POSITIONING INVENTORY IN DISTRIBUTION SYSTEMS WITH STOCHASTIC DEMAND

by

Shitao Yang

A Dissertation submitted to the faculty of The University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Kenan-Flagler Business School.

Chapel Hill

1999

Approved by:

O-Reg Whybark Adviser

J. Morgan Jones Reader

Reader

Ann E Marucheck Reader

John P. Evans Reader

UMI Number: 9943273

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

SHITAO YANG: Positioning Inventory in Distribution Systems with Stochastic Demand (Under the direction of DAVID CLAY WHYBARK)

ABSTRACT

This dissertation treats an important theoretical and practical problem: whether inventory in a distribution system should be positioned at retail facilities near the customer or at a warehouse closer to the outside supplier (e.g., a factory). In the extreme, distribution managers can position all of their inventory at the retail facilities (using the warehouse simply to break bulk). Having some inventory positioned at the warehouse, however, enables managers to send it out to the retail stores later. Intuitively, the idea of holding some inventory at the warehouse is attractive since it provides the ability to react to changing conditions at the retail level. On the other hand, inventory held at the warehouse is not available to meet the instant demand at the retail stores where it occurs.

In this dissertation, we conducted a simulation experiment that carefully controlled the conditions studied to gain clear insights into the question. For companies that must fill customer demand from inventory, the results indicate that higher levels of customer fill-rates are achieved by positioning the inventory near the customer. The results also show that given a set of inventory, transportation, and control system resources, the difference between the maximum fill-rate achievable and that for positioning inventory as close to the customer as possible, is very small. The general finding that inventory should be positioned near the customer to get high levels of customer fill-rate does not change materially with any of the experimental factors investigated. This observation holds even as demand uncertainty increases to extremely high levels. We also investigated the effect of inventory control systems on the positioning of inventory in the distribution system. The experiment for this investigation involved four control systems with a similar form of decision rules but different information requirements. The design was intended to reflect various information-sharing schemes we observed in industry. We found that information-sharing shifted inventory from the warehouse to the retail stores and improved customer fill rate. The other experimental factors investigated included system-wide inventory

levels, transit lead times, predetermined shipment frequencies, and the number of retail stores. This dissertation ends with concluding remarks on research contributions, managerial implications, and directions for future research.

ACKNOWLEDGEMENTS

I owe a great deal to my advisor, Professor D. Clay Whybark, for his guidance, encouragement, and patience. He has taught me a framework for conducting research that will influence any future research I undertake. His enthusiasm in developing elegant approaches to important managerial issues have provided me with a role model and significantly heightened my scientific and personal aspirations.

Each of my other committee members has also been instrumental in my development. I appreciate the substantial amount of time Professor Morgan Jones has devoted to the doctoral program in Operations Management area and to myself in particular. Without his help, it would be impossible to complete this dissertation. Professor John P. Evans has also helped me immensely. Professor Marucheck has always been there and helped me out whenever I had difficulty. I would thank Professor Dennis Rondinelli not only for serving in my dissertation committee bust also for providing me the opportunity to serve as his research assistant at the Kenan Private Enterprise Center.

I would also like to thank Professor Harvey M. Wagner for sharing his extensive knowledge in the fields of inventory control, optimization, simulation, and statistical analysis. I thank Professor Curtis P. McLaughlin for co-authoring two journal articles and teaching me how to be a team player. Several other faculty members also served to influence and encourage me. I would like to thank Professors Douglas A. Elvers, Kyle D. Cattani, Ronald T. Pannesi, and Professor Aleda V. Roth. I also want thank Dr. Edward Neal, the director of UNC Center for Teaching and Learning, for helping me start my teaching career. Special thanks are reserved for Professor William A. Fischer for helping me start my education at the University of North Carolina and to Professor Jack N. Behrman and Mrs. Louise Behrman for being my mentors since I came to the United States.

Finally, I would like to thank my grandmother and my parents for providing an environment conducive to learning early in my life. After I came to the University of North

Carolina at Chapel Hill, my grandmother, a retired teacher of Chinese literature, and my father, a surgeon, passed away. I am very grateful to my mother, an emeritus professor of pediatrics, for her love and support during this very difficult time and along the long path of my education.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

α	Percentage of Incoming Stocks Immediately Allocated in the "Push" System
ASN	Advanced Shipment Notice or Advance Shipment Notice Control System
ASNS	Advanced Shipment Notice Control System
CV	Coefficient of Variation
Dj(t)	Customer Demand at the jth Retail Store in Period t
DRP	Distribution Requirement Planning
EROP	Echelon Reorder and Order-up-to Level Control System
F	Customer Fill Rate
F*	Maximum Fill Rate
Fc	Fill Rate for Positioning Close to the Customer
FCFS	First-Come-First-Served
Fs	Fill Rate for the "Ship-All"
FS(t)	Balanced Level of the Retail Stores in Period t
Δ F*	Change in Maximum Fill Rate
Δ (F*-Fc)	Fill Rate Improvement due to the best Positioning or Penalty for Positioning Close to the Customer
Δ (Fc-Fs)	Fill Rate Improvement Attributable to the Increase in SHIPR
INV	System-wide Inventory Level (i.e., the sum of the average on-hand inventory levels at the warehouse and that at the retailer stores.)
INV#	Predetermined System-wide Inventory
IPj(t)	Inventory Position of the jth Retail Store in Period t
L_r	Transit Lead Time from the Warehouse to Any of the Retail Stores
Δ L_r	Change in L_r
LROP	Local Information Reorder and Order-up-to Level Control System
L_w	Transit lead time from the Outside Supplier to the Warehouse
Δ L_w	Change in L_w
N	The Number of the Retail Stores Serviced by the Warehouse
Δ N	Change in N
$N(\mu,\sigma^2)$	Normal Distribution with Mean μ and Variance σ^2

$N_j(t)$	The Net Inventory of the jth Retail Store at the Beginning of the Period t After Receiving the Stocks Allocated at Period t-Lr-1 before $D_j(t)$ Occurs
NL(μ,σ^2)	Lognormal Distribution with Location Parameter μ and Shape Parameter σ
$O(t)$	Order Quantity Placed by a Stocking Location in Period t
$O_j(t)$	Order Quantity Placed on the Warehouse by the jth Retail Store in Period t
P	Inventory Positioning Ratio
P*	The Best Inventory Positioning Ratio
Δ P*	The change in P or Percentage of INV Shifted from the Warehouse to the Retail Stores
Pc	Positioning Ratio of Positioning Close to the Customer
PO	Planned Order
Ps	Positioning Ratio for the Ship-all (which is always equal to 1)
"Push"	"Push" Control System
RIP	Information about Retailers' Inventory Positions Known to the Warehouse
ROP	Reorder and Order-up-to Level Control System
R	Generic Inventory Review Period(s)
R_w	Warehouse Inventory Review Period(s) or the Warehouse Replenishment Cycle
s	Generic Reorder Point
S	Generic Order-up-to Level
SHIPR	Shipment Frequency to Each of the Retail Stores (measured by the average number of shipments per retail store per period)
SHIPR#	Predetermined Shipments Frequency to Each of the Retail Stores
SHIPW	Shipment Frequency to the Warehouse (measured by the average number of shipments per period)
SHIPW#	Predetermined Shipment Frequency to the Warehouse
s_r	Retailer's Reorder Point
Sr	Retailer's Order-up-to Level
s_w	Warehouse's Reorder Point
S_w	Warehouse's Order-up-to Level
t'	The First Period in which Simulation Data were Collected
T	The Delay Time for Making the Secondary Allocation in the "Push" System
t' + T'	The Last Period in which Simulation Data were Collected

TPPD	Time-Phased Projected Demand (or Request)
$UF_j(t)$	Unsatisfied Demand at the jth Retail Store in Period t
$X(t)$	Inventory Status in Period t
$X_j(t)$	Inventory Status of the jth Retail Store in Period t
$WI(t)$	Warehouse On-Hand Inventory in Period t

CHAPTER 1
INTRODUCTION

1.1 THE PROBLEM

This dissertation treats an important practical and theoretical problem: whether inventory in a distribution system should be positioned at retail facilities near the customer or at a warehouse closer to the outside supplier (e.g., a factory). The importance of this problem has grown recently as companies are increasingly caught between the pressures of the market and the high cost of inventory. Global competition and the growing emphasis on customer satisfaction have underscored the need to improve customer service levels. At the same time, capital, space and obsolescence costs of carrying inventory have increased, necessitating prudent management of inventory.

There are several methods of improving customer service levels in the presence of uncertain demand. The traditional method, increasing inventory, may not be prudent, given the costs mentioned above. Decreasing uncertainty may be possible to achieve in those markets where customer alliances allow the development of information-sharing between suppliers and customers, but this is difficult to bring about in the consumer market. Increased shipment frequency may improve service levels, but at some cost. All of these alternatives can be expensive and some, as indicated, may not even be feasible, which reduces the issue to one question: how can distribution managers use existing resources more efficiently?

That issue--how to make better use of existing resources--is the problem this dissertation will be addressing. Specifically, we answer the question of where inventory should be positioned in a distribution system to get the best customer service level, given a set of inventory, transportation, and inventory control system resources. The market studied is the consumer market where demand uncertainty is present, and the customer service level is defined as the percentage of demand that is satisfied immediately from inventory, the fill-rate.

In the extreme, distribution managers can position all of their inventory at the retail facilities (using the warehouse simply to break bulk). Having some inventory positioned at the warehouse, however, enables managers to send it out to the retail stores later. Intuitively, the idea of holding some inventory at the warehouse is attractive since it provides the ability to react to changing conditions at the retail level. On the other hand, inventory held at the warehouse is not available to meet the instant demand at the retail stores where it occurs.

In this dissertation, we conducted a simulation experiment that carefully controlled the conditions studied to gain clear insights into the question. The dissertation starts with a description of the practical interest in the positioning problem and some conflicting results from the literature. Next, we explain the research methodology used, and follow this with a chapter devoted to the description of inventory control systems used. A detailing of our experimental design precedes the presentation of the results of the experiments. The dissertation ends with concluding remarks on research contribution, managerial implications, and directions for future research.

1.2 BACKGROUND

Positioning inventory is a problem that confronts distribution managers in virtually all companies with distribution networks. Practical solutions to the problem include the concept of "national safety stock" advocated by R.G. Brown (1977) and still incorporated into the Baxter's distribution inventory management philosophy for holding inventory at central warehouses (Baxter International Inc., Annual Report, 1992). At the other extreme, Wal-Mart uses the concept of "cross-docking" to minimize the level of inventory held in the company's distribution centers (G. Stalk et al. 1992). With "cross-docking," goods are continuously delivered to Wal-Mart's distribution centers, where they are selected, repacked, and then dispatched to stores, often without ever sitting in inventory. The diversity of positioning practices reflects the fact that very little consensus on where to position inventory exists among practitioners. The question is usually settled by executive decision based largely on intuitive grounds.

Bowersox (1963), among others researchers, cautions against the tendency among firms to maintain inventories in each customer's backyard, saying that "Maintaining large numbers of localized field inventories is an expensive habit, which, in the final analysis, may result in substantial losses." On the other hand, the Boston Consulting Group argues that Wal-Mart's remarkable success lies in its "cross-docking" practice (G. Stalk et al. 1992). Unfortunately, Bowersox has never provided his "final analysis" to support his argument for positioning inventory at central facilities. Similarly, some of the business literature that advocates cross-docking for streamlining of supply chains often simply states the potential benefits (ranging from $30 billion for the grocery industry by Kurt Salmon Associates, 1993, to $14 billion for the food-service industry by Troyer, 1996) without detailed analyses. As a result, the practitioners often have to take the advice on the positioning problems by faith.

Traditionally, distribution systems have been managed by adaptations of methods designed for managing inventory at individual stocking sites. The justification for this approach is that there has been little in the way of an information infrastructure available for managing system-wide inventories. Even if the information for the coordination were available, managers often would have to be assigned to individual stocking sites for organizational reasons (Hausman and Erkip, 1994). With this management philosophy, the question concerning where to position system-wide inventories has not drawn much managerial attention.

In the 1990s, however, both technology and management philosophies have changed. The rapid development of information technology provides distribution managers greater and faster access to information than ever before. New technologies such as bar coding, point-of-sale screening, computer-based merchandise allocation systems, ASN (Advanced Shipment Notice), and EDI (electric data interchange) have rapidly increased the visibility of the inventory flows throughout manufacturing and distribution networks or supply chains. With such visibility, distribution managers are increasingly being given the power and the responsibility to manage a large-scale process, cutting across established boundaries of functions, divisions, and even firms. The tradition of distribution managers taking care of stocking sites individually in isolation no longer prevails.

Many companies are now re-examining their positioning practices. Some are shocked to find out how much inventory has accumulated throughout their distribution networks. For instance, in 1992, an estimated $75 billion to $100 billion in grocery products sat at any one time on trucks and railcars, or were stocked inside distribution centers, caught in a gruesome gridlock (Seller, 1992). This shocking observation led the grocery industry to believe that streamlining their supply chains alone might eliminate $30 billion or nearly 10% of its annual operating costs (Henkoff, 1994). Others have found that their practice of positioning inventories by trial-and-error have gone through some unwelcome circles (Chain Store Age Executive, April 1992, p31A) and are beginning to wonder how much money they could have saved if they had been able to deal with the positioning problem more intelligently. Companies across industries seem to have reached the same conclusion: "You have far more opportunities to get cost out of the supply chain than you do out of manufacturing. There's so much duplication and inefficiency" (Henkoff, 1994). Consequently, logistics, or supply chain management, long an unsung, operations-intensive area, has suddenly become very strategic. So also has the inventory positioning problem, long identified as one of the fundamental issues for logistics management (Magee, et al. 1985), gained stature as a key strategy for streamlining supply chains.

It is of note that the globalization process has also heightened the importance of the inventory positioning problem. Some truly multinational firms now operate facilities in virtually every corner of the world, causing their sourcing, production, warehousing, and distribution decisions to be made on a global basis. The proper positioning of inventory throughout those global distribution networks potentially could save millions of dollars. While many companies are restructuring their supply chains, they find that a fresh perspective on the capacities of their restructured logistics /distribution systems is often associated with new strategies for positioning system-wide inventories (Gopal, 1992). An important feature of the single European market, for instance, is found to be the improved possibility for proper positioning inventory on the basis of an all Europe logistics consideration rather than on the common basis of national territory (Van der Hoop, 1992). Says Harold Sirkin, a vice president of the Boston Consulting Group, "As the economy changes, as competition becomes more global, it is no longer company vs. company, but supply chain vs. supply chain" (Henkoff, 1994).

While the importance of the inventory positioning problem has increasingly been recognized, how to solve the problem by and large remains an open issue in the research literature as well as with distribution managers (Baker 1993, Nahmias 1997, Vollmann, Berry, and Whybark 1997, and Zipkin 1998). The lack of normative results in the research literature on the proper positioning of inventory and the on-going efforts of many companies to streamlining their supply chain operations motivated this dissertation study. Our goal is to extend the current understanding of the inventory positioning problem and to provide distribution managers with theoretically sound, yet practical, guidelines for determining where inventory should be positioned in distribution systems with stochastic demand.

1.3 PREVIEW

For this dissertation study, we have considered a distribution system consisting of a central warehouse supplied by a completely reliable outside supplier (e.g., a factory) and serving several retail stores, which, in turn, supply stochastic customer demand in a periodic review environment. The question addressed is where inventory should be positioned in the system to get the highest level of customer fill rate. We based the study on a simulation experiment. One of the key features of this study is that we carefully controlled the frequency of shipments from the outside supplier to the warehouse and from the warehouse to the retail stores as we simulated and evaluated different strategies for positioning system-wide inventories. The approach ensured the comparability of the alternative positioning strategies, an element largely ignored in previous studies.

Another important feature of this study is that it involved multiple inventory control systems with similar decision rules, but different information requirements. We carefully designed these control systems to reflect various information-sharing schemes we observed in industry. Such a design enabled us to investigate the effect of information sharing on the positioning of inventory in the distribution system.

The key findings of our analysis can be summarized as follows. First, inventory should be positioned near the customer to get the high levels of customer fill rate. Second, the warehouse needs to keep a relatively small amount of inventory to get the maximum fill rate. However, the difference between the maximum fill rate and that for positioning inventory as close to the customer as possible is fairly small. Third, information sharing shifts inventory from the warehouse into the retail stores and improves the maximum fill rate. Fourth, as long as the fill-rate is an appropriate service criterion, the general finding that inventory should be positioned near the customer to get high levels of customer fill rate does not change materially with any of the experimental factors we investigated. These factors include three resource factors (system-wide inventory level, predetermined shipment frequency, and inventory control system used) and three environmental factors (demand uncertainty, transit lead times, and the number of retail stores).

The remainder of this dissertation is organized as follows: Chapter 2 reviews the literature related to the inventory-positioning problem. Section 2.1 provides an introduction. Sections 2.2 to Section 2.4 review key studies on the inventory positioning problem, highlighting their assumptions, methodologies, and conclusions. Section 2.5 offers remarks on what we know and what we do know. The open issues and conjectures represent opportunities for our investigation.

Chapter 3 describes the methodology we used. Section 3.1 explains why we chose to use computer simulation, rather than analytical approach, as our primary research tool. Section 3.2 features the new approaches developed. Section 3.3 provides the simulation model framework and defines three positioning strategies to be identified for each design points.

Chapter 4 gives a detailed description of inventory control systems used as well as the search procedures for identifying the three positioning strategies defined in Chapter 3.

Chapter 5 is devoted to the experimental design. Section 5.1 specifies the stochastic process representing customer demand and explains customer service criterion used in our simulation experiment. Section 5.2 gives some detailed description about our simulation models so that the experimental factors we considered could be defined precisely. Section 5.3 highlights the design of our experiments. Section 5.4 details how we implemented and validated our simulation.

Chapter 6 presents the simulation results and analysis. Chapter 7 provides concluding remarks.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

When customers are located over an extensive geographical region, it may not be easy to provide efficient service support by having only one central inventory stocking point. Instead, many companies have multiple stocking points, each serving a specific region. These stocking points are, in turn, served by warehouses, each of which supports a group of stocking points that are usually in close proximity to one another. The warehouses themselves can be served by some distribution centers, etc. Such a hierarchy of stocking locations is termed a multi-echelon inventory system. When each stocking location has only one other location that serves as its supplier, at a higher level of the hierarchy, the multi-echelon inventory system is said to have an arborescent structure. A serial system is a special type of the arborescent system, in which every stocking point at a higher echelon has only one successor. Most consumer and industrial finished goods are distributed through multi-echelon inventory systems of one sort or another. Spare parts for office equipment, computer, automobiles, and military hardware are commonly provided through multi-echelon systems. Because of the practical importance of such systems, there has been a rapid increase of the literature on multi-echelon inventory models, most of which are on two echelons.

Figure 2.1 shows a special type of the two-echelon arborescent inventory system, called a one-warehouse multi-retailer distribution system. Demand is assumed to originate at the lowest echelon (stores) and is transmitted to the higher echelon (warehouse). Stocks flow in the opposite direction, from the outside supplier (e.g., factory) to the warehouse and then from the warehouse to the retail stores. If there is only one store at the retail level, then it becomes a serial distribution system (see Figure 2.2).

Guillermo and Zipkin (1994) have proved that for a serial system, a necessary condition for a stocking point at a higher echelon to hold stock is that it is cheaper to hold it there than at any downstream echelons. If there is no holding cost advantage to holding inventory at any particular stocking point, then all stock in a serial system should be positioned at the retail level near the customer. For a one-warehouse multi-retailer distribution system with deterministic demand, Roundy (1985) and Maxwell and Muckstadt (1985) have showed that it is desirable to immediately allocate all inventory to the stores. For a one-warehouse multi-retailer distribution system with stochastic demand, however, determining the proper positioning of system-wide inventories seems to be much harder. Despite of many years of research, the structure of an exact optimal solution has not been identified (Federgruen, 1993a). Several approximate models have been developed that lead to somewhat different conclusions regarding where inventory should be positioned in the distribution system. Our understanding of these different conclusions is still limited.

This literature-review chapter is organized as follows. In section 2.2, we cover Simpson's model (1958) and its derivatives. The common feature shared by these models is that they avoid considering the consequences of shortages at the warehouse. Simpson's model generates a so-called "all-or-nothing" solution, namely either positioning sufficient stocks at a specific location or nothing at all. Next, in Section 2.3, we review the models in which the warehouse follows the first-come-first-served allocation rule and when the warehouse stocks out, the delivery of the shortfalls is assumed to be delayed until the upstream location has sufficient stock. These models seem to indicate that virtually all inventory should be positioned at the retail level near the customer. In Section 2.4, we survey studies on holding inventory at the warehouse for better-informed allocations. In these models, the motive for holding inventory at the warehouse is to postpone allocation decisions until more information becomes available. These models and their empirical results seem to suggest that there is a benefit from holding some inventory at the warehouse but there are mixed results on whether or not the benefit is significant. We finish this Chapter with a discussion of the limitations of the models we reviewed and open issues that remain to be addressed.

2.2. KEY STUDIES THAT AVOID CONSIDERING CONSEQUENCES OF WAREHOUSE SHORTAGES

This section consists of three subsections. First, we review Simpson's model and its derivatives. Next, we discuss whether or not Simpson's model has provided analytical support for Brown's concept of "national safety stock" (Brown 1977). This concept has been known for some time and many practitioners usually take it for granted. We show that Simpon's model has provided no analytical support for the concept of the "national safety stock." In Section 2.2.3 we provide a brief discussion on the strength and the weakness of Simpson's model.

2.2.1 SIMPSON'S MODEL AND ITS DERIVATIVES

The earliest work on the positioning of system-wide inventories is that of Simpson (1958), who studied the location of "in-process" inventories in a serial production system (see Figure 2.3). "In-process" inventories are those inventories that separate one production stage from another. "In-process" inventories are also called buffer stocks. The control system Simpson considered is a "base-stock" system; it works as follows: When an order for an item is placed against any inventory, it is filled from the inventory, if the inventory is not zero. If the inventory is zero, the order is placed in a backorder file, to be filled when the item arrives. In any event, a manufacturing order is immediately placed with the most recent manufacturing operation, to produce a replacement item for the item already consumed. This is similar to the "kanban" aspect of JIT (Just-in-time) systems of today.

Simpson assumes that sufficient input material is not always immediately available to accomplish the desired release, except for the raw materials. Rather, between every pair of adjacent stages he specifies a "service time," which is a decision variable. He assumes that the upstream stage can always satisfy a request from the downstream stage within the service time; with this assumption, Simpson in effect avoids considering the consequences of inter-stage shortages. The justification for this assumption seems to be the supposition that the purpose of "in-process" inventories, or buffer stocks, is to protect against normal variability (i.e., the maximum reasonable demand); buffer stocks permit the system to function routinely in the face of normal variability. Buffer stocks should not be held for protection against abnormal or excessive variability; rather the organization maintains some slack capacities to respond to abnormal variability. Interestingly, Simpson's model produces an "all-or-nothing" solution: that is, between any two stages either there is no inventory (i.e., the service time promised by upstream stage is equal to the time for the upstream stage to replenish its inventory) or there is sufficient inventory to de-couple the two stages completely (i.e., the service time is equal to zero).

Inderfurth (1991) extends Simpson's model in the general serial and arborescent production/distribution networks. For the three-stage production system shown in Figure 2.4, Inderfurth finds that the optimal service time for stage 1 is zero and for stage 2 is equal to the time for stage 2 to replenish the inventory from stage 1. In other words, while sufficient inventory should be positioned at stage 1, no inventory should be held at stage 2. Note that stage 2 is an intermediate stocking location, which receives stock from its predecessor but also serves as a supplier to its downstream stage. This example shows that optimal positioning strategy may require no inventory positioned at the intermediate stocking location. Inderfurth (1991) also takes into account the impact of an end-item demand correlation on the positioning of the "in-process" inventories in the production system. For further extension and detailed discussion, see Inderfurth (1994a, 1995) and van Houtum et al (1996).

Graves and Willems (1996) extend Simpson's model to address the inventory positioning problem in assembly networks. An assembly network is a special type of multi-echelon inventory systems in which each stocking location has at most one downstream location as its successor (see Figure 2.5). Graves and Willems (1996) also discuss the idea for extending Simpson's model to general multi-echelon inventory systems.

Simpson's concept of "service time" provides an interesting modeling framework. However, as Graves (1988) points out, this model is not rigidly specified because we no longer have a clear description of how the system works when subjected to extreme or abnormal variability. Also note that Simpson's model can be applied only to multi-echelon inventory systems controlled by a "base-stock" system. With lot sizing, the buffer stock at the upstream stocking location could become negative. Then, the nice property associated with Simpson model's (i.e., "all-or-nothing" solution) collapses. Even for the distribution system controlled by "base-stock" system, there are open questions as well. For instance, there has been no published study dealing with the case in which a single stocking location quotes different service times to its downstream stocking locations.

2.2.2 BROWN'S CONCEPT OF "NATIONAL SAFETY STOCK"

Simpson's model seems to have had an impact on the positioning practice, not directly, but through the work of other researchers and business consultants in particular. Most notable is the work of Brown (1977), who advocated the concept of "national safety stock" for holding inventories at every upstream stocking location. Brown seems to suggest that Simpson's model has provided analytical support for the concept of "national safety stock." But a closer look shows that Brown may have misinterpreted the "all-or-nothing" solution generated by Simpson's model.

In his original article, Simpson summarized his main theorem as follows: "...for an optimal system, the service time for every non-vanishing inventory is zero" (Simpson, 1958, p. 868). Here, "non-vanishing" is a key word. Whether an inventory should be "vanishing" or "non-vanishing" is determined by Simpson's mathematical programming model. Simpson's main theorem has in no way suggested that every upstream stocking location should have "non-vanishing" inventory or zero service time. But Brown (1977) seems to believe otherwise.

"The linear programming model Simpson developed shows that if some intermediate investment is reduced to give more stock at the end of the pipeline next to the customer, service will of course go up in terms of the ability of the warehouse to cope with unforecasted demands. But when they have exhausted their safety stocks, the fact that they can't get what they need from the next echelon back results in a longer period of being out of stock, and on the average the total service will be worse," Brown wrote (Brown, 1977, p177). After introducing his concept of "national safety stock," Brown argued, "If the national safety stock were not there, then the satellite stocking locations could not be sure of getting replenished in the normal lead time—there might be a shortage at the source, which would increase the effective lead time. Even worse, the lead times would appear variable and unpredictable to the stocking locations. That would increase their need for safety stock to achieve any particular desired level of service." For theoretical justification, he explained, "Kenneth F. Simpson showed in his article 'In-Process Inventories' that to give a desirable level of service to the ultimate customer with the minimum inventory, all risks should be taken at the last stage of the distribution. At every intermediate echelon perfect service should be given" (Brown, 1977, pp. 348-349).

It seems that Brown has changed the "all-or-nothing" solution of Simpson's model to an "all" solution without giving any justification. As we pointed out in the previous section, Section 2.2.1, Inderfurth (1991) has demonstrated that the optimal solution generated by Simpson's model for a three-echelon production system (shown in Figure 2.4) requires that one of the intermediate stocking location (i.e., stage 2) hold no inventory at all. This counter example is enough for us to conclude that Brown's interpretation of Simpson's model is not accurate.

In Simpson's model, it is the service time that captures the interactions between stocking locations across echelons. By assuming zero service times for all intermediate stocking locations, Brown in effect has de-coupled a multi-echelon inventory system into individual stocking locations where safety stocks are determined independently. Indeed, Brown has provided a formula for safety stock calculation, which is supposed to work at every stocking location, regardless of whether it is intermediate stocking location at an upstream echelon or a location at the downstream echelon near the customer. The simplicity of Brown's formula has made it quite popular among practitioners for sometime and even today much commercial software for inventory management is still based on Brown's formula for the calculation of safety stocks directly or indirectly. In some cases, Brown's formula may turn out to be a good heuristic. But Brown's formula, particularly when it applied to calculating safety stocks required at upstream stocking locations, does not seem to have the analytical support of Simpson's model. While Nahmias (1997) has pointed that "it is not recommended to include independent safety stock at all levels of the system," (as suggested by Brown 1977), many other researchers still take Brown's formula for granted (e.g., Stenger and Cavinato 1979.)

2.2.3. DISCUSSION

Despite its limitations, Simpson's model provides a useful framework by which we can explain why some practical solutions to the inventory positioning problem, such as the concept of "national safety stock" advocated by Brown (1977), are actually without analytical support. Another well-know heuristic for determining the positioning of safety stocks in multi-echelon inventory systems is Miller's (1979) concept of "hedging." By comparing Miller's model with Simpson's, Graves (1988) has also concluded that Miller's concept of "hedging" does not seem to have analytical support.

The conventional wisdom about the location of safety stock suggests that safety stock is needed only at the echelon near the customer. Orlicky (1975) says of safety stock in a Material Requirement Planning (MRP) system: "Safety stock is properly applied only to inventory items subject to independent demand". This conventional wisdom may be described a "nothing" solution as compared to Brown's "all" solution for "national safety stock." Rosenfield and Pendrock (1980) provide a general discussion on these two extreme positioning strategies. It is of note that neither of these extreme strategies seems to have the analytical support of Simpson's model, which generally produces an "all-or-nothing" solution, rather than an "all" solution or a "nothing" solution. Indeed, Baker (1993) surveys the literature on the location of safety stocks in MRP systems, showing that the results are mixed (e.g., De Bodt and Van Wassenhove 1983, McClelland and Wagner 1988, Yano and Carlson, 1988). Vollmann, Berry, and Whybark (1997) indicate that safety stocks can be located anywhere in Distribution Requirement Planning (DRP) systems.

The literature reviewed in this section shows that the inventory-positioning problem is still not well understood. Many plausible qualitative arguments for or against holding inventory at upstream locations do not seem to have analytical support and, therefore, should not be taken for granted. Simpson's model (1958) has demonstrated that positioning inventory based on intuitive grounds could be costly. Systematic research on the positioning problem is very much needed.

2.3 KEY STUDIES IN WHICH THE WAREHOUSE FOLLOWS THE FIRST COME FIRST SERVED ALLOCATION RULE

Having commented on the well-know practical approaches to the positioning problem, in the remaining of this chapter, we would focus our attention on the research literature exclusively. Key studies reviewed in this section assume that the orders from the downstream echelon are filled on a first-come-first–served (FCFS) basis with the shortfalls being kept in backorder files to be satisfied as soon as stock becomes available. The consequence of the shortages is assumed to be an increase in the downstream location's replenishment lead time. Replenishment lead time is the elapsed time from the moment an order is placed until the moment the full quantity ordered is received. That definition is purposefully vague because of the possibility that portions of an order may be delivered at different times. The time needed to physically move stocks from one location to another is called transit lead time, which is equal to the replenishment lead time, provided that there is no shortage at the upstream location.

To facilitate the review, we need to define several terms that are commonly used in inventory theory. Inventory of a product that is physically available is on-hand stock. The quantities of product that have been ordered but not yet received are stock on-order. Stock on-hand minus backorders is the net stock. The average level of net stock just before an order arrives (Silver and Peterson, 1985) is safety stock, which can also be expressed as the difference between the reorder point and the expected demand (or request from the downstream locations) during the replenishment lead time. The net stock plus stock on-order is called inventory position, which represents the amount of inventory that is available to meet future demand without placing further orders.

The remainder of this section is organized as follows. We first review Hanssmann's (1959) model and other key studies on the positioning problem, then discuss why the models reviewed in this section consistently show that virtually all inventory should be positioned at the retail level near the customer.

2.3.1 HANSSMANN'S MODEL AND ITS DERIVATIVES

Hanssmann (1959) has considered a scenario very similar to that of Simpson (1958), but with some significant differences in assumptions. He studies a two-stage serial production system (one semi-finished product and one finished product) facing normally distributed customer demand for the finished products in a periodic-review environment. The control system Hanssmann considered is one in which orders are placed each period to bring inventory position at each stage to a predetermined target level each period. When an upstream stage has insufficient stock, the delivery of the shortfall is delayed until the upstream stage has sufficient stock. Although the length of this delay is a random variable, Hanssmann approximates it as a deterministic delay equal to its expected value. This deterministic delay from an upstream stage is added to the fixed transit lead time for the downstream stage. Hanssmann also assumes that customer demand is a function of the expected backorders at the finished product level. The object is to determine the target level of inventory at both semi-finished product and finished product stages to maximize expected system profit given a known contribution margin for each unit sold and marginal holding costs for inventory on hand at each stage. Although the model is only approximate in some respects, its analysis does point out the trade of relationship in positioning fixed amount of inventories in the multi-echelon production system. The less inventory at the upstream stage, the longer the replenishment lead time for the downstream stage will be.

Hanssmann's approach for modeling the consequences of the shortages at the upstream location is important because a vast literature on the positioning problem is based on and motivated by it. For a comparison between Simpson's and Hanssmann's model in general, see Graves (1988). For insight on the positioning problem offered by Hanssmann's model in particular, we refer to Schwarz (1981b).

Next, we review key studies that have followed Hanssmann's approach for modeling the consequences of the shortages at the warehouse. The most notable work was provided by Deuermeyer and Schwarz (1981). These two authors are the first researchers to consider adapting a continuous (Q, r) replenishment rule in the context of one-warehouse multi-retailer distribution systems. A continuous (Q, r) is implemented as follows: the inventory position is reviewed continuously; when it falls to or below reorder point r, an order for Q is placed. The demands at the retail stores form the independent Poisson process. The stores independently place orders on the warehouse according to their (Q, r) replenishment rules. The warehouse in turn replenishes stock from an outside supplier according to its own (Q, r) replenishment rule. The goal of the study is to determine approximate expressions for the expected service level as a function of that system's decision parameters.

It is difficult to analyze the process when the warehouse is out of stock. Following Hanssmann's approach, Deuermeyer and Schwarz assume that orders from the stores queue up at the warehouse and are filled on a first-come-first-served (FCFS) basis and the consequence of a warehouse shortage is an increase in the retailers' replenishment lead time. Deuermeyer and Schwarz (1981) further assume that if the warehouse is unable to fill a retailer order in full, the entire retailer order is backordered until it can be filled in full. No partial filling is permitted. Simulation tests showed a close match between observed simulated measures and those computed with the model, indicating the efficacy of the model.

Later, Schwarz, Deuermeyer, and Badinelli (1984) considered an extension of the

Deuermeyer and Schwarz (1981) model, which incorporated fill-rate optimization to determine where safety stock should be held in the system. The near-optimal solution found through a search shows that the warehouse should hold negative safety stock with very little on- hand inventory. Virtually all inventories are positioned at the retail level near the customer.

Badinelli and Schwarz (1988) also study the problem of minimizing expected backorders for the one-warehouse multi-retail distribution system previously studied by Deuermeyer and Schwarz (1981). Expected backorders, or time-weighted backorders, are the average backorders expressed as a fraction of an average period demand. Assuming identical retail stores, the authors formulate three optimization problems in which the goal is to choose safety stock at the stores and at the warehouse (1) to minimize system backorders (i.e., expected backorders across all the retail stores) subject to a budget constraint on the average system inventory, (2) to minimize average system inventory subject to a constraint on expected system backorders, or (3) to minimize system backorders subject to a constraint on system-wide safety stock. Consistent with the results of Schwarz, et al.(1984), Badinelli and Schwarz (1988) find that the motive for holding inventory at the warehouse is a very weak one and the average on-hand warehouse inventory for the system modeled should be close to zero.

While Schwarz, Deuermeyer, and Badinelli(1984) and Badinelli and Schwarz (1988) use different customer service criteria (i.e., fill-rate and expected backorders), they have found that the optimal strategy for positioning system-wide inventories remains unchanged: virtually all inventories should be held at the retail level. The results seem to indicate that the positioning of safety stocks in the distribution system is insensitive to changes in the service criterion used. This should not be a surprise because the FCFS sets allocation priorities according to the timing of the retailers' replenishment orders. The warehouse cannot change these allocation priorities regardless of what service criterion is used.

Rosenbaum (1981) describes another inventory control system which was developed to aid in determining the safety stock positioning problem in the Eastman Kodak Company. The warehouse orders from an outside supplier (the factory) using the order-up-to-level replenishment decision rule (S-1, S) on a periodic basis (e.g., weekly). The replenishment decision rule is implemented as follows: an order is placed to bring the inventory position to the predetermined order-up-to level S whenever the inventory position is below S. The stores order from the warehouse using a continuous (only approximately--i.e. daily) review (Q, r) replenishment decision rule. Rosenbaum assumes that when there is a shortage in the central warehouse, the retailer order can be split into two orders. The portion that can be filled immediately has a "normal" replenishment lead time. Only that portion that cannot be filled immediately has an extended "long" replenishment lead time.

Rosenbaum (1981) defines fill rate not only for retail stores, but also for central warehouses. He focus on the service-level relationship, rather than the location of safety stocks. Rosenbaum finds that there is no direct, computable relationship between the service levels used to calculate the system parameters and the actual service level the customers received. He reports the results of a field test showing that a substantial decrease in the central warehouse (Distribution Center) inventory (37%) accompanied by a smaller increase in total retailer store (Regional Distribution Center) inventory (11%), results in a net decrease in average total company inventory of the products tested. Since the inventory was shifted further into the filed, the results of Rosenbaum (1981)'s model to some degree are consistent with results of Schwarz et al. (1984) as well.

Ehrhardt, Schultz, and Wagner (1981) conduct the first study of periodic (s, S) replenishment rules in one-warehouse multi-retailer distribution system setting. Under an (s, S) replenishment rule, every period the inventory position is reviewed. An order is placed to bring the inventory position to the predetermined order-up-to-level (S) whenever the inventory status reaches at or below a reorder point (s). It is assumed that each store and the warehouse follows an (s, S) replenishment decision rule. Demands at the stores are assumed to be independent, identically distributed random variables with arbitrary probability distributions. Full backorder of excess demand is assumed at all stocking locations. The primary contribution of this study is the observation that if the stores follow identical periodic review (s, S) replenishment decision rules, the demand pattern at the warehouse will be correlated in time. Significant errors arise if this correlation is ignored. Because of the correlation, the optimal form of the replenishment policy at the warehouse is not a periodic review (s, S) decision rule, but that form is still adopted because of its widespread popularity in practice. The authors suggest that the warehouse replenishment policy be computed by a modification of the so-called power approximation (Ehrhardt, 1979), which takes the period-to-period correlation into account.

Schneider and Ringuest (1990) consider the same model studied by Ehrhardt et al. (1981), but focus their attention on the inventory positioning problem. Like Rosenbaum (1981), Schneider and Ringuest define a service level for the warehouse. Specifically, theirservice level is defined as the fraction of periods for which the retailerdemand is completely satisfied by the warehouse. However, no service level is defined for the retail stores. The authors construct an approximation cost function for the total long-run cost per period. The theoretical feature of their model is that the mean and variance of the demand distribution during the replenishment lead time from warehouse to stores is computed as a function of the service level provided by the warehouse. The inventory positioning problem thus becomes a problem of correctly specifying the service level for the warehouse. The replenishment lead time is approximated by a negative binomial distribution. Power approximations are then used to provide estimates for the periodic (s, S) decision rule parameter values at the retail stores and at the warehouse.

Schneider et al. (1995) conduct empirical studies of the procedure they developed earlier (Schneider and Ringuest, 1990) and conclude that the total cost is quite sensitive to the warehouse service level. The interesting result is that if the service level at the warehouse tends to be lower than optimal, the increase in cost can be balanced by the adjustment of the retailers' replenishment lead time parameters. For the optimal and higher than optimal service level, the improvement achieved by the adjustment is marginal. If the lead time from the warehouse to the store is adjusted in an optimal manner, the cost function of the system tends to be steep as service level goes up higher than the optimal level but is rather flat as α-service level becomes lower than the optimal level. The implication is that holding more than the optimal amount of inventory at the warehouse causes costs to go up substantially. Yet, there is little penalty for holding less than an optimal amount of inventory at the warehouse provided that the replenishment lead time parameters are adjusted optimally.

Newsboy-style results for the backorder minimization problem in a one-warehouse multi-retailer distribution system is provided by Rogers and Tsubakitani (1991). They model the consequence of warehouse shortages as an increase in retailer's replenishment lead time. Rogers and Tsubakitani find that their model prescribes relatively little inventory and relatively large expected backorders at the warehouse. One of the interesting results from the Rogers and Tsubakitani (1991) model is that increased uncertainty of customer demand, measured by the coefficient of variation (i.e., the standard deviation of the demand divided by the mean of the demand) is offset by more inventory being kept at the warehouse.

More recently, Graves (1996) develops an interesting multi-echelon inventory model in which each stocking location orders at preset times according to the order-up-to replenishment rule, and the upstream location follows a unique allocation decision rule, referred to as "virtual allocation." Specifically, the warehouse commits its inventory to the stores according to the FCFS on a virtual basis, using demand rather than retailer orders. The warehouse is assumed to have on-line information about the demand at the retail level. However, the committed stock can only be delivered to the retail stores according to a fixed schedule. Graves finds that all safety stocks should be positioned at the retail stores.

The models reviewed so far seem to be consistent with the results reported by Schwarz, Deuermeyer, and Badinelli (1984). That is, if the warehouse follows the FCFS allocation rule, there is little incentive to hold inventory at the warehouse. Virtually all inventories should be positioned at the retail level near the customer. To our best knowledge, the only published study that seems to suggest otherwise is a simulation study conducted by Chakravarty and Shtub (1986).

Chakravarty and Shtub (1986) simulate a system similar to the one described by Rosenbaum (1981). Two experimental parameters are used, labeled TSSF and CWSSF. TSSF is defined as the proportion of a maximum safety stock quantity permitted in an experiment. CWSSF is the proportion of the system-wide safety stock retained centrally. Each of these two experimental variables takes on a value of between 0 and 1. The maximum amount of system-wide safety stock for any experiment is labeled as the base safety stock (BSS). BSS is calculated as the sum of the safety stock at each retail store required for some predetermined level of fill-rate. Therefore, the amount of safety stock in the system for any given value of TSSF is TSSF*BSS. The quantity of safety stock positioned at the warehouse is TSSF*BSS*CWSSF. If the retail stores are identical, the safety stock at each store is given as TSSF*BSS*(1-CWSSF)/N, where N is the number of retail stores. A reference system fill-rate is specified by setting TSSF to equal 1 and CWSSF to equal 0. In this case, no safety stock is retained centrally.

Chakravarty and Shtub (1986) find that fill-rates higher than the reference can be achieved by holding some safety stock at the warehouse. It is difficult to duplicate Charkravarty and Shtubs simulation results because they did not report the demand distribution parameters they used. However, we could not conclude that the work of Charkravarty and Shtubs is inconsistent with the results of Schwarz et al. (1984), because Charkravarty and Shtubs have not considered a scenario in which the warehouse holds negative safety stocks.

2.3.2 DISCUSSION

The results of Schwarz, Deuermeyer, and Badinelli (1984) have come as a surprise to many researchers in the field. For instance, Nahamias and Smith (1993) comment on the results: "this finding seems contrary to the 'portfolio effect' observed by others for slightly different types of systems (for example, Eppen, 1979)."

McGavin et al. (1993) speculate that the fixed retailer order-quantities and the FCFS allocation rule in Deuermeyer and Schwarz (1981) model do not allow the warehouse to balance retail stock levels and, therefore, preclude any potential benefit from holding inventory at the warehouse.

Indeed, we believe that the FCFS is the main reason why there is little incentive to hold inventory at the warehouse. When the warehouse follows the FCFS allocation rule, allocation priority is always given to the oldest outstanding orders. The FCFS allocation scheme is not optimal in that it does not account for the relative need of the stores for inventory. For instance, it may be desirable not to commit an inventory unit which had been destined for one store with ample safety stock, but rather to redirect it to another store with a more critical need for replenishments. Unfortunately, the FCFS precludes the possibility for such a maneuver.

Note that under the FCFS studied by Schwarz, Deuermeyer, and Badinelli (1984) and others, it was the timing of the retailers' replenishment orders that created allocation priorities. Under the "virtual allocation" proposed by Graves (1996), the FCFS allocation priorities were created even earlier as the demand occurred, rather than as the retailer's order was triggered. In both cases, once the allocation priorities were created, no changes could be made. Since the FCFS allocation scheme prevents the warehouse from reacting to the changing conditions at the retail level, not surprisingly, there is little incentive to hold inventory at the warehouse.

Based on the literature reviewed in this section, we have excluded that the FCFS allocation rule is not optimal. In this study, therefore, we consider only these control systems that allow the warehouse to use other allocation rules.

2.4 KEY STUDIES ON HOLDING INVENTORY AT THE WAREHOUSE FOR MORE INFORMED ALLOCATIONS

The studies reviewed in this section all assume that it is the warehouse that sets the allocation priorities. With this assumption, the warehouse could delay committing stock to the stores until it physically arrives at the warehouse, and then further postpone allocation decision by holding some inventory at the warehouse for sending it out later to "rebalance" the retailer inventories that may have become "unbalanced."

This section is organized as follows. First, we introduce the concept of "risk pooling over the supplier lead time" and the "depot effect" for holding inventory at the warehouse for more informed allocations. Second, we review the key studies on "depot effect." We end this section with comments on the key assumptions, open issues, and the conjectures.

2.4.1 EPPEN AND SCHRAGE'S MODEL AND ITS EXTENSIONS

Eppen and Schrage (1981) examine a one-depot multi-retailer distribution system in which the depot receives inventory from an outside supplier each replenishment cycle (i.e., the depot can issue orders to the outside supplier in a fixed time interval, called "replenishment cycle.") and immediately allocates all the inventory to the retail stores. Eppen and Schrage demonstrate that backorder costs are reduced in such system, if the depot acts as a centralized ordering facility and delays the assignment of stock to the retailers from the time at which an order is placed with the outside supplier to the time at which the depot actually receives the stock. This phenomenon is referred to as "risk pooling over the supplier lead time." Eppen and Schrage also identify (but do not investigate) a second possible advantage of postponing the allocation decision: the depot can hold inventory and allocate it to the stores between the times at which orders arrive at the depot. They study the question: Should stock be immediately allocated to the retail stores upon receipt at the depot or should some depot inventory be held in reserve, to be distributed to the retail stores later in the cycle? Eppen and Schrage coin the phrase "depot effect" for any possible advantages for holding some inventory at the depot.

The work of Eppen and Schrage (1981) is important in that it provides a new framework for making positioning trade-off decisions. This framework shifts our attention away from focusing on material flows exclusively to both inventory flow and information flow. Holding inventory at the depot is now considered as an effort to postpone the allocation decision more so that the depot can observe more materialized demands and make more informed allocation decisions. The control system Eppen and Schrage (1981) considered is quite different from the ones we reviewed in the previous section, Section 2.2. Their system assumes that the retail stores do not place orders on the depot. When and how much inventory should be allocateed to the retail stores is determined by the depot, which is assumed to be able access continuously or periodically updated information about the retailers' inventory positions. Without orders from the retail stores, depot shortages seem to have been avoided. But this is not true. The allocation rule Eppen and Schrage (1981) considered is one that equalizes the probabilities of stockout at the retail stores, called the balancing allocation rule. The implementation of this allocation rule critically depends on whether there is enough stock available at the depot.

The uncertainty as to the availability of stock at the depot creates interdependence between the retail stores and their supplier, the depot, which in turn induces interdependence among the retail stores themselves. Such interdependence could tremendously complicate inventory analysis and easily make Eppen and Schrage's model analytically intractable. To avoid dealing with the consequences of depot shortages, Eppen and Schrage (1981) made a bold assumption: in each allocation period, the depot receives sufficient goods from the outside supplier that each retail store can be allocated goods in sufficient quantities to ensure that the probability of stockout is the same at all retail stores.

Eppen and Schrage coin the phrase "Allocation Assumption," and established conditions under which this assumption is likely to hold: when the coefficient of variation of demand (i.e., the standard deviation of the demand divided by the mean of the demand) at the retail stores is moderate, the fixed cost of ordering from the depot is high, and the number of retail stores is small.

With the Allocation Assumption, Eppen and Schrage(1981) build a dynamic model for one-depot multi-retailer distribution systems. Assuming the depot holds no inventory, Eppen and Schrage demonstrate that if the transit lead times to the depot (L_w) and to the retail stores (L_r) are not negligible, and if demand at each retail store follows a normal distribution, then the total system inventory on hand plus the stock on order is greater for a decentralized system, where the retail stores order independently on their own, than for a centralized system, where the individual retail stores are treated as one single stocking site. The one-depot multi-retailer distribution system lies between the two. Specifically, they find that a part of the total inventory on hand plus the stock on order for the one-depot multi-retailer system appears to consist of an (L_w) period centralized system and an (L_r+1) period decentralized system. The authors refer to the reduction in total system inventory on hand and on order from adopting the central warehouse system rather than the decentralized system as "risk-pooling effect over the supplier lead time."

The work of Eppen and Schrage (1981) on the "risk-pooling effect over the supplier lead time" is extended by Schwarz (1989). Schwarz considers two systems: one in which the external suppler ships directly to the retail stores (i.e., Eppen and Schrage's decentralized system) and one in which the depot receives the incoming stock and then redistributes the stock to the retail stores. Similar to Eppen and Schrage's analysis, Schwarz assumes that the depot holds no inventory. The key feature of Schwarz's work is that it assumes that the system with the depot has additional lead times--receiving lead time at the depot, which increases the lead time from the supplier to the depot, and repackaging lead time added to the lead time from the depot to the retail stores. The question Schwarz addressed is: What would be the additional lead time when a depot that holds no inventory is included and the system-wide fill rate and safety stock are the same? The primary result is that the benefit from adding the central depot to "pool risk over the supplier lead time" depends critically on these additional lead times. Schwarz's analysis suggests that there is a substantial benefit to including the depot between the outside supplier and the retail stores, if the depot can be located near enough to the retail stores (i.e., if the lead time from the depot to the retail stores is small).

Note that the one-depot multi-retailer distribution system studied by Eppen and Schrage (1981) is identical to the one-warehouse and multi-retailer distribution system we described earlier. Since the difference is one of terminology only, in the reminder of this chapter we will continue to use our term "the one-warehouse multi-retailer distribution system." But we will also use the special phrase "depot effect," coined by Eppen and Schrage (1981) for the benefit from holding some inventory at the warehouse because of its wide usage in the inventory literature.

Several researchers have investigated the optimality of the proposed balancing allocation rule appearing in the work of Eppen and Schrage (1981). Zipkin (1984) proves that for an m-period multi-retailer model, the allocation decision rule which maximizes a proposed measure of balance in every period, minimizes an approximation of a dynamic program cost function describing the multi-period newsboy problem. Federgruen and Zipkin (1984c) define the concept of "inventory balance" as the situation when all retailer inventory positions are at the same fractile of demand (i.e., at the same normalized level). To obtain the normalized inventory for a retailer in a given time interval requires dividing the difference between its inventory position and the mean interval demand by the standard deviation of the interval demand.

The optimality of a decision rule that equalizes the normalized inventory levels among retailers can be easily verified from the Karush-Kuhn-Tucker condition. When the normalized inventory levels are exactly equalized, the inventories are said to be "fully balanced." Any deviation from the fully balanced inventories is referred to as "imbalance." More recently, McGavin, Schwarz and Ward (1992) show that for multiple-identical retailers, this balancing allocation minimizes expected lost sales and backorders over multiple independent time intervals.

There are three alternative allocation decision rules that can be used to balance retailer inventories. First, the myopic balancing allocation decision rule is one possible choice. Stock is allocated so as to ensure that the probability of stockout in the very first period in which the allocations have an impact (i.e., in the period when the shipments arrive at their destinations) is the same at all retail stores. Second, let us consider the following situation: the central warehouse is replenishing its stock on a fixed schedule, say every m periods (referred to as warehouse replenishment cycle). If all incoming stock must be immediately allocated (i.e., the warehouse holds no inventory), it is generally better for the warehouse to allocate its stock so as to equalize the probability of stockout at all retail stores over the entire m periods starting from the very first period in which the allocations have an impact, not just in the period when the allocation arrives at their destinations. Such an allocation decision rule is called the cycle balancing allocation decision rule. Third worthy choice is to equalize the probability of stockout at all retail stores in selected period(s) in which the allocations have an impact (e.g., the period[s] in which the imbalance is most likely to occur), as opposed to the entire m periods (cycle balancing allocations) or the very first such period (myopic balancing allocations).

Several researchers have also tried to relax the "Allocation Assumption." For instance, Jackson (1988) proposes the so-called "run-out allocation rule." The idea is to implement the myopic balancing allocation decision rule to the extent that the available stock at the central warehouse would allow. In case of identical retail stores, the "run-out allocation rule" becomes the "ax-min allocation rule;" that is, the warehouse would allocate its available stocks so as to the retail stores to maximize the minimum inventory position at the retail level. It is easily shown that for identical retail stores the "max-min allocation rule" is equivalent to equalizing the inventory positions at the retail stores as much as possible. So, for this study the "max-min allocation rule" is used as the warehouse's allocation rule.

2.4.2 KEY STUDIES ON THE "DEPOT EFFECT"

The Eppen and Schrage (1981) model is extended by Federgruen and Zipkin (1984a) in a variety of significant ways (e.g., a finite horizon, other-than-normal demand distributions, and non-identical retailers). Nevertheless, Federgruen and Zipkin (1994a) still assume that the warehouse holds no inventory. One of the interesting results is that they show empirically that if the system starts with nearly balanced inventories at the retail stores and if for the remainder of the warehouse replenishment cycle the differences among these normalized inventories are generated by the customer demand process only, then the cost effects of "imbalances" at the retail level are quite small. Note that holding inventory at the warehouse is for the ability to reacting to the "imbalances" at the retail level. Since the cost effects of "imbalances" at the retail level are quite small in the case of no central stock at the warehouse, they must be small in the system with central stock as well. In other words, the "depot effect" must be a weak one.

Federgruen and Zipkin (1984a)'s approximations are shown to be very good in the case of the linear order cost, "low" demand variance, and a short warehouse replenishment cycle, but they deteriorate significantly when the warehouse replenishment cycle includes many periods, a fixed order cost at the warehouse is allowed, or the coefficients of variation of demand at the retail stores are large and unequal. Federgruen and Zipkin (1984b) argue that the myopic balancing allocation decision rule performs, if anything, even better in the system with central stock (and equal coefficients of variation). However, with significant unequal coefficients of variation, alternative allocation decision rules are needed.

The "imbalance" of stocks at the retail level has been examined by Jonsson and Silver (1987a). They show the desirability of redistribution during the warehouse replenishment cycle to correct the imbalance of stocks at the retail level. The redistribution is accomplished through making lateral transshipment between the retail stores. Lateral transshipment is just another way in which the distribution system can react to the changing conditions at the retail level. Jonsson and Silver's computational tests show that with a considerably reduced inventory investment, a system with redistribution can provide the same level of customer service as a system without redistribution. The study is important because it demonstrates that the "imbalance" at the retail level could have significant impact on the system performance. The benefit of the reduction in inventory investment, however, must be balanced with the cost for making the lateral transshipment. In this study we do not investigate the effect of transshipment.

Later, Jonsson and Silver (1987b) study the "depot effect" as identified by Eppen and Schrage (1981). Specifically, Jonnson and Silver consider the effect of holding some of the system inventory at the warehouse to allocate to the retail stores exactly one period before the start of the next warehouse replenishment cycle (which consists of H periods). Jonsson and Silver denote the system-wide inventory and the inventory positioned at the warehouse by Io and Ic respectively and assume that at the beginning of the cycle, called time 0, each of N retail stores receives (Io-Ic)/N stock in the initial allocation. Ic will be allocated to the retail stores in period H-1 to maximize the lowest inventory level at the retail stores (i.e., maxi-min allocation rule). The transit lead times from the warehouse to the retail stores are assumed to be zeroes. With these assumptions, Jonsson and Silver (1987b) develop a simple procedure for searching over Ic to minimize the expected unit shortages.

To quantify the "depot effect," Jonsson and Silver (1987b) compare the performance for holding some inventory at the warehouse with that for a "ship-all" and an "extreme push" positioning. The "ship-all" allocates all available system inventory to the retail stores at the beginning of the cycle; then no further shipment of units between locations is carried out until the next warehouse replenishment arrives. Following the "extreme push," all available system inventory is allocated to the retail stores as in the "ship-all." However, at the end of the period H-1, the inventories at the retail stores are completely balanced by carrying out transshipment among the retail stores. The authors define service level as 1-(expected shortages in the last two periods of the warehouse replenishment cycle). For a given service level, obviously, the "ship-all" requires the largest amount of system inventory while the "extreme push" requires the least. They evaluate the performance of these three strategies using the same 48 numerical examples for each positioning strategy. The results show that holding some inventory at the warehouse achieves an average 64% of the inventory reduction achieved by the "extreme push" relative to the "ship-all." Thus, they conclude that a substantial portion of the benefit of complete redistribution as shown by Jonsson and Silver (1987a) can be achieved by holding some inventory at the warehouse. Specifically, they show that on average approximately 6% reduction in inventory can be achieved by holding approximately 5% of the system-wide inventory at the warehouse in comparison with the "ship-all." This is the first empirical evidence suggesting that the "depot effect" can be significant.

However, the work of Jonsson and Silver (1987b) has several limitations. First, they examine only a single warehouse replenishment cycle. While single cycle models sometimes can be used in a dynamic environment, they are often inadequate for handling system dynamics. Second, by assuming that the system starts with zero inventory, the authors in effect have invoked “Allocation Assumption” at the beginning of the warehouse replenishment. Third, they ignore lead time. And last, but not least, they directly compare the system performance of holding some inventory at the warehouse for making secondary allocation with the that of the "ship-all" without taking into account the difference in the shipment cost to the retail stores.

To investigate the "depot effect," Jackson (1988) has developed a well-structured approximate model. Similar to the Jonsson and Silver (1987b)'s work, Jackson (1988) has invoked Eppen and Schrage's "Allocation Assumption" at the beginning of the warehouse replenishment cycle. The system is assumed to start at time 0 with the net inventory at each retail store below S; that is, the probability of having inventory "imbalance" at the beginning of the warehouse replenishment cycle (which consists of m periods) is negligible. At the beginning of the warehouse replenishment cycle, the warehouse makes its initial allocation by bringing the inventory position at each store to level S. In subsequent periods of the cycle, the warehouse makes shipments to each retail store equal to the demand observed at that location in the previous period until the warehouse runs out of stock. In the run-out period, the warehouse follows the max-min Allocation Rule as discussed earlier.

Assuming general demand distribution, Jackson (1988) formulates an exact cost function and obtains a simple approximation when the distribution of demand at the retail stores is Poisson or normal. His approximation gives good results, but deteriorates as the number of retail stores increases. In Jackson's model, S is the only decision variable.It determines how much inventory on average will be positioned at the central warehouse. He tests 16 scenarios for independent cycles of four periods. Only 2 of the 16 scenarios have positive lead time. The others are zero. He compares the best ship-up-to strategy in each scenario with a "ship-all" strategy with the identical total net inventory in the system at the beginning of the cycle. The results show that considerable improvements can be achieved by positioning some inventory at the warehouse in comparison with the "ship-all" strategies; that is, the "depot effect" is significant. However, Jackson reports that the objective function is quite flat in the neighborhood of the optimum, so it is difficult to estimate the true optimum with precision.

Jackson notices that the observed "depot effect" almost tripled when the warehouse changed its allocation rule from the largest-order-first to max-min allocation rule. Thus, he concludes that the allocation rule used is critical for a distribution system to achieve the benefit of "depot effect." However, it is of note that Jackson's investigation was conducted with respect to independent warehouse replenishment cycles. Consequently, the change in the allocation rule used would have no effect on the performance of the "Ship-all." Being unable to observe the changes in the performance of the "Ship-all" in response to the changes in the allocation rule indicates that the single warehouse-replenishment-cycle model could be misleading.

Similar to the work of Jonsson and Silver (1987b), Jackson (1988) also ignores the shipment frequencies. He directly compares the ship-up-to-S strategy, which may require up to 4 shipments per warehouse replenishment cycle, with the "ship-all" that needs only one shipment per warehouse replenishment cycle. In an extension of Jackson's (1988) work, Jackson and Muckstadt (1989) show that the strategy of positioning some inventory at the central warehouse appears to be attractive, even though there are only two retail stores in the distribution system.

Based on limited results from his own study and Jonsson and Silver's (1987b) work, Jackson (1988) conjectures that the benefit of the "depot effect" increases with the coefficient of variation (i.e., the standard deviation of the demand divided by the mean of the demand). Having observed data from industry on B-and C-type items (Silver and Peterson, 1985) exhibiting a demand pattern with standard deviation-to-mean ratio (i.e., the coefficient of variation, of two, three and higher), Jackson (1988) has predicted that the strategy of having some inventory positioned at the warehouse, such as the ship-up-to-S strategy, will become increasingly attractive for these high coefficient of variation items.

Acknowledging that his conclusion is in sharp contrast to that of Schwarz et al. (1984), Jackson (1988) writes, " It is beyond the scope of this paper to reconcile the two sets of findings, except to point out that the two systems being studied differ greatly in the order and allocation rules employed, in the timing of decisions, in the information (local vs. global) upon which the decisions are based, and in the parameters used for empirical study."

More recent study on the "depot effect" is that of McGavin, Schwarz and Ward (1993). These three researchers characterize the allocation rule for the warehouse by a set of four decision parameters: (1) the number of withdrawals from warehouse stock; (2) the times between successive withdrawals, which divide the warehouse replenishment cycle into intervals; (3) the quantity of stock to be with drawn from the warehouse for each withdrawal; and (4) the division of withdrawn stock among the retail stores, for each withdrawal. This framework includes the allocation decision rule used by Jonsson and Silver (1987b), and Jackson (1988) as special cases. McGavin, Schwarz and Ward (1993) investigate two withdrawals. They later compare their results with those of Jackson (1988) where the number of withdrawals can be up to four. Instead of equally dividing the warehouse replenishment cycle into periods, McGavin, Schwarz and Ward (1993) try to optimally divide the cycle into two withdrawal intervals and set a quantity for each withdrawal. The partitioning scheme is the max-min allocation rule.

Assuming that customer demand is a Gamma process and shortages at the retail level are lost sales, McGavin, Schwarz and Ward (1993) construct a model with an infinite number of stores. Based on this model, they develop two heuristic allocation decision rules. Their "infinite-retailer heuristic" seems to perform well. The results indicate that the scenario with very high demand uncertainty (i.e., the large coefficient of variation) and high service levels usually displays significant "depot effect." However, the infinite retailer model is not a good predictor of the lost sales per retailer incurred in a finite-retailer scenario: it consistently under-predicts the lost sales per retailer. Nevertheless, McGavin, Schwarz and Ward (1993) conclude that policies with two "well-chosen" withdrawals can provide "depot effect" benefits comparable to four equal interval withdrawals as observed by Jackson (1988). Since the costs to operate the distribution system are likely to increase in the number of warehouse withdrawals, the allocation rule with a few "well-chosen" withdrawals, on an overall cost basis, would out-perform equal-interval policies with more withdrawals, even though more withdrawals have fewer lost sales per retailer.

Recently, Nahmias and Smith (1994) consider a one-warehouse multi-retailer distribution system where the retail stores follow an order-up-to-level S in the period in which they need to replenish their inventories from the warehouse. "Allocation Assumption" holds except for the last period of the warehouse replenishment cycle. In the last period of the cycle, each unit of excess demand from the retail stores is special-ordered with probability (1-u) and lost sale with probability u to avoid the run-out allocation decision. Customer demand is assumed to follow a negative binomial distribution. Lead times are zeroes. The total retailer demand per period as observed at the warehouse is assumed to be normally distributed. Under these assumptions, Nahmias and Smith (1994) build a model that yields close form formulas for the total expected cost per cycle at both echelons.

Nahmias and Smith (1994) find that "positioning inventory at the central warehouse may or may not achieve a significant savings depending on the item. Similarly, depending upon the item, an additional store replenishment between the warehouse replenishments may or may not have a significant impact." The benefit of positioning inventory at the warehouse depends on item characteristics. Specifically, items with low target service levels (i.e., low inventory investment) at the retail stores derive the greatest benefit from holding inventory at the warehouse. It is of note that this is inconsistent with McGavin, Schwarz, and Ward (1993), who find that the high service levels usually display significant benefits. Increasing the frequency of shipment from the warehouse to the retail stores produces a significant reduction in total cost for certain types of items. Items with high optimal target service levels (hence large safety stocks) at the retail stores benefit most, since more frequent shipments allow their safety stock to be reduced.

2.4.3 DISCUSSION

The key studies reviewed in this section seem to suggest that a "depot effect" could be significant if the warehouse follows the balancing allocation rule. However, the model built by Federgruen and Zipkin (1984a) suggests that the "depot effect" is a weak one. The conclusions of recent studies seem to be more conservative, emphasizing that the "depot effect" is "scenario specific" (McGavin et al., 1993) or "depends on the item characteristics" (Nahmias and Smith, 1994).

It is of note that published studies on the "depot effect" often compare the performance of the best positioning strategies with that of the "ship-all" without taking into account their difference in shipment frequencies. Clearly, whether or not the observed benefits of holding some inventory at the warehouse are sufficient to outweigh the additional fixed shipment costs depends on the cost parameters involved. However, for distribution managers, it is still difficult to make the positioning trade-off decisions because many of the published studies have not reported, or even recorded the shipment frequency or fixed shipment costs required to achieve that reported "depot effect" benefit.

Also note that individual studies are typically conducted with respect to a single inventory control system. While various inventory control systems have been used for studying the inventory positioning problem, the results of these individual studies are hard to compare because they differ not only greatly in the inventory control system used but also in the parameters used for empirical study. The majority of the studies reviewed in this section assume zero lead times and model system behaviors limited to a single warehouse replenishment cycle, rather than in a truly dynamic environment. The range of the coefficient of variation of demand investigated has been very limited, typically no greater than 1. Jackson (1988) conjectures that the benefit from holding some inventory at the warehouse becomes increasingly significant as the coefficient of variation of demand increases to levels of two, three, and higher (Jackson 1988). This conjecture, which has been widely used to support the argument for holding inventory at the warehouse, has never been tested.

The studies reviewed in the previous section, Section 2.3, seem to be fairly consistent, showing that when the warehouse follows the FCFS allocation rule, virtually all inventory should be positioned near the customer. The results reported in this section, however, do not appear to be very consistent. For example, McGavin, Schwarz, and Ward (1993) observe that the "depot effect" becomes significant at high target service level. On the other hand, Nahmias and Smith (1994) report that items with low target service levels (i.e., low inventory investment) at the retail stores derive the greatest benefit from holding inventory at the warehouse. More research needs to be down before we can reconcile these seemly conflicting observations.

These concerns motivated us to develop new approaches to address the inventory positioning problem. These new approaches are highlighted in the next chapter, Chapter 3.

2.6 REMARKS

In this chapter, we have gone through some the literature that are relevant to the inventory positioning problem. Four conclusions can be drawn from this literature review. First, at this stage, there are no normative results on the proper positioning of inventories in a one-warehouse multi-retailer distribution system subject to stochastic demand. Second, the concept of "national safety stock" for positioning inventory at every intermediate stocking location does not seem to have the analytical support as Brown (1977) has suggested and, therefore, should not be taken for granted. Third, when the warehouse follows the FCFS allocation rule, the literature seems to indicate that virtually all inventory should be positioned at the retail level near the customer. Fourth, the idea of holding inventory at the warehouse for more informed allocations seems to be justifiable. There seem to be some evidence that the "depot effect" exists. But at this stage, it is difficult, if not impossible, to draw general conclusions about how significant it is or could be. There are open issues and untested conjectures. To gain clear insight, more research is needed.

Before ending this chapter, some comments on the research methodologies seem to be warranted. Within the literature we reviewed, one differentiating feature seems to be the treatment of shortages at the upstream echelon, or more precisely, the treatment of the uncertainty as to the availability of stocks at the upstream echelon. The models reviewed in this chapter seem to follow three different approaches: to avoid the consequences of warehouse shortages, to assume the consequences of warehouse shortages is a delay in the stores' replenishment lead time, or simply to invoke the Allocation Assumption.

The consequences of the shortages at an upstream echelon are difficult to predict and evaluate (Schwarz, 1981a). For example, in a one-warehouse multi-retailer inventory system, the consequences of a shortage at the warehouse may be a shortage at the retail level, or nothing at all, depending on the stores' inventory status. If we consider positioning inventory as an effort to strike a balance between shortages at the retail level and at the warehouse, then it is the difficulty in assessing the consequences of the warehouse shortages that makes the positioning problem an unusually challenging one for theoreticians and management alike.

About four decades ago, Clark and Scarf (1960) built the first multi-echelon inventory model by constructing a penalty cost function for the upstream echelon for not being able to provide adequate stock to the downstream echelon. This classical work has stimulated considerable interest in multi-echelon inventory models. Yet, various modeling approaches that have appeared in the literature seem to be focused on how to avoid the consequences of the shortages at the upstream echelon. The typical approach is to build a model for a single location system first, and then use this model as a "building block" for constructing a model for a multi-echelon inventory system (Graves 1988, Federgruen, 1993a). Occasionally, researchers may allow shortages to occur at the upstream echelon, but assume that the shortfall is filled by a special order from an outside supplier with a known cost (e.g., see Nahmias and Smith, 1994). An outside supplier is assumed to be completely reliable, an assumption commonly used, but rarely discussed in the literature. The difficulty involved in assessing the consequences of shortages at the upstream echelon is avoided by assuming "a known cost."

Whether various modeling approaches to avoid the consequences of shortages at the upstream echelon will succeed and be productive remains to be seen. Despite many years of research, even for a one-warehouse multi-retailer distribution system, the structure of an exact optimal control has not been identified. More importantly, it is clear by now that this structure is exceedingly complex. A fully optimal control system is therefore unattractive even if it could be computed efficiently (Federgruen, 1993a). In the last 15 years, we have experienced rapid progress towards the development of approximate models for multi-echelon inventory systems. More recently, Chen and Zheng (1994a, 1994b) obtained clever and tractable lower bounds for one-warehouse multi-retailer distribution systems with stochastic demand. Yet, numerical result show that the gap between the lower bound and the "optimal" cost is small when there is only one store, but widens as number of stores increases.

Numerous surveys of multi-echelon inventory theory exist. Earlier work has been summarized in A. J. Clark's well-known informal survey on multi-echelon inventory theory (Clark, 1972). Wagner (1974 & 1980) provides excellent research portfolios for inventory management. Porteus (1990) gives a comprehensive review on stochastic inventory theory. Graves (1988) furnishes a survey on safety stocks in manufacturing systems. Nahmias and Smith (1993) cover mathematical models of retailer inventory systems. For multi-echelon inventory systems with stochastic demand, Federgruen (1993a, 1993b) provides a review on centralized planning models and Axsater (1993) surveys continuous review models. For MRP and DRP systems, see Orlicky (1975), Baker (1993), and Vollmann, Berry, and Whybark (1997). For recent advances in supply chain management, see Lee (1997), Lee, Padmanabhan and Whang (1997), Fisher M.L., J.H. Hammond, W.R. Obermeyer, A. Raman (1994), Davis (1993), and Lee and Billington (1992).

CHAPTER 3

METHODOLOGY

Determining where inventory should be positioned in a distribution system with stochastic demand is a very complex problem. Without restrictive assumptions, analytical approaches to the problem often turn out to be intractable. The use of computer simulation to evaluate alternative positioning strategies seems to be the best way forward currently for most practical situations.

The purpose of this chapter is three-fold: First, we discuss the pros and cons associated with analytical approaches and the use of computer simulation. We argue that both analytical approaches and the use of simulation are needed for studying the inventory positioning problem. However, as our starting point, we selected simulation as our primary research tool. Next, we outline the new approaches developed in this dissertation study. Specially, we treat shortages at the retail stores and at the warehouse differently: while shortages at the retail stores are backordered, no backorder is allowed at the warehouse. To ensure comparability for the alternative positioning strategies simulated, we carefully controlled the frequency of shipments. We also treated the inventory control system as an experimental factor so as to be able to compare the best positioning strategies across different inventory control systems. These new approaches make this study quite unique in the context of existing research literature. Finally, we preview our simulation model and explain how we simulate and characterize alternative positioning strategies.

3.1 INTRODUCTION

Determining where inventory should be positioned in a one-warehouse multi-retailer distribution system with stochastic demand is an unusually challenging problem both in theory and in practice. Traditionally, researchers approach the positioning problem by building a simplified analytical model with restrictive assumptions and then use simulation to demonstrate the efficacy of their model. It is possible to gain some theoretical insight from such exercises. Unfortunately, these simplified analytical models often do not resemble real world scenarios. The technical justification for invoking restrictive assumptions is understandable. As we explained earlier in Section 2.4.1, the uncertainty as to the availability of stock in the warehouse creates interdependence between the retail stores and their supplier, the warehouse, which in turn induces interdependence among the retail stores themselves. Without invoking some restrictive assumptions to avoid the consequences of warehouse shortages, analytical approaches often turn out to be intractable. Those restrictive assumptions, however, may have prevented us from observing some important system behaviors with regard to the positioning problem.

Rather than following the traditional approach, we attempt to pursue a new research strategy by reversing the order in which analytical approaches and the use of computer simulation are applied. We start with a well-controlled simulation model without restrictive assumptions to observe system behaviors under conditions that resemble real world scenarios. These observations would provide the basis for analytical explanations. It might not be possible to build a tractable analytical model for the entire system, but analytical explanations for some specific phenomena observed might still be derived.

Simulation has proved to be a powerful tool in analyzing multi-echelon inventory systems and has been applied to many aspects of the system. Computer simulation seemed to be a particularly useful tool for our investigation for the following reasons. First, we no longer needed to invoke restrictive assumptions to avoid the consequences of warehouse shortages. Second, we were able to perform a diagnostic evaluation on alternative positioning strategies quickly, while at the same time, characterizing the complexities and subtleties inherent within the various positioning strategies. Third, the use of computer simulation allowed us to maintain much better control over experimental conditions than would generally have been possible either in analytical approaches or when experimenting with a distribution system itself.

However, we remain aware that simulation models might not be as accurate as analytical approaches in identifying the best inventory positioning strategies. Indeed, each run of a stochastic simulation produced only estimates of the model's true characteristics for a particular set of input parameters. In general, multiple independent runs of the model are required before we can draw any meaningful conclusions. On the other hand, an analytical model, if appropriate, can often easily produce the exact characteristics of that model for a variety of sets of input parameters. Unfortunately, at this stage, there did not seem to be any appropriate and efficient algorithms that would give good estimates of the costs and customer services associated with different strategies for positioning system-wide inventories in a one-warehouse multi-retailer distribution system with stochastic demand. While there may be an element of truth to pejorative old saws such as "method of last resort" sometimes used to describe simulation, the fact is that we are very quickly led to simulation in many situations as an only resort, due to the sheer complexity of the systems of interest and of the models necessary to represent them in a valid way. Our investigation of the inventory positioning problem seems to have been a case in point.

While simulation was our primary research tool in this dissertation study, we viewed our simulation model only as a starting point in the journey to gain clear insight into the question of where inventory should be positioned in distribution systems with stochastic demand. For future research, we believe that both analytical approaches and the use of computer simulation are needed. While integrative analytical models appear to be well beyond the capability of the current state of the art, the research strategy we propose seems to be an alternative worth pursuing. We believe that analytical approaches may supplement a well-controlled simulation model on a piecemeal basis, carving out one area at a time, and providing analytical support for decision makers.

3.2 NEW APPROACHES

This section consists of three subsections in which we have outlined the new approaches we developed for the investigation of the inventory positioning problem.

The intent is twofold: first, to reveal the unique features of this dissertation study and place our investigation in the context of the existing research literature as reviewed in Chapter 2; and second, to provide a preview of the simulation model and experimental design which will be discussed in detail in Chapters 4 and 5.

3.2.1 TREATMENT OF SHORTAGES

The use of computer simulation allowed us to avoid restrictive assumptions concerning warehouse shortages, but how warehouse shortages were handled in our simulation model warrants further explanation.

At the retail level, shortages become backorders if customers are willing to wait, and become lost sales if not. In this study, we assume that unsatisfied customer demands are fully backordered at the retailer stores. The question is: should we treat shortages at the warehouse in the same way as the shortages at the retail stores are treated?

Of course, we can backorder warehouse shortages as well. But we would argue that backorder is not the only way to handle the warehouse shortages. Before proposing our new approach, it is worth pointing out that retailer's replenishment orders differ from customer demand in several important aspects. First, retailers may request more than what they currently need in order to exploit the economies of scale in ordering and in transportation. Also note that retailers usually do keep inventories. Consequently, warehouse shortages may or may not result in a shortage at the retail level. Backorder of the warehouse shortages, therefore, cannot be as easily justified as in the case of shortages at the retail level. Second, retailers are internal customers, who stay in the distribution system even if the warehouse from time to time cannot satisfy their orders in full. Therefore, the lost sales assumption can hardly apply to warehouse shortages. Since they place replenishment orders on the warehouse repeatedly, the retailers generally would prefer to make their replenishment decision based on the most updated information. If warehouse shortages are backordered, the allocation priorities are usually given to orders placed earlier. As we explained in Section 2.3, the first-come-first-served allocation rule is not optimal.

What has often been ignored, however, is the fact that backorder shortages at the warehouse could be undesirable for the retailers in making their replenishment decisions as well. For illustration, let us consider the following scenario: A retail store has placed an order consisting of the quantity supposedly needed for satisfying short term demand (X) and an additional quantity (Y) added on to exploit the economies of scale in transportation. If the warehouse allocates by backorders and can allocate only X units of stock to the retail store, then the shortfall (Y units) will be kept on a backorder file to be delivered to the store as soon as stock at the warehouse becomes available. While the delivery of the shortfall (Y units) to the retail store will not be available to meet short-term demand, additional fixed shipment costs have occurred. This surely is not what the retail store had been hoping for. In fact, saving fixed shipment costs was the reason that the retailer added additional Y units to the replenishment order in the first place. Knowing that there was not enough stock at the warehouse, the retailer may wish to withdraw the original request for additional Y units to avoid the unwanted fixed shipment costs. The store may instead add this shortfall into its next replenishment order to exploit the economies of scale in the future. Unfortunately, backorders at the warehouse will have precluded the possibility for such a maneuver.

In another scenario, we may consider the case where no partial filling is permitted. In this case, if the warehouse cannot provide the additional Y units, the entire retailer order is backordered until it can be filled in full. While this approach has avoided the unwanted fixed shipment costs, the delay of the delivering the X units to the store could have caused customer service to deteriorate. If the retailer could withdraw the original request for additional Y units, it would be done. But again when shortages are backordered at the warehouse, such a maneuver is not allowed. Finally, we may not be allowing the retailer to add the additional Y units in the first place, but such restriction would be hard to justify because managers do wish to exploit the economies of scale in ordering and transportation if they can.

In this study, we treat shortages at the warehouse and at the retail store differently. While shortages at the retail stores are backordered, no backorder is allowed at the warehouse. The warehouse shortage information is handled in a decentralized fashion. As soon as the retail stores find out that the stocks delivered to them differ from what they have ordered, they will adjust their inventory status and make their subsequent replenishment decisions based on the updated inventory status. The warehouse responds only to the orders issued in the current period and allocates available stocks to the retail stores to equalize the probabilities of stockout at the retail level as much as possible. However, if the shortages at the warehouse are not backordered, can we still incorporate the warehouse shortage information into the warehouse replenishment decisions?

We know that shortages at the retail level are backordered. If the warehouse shortage information were not incorporated into warehouse replenishment decisions, then the system may not work properly. Backorder at the warehouse is one way to incorporate the warehouse shortage information into the warehouse replenishment decision. But it is not the only way. In fact, the shortage information can be incorporated into the warehouse replenishment decisions without resorting to backorders. The solution is to allow negative net inventory at the warehouse for the purpose of triggering replenishment orders and have the warehouse make its replenishment decision after rather than before receiving the replenishment orders from the retailers.

The net inventory at the warehouse was simply on hand inventory at the warehouse minus all the retailers' replenishment orders currently received. The negative net inventory signals that there was a shortage at the warehouse. Although the shortage would not be backordered, that would not prevent the net inventory at the warehouse from becoming negative for the purpose of triggering warehouse replenishment orders.

The problem actually can be avoided by having the warehouse make its replenishment decisions based on system-wide inventory status, rather than the inventory status at the warehouse exclusively. Specifically, the warehouse can make its replenishment decisions based on echelon inventory position, defined as the sum of inventory positions at the retail stores and on-hand inventory at the warehouse plus stock in transit from the outside supplier to the warehouse. This approach has been detailed in the next chapter, Section 4.1.5.

The proposed decentralized approach to warehouse shortages was facilitated by a newly developed information technology, called "Advance Shipment Notice" (ASN).

With ASN, the retail stores know how much stock is in transit as soon as the stock leaves the warehouse. Thus the retailers can update their inventory status immediately, rather than having to wait until the shipments physically arrive. ASN is gaining acceptance in practice. According to a 1992 survey among "quick response technologies" that companies plan to invest in, ASN was ranked number two, next only to Electronic Date Interchange (EDI) (Chain Store Age Executives, April 1992, p31). Based on those observations, we believe that the decentralized approach to warehouse shortages proposed in this dissertation could find wide applications in industry.

As we explained earlier (in Section 2.3.1), the inventory position represents the amount of stock that is available to meet future demand without placing further orders. When shortages at the upstream stocking location are backordered, the inventory position is defined as the sum of net inventory plus stock on order. When shortages are not backordered at the upstream echelon, the inventory position needs to be redefined as the sum of net inventory plus stock in transit. Quantities of the goods that have left the supplier's facility, but have not yet been received are stock in transit. Clearly, the retailers with ASN know what stock is in transit. If ASN is not available, then the retailers may use stock on order to estimate stock in transit. Under the assumption that no shortages are backordered at the warehouse, the sum of the net inventory plus stock on order no longer always represents the amount of stock that is available to meet future demand without placing further orders. For this reason, in the remainder of this dissertation, the sum of net inventory plus stock on order has been referred to as the nominal inventory position. The term "inventory position" has been reserved for the sum of the net inventory plus stock in transit.

Our treatment of warehouse shortages invoked no restrictive assumptions, which allowed us to observe the system behavior in a dynamic environment, rather than be limited to a single warehouse replenishment cycle. Under the new approach, the warehouse followed balancing allocation rules and the retail stores made their replenishment decisions all based on the most updated information. The process of updating inventory status at the retail level also provided flexibility needed for the retailers to explore the economies of scale in ordering and in transportation.

3.2.2 THE CONTROL OF SYSTEM WORKLOAD

Customer service, total inventory, and system workload are three primary dimensions on which we evaluate the performance of an inventory system. In the inventory theory literature, ordering frequency is often used to measure the system workload. For a single location system in which a single stocking location is supplied directly by an outside supplier, such modeling is adequate because ordering and shipment frequencies are identical. It is no longer so for a multi-echelon inventory system where an upstream stocking location may be out of stock and each replenishment order does not necessarily trigger a shipment. Therefore, for a multi-echelon inventory system, we need to distinguish between ordering and shipment frequencies.

Thanks to information technology, ordering costs have declined sharply in recent years. On the other hand, the transportation cost has remained relatively stable and still counts the largest percentage of average distribution cost (Davis and Drumm, 1994). The change in the cost structure further indicates that the traditional modeling approach to system workload by focusing on the ordering frequency is inadequate in today's distribution practice. In this study, we have measured the system workload by shipment frequency and assumed the ordering costs are very small.

Published studies on the positioning problem, particularly those on the "depot effect," have by and large ignored the system workload. The shipment frequencies associated with different positioning strategies have rarely been reported or even recorded. To directly compare alternative inventory positioning strategies without taking into account their difference in shipment frequencies is problematic because fixed shipment costs can rarely be ignored both in theory and in practice.

To a researcher, however, dealing with fixed shipment costs is an unusually challenging task. Despite many years of research, no satisfactory planning methods are available for a distribution system with fixed shipment costs between the warehouse and the retail stores. It is well known that the first multi-echelon inventory model built by

Clark and Scarf (1960) can handle fixed shipment costs only at the highest echelon, where the outside supplier is assumed to be completely reliable. It has also become clear by now that in the presence of fixed shipment costs, even for a serial distribution system, no exact decomposition of the dynamic program into two separate "single location" problems as suggested by Clark and Scarf (1960) can be achieved (Federgruen 1993a). Chen and Zheng (1994) recently obtained a tractable lower bound for a one-warehouse multi-retailer distribution system with fixed plus variable shipment costs. A complete methodology, however, is still in progress.

In this study, we develop a new approach to control the system workload. We assumed that the frequencies of shipments to the warehouse and to the retail stores were predetermined. We carefully controlled these shipment frequencies as alternative inventory positioning strategies were simulated and evaluated. When positioning strategies with different shipment frequencies need to be compared, we develop a method to decompose their performance difference into two components: One is due to differences in the positioning strategy, subject to the identical constraint on the system workload. Another is attributable to the changes in system workloads for pursuing the same inventory positioning strategy. The decomposition method is detailed in Section 3.3. Finally, we also investigated the changes in the predetermined shipment frequencies and their effect on the proper positioning of inventory in the distribution system. These new approaches ensured that alternative positioning strategies simulated are comparable and the results are meaningful.

3.2.3 INVENTORY CONTROL SYSTEM AS AN EXPERIMENTAL FACTOR

Research on the inventory positioning problem typically is conducted for a single inventory control system. The results of individual studies are hard to compare because they come from inventory control systems. The lack of comparability among the strategies for the proper positioning of inventory under different inventory control systems has prevented researchers from being able to reach general conclusions on the inventory positioning problem.

To expand the current understanding of inventory positioning problem, we could choose to investigate the positioning problem for another inventory control system or to conduct a much-expanded experiment in which the inventory control system is treated as an experimental factor. We made a strategic choice to treat the inventory control system as an experimental factor for two reasons. First, we want to draw conclusions that are not limited to a specific inventory control system. Second, we want to study the effect of different control systems on the proper positioning of inventory in the distribution system.

As an experimental factor, the inventory control system is a categorical variable. To investigate the effect of alternative inventory control systems on the proper positioning of inventory, multiple inventory control systems had to be included in our experimental design. We designed several inventory control systems with similar decision rules, but different information requirements. Their information-flows design imitated different information-sharing schemes we observe in industry. (In Chapter 4, we have provided full descriptions of the inventory control systems we studied.) Such a design enabled us to address one important question: as more information becomes available and is utilized by the inventory control system, should more or less inventory be held back at the warehouse?

3.3 THE SIMULATION MODEL FRAMEWORK

Following the new approaches just outlined, we built a simulation model for a one-warehouse multi-retail distribution system operated in a periodic review environment. For this study, the retail stores are assumed to be identical and the warehouse that follows a balancing allocation principle allocating available stock to equalize the probabilities of stockout among retail stores. No backorders are kept at the warehouse.

We defined the system-wide inventory level as the sum of the average on hand inventory level at the warehouse and the average on hand inventory at the stores. We characterized the inventory positioning strategy by a positioning ratio P, the average on hand inventory at the retail level divided by the system-wide inventory level. If P was equal to 1, it indicated that all inventory was positioned at the retail level near the customer. In general, $0<P<1$ indicated that only a proportion of system-wide inventory was positioned at the retail level.

We simulated inventory positioning strategies by manipulating the inventory control system parameters in a way that changed the positioning ratio, P, while holding all other factors under experimental control; in particular, shipment frequencies and the system-wide inventory level were consistently checked to ensure that they were controlled at the predetermined levels.

For a given set of inventory, transportation, and inventory control system resources, the positioning ratio that provided the highest level of customer service was identified as the best inventory positioning. The positioning ratio for the best inventory positioning enabled us to answer our basic research question: Where should inventory be positioned in a one-warehouse multi-retailer distribution system with stochastic demand?

The best positioning ratio, however, would not have provided information for us to answer another important question: What is the benefit from pursuing the best positioning? To answer this second question, we defined a reference point for comparison. The traditional approach would have been to compare directly the service level for the best positioning strategy with that for the "ship-all" without taking their difference in the frequency of shipments into account. In this study, we choose a new reference point, called positioning inventory close to the customer. These three positioning strategies, the "best" positioning, the positioning close to the customer, and the "Ship-all" positioning are found as the solutions to the three problems described below.

Given a system-wide inventory level (INV#) and a frequency of shipments to the warehouse (SHIPW#) and to the retailers (SHIPR#), we characterize a positioning strategy by two parameters: positioning ratio P and the service level it provided as F, we denoted it as (P, F). The best inventory positioning strategy (P^*, F^*) corresponds to the solution to the following problem:

"Best" Find the positioning ratio, P*, that

Max F (3.0)

Subject to INV=INV# (3.1)

SHIPR=SHIPR# (3.2)

SHIPW=SHIPW# (3.3)

Where F= Customer fill rate

INV= System-wide inventory level

SHIPR= Frequency of shipment to each of the retail store

SHIPW= Frequency of shipment to the warehouse

The "ship-all" positioning strategy (Ps, Fs) was the solution to the following problem:

"Ship-all" Find the Fs that

Max P (3.0')

Subject to INV=INV# (3.1)

SHIPR$\leq$SHIPW#[1] (3.2')

SHIPW=SHIPW# (3.3)

Where P= Inventory positioning ratio

INV= System-wide inventory level

SHIPR= Frequency of shipment to each of the retail store

SHIPW= Frequency of shipment to the warehouse

[1] Some explanation for the constraint (3.2') (i.e., SHIPR$\leq$SHIPW#) seems to be warranted. Under the "Ship-all" positioning strategy, the warehouse allocates all incoming stocks from the outside supplier to the retail stores immediately. That's why the shipment frequency from the warehouse to the retail stores on average cannot be greater than the frequency in which the warehouse receives its incoming stocks. In case there is an unexpected large demand at one retail store, a large proportion of the incoming stocks will be allocated to that store according to the balancing allocation rule. As a result, some other stores that experienced low demand may not get any incoming stocks in that particular allocation decision. When this occurs, we may have SHIPR<SHIPW#. In general, SHIPR$\leq$SHIPW#.

Direct comparison between F* and Fs was problematic in other research because the difference between the constraint of (3.2) and (3.2') had not been taken into account.

We overcame the lack of comparability in this study by defining another inventory positioning strategy (Pc, Fc), called "close" to the customer. (Pc, Fc) was the solution to the following problem:

"Close"	Find Fc that		
		Max P	(3.0')
	Subject to	INV=INV#	(3.1)
		SHIPR=SHIPR#	(3.2)
		SHIPW=SHIPW#	(3.3)

Where P = Inventory positioning ratio

INV = System-wide inventory level

SHIPR= Frequency of shipment to each of the retail store

SHIPW= Frequency of shipment to the warehouse

Note that the actual inventory positioning ratios P, average inventory levels INV, and shipment frequencies SHIPR and SHIPW were all ex post facto measures of random variables. Therefore, the three deterministic problems should be understood only as conceptual presentations of the problems we solved, rather than the definition of the problems.

Clearly, the differences between "Best" and "Close" lay in the objective function (3.0) and (3.0'). They had the identical inventory and transportation resource constraints. Therefore, the service level difference between "Best" and "Close" was due to the difference in the positioning strategy subject to the identical resource constraints. On the other hand, the problems "Ship-all" and "Close" had the identical objective function (3.0'), but differed in the constraint (3.2) and (3.2') on the frequency of shipments to the retail stores. The difference was due to the change in the shipment frequency.

Thus, we could decompose the fill rate difference between "Best" and "Ship-all" into to two components as follows:

Fill rate for the best positioning – Fill rate for the "Ship-all"
$=F^*-Fs$
$=(F^* - Fc) + (Fc - Fs)$
=[Fill for "Best"- Fill rate for "Close"]
+ [Fill rate for "Close" – Fill rate for "Ship-all"]
=[Fill rate improvement due to best positioning]
+ [Fill rate improvement due to the increase in the shipment frequency SHIPR] (3.4)

What we are really concerned with is the fill rate improvement due to the best positioning, (F^*-Fc). It represents the penalty for simply positioning inventory as close to the customer as possible, rather than pursuing the best inventory positioning. In our simulation experiments, the fill rate improvement due to the increase in frequency of shipments to the stores was also recorded (so that our results could be compared with the results reported by other researchers).

In this study, we solved the three problems –the "Best", the "Close", and the "Ship-all"-- by searching control system parameters via simulation. For different control systems, different search procedures are required. We describe these procedures in the next chapter, Chapter 4, following the description of inventory control systems used.

CHAPTER 4

INVENTORY CONTROL SYSTEMS

In this chapter, we describe the inventory control systems used and the simulation search procedures we developed for identifying the "best," the "close," and the "Ship-all" positioning for each of the inventory control systems proposed. Their complexity and central role in the study warrant detailed treatment.

4.1 THE DESCRIPTION OF INVENTORY CONTROL SYSTEMS

This section consists of five subsections. First, we outline the common features shared by the inventory control systems we used. Second, we describe information-availability scenarios we considered. Third, we show that given information availability, there could be different schemes for utilizing the available information. Fourth, we provide a detailed description of the four inventory control systems used. Finally, in Section 4.1.5, we present a framework for designing multi-echelon inventory control systems.

4.1.1 COMMON FEATURES

The inventory control systems we designed share a common rule for determining when and how much additional inventory is needed. The general form of this replenishment rule is (R, s, S), implemented as follows. There is a review of inventory status every R periods. If the inventory status is less than or equal to s, an order is placed to bring the inventory status up to level S; if not, nothing is done until the next review period. Let O(t) be the order quantity placed by a stocking location in period t.

$$O(t) = S - X(t), \text{ if } X(t) \le s \qquad (4.0)$$
$$= 0, \text{ otherwise.}$$

Where s and S are reorder point and order-up-to level, respectively. $X(t)$ is the inventory status reviewed at that stocking location in period t. The term "inventory status" is purposely vague because it can be measured in different ways, depending on the availability of information and the scheme for utilizing the available information. The discussion on the information availability scenarios and the different schemes for utilizing the available information appear in the following two subsections (4.1.2 and 4.1.3.)

Whenever there is a need to allocate warehouse inventory to the retail stores, a common principle is applied. This need arises as part of the normal operation of the decision rule or when there is insufficient inventory to satisfy the orders from the retail stores. The general allocation principle is to equalize, as much as possible, the probability of stockout at each of the retail stores, called the balancing allocation rule. As pointed out in Section 2.4, for identical retail stores, this allocation principle is equivalent to maximizing the minimum inventory position at the retail level, called the Max-min Allocation Rule.

The implementation of the Max-min Allocation Rule required calculating the balanced stock level first, then determining the quantities of warehouse inventory to be committed to specific retail stores, and finally deciding whether or not the committed warehouse stocks should be delivered to the stores. To illustrate the calculation method of the balanced stock level, we assigned N to be the number of retail stores; S_r the order up to level of each retail store; $IP_j(t)$ the inventory position of the jth retail store in period t; $WI(t)$ the on-hand inventory of the warehouse at the time t; and $FS(t)$ the balanced level of the retail stores in period t. Upon receiving orders from any retailer, the warehouse first calculates the balanced level of the retail stores in period t, $FS(t)$, according to the following equation

$$FS(t) = \text{Min}\{ S_r, [WI(t) + \sum_j IP_j(t)]/N \} \qquad (4.1)$$

where S_r is order-up-to level used by each of retail stores. If $FS(t) \geq IP_j(t)$, for all j, then the calculation is finished. Otherwise, the calculation continues by removing those retailers with $IP_j(t) > FS(t)$, and then updating N, $IP_j(t)$ and $FS(t)$ to refer to only those retailers remaining. The process will be repeated until $FS(t) \geq IP_j(t)$ for all those remaining retailers, j.

Note that the retailer's order-to-level, S_r, sets an upper bound for how much stock can be allocated to any of the retail stores. $FS(t)$ - $IP_j(t)$ is stock committed to the jth store at time t. Let $O_j(t)$ be the jth store's replenishment order placed on the warehouse in period t. According to Equation (4.0),

$$O_j(t) = S_r - X_j(t), \text{ if } X_j(t) \leq s_r$$
$$= 0, \text{ otherwise.}$$

Where s_r and S_r are reorder point and order-up-to level used by each of the retail stores, respectively. $X_j(t)$ is the inventory status of the jth store in period t. Let $A_j(t)$ be the stock allocated to the retailer j in period t, then

$$A_j(t) = \text{Min } \{ O_j(t) , FS(t) - IP_j(t) \} \tag{4.2}$$

Equation (4.2) implies that if there is no order, then there is no delivery. Because no retailer order could be triggered unless $X_j(t) \leq s_r$, the retailer's reorder point s_r sets the timing for allocation decision.

In case $IP_j(t)$ is unknown to the warehouse, the warehouse could still follow the allocation principle approximately. A full description of the allocation rules used by individual control systems are presented in Section 4.2.

It also worth pointing out that after calculating $FS(t)$, there may be some residual stocks (less than the most updated N) left over at the warehouse. In this case, we would randomly assign the residual stocks to the store j if $A_j(t)$ >0 until nothing left.

4.1.2 INFORMATION AVAILABILITY SCENARIOS

We assumed that both the warehouse and the retail stores make their replenishment decisions using the (R , s , S) decision rules on their own. When an allocation decision has to be made, the warehouse follows the Max-min Allocation Rule as described in the previous section. The implementation of these decision rules critically depends on the availability of information and the scheme for utilizing the available information. In this section, we describe four information scenarios we considered.

Scenario 1: From time to time, the retail stores and the warehouse place replenishment orders with their supplier. Besides the replenishment orders, there is no information exchange between the warehouse and the retail stores. In this sense, there is only local information available (See Figure 4.1). Because the retail stores do not know the amount of stock in transit, they have to monitor their inventory status by the nominal inventory position, defined as the sum of the net inventory plus stock on order. If there is a shortfall at the warehouse, the retail stores will adjust their inventory status only when shipments from the warehouse physically arrive at the retail stores. The delay will force the retail store to make subsequent replenishment decisions based on the distorted inventory status, the nominal inventory position.

The warehouse allocates stocks to the retail stores based on replenishment orders received. The warehouse tries to balance inventories among those stores that have placed orders in the current period. The store that has not issued a replenishment order will not be considered.

Scenario 2. Everything is assumed to be the same as in scenario 1, except that now we assume that Advance Shipment Notice (ASN) has become available. With ASN, the retail stores receive information about stocks in transit from the warehouse as soon as the stocks leave the warehouse. Thus, the retail stores are able to monitor their inventory status by inventory position, defined as the sum of net inventory plus stock in transit, which avoids the information distortion caused by warehouse shortages. As a newly developed information technology, ASN has rarely been studied in the literature (See Figure 4.2).

Scenario 3. Everything is assumed to be the same as in scenario 2, except that now we assume that the warehouse knows the retailers' inventory positions (RIP) regardless of whether or not they have placed a replenishment order. With this information, the warehouse can allocate stocks to balance inventories of all retail stores, rather than only to those have placed orders in the current period. Sharing of inventory status information with suppliers has been advocated for managing multi-echelon inventory systems for some time (See Figure 4.3).

Scenario 4. Everything is assumed to be the same as in Scenario 3, except that the time-phased projected demand is now available and the retail stores make their replenishment decisions based on the Distribution Requirement Planning (DRP) ordering logic. DRP ordering logic can be described as follow: replenishment orders are planned to prevent the projected inventory balance (i.e., projected net inventory) from falling below the level s. Whenever the projected inventory balance is at or below s, a shipment is planned to arrive in that period to bring the projected inventory balance up to S. Now, the retail stores not only from time to time place replenishment orders with the warehouse, but also share their planned order information with the warehouse. Similarly, the warehouse follows DRP ordering logic for its own replenishment decisions and shares its planned orders with the outside supplier as well.

One of the important features of DRP is that it can incorporate time-phased projected demand (TPPD) into its replenishment decisions. If the average demand is used as time-phased projected demand, the features associated with DRP are reduced to one-- that is, the sharing of planned orders (PO) information with the supplier (Figure 4.4).

4.1.3 SCHEMES FOR UTILIZING AVAILABLE INFORMATION

Given the information availability, there can be more than one way to utilize the available information. Scenario 3 provides a good example. As described in the previous section, under scenario 3, the retail stores' inventory positions are known to the warehouse. Thus, the warehouse can use this information to make better allocation decisions; that is, to allocate stocks for balancing inventories among all retail stores, rather than among just those that have placed orders in the current period exclusively. Besides using this information for allocation decisions, the warehouse can also incorporate it into its replenishment decisions.

Instead of monitoring its installation-inventory position (i.e., inventory position at the warehouse), the warehouse can make its replenishment decisions based on echelon inventory position, defined as the sum of the inventory positions of all retail stores plus on-hand inventory at the warehouse and stock in transit from the outside supplier. This represents the amount of stock that is available in the whole distribution system to meet future demand without placing further orders on the outside supplier. Note that when echelon inventory position is used, a warehouse shortage has no direct effect on the warehouse's replenishment decisions.

Clark and Scarf (1960) first introduced the concept of the echelon inventory position. It was based on echelon inventory position that they construct a penalty-cost-function for the upstream echelon for not being able to provide adequate stock to the downstream echelon. For comparative studies on installation vs. echelon-based inventory control systems, see Axsater and Rosling (1993b, 1994), and Chen and Zheng (1994).

4.1.4 A FRAMEWORK FOR DESIGNING MULTI-ECHLON INVENTORY CONTROL SYSTEMS

Concerning the possible combinations of the information- availability scenarios and the different schemes for utilizing the available information (i.e., installation vs. echelon), various inventory control systems could be designed.

First, we considered installation-based inventory control systems. We developed four control systems, corresponding to the four scenarios described in Section 4.1.2. When there was only location-information available as described in scenario 1, the retail stores could follow the (R, s, S) replenishment decision rules based on the nominal inventory position, and the warehouse could make allocation decisions based on replenishment orders received from the retail stores. We called this control system the Location Information Reorder Point and Order-up-to-Level System, denoted as LROP. In scenario 2, ASN had become available. The stores made their replenishment decisions based on their inventory position, rather than on a nominal inventory position. Yet, the warehouse still had to make its allocation decision based on replenishment orders from the retail stores because the warehouse did not know the inventory status at the retail level. We called such a control system Advanced Shipment Notice System (ASNS). In scenario 3, the retail stores shared their inventory status information with the warehouse. Thus, the warehouse could allocate stocks to balance inventory positions of all retail stores, rather than only those that had placed orders in the current period. We called this control system Reorder Point and Order-up-to Level System (ROP). Finally, in scenario 4, we had a Distribution Requirement Planning (DRP) system, in which inventory status was measured by projected inventory balance.

Next, we considered the echelon-based inventory control systems. In scenario 3, everything was assumed to be the same as ROP except where the warehouse used an echelon based (R , s , S) replenishment decision rule. We called this control system an Echelon Reorder Point and Order-up-to Level System (EROP). For all control systems described so far, we assumed that the inventory review period $R = 1$, which meant that both the warehouse and the retail stores reviewed their inventory status every period and made their replenishment decisions according to their replenishment decision rules. Now, we considered a special EROP system in which the warehouse would review its echelon inventory position once every R_w periods, where $R_w > 1$. In terms of information availability, this is clearly a more restrictive scenario. In addition, we assumed that the retail stores would not make replenishment decisions. Whenever the shipment arrived at the warehouse, α percentage of the incoming stock would be allocated to the retail stores immediately. The remaining incoming stock would be allocated T periods later, where $T < R_w$. In this study, we refer to this control system as the "Push" system.

Table 4.1 **A PRAMEWORK FOR DESIGNING MUTI-ECHLON CONTROL SYSTEMS**

Inventory Status/ Information.	RESTRICTIVE	[redacted]		
INSTALLATION	Local Info. **LROP**	[redacted] ASNS	[redacted] ROP	[redacted] **DRP**
[redacted]	[redacted] $R_w > 1$ **"PUSH"**		[redacted] **EROP**	
PROJECTED DEMAND	Average			Time- Phased

ASN: Advanced Shipment Notice
RIP: Retailers' Inventory Positions known to the Warehouse
PO: Planned Order Information
TPPD: Time-Phased Projected Demand

We simulated all the control systems just described. But our investigation was primarily focused on four control systems: LROP, EROP, DRP, and "Push" systems. We described all the control systems simulated in this section with the intent of placing these control systems in a framework that could help us understand their differences and relationships to each other. While we did not include ASNS and ROP in our major simulation experiments, we found out that they were quite useful in explaining some important results. For instance, ASNS and ROP allowed us to decompose the observed performance differences between LROP and DRP. The decomposition provided evidence consistent and compelling enough to draw a general conclusion on the effect of information-sharing on the proper positioning of inventory in the distribution system. Chapter 6 provides a detailed discussion of this decomposition approach.

The framework we just presented shows that multi-echelon inventory control systems need to be characterized and differentiated in multiple dimensions. As information sharing schemes become more complicated, the traditional classification schemes such as decentralized vs. centralized control systems, or local information vs. global information control systems, no longer seem adequate or capable of differentiating the newly developed multi-echelon inventory control systems. The framework presented here is just a starting point. More research needs to be done before it would be possible to present a comprehensive framework for designing multi-echelon inventory control systems that fits today's technology and business environment.

4.1.5 DESCRIPTION OF INVENTORY CONTROL SYSTEMS USED

In this section, we gave the detailed description of the four inventory control systems we investigated. As noted earlier, these control systems are distinguished, primarily, by the information that is used to implement their decision rules. A description of each of the systems follows in detail below.

LOCAL INFORMATION BASED REORDER POINT AND ORDER-UP-TO LEVEL CONTROL SYSTEM (LROP)

LROP represents independent management of the facilities in the distribution system. It is characterized by the use of information available at the facility only, called local information. The retail stores monitor their inventory status by the nominal inventory position, which is defined as the sum of the net inventory plus stock on order. Shortages at the retail stores are backordered. There is no backorder at the warehouse. However, before triggering a replenishment order on the outside supplier, the inventory position at the warehouse could be negative if the retail orders exceeded the sum of the supply at the warehouse plus stock on order from the outside supplier. The (R , s , S) decision rule is implemented with $R = 1$ at the warehouse and at all retail stores. Yet, the reorder point s and order-up-to level S used by the retailers could be different from those used by the warehouse.

The warehouse fills all store orders if there is enough inventory available. If not all orders can be filled in full, the inventory is allocated to the retail stores by minimizing the maximum unfilled portion of the retail order. This equalizes the probability of stockout (to the extent possible) for those retailers who have placed orders. Since the only retail information available at the warehouse is the orders, stores that have not placed orders cannot be considered.

To explain the allocation rule, note that,

$$IP_j(t) = S - O_j(t),\ \text{j for } O_j(t) > 0. \tag{4.3}$$

where $O_j(t)$ is the jth retail store's replenishment order placed on the warehouse in period t. This information is known to the warehouse. Substitute (4.3) into (4.1) and update N for the number of stores that have placed orders in the current period only. Then, it can be shown that the Max-min Allocation Rule in this case minimizes the maximum shortfall of the retailer order. That is, the balanced level of the retail stores in period t, $FS(t)$,

$$\begin{aligned} FS(t) &= \text{Min } \{ S_r, [WI(t) + \sum_j (S_r - O_j(t))]/ N \} \\ &= \text{Min} \{ S_r, S_r + [WI(t) - \sum_j O_j(t)]/ N \} \end{aligned}$$

If $WI(t) \geq \sum_j O_j(t)$, then the warehouse simply allocate what the retailers have requested. If $WI(t) < \sum_j O_j(t)$, then $[WI(t) - \sum_j O_j(t)]/ N$ is the average shortfall. Any

$O_j(t)$ that is smaller than the average shortfall will be removed from the calculation. Update N, the calculation continues until $O_j(t)$ is greater than the average shortfall for all remaining retail stores j. The stock to be allocated to the jth store is then determined by (4.2) as follows.

$$
\begin{aligned}
A_j(t) &= \text{Min}\ \{ O_j(t),\ FS(t) - IP_j(t) \} \\
&= \text{Min}\ \{ O_j(t),\ S_r + [WI(t) - \sum_j O_j(t)]/\ N\ -\ IP_j(t) \} \\
&= \text{Min}\ \{ O_j(t),\ S_r + [WI(t) - \sum_j O_j(t)]/\ N\ -[S_r - O_j(t)] \} \\
&= \text{Min}\ \{ O_j(t),\ O_j(t) + [WI(t) - \sum_j O_j(t)]/\ N \} \\
&= O_j(t) + [WI(t) - \sum_j O_j(t)]/\ N \qquad (4.4)
\end{aligned}
$$

Equation (4.4) shows that the warehouse does not need to know the order-up-to level S_r used by the retailers to make its allocation decisions. When the warehouse stocks out, the stock in transit to a retail store could be less than what the store has requested. However, the retail stores do not immediately know this information. The delay forces the retail store to make subsequent replenishment decisions based on the nominal inventory status. If what they get is less than what they have ordered, the retail stores will adjust their inventory status to reflect the shortfall and make their subsequent replenishment decisions based on the updated inventory status. Note that the order triggered by the distorted inventory status information—the nominal inventory position-- could in turn mislead the warehouse allocation decisions.

ECHELON REORDER POINT AND ORDER-UP-TO LEVEL CONTROL SYSTEM (EROP)

In this control system, the degree of integration is achieved by assuming that ASN (Advanced Shipment Notice) is in place and the retail stores share their inventory status information with the warehouse. With ASN, the retail stores monitor their inventory status by the inventory position (i.e., the sum of net inventory plus inventory in transit), rather than that nominal inventory position. Stock in-transit could differ from what was ordered, but the retail store learns this information as soon as the allocation decision is made at the warehouse. As in control system LROP, the (R , s , S) replenishment decision rules are used with (R =1) for the replenishment of warehouse inventories and retail store orders.

In determining whether to place a replenishment order at the warehouse, a distribution system-wide view is taken. Inventory status at the warehouse is monitored by the echelon inventory position (i.e., the sum of the net inventory and stock in-transit at all retail stores plus on-hand inventory and in-transit inventory at the warehouse). Whether or not the warehouse would trigger a replenishment order on the outside supplier depends on the amount of stock that is available in the entire distribution system to meet future demand with placing further orders on the outside supplier, rather than the inventory position at the warehouse exclusively. It is of note that in this control system, warehouse shortages have no direct impact on the warehouse's replenishment decisions.

The availability of retailer inventory status information changes the details of allocating inventory, if all retail orders cannot be filled. All retail stores are considered (not just those that have placed an order) in the allocation. The quantity required by each retailer is determined to maximize the minimum inventory status among all retailers. That quantity is sent to any retailer that placed an order in the period (i.e., no order, no shipment), while any remaining inventory stays in the warehouse to be used later. The calculation is done according to equation (4.1) and equation (4.2).

DRP (DISTRIBUTION REQUIREMENT PLANNING)

This is the most highly integrated system. It uses time-phased projected demand information (as opposed to just current information) to calculate inventory status. At each period, each retail store projects its inventory balance for several periods into the future, using the DRP ordering logic. That is, replenishment orders are planned to prevent the projected inventory balance from falling below the level s. Whenever the projected inventory balance is below s , a shipment is planned to arrive in that period to bring the projected inventory balance up to S .

The warehouse uses retail-planned orders as the demand in calculating its planned orders. It places a replenishment order with the outside supplier any time there is a planned order in the current period. The warehouse fills all current period planned retail orders if possible. If not, the warehouse uses the same allocation decision rule it used in EROP. Whenever allocation is necessary, the planned orders from the retail stores are re-projected to take the stock in transit from the warehouse into account. We observed cases where the initial retail-planned order did not trigger a warehouse replenishment but the re-projected retail-planned orders did.

"PUSH" SYSTEM

In this control system, the warehouse monitors its inventory status exactly as it did in EROP (i.e., echelon inventory position). The information-availability scenario 3 is assumed to be in place, except that the warehouse can now review its echelon inventory position once every six periods, instead of every period as in EROP. Therefore, information used by the "Push" system is more restrictive than that for EROP.

A different replenishment philosophy is used in this system. The warehouse replenishes its inventory on a fixed replenishment cycle of six periods in order to bring its echelon inventory position up to S_w. This is equivalent to using the (R_w, s_w, S_w) decision rule with $s_w = S_w - 1$ and $R_w = 6$. In this system, retail replenishments are taken care of by the warehouse and no orders are calculated by the retail stores. This assumption is, however, equivalent to the scenario in which every retail store uses a (1, s_r, S_r) replenishment decision rule with $s_r = \infty$ and $S_r = \infty$. Because the inventory position at any store at any time is a finite number, the (1, ∞, ∞) replenishment decision rule forces the store to place a "pseudo" replenishment order of a huge amount every period that can never be filled by the warehouse in full. As a result, the retailers' replenishment orders contain no information at all. This interpretation allowed us to consider "no retailer order" as a special case of reorder point order-up-to level replenishment decision rule.

Given the "pseudo" replenishment order, "no order, no delivery" as specified by (4.2) is no longer a constraint for the "Push" system. The "upper bound" imposed by the retailers on how much inventory could be allocated to any of the retail stores is not active either because the retailers' order-up-to level $S_r = \infty$ (see Equation 4.1). In order to keep some portion of the inventory at the warehouse, two new parameters α (substitutes for S_r) and T (substitutes for s_r) have to be introduced. The control system works as follows. When replenishment stock is received at the warehouse α% of the stock will be sent immediately to the retail stores. The remaining stock is retained at the warehouse to be distributed T periods later in the warehouse replenishment cycle, where $T < 6$. The warehouse allocates stock to the retail store according to the Maximum Allocation Rule. The calculation follows the same procedure as described in Section 4.1.1.

4.2 SIMULATION SEARCH PROCEDURES

Since the retail stores are assumed to be identical, we can characterize control system LROP, EROP, and DRP by four parameters (s_r, S_r, s_w, S_w). Each retail store uses the s_r and S_r, reorder point and order-up-to level, respectively, and the warehouse uses the s_w and S_w, reorder point and order-up-to level, respectively. R is fixed at 1 period for all these three control systems. Since they can be characterized by the same parameters, we say that LROP, EROP, and DRP have the same form of decision rules, but different information requirements.

The "Push" system has to be characterized by a different set of parameters (α, T, R_w, S_w). Parameter α specifies the percentage of incoming stock at the warehouse to be used for the primary allocation (i.e., to be allocated immediately). Parameter T specifies the timing for the secondary allocation. R_w specifies the length of the warehouse review period, which is also the length of the warehouse replenishment cycle. S_w specifies the order-up-to level for the warehouse to monitor its echelon inventory position.

In our simulation experiments, the values of R for both the retail stores and the warehouse were fixed, while the s_r, S_r, s_w, S_w values for the LROP, EROP, and DRP systems, and the α, T and S values for "Push" system were manipulated to vary the positioning ratio, P, while holding the other factors constant. For each control system, we needed to identify the parameter values for the best positioning, positioning inventory "close" to the customer, and the "Ship-all." In other words, we needed to find solutions for the three problems, the "Best", the "Close", and the "Ship-all" as described in Section 3.4.

In this study, we solved the three problems using simulation search procedures. Different control systems required different search procedures. Here, we first describe the search procedures for the EROP system, which are relatively easily implemented because warehouses make their replenishment decisions based on echelon inventory position. Given the reorder point (s_w) and the order-up-to level (S_w) used by the warehouse, the system-wide inventory INV is almost fixed. Manipulating the retailer's reorder point (s_r) and the order-up-to level (S_r) would change the inventory positioning ratio but would have relatively little impact on the system-wide inventory level, INV. Procedures for LROP and DRP are similar to the one for EROP but are more complicated and difficult to implement because as we manipulated the parameter values of s_r, S_r, s_w, and S_w, we had to consistently check on whether the constraints on the system-wide inventory level (3.1), and on the frequency of shipments (3.2) and (3.3) were violated.

SEARCH PROCEDURES FOR EROP

Given SHIPW#, SHIPR#, INV#, and control system EROP, we used the following procedures for finding solutions for the three problems: the "Best," the "Close," and the "Ship-all."

1. Initiating s_w=a positive value. Letting $(s_r, S_r, s_w, S_w) = (\infty, \infty, s_w, s_w + \Delta w)$, search for Δw so that SHIPW = SHIPW#. Denote this value Δw_1.
2. Using Δw_1 obtained in step1, search for s_w such that INV=INV#, record customer fill rate, Fs, which is the fill rate for the "Ship-all."
3. Letting $s_r = S_r$, search for the maximum S_r so that SHIPR > SHIPR#.
4. Using S_r obtained in 3, search $s_r = S_r$-1, S_r-2, ...until SHIPR=SHIPR#, record the customer fill rate Fc, which is the fill rate for the positioning inventory "close" to the customer. Also record the positioning ratio Pc.
5. Letting $S_r = S_r$-1, search for $s_r = S_r$-1, S_r-2, ...until SHIPR=SHIPR#, record the customer fill rate F^ and the positioning ratio P^.
6. Go to step 5, record the customer fill rate F, and the positioning ratio P. If the fill rate F>F^, go to 5; otherwise let F*=F^ and P*=P^, where (P*, F*) represent the best positioning. Stop.

If INV changes during the search procedure, some manipulation may be needed to make sure that constraint on the inventory resource (3.1) is not violated.

SEARCH PROCEDURES FOR THE "PUSH" SYSTEM

The search procedures for the "Push" system are relatively simple because we need to manipulate only three parameters α, T, and S. Given SHIPW# (i.e., 1 shipment every R_w periods), SHIPR#, INV#; for the "Push" system, the procedures we used for finding solutions for the "Best," the "Close," and the "Ship-all" are as follows.

1. Δ=0.001;
2. Letting (α, T, R_w, S_w) = (1, 0, 6, S_w), search for S_w such that INV=INV#, record the fill rate, Fs, which is the fill rate for the "Ship-all."
3. Let $\alpha=\alpha-\Delta$, if $\alpha>0$; otherwise, go to step 5.
4. For T=1, 2, 3, 4, and 5, record the service level F^ and the positioning ratio P^ if SHIPR = SHIPR#. Go to step 2.
5. Let the very first observed (P^, F^) = (Pc, Fc), which represents the strategy for positioning inventory "close" to the customer. Let the best positioning (P*, F*) = (F^, P^) where F^ is the maximum fill rate observed. Stop.

CHAPTER 5

EXPERIMENTAL DESIGN

This chapter is devoted to experimental design. It consists of four sections. Section 5.1 discusses the specification of customer demand and customer service. Section 5.2 describes the simulation model and experimental factors we considered. Section 5.3 gives the ranges of the experimental factors we investigated and describes the design of the experiments. Finally, Section 5.4 explains how we implemented and validated the simulation.

5.1 CUSTOMER DEMAND AND SERVICE CRITERION

In the preparation of this dissertation study, we collected some data from pharmaceutical companies. To simulate customer demand, one possibility was to construct an empirical distribution based on the data collected. This approach would resemble the demand pattern observed in the real world, but would not have allowed random demand to be generated beyond the range of the observed data. This would have been unfortunate because a very large demand could have a significant impact on the disposition of a simulation run. For example, a very large demand at one retail store and/or a very small demand at another can cause extremely "unbalanced" inventories at the retail level. Being unable to observe the system behavior in response to the occurrence of the extremely "unbalanced" retail inventories could have reduced the effectiveness of our simulation model to differentiate among alternative inventory positioning strategies. In order to allow values larger than the largest observation to be generated, we could have appended an exponential distribution to the right side of the empirical distribution as suggested by Bratley, Fax, and Schrage (1987). But in doing so, we would have had to invoke additional assumptions. Therefore, we decided not to generate demand based on an empirical distribution but rather to search for a theoretical distribution that would reflect the empirical data.

In our search for a theoretical distribution, we followed two guidelines. First, we wanted it to closely resemble demand distributions frequently observed in practice. Second, we wanted it to allow demand uncertainty to change in a wide range. (We measured the demand uncertainty by the standard deviation to mean ratio, called the coefficient of variation, or simply CV.)

In practice, many distributions are skewed to the right and have a density function with a shape similar to that of Figure 5.1 (Law and Kelton, 1991). The empirical data we collected from the pharmaceutical industry supported this observation. The data we collected have CVs ranging from less than one to slightly larger than one. The empirical data Jackson (1988) collected had a CV as high as 2 and 3. Johnson, Davis, and Waller (1998) have also observed a wide range of CV. They report that "At Hewlett-Packard, we often see product CVs of 1.0 and sometimes as high as 2.0."

We considered several theoretical distributions. The normal distribution was a candidate because it could have an arbitrary CV. But the normal distribution is a poor choice for representing customer demand because it can generate negative value for the demand. Indeed, for representing the distribution of customer demand, Jackson (1988) has pointed out that the normal distribution often begins to lose its validity for CV > 0.3. For Gamma and Weibull distributions, the CV is greater than, equal to, or less than 1, when their shape parameter α is less than, equal to, or greater than 1, respectively. However, these distributions will have a shape similar to the density function as shown in Figure 5.1 only when $\alpha > 1$, which implies CV < 1. They do not fit our second guideline (to allow demand uncertainty to vary over a wide range) for selecting the demand distribution. Fortunately, the lognormal distribution always has a density function with a shape similar to that in Figure 5.1, but its CV can be any positive real number. Since the lognormal distribution fits both of our guidelines, we selected it to generate demand for our investigation of the inventory positioning problem.

The lognormal distribution can be specified by two parameters, scale parameter μ and shape parameter σ, denoted by NL (μ, σ^2). Its density function is:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\frac{-(\ln x-\mu)^2}{2\sigma^2}, \quad \text{if } x>0$$

$$= 0, \qquad\qquad \text{otherwise}$$

Where the shape parameter $\sigma > 0$, the scale parameter $\mu \in (-\infty, \infty)$.

The mean of the lognormal distribution is $e^{\mu+\frac{\sigma^2}{2}}$ and its variance is $e^{2\mu+\sigma^2}(e^{\sigma^2}-1)$. Notice that its CV, $\sqrt{(e^{\sigma^2}-1)}$, does not depend on the scale parameter μ and that $X \sim$ NL (μ, σ^2) if and only if $\ln X \sim$ N (μ, σ^2), where N (μ, σ^2) is normal distribution with mean μ and variance σ^2.

The lognormal distribution is, in at least one important respect, a more realistic representation of the distribution of characteristics like demand, weight, and size than is the normal distribution. These quantities cannot take negative values, but a normal distribution ascribes positive probability to such events, while the lognormal does not. Furthermore, by taking small enough σ, it is possible to construct a lognormal distribution closely resembling any normal distribution. Hence, even if a normal distribution is felt to be really appropriate, it might be replaced by a suitable lognormal distribution.

The fields of application of the lognormal distribution are very broad. Gilbrat (1930, 1931) found the distribution useful in representing the distribution of size for various kinds of "natural" economic units. Magee, Copacino, and Rosenfield (1985) observe that the demand in numerous industries and businesses follows a pattern represented by the lognormal distribution. Indeed the empirical data we collected from the pharmaceutical industry closely fits a lognormal distribution as well. For applications of the lognormal distribution in other fields, see Johnson, Kotz, and Balakrishnan (1994).

The lognormal distribution is known for its mathematical complexity. That's probably part of the reason why it has not been widely used in the inventory research literature. For example, we know that the convolution (i.e., the sum of independent random variables) of normal distributions is still a normal distribution. But the convolution of lognormal distributions is no longer a lognormal distribution. This fact certainly would complicate analytical approaches to inventory problems. Fortunately, mathematical complexity was not of concern to us in this study because our primary research tool was simulation.

The lognormal is a continuous distribution. The demand it generates is a real number, rather than an integer. Since we were measuring demand in units, we rounded demand simulated for each of the retail stores in each period to an integer. If there was not enough stock at a store to satisfy the customer demand, the shortage was backordered at the store.

There are different criteria available to measure the customer service level. One is fill-rate, defined as the average fraction of demand satisfied, or "filled," directly from the stock on hand. Other service criteria include the time-weighted average number of backorders, called expected backorders, and the demand-weighted average time that a customer must wait before his or her demand is filled from on-hand inventory, called expected delay. It is easily shown that expected delay = (expected backorders) / (expected demand). In this study, we used the fill- rate as our service criterion.

5.2 SIMULATION MODEL AND EXPERIMENTAL FACTORS

The distribution system we considered is a one-warehouse multi-retailer distribution system. Lateral transshipments between retail stores were not allowed. No shipments directly from the outside supplier to retail stores. Returns from the retail stores to the central warehouse were not allowed either. It was also assumed that there was no cost advantage to holding inventory at any particular stocking location, either at the warehouse or at any retail store. The outside supplier was assumed to be completely reliable, always able to fill orders placed by the warehouse in any quantity and at any time without delay. Transit lead times from the outside supplier to the warehouse and from the warehouse to the retail stores were deterministic and known in advance. All the retail stores were identical in their demands, transit lead times, decision rules and the information upon which the decision rules were implemented. Customer demand occurred only at the retail level. The distribution system was operated in a periodic review environment.

The justification for choosing a one-warehouse multi-retailer distribution system was twofold. First, it is the most frequently observed multi-echelon distribution system in industry (Chain Store Age Executive, April 1992). Second, under the assumption that there was no inventory holding cost advantage, it is the simplest multi-echelon distribution system where the positioning problem exists.

For our simulation model, we assumed that stocks were received at the beginning of a period, while demand occurred later during the period. Inventory status before ordering was calculated after the demand in that period had occurred. Replenishment decisions were made based on the inventory status calculated. If there was not enough stock available to satisfy customer demand, the shortages were backordered at the retail stores. On the other hand, if the warehouse was unable to fill replenishment orders from the retailers in full, no shortages were backordered. The warehouse followed the balancing allocation principle to allocate the available stock to these retail stores.

The outside supplier takes L_w periods to move stock physically to the warehouse. The warehouse takes L_r periods to deliver stock to any of the retail stores. If $L_r = 0$, then stock allocated in period t would be available to satisfy the customer demand in period t+1, rather than in period t because the order and, therefore, the allocation in period t were triggered after demand in that period occurred. In general, stock allocated in period t would be available to satisfy customer demand in period t + L_r +1. The same logic applied to the transit lead time L_w as well.

Let $N_j(t)$ be the net inventory of the jth store at the beginning of period t after receiving the stocks allocated by the warehouse in period t-L_r-1 but before the customer demand at that store in period t, $D_j(t)$, occurred. Then unsatisfied demand in period t at the store j, denoted by $UF_j(t)$, was calculated as follows.

$$UF_j(t) = -MIN\{MAX(0, N_j(t)) - D_j(t), 0\}$$

$MAX(0, N_j(t))$ is the on hand inventory before the demand occurred. It is worth pointing out that the calculation of unsatisfied customer demand was based on the on-hand inventory, rather than the net inventory. If no correction, $MAX(0, N_j(t))$, was made for a possible shortage at the beginning of the period, the calculation would lead to an overestimation of the unfilled customer demand. For a detailed discussion, refer to de Kok (1990).

Let N be the total number of the retail stores served by the warehouse. Let t' and t' + T' be the first and the last period in which simulation data were collected respectively. The system-wide fill rate was calculated as follows.

$$F = 1 - \left[\frac{\sum_{j=1}^{N} \sum_{t=t'}^{t'+T'} UFj(t)}{\sum_{j=1}^{N} \sum_{t=t'}^{t'+T'} Dj(t)} \right]$$

We recorded period-ending on-hand inventory at the warehouse and at the retail stores, respectively. At the end of the period, we also recorded backorder levels and unsatisfied customer demand at the individual retail stores. We aggregated on-hand inventory at individual stocking locations to provide a system-wide inventory, that is, on-hand inventory at the warehouse and all the retail stores, denoted by INV. The system workload was measured by the shipment frequencies to the warehouse and to the retail stores, denoted by SHIPW and SHIPR, respectively. SHIPW = (Number of shipments to the warehouse from period t' to period t'+T')/(T'+1), that is, the average number of shipments to the warehouse per period. Similarly, SHIPR = the average number of shipments to the retail level per store per period , defined as (Total number of shipments to the retail stores from period t' to period t' + T') / [(T'+1)N]. (For simplicity, in the remainder of this dissertation, we would describe SHIPR as "shipments per period," instead of "shipments per store per period.)

In our simulation experiments, we considered six experimental factors, including three resource factors (system-wide inventory level, shipment frequencies, and inventory control system) and three environmental factors (demand uncertainty, the transit lead times, and the number of retail stores), as listed in Table 5.1.

Table 5.1 **EXPERIMENTAL FACTORS**

RESOURCE FACTORS	ENVIRONMENTAL FACTORS
System-wide Inventory Level	Demand uncertainty
Shipment Frequencies	Transit lead times
Inventory Control System	The number of the retail stores

The inventory positioning ratio, denoted by P, is the average inventory at the retail stores divided by the average system-wide inventory level. The actual inventory positioning ratios, average inventory levels, and shipment frequencies are all ex post facto measures. In each of our simulation experiments, we simulated alternative inventory positioning strategies by manipulating control system parameters to change the positioning ratio P, while keeping all the experimental factors unchanged.

5.3 DESIGN OF EXPERIMENTS

The simulation experiments we conducted consisted of two parts. First there was a baseline study in which alternative positioning strategies were simulated and evaluated for a given set of inventory, transportation, and control system resources. The second part was devoted to the investigation of the effects of changing experimental factors.

5.3.1 BASELINE STUDY

For this baseline study, we considered a distribution system consisting of one warehouse and eight identical retail stores where external customer demand occurred. We simulated the demand for each retail store using a lognormal distribution with the scale parameter $\mu = 1.125$ and the shape parameter $\sigma = 1$. These parameters provided a mean demand of 5.08 units per period, a standard deviation of 6.657 units per period, and a coefficient variation (i.e., the standard deviation of the demand divided by the mean of the demand) of 1.31.

We set the transit lead times from the outside supplier to the warehouse (L_w) and from the warehouse to any of the retail stores (Lr) at 1 period each. We fixed the average system-wide inventory level (i.e., the sum of the on-hand inventory levels at the retail stores and the warehouse per period) at 150 units. This level was chosen based on a pilot simulation run, which indicated that the potential benefit from holding some inventory at the warehouse was relatively larger at this level than at any other levels we tested.

The shipment frequencies from outside supplier to the warehouse (SHIPW) and from the warehouse to each of the retail stores (SHIPR) were set at 0.1667 shipments per period (i.e., 1 shipment every 6 periods) and 0.2315 shipments per period (i.e., 1 shipment every 4.32 periods), respectively. The shipment frequency from the outside supplier to the warehouse was held to 0.1667 shipments per period, or one shipment every 6 periods. 0.1667 shipments per period were chosen to accommodate the "Push" system in which the warehouse replenished its inventory once every six periods. Of course, we could have redesigned the "Push" system by changing the warehouse's review cycle, R_w. But for simplicity, throughout our simulation experiments we kept $R_w = 6$.

For the baseline study, we assumed that the distribution system was controlled by Local Information Based Reorder Point and Order-up-to Level System (LROP). With this control system, both the retail stores and the warehouse used a (R, s, S) replenishment decision rule with $R = 1$. It works as follows: There is a review of inventory status every period. If the nominal inventory position is less than or equal to s, an order is placed to bring the inventory status up to level S; if not, nothing is done until the next review period. When the warehouse is unable to fill all outstanding retail orders (i.e., stockout), the available inventory at the warehouse was allocated to minimize the maximum unsatisfied portion of the retailer' orders, which is equivalent to equalizing the probabilities of stockout as much as possible among those retail stores that had placed orders in the current period. There is no communication between the warehouse and the retail stores beside the retailers' replenishment orders placed on the warehouse. The warehouse shortage information was unknown to the retailers until they physically received the shipments from the warehouse. While shortages at the retail level were backordered, shortages at the warehouse are not backordered. Because stock in transit was unknown to the retail stores, the retailers monitored their inventory status by nominal inventory position, defined as on-hand inventory minus backorders plus stock on order.

We simulated alternative positioning strategies that shared the same system workload. Using the search procedures described in Section 4.2, we identified the positioning ratios and fill rates for the best positioning, the positioning "close" to the customer, and the "Ship-all." The baseline parameter settings are listed in Table 5.2.

Table 5.2 **BASELINE PARAMETER SETTING**

Experimental Factor	**Parameter Value for Baseline Study**
System-wide Inventory Level	INV = 150 units per period
Shipment Frequency to the Warehouse	SHIPW# = 0.1667 shipment per period
Shipment Frequency to the Retail Stores	SHIPR# = 0.2315 shipment per period
The Coefficient of Variation of Demand	CV= 1.31
Transit Lead Time to the Warehouse	L_w = 1 period
Transit Lead Time to any of the stores	Lr = 1 period
The number of stores	N = 8

5.3.2 CHANGING FACTORS

Next we investigated the effect of the experimental factors. The purpose of this investigation was to test the robustness of our baseline study conclusions and to gain clear insights on the effect of the experimental factors on the positioning of inventory in the distribution system. We investigated changes in the three resource factors as well as the three environmental factors. The ranges and the levels of the experimental factors we studied appear in summary form in Table 5.3.

From Table 5.3, we see that there are 19440 possible combinations. For each, three positioning strategies need to be identified. This implies that potentially we have to search for 58320 inventory positioning strategies. Because of the time-consuming nature of the simulation runs, we concluded that it was not practical to have a full factorial design. Even a 2^k fractional factorial design was still too large to implement. Furthermore, the 2^k fractional factorial design would have limited observation to only 2 levels for each experimental factor. If the effect of any experimental factor had a non-monotonic relationship, observations based on only two levels would have led us to draw false conclusions.

Table 5.3 **CHANGING FACTORS**

Factors	Range	Levels	# of levels
INV	70 to 400	70, 100, 150, 200, 250, 300, 350, 400, 84, 103, 120, 203.	12
SHIPR#	0.1667 to 0.232	0.1667, 0.232, 0.3	3
SHIPW#	0.1667	0.1667	1
Control System	Categorical	LROP, EROP, DRP, and "Push" system	4
CV	0.53 to 1.8	0.53, 0.8. 1.0, 1.31, 1.8	5
Lr	0 to 2	0, 1, 2	3
L_w	0 to 2	0, 1, 2	3
N	4 to 32	4, 8, 32	3

The main effect of a factor is defined as the change in response produced by a change in the level of the factor. Note that there could be a case in which the effect of factor A depended on the level of factor B. When this occurs, there is an interaction between the two factors A and B. A significant interaction can mask the significance of a main effect. To capture these interactions a full factorial design is often needed.

In our investigation, we focused our attention primarily on the main effects for practical reasons. However, we screened the possible two-way interactions by plotting the results against the levels of factor A for different levels of factor B. If the B lines at different levels are roughly parallel, this would indicate that factors A and B did not interact significantly (See Figure 5.2). If they were not parallel, that would have indicated the interactions between factors A and B. (See Figure 5.3). If these graphs indicate possible interactions, additional experiments might have to be conducted before we can draw any meaningful conclusions. The method we used to deal with observed interactions was factor specific. Therefore, we have left the detailed description of these analyses to Chapter 6 where our simulation results are presented and analyzed.

5.4. SIMULATION IMPLEMNTATION AND VILIDATION

In order to represent faithfully the specific, detailed logic of the inventory control systems we designed, we wrote our entire simulation program in a general-purpose language, FORTRAN, rather than a more high-level simulation language, such as SIMSCRIPT II.5, GPSS, or SIMAN. Customer demand was simulated using IMSL's FORTRAN subroutine RNLNL. The program was run on a VAX6620 running VM5.5.

The simulation modeled the behavior of the distribution system on a period by period basis. We truncated data collection after the first 999 periods to mitigate start-up effects. We collected our data from period 1000 to period 5000, with 40 replications. We facilitated statistical comparisons between positioning strategies by using the well-established approach of common random numbers. That is, we used a common random number generator for all experimental conditions with different seeds for each replication. We constructed 95% confidence intervals for both the inventory positioning ratio and the customer-fill rate observed. When we needed to differentiate the results further, we conducted paired t-tests using the 40 replications.

The validation of the simulation model was relatively straightforward. We evaluated each of the retail store's mean and the standard deviation of simulated demand. We were also able to print out detailed information about the distribution system's operations at any time to check whether or not it had operated the way we designed it to. Our tests indicated that our simulation model performed as expected.

CHAPTER 6
RESULTS AND ANALYSIS

We obtained very consistent inventory positioning results for a one-warehouse multi-retailer distribution system with stochastic demand. For all our experiments, in no case was the positioning ratio (i.e., the proportion of inventory positioned at the retail level) that maximized the fill-rate less than 0.8 and most often the positioning ratio was substantially higher. The increase in the fill-rate achieved by holding some inventory at the warehouse, as opposed to having it as close to the customer as possible given a set of inventory, transportation, and control system resources, was fairly small. It never exceeded 1.5 percentage points and in most cases was less than 1 percentage point. The results indicate that a majority of the inventory should be near the customer to get high levels of customer fill-rate. As long as fill-rate is the appropriate service criterion, the preference for holding most inventory near the customer held through changes in all experimental factors we investigated. These factors include three resource factors (system-wide inventory level, predetermined shipment frequency, and inventory control system used) and three environmental factors (demand uncertainty, transit lead times, and the number of retail stores).

To present our simulation results, we have divided this chapter, Chapter 6, into eight sections. Presentation of the results for the baseline study precedes the results for changes in the six experimental factors: Effect of Inventory Control Systems, Effect of System-wide Inventory Level, Effect of Demand Uncertainty, Effect of Transit Lead Times, and Effect of the Number of Retail Stores. (Note: The numerical data presented in this chapter are statistically significant at 0.01 level using paired t-tests. If not, the data is denoted by #.) We end this chapter with a summary of the main results.

6.1 BASELINE STUDY

We ran a baseline study to determine where inventory should be positioned in a distribution system so as to get the highest level of customer fill-rate (i.e., the percentage of demand that is satisfied immediately from inventory), given a set of inventory, transportation, and control system resources. The baseline parameter settings are described in Table 5.2 in the previous chapter, Chapter 5.

We defined inventory positioning by a positioning ratio, P, the average on-hand inventory level at the retail stores divided by the sum of the average inventory level at both the retail stores and the warehouse. Clearly, $0 \leq P \leq 1$. Note that the actual inventory positioning ratios, average inventory levels, and shipment frequencies were all ex post facto measures. By manipulating the control system parameters (s_r and S_r, reorder point and order-up-to level used by each of the retail stores, and s_w and S_w, reorder point and order-up-to level used by the warehouse), we changed the positioning ratio P while keeping all the other experimental factors unchanged. We recorded the average fill-rate (across 40 replications) and the average positioning ratio (across 40 replications) associated with the three different positioning strategies. The fill rates in response to the changes in the positioning ratio are plotted in Figure 6.1 for this baseline case.

The curve shown in Figure 6.1 represents a set of alternative positioning strategies that shared the same set of inventory, transportation, and control system resources. We were able to draw four important observations from this figure. First, the customer fill-rate does depend on where inventory is positioned in the distribution system. When the positioning ratio dropped from 0.8424 to 0.2080, the customer fill-rate deteriorated from 84.3% to 53.1%, showing that the choice of positioning ratio makes a substantial difference. Second, the positioning ratio that achieved the maximum fill rate was 0.8424, indicating that the major portion of the system-wide inventory should be positioned at the retail level near the customer. Third, the positioning strategies in the neighborhood of the maximum fill-rate had very similar profiles. Fourth, the increase in the fill-rate by holding some inventory at the warehouse, as opposed to positioning inventory as close to the customer as possible given the resource constraints, was fairly small. The observed fill rate only improved by 1.47% from 81.93% to 83.4%.

Having examined the pattern of response to the changes in the positioning ratio, we now focus our attention on three specific positioning strategies: the best positioning, the positioning "close" to the customer, and the "Ship-all" positioning. As defined in Chapter 3, Section 3.3, the best positioning is the one that provides the maximum fill-rate. The positioning close to the customer is the one that has a maximum positioning ratio but does not violate the retail shipment frequency constant. The "Ship-all" positioning is one in which the warehouse allocates all incoming stocks to the retail stores immediately. Because there is no stock left at the warehouse, the positioning ratio for the "Ship-all" is equal to 1.0.

We earlier denoted the positioning ratios and the fill-rates associated with the "best" positioning, the positioning "close" to the customer, and the "Ship-all" positioning by (P*, F*), (Pc, Fc), and (Ps, Fs), respectively. We obtained the solutions for the above three problems by searching over the space of the control system parameters via simulation. The search procedures we used are described in Chapter 4, Section 4.2. The results, including the values of the control system parameters and the average values (across 40 replications) and 95% confidence intervals for the positioning ratios and the fill-rates are presented in Table 6.1.

Table 6.1. **THREE POSITIONING STRATEGIES IDENTIFIED FOR BASELINE STUDY**

Strategy	(s_r, S_r, s_w, S_w)	Positioning Ratio P (95% Conf. Interval)	Fill-Rate F (95% Conf. Interval)
"Best"	(24, 35, -55, 159)	0.8424 ± 0.0006	0.8340 ± 0.0015
"Close"	(46, 48, -150, 67)	0.9697 ± 0.0002	0.8193 ± 0.0013
"Ship-All"	(60, 61, -215, 0)	1 ± 0.0	0.7958 ± 0.0016

It is of note that the best positioning prescribed a negative reorder point for the warehouse (i.e., $s_w = -55$). This observation provided additional evidence that the warehouse held relatively little on-hand inventory. Note that there is no conflict between the negative reorder point and the assumption that no backorder is allowed at the warehouse. The negative reorder point, $s_w = -55$, simply means that no order would be triggered unless the on-hand inventory at the warehouse minus the orders received from the retailers plus the stock on order from the outside supplier reached a level lower than -55 units. For detailed discussion, see Chapter 3, Section 3.2.1.

By definition, the warehouse safety stock is the difference between the warehouse's reorder point and the average demand from the retailers during the lead time from the outside supplier to the warehouse ($L_w = 1$) plus the review time ($R_w = 1$). With a negative reorder point, the warehouse must have kept a negative safety stock. That the warehouse should keep a negative safety stock to get the highest level of customer fill rate was first reported by Schwarz, Deuermeyer, and Badinelli (1984). As reviewed in Chapter 2, the control system Schwarz, Deuermeyer, and Badinelli (1984) studied is one in which the warehouse implements the FCFS allocation rule and shortages at the warehouse are backordered. We believe that the results shown in Table 6.1 provide the first evidence that the warehouse should keep a negative safety stock to get the maximum fill-rate, even if the balancing allocation rule is used at the warehouse and no warehouse shortages are backordered. The conclusion that the warehouse should keep a negative safety stock to get the highest level of customer fill rate seems to apply more broadly than Schwarz, Deuermeyer, and Badinelli (1984) have suggested.

To analyze the data presented in Table 6.1 more systematically, we decomposed the maximum fill- rate into three components: the fill-rate for the "Ship-all" strategy (Fs); the difference between the fill-rate for the "close" positioning (Fc) and the fill rate for the "Ship-all" (Fs); and finally the difference between the fill-rate for the best positioning (F*) and the fill rate for the "close" positioning (Fc). Clearly,

$$\text{Fs} + (\text{Fc} - \text{Fs}) + (\text{F*} - \text{Fc}) = \text{F*} \qquad (6.4)$$

A discussion of each of the components expressed in the Equation (6.4) is warranted. Fs is the fill rate that one would be expected if no inventory were held at the warehouse. The positioning "close" to the customer and the "Ship-all" positioning correspond to the problems of maximizing positioning ratio, P, using the same inventory resource but subject to different constraints on the shipment frequency to the retail stores. Specially, the warehouse on average sends 0.2315 shipments per period to each of the retail stores for positioning inventory "close" to the customer. For the "Ship-all" strategy, the warehouse cannot send shipments to any of the retail stores more frequently than it receives the coming stocks from the outside supplier, on average 0.1667 shipments per period. Their fill-rate difference (Fc-Fs) is, therefore, primarily due to their difference in the shipment frequency SHIPR. For this reason, we call (Fc-Fs) the fill rate improvement due to the increase in the shipment frequency (relative to the "Ship-all").

Next, let us look at the difference between the fill rate for the best positioning and that for positioning "close" to the customer, (F*-Fc). While the positioning close to the customer seeks the maximum positioning ratio without violating the constraint on the retail predetermined shipment frequency, the best positioning tries to maximize the customer fill-rate subject to the same resource constraints. Because these two positioning strategies use the same inventory, transportation, and control system resources, their fill rate difference (F*-Fc) is, therefore, attributable to their difference in positioning strategy only. For this reason, in the remainder of this dissertation, we call (F*-Fc) the fill rate improvement due to the best positioning, which also represents the penalty for positioning inventory close to the customer, as opposed to pursuing the best positioning.

Note that the fill rate improvement due to the increased in SHIPR is relative to the "Ship-all." (i.e., Fc-Fs). On the other hand, the fill rate improvement due to the best positioning (i.e., F*-Fc) is relative to the fill rate for the positioning inventory close to the customer, Fc. Please note that while we have listed the numerical results in the tables in decimal, we interpret the fill rate improvement by percentage points. For instance, if the fill rate increased from 50% to 60%, we say that fill rate increased by 10% from 50% to 60%, where 10% means 10 percentage points, rather than 10 percent of the initial fill rate.

We summarize the data contained in Table 6.1 in a more informative format in Table 6.2. By Equation (6.4), we know that the sum of the columns 1, 2, and 3 is equal to the column 4. P* appearing in column 5 is the positioning ratio that provided the maximum fill rate F* appearing in the column 4.

To visually present the data appearing in Table 6.2, we first plotted the fill-rate for the

"Ship-all," then added the fill-rate improvement due to the increase in the shipment frequency, and finally attached the additional fill-rate improvement due to the best positioning. We also recorded the best positioning ratio. Figure 6.2 summarizes the results presented in Figure 6.1.

Table 6.2 **BEST POSTIONING RATIO, MAXIMUM FILL-RATE AND ITS DECOMPOSITION**

Fs	Fc – Fs	F* - Fc	F*	P*
0.7958	0.0235	0.0147	0.8340	0.8424

Now, we analyze the results of the baseline study presented in Figure 6.2: Given a set of inventory, transportation, and control system resources, the "Ship-all" positioning did not fully utilize the available transportation resource but still provided a 79.58% customer fill-rate. When the frequency of shipment to the stores SHIPR increased from no greater than 0.1667 shipments per period to the predetermined shipment frequency, 0.2315 shipments per period, the fill-rate increased by 2.35 % from 79.58% to 81.93%. This is the strategy for positioning "close" to the customer. The best positioning strategy uses a positioning ratio of 0.8424 (i.e., positioning 84.24 % of system-wide inventory at the retail stores). The customer fill-rate is increased by an additional 1.47% from 81.93% to 83.4%, the maximum fill-rate, over the "close" positioning strategy..

The conclusions we draw from the baseline study are as follows: the correct positioning of inventory is very important to achieving high levels of customer fill-rate (see Figure 6.1). While there is a benefit from holding some inventory at the warehouse, the penalty for positioning more inventory near the customer (e.g., the observed fill rate improvement due to the best positioning) is fairly small. Since the positioning ratio that provided the maximum fill-rate is close to 1 (P*=0.8424) and the positioning strategies in the neighborhood of the maximum fill-rate have very similar profiles (see Figure 6.1), we conclude that inventory should be positioned near the customer to get high levels of customer fill-rate.

Eppen and Schrage (1981) coined the term "depot effect" for the benefit from holding some inventory at the warehouse, as opposed to using the "Ship-all." Many published studied on the positioning problem directly compare the maximum fill-rate, F*, with the fill-rate for the "Ship-all," Fs, without taking their difference in the shipment frequency into account. If we had followed this traditional approach, then we would report that by holding 15.76% (i.e., 1- P*=1-0.8424=0.1576) of the system-wide inventory at the warehouse, the customer fill-rate increased by 3.82 % from 79.58% to 83.4% (F*-Fs=0.0382). This would have been misleading, however, because it would have ignored the fact that 61.42% of the reported "depot effect" was actually due to an increase in the frequency of shipments to the retail stores [i.e., (Fc-Fs)/(F*-Fs)=0.0235/0.382=61.42%]. Also we must note that the warehouse could hold the same amount of inventory with very different shipment frequencies (For a detailed discussion, see Section 6.6). Yet, in many published studies, the shipment frequencies, or fixed shipment costs, have been ignored. It is unclear how much of the observed "depot effect" for holding inventory at the warehouse was actually due to the increase in the shipment frequency, SHIPR. Recording the fill-rate for the "Ship-all" and the fill-rate improvement due to the increase in the shipment frequency enabled us to answer this question clearly and precisely.

We devote the remainder of this chapter to showing that the general finding from the baseline study would not change materially with any of the experimental factors we considered. Changes in any of the experimental factors bring changes in the best positioning ratio P* as well as in the maximum fill rate F*. We denoted these changes by ΔP* and ΔF*, respectively. $\Delta P^* > 0$ indicated some inventory had been shifted from the warehouse to the retail stores. $\Delta F^* > 0$ indicated that the change in that experimental factor improved customer fill-rate. By Equation (6.4), we can decompose the change in the maximum fill rate (ΔF*) into three components: the changes in the fill rate for the "Ship-all" ΔFs, the change in the fill rate improvement due to the increase in the shipment frequency Δ(Fc –Fs), and the change in the fill rate improvement due to the best positioning Δ (F*-Fc). The relationships among those changes are as follows.

$$\Delta Fs + \Delta\ (Fc - Fs) + \Delta\ (F^*\text{-}Fc) = \Delta F^* \qquad (6.5)$$

The baseline study has provided us with a set of values for P*, F*, (F*-Fc), (Fc-Fs), and Fs. Because the best positioning ratio P* was high and the fill-rate improvement due to the best positioning over having inventory "close" to the customer (F*-Fc) was small, we have concluded that inventory should be positioned near the customer to get the high levels of customer fill-rate. The robustness of this conclusion critically depends on the signs and magnitudes of Δ P* and Δ (F*-Fc) in response to the changes in the experimental factors we investigated. The results of our investigation on the robustness of our findings in the baseline study are presented in the next six sections (Sections 6.2-6.7).

6.2. EFFECT OF INVENTORY CONTROL SYSTEMS

One of the key features of this dissertation study is that it treated the inventory control system as an experimental factor. Such an experiment was designed to address an important question: Would the choice of control system change the best positioning of inventory in the distribution system, and if so, how?

To study the effect of the inventory control system on the positioning of inventory in the distribution system, we ran simulation experiments for all control systems using the same baseline parameter settings. The new control systems included were Echelon Based Reorder Point and Order-up-to Level Control system (EROP), Distribution Requirements Planning System (DRP), and the "Push" system. These three new control systems required sharing information between the warehouse and the retail stores in various ways. Information sharing refers to the exchange of information between the warehouse and the retail stores in addition to issuing replenishment orders. For instance, the warehouse may use ASNs to tell the retailers how much stock is on the way before the shipments physically reach at the retail stores. The retailers may also share their inventory status information with the warehouse so that the warehouse can make better allocation decisions. A detailed description of these new control systems, including their information-sharing schemes, was provided in Chapter 4, Section 4.1.5.

The simulation results in conjunction with the results of the baseline study where LROP was used are presented in Figure 6.3 and Figure 6.4. Let us look at Figure 6.3 first. There are four curves, each representing a set of alternative positioning strategies for one of the control systems we investigated. The results show that all control systems achieve high levels of fill-rate at high levels of the positioning ratio. At low levels of the positioning ratio, the "Push" system stands out. It provides much lower fill-rates than those provided by other control systems. At the higher levels of the positioning ratio, all control systems provide high levels of fill-rate, although EROP and DRP dominate the other control systems. When the positioning ratio P>0.9, all three control systems with information sharing, EROP, DRP, and the "Push" system, dominate LROP that requires local information only. The differences among the maximum fill-rates are statistically significant at least at a 0.01 level using a paired t-test, except between EROP and DRP. With the exception of the "Push" system, however, the patterns of response to changes in the positioning ratio are very similar for all control systems.

It is of note that LROP, EROP, and DRP share the same form of decision rules (s_r, S_r, s_w, S_w) where s_r and S_r, s_w and S_w, are reorder point and order-up-to level used by each of the retail stores and by the warehouse, respectively. The changes from LROP to EROP or DRP were, therefore, primarily due to changes in the information requirement. The change from LROP to the "Push" system was more complicated because those two control systems differed not only in their information requirements but also in the form of the decision rules used. As explained in Section 4.2, the "Push" system uses decision rules that are characterized by different parameters (α, T, R_w, S_w) where α specifies the percentage of the incoming stocks at the warehouse that would be allocated to the stores immediately and the remaining stocks would be allocated T periods later but before reaching the end of the warehouse replenishment cycle R_w. The difference in the form of decision rules may explain why the "Push" system behavior somewhat differently.

For each of the control systems, we identified three positioning strategies: the best positioning, the positioning "close" to the customer, and the "Ship-all." The values of the control system parameters, the average values (across 40 replications) and 95% confidence intervals for the positioning ratios and the fill-rates corresponding to each of these positioning strategies are shown in Table 6.3.

For each of the control systems we simulated, we recorded the best positioning ratio as well as the fill-rate provided by "Ship-all." We also calculated the fill-rate improvement due to the increase in the shipment frequency and the fill-rate improvement due to the best positioning. By Equation (6.4), the sum of the fill-rate for the "Ship-all" and the fill-rate improvements gave the maximum fill-rate. The results presented in Figure 6.4 show that the changes in inventory control system have changed the best positioning ratio and the maximum fill-rate, as also do the fill-rate for the "Ship-all" and the fill-rate improvements, due both to the increase in the shipment frequency and the best positioning.

Table 6.3. **POSITIONING STRATEGIES ASSOCIATED WITH FOUR INVENTORY CONTROL SYSTEMS**

Control System	Strategy	(s_r, S_r, s_w, S_w) or (α, T, R_w, S_w)	Positioning Ratio P (95% Conf. Interval)	Fill-rate F (95% Conf. Interval)
LROP	"Best"	(24, 35, -55, 159)	0.8424 ± 0.0006	0.8340 ± 0.0015
	"Close"	(46, 48, -150, 67)	0.9697 ± 0.0002	0.8193 ± 0.0013
	"Ship-All"	(60, 61, -215, 0)	1 ± 0.0	0.7959 ± 0.0016
EROP	"Best"	(24, 36, 180, 399)	0.9218 ± 0.0005	0.8615 ± 0.0013
	"Close"	(35, 38, 178, 397)	0.9769 ± 0.0002	0.8527 ± 0.0013
	"Ship-All"	(100, 100, 177, 396)	1 ± 0.0	0.8455 ± 0.0013
DRP	"Best"	(14, 26, -152, 68)	0.9208 ± 0.0005	0.8615 ± 0.0014
	"Close"	(15, 28, -206, 13)	0.9769 ± 0.0002	0.8527 ± 0.0013
	"Ship-All"	(32, 35, -263, -44)	1 ± 0.0	0.8455 ± 0.0013
"Push"	"Best"	(0.9154, 2, 6, 400)	0.9531 ± 0.0002	0.8481 ± 0.0018
	"Close"	(0.94, 1, 6, 398)	0.9832 ± 0.0001	0.8453 ± 0.0018
	"Ship-All"	(1.0, 0, 6, 397)	1 ± 0.0	0.8373 ± 0.0018

By design, the three new control systems utilized more information than LROP did because they all required sharing information between the warehouse and the retail stores in one way or another. With information-sharing, better allocation decisions could be made at the warehouse, and, therefore, one might have expected that a larger proportion of the system-wide inventory ought to have been positioned at the warehouse. Surprisingly, our simulation results show that the intuitive feelings in this case were quite wrong; the change from LROP which used local information only to any of the new control systems with information sharing would require not more but less inventory to be held back at the warehouse.

As an example, see in Table 6.3 and Figure 6.4, the change from LROP to EROP increases the best positioning ratio from 0.8424 to 0.9218, shifting 7.94% of system-wide inventory from the warehouse to the retail stores. While the maximum fill-rate increases by 2.75% from 83.4% to 86.15%, the fill-rate improvement due to the best positioning decreased by 0.59% from 1.47% to 0.88%.

Very consistent results were obtained for the change from LROP to DRP. The change increased the best positioning ratio from 0.8424 to 0.9208, shifting 7.84% of system-wide inventory from the warehouse to the retail stores near the customer. The changes in the maximum fill-rate and in the fill-rate improvement due to the best positioning are statistically identical to these in the case of the change from LROP to EROP. (Note that in our baseline study EROP and DROP looked almost identical. But they were different in terms of how their warehouse replenishment orders were triggered. In the presence of time-phased projected demand, the fill rates of EROP and DRP were very different. These results are detailed in the last section of this Chapter, Section 6.8.)

TABLE 6.4 EFFECT OF CHANGES IN

INVENTORY CONTROL SYSTEM USED (RELATIVE TO LROP)

Changes	Δ Fs	Δ (Fc-Fs)	Δ (F*-Fc)	Δ F*	Δ P*
LROP to EROP	0.0496	-0.0162	-0.0059	0.0275	0.0794
LROP to DRP	0.0496	-0.0162	-0.0059	0.0275	0.0784
LROP to "Push"	0.0414	-0.0154	-0.0119	0.0141	0.1107
Average	0.0469	-0.0159	-0.0079	0.0230	0.0894

The report on the change from LROP to the "Push" system was very similar. The change increased the best positioning ratio from 0.8424 to 0.9531, shifting 11.07% of system-wide inventory from the warehouse to the retail stores. While the "Push" system shifted the largest percentage of inventory from the warehouse to the retail stores, its maximum fill-rate increased only by 1.41 % from 83.4% to 84.81%.

Table 6.4, Effect of Changes in Inventory Control System Used, summarizes the observed changes brought by switching from LROP to each of the three new control systems.

The results were very consistent. As more information becomes available, the best positioning ratio increases, on average (across the control systems) shifting 8.94% of the system-wide inventory from the warehouse to the retail stores near the customer. In the meantime, the fill rate improvement due to the best positioning as opposed to positioning inventory close to the customer decreased on average by 0.79 percentage points. In other words, information sharing reduced the penalty for the positioning close to the customer, rather than pursuing the best positioning, on average by 0.79 percentage points. While the decrease seemed to be small, it cut the penalty as we observed in the baseline study (i.e., an decrease in the fill rate by 1.47 percentage points as shown in Figure 6.4) more than in half! Note that information sharing also reduced the fill-rate improvement due to the increase in the shipment frequency SHIPR.

Next, let us look at the changes in the fill rate for the "Ship-all" strategy. As we switched the inventory control system from LROP to any of the three new control systems, EROP, DRP, or the "Push" system, the largest increases in the fill-rate was for the "Ship-all" positioning strategies. On average (across the control systems), the fill rate for the "Ship-all" increased by 4.69 percentage points. Offsetting the decreases in the fill-rate improvements due to the best positioning and due to the increase in the shipment frequency SHIPR, the new control systems still improved the maximum fill-rate on average by 2.3 percentage points. As shown in Table 6.4, in all cases, $\Delta F^* > 0$, indicating the information utilized by the control system has a positive impact on the system performance, the maximum fill rate.

The increase of the fill-rate for the "Ship-all" and the decrease of the fill-rate improvement due to the best positioning seem to indicate that as more information becomes available, the value for postponing allocation decisions by holding some inventory at the warehouse for more informed allocations diminishes. Since the best positioning ratio increased (i.e., $\Delta P^* > 0$), a larger proportion of inventory had been positioned at the retail stores near the customer.

To understand this interesting phenomenon, we conducted additional experiments designed to reveal the effect of incremental changes in transforming the control system used in the baseline study, LROP, to each of the these new control systems, EROP, DRP, and the "Push" system.

As discussed in Chapter 4 and shown in Table 4.1 (A Framework for Designing Multi-Echelon Inventory Control Systems), three incremental changes can transform LROP into

EROP. The first incremental change is to add an Advanced Shipment Notice (ASN) to LROP. ASN is an electronic message from the warehouse telling the retail stores what merchandise is on the way. With ASN, the retailers should know the amount of stock in transit immediately as the stock left the warehouse. Thus, the retailers were able to make their replenishment decisions based on inventory position (i.e., on-hand inventory minus backorders plus stock *in transit*), rather than on nominal inventory position (i.e., on-hand inventory minus backorders plus inventory *on order*) used by LROP. Using the terms defined in Section 4.1.4, we called this new control system an Advance Shipment Notice System (ASNS). For the next incremental change, the retailers share the information about their inventory positions with the warehouse. With this new information, the warehouse was able to make its allocation decision to balance inventories for *all* retail stores, not just for those that had placed an order in the current period. Still the warehouse made its replenishment decision based on its installation inventory position (i.e., inventory position at the warehouse). We called this control system the Reorder Point System (ROP). Finally, we had the warehouse make its replenishment decision based on its echelon inventory position (i.e., the sum of the inventory positions at the retail stores and the on hand inventory at the warehouse plus inventory in transit from the outside supplier). This transformed the ROP into the EROP. It is of note that the final incremental change only modified the way in which the information about retailers' inventory positions was used. No additional information resource was required.

Table 6.5. **EFFECT OF INFORMATION-SHARING**

Control System	(s_r, S_r, s_w, S_w)	**Positioning Ratio** P* **(95% Conf. Interval)**	**Fill-Rate** F* **(95% Conf. Interval)**
LROP	(24, 35, -55, 159)	0.8424 ± 0.0006	0.8340 ± 0.0015
ASNS	(24, 35, -76, 136)	0.9083 ± 0.0005	0.8596 ± 0.0013
ROP	(24, 36, -81, 130)	0.9197 ± 0.0006	0.8607 ± 0.0013
EROP	(24, 36, 180, 399)	0.9218 ± 0.0005	0.8615 ± 0.0013
DRP	(14, 26, -152, 68)	0.9208 ± 0.0005	0.8615 ± 0.0014

The change from LROP to DRP can be decomposed into incremental changes as well. The first was a change from LROP to ASNS. The second was a change from ASNS to ROP. These incremental changes were identical to those in transforming LROP to EROP. Finally, using DRP ordering logic at both the retail stores and at the warehouse would transform ROP into a DRP system.

We recorded the values of the control system parameters, the best positioning ratios, and the maximum fill-rates associated with LROP, ASNS, ROP, EROP, and DRP in Table 6. 5. The incremental changes in the maximum fill-rate Δ F* and the best positioning ratio Δ P* are summarized in Table 6.6, Effect of Incremental Changes in Transforming LROP to EROP and DRP.

Table 6.6 **EFECT OF INCREMENTAL CHANGES IN TRANSFORMING LROP TO EROP and DRP**

Incremental Changes From LROP to EROP	**Δ P***	**% of total Δ P***	**Δ F***	**% of total Δ F***
LROP to ASNS	0.0659	83.0	0.0256	93.1
ASNS to ROP	0.0114	14.4	0.0011	4.0
ROP to EROP	0.0021	2.6	0.0008	2.9
Total	0.0794	100	0.0275	100
Incremental Changes From LROP to DRP	**Δ P***	**% of total Δ P***	**Δ F***	**% of total Δ F***
LROP to ASNS	0.0659	84.1	0.0256	93.1
ASNS to ROP	0.0114	14.5	0.0011	4.0
ROP to DRP	0.0011	1.4	0.0008	2.9
Total	0.0784	100	0.0275	100

The effect of the incremental changes in the information-sharing scheme are remarkably consistent, showing that as more information was shared between the warehouse and the retail stores, both the best positioning ratio and the maximum fill-rate increased. The largest incremental changes occurred after introducing ASN to LROP, which shifted 6.59% of system-wide inventory from the warehouse to the retail stores and increased the maximum fill-rate by 2.56 percentage points. In transforming the LROP to EROP, ASN explains 83% of total increase in the best positioning ratio and 93.1% of increase in the maximum fill-rate. A very similar story can be told about the transforming of the LROP to DRP.

The effect of ASN on the positioning of inventory has drawn our attention. Recall that the LROP system uses local information only. There is no information-sharing between the warehouse and the retail stores besides the replenishment orders placed by the retailers. When the warehouse stocked out, the retailers would not know the warehouse shortage information until the shipments physically arrived at the retail stores. The lack of information about stocks in transit forced the retailers to monitor their inventory status by the nominal inventory position (i.e., on hand inventory minus backorders plus stock *on order*), rather than inventory position defined as on hand inventory minus backorders plus stock *in transit*. Inventory position represents the amount of inventory that is available to meet future customer demand before placing further orders. Whenever the nominal inventory differs from the inventory position, we say that inventory status information has been "distorted." Furthermore, since the warehouse made its allocation decisions based on the orders received from the retailers, the orders triggered by the distorted retailers' inventory status information could distort the warehouse's allocation decisions, which in turn could further distort the retailers' replenishment decisions, and so on. The information distortion caused by the warehouse shortages could mislead inventory decisions and deteriorate in the customer fill-rate. In an information-poor environment such as our baseline study, there seemed to be an incentive to hold some inventory at the warehouse, in an effort to contain the information distortion caused by the warehouse shortages. When the ASN became available, such a motive for holding inventory at the warehouse disappeared. Indeed, as shown in Table 6.6, ASN caused the shifting of 6.59% of system-wide inventory from the warehouse to the retail stores and increased the maximum fill-rate by 2.56%. The information shifted inventory and improved customer fill-rate! After introducing ASN, the best positioning strategy required more than 90% of the system-wide inventory to be positioned at the retail level near the customer.

Rosenfield and Pendrock (1980) and Zipkin (1993) among others have noticed before that holding inventory at the warehouse permitted some decentralization of decision-making. However, we believe that the results presented in Table 6.6 provide the first evidence that the decentralization of decision-making would have been effective if the information itself were not seriously distorted by the shortages at the warehouse. While the literature is now rich on the distortion of demand information and demand variance amplification caused by the "Bullwhip" effect (i.e., the phenomenon where orders to the supplier tend to have larger variance than end customer demand, and the distortion propagates upstream in an amplified form. See Lee, Padmanabhan, and Whang, 1997), the distortion of inventory status information caused by the warehouse shortages often give rise to much more difficult technical problems and as a result, the literature is scant on this topic. Most researchers avoid the problem by assuming either the shortages at the warehouse are backordered or inventory in transit is known. But the problem does exist in distribution practice. As Lee and Billington (1992) reported, suppliers often treat the internal and external customers differently[1]. While unsatisfied demands are often backordered for external customers, unfilled orders from internal customers are rarely backordered. That is a part of reason why Advanced Shipment Notice (ASN) is gaining momentum. ASN is an electronic message from the suppliers telling the downstream distributors what merchandise is on the way. However, at this stage, ASNs are not always sent on time. Says one distribution manager, "In many cases, we received ASNs after the truck gets here." For managers' concerns about the distorted inventory status information caused by shortages at source, see *Progressive Grocer* (October 1994, p64.)

Next, let us look at the changes brought about by other incremental changes. If the retailers also shared the information about their inventory positions with the warehouse, then the warehouse was able to make better allocation decisions by balancing inventories among *all* retail stores, rather than just those that had placed an order in the current period. Table 6. 6 shows that sharing of this additional information further shifted 1.14% of system-wide inventory

[1] Note that differentiating internal and external customers may not be a pitfall as Lee and Billington (1992) suggested. As we discussed earlier in Chapter 3, there are good reasons to treat the internal and external customers differently. Not differentiating internal and external customers is a feature of the traditional modeling approach for multi-echelon inventory problems. That is, build a single location model first and then use this model as a "building block" for constructing a multi-echelon inventory model. The technical justification for this modeling approach is understandable. But the

from the warehouse to the retail stores and improved the maximum fill-rate by an additional 0.11 % from 85.96% to 86.07%. The final incremental change from ROP to EROP had a similar but smaller effect on the positioning of inventory in the distribution system and the maximum fill-rate.

As pointed earlier, the final incremental change only modified the way in which the information about retailers' inventory positions was used. No additional information resource was required. Table 6.6 shows that the change still made a difference. It shifted 0.21% system-wide inventory from the warehouse to the retail stores and increased the maximum fill-rate by 0.08 % from 86.07% to 86.15%. Although the changes were small, they were still statistically significant at 0.01 level. These results have important theoretical implications, demonstrating that the performance of a control system depends on not only the amount information utilized but also how the available information is utilized. Specifically, it shows that in the multi-echelon distribution system setting, the echelon based inventory control system, EROP, outperformed the installation based inventory control system, ROP. The result echoes the study on installation vs. echelon inventory control systems reported by Axsater and Rosling (1993, 1994).

Finally, we took a closer look at the "Push" system. Conceptually, we can transform LROP to the "Push" system by first transforming LROP to EROP and then restricting the warehouse inventory review period to once every 6 periods, and then changing the form of decision rules by having the retailer's reorder point s_r and order-up-to level S_r fixed at infinity (i.e., letting $s_r = S_r = \infty$) and introducing two new control system parameters α (i.e., the percentage of incoming stock that would be allocated to the retail stores immediately) and T (i.e., the delay time after receiving the incoming stock for making secondary allocation).

As shown in Table 6.3, the change from LROP to the "Push" system shifted 11.4% of system-wide inventory from the warehouse to the retail stores, the largest percentage among the three added systems. However, its maximum fill-rate increased only by 1.41 percentage points from 83.4% to 84.81%. The explanation seemed to lie in the amount of information that was utilized by the "Push" system. In the "Push" system, the warehouse was aware of the inventory positions at the retail level in each allocation period. Such an information-sharing scheme was not available in LROP that required local information only. Compared

---

implicit assumption that external and internal customer can be, or should be, treated in the same way seems to be hardly justifiable in the real world.

with EROP or DRP, however, the information utilized by the "Push" system was more restrictive because the "Push" system allowed the warehouse to review its inventory status once every six periods, instead of every period. In other words, the information utilized by the "Push" system was more intensive than that utilized by LROP, but more restrictive than that utilized by EROP or DRP. Not surprisingly, the maximum fill-rate achieved by the "Push" system was higher than that of LROP but lower than that of EROP or DRP.

To get the maximum fill rate, the "Push" system required 95.32% of the system-wide inventory to be positioned at the retail stores near the customer. This is the largest best positioning ratio and several explanations can be offered. First, in the "Push" system, there was no need to hold inventory at the warehouse to prevent information from being distorted by the warehouse shortages. Second, the warehouse was able to balance inventories among *all* retail stores, rather than just those that had placed orders in the current period. As the results in Table 6.6 showed, better allocation decisions reduced the incentive for holding inventory at the warehouse to postponing subsequent allocation decisions. Finally, the explanation had to do with the form of the decision rules used. In the "Push" system, we assumed that the retailers placed huge "pseudo" replenishment orders on the warehouse every period. As a result, neither the constraint on when the inventory could flow to the retail store (i.e., "no order, no delivery" as expressed by the constraint 4.2, see Chapter 4, Section 4.11) nor the constraint on how much could be delivered to the stores (i.e., the "upper bound" for allocation as expressed by the constraint 4.1 became infinity because the retailer's order-up-to level $Sr = \infty$) was active. Without the retailer-imposed constraints on the inventory flows, the warehouse could easily "push" inventory into the field near the customer. We believe that the lack of the retailer-imposed constraints can partially explain why the "Push" system had the largest best positioning ratio among the control systems we simulated.

The effect of the inventory control system on inventory positioning is an area where relatively little research has been done. While it has been a general belief that the choice of inventory control system could have an effect on the positioning of inventory in the distribution system (e.g., Jackson 1988), very little is known about "how." The results presented in this section have provided some insights into this important question, which expand our current understanding about "how," particularly with respect to the effect of information-sharing on the positioning of inventory in the distribution system we investigated.

To our best knowledge, there is no published study that has explored the effect of information sharing on the positioning of a fixed amount of inventory in a one-warehouse multi-retailer distribution system. Only recently have researchers begun to study information sharing in serial supply chains. Gavirneni, Kapuscinski, and Tayur (1996) have reported on a study of a two-firm supply chain with stationary consumer demand, a capacity-restricted upstream firm and a downstream firm that implements a (R, s, S) decision rule with R=0, namely the inventory status is monitored continuously. For this serial supply chain with stationary stochastic demands, they found that information-sharing reduced the upstream firm's costs up to about 35%. However, Gavirneni, Kapuscinski, and Tayur did not investigate how information-sharing would change the downstream firm's costs. Lee, So, and Tang (1996) also have reported on a study of a serial supply chain with non-stationary demand and controlled by base stock policies (for explanation, see Chapter 2, Section 2.3.1). They found that the retailer did not gain from information sharing, but that the supplier reduced costs significantly. One study that seems to echo our simulation results is Bourland, Powell, and Pyke (1996). These authors have studied information-sharing in serial supply chains with stationary stochastic demands. They assume that firms use base stock policies but orders are not placed every period. They found that information-sharing reduced the upstream firm's inventory between 0% and 62%, but increased the downstream firm's inventory between 0% and 4.2%.

Information-sharing lies at the core of many of the 1990s' most touted supply chain management initiatives. Reengineering, quick response, efficient consumer response, vendor-managed inventory and continuous replenishment programs, among others, all require sharing of information among supply chain participants in one way or another. Our simulation results seem to suggest that those industrial initiatives should also shift inventory to the end of the supply chain where customers are. We have shown that information-sharing not only increased the best positioning ratio but also reduced the penalty for positioning close to the customer (see Table 6.4). This observation has managerial implications. As an example, it is well known that Wal-Mart has made a strategic investment in its information-capacity while practicing "cross-docking" (i.e., positioning practice that minimizes the level of inventory positioned at the intermediate stocking locations, such as warehouses or distribution centers) (G. Stalk et al. 1992). Knowing the inherent relationships between information sharing and penalty for the positioning inventory close to the customer, it is safe to say that if Wal-Mart had not made that strategic investment, it might not have been able to successfully practice "cross-docking" and enjoy its remarkable success.

6.3 EFFECT OF SYSTEM-WIDE INVENTORY LEVEL

To study the positioning of inventory in response to the changes in the system-wide inventory level (INV), we repeated the simulation procedures used in the baseline study for all four control systems (i.e., LROP, EROP, DRP, and the "Push" system) but this time we held the system-wide inventory level at 200 units (rather than 150 units as specified in the baseline study). The results are presented in Figure 6.5. Compared to Figure 6.3 (The Effect of Control Systems: Baseline Parameter Settings), we see similar patterns in response to the changes in the positioning ratio for all the control systems. One difference is that the curves representing alternative positioning strategies shown in Figure 6.5 appear to be flatter, indicating that at the higher level of system-wide inventory, it would be more difficult to identify the best positioning ratio, P*, with precision.

Table 6.7 **EFFECTS OF INV AND CONTROL SYSTEMS**

Control System	Positioning Ratio P*		Fill Rate F*	
	INV=150	INV=200	INV=150	INV=200
LROP	0.8424	0.8260	0.8340	0.9050
EROP	0.9218	0.9117	0.8615	0.9165
DRP	0.9208	0.9130	0.8615	0.9210
"Push"	0.9531	0.9380	0.8481	0.9100
Average	0.9095	0.8972	0.8513	0.9131

The comparison between these four control systems at two different levels of system-wide inventory is presented in Table 6.7. As expected, the maximum fill-rate increased with the system-wide inventory level. For all four of the control systems, when the system-wide inventory level increased from 150 units to 200 units, the best positioning ratio decreased slightly, on average shifting 1.23% (=90.95% - 89.72%) of the system-wide inventory from the retail stores back to the warehouse.

Note that the choice of inventory control systems becomes less important as the system-wide inventory level increases to the high levels. This conclusion is supported by Table 6.8, which shows that as the system-wide inventory level increased from 150 units to 200 units, the differences between the maximum fill-rate for LROP and those for the other three control systems, EROP, DRP, and the "Push" system, were reduced. On average, the difference was cut more than in half from 2.3 percentage points to 1.08 percentage points.

Table 6.8. **EFFECT OF INV ON PERFORMANCE DIFFERENCES OF ALTERNATIVE CONTROL SYSTEMS**

	Change in Fill Rate ΔF^*	
Changes	INV=150	INV=200
LROP to EROP	0.0275	0.0115
LROP to DRP	0.0275	0.0160
LROP to "Push"	0.0141	0.0050
Average	0.0230	0.0108

To gain more insights on the effect of a system-wide inventory, we also varied the system-wide inventory level --from 70 units to 100 units, then from 100 units to 150 units, continuing at intervals of 50 units, until the system-wide inventory level reached 400 units. For each inventory level we identified the three positioning strategies: the best positioning, the positioning "close" to the customer, and the "Ship-all" positioning. The investigation, however, was conducted only for EROP. We focused our attention on EROP largely because the simulation search procedures as described in Section 4.2 were relatively easy to implement for EROP. Another justification for focusing on this control system is that there do not seem to have been significant two-way interactions between the system-wide inventory level and the control systems we considered. (This observation is supported by Table 6.7, in which we see that the maximum fill-rates for all four of the control systems increased in response to the increase in the system-wide inventory level.) The results of varying INV are presented in Figure 6.6, 6.7, 6.8, and 6.9.

If we first look at Figure 6.6 and Figure 6.7, we see that the maximum fill-rate increased with additional inventory. The positioning ratio for the maximum fill-rate remained relatively constant and high across the inventory range we investigated. The highest level of the positioning ratio 0.9307 was reached at the lowest system-wide inventory level evaluated, 70 units. The positioning ratio dropped slightly with increasing inventory and then leveled off at about 0.90. Figure 6.8 shows the result of linear regression. The best positioning ratio decreased slightly in response to the increase in the system-wide inventory level.

Table 6.9 **EFFECT OF CHANGES IN INV**

Δ INV	Δ Fs	Δ (Fc-Fs)	Δ (F*-Fc)	Δ F*	Δ P*
70-100	0.0965	0.0016	0.0040	0.1021	-0.0306
100-150	0.0984	-0.0010	0.0030	0.1004	0.0217
150-200	0.0563	-0.0014	0.0001#	0.0550	-0.0101
200-250	0.0343	-0.0021	-0.0019	0.0303	-0.0026
250-300	0.0213	-0.0011	-0.0023	0.0179	0.0096
300-350	0.0136	-0.0005	-0.0020	0.0111	0.0010
350-400	0.0085	-0.0010	-0.0006	0.0069	-0.0138

Not statistically significant at the 0.01 level of paired t-test.

A closer look at the changes in the fill-rate improvements shows that when the system-wide inventory level increased from 70 units to 400 units, the fill-rate improvements first increased and then decreased (see Figure 6.9 and Table 6.9). This interesting phenomenon reflected the nature of the positioning problem. Positioning of inventory is an effort to make better use of the existing inventory resources. If there is little inventory available, to get maximum fill-rate, managers have no choice but to keep the inventory near the customer. This can be explained by considering an extreme case: Supposing there was only one unit of stock available, to maximize fill-rate, managers definitely would have to hold this one unit of stock at the retail level, rather than at the warehouse (because the stock positioned at the warehouse could not provide immediate service to the customer and, therefore, would result in zero fill-rate). Generally speaking, at low levels of system-wide inventory, there is not much room for maneuverability because the difference between the best positioning and the positioning close to the customer was small. On the other hand, if there were a huge amount of inventory available, the positioning of inventory would not make much difference. Excessive inventories could mask the benefit from positioning inventory. This observation is supported by Table 6.9, which shows that as the system-wide inventory increased, the best positioning ratio required the warehouse to keep a larger proportion of system-wide inventory, but the penalty for using the "close" to the customer positioning strategy began to diminish.

Figure 6.9 shows that the best positioning ratio remained of greater than 0.9 as the system-wide inventory level varies from 70 units to 400 units. The largest penalty for positioning "close" to the customer (F*-Fc) was a decrease in the customer fill-rate by 0.89 percentage points founded at INV=200. On average, across all the simulation runs from INV=70 to INV= 400, the "close" positioning strategy only decreased the maximum fill rate by 0.52 percentage points. Based on these results, we can make an even stronger case for positioning inventory near the customer than we could have from the baseline study.

Note that for a given set of inventory, transportation, and control system resources, it is reasonable to measure the penalty for positioning "close" to the customer by the difference between the maximum fill rate and the fill rate for the maximum positioning ratio. However, preventing a decrease in the customer fill-rate in the high level range, say from 95% to 90%, would require much more inventory than that required for preventing a decrease in the fill-rate the same amount but at the low level range, say from 65% to 60%. In other words, the same magnitude of changes in fill-rate could have very different inventory implications depending on the base. To study the effect of the system-wide inventory (INV), it maybe better to measure the penalty for positioning close to the customer by additional inventory required for positioning "close" to the customer to match the fill-rate of the best positioning. The question is whether or not the pattern of the changes in response to the increase in INV still holds as we reported before.

To find out additional inventory required, additional simulation runs have to be conducted. Clearly, the difference between the maximum fill rate and the fill rate for positioning "close" to the customer is bounded by the difference between the maximum fill rate and the fill rate for "Ship-all" strategy. Consequently, the additional inventory required for the "close" positioning to match the best positioning was bounded by the additional inventory required for the "Ship-all" to match the best positioning. Because it is easier to identify the "Ship-all" positioning than to identify the positioning "close" to the customer, we decided to identify the additional inventory required for the "Ship-all," rather than for the "close" positioning, to match the best positioning to simplify our simulation procedures. Since we were concerned only with the pattern of the changes in response to the increase in the system-wide inventory level, such an approximation should not cause any major concern.

Figure 6.10 shows the results about the fill rate improvement and the additional inventory required in response to the changes in the system-wide inventory level. While the additional inventory required began to decline at a higher level of the system-wide inventory level than the fill-rate improvement did, their responses to the increase in the system-wide inventory level shared a similar pattern: increased first and then decreased. Based on these results, we conclude that the general picture did not change materially when the value for pursuing the best positioning was measured by the additional inventory required, rather than by the fill-rate improvement. Our observation based on the changes in fill rate still holds: the positioning of inventory appears to be most important when the system-wide inventory is in the middle levels.

We should note that both the system-wide inventory level and the inventory control system are resource factors. As the system-wide inventory increased to high levels, the positioning of inventory (as shown in Table 6.9) as well as the choice of inventory control systems (as shown in Table 6.8) became less important. It is of interest to note that both the positioning problem and the choice of control systems did not draw much managerial attention when there was a huge amount of inventories accumulated throughout distribution networks. The distribution practice has changed, however. In 1990s, the intensified global competition has necessitated prudent management of inventory and rapid development in information technology has provide distribution managers infrastructure and tools to take actions for improving distribution operations. As more and more companies have begun to devote their efforts to getting rid of redundant inventories and streamlining their supply chains, we believe that the positioning of inventory and the choice of inventory control system will increasingly be recognized as important levers for companies to gain competitive advantage in today's highly competitive marketplace.

6.4 EFFECT OF DEMAND UNCERTAINTY

Demand uncertainty is often measured by the coefficient of variation (CV), the standard deviation of the demand divided by the mean of the demand. The general belief is that the benefit from holding inventory at the warehouse increases with the demand uncertainty. Indeed, Jackson (1988) has predicted that holding some inventory at the warehouse, as opposed to the "Ship-all" positioning, becomes increasingly attractive as the CV increases to the level of two, three, and higher. If this conjecture were true, then the robustness of our general finding in the baseline study (i.e., that inventory should be positioned near the customer to get the high levels of customer fill-rate) could be challenged.

To study the effect of the CV on the positioning of inventory in the distribution system, we varied the value of CV while holding the values of the other experimental factors constant. The results, using EROP, are presented in Figure 6.11, 6.12, and 6.13. When demand uncertainty increased, the maximum fill-rate that provided by the same amount of system-wide inventory declined as expected (see Figure 6.11). In the meantime, the trend of the positioning ratio for the maximum fill-rate is downward (see Figure 6.12), suggesting that a larger proportion of the system-wide inventory should be held back at the warehouse in response to the increase in the demand uncertainty. (The number are close though.).

Some especially interesting results are those presented in Figure 6.13 and Table 6.10. The fill-rate improvement due to best positioning (F*-Fc) increased as the CV approached 1, and then decreased as the CV increased to the high levels of greater than 1. This is a surprising result, suggesting that Jackson's (1988) conjecture on the effect of CV may not be accurate.

To make sure that the observed effect of the CV on the positioning of inventory in the distribution system would not be masked by potential interactions with other experimental factors, we carefully screened for possible two-way interactions. We found that there were significant interactions between the CV and the system-wide inventory level (INV). We repeated the simulation procedures used for investigating the effect of system-wide inventory level but held the CV at different levels. In doing so, we generated several curves similar to the one shown in Figure 6.7 (i.e., Effect of System-Wide Inventory Level, see the previous section 6.3). Since the curves shown in Figure 6.14 were not parallel, we concluded that interactions between CV and INV did exist.

Table 6.10 **EFFECT OF CHANGES IN CV (INV=150)**

Δ CV	Δ Fs	Δ (Fc-Fs)	Δ (F*-Fc)	Δ F*	Δ P*
0.53 to 0.80	-0.0409	0.0027	0.0010	-0.0372	0.0029
0.80 to 1.00	-0.0338	0#	0.0029	-0.0309	-0.0131
1.00 to 1.31	-0.0484	0.0003	-0.0025	-0.0506	0.0137
1.31 to 1.80	-0.6530	-0.0013	-0.0019	-0.0685	-0.0205

Not statistically significant at 0.01 level using paired t-test.

As shown in Figure 6.11, when demand uncertainty was low (e.g., CV=0.53), the 150 units of inventory provides a relatively high level of customer fill-rate, 98.02%. When the demand uncertainty increased to the high level (e.g., CV=1.8), the same amount of inventory resource provided only a fairly low level of customer fill-rate, 76.84%. In other words, the fixed system-wide inventory level of 150 units provided a high fill rate at the low demand uncertainty (CV=0.53) and a low level fill rate at the high demand uncertainty (CV=1.80). We know from the results presented in the previous section (Section 6.3) that the fill-rate improvement would be small when the inventory resource was at either low or high levels. Thus, one might question whether or not the phenomenon we observed as shown in Figure 6.13 was actually due to the strong interactions between CV and the system-wide inventory level (INV).

In an effort to isolate the effect of the CV from the effect of INV, we changed the base upon which the effect of the CV on the positioning of inventory was measured and compared. Instead of fixing the system-wide inventory level, we fixed the fill-rate for "Ship-all" strategies at 0.845. This is the fill rate found for the "Ship-all" strategy in the baseline parameter settings. Given a specific value of the CV, we recorded the inventory level required for the "Ship-all" to provide the predetermined customer fill-rate. Then, fixing this inventory level, we simulated alternative positioning strategies and recorded the best positioning ratios, and the fill-rate improvements both due to increase in the shipment frequency SHIPR and the best positioning, and the maximum fill rate. For control, we also recorded the realized fill rate for "Ship-all" strategies associated with different values of CV. This approach allowed us to compare the fill-rate improvement for different values of CV with an equal starting point, in terms of the fill-rate for "ship-all." For different CV settings, different amounts of inventory were needed for the "Ship-all" positioning to provide the same predetermined customer fill-rate. As shown in Figure 6.15, for CV values of 0.53, 0.80, 1.00, 1.31, and 1.80, the distribution system needed to have a system-wide inventory of 84, 103, 120, 150, and 203 units, respectively, for the "Ship-all" to provide the predetermined fill-rate, 84.55%.

Figure 6.15 shows the effect of the CV on the positioning of inventory for a predetermined customer fill-rate of 0.845 for the "Ship-all" strategies. Surprisingly, we observed the same pattern as the one shown in Figure 6.13 (where the system-wide inventory level was fixed at 150 units). The largest fill-rate improvement for the best positioning ratio still occurred at CV=1. The fill-rate improvement decreased as the CV went either down to the levels of lower than 1 or up to the levels of greater than 1. In both Figure 6.13 and Figure 6.15, the fill-rate improvement due to the increase in the shipment frequency (Fc-Fs) was relatively constant with the changes in the CV. The effect of demand uncertainty was mainly reflected in the fill-rate improvement due to the best positioning. The detailed results about the changes in response to the changes in the CV are shown in Table 6.11.

Table 6.11. **EFFECT OF CHANGES IN CV**

(Fill-rate for "Ship-all" $\approx$ 0.845)

ΔCV	ΔINV	Δ Fs	Δ (Fc-Fs)	Δ (F*-Fc)	Δ F*	Δ P*
0.53 to 0.80	19	0.0007	0.0007	0.0047	0.0061	0.0666
0.80 to 1.00	17	0.0001	-0.0001#	0.0029	0.0029	-0.0379
1.00 to 1.31	30	0.0001	0.0003	-0.0023	-0.0019	0.0218
1.31 to 1.80	53	-0.0002	-0.0011	-0.0027	-0.004	-0.0057

Not statistically significant at 0.01 level using paired t-test.

We performed some sensitivity analysis to evaluate the generality of the effect of CV. We obtained consistent results when we changed the demand distribution from lognormal to

Gamma or Weibull. We also tested the effect of the CV using the "Push" control system. The same pattern as shown in Figures 6.13 and 6.15 was observed. The change in the predetermined fill-rate for the "Ship-all" would not change the pattern either. This conclusion is supported by Figure 6.16, which shows that the fill-rate improvement at the best position ratio for CV=1 envelops the fill-rate improvements observed when the CV was set at a level either smaller or greater than 1. Finally, we identified the additional inventory required for the fill-rate of "Ship-all" to match the maximum fill-rate. The largest additional inventory required divided by the system-wide inventory level used also occurred as the CV approached the level of 1 (see Figure 6.17).

There seems to be something occurring at CV=1.0. When demand uncertain is at low levels, we know that inventory should be positioned close to the customer because there is little chance for the "unbalanced" inventories to occur at the retail level. In the absence of "unbalanced" retail inventories, there is no need to hold inventory at the warehouse. Indeed, Roundy (1985) and Maxwell and Muckstadt (1985) show that if there were no demand uncertainty, then all inventory should be immediately sent to the retail stores. In general, at the low levels of demand uncertainty, there is not much difference between the best positioning and the positioning close to the customer. As expected, when demand uncertainty changes in a low-level range, the fill rate improvements due to the best positioning were small. As demand uncertainty increases, holding some inventory at the warehouse for sending it out later to balance to the "unbalanced" inventories at the retail level became justifiable. Figures 6.13 and 6.15 show that the fill-rate improvement due to the best positioning did increase as the CV increased approaching 1. The question is, when the demand uncertainty increases to levels of greater than 1.0, why should the fill-rate improvement for holding some inventory at the warehouse go down?

Here are two possible explanations. First, Eppen and Schrage (1981) have shown that the probability that the warehouse is able to send stocks to balance the retail inventories decreases as the CV increases (i.e., "allocation Assumption" no longer holds with high probability as CV goes up above 1.0.) The implication is that as demand uncertainty went up above 1.0, balancing supply is more difficult. The second explanation focuses on the demand side. We may hold some inventory at the warehouse with the intention of sending it out later to balance the retail inventories. But then we must ask ourselves: Why should we want to balance the supply? The premise is that while demands in each period are stochastic, the cumulative demand at the retail stores are somehow "balanced." When the CV increases to extremely high levels, however, this premise may no longer hold. If demands at the individual retail stores were extremely chaotic and "unbalanced," balancing supply (assuming it is possible) would not necessarily lead to a better match between the demand and the supply. In such a case, one should not expect that holding inventory at the warehouse could significantly improve the customer fill-rate.

These intuitive explanations, however, offer no explanation on why the observed fill-rate improvement due to the best positioning would occur as the CV goes up *just* past 1. It is beyond the scope of this dissertation to give a detailed analytical explanation for this interesting phenomenon. For this dissertation study, suffice it to say that our general finding that inventory should be positioned near the customer to get a high level of customer fill-rate still holds, even if the demand uncertainty increases to the levels of CV>1.0.

The results presented here also provide the first evidence that Jackson's (1988) conjecture (i.e., the benefit of holding inventory at the warehouse, as opposed to "Ship-all" increases with the CV) may not apply as the CV goes above 1.0. Jackson's conjecture is well known and has been widely used to support the argument for positioning inventory at the warehouse. The conjecture, however, to our best knowledge, has never been tested before.

Before we end this section, let us look at the Figure 6.13, and 6.15 again. As the CV increases from 0.53 to 1.8, for the fixed system-wide inventory level of 150 units, the largest observed fill-rate improvement due to the best positioning was a 1.13 percentage points, with an average (across all experimental points) fill rate improvement of 0.81 percentage points (Figure 6.13); for INV fixed to equal predetermined fill-rate of 0.845 for the "Ship-all" strategies, the largest observed fill-rate improvement due to the best positioning was a 1.11 percentage points, with an average fill rate improvement of 0.754 percentage points (figure 6.15). In both cases, the penalty for positioning inventory "close" to the customer remained fairly small. With these results, the robustness of our general finding that inventory should be positioned near the customer to get the high levels of customer fill rate has been extended in one of the key dimensions—the demand uncertainty.

6.5 EFFECT OF TRANSIT LEAD TIMES

In this section, we focus our attention on the effects of transit lead times. Here, our experiments were conducted only for EROP. We considered three scenarios about the changes in the transit lead times: First, the transit lead times to both the warehouse (L_w) and the retail stores (L_r) increase from 0 to 1 to 2 periods. Second, having fixed the transit lead time to the warehouse (L_w) at 1 period, the transit lead time to the stores (L_r) increases from 0 to 1 to 2 periods. Third, having fixed (L_r) at 1 period, (L_w) increases from 0 to 1 to 2 periods.

Table 6.12 **EFFECT OF PROPORTIONAL INCREASES IN TRANSIT LEAD TIMES (L_w and L_r)**

Changes	Δ Fs	Δ (Fc-Fs)	Δ (F*-Fc)	Δ F*	Δ P*
$L_w = L_r$=0 to $L_w = L_r$=1	-0.0197	-0.0018	-0.004	-0.0255	0.0151
$L_w = L_r$=1 to $L_w = L_r$=2	-0.0174	-0.0011	-0.0043	-0.0228	0.007
$L_w = L_r$=0 to $L_w = L_r$=2	-0.0371	-0.0029	-0.0083	-0.0483	0.0221

The results for the first scenario (i.e., where the transit lead times to the warehouse and to the retail stores both increase from 0 to 1 to 2 periods) are presented in Figure 6.18 and Table 6.12, which show that as the lead times to the warehouse and to the stores increased equally, the fill-rate improvement due the best positioning decreases and the positioning ratio increases. Note that in response to the increase in the lead times, the incentive for holding inventory at the warehouse diminished. The fill rate improvement due to the increase in the shipment frequency to the stores SHIPR, (Fc-Fs), decreased as well. Note that the sum of the fill rate improvements due to the increase in the shipment frequency (Fc-Fs) and due to the best positioning (F*-Fc) represents the so called "depot effect" (i.e., F*-Fs). Many published studies on the "depot effect" have assumed zero transit lead times. Based on our simulation results, we expect that the reported "depot effect" from holding inventory at the warehouse would become substantially smaller if non-zero lead times were used.

The results for the second scenario (i.e., having fixed the transit lead time to the warehouse, L_w, at 1 period, the transit lead time to the stores, L_r, increases from 0 to 1 to 2 periods) are shown in Figure 6.19 and Table 6.13. They are very similar to the results for the first scenario except that the changes in the best positioning ratio, Δ P*, are large. The longer the transit lead time to the stores (L_r), the less effectively the warehouse could react to changing conditions at the retail level. Consequently, it means less advantageous for holding back inventory at the warehouse.

Table 6.13 **EFFECT OF CHANGES IN LEAD TIME TO STORES (Lr)**

L_w=1	Δ Fs	Δ (Fc-Fs)	Δ (F*-Fc)	Δ F*	Δ P*
L_r=0 to L_r=1	-0.0161	-0.0036	-0.0027	-0.0224	0.0736
L_r=1 to L_r=2	-0.015	-0.0015	-0.0033	-0.0198	0.0115
L_r=0 to L_r=2	-0.0311	-0.0051	-0.006	-0.0422	0.0851

For the third scenario (i.e., having fixed Lr at 1 period, L_w increases from 0 to 1 to 2 periods), Figure 6.20 and Table 6.14 show that the results are different from the first two scenarios in two aspects. First, the changes brought about by the increase in the lead time to the warehouse (L_w) are relatively small. The fill rate improvement due to the increase in the shipment frequency to the stores is almost unchanged or increased only very slightly. Second, in response to the increase in L_w, the best positioning ratio, P*, decreased, rather than increasing as it had in the first two scenarios. The fill-rate improvement due the best positioning still decreased as in the first two scenarios, indicating that more inventory should be held back at the warehouse even though the benefit for doing so has begun to diminish.

The second observation about the third scenario (i.e., having fixed L_r at 1 period, L_w increases from 0 to 1 to 2 periods) is important because it provided us an example that the changes in best positioning ratio and in the fill-rate improvement do not always follow the same pattern. (Other examples we have seen are the changes in the fill rate improvement in response to the increases in the system-wide inventory level at the high experimental levels, see Section 6.3) That is why we have recorded both the best positioning ratio and the fill-rate improvement in this study. The fill-rate improvement seems to be more important for distribution managers to consider in deciding whether or not some inventory should be held at the warehouse. If the answer is "yes," then the best positioning ratio becomes important because it provides the guideline for what actions should be taken. Our argument for positioning inventory near the customer is based on the fact that the observed fill-rate improvement was so small, there was little penalty for not taking the actions for holding inventory at the warehouse as required by the best positioning.

Table 6.14 **EFFECT OF CHANGES IN LEAD TIME TO WAREHOUSE (L_w)**

L_r=1	Δ Fs	Δ (Fc-Fs)	Δ (F*-Fc)	Δ F*	Δ P*
L_w=0 to L_w=1	-0.0018	0.0002	-0.0009	-0.0025	-0.0047
L_w=1 to L_w=2	-0.0019	0.0001#	-0.0009	-0.0027	-0.0045
L_w=0 to L_w=2	-0.0037	0.0003	-0.0018	-0.0052	-0.0092

Not statistically significant at 0.01 level using paired t-test

Considering all three scenarios together, the results of transit lead times seem to suggest that the longer the transit lead time is to a stocking location, the more inventory should be positioned at that location. This to some degree reflects the well-known results from single-location inventory models.

Table 6.15 **EFFECT OF CHANGES IN LEAD TIMES (L_w + L_r =2)**

L_w+L_r=2	Δ Fs	Δ (F*-Fs)	Δ (Fc-Fs)	Δ (F*-Fc)	Δ F*	Δ P*
L_w=0 to 1	0.0133	0.0043	-0.0007	0.0050	0.0176	-0.0164
L_w=1 to 2	0.0142	0.0055	-0.0003	0.0058	0.0197	-0.0789
L_w=0 to 2	0.0275	0.0098	-0.0010	0.0108	0.0373	-0.0953

Next, we studied a scenario in which the sum of the lead times to the warehouse and to the retail stores was held constant at 2 periods, but the relative magnitudes of the lead times varied from 0 to 2. The results are shown in Figure 6.21. Note that different transit lead time scenarios represent different distribution system configurations. For instance, when the transit lead time to the warehouse (L_w) changes from 0 to 1, the warehouse can postpone the allocation decision for 1 period until the stock from the outside supplier physically arrives at the warehouse. As a result, there should be a "risk-pooling effect" over the transit lead time L_w=1. The change in the fill-rates for "Ship-all," denoted by Δ Fs, in response to the change in L_w provided us the information needed for assessing the value of the "risk pooling effect over the supplier lead time." The results are detailed in Table 6.15

Note that in Table 6.15 we have added a new column Δ (F*-Fs), which represents the changes in the "depot effect" (i.e., the difference between the fill rate of the best positioning and the fill rate for the "Ship-all") in response to the changes in the transit lead times. Clearly, Δ (F*-Fs) = Δ (Fc-Fs) + Δ (F*-Fc), that is, the numbers appearing in this new column, column 3, are equal to the sum of these listed in column 4 and column 5. As explained above, the column 2 represents the "risk pooling effect over the transit lead time to the warehouse," Δ Fs.

Figure 6.21 and Table 6.15 show that as the warehouse moves closer to the retail stores, both the "risk pooling effect over the supplier lead time" and the "depot effect" increase. However, the changes in the "risk pooling effect over the supplier lead time" seem to be greater than the changes reflected in the "depot effect." Since Δ P* < 0, more inventory should be held back at the warehouse. But, the smallest best positioning ratio observed was still 0.8429, and the largest fill rate improvement due to the best positioning remained fairly small at 1.46 percentage points (see Figure 6.21 for the case where L_w = 2 and L_r=0).

Schwarz (1989) has suggested that there is no known study on the relative magnitudes of "risk pooling effect over the supplier lead time" and the "depot effect" of holding some inventory at the warehouse. Using the data presented in Figure 6.21 and Table 6.15, we can explore this interesting topic easily.

For illustration, let us look at an example (see Figure 6.22). Suppose that at the starting point L_w=0 and L_r=2 and the warehouse holds no inventory. To improve the customer fill-rate, we assume that distribution managers can take two recourse actions: One is to re-locate the warehouse to a location closer to the stores with L_w=1 and L_r=1. Another is to hold some inventory at the warehouse, which would increase the shipment frequency SHIPR from no greater than 0.1667 shipments per period to the predetermined shipment frequency, 0.2315 shipments per period. The question is: Which action will provide the greatest improvement?

As shown in Figure 6.21, holding some inventory at the warehouse without changing the configuration of the distribution system (i.e., L_w=0 and L_r=2) could increase the customer fill-rate by 1.17 percentage points (i.e., "depot effect," (F*-Fs) = (F*-Fc)+(Fc-Fs) = 0.0038 +0.0079 = 0.0117). Alternatively, relocating the warehouse to a location closer to the stores while holding no inventory at the warehouse would increase the customer fill-rate by 1.33 percentage points (i.e., "risk-pooling effect over the supplier lead time," Δ Fs=0.0133 as shown in Table 6.15). Since Δ Fs = 0.0133 > 0.0177 = (F*-Fs) (i.e., the fill rate improvement by relocating the warehouse is greater than the fill rate improvement due to the "depot effect" of holding inventory at the warehouse without changing distribution system configuration), we conclude that the relocation is the preferred action. This example is interesting because it shows that the "risk-pooling effect over the supplier lead time" could be greater than "depot effect" for holding inventory at the warehouse.

To further demonstrate the importance of preserving the "risk-pooling effect over the supplier lead time," let us consider another scenario. Suppose that at the starting point L_w=2, L_r=0 and the warehouse is allowed to hold inventory. From Figure 6.21, we know that the penalty for positioning inventory close to the customer was a decrease in the fill rate by 1.46 percentage points. On the other hand, if we go one step further to eliminate the warehouse, or more precisely, to consolidate the warehouse with the outside supplier's stocking facility (i.e., change from starting point L_w=2 and L_r=0 to the point where L_w=0 and L_r=2), the maximum fill-rate would decrease by 2.65 percentage points (Δ Fc= Δ Fs + Δ (Fc –Fs) = -0.0275 + 0.001 = -0.0265, see Table 6.15). The results show that the penalty for going to the extreme by eliminating the warehouse could be as high as twice of the penalty for positioning inventory close to the customer. It is of interest to note that Wal-Mart practices "cross-docking" but keeps 85% of its products going through its distribution centers. On the other hand, K-Mart has only 50% of its products going through its distribution centers where it keeps inventories (Stalk et al. 1992). Again, our results seem to support Wal-Mart's distribution practice.

6.6 EFFECT OF SHIPMENT FREQUENCY TO THE RETAIL STORES

All simulation experiments reported in the previous sections (Sections 6.1-6.5) were conducted with the predetermined shipment frequency to any of the stores (SHIPR) fixed at 0.2315 shipments per period. In this section, we examine the changes in response to the increase in the shipment frequency SHIPR while keeping all the other experimental factors constant as specified in the baseline study. The results for EROP are shown in Figure 6.23 and Table 6.16.

Table 6.16 **EFFECT OF SHIPMENT FREQUENCY (SHIPR)**

SHIPR (SHIPW= 0.1667)	**Strategy**	(s_r, S_r, s_w, S_w)	**Positioning Ratio** P **(95% Conf. Interval)**	**Fill-rate** F **(95% Conf. Interval)**
$\leq$0.1667	"Ship-All"	(100, 100, 177, 396)	1 ± 0.0	0.8455 ± 0.0013
0.2315	"Best"	(24, 36, 180, 399)	0.9218 ± 0.0005	0.8615 ± 0.0013
	"Close"	(35, 38, 178, 397)	0.9769 ± 0.0002	0.8527 ± 0.0013
0.3	"Best"	(25, 32, 181, 400)	0.8626 ± 0.0006	0.8637 ± 0.0014
	"Close"	(35, 36, 179, 398)	0.9572 ± 0.0003	0.8565 ± 0.0013

When the shipment frequency SHIPR increased from 0.2315 to 0.3 shipments per period, as expected, the maximum fill-rate increases and the best positioning ratio decreased. The decrease in the best positioning ratio indicates that more inventory should be held back at the warehouse. However, the fill-rate improvement due to the best positioning becomes smaller as the shipment frequency SHIPR increases from 0.2315 to 0.3 shipments per period. As a result, the observed increase in the maximum fill-rate, Δ F*, is attributed only to the increase in the fill rate improvement due to the increase in the shipment frequency Δ (Fc-Fs). The changes in response to the increase in the shipment frequency are detailed in Table 6.17.

Table 6.17 **EFFECT OF CHANGES IN SHIPMENT FREQUENCY (SHIPR)**

Δ SHIPR	Δ Fs	Δ (Fc-Fs)	Δ (F*-Fc)	Δ F*	Δ P*	Δ Pc
0.1667 to 0.2315	0	0.0072	0.0088	0.0160	-0.0782	-0.0231
0.2315 to 0.3	0	0.0038	-0.0016	0.0022	-0.0592	-0.0197
0.1667 to 0.3	0	0.0110	0.0072	0.0182	-0.1374	-0.0428

The decrease in the fill rate improvement due to the best positioning Δ (F*-Fc) indicates that the penalty for positioning close to the customer was reduced in response to the additional increase in the shipment frequency SHIPR. The convergence between the best positioning and the positioning close to the customer provided evidence that the additional increase in the shipment frequency SHIPR would not materially change the conclusion we reached in the baseline study: inventory should be positioned near the customer to get high levels of customer fill-rate.

It is of note that in Figure 6.23 and in Table 6.16 and 6.17, we have recorded the positioning ratio associated with the positioning inventory close to the customer (Pc). By definition, Pc is the maximum positioning ratio that would be possible for a given set of inventory and transportation resources. Consequently, 1-Pc should be explained as the minimum percentage of the system-wide inventory that should be held at the warehouse. For the "Ship-all," Pc=1. No inventory was held at the warehouse. However, when the shipment frequency SHIPR increased to a level of greater than the shipment frequency to the warehouse (SHIPW), some proportion of inventory (which must be ≥ 1-Pc) had to be held at the warehouse. Otherwise, it would not be possible for the warehouse to send stock to the retail stores more frequently than the warehouse received its incoming stock from the outside supplier. For instance, to maintain the predetermined shipment frequency SHIPR at 0.2315 shipments per period, the warehouse had to hold at least 2.31% (i.e., 1-Pc=1-0.9769) of the system-wide inventory (see Figure 6.23). It is important to note, however, that the warehouse could hold more than 2.3% of the system-wide inventory while keeping the predetermined SHIPR =0.2315 shipments per period unchanged (as shown in Figure 6.3). When the predetermined shipment frequency SHIPR further increased from 0.2315 to 0.3 shipments per period, at least 4.28% of the system-wide inventory had to be kept at the warehouse. But again, the warehouse could hold more than 4.28% of the system-wide inventory while keeping the predetermined SHIPR=0.3 shipments per period unchanged.

As shown in Figure 6.23, when SHIPR=0.3, the best positioning ratio is 0.8626, indicating that by holding13.74% of the system-wide inventory at the warehouse, the distribution system could reach the maximum fill rate, 86.37%. However, when the shipment frequency SHIPR was set at 0.2315 shipments per period, the warehouse could still hold 13.74% of the system-wide inventory. Therefore, determining whether or not holding 13.78% of the system-wide inventory at the warehouse was the best positioning would depend on the shipment frequency SHIPR. Clearly, the maximum fill-rate and the magnitude of the fill-rate improvements due to the best positioning and due to the increase in the shipment frequency SHIPR depend on the shipment frequency SHIPR as well. Although the changes brought about by the additional increase in the shipment frequency seemed to be relatively small, they were enough to show why a lack of control of the shipping frequency SHIPR could cause different conclusions for studies conducted in very similar environments.

While the additional increase in the shipment frequency SHIPR would not change our general finding about where inventory should be positioned in the distribution system, we have demonstrated the importance of controlling the shipment frequency SHIPR when different positioning strategies are simulated and evaluated. The final point we would like to make about the shipment frequency is that it is a resource factor, representing the part of transportation resource utilized by the distribution operations. Resource factors seem to share some common features. When additional system-wide inventory becomes available (see Section 6.3), an information-intensive control system is used (see Section 6.2), or the shipment frequency SHIPR is increased (i.e., to levels greater than the shipment frequency to the warehouse SHIPW), the observed fill-rate improvement due to the best positioning almost always decreased. Positioning of inventory is an effort to make better use of the existing resources. The observed convergence between the best positioning and the positioning "close" to the customer in response to the increase in each of these resource factors seems to indicate that only when prudent management of all the existing resources is necessary is the best positioning an issue of managerial importance.

6.7 EFFECT OF THE NUMBER OF RETAIL STORES

Our final experiments were designed to investigate the effect of the number of retail stores supplied by the warehouse. Again, our investigation was conducted only for EROP. In this study, we kept the ratio of the system-wide inventory level (INV) divided by the number of the retail stores (N) constant at 18.75 units per store. We kept all the other experimental factors unchanged at the levels as specified in the baseline study.

Figure 6.24 shows that as the number of retail stores (N) increased, the best positioning ratio decreased, indicating that more inventory should be positioned at the warehouse. In response to the increase in N, the fill-rate for the "Ship-all" declined. This should not be considered surprising, because the chance of having "unbalanced" inventories at the retail level increased with the increase in the number of retail stores. Despite the greater chance of having "unbalanced" inventories at the retail level, the maximum fill-rate increased with the increased number of retail stores. To understand the reason behind this interesting phenomenon, let us consider the difference between a one-warehouse 8-retailer distribution system and two distribution systems in which each warehouse supplies 4 retail stores. Under the assumption that all the retail stores are identical, the one-warehouse 8-retailer distribution can be considered as a result of consolidating warehouse operations of the two distribution systems, each with 4 retail stores. Similarly, a distribution system with 32 identical retail stores can be consider as a result of consolidating warehouse operations of four distribution systems, each with 8 identical retail stores.

Table 6.18 **EFFECT OF CHANGES IN STORE NUMBER (N)**

ΔN	$(N+\Delta N)/N$	Δ Fs	Δ (Fc-Fs)	Δ (F*-Fc)	[(F*-Fc) + Δ (F*-Fc)] /(F*-Fc)	Δ F*	Δ P*
4 to 8	2	-0.0027	0.0034	0.0018	$<\sqrt[3]{2}$	0.0025	-0.003
8 to 32	4	-0.0033	0.0059	0.0013	$<\sqrt[10]{4}$	0.0039	-0.0129

The effect of centralization of stocking locations at the same echelon is well known. Eppen (1979) has demonstrated that for stocking locations facing independent demands, the expected holding and shortage cost increases with the square root of the stocking location size, $\sqrt{N}$. In the distribution system we considered, the retailers' replenishment orders are not independent either as assumed by Eppen (1979) across periods (they are correlated as shown by Ehrhardt, Schultz, and Wagner, 1981) or across the retail stores (because of interdependence induced by the balancing allocation rule used at the warehouse), and more importantly, the warehouse shortages do not necessarily lead to shortages at the retail stores (for detailed discussion, see Schwarz, 1981a). The benefit for consolidating the warehouse operations, as reflected in the fill-rate improvement at the retail level, seemed to be much smaller than theory suggests. Indeed, the observed fill-rate improvement due to the best positioning increased with N but not proportionally to $\sqrt{N}$. Table 6.18 shows that when N was doubled from 4 to 8, the fill-rate improvement due to the best positioning increased only by 0.18% (from 0.7 percentage points to 0.88 percentage points), which was less than $\sqrt[3]{2}$ times. When N increased 4 times from 8 to 32, the fill-rate improvement increased only by 0.13 % (from 0.88 percentage points to 1.01 percentage points), which was less than $\sqrt[10]{4}$ times.

Another explanation of the data shown in Table 6.18 would be that the number of stores had to increase more than 4 times from 8 to 32 to achieve nearly the same magnitude of fill-rate improvement as when the number of stores was doubled from 4 to 8. Since the increase in the penalty for positioning close to the customer quickly became marginal, the number of stores does not seem to have materially changed our general finding that inventory should be positioned near the customer to get the high levels of customer fill-rate.

6.8 SUMMARY AND DISCUSSION

In the preceding text and charts, we have presented the results of our simulation experiments designed to answer the question of where inventory should be positioned to get the best customer fill-rate in a distribution system with stochastic demand. Our study differs from many previously published studies on the positioning problem in one important aspect—we concerned ourselves with only those alternative positioning strategies that shared the same inventory, transportation, and control system resources. This approach ensured that the alternative positioning strategies were comparable and that the results were meaningful.

Our baseline study and our investigation on the effect of six experimental factors produced remarkably consistent results, showing that a high portion of the available inventory should be positioned near the customer to maximize levels of customer fill-rate. This is not an intuitively appealing finding and is counter to much previous research and management practice. Many managers may still hold the belief that having some inventory at the warehouse will provide them with more flexibility to respond to changes in the field. This study does not support that contention for firms that high availability needs.

The investigation on the effect of the three resource factors and three environmental factors also provide us insight on the conditions in which managers may consider pursuing the best positioning, rather than simply positioning inventory "close" to the customer. Positioning of inventory is an effort to make better use of the existing resources. We have demonstrated that if there were plenty of resources available, either with high levels of the system-wide inventory (section 6.3), shipment frequencies (Section 6.6), or information-sharing schemes (Section 6.2), the difference between the maximum fill rate and the fill rate for the positioning "close" to the customer is small. Table 6.19 summarizes the conditions in which relatively large penalty for positioning inventory "close" to the customer: (1) lack of information-sharing schemes; (2) infeasible for making frequent shipments to the retail stores; (3) moderate available inventory resource; (4) negligible transit lead times; (5) larger number of the retail stores; and (6) demand uncertainty of CV approaching1.0.

It is worth pointing out that our well-controlled simulation results seem to support Wal-Mart's "cross-docking" practice. For companies that would like to follow Wal-Mart's footsteps, we have pointed out two possible pitfalls: "cross-docking" without information sharing (see Section 6.2); or going to the extreme to eliminate the warehouse itself (see Section 6.5).

We are convinced that this dissertation work is important not only because it has expanded the current understanding of the inventory positioning problem but also because it provides a new ground for conducting meaningful comparative studies on multi-echelon inventory control systems. Currently, comparative studies on multi-echelon inventory control systems are often based on the relative performance of some "reasonable" realizations of the

Table 6.19 **SUMMARY**

		Best Positioning Ratio, P*	Maximum Fill Rate, F*	Maximum Positioning Ratio with SHIPR, Pc	Penalty for Positioning Close to the Customer, (F*-Fc)	Conditions in which Large Penalty for Positioning Close to the Customer Observed
Baseline Study with LROP		0.8424	0.8340	0.9697	0.0147	
Resource Factors	Information ↑	Increases ↑	Increases ↑	Increases ↑	Decreases ↓	Lack of Information Sharing
	Retailer Shipment Frequency (SHIPR) ↑	Decreases ↓	Increases ↑	Decreases ↓	Decreases ↓	Frequent Shipments to Retail Stores are not feasible
	System-wide Inventory Level (INV) ↑	Decreases ↓	Increases ↑	Increases ↑	First Increases and then Decreases	Moderate Inventory Resource
Environmental Factors	Demand Uncertainty (CV) ↑	Decreases ↓	Decreases ↓	Decreases ↓	Increase When CV ≤ 1 and Then Decreases as CV>1	CV approaches 1
	Transit Lead Time to the Warehouse (L_w) ↑	Decreases ↓	Decreases ↓	Decreases ↓	Decreases ↓	Lead Times are Negligible
	Transit Lead Time to Each of Retail Stores (L_r) ↑	Increases ↑	Decreases ↓	Increases ↑	Decreases ↓	
	Number of Retail stores (N) ↑	Decreases ↓	Increases ↑	Decreases ↓	Increases ↑	Many Retail Stores Served by one Warehouse
Comments		P* decreases only slightly. Smallest observed P*> 0.8		P* and Pc share the similar pattern except for INV. 0.96<Pc<0.999	The Penalty almost always decreases expect with store number N	

control systems compared. As we show, the system performance in a multi-echelon inventory system setting does depend on where inventory is positioned. Not surprisingly, comparative studies without controlling inventory positioning have mixed results (e.g., Krajewski et al. 1987, Pyke and Cohen 1989, and Rees et al. 1989).

After identifying best positioning points for multiple inventory control systems, we were able to compare the performances of these control systems at their best inventory positioning points. Table 6.20 summarizes the maximum fill-rates observed at the best inventory positioning points for the alternative control systems we studied. The results show that information-sharing almost always improves the customer fill-rate. Using the baseline parameter settings, the fill-rates for EROP and DRP were not statistically differentiable. With time-phased projected demand, the behaviors of EROP and DRP were significantly different. For example, if we know that an additional demand of X unites will occur every X periods at each of the retail stores, DRP can incorporate this information into its replenishment decisions while other control systems can only increase their estimated average demand by one unit and/or absorb extra variances. In this case, DRP performed much better than EROP could have done. (The results for X=10 are shown in Table 6.20. The fill-rate differences were statistically significant at 0.01 level using paired-t test.) This observation echoes a study reported by Jacobs and Whybark (1992). Also note that the ranking list we established for the alternative control systems that we studied at their best positioning points are consistent with the ranking list established by Axsater and Rosling (1994) for similar control systems.

Figure 6.20 **PERFORMANCES OF INVENTORY CONTROL SYSTEMS AT BEST INVENTORY POSITIONING POINTS**

Projected Demand	LROP	ROP	EROP	DRP	"Push"
Average (X=0) INV=150	0.8340	0.8596	0.8615	0.8615	0.8481
Lumpy (X=10) INV=200	N/A	0.9026	0.9064	0.9173	N/A

N/A: Not available.

CHAPTER 7

CONCLUDING REMARKS

This dissertation presents a simulation study that answers the question of where inventory should be positioned to get the best customer fill-rate in a distribution system, given a set of inventory transportation and control system resources. For a one-warehouse multi-retailer distribution, we built a simulation model for identifying the proportion of system-wide inventory required to be positioned at the warehouse to get the maximum customer fill-rate. The model also enabled us to assess the value of the best positioning, as opposed to simply positioning inventory as close to the customer as possible.

The key findings of our study can be summarized as follows. First, inventory should be positioned near the customer to get the high levels of customer fill-rate. Second, the warehouse needs to keep a relatively small proportion of system-wide inventory to get the maximum customer fill-rate. However, the difference between the maximum fill-rate and that of the maximum positioning ratio is very small. Third, information-sharing consistently shifts inventory closer to the customer and improves the maximum customer fill rate. Fourth, our general finding that inventory should be positioned near the customer to get high levels of fill-rate does not change materially with any of the experimental factors we investigated. This finding holds true even as demand uncertainty increases.

These research findings have important theoretical and managerial implications, which we shall discuss in this final chapter as concluding remarks. We have divided this chapter into three sections. Section 7.1 highlights this study's research contributions, placing our study in the context of the existing research literature. Section 7.2 discusses major conclusions as they pertain to managerial implications and applications. Section 7.3 identifies several directions for future research

7.1 RESEARCH CONTRIBUTIONS

This study differs from the traditional approach to the positioning problem in several important aspects. First, we argued that there are good reasons to differentiate the stocking location near the customer (e.g., stores) from the intermediate stocking location (e.g., the warehouse). The traditional approach to multi-echelon inventory problems is to build a single stocking location model first, and then to use this model as a common "building block" for constructing multi-echelon inventory models. We departed from this traditional approach. For example, in our study, warehouse shortages were not treated in the same way as shortages were treated at the retail stores. While shortages at the retail level were backordered, no backorder was allowed at the warehouse. The warehouse shortage information was handled in a decentralized fashion, which allowed both the warehouse and the retailer stores to make their inventory decisions based on the most updated information. This approach closely resembled the distribution practice we observed in industries where Advanced Shipment Notice (ASN) plays an increasingly important role in distribution operations.

Second, we carefully controlled the system workload when different inventory positioning strategies were simulated and evaluated. Previous research on the positioning problem by and large has ignored the system workload. When different positioning strategies were evaluated, their shipment frequencies, or fixed shipment costs, were rarely reported or even recorded. This lack of comparability makes it difficult to draw meaningful conclusions with regard to the benefit from holding inventory at the warehouse. To overcome this problem, we developed a method to decompose the performance difference between two positioning strategies with different shipment frequencies. The decomposition allowed us to assess the penalty for positioning inventory close to the customer more precisely than the results we reported earlier (Whybark and Yang, 1996).

Third, we treated inventory control systems as an experimental factor. Previous studies have largely concentrated on solving the positioning problems for single inventory control systems. Researchers often picked an inventory control system arbitrarily and then tried to identify the best positioning for that specific control system. Many different inventory control systems have been studied, but there has been no experimental design behind these individual efforts. As a result, it is difficult, if not impossible, to compare or reconcile the results of those individual studies. In this dissertation study, we developed a much-expanded experimental design in which inventory control systems were treated as an experimental factor. This approach enabled us to draw conclusions sufficiently generalized that they can be broadly applied.

Fourth, we characterized an inventory control system not only by the form of decision rules used but also by the information with which the decision rules were implemented. The design of the alternative control systems was intended to reflect various information-sharing schemes we observed in industry. This approach shifted emphasis away from the inventory flow exclusively to include both inventory and information flows. Such a design also allowed us to investigate the effect of information sharing, an issue of increasing importance in practice, but one that has rarely been addressed in the existing research literature.

Fifth, we chose the lognormal distribution to generate customer demand, which allowed various demand uncertainty scenarios to be included in our simulation experiments. We find that the benefit from holding some inventory at the warehouse began to diminish as the coefficient of variation of demand past 1. This was a surprising result, indicating that Jackson's (1988) conjecture on the effect of demand uncertainty (i.e., the benefit of holding some inventory at the warehouse becomes increasingly significant as the CV goes up to the levels of two, three, and higher) might not be accurate.

Using these approaches, we demonstrated that as long as the fill-rate is an appropriate service criterion, inventory should be positioned near the customers. This general finding is remarkably robust with respect to the changes in three resource factors (system-wide inventory level, predetermined shipment frequencies, and the control system used), and three environmental factors (demand uncertainty, the transit lead times, and the number of retail stores served by the warehouse).

It is worth pointing out that the theoretical implications of this study are not limited to the positioning problem itself. The question of which inventory control system should be used to manage inventory in a multi-echelon distribution system setting by and large remains an open issue in the research literature. Using the modeling framework developed in this dissertation, we were able to compare different multi-echelon inventory control systems at their best inventory positioning points. We found that DRP is the best.

7.2 MANAGERIAL IMPLICATIONS

In this study, we demonstrated that excessive inventory can mask the need to address the inventory positioning problem. As more and more companies streamline their supply chains, we believe that the proper positioning of inventory will become an increasingly important lever for companies to gain competitive advantages in the logistics battleground.

Our results show that the positioning ratio that provided the maximum fill-rate was at the high levels of greater than 0.8, and most often substantially higher. More than 80% or more of the inventory should be positioned near the customer. Holding less inventory at the warehouse than the amount required by the best positioning resulted in only a fairly small decrease in the fill-rate. On the other hand, positioning more inventory at the warehouse than the amount required by the best positioning could lead to substantial deterioration in customer fill-rate.

The guidelines for managers are clear: never position too much inventory at the warehouse; positioning inventory close to the customer at retail stores is the practical approach to getting high levels of customer fill-rate. As long as the fill-rate is an appropriate service criterion, the rule of thumb is that positioning inventory close to the customer is a good strategy. This general finding is remarkably robust to the changes in various experimental factors we investigated. We believe that the insight from this study should help distribution managers deal with their positioning problems easily and effectively.

Another important finding of considerable managerial importance is the effect of information-sharing on the positioning of inventory in the distribution system. Contrary to general belief, our simulation results consistently show that the more information that becomes available and is utilized by the control system used, the less inventory needs be held back at the warehouse.

Interestingly, our simulation results seem to be consistent with what is happening in industry. Information sharing lies at the core of many of the 1990s highly touted supply chain management initiatives. Reengineering, quick response, efficient consumer response, vendor-managed inventory, and continuous replenishment programs all require sharing of information among supply chain participants in one way or another. Armed with newly developed information technologies, positioning practices seem to have already begun to change. "We've come full circle with the distribution center where 10 years ago you saw people having everything. The distribution center is now getting to the point that you are trying to move everything directly to the store," says David Lynn, distribution vice president at McCrory Stores (Chain Store Age Executive, April, 1992, P31A). As noticed earlier, in the grocery industry, one of the principal directions identified for Efficient Consumer Response (ECR) is "to focus the industry on distributing products rather than warehousing them" (Progressive Grocer, January 1994, p.6). That less inventory is positioned at distribution centers is also evident when we look at the aggregate square footage of distribution centers. According to a 1992 survey, one in five retailers (20%) planned to shrink its square footage in distribution center space by 1993. One in eight (12%) planned to cut more than 100,000 square footage or more during that period (Chain Store Age Executives, April 1992). These industrial observations should come as no surprise. The managers' learning process is based on a real world simulation where the measure of service criteria happens most frequently to be the fill-rate.

7.3 DIRECTIONS FOR FUTURE RESEARCH

Upon completion of this dissertation study, we have become convinced that there is need for even more research on the inventory positioning problem. Specific topics that immediately come to mind include research of non-identical retail stores. Another research project could concern itself with different demand distribution assumptions. However, there is presently research (Wheng, 1994) that shows that the benefit from positioning inventory at the warehouse with identical retail stores is always larger than that with non-identical retail stores. It is also well known that lognormal density function appears to be thicker in the right tail than most other distribution density functions with the same coefficient of variation. The implication is that the probability of having to deal with the "unbalanced" retail inventories under lognormal distribution is larger than that under other demand distributions. Therefore, we suspect that the major conclusions reached in this study will not change materially even with these suggested research expansions.

We believe that important directions for the future research lie in three areas. First, we should expand our simulation model to include (1) different service criteria, full factorial design in conjunction with the new service criteria, (3) stochastic transit lead times, and (4) additional inventory control systems. Using the simulation framework developed in this study should help us gain invaluable insights on the effect of many new industrial initiatives on the positioning of inventory in the distribution networks. The industrial initiatives with high priorities in our research agenda include the Vendor Managed Inventory, Just-in-Time Purchasing, and Continuous Replenishment Programs.

Another area for future research is to develop analytical explanations for helping to understand why the fill-rate improvement increases as the coefficient of variation (CV) approaches 1.0 and decreases as the CV exceeds 1.0. We also believe that the effect of information sharing on the positioning of inventory can be demonstrated analytically.

Finally, future research should be directed to one important area: comparative study on multi-echelon inventory control systems at their best positioning points. It is important to measure and control the resources that are consumed in an inventory system (Yang and Whybark 1994). That is also the only way we can compare and understand other researchers' work.

FIGURES

Figure 2.1 A One-Warehouse Multi-Retailer Distribution System

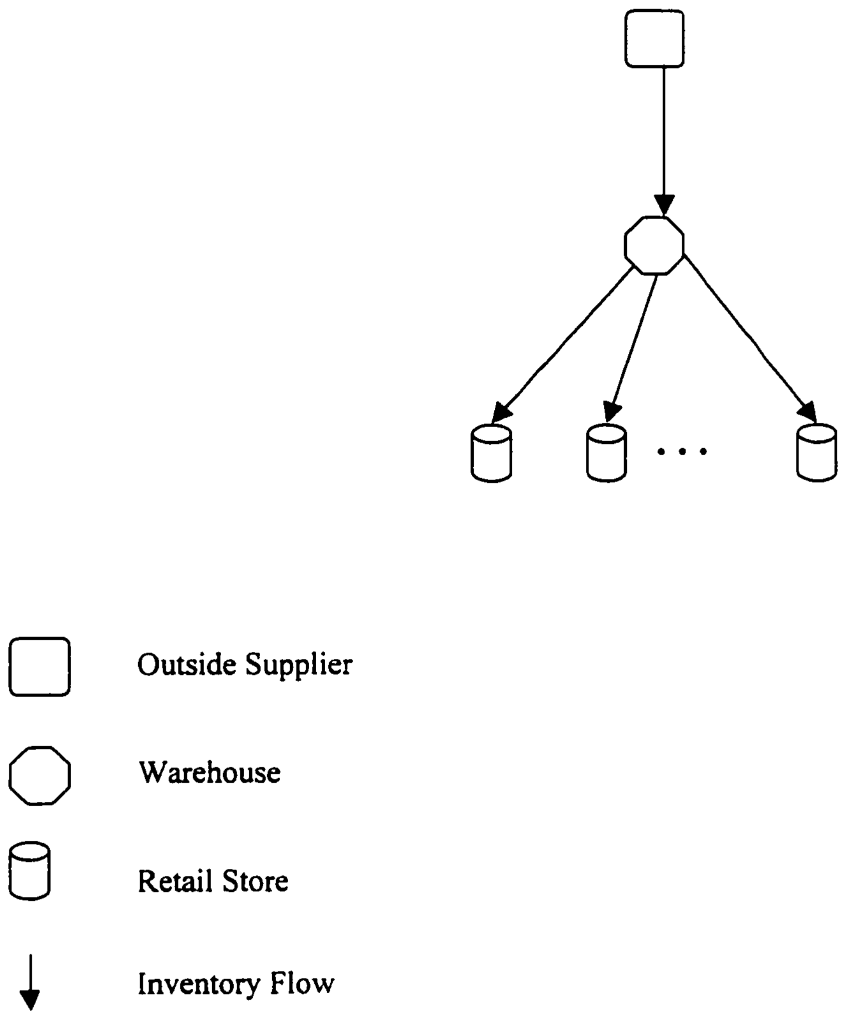
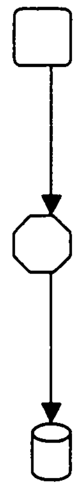


Figure 2.2 Serial Distribution System



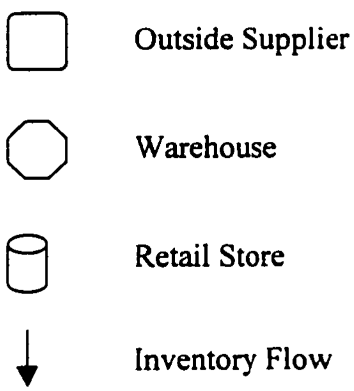


Figure 2.3 Serial Production System

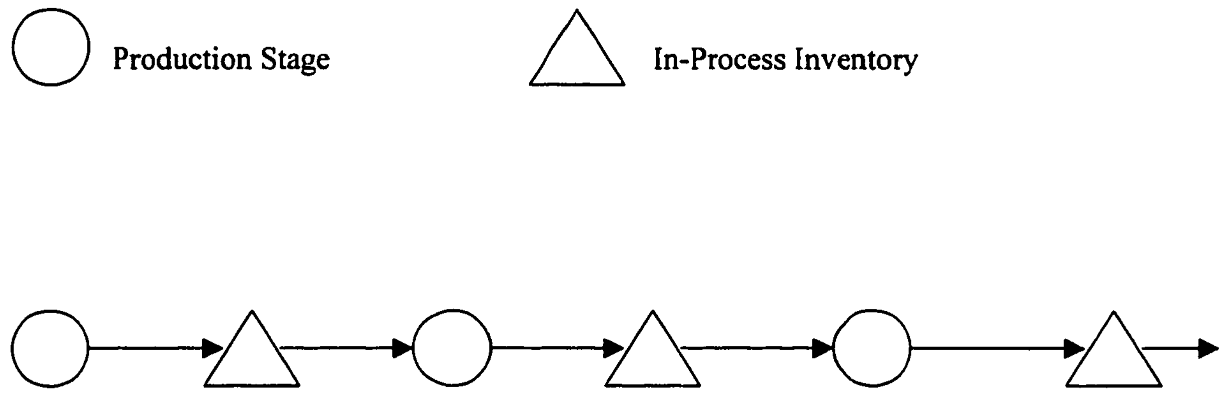


Figure 2.4 Three stage production/inventory system

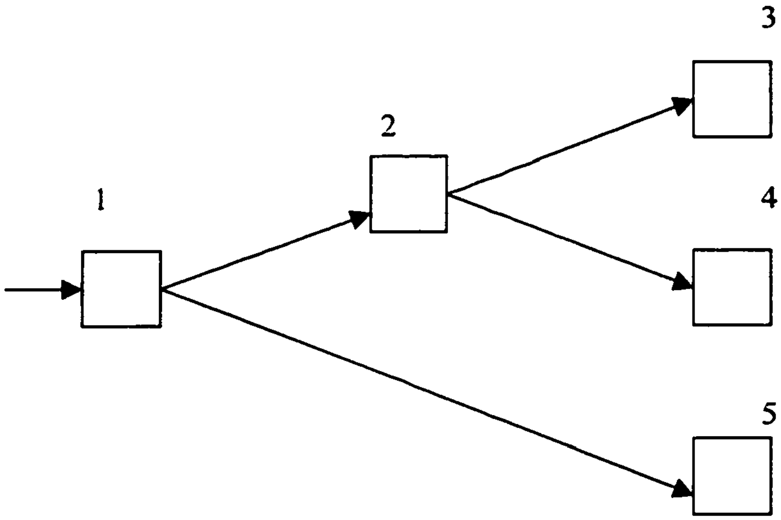


Figure 2.5 Two-Level Assembly System

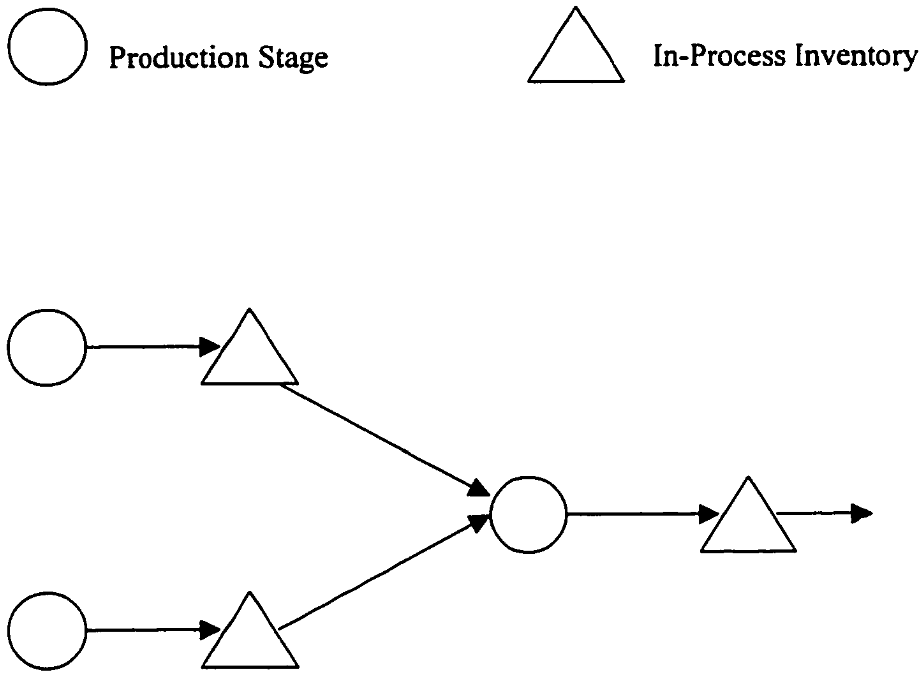


Figure 4.1 Information Availability Scenario 1

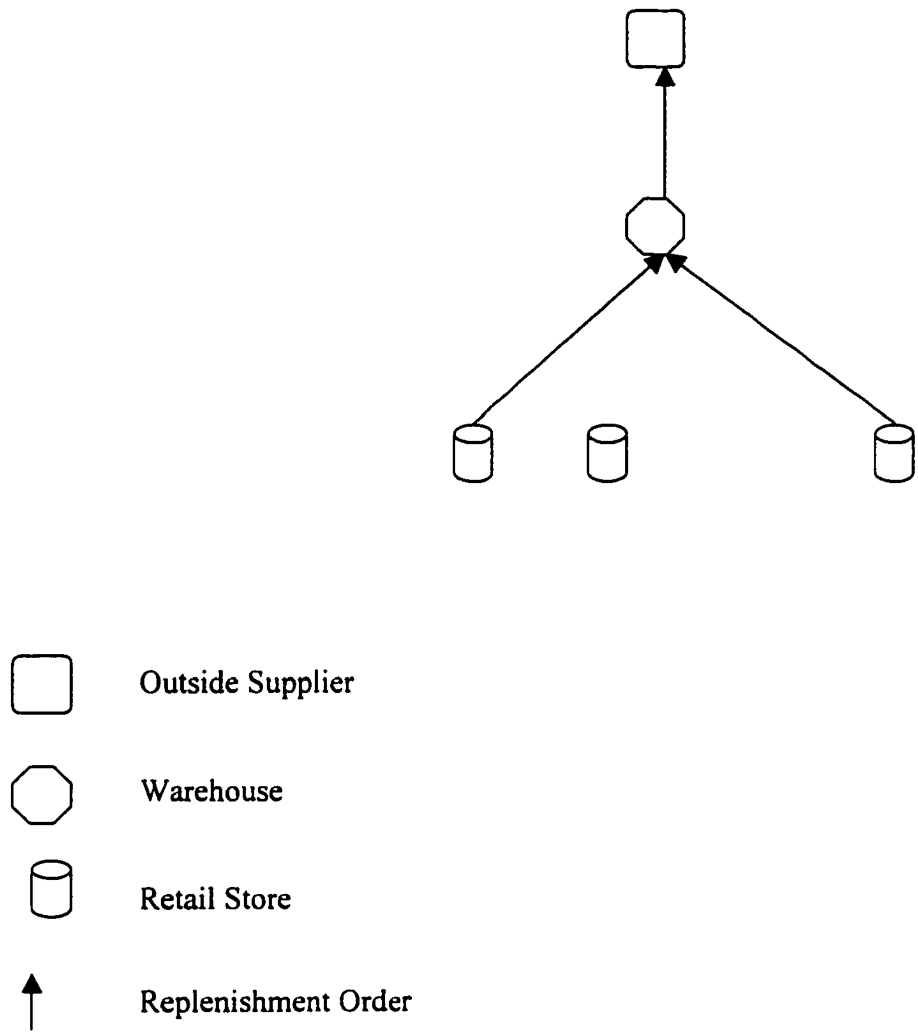
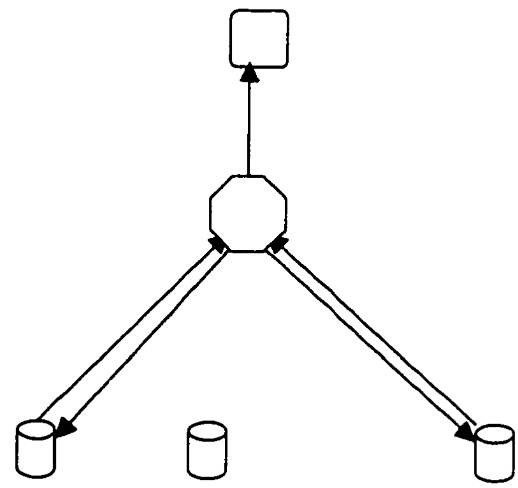


Figure 4.2. Information Availability Scenario 2



Outside Supplier

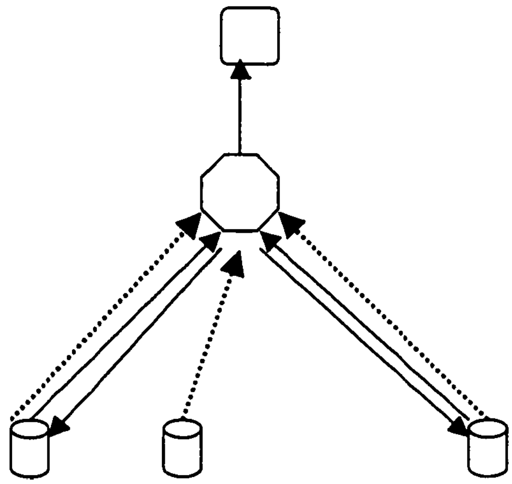
Warehouse

Retail Store

Replenishment Order

Advanced Shipment Notice (ASN)

Figure 4.3. Information Availability Scenario 3



Outside Supplier

Warehouse

Retail Store

Replenishment Order

Advanced Shipment Notice (ASN)

Retail Inventory Positions (RIP)

Figure 4.4. Information Availability Scenario 4

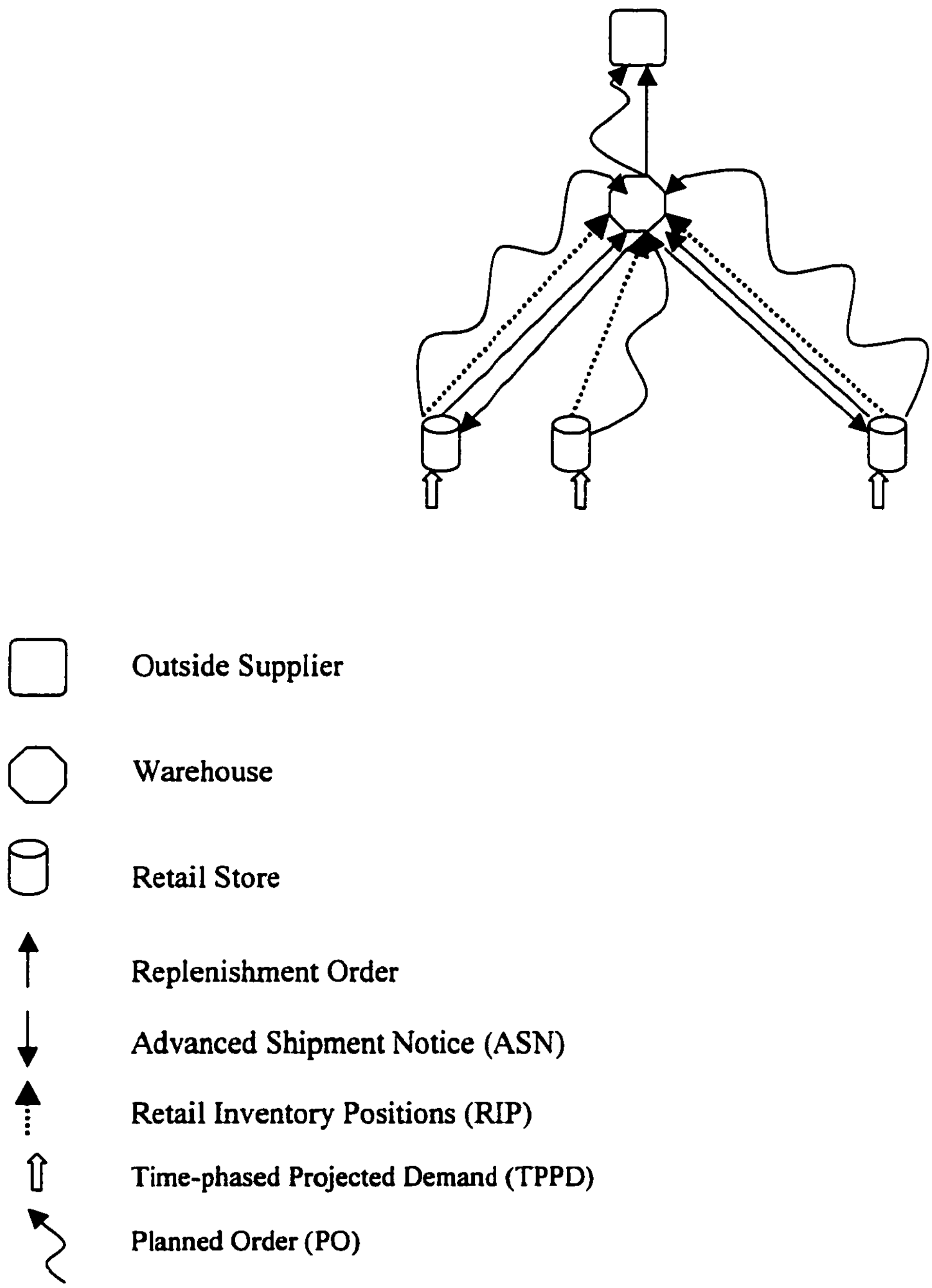


Figure 5.1: TYPICAL DENSITY FUNCTION EXPERIENCED IN PRACTICE

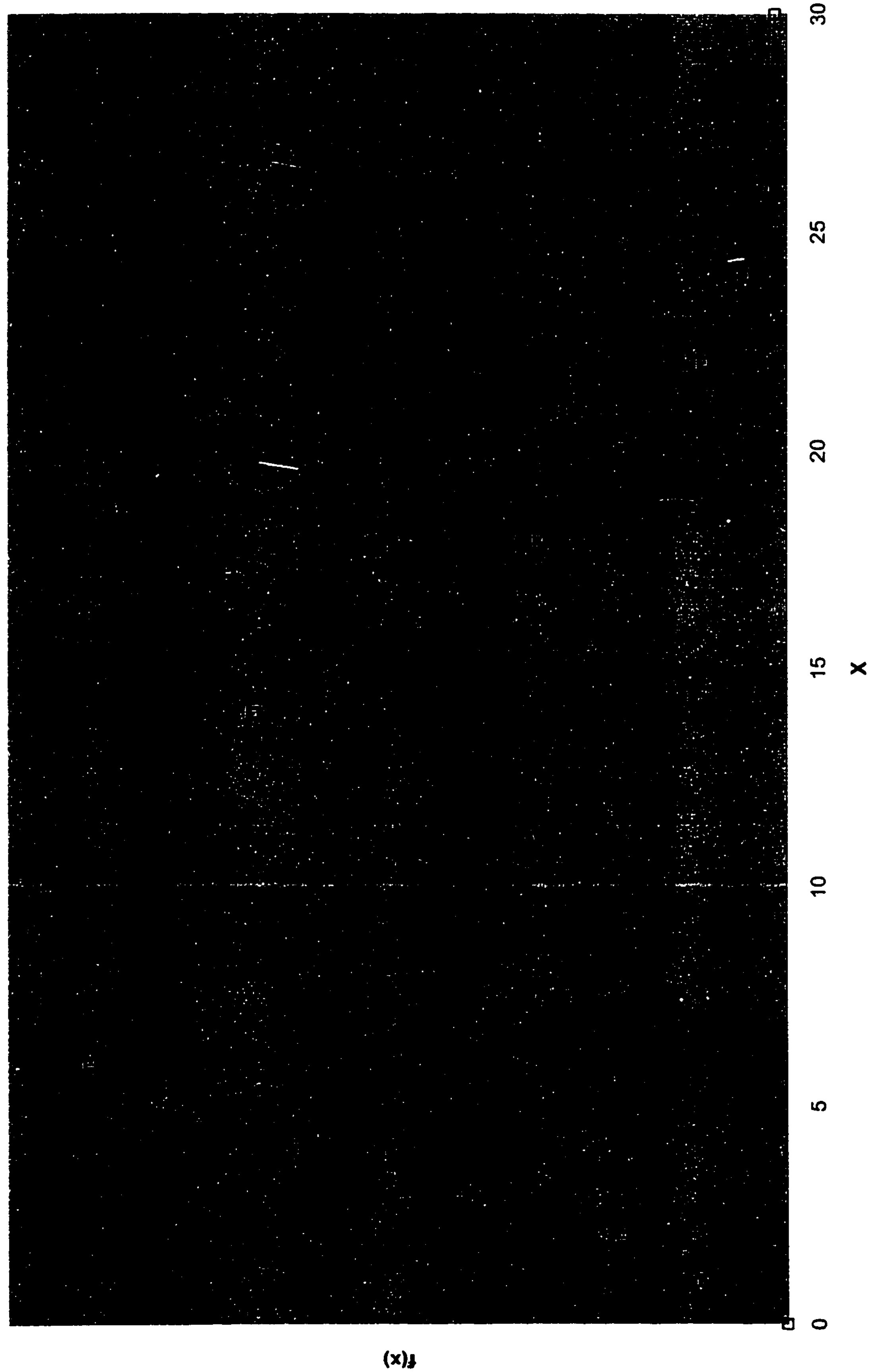


Figure 5.2 Factorial Experiment, No Interaction

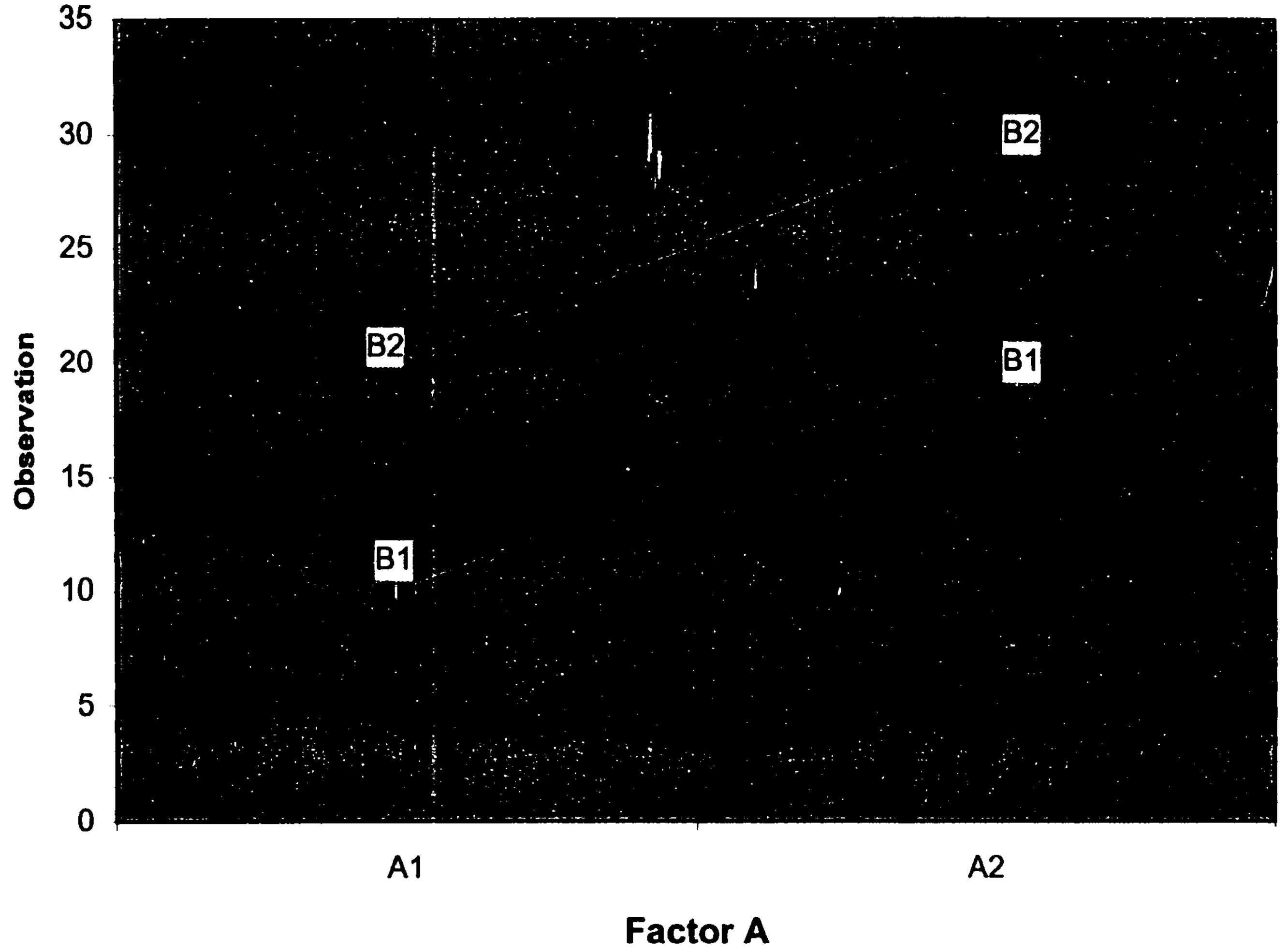


Figure 5.3 Factorial Experiment, With Interactions

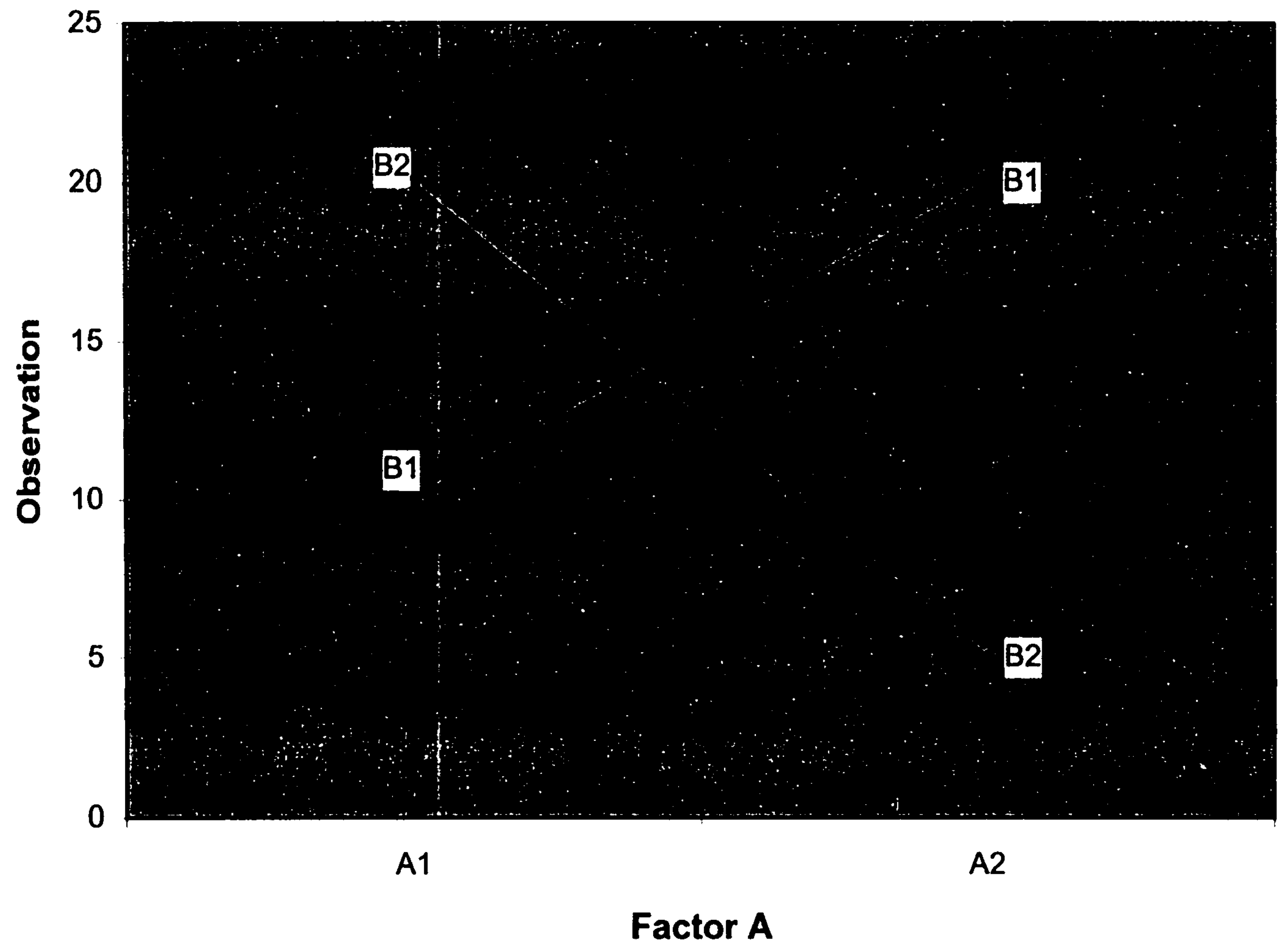


Figure 6.1: BASELINE STUDY
(Fill Rate vs. Positining Ratio)

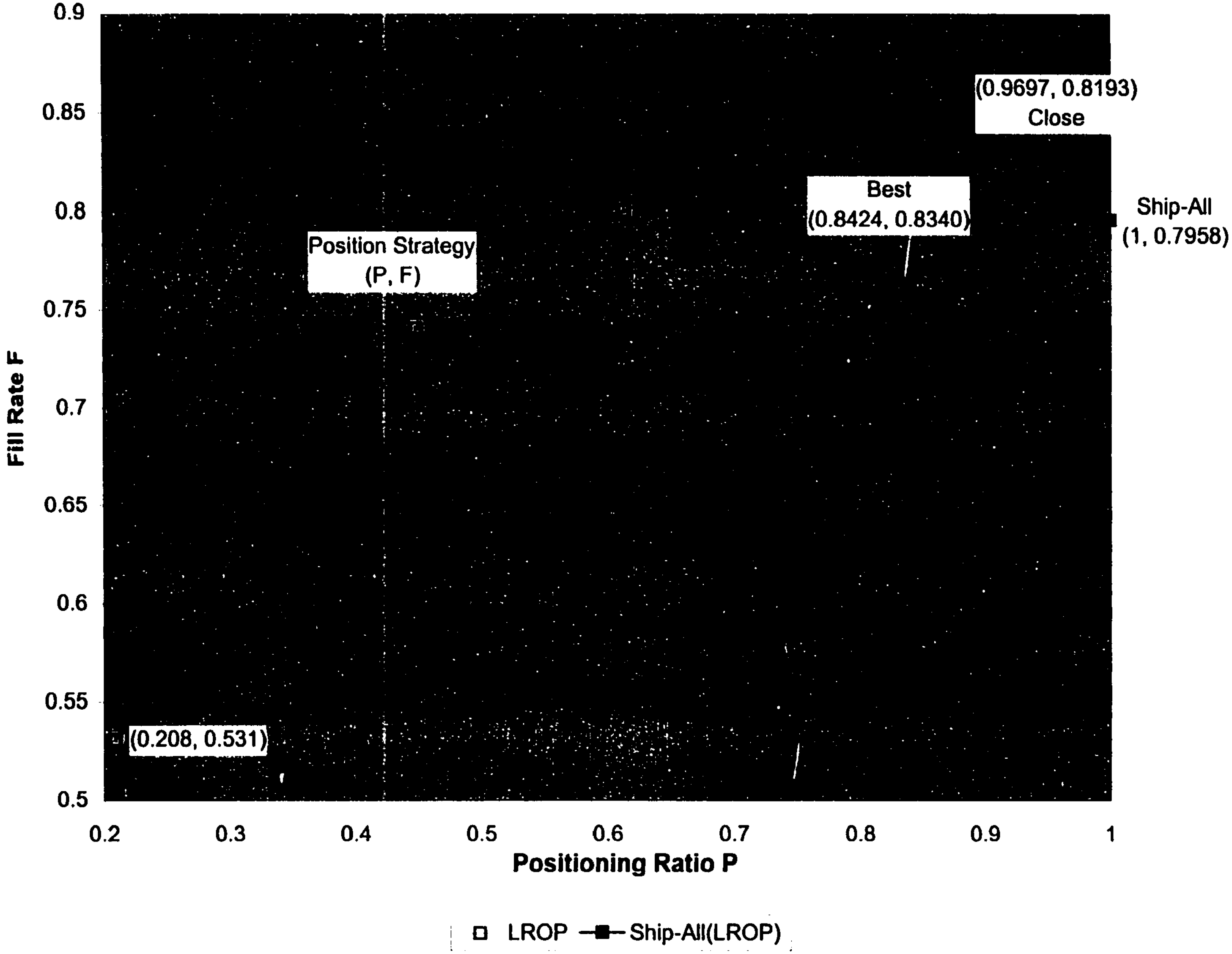
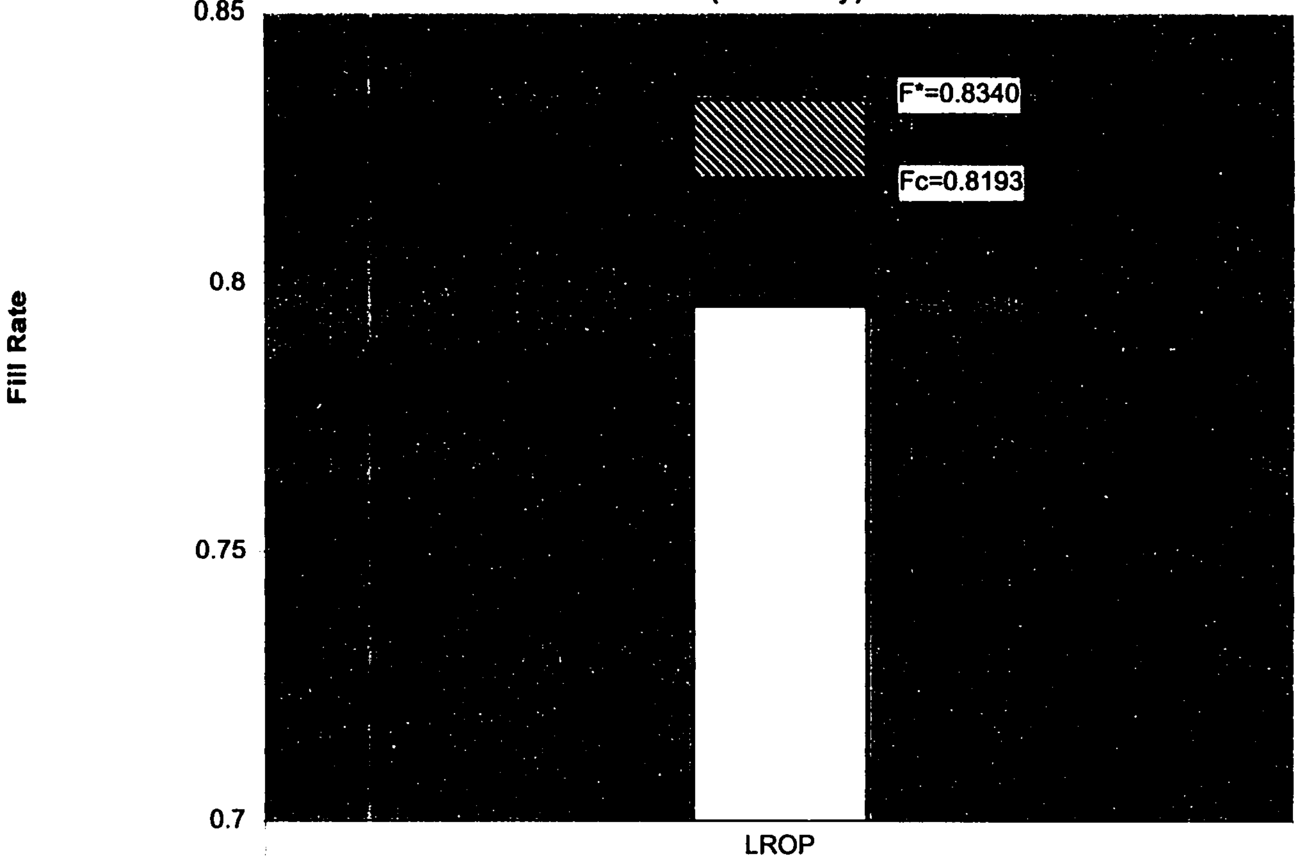


Figure 6.2: BASELINE STUDY
(Summary)



	LROP
Best Positioning Ratio	0.8424
Due to Best Positioning	0.0147
Due to Increase in SHIPR	0.0235
Ship-All	0.7958

Figure 6.3: EFFECT OF CONTROL SYSTEMS
(Baseline Parameter Settings)

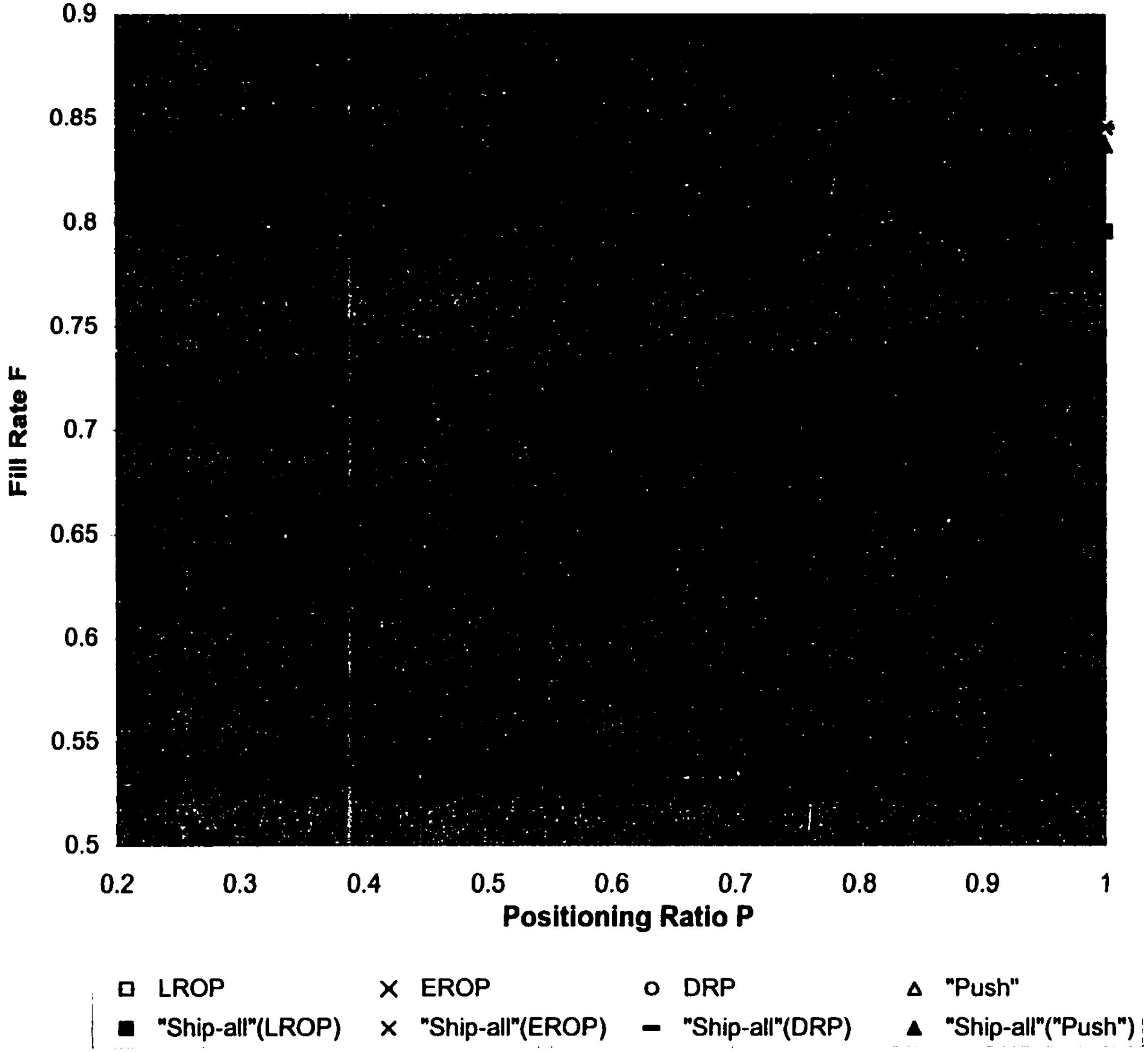
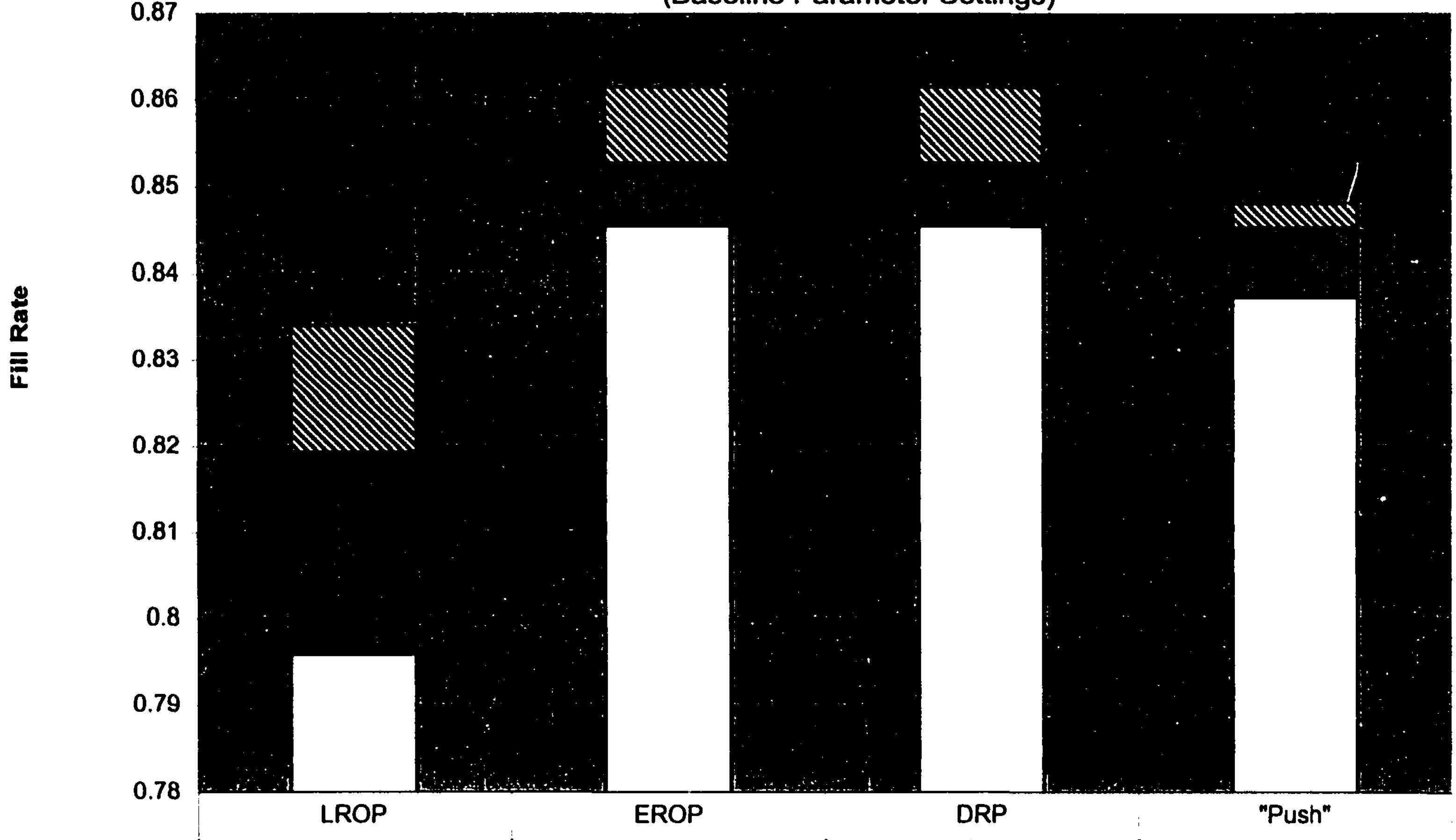


Figure 6.4: CONTROL SYSTEM COMPARSION
(Baseline Parameter Settings)



	LROP	EROP	DRP	"Push"
Best Positioning Ratio	0.8424	0.9218	0.9208	0.9531
▧ Due to Best Positioning	0.0147	0.0088	0.0088	0.0028
■ Due to Increase in SHIPR	0.0235	0.0072	0.0072	0.008
□ Ship-All	0.7958	0.8455	0.8455	0.8373

Figure 6.5: EFFECT OF CONTROL SYSTEMS
(INV=200)

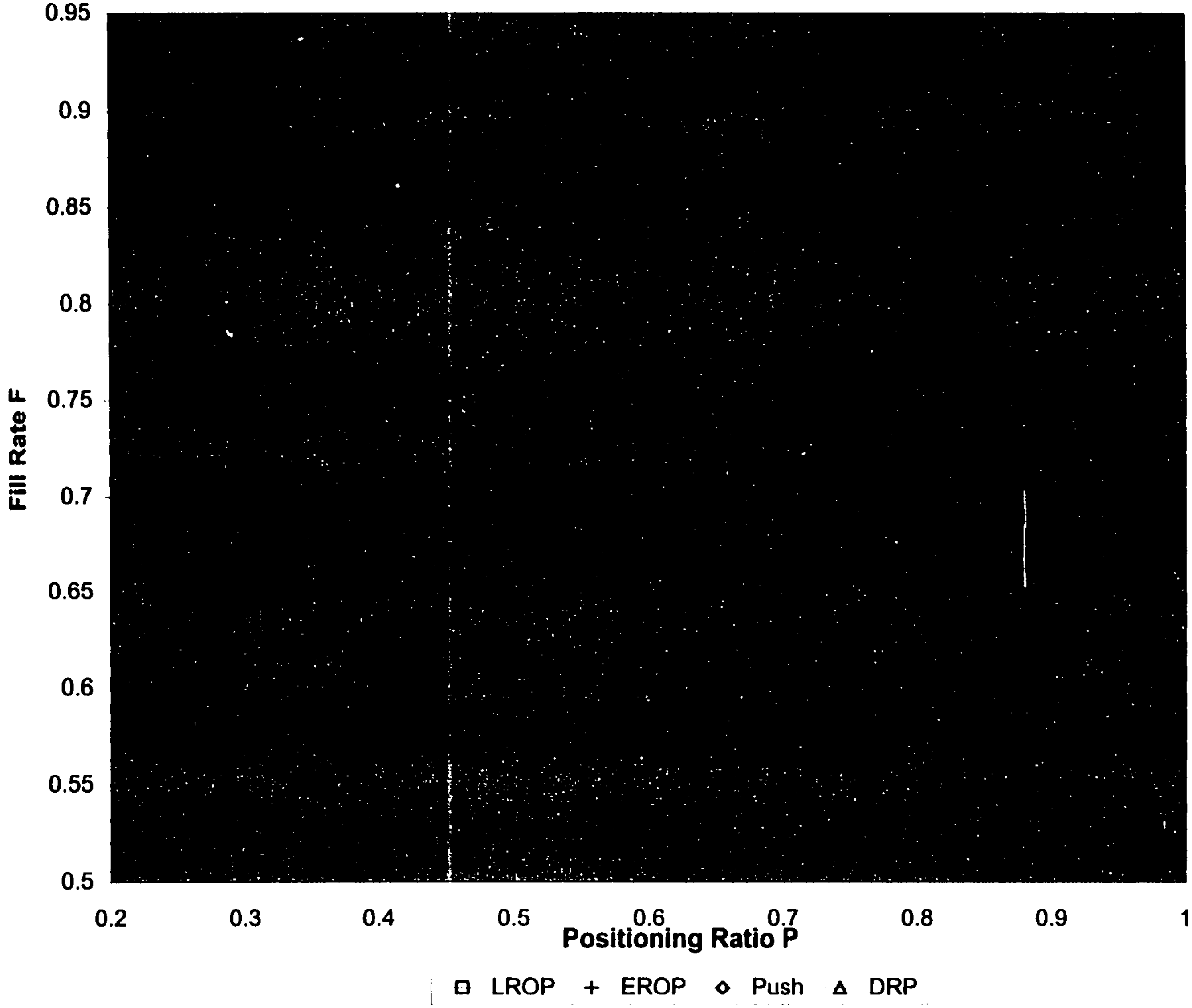
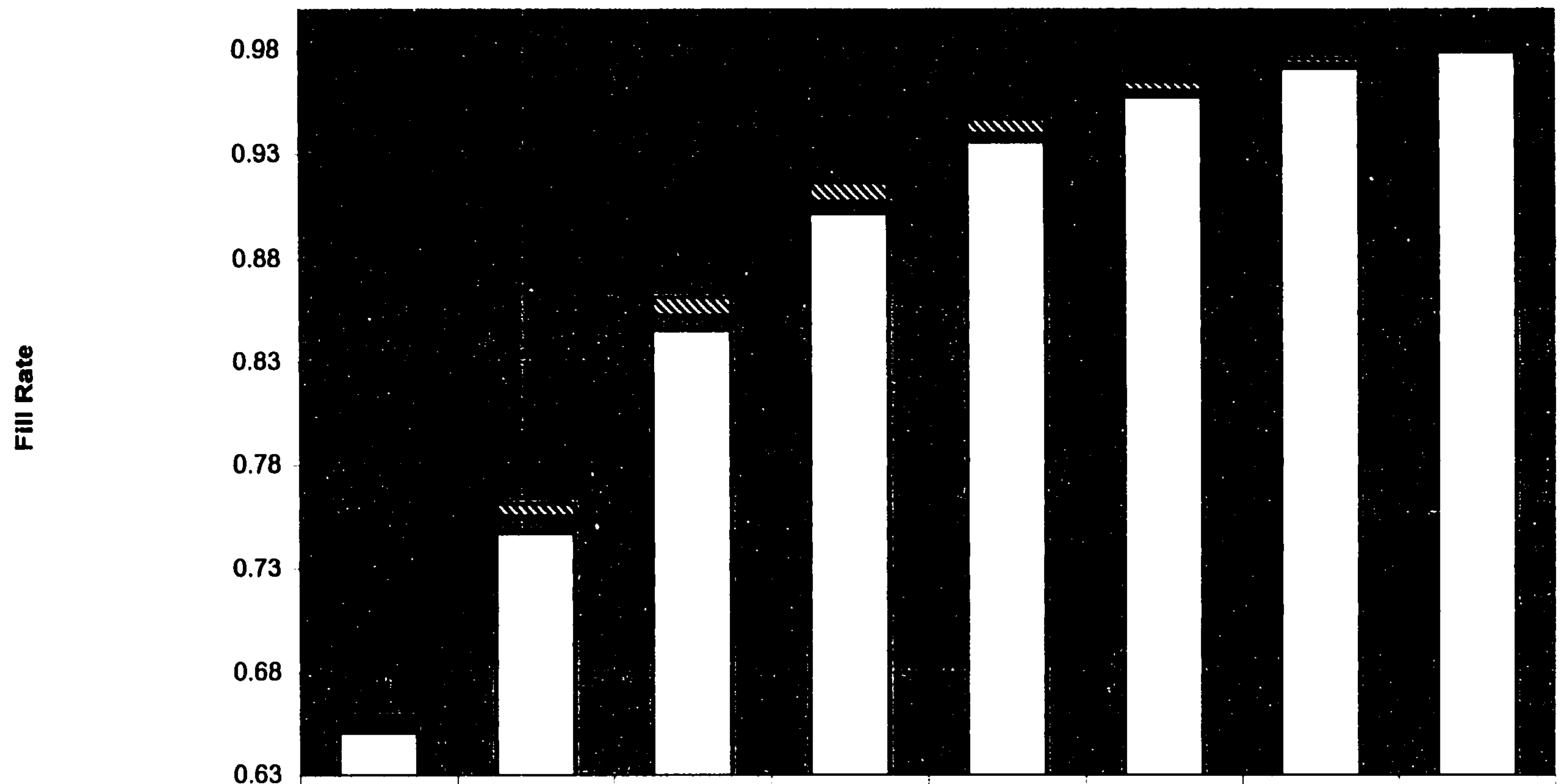


Figure 6.6: EFFECT OF INV



	70	100	150	200	250	300	350	400
Best Positioning Ratio	0.9307	0.9001	0.9218	0.9117	0.9091	0.9187	0.9197	0.9059
Due to best positioning	0.0018	0.0058	0.0088	0.0089	0.007	0.0047	0.0027	0.0021
Due to increase in SHIPR	0.0066	0.0082	0.0072	0.0058	0.0037	0.0026	0.0021	0.0011
Ship-All	0.6506	0.7471	0.8455	0.9018	0.9361	0.9574	0.971	0.9795

System-wide Inventory Level (INV)

Figure 6.7: EFFECT OF SYSTEM-WIDE INVENTROY LEVEL

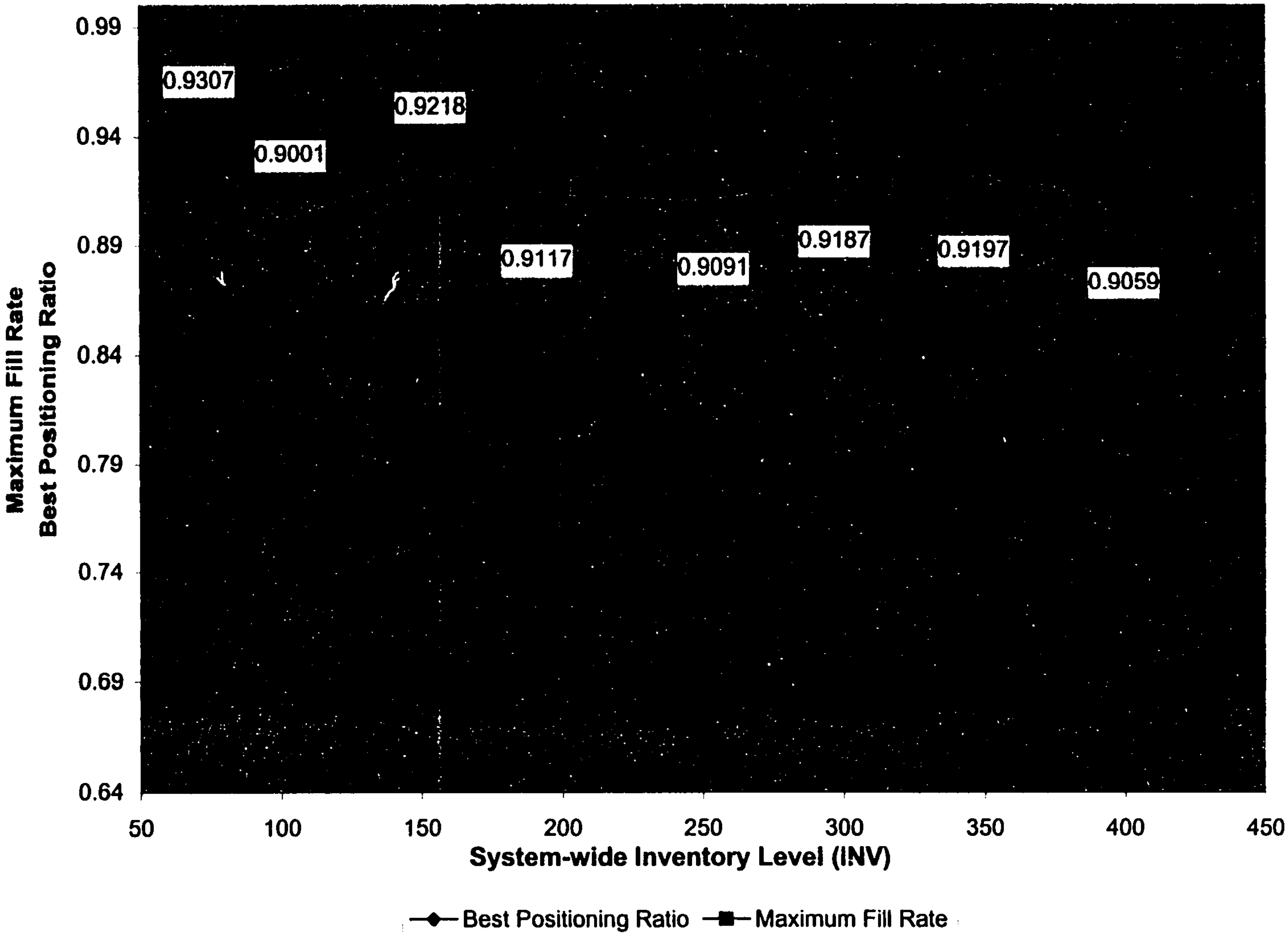


Figure 6.8: BEST POISTIONING RATIO vs. INV

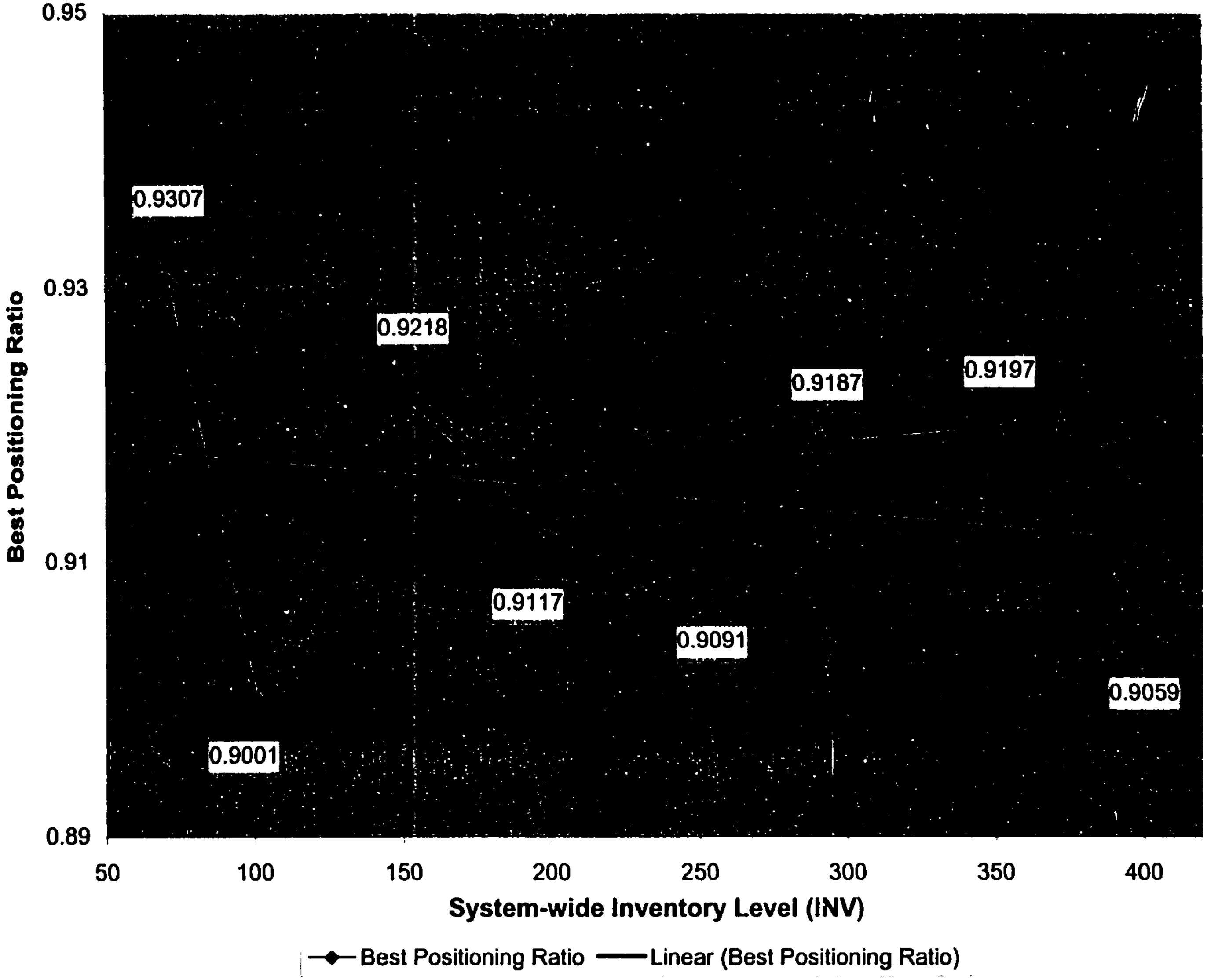
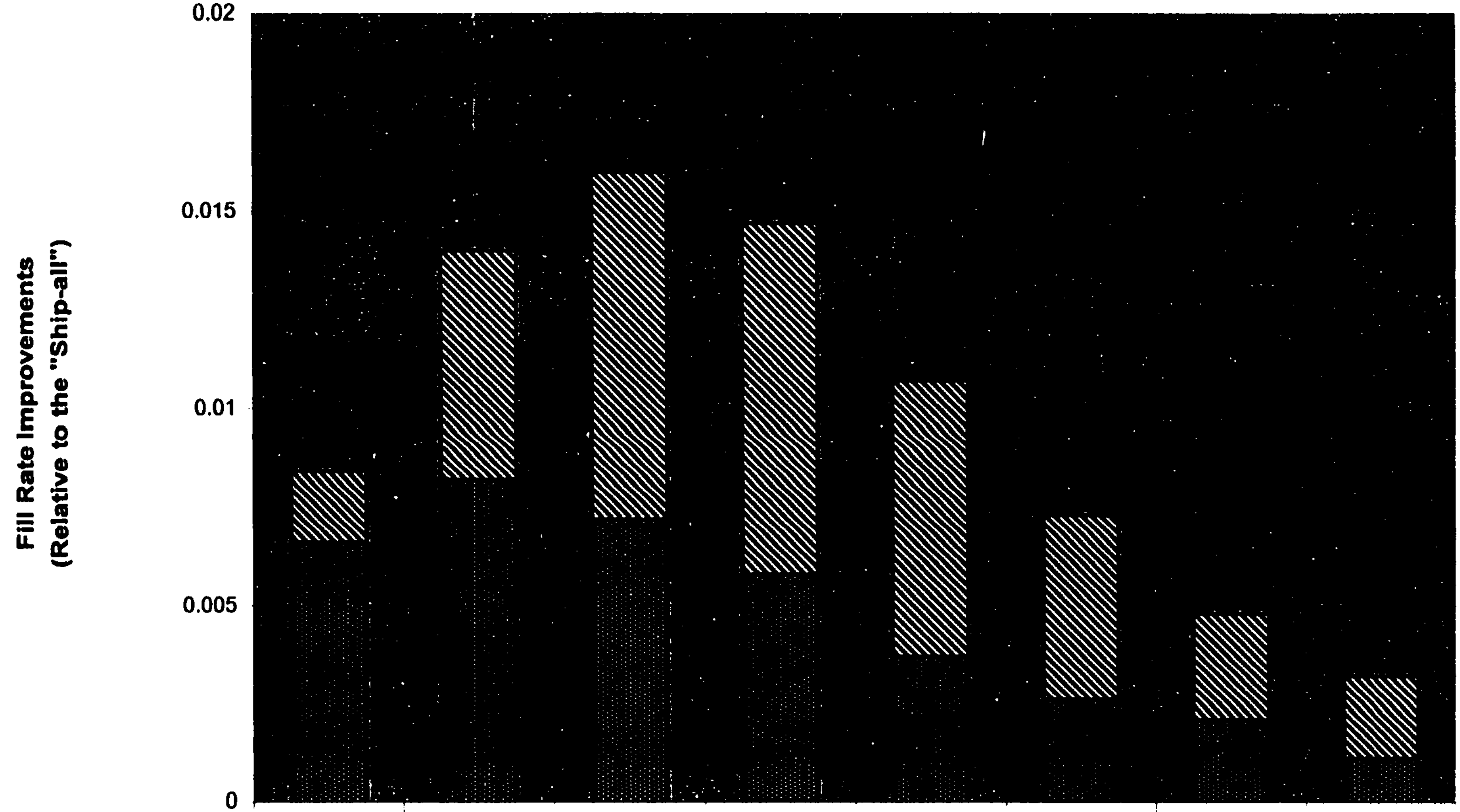


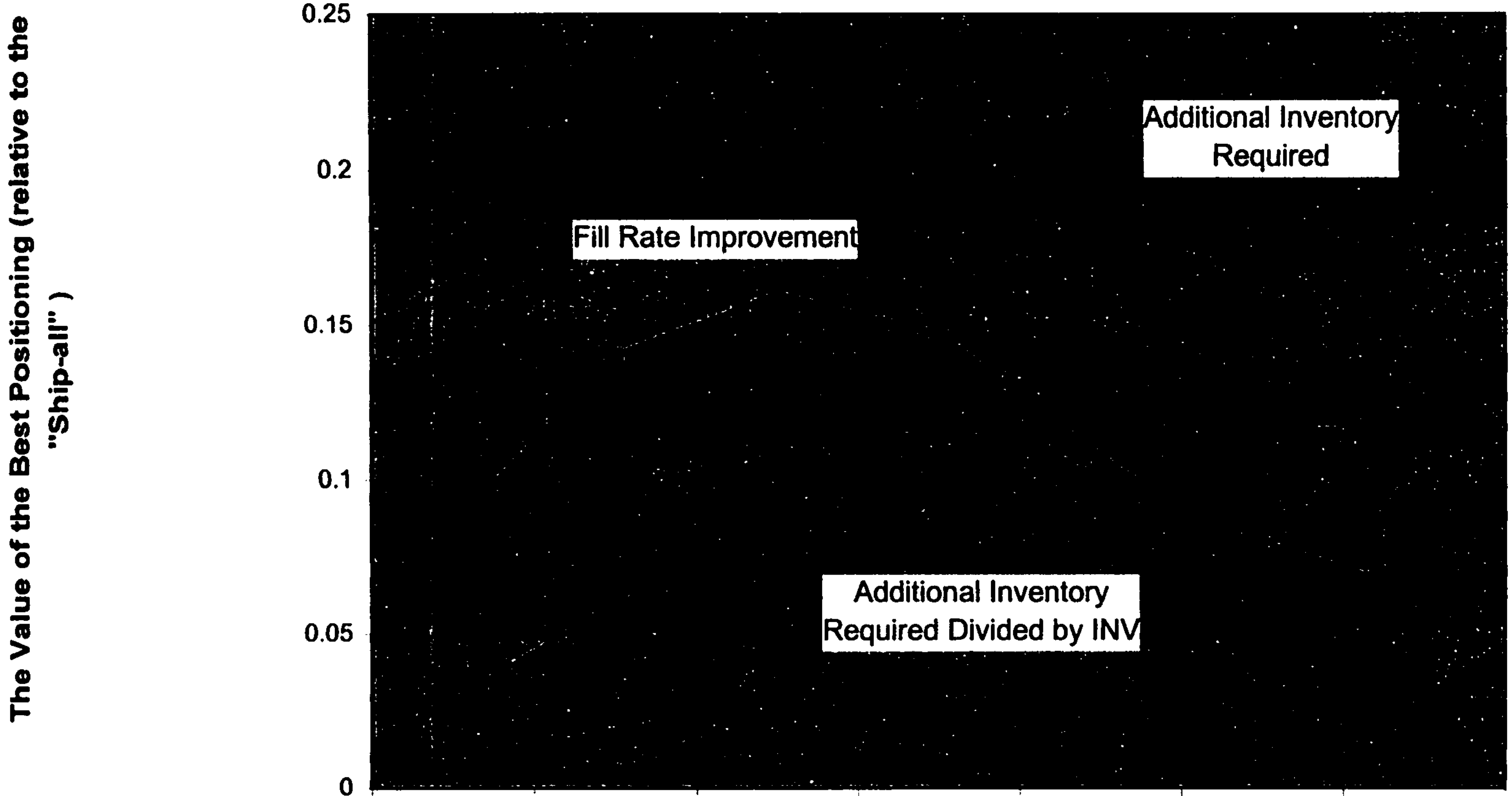
Figure 6.9: EFFECT OF SYSTEM-WIDE INVENTORY LEVEL (INV)
(Fill Rate Improvements relative to the "Ship-all")



	70	100	150	200	250	300	350	400
Best Positioning Ratio	0.9307	0.9001	0.9218	0.9117	0.9091	0.9187	0.9197	0.9059
Due to best positioning	0.0018	0.0058	0.0088	0.0089	0.007	0.0047	0.0027	0.0021
Due to increase in SHIPR	0.0066	0.0082	0.0072	0.0058	0.0037	0.0026	0.0021	0.0011

System-wide Inventory Level (INV)

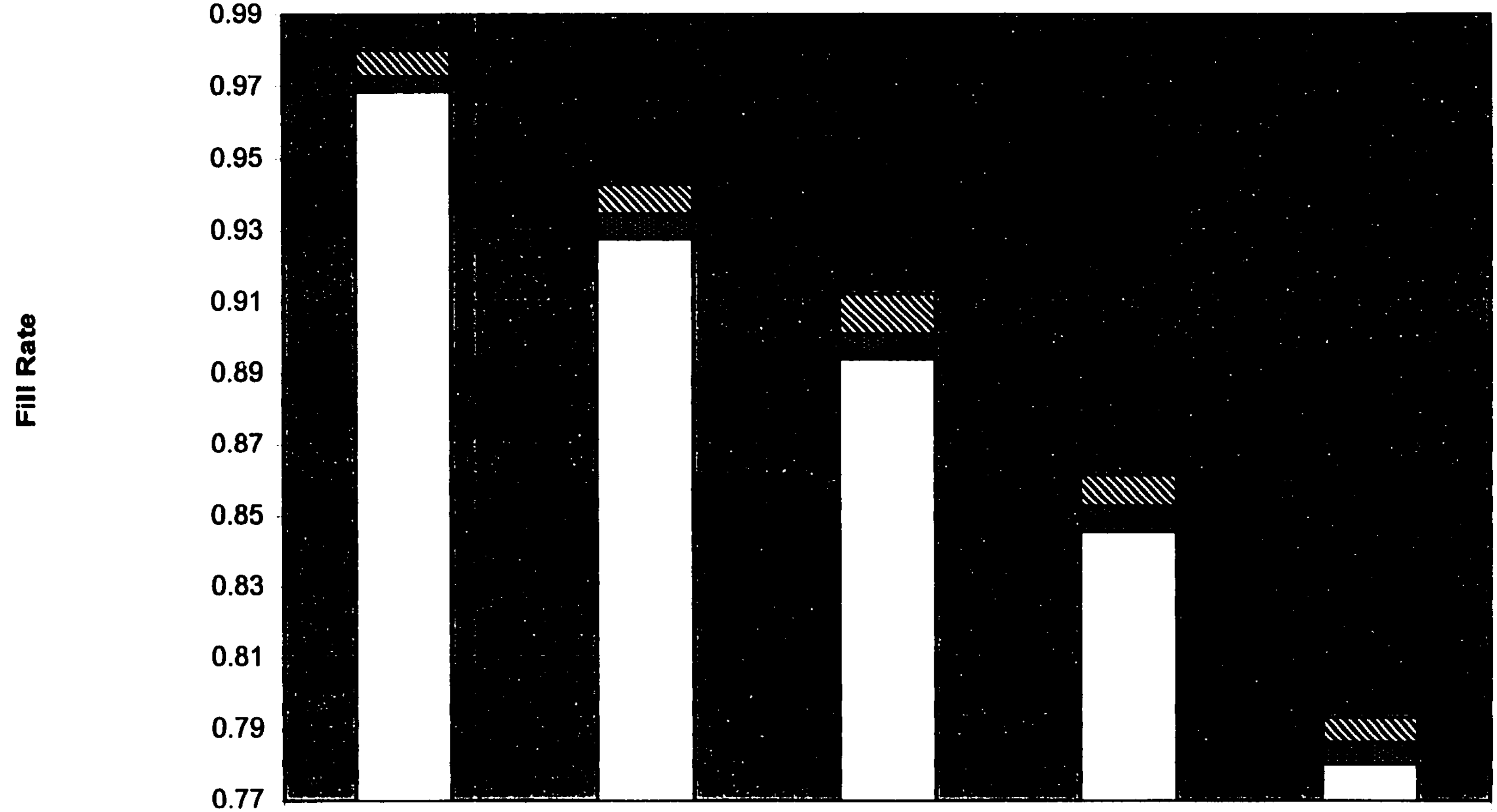
Figure 6.10: FILL RATE IMPROVEMENT vs. ADDITION INVENTORY REQUIRED
(for the fill rate of the "Ship-All" to match that of the best positioning)



	70	100	150	200	250	300	350
Fill Rate ImprovementX10	0.084	0.14	0.16	0.147	0.107	0.073	0.048
Additional Inventory Required X0.01	0.02	0.06	0.1147	0.18	0.23	0.24	0.22
Additional Inventory Required/INV	0.0286	0.06	0.0765	0.09	0.092	0.08	0.0629

System-wide Inventory Level (INV)

Figure 6.11: EFFECT OF CV
(INV=150)



	CV=0.53	CV=0.80	CV=1.00	CV=1.31	CV=1.80
Best Positioning Ratio	0.9183	0.9212	0.9081	0.9218	0.9013
Due to "best" poistioning	0.0074	0.0084	0.0113	0.0088	0.0069
Due to increase in SHIPR	0.0042	0.0069	0.0069	0.0072	0.0059
Ship-All	0.9686	0.9277	0.8939	0.8455	0.7802

Figure 6.12: EFFECT OF CV
(INV=150, Trend of Changes in Best Position Ratio P*)

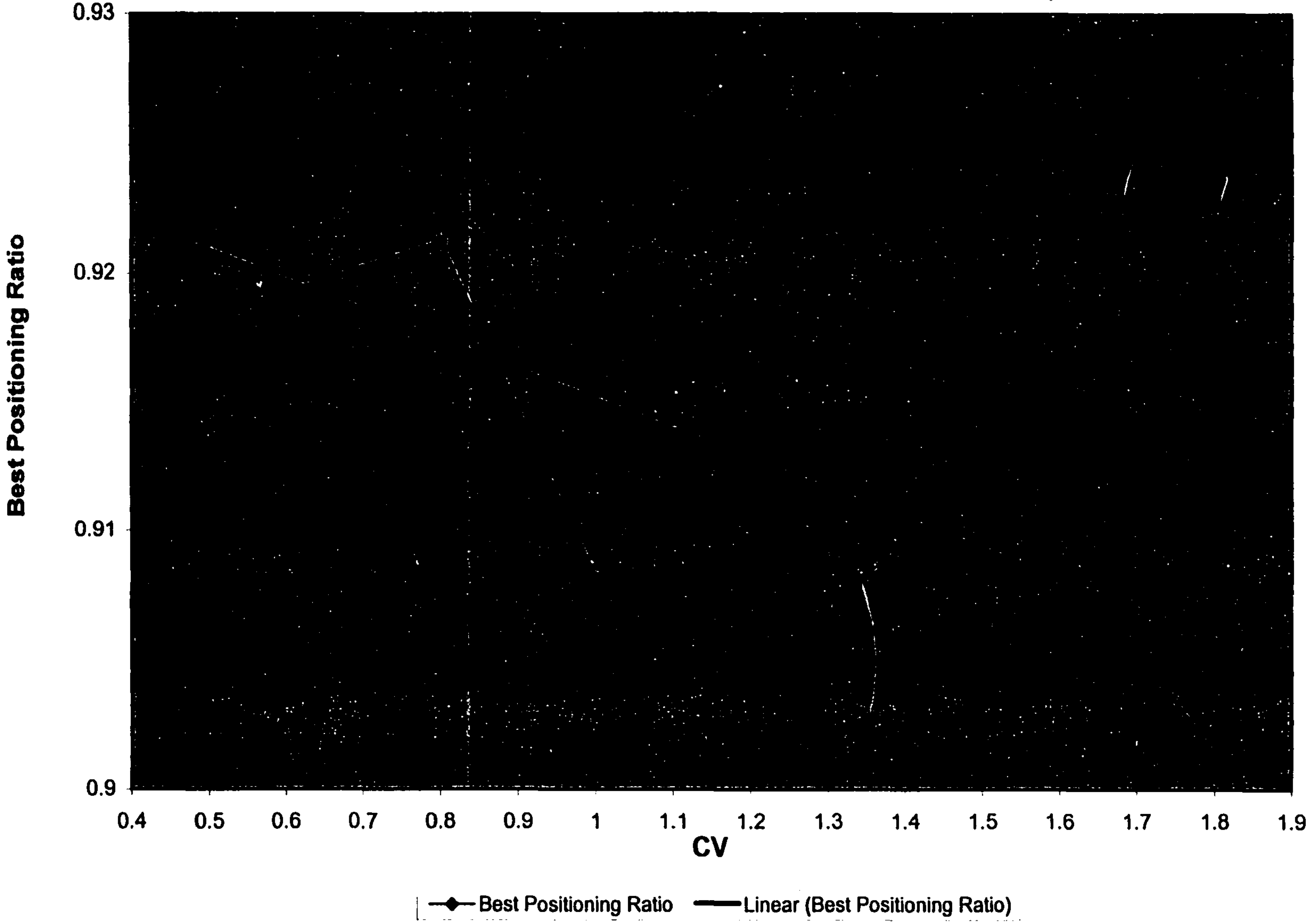
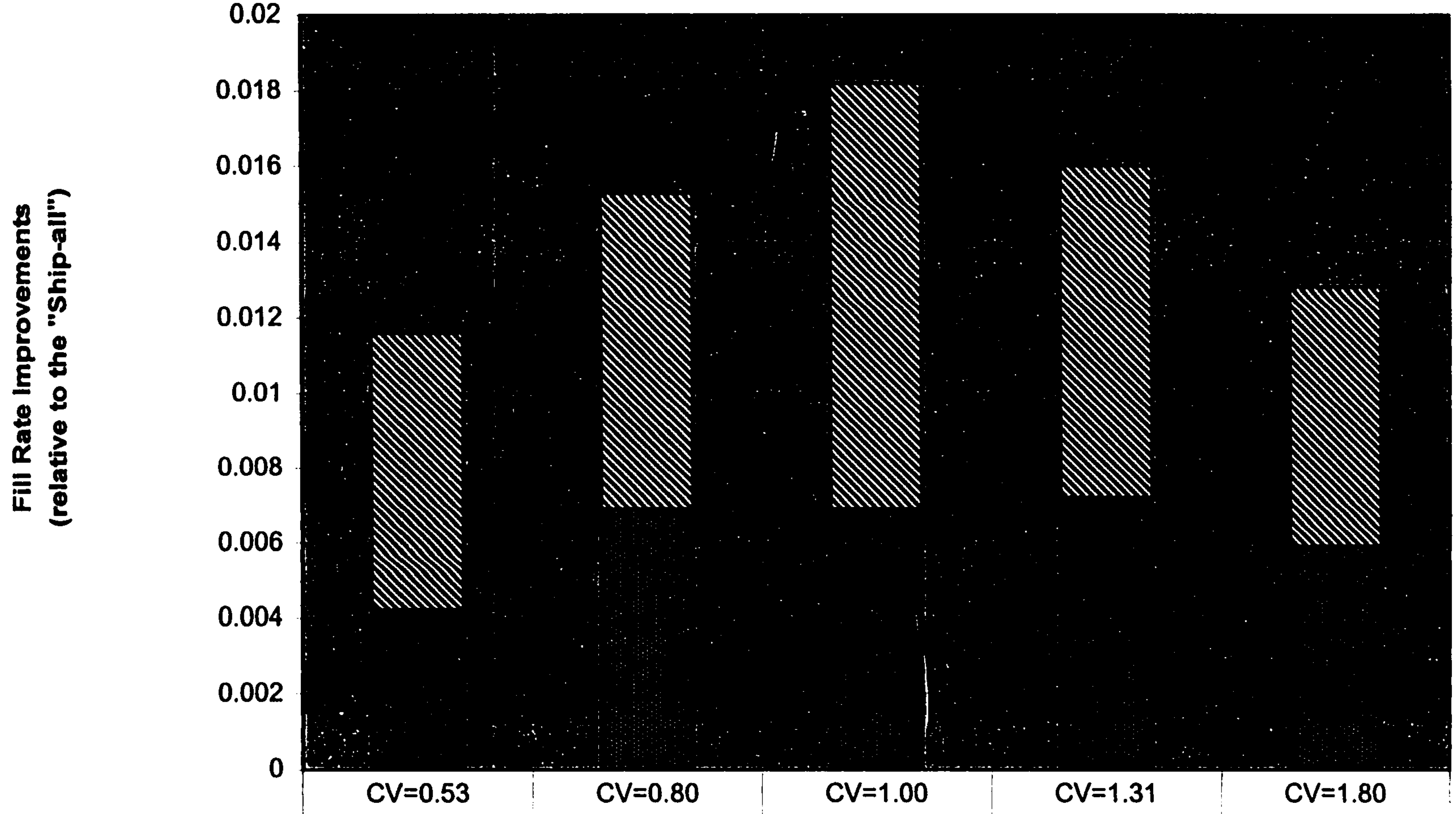


Figure 6.13:EFFECT OF CV
(INV=150, Detailed Fill Rate Improvements)



	CV=0.53	CV=0.80	CV=1.00	CV=1.31	CV=1.80
Best Positioning Ratio	0.9183	0.9212	0.9081	0.9218	0.9013
Due to "best" poistioning	0.0074	0.0084	0.0113	0.0088	0.0069
Due to increase in SHIPR	0.0042	0.0069	0.0069	0.0072	0.0059

Figure 6.14: SCREEN FOR TWO-WAY INTERACTIONS
Between CV and INV

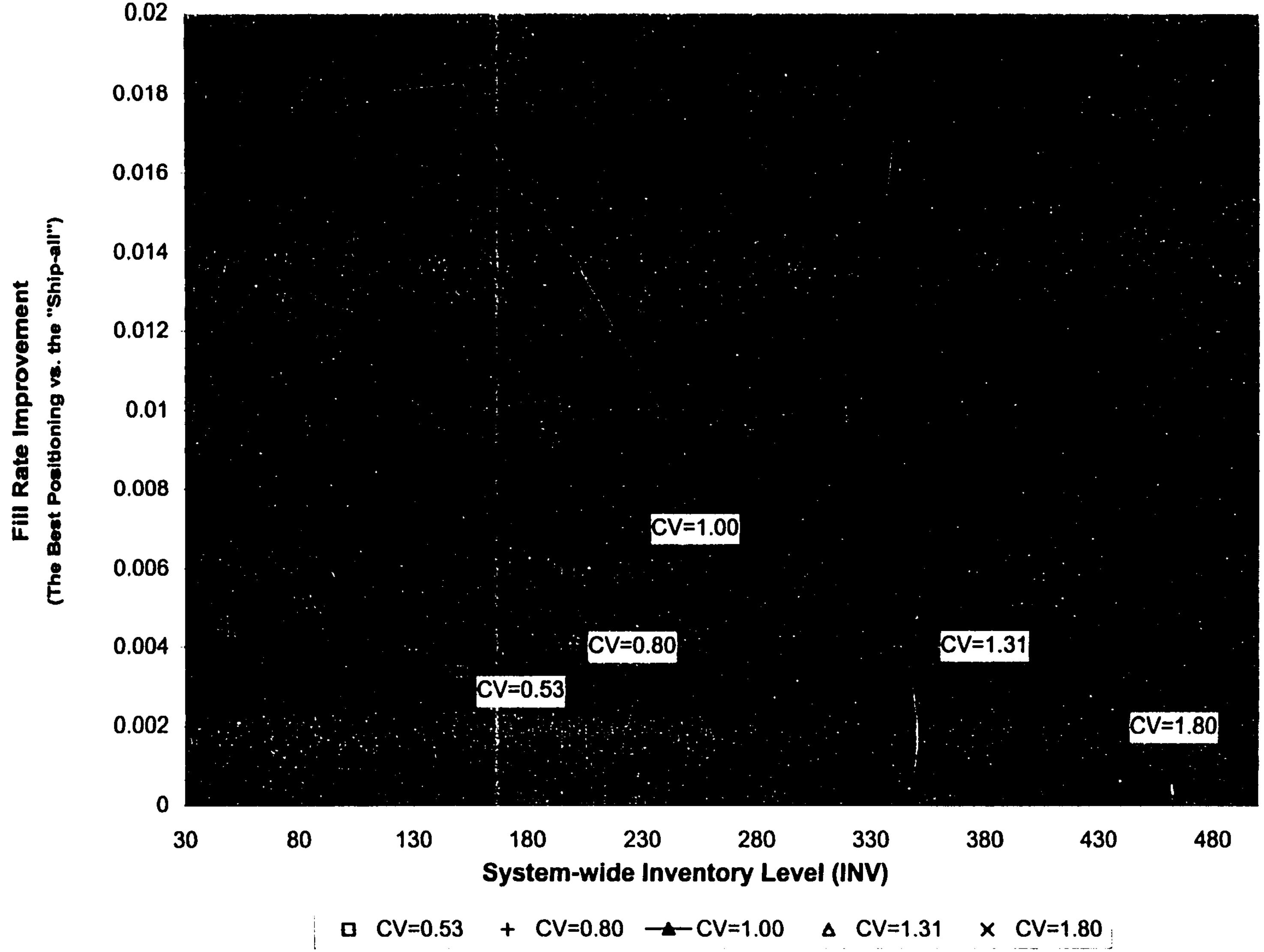
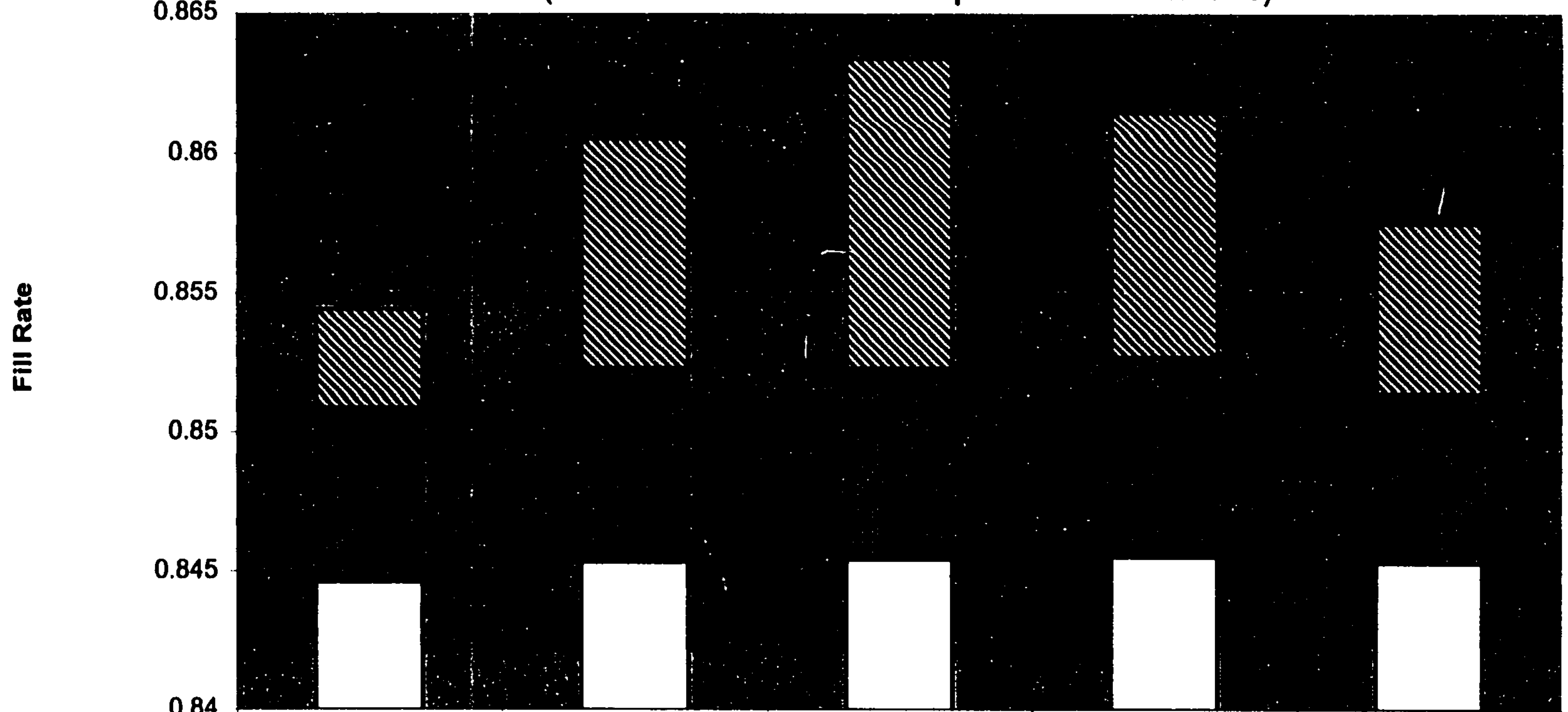


Figure 6.15: EFFECT OF CV
(Fixed the Fill Rate of the "Ship-all" for Different CVs)



	CV=0.53	CV=0.80	CV=1.00	CV=1.31	CV=1.80
INV	84	103	120	150	203
Best Positioning Ratio	0.8713	0.9379	0.9	0.9218	0.9161
Due to "best" positioning	0.0035	0.0082	0.0111	0.0088	0.0061
Due to increase in SHIPR	0.0063	0.007	0.0069	0.0072	0.0061
Ship-All	0.8446	0.8453	0.8454	0.8455	0.8453

Figure 6.16: THE EFFECT OF CV
(The fill rate improvement vs. the fill rate for the "Ship-all")

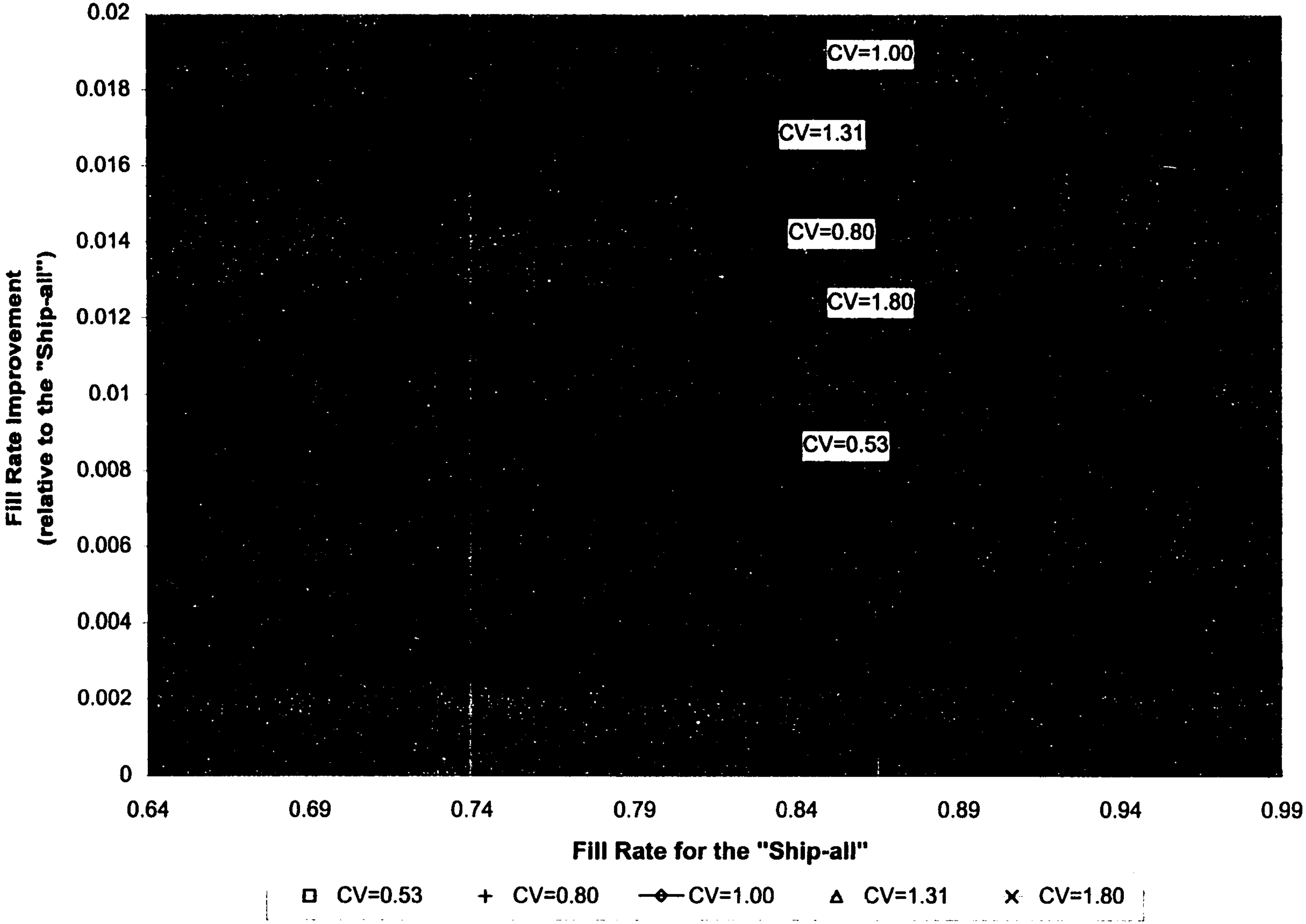
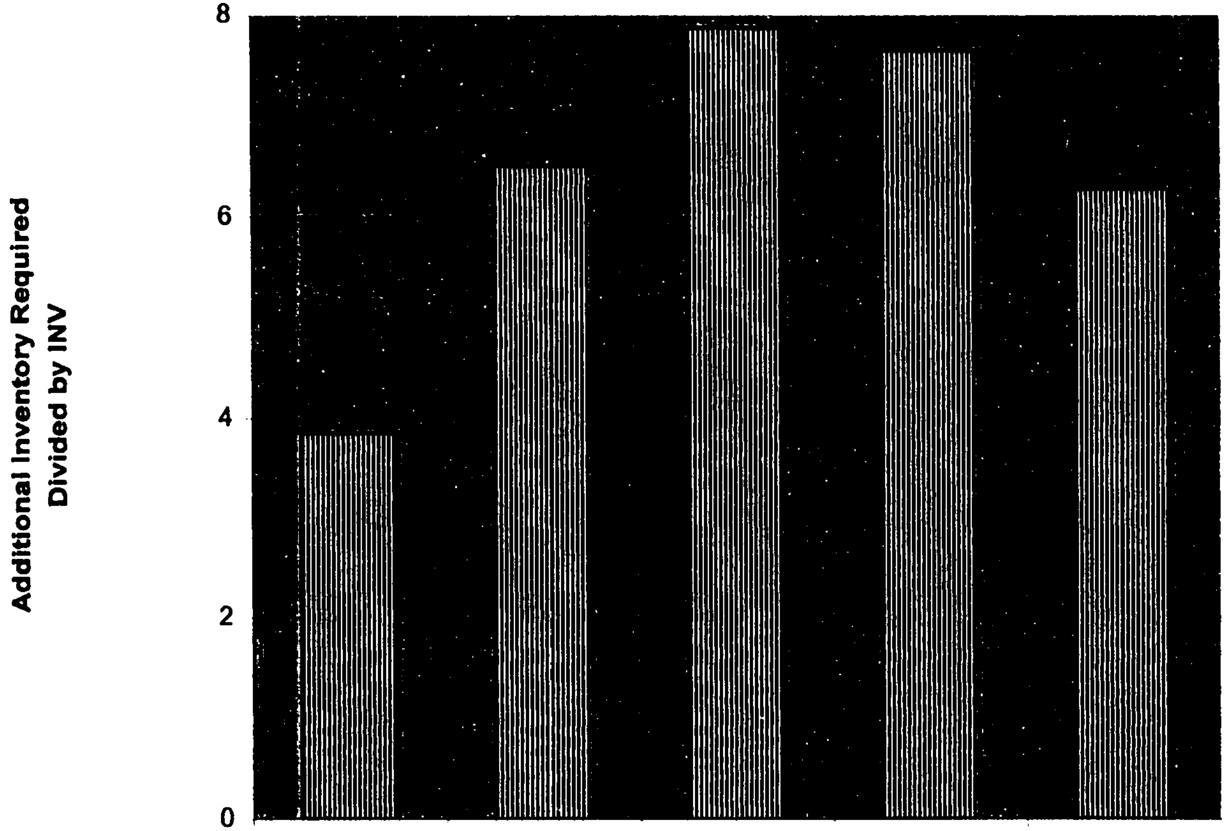
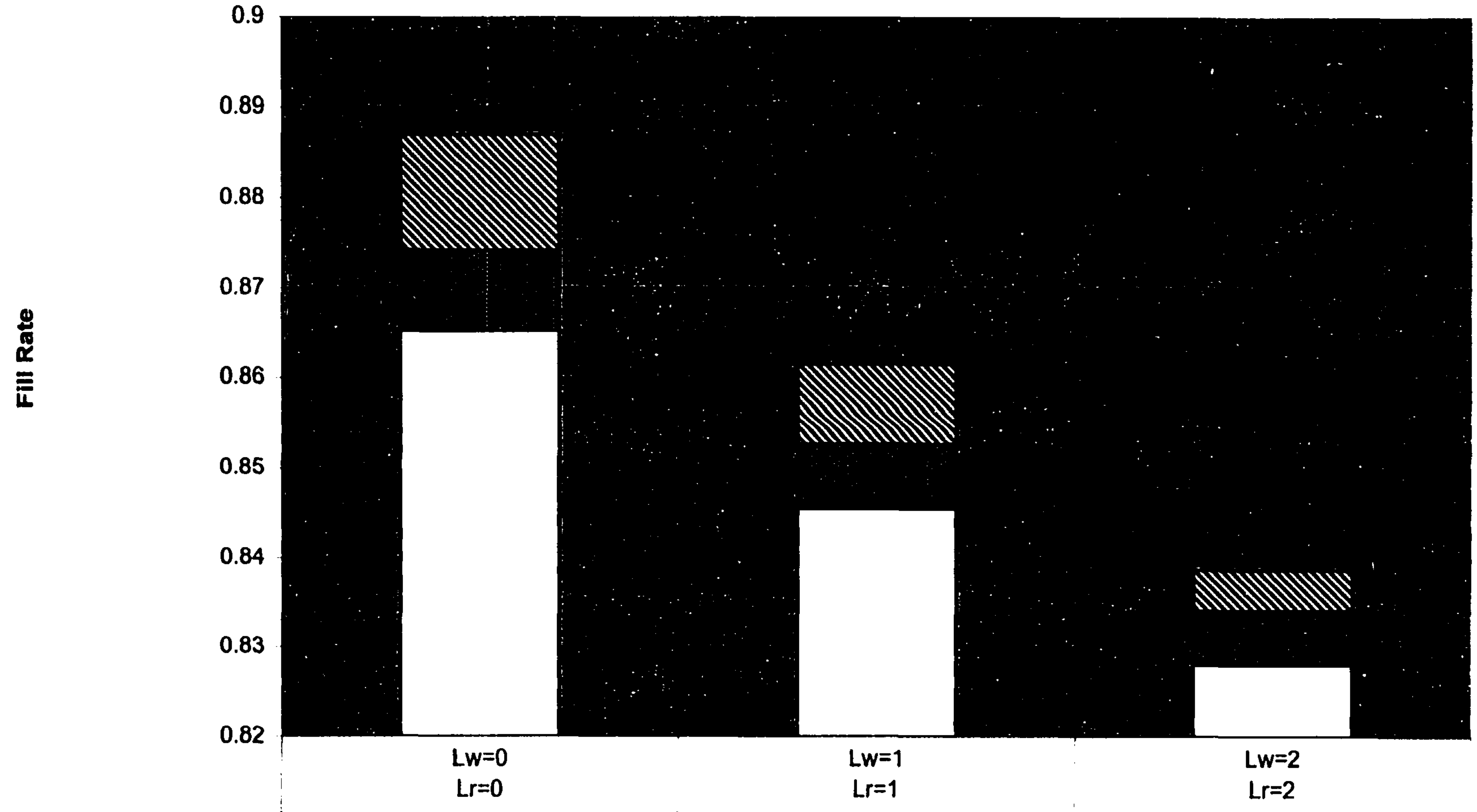


Figure 6.17: EFFECT OF CV ON ADDITIONAL INVENTORY REQUIRED
(for the "Ship-all" to match the fill rate of the best positioning)



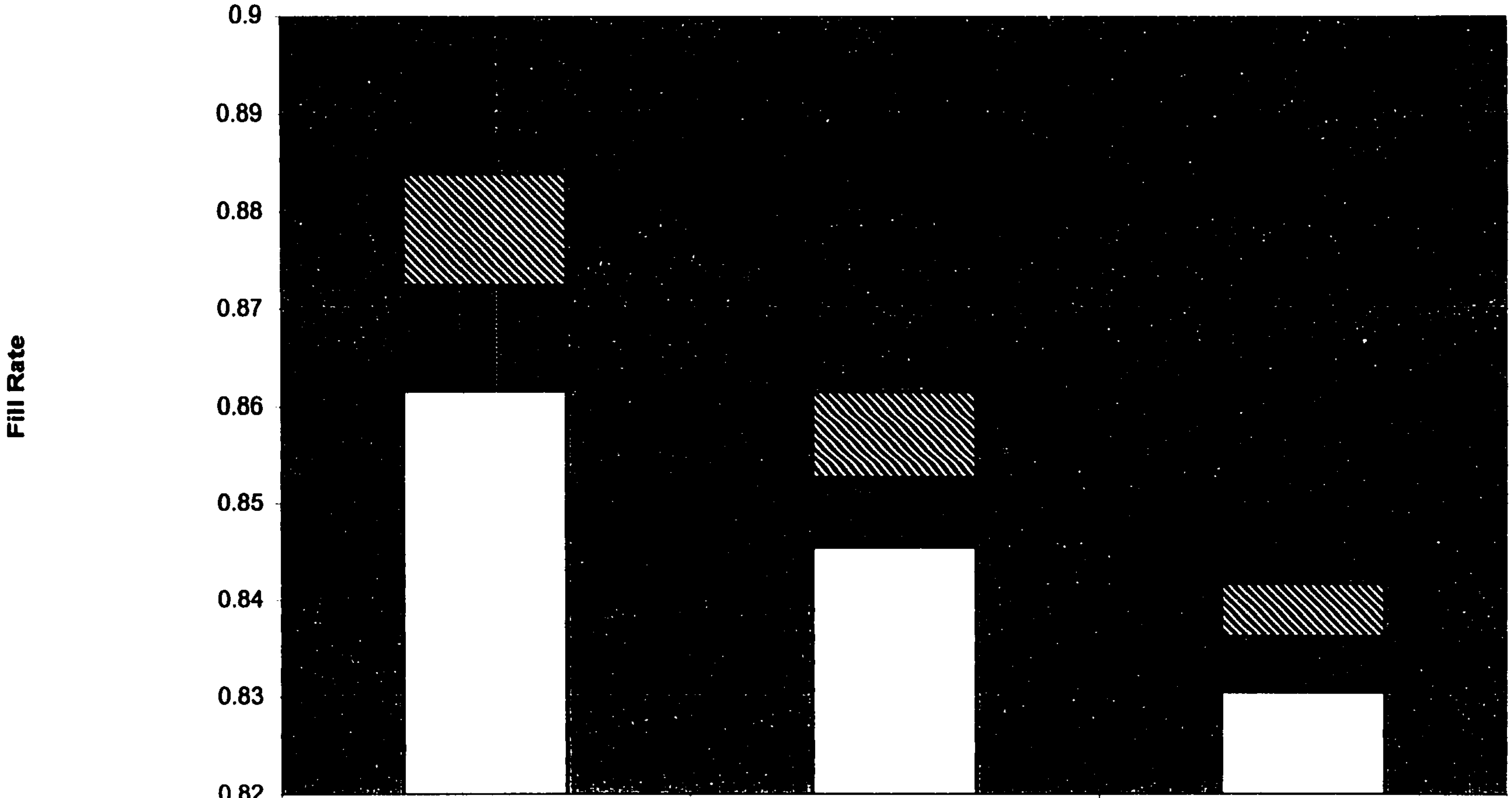
	CV=0.53	CV=0.80	CV=1.00	CV=1.31	CV=1.80
INV	84	103	120	150	203
Maximum Fill Rate	0.8544	0.8605	0.8634	0.8615	0.8577
Additional Inventory Required	3.23	6.71	9.45	11.47	12.74
Additional Inventory Required Divided by INV (%)	3.85	6.51	7.87	7.65	6.28

Figure 6.18: EFFECT OF LEAD TIME
(Lw and Lr increase proportionally)



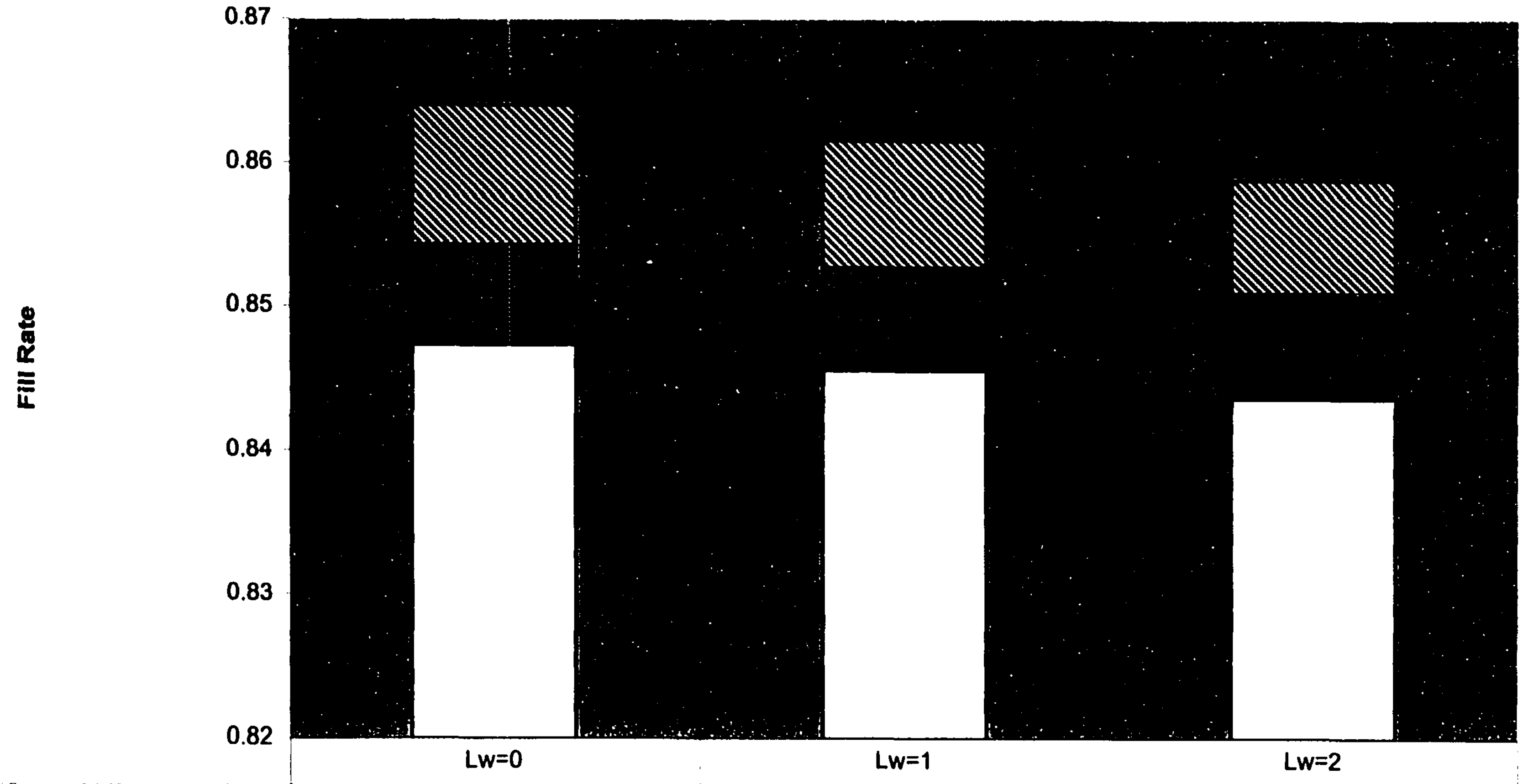
	Lw=0 Lr=0	Lw=1 Lr=1	Lw=2 Lr=2
Best Positioning Ratio	0.9067	0.9218	0.9288
Due to "best" positioning	0.0128	0.0088	0.0045
Due to increase in SHIPR	0.009	0.0072	0.0061
Ship-All	0.8652	0.8455	0.8281

Figure 6.19: EFFECT OF LEAD TIME Lr
(Lw=1)



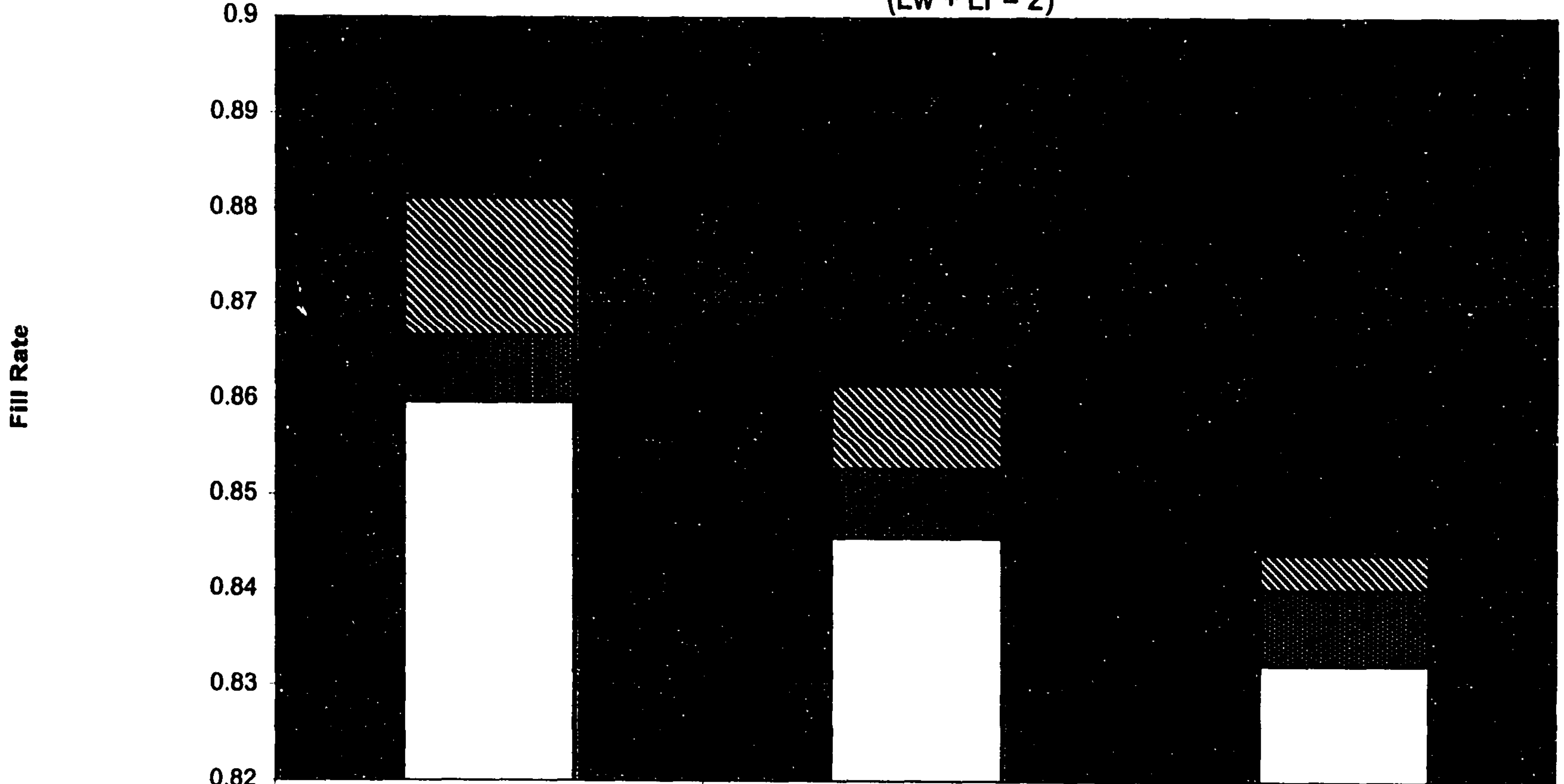
	Lr=0	Lr=1	Lr=2
Best Positioning Ratio	0.8482	0.9218	0.9333
Due to "best" positioning	0.0115	0.0088	0.0055
Due to increase in SHIPR	0.0108	0.0072	0.0057
Ship-All	0.8616	0.8455	0.8305

Figure 6.20: EFFECT OF LEAD TIME Lw
(Lr=1)



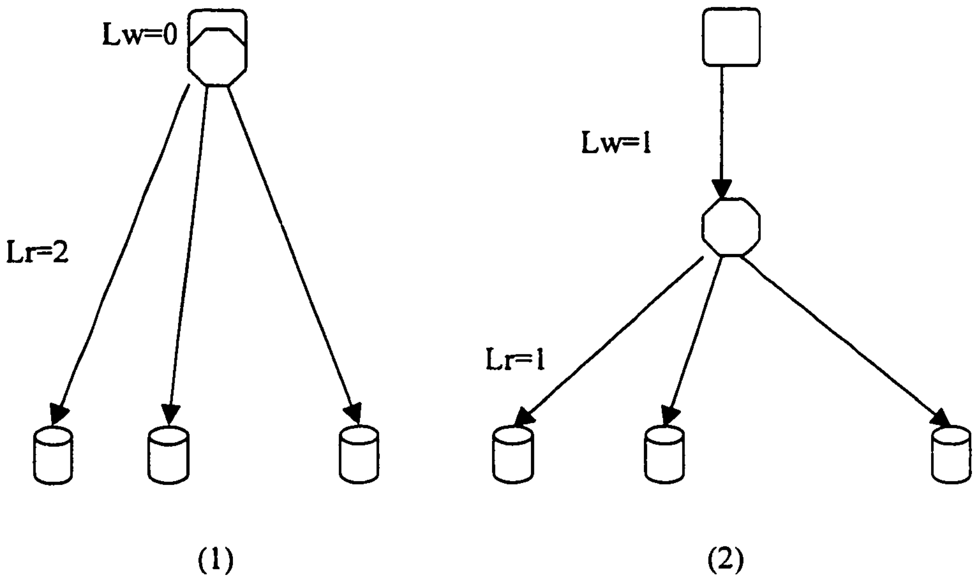
	Lw=0	Lw=1	Lw=2
Best Positioning Ratio	0.9265	0.9218	0.9173
Due to "best" positioning	0.0097	0.0088	0.0079
Due to increase in SHIPR	0.007	0.0072	0.0073
Ship-All	0.8473	0.8455	0.8436

Figure 6.21: EFFECT OF LEAD TIMES
(Lw + Lr = 2)



	Lw=2 Lr=0	Lw=1 Lr=1	Lw=0 Lr=2
Best Positioning Ratio	0.8429	0.9218	0.9382
Due to "best" positioning	0.0146	0.0088	0.0038
Due to increase in SFr	0.0069	0.0072	0.0079
Ship-All	0.8597	0.8455	0.8322

Figure 6.22 "Depot effect" vs. "risk-pooling effect over the supplier lead time"



Where 

Outside Supplier

Warehouse

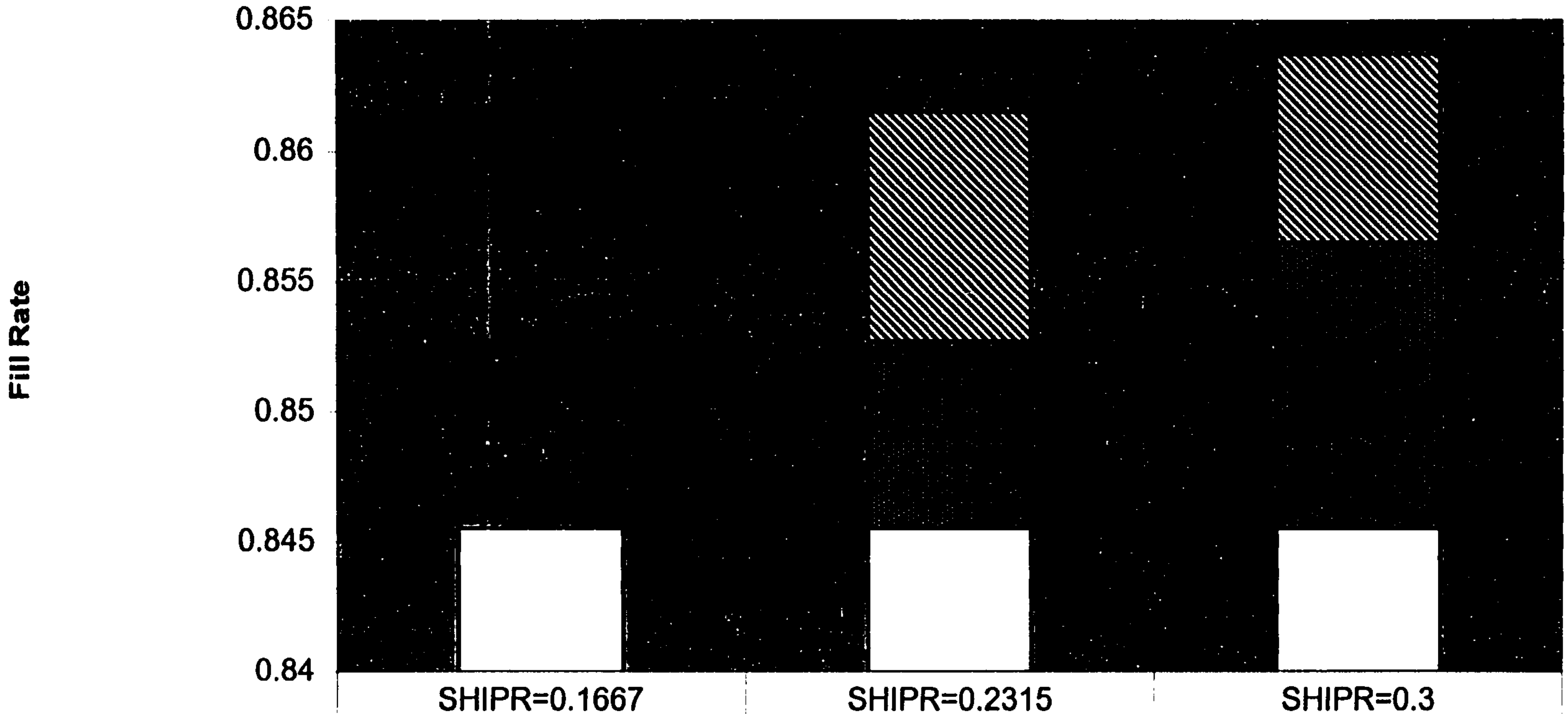
Retail Store

Two choices: (1) To hold inventory at the warehouse (which will increase the shipment frequency to the stores, SHIPR, from no greater than 0.1667 to 0.2315 shipment per period) without reallocation of the warehouse. The fill rate improvement due to the "depot effect": (F*-Fs)=0.0117=1.17%.
(2) To reallocate the warehouse (L_w changes from 0 to 1) without holding inventory at the warehouse (with SHIPR unchanged at 0.1667 shipments per period). The fill rate improvement due to the "risk-pooling over the supplier lead time (L_w =1 period): ΔFs=0.0133=1.33%.

$\because$ ΔFs > (F*-Fs), (i.e., the "depot effect" > the "risk-pooling effect over L_w")

$\therefore$ Reallocation is preferred for improving the customer fill rate.

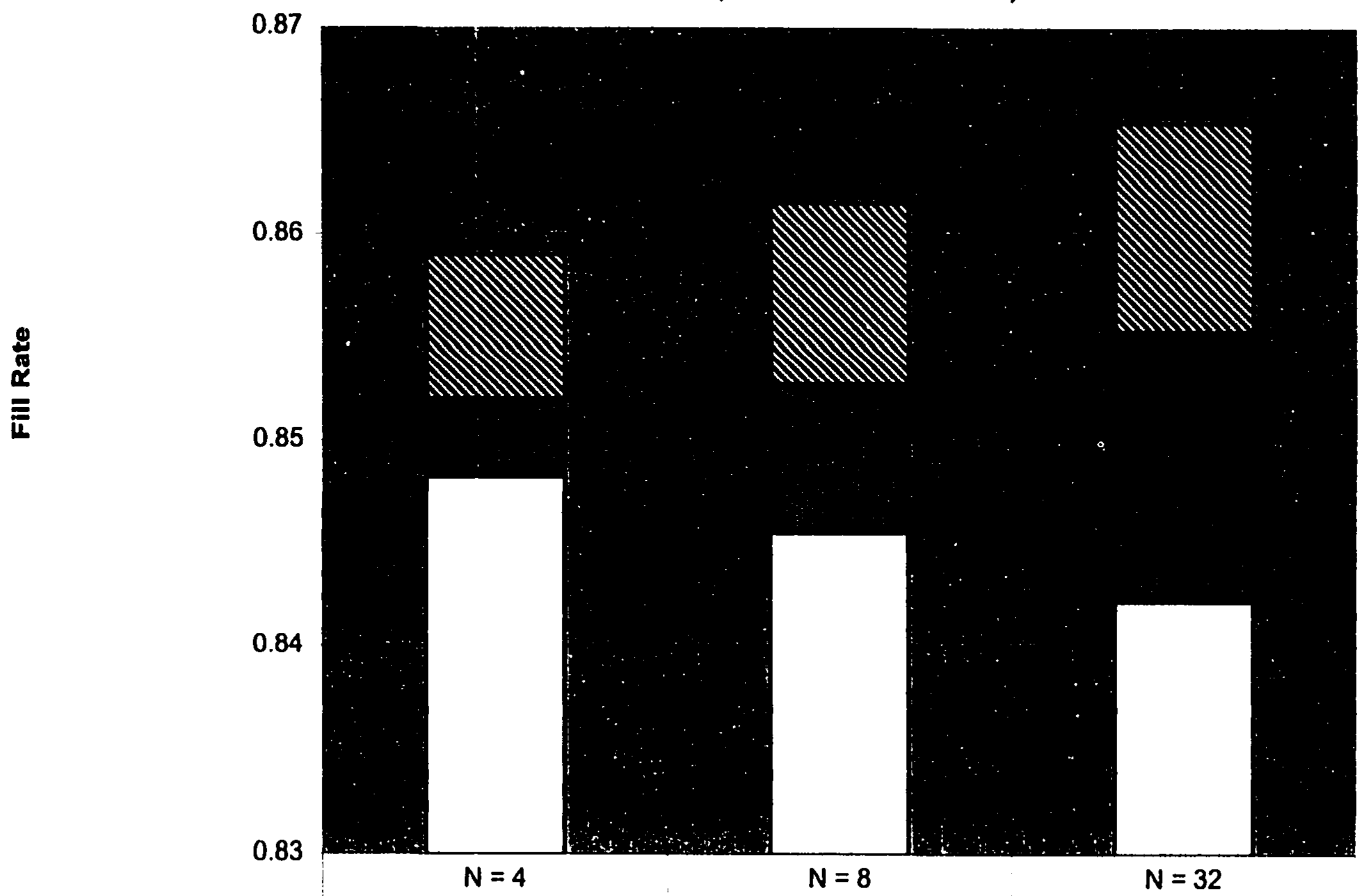
Figure 6.23 EFFECT OF SHIPMENT FREQUENCY (SHIPR)
(SHIPW=0.1667)



	SHIPR=0.1667	SHIPR=0.2315	SHIPR=0.3
Max Positioning Ratio (Fc)	1	0.9769	0.9572
Best Positioning Ratio	1	0.9218	0.8626
Due to "best" positioning	0	0.0088	0.0072
Due to increase in SHIPR	0	0.0072	0.011
"Ship-all"	0.8455	0.8455	0.8455

Shipment Frequency to Any of the Retail Stores (SHIPR)

Figure 6. 24: EFFECT OF STORE NUMBER (N)
(INV / N =18.75 units)



	N = 4	N = 8	N = 32
Best Positioning Ratio	0.9248	0.9218	0.9089
Due to "best" positioning	0.007	0.0088	0.0101
Due to increase in SHIPR	0.0038	0.0072	0.0131
Ship-All	0.8482	0.8455	0.8422

REFERENCES

Axsater, S. (1993), "Continuous review policies for multi-level inventory systems with stochastic demand," in: Graves, S.C., Rinnooy Kan, A.H.G., and Zipkin, P.H., eds., *Logistics of Production and Inventory,* Handbooks in Operations research and Management Science *4,* Elsevier (North-Holland), Amsterdam, Chapter 4, 175-197.

Axsater, S., and Rosling, K. (1993), "Notes: Installation vs. echelon stock policies for multilevel inventory control," *Management Science* 39, 1274-1280.

Axsater, S., and Rosling, K. (1994), "Multi-level production-inventory control: Material requirements planning or reorder-point policies?" *European Journal of Operational Research* 75, 405-412.

Baxter International Inc., *Annual Report*, 1992.

Brown, R. G. (1977), "Safety stock strategies," and "Central control," in *Materials Management Systems*, New York: John Wiley and Sons, Chapter 10 and Chapter 17, 176-177 and 346-349.

Bourland, K. S., S. Powell, and D. Pyke (1996), "Exploiting timely demand information to reduce inventories," *European Journal of Operational Research,* 92, 239-253.

Badinelli, R. and L.Schwarz (1988), "Backorders optimization in a one-warehouse N-identical retailer distribution system," *Naval Research Logistics Quarterly*, 35, 5, 427-440.

Bowersox, D.J. (1963), "The role of the marketing executive in physical distribution," *Proceedings of the Forty-fourth National Conference of the American Marketing Association*, 393-399.

Cachon, G. and M. Fisher (1997), "Campbell Soup's continuous replenishment program: evaluation and enhanced inventory decision rules," *Production and Operations Management*, 6 (3), 266-279.

Chakravarty, A. K. and Shtub, A. (1986), "Simulation safety stock allocation in a two-echelon distribution system," *International Journal of Production Research*, 24 (5), 1245-1253.

Carlson, R. C., and C. A.Yano. (1986), "Safety stocks in MRP -- systems with emergency setups for components," *Management Science* 32(4): 403-12.

Chain Store Age Executive, April 1992.

Chen, F., and Zheng, Y.S. (1994a), "Evaluating echelon stock *(R, nQ)* policies in serial production/inventory systems with stochastic demand," *Management Science*, 40, 1262-1275.

Chen, F., and Zheng, Y.S. (1994b), "Lower bounds for multi-echelon stochastic inventory systems," *Management Science*, 40, 1426-1443.

Clark, A.J., and Scarf. H. (1960), "Optimal policies for a multi-echelon inventory problem," *Management Science* 6, 475-490.

Clark, A.J., and H. Scarf (1962). "Approximate solutions to a simple multi-echelon inventory problem," in: K.J. Arrow, S. Karlin and H. Scarf (eds.), *Studies in Applied Probability and Management Science*, Stanford University Press, Stanford, CA, pp. 88-110.

Clark, A.J. (1972), "An informal survey of multi-echelon inventory theory," *Naval. Research Logistics Quarterly*, 19, 621-650.

Clark, T.H. and J. Hammond (1995), "Re-engineering channel reordering processes to improve total supply-chain performance," *Working Paper*, Hong Kong University of Science and Technology.

de Bodt, M.A., and L.N. Van Wassenhove (1983) "Lot sizes and safety stocks in MRP: a case study," *Production Inventory Management*, 24, 1-16

de Kok, A.G. (1990), "Hierarchical production planning for consumer goods," *European Journal of Operational Research* 45, 55-69.

Davis, T. (1993), "Effective supply chain management," *Sloan Management Review*, Summer 1993, 35-45.

Davis, H.W. and W.H. Drumm (1994), "Physical distribution cost and service," *Council of Logistics Management Annual Conference Proceedings*, Cininnanti, Ohio, October 16-19, 1994, 117-127.

Deuermeyer, B.L., and Schwarz, L.B. (1981), "A model for the analysis of system service level in warehouse-retailer distribution systems: the identical retailer case," in: Schwarz, L.B., ed., *Multi-level Production/Inventory Control Systems: Theory and Practice*, TIMS Studies in the Management Sciences *16*, North-Holland, Amsterdam, 163-193.

Ehrhardt, R., Schultz, C. R., and Wagner, H. M. (1981), "(s, S) policies for a wholesale inventory system," in: Schwarz, L.B., ed., *Multi-level Production/Inventory Control Systems: Theory and Practice,* TIMS Studies in the Management Sciences *16,* North-Holland, Amsterdam, 145-161.

Eppen, G. (1979), "Effect of centralization on expected costs in a multi-location newsboy problem," *Management Science* 25 (5), 498-501.

Eppen, G., and Schrage, L. (1981), "Centralized ordering policies in a multi-warehouse system with lead times and random demand," in: Schwarz, L.B., ed., *Multi-level Production/Inventory Control Systems: Theory and Practice,* TIMS Studies in the Management Sciences *16,* North-Holland, Amsterdam, 51-67.

Erkip, N., W. Hausman and S. Nahmias (1984). "Optimal and near-optimal policies in multiechelon inventory systems with correlated demands," Presented at the *TIMS XXVI International Meeting*, Copenhagen, Denmark, June 17-21.

Exective Report: ECR (1994), "The ERC road map," *Progressive Grocer*, January 1994, 5-6.

Federgruen, A. (1993a), "Recent advances in production and distribution management," in: R. K., Sarin eds., *Perspectives in Operations Management: Essays in honor of Elwood S. Buffa*, Boston MA: Kluwer Academic Publishers, 297-300.

Federgruen, A. (1993b), "Centralized planning models for multi-echelon inventory systems under uncertainty," in: Graves, S.C., Rinnooy Kan, A. H. G., and Zipkin, P.H., eds., *Logistics of production and inventory*, Handbooks in Operations Research and Management Science 4, Elsevier (North-Holland), Amsterdam, Chapter 3, 133-173.

Federgruen, A. and P. Zipkin (1984a), "Approximations of dynamic, multi-location production and inventory problems," *Management Science,* 30, 1, 69-84.

Federgruen, A., and P. Zipkin (1984b), "Computational issues in an infinite-horizon, multi-echelon inventory model," *Operations Research* 32, 818-836.

Federgruen, A., and P. Zipkin (1984c), "Allocation policies and cost approximation for multilocation inventory systems," *Naval. Research Logistics Quarterly,* 31, 97-131.

Fisher, M. L. and J.H. Hammond, W.R. Obermeyer, and A. Raman (1994), "Making supply meet demand in an uncertain world," *Harvard Business Review,* May-June 1994, 83-93.

Gallego, G. and P. Zipkin (1994), "Qualitative analysis of multi-stage production-transportation systems: stock positioning and performance estimation," *Working Paper*, Columbia University.

Gavirneni, S., R. Kapuscinski and S. Tayur (1996), "Value of information in capacitated supply chains," *Working Paper*, Carnegie Mellon University.

Gibrat, R. (1930), see Johnson, N.L., S. Kotz, and N. Balakrishnan (1994), 207-252.

Gibrat, R. (1931), see Johnson, N.L., S. Kotz, and N. Balakrishnan (1994), 207-252.

Gopal, C. (1992), "Manufacturing logistics systems for a competitive global strategy," in: Christopher, M. (ed.), *Logistics: strategic issues*, New York: Chapman & Hall, Chapter 11, 130-139.

Graves, S. (1996), "A multi-echelon inventory model with fixed replenishment intervals," *Management Science*, 42, 1-18.

Graves, S. C. (1988), "Safety stocks in manufacturing systems," *Journal of Manufacturing and Operations Management*, 1, 67-101.

Graves, S. C. and S. P. Willems (1996), "Strategic safety stock placement in supply chains," *1996 Manufacturing and Service Operations Management Conference Proceedings*, 299-304.

Guerrero, H. H., K. R. Baker, and M. H. Southard (1986), "The dynamics of hedging the master schedule." *International Journal of Production Research* 24(6): 1475-1483.

Hadijian, A. (1994), Corporate Performance, *Fortune*, March 7, 1994: 75-82.

Hanssmann, F. (1959), "Optimal inventory location and control in production and distribution networks," *Operations Research*, 7(9), 483-498

Hausman, W.H. and N.K. Erkip (1994), "Multi-echelon vs. single-echelon inventory control policies for low demand items," *Management Science*, 40(5), 579-602.

Henkoff, R. (1994), "Delivering the good, " *Fortune*, Nov. 28, 1994, 37-47.

Inderfurth, K. (1991), "Safety stock optimization in multi-stage inventory systems," *International Journal of Production Economics* 24, 103-113.

Inderfurth, K. (1994a), "Safety stocks in multistage divergent inventory systems: A survey", *International Journal of Production Economics* 35, 321-329.

Inderfurth, K. (1994b), "Erratum to: 'Safety stocks in multistage divergent inventory systems: A survey," *International Journal of Production Economics* 37, 287-288.

Inderfurth, K. (1995), "Multistage safety stock planning with item demands correlated across products and through time," *Production and Operations Management* 4, 127-144.

Jackson, P.L. (1988), "Stock allocation in a two-echelon distribution system or 'what to do until your ship comes in'," *Management Science* 34, 880-895.

Jacobs, F.R. and D.C. Whybark (1992), "A comparison of reorder point and material requirements planning inventory control logic," *Decision Science*, 23, 332-342.

Johnson, M. Eric, T. Davis, and M. Waller (1998), "Vendor Managed Inventory in the Retail Supply Chain," *Working Paper*, Owen Graduate School of Management, Vanderbilt University.

Johnson, N.L., S. Kotz, and N. Balakrishnan (1994), "Lognomal Distribution," *Continuous Univariate Distribution*, Chapter 14, 2nd ed., New York: Wiley & Sons, 207-252.

Jonsson, H., and E.A. Silver (1987a), "Stock allocation among a central warehouse and identical regional warehouses in a particular push inventory control system," *International Journal of Production Research* 25, 191-205.

Jonsson, H. and E. A. Silver (1987b), "Analysis of a Two-Echelon Inventory Control System with Complete Redistribution," *Management Science*, 33, 2, 215-227.

Krajewski, L.J., B.E. King, L.P. Ritzman, and D.S. Wong (1987), Kanban, MRP and shaping the manufacturing environment," *Management Science*, 33, 39-57.

Kurt Salmon Associates Inc. (1993), *Efficient Consumer Response: Enhancing Consumer Value in the Grocery Industry*, Food Marketing Institute, Washington D. C.

Law, M. A. and W. D. Kelton (1991), *Simulation Modeling and Analysis*, New York: McGraw-Hill, 2nd ed.

Lee, H and C. Billington (1992), "Managing supply chain inventory: pitfalls and opportunities," *Sloan Management Review*, Spring 1992, 65-74.

Lee, H (1997), *Lecture Notes for Research Seminars on Recent Advances in Supply Chain Management*, Department of Information and Systems Management, Hong Kong University of Science and Technology.

Lee, H., K. So, and C. Tang (1996), "Supply Chain Re-engineering Through Information Sharing and Replenishment Coordination," *Working Paper*, Stanford University.

Lee, H., V. Padmanabhan and S. Whang (1997), "Information Distortion in a Supply Chain: The Bullwhip Effect," *Management Science* 43(4), 546-558.

Magee. J.F., W.C. Copacino, and D. B. Rosenfield (1985), *Morden Logistics Management*, New York: Wiley & Sons, 59-69.

Maxwell, W.L. and Muckstadt, J.A., "Establishing consistent and realistic reorder intervals in production-distribution systems," *Operations Research*, 33 (6), 1316-1341.

McClelland, M.K. and H.W. Wagner (1988), "Location of inventories in an MRP environment," *Decision Science*, 19(3), 535-553.

McGavin, E. (1990), "Determining inventory allocation policies in a one-warehouse N-identical retailer distribution system," *Ph.D. dissertation*, Krannert Graduate School of Management, Purdue University, West Lafayette, IN.

McGavin, E., L. Schwarz and J. Ward (1992), "Balancing Retailer Inventories," *Working Paper,* Krannert Graduate School of Management, Purdue University, West Lafayette, IN.

McGavin, E., L. Schwarz and J. Ward (1993), "Two-interval inventory allocation policies in a One-warehouse N-identical retailer distribution system," *Management Science,* 39, 1092-1107.

Miller, J.G. (1979), "Hedging in master schedule," in: Desegregated problems in manufacturing and service organizations, L.P. Ritzman, et al. (eds.) Boston MA: Martinus Nifhoff, 237-256.

Nahmias, S. and S. A. Smith (1993), "Mathematical models of retailer inventory systems: a review," in: R. K., Sarin (ed.), *Perspectives in Operations Management: Essays in honor of Elwood S. Buffa*, Boston MA: Kluwer Academic Publishers, 249-278.

Nahmias, S. and S. A. Smith (1994), "Optimal inventory levels in a two-echelon retail system with partial lost sales," *Management Science* 40(5), 582-596.

Nahmias, S. "Shortcomings of MRP," *Production and Operations Analysis*, 3rd edition, Chicago: Irwin 1997, Chapter 6, 361.

Orlicky, J. (1975), *Material Requirement Planning*, New York: McGraw-Hill.

Porteus, E.L. (1990), "Stochastic inventory theory," in: D.P. Heyman, M.J. Sobel, eds., *Handbooks in Operations Research and Management Science 2*, Elsevier (North-Holland), Amsterdam, Chapter 12, 605-652.

Pyke, D.F. and M.A. Cohen (1990), "Push and Pull in Manufacturing and Distribution Systems," *Journal of Operations Management*, 1, 1-10.

Rees, L.P., P.Y. Huang, and B.W.Taylor III (1989), "A comparative analysis of an MRP lot-for-lot system and a Kanban system for a multistage production operation," *International Journal of Production Research*, 37, 565-579.

Riggins, F. and T. Mukhopadhyay (1994), "Interdependent benefits from inter-organizational systems: opportunities for business partner re-engineering," Journal of Management Information Systems, 11, Fall, 37-57.

Rogers D.F. and S. Tsubakitani (1991), "Newsboy-style results for multi-echelon inventory problems: backorders optimization with intermediate delays." *Journal of Operational Research Society*, 42 (1), 57-68.

Rosebaum, B.A. (1981), "Inventory placement in a two-echelon inventory system: an Application," in: Schwarz, L.B., ed., *Multi-level Production/Inventory Control Systems: Theory and Practice,*, TIMS Studies in Management Sciences *16*, North-Holland, Amsterdam, 195-207.

Rosenbaum, B.A. (1981), "Service level relationships in a multi-echelon inventory system," *Management Science* 27, 926-945.

Rosenfield, D., and M. Pendrock (1980), "The effects of warehouse configuration design on inventory levels and holding costs," *Sloan Management Rev.* 24 (4), 21-33.

Rosenbaum, B.A. (1981), "Service level relationships in a multi-echelon inventory system," *Management Science* 27, 926-945.

Roundy, R.O., ". and J.L. Ringuest (1990), "98%-effective integer-ratio lot-sizing for one-warehouse multi-retailer systems, " *Management Science*, 31 (11), 1416-1430.

Schneider, H., D.B. Rinks, and P. Kelle (1995), "Power approximations for a two-echelon inventory system using service levels," *Production and Operations Management*, 4 (4), 381-400.

Schwarz, L. (1989), "A model for assessing the value of warehouse risk-pooling: risk pooling over outside-supplier leadtimes," *Management Science, 3 5(7), 1989, 828-842.*

Schwarz, L. B. (1981a), "Introduction," in Schwarz, L. B. (ed.), *Multi-Level Production/Inventory Control Systems: Theory and Practice*, TIMS Studies in the Management Sciences, North Holland, Amsterdam, The Netherlands, 1-9.

Schwarz, L. B. (1981b), "Physical distribution: the analysis of inventory and location," *AIEE Transaction*, 13 (2), 138-149.

Schwarz, L., B. Deuermeyer and R. Badinelli (1984), "Fill-rate optimization in a one-warehouse N-identical retailer distribution system," *Management Science*, 31(4), 488-498.

Seller, P. (1992), "The Dumbest Marketing Ploy," *Fortune*, Oct. 5, 1992, 88-94.

Silver, E.A., and Peterson, R.(1985), *Decision Systems for Inventory Management and Production Planning*, Wiley, New York, *1985*.

Simpson, K.F. (1958), "In-Process Inventories," *Operations Research*, 6, 863 -873.

Stalk G., P. Evans, and L. E. Shulman (1992), "Competing on capacities," *Harvard Business Review*, March-April 1992, 57-69.

Stenger, A.J. and J.L. Cavinato (1979), "Adapting MRP to the outbound side—Distribution Requirements Planning," *Production and Inventory Management*, Fourth Quarter, 1-13.

Troyer, C. (1996), "EFR: Efficient Food-Service Response," paper presented at *Conference on Logistics, GMA*, Palm Springs, California, May 21-23, 1996.

van der Hoop, J.H. (1992), "The single European Market: optimizing logistics operations in the post-1992 Europe," in: Christopher, M. (ed.), *Logistics: strategic issues*, New York: Chapman & Hall, Chapter 17, 260-267.

van Houtum, G.J., K. Inderfuth, K. and W.H.M. Zijim (1996), "Materials coordination in stochastic multi-echelon systems," *European Journal of Operational Research*, 95, 1-23.

Vollmann, T., W. Berry and D. Whybark (1997). "Distribution requirements planning," *Manufacturing Planning and Control Systems, 4^{th} ed.*, Now York: McGraw-Hill, Chapter 18, 744-776.

Wagner, H.M. (1974), "The design of production and inventory systems for multi-facility and multi-warehouse companies," *Operations Research*, 22, 278-291.

Wagner, H.M. (1980), "Research portfolio for inventory management and production planning systems," *Operations Research*, 28, 445-475.

Whang, S (1994), *Abstract*, TIMS/ORSA Joint National Meeting, Boston, MA, April 1994.

Whybark, C. D., and S. Yang (1996), "Positioning inventory in distribution systems," *International Journal of Production Economics* 45, 271-278.

Yang, S. and Whybark, C. D. (1994), "Managing inventory in multi-echelon distribution systems with stochastic demand," *Proceedings of the Twenty-fifth Annual Meeting of Decision Science Institute*, Honolulu, Hawaii, 1080-1810.

Zipkin, P. (1984), "On the imbalance of inventories in multi-echelon systems". *Mathematics of Operations Research* 9, 402-423.

Zipkin, P. (1998), "Supply-Chain Management: Convergences of Theory and Practice," *Abstract of Seminar Presentations*, Department of Operations Research, the University of North Carolina at Chapel Hill.