL EXPERIMENT USING LOST SALES MODEL
FACTORS LABOR COSTS AND HOLDING COSTS

Source	DF	Mean-Square	F-Ratio	Prob > F
Labor Costs	1	0.2463	0.23	0.6350
Holding Costs	1	0.0935	0.09	0.7697
Interaction	1	0.4080	0.38	0.5414
Error	76	1.0840		
Total Sum-Squares	79	83.1317		

TABLE C-3
ANALYSIS OF VARIANCE FOR TARGET INVENTORY 3-5
2^2 FULL FACTORIAL EXPERIMENT USING LOST SALES MODEL
FACTORS LABOR COSTS AND HOLDING COSTS

Source	DF	Mean-Square	F-Ratio	Prob > F
Labor Costs	1	2.1255	1.04	0.3121
Holding Costs	1	10.5893	5.16	0.0259
Interaction	1	1.2722	0.62	0.4335
Error	76	2.0523		
Total Sum-Squares	79	169.9620		

* Significant at the 1% level.

TABLE C-4
ANALYSIS OF VARIANCE FOR ACTUAL INVENTORY 1-3
2^2 FULL FACTORIAL EXPERIMENT USING LOST SALES MODEL
FACTORS LABOR COSTS AND HOLDING COSTS

Source	DF	Mean-Square	F-Ratio	Prob > F
Labor Costs	1	1.2248	0.50	0.4829
Holding Costs	1	47.9681	19.47	0.0000*
Interaction	1	0.0090	0.00	0.9519
Error	76	2.4633		
Total Sum-Squares	79	236.4105		

TABLE C-5
ANALYSIS OF VARIANCE FOR ACTUAL INVENTORY 2-3
2^2 FULL FACTORIAL EXPERIMENT USING LOST SALES MODEL
FACTORS LABOR COSTS AND HOLDING COSTS

Source	DF	Mean-Square	F-Ratio	Prob > F
Labor Costs	1	5.8405	3.06	0.0844
Holding Costs	1	43.9455	23.01	0.0000*
Interaction	1	0.5476	0.29	0.5939
Error	76	1.9099		
Total Sum-Squares	79	195.4803		

TABLE C-6
ANALYSIS OF VARIANCE FOR ACTUAL INVENTORY 3-4
2^2 FULL FACTORIAL EXPERIMENT USING LOST SALES MODEL
FACTORS LABOR COSTS AND HOLDING COSTS

Source	DF	Mean-Square	F-Ratio	Prob > F
Labor Costs	1	1.4545	2.21	0.1413
Holding Costs	1	2.3115	3.51	0.0648
Interaction	1	0.1283	0.19	0.6601
Error	76	0.6584		
Total Sum-Squares	79	53.9351		

* Significant at the 1% level.

TABLE C-7
ANALYSIS OF VARIANCE FOR ACTUAL INVENTORY 3-5
2^2 FULL FACTORIAL EXPERIMENT USING LOST SALES MODEL
FACTORS LABOR COSTS AND HOLDING COSTS

Source	DF	Mean-Square	F-Ratio	Prob > F
Labor Costs	1	4.4874	1.48	0.2269
Holding Costs	1	17.8889	5.92	0.0174
Interaction	1	0.7742	0.26	0.6143
Error	76	3.0236		
Total Sum-Squares	79	252.9446		

TABLE C-8
ANALYSIS OF VARIANCE FOR RESTORATION COEFFICIENT 1-4
2^2 FULL FACTORIAL EXPERIMENT USING LOST SALES MODEL
FACTORS LABOR COSTS AND HOLDING COSTS

Source	DF	Mean-Square	F-Ratio	Prob > F
Labor Costs	1	0.0004	0.21	0.6513
Holding Costs	1	0.0003	0.15	0.6960
Interaction	1	0.0012	0.68	0.4131
Error	76	0.0017		
Total Sum-Squares	79	0.1329		

TABLE C-9
ANALYSIS OF VARIANCE FOR RESTORATION COEFFICIENT 1-5
2^2 FULL FACTORIAL EXPERIMENT USING LOST SALES MODEL
FACTORS LABOR COSTS AND HOLDING COSTS

Source	DF	Mean-Square	F-Ratio	Prob > F
Labor Costs	1	0.0096	2.94	0.0903
Holding Costs	1	0.0289	8.83	0.0040*
Interaction	1	0.0000	0.00	0.9579
Error	76	0.0033		
Total Sum-Squares	79	0.2877		

* Significant at the 1% level.

TABLE C-10
ANALYSIS OF VARIANCE FOR RESTORATION COEFFICIENT 2-4
2^2 FULL FACTORIAL EXPERIMENT USING LOST SALES MODEL
FACTORS LABOR COSTS AND HOLDING COSTS

Source	DF	Mean-Square	F-Ratio	Prob > F
Labor Costs	1	0.0006	0.99	0.3238
Holding Costs	1	0.0004	0.69	0.4080
Interaction	1	0.0005	0.96	0.3304
Error	76	0.0006		
Total Sum-Squares	79	0.0446		

TABLE C-11
ANALYSIS OF VARIANCE FOR RESTORATION COEFFICIENT 2-5
2^2 FULL FACTORIAL EXPERIMENT USING LOST SALES MODEL
FACTORS LABOR COSTS AND HOLDING COSTS

Source	DF	Mean-Square	F-Ratio	Prob > F
Labor Costs	1	0.0075	3.40	0.0692
Holding Costs	1	0.0173	7.81	0.0066*
Interaction	1	0.0005	0.24	0.6240
Error	76	0.0022		
Total Sum-Squares	79	0.1936		

TABLE C-12
ANALYSIS OF VARIANCE FOR RESTORATION COEFFICIENT 3-4
2^2 FULL FACTORIAL EXPERIMENT USING LOST SALES MODEL
FACTORS LABOR COSTS AND HOLDING COSTS

Source	DF	Mean-Square	F-Ratio	Prob > F
Labor Costs	1	0.0033	1.42	0.2379
Holding Costs	1	0.0093	4.04	0.0480
Interaction	1	0.0004	0.17	0.6830
Error	76	0.0022		
Total Sum-Squares	79	0.1875		

* Significant at the 1% level.

TABLE C-13
ANALYSIS OF VARIANCE FOR RESTORATION COEFFICIENT 3-5
2^2 FULL FACTORIAL EXPERIMENT USING LOST SALES MODEL
FACTORS LABOR COSTS AND HOLDING COSTS

Source	DF	Mean-Square	F-Ratio	Prob > F
Labor Costs	1	0.0021	4.67	0.0338
Holding Costs	1	0.0040	9.04	0.0036*
Interaction	1	0.0002	0.44	0.5081
Error	76	0.0004		
Total Sum-Squares	79	0.0403		

* Significant at the 1% level.

APPENDIX D

ANALYSIS OF VARIANCES FOR TARGET INVENTORIES, ACTUAL INVENTORIES, AND RESTORATION COEFFICIENTS SINGLE FACTOR EXPERIMENT FACTOR COEFFICIENT OF VARIATION

TABLE D-1
ANALYSIS OF VARIANCE FOR TARGET INVENTORY 2-3
SINGLE FACTOR EXPERIMENT USING LOST SALES MODEL
FACTOR COEFFICIENT OF VARIATION

Source	DF	Mean-Square	F-Ratio	Prob > F
Coefficient of Variation	3	148.65	89.29	0.0000*
Error	76	1.67		
Total Sum-Squares	79	572.46		

TABLE D-2
ANALYSIS OF VARIANCE FOR TARGET INVENTORY 3-4
SINGLE FACTOR EXPERIMENT USING LOST SALES MODEL
FACTOR COEFFICIENT OF VARIATION

Source	DF	Mean-Square	F-Ratio	Prob > F
Coefficient of Variation	3	92.75	73.73	0.0000*
Error	76	1.26		
Total Sum-Squares	79	373.86		

TABLE D-3
ANALYSIS OF VARIANCE FOR TARGET INVENTORY 3-5
SINGLE FACTOR EXPERIMENT USING LOST SALES MODEL
FACTOR COEFFICIENT OF VARIATION

Source	DF	Mean-Square	F-Ratio	Prob > F
Coefficient of Variation	3	114.92	66.80	0.0000*
Error	76	1.72		
Total Sum-Squares	79	475.50		

TABLE D-4
ANALYSIS OF VARIANCE FOR ACTUAL INVENTORY 1-3
SINGLE FACTOR EXPERIMENT USING LOST SALES MODEL
FACTOR COEFFICIENT OF VARIATION

Source	DF	Mean-Square	F-Ratio	Prob > F
Coefficient of Variation	3	123.2598	56.09	0.0000*
Error	76	2.1976		
Total Sum-Squares	79	536.7980		

* Significant at the 1% level.

TABLE D-5
ANALYSIS OF VARIANCE FOR ACTUAL INVENTORY 2-3
SINGLE FACTOR EXPERIMENT USING LOST SALES MODEL
FACTOR COEFFICIENT OF VARIATION

Source	DF	Mean Square	F-Ratio	Prob > F
Coefficient of Variation	3	130.9755	86.75	0.0000*
Error	76	1.5099		
Total Sum-Squares	79	507.6753		

TABLE D-6
ANALYSIS OF VARIANCE FOR ACTUAL INVENTORY 3-4
SINGLE FACTOR EXPERIMENT USING LOST SALES MODEL
FACTOR COEFFICIENT OF VARIATION

Source	DF	Mean-Square	F-Ratio	Prob > F
Coefficient of Variation	3	48.1149	46.22	0.0000*
Error	76	1.0409		
Total Sum-Squares	79	223.4559		

TABLE D-7
ANALYSIS OF VARIANCE FOR ACTUAL INVENTORY 3-5
SINGLE FACTOR EXPERIMENT USING LOST SALES MODEL
FACTOR COEFFICIENT OF VARIATION

Source	DF	Mean-Square	F-Ratio	Prob > F
Coefficient of Variation	3	88.7071	39.46	0.0000*
Error	76	2.2479		
Total Sum-Squares	79	436.9596		

TABLE D-8
ANALYSIS OF VARIANCE FOR RESTORATION COEFFICIENT 1-4
SINGLE FACTOR EXPERIMENT USING LOST SALES MODEL
FACTOR COEFFICIENT OF VARIATION

Source	DF	Mean-Square	F-Ratio	Prob > F
Coefficient of Variation	3	0.0339	11.74	0.0000*
Error	76	0.0029		
Total Sum-Squares	79	0.3213		

* Significant at the 1% level.

TABLE D-9
ANALYSIS OF VARIANCE FOR RESTORATION COEFFICIENT 1-5
SINGLE FACTOR EXPERIMENT USING LOST SALES MODEL
FACTOR COEFFICIENT OF VARIATION

Source	DF	Mean-Square	F-Ratio	Prob > F
Coefficient of Variation	3	0.0705	13.92	0.0000*
Error	76	0.0051		
Total Sum-Squares	79	0.5969		

TABLE D-10
ANALYSIS OF VARIANCE FOR RESTORATION COEFFICIENT 2-4
SINGLE FACTOR EXPERIMENT USING LOST SALES MODEL
FACTOR COEFFICIENT OF VARIATION

Source	DF	Mean-Square	F-Ratio	Prob > F
Coefficient of Variation	3	0.0268	29.13	0.0000*
Error	76	0.0009		
Total Sum-Squares	79	0.1501		

TABLE D-11
ANALYSIS OF VARIANCE FOR RESTORATION COEFFICIENT 2-5
SINGLE FACTOR EXPERIMENT USING LOST SALES MODEL
FACTOR COEFFICIENT OF VARIATION

Source	DF	Mean-Square	F-Ratio	Prob > F
Coefficient of Variation	3	0.0422	14.03	0.0000*
Error	76	0.0030		
Total Sum-Squares	79	0.3554		

TABLE D-12
ANALYSIS OF VARIANCE FOR RESTORATION COEFFICIENT 3-4
SINGLE FACTOR EXPERIMENT USING LOST SALES MODEL
FACTOR COEFFICIENT OF VARIATION

Source	DF	Mean-Square	F-Ratio	Prob > F
Coefficient of Variation	3	0.0529	20.22	0.0000*
Error	76	0.0026		
Total Sum-Squares	79	0.3574		

* Significant at the 1% level.

TABLE D-13
ANALYSIS OF VARIANCE FOR RESTORATION COEFFICIENT 3-5
SINGLE FACTOR EXPERIMENT USING LOST SALES MODEL
FACTOR COEFFICIENT OF VARIATION

Source	DF	Mean-Square	F-Ratio	Prob > F
Coefficient of Variation	3	0.0010	14.60	0.0000*
Error	76	0.0007		
Total Sum-Squares	79	0.0819		

* Significant at the 1% level.

VITA

Gary Clendenen received a B.S. in mathematics from New Mexico State University in 1972; a M.A. in Mathematics from the University of New Mexico in 1973; and a M.S. in Forest Resources from the University of Washington in 1975. He then worked in the insurance industry until 1980 at which time he started his own oil and gas exploration company. In 1989, Mr. Clendenen began the Ph.D. program in Quantitative Business Analysis at Louisiana State University. He is currently an Assistant Professor at the University of Texas at Tyler, Texas.

DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: Gary Wayne Clendenen

Major Field: Business Administration

Title of Dissertation: A Restoration Based Model for Materials Management in a Global Manufacturing Environment

Approved:

Dan B. Rinks

Major Professor and Chairman

Daniel Fogel

Dean of the Graduate School

EXAMINING COMMITTEE:

James M Pruett

[illegible]

[illegible]

[illegible]

Ralph W. [illegible]

Date of Examination:

July 22, 1993