Lead Time Management in Supply Chains

by

Saibal Ray

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Management Sciences

Waterloo, Ontario, Canada, 2001

0-612-65257-2

The University of Waterloo requires the signature of all persons using or photocopying this thesis. Please sign below, and give address and date.

Abstract

In recent years speed and cost have emerged as important competitive priorities in supply chains. Firms are now investing substantially in lead time reduction; however, the focus of such investments has been quite different for make-to-stock (MTS) and make-to-order (MTO) firms. The demand for MTS items tends to be deterministic but price-sensitive, while demand for MTO items is more variable and sensitive to both price and delivery lead time. These differences in market characteristics require that MTS firms focus on supplying predictable demand at the lowest possible cost while MTO firms focus on reducing the delivery lead time. Our research deals with the costs and benefits of lead time management in supply chains, taking into account the differences in competitive environments. In particular, we develop separate lead time management models for profit-maximising MTS and MTO firms.

For the MTO firm, we assume that customer demand is stochastic and the mean demand rate is decreasing in both price and a uniform guaranteed delivery lead time offered by the firm. To further model the premium for lower delivery lead times, we assume that price is decreasing in the length of the guaranteed delivery lead time. We also capture economies of scale by assuming the unit operating cost to be a decreasing convex function of the demand rate. The MTO firm may invest in increasing capacity in order to reduce delivery lead time, but must be able to satisfy customers according to a pre-specified service level. Our analytical model for delivery lead time management of such MTO firms trades off the costs of investment against the resultant benefits. Our model allows a MTO firm to determine the optimum level of the guaranteed delivery time, processing rate and investment that maximise its profit. We show that ignoring - i) the dependence of market price on the lead time offered and economies of scale, when they exist, and ii) the inherent preference of customers for price or lead time - can lead to potentially large profit losses.

Normally MTS firms invest in developing more efficient processes that reduce operating costs. While the process-improving investments can be of various types, we focus on investments in reducing supplier lead time and develop models for supply lead time

management for MTS firms. We show that such investments in lead time reduction can, after accounting for all the associated costs and benefits, result in substantial reduction of inventory costs. We examine different types of investment and amortisation schemes in supplier lead time reduction and the different cost models they generate. We compute the cost-minimising inventory and supply lead time levels for each type of model. We also perform comparative statics with respect to model parameters, and find several "apparently" counter-intuitive results.

We then assume that a MTS firm sets its price as a percentage mark-up over its total operating costs per unit. In that case, any investment in reducing operating costs can lower price and help the firm to gain a greater market share. For the case of investment in set-up time (cost) reduction, we are able to formulate an integrated production-marketing model for a profit-maximising MTS firm where price and demand, and hence profit, are functions of the firm's operating variables. We show that when demand depends on the operating variables in a profit maximisation model, some of the best known properties from classical inventory management no longer hold. We are also able to show that if a MTS firm ignores the explicit dependence by either assuming demand to be constant or price to be an independent decision variable, sub-optimality occurs and the firm can loose substantial profits. For the case of investment in supply lead time reduction, we are also able to formulate the profit-maximising problem in terms of the operating variables of the firm and to indicate how it can be solved.

Acknowledgements

This thesis is my contribution to a three year joint research by my supervisors - Professor Elizabeth Jewkes and Professor Yigal Gerchak - and myself. I wish to convey my heartfelt thanks and gratitude to both of my supervisors for their continuous moral, intellectual and financial support during the course of my study at the Department of Management Sciences at the University of Waterloo. They have been kind enough to pay for everything that I ever asked for, including a number of trips to INFORMS and CORS conferences. My meetings with them have greatly enhanced my understanding of the subject.

Sincere appreciation and thanks also goes to Professor Raymond Vickson, Professor Paul Iyogun and Professor Anthony Atkinson for their special interest and valuable comments on this research, in spite of their busy schedule. I would also want to acknowledge here my indebtedness to Professor Rick So for kindly consenting to be the external examiner for this research.

Now comes the difficult part. There are so many other friends (past and present), teachers and countless other people who during any part of my life have helped me in anyway for shaping it. All the guidance, support, help and encouragement from them are deeply appreciated.

My whole family has been solidly behind me during the good and bad periods of my life and I want to express my sincere appreciation for that. My very special appreciation goes to Taweewan whose constant love and encouragement made this thesis possible and who suffered with me during it. Lastly I would pray to my idols to guide me through the rest of my life in the same way they have done so far.

Anything that is good in this report is because of all the above persons and all faults and omissions are exclusively mine. For all those, my apologies.

DEDICATION

To Taweewan, my parents, Riya, Raja and Riju

Table of Contents

Chapter 3: Delivery Lead Time Management for Make-to-Order Firms

Chapter 4: Supply Lead Time Management for Make-to-Stock Firms

Chapter 5: Lead Time Management for Make-to-Stock Firms

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

1.1 Motivation

Speed is one of the most important competitive elements in many modern business environments. With this in mind, many supply chains are investing in reducing lead times throughout their operations. The focus of lead time reduction in supply chains varies depending on their product characteristics. Make-to-stock supply chains typically strive for cost reduction. On the other hand, the aim for make-to-order supply chains is normally to reduce their delivery times. One of the most popular ways of cost reduction in make-to-stock supply chains is through investment in supply lead time reduction.

In this thesis we attempt to gain insights into the impact of investing in lead time reduction for supply chains that operate in different competitive environments taking into account both the costs and benefits associated with such reductions. We develop three analytical models to investigate the issue. One model deals with delivery lead time management for make-to-order supply chains; the other two models deal with set-up time and supply lead time management for make-to-stock supply chains.

The initial motivation for this research came from discussions with an Electronics Manufacturing Service (EMS) company in Toronto. In general, EMS companies provide everything from a simple cable harness to a complete high-end server or workstation for the supply chains of large original equipment manufacturers (OEMs). Recently, OEMs have recognised EMSs capabilities and are outsourcing an increasing amount of production to them. The EMS industry is forecast to grow from $78 billion in 1999 to $260 billion in 2004, a 28% compound annual growth rate (CAGR), compared to CAGR of about only 8% for the electronics industry as a whole during the same time (Carbone 2000). This particular EMS specialises in supplying electronics components for a number of international OEMs primarily in the computer and communications industry. The EMS was involved in a number of supply lead time reduction initiatives for the OEMs and somewhat surprisingly, the OEMs were ready to pay for many of those initiatives. Our research began by trying to gain a better understanding of why OEMs were willing to invest in lead time reduction.

In this case, one of the OEMs that specialised in customised servers indicated that the reason for their interest in supply lead time reduction was to reduce the delivery time of customised server products to their customers. The servers were mainly being built for dot-com companies that were ready to pay a significant price premium for early delivery. The situation for another OEM that specialised in standard, off-the-shelf computers was different. The reason for their focus on supply lead time reduction was to reduce inventory costs and ultimately to cut prices. The market for such standard computers is extremely competitive and price is a key selling feature.

This discussion led us to realise that while investment in lead time reduction within the supply chain can lead to cost reduction and/or faster delivery, such investments are motivated by different factors depending on the marketplace. To set the broader stage for this research, in the next section we provide an overview of the relevant supply chain environment issues.

1.2 Background

1.2.1 Costs and Benefits of Lead Time Reduction in Supply Chains

The 1980s brought a widespread recognition of the importance of effective analysis and improvement of manufacturing practices in maintaining a firm's competitiveness (Hayes, Wheelwright and Clark 1988). It was also during the 1980s that many firms started outsourcing a large part of their business (refer to October 2000 issue of Fortune for more on the importance of outsourcing in modern business). The move towards large-scale outsourcing led to the break-up of vertically integrated companies which began to give way to a network of loosely integrated companies that use each others' capabilities to form beneficial, win-win partnerships called supply chains. Companies started recognising that they had to rely on effective management of their supply chains for competitive advantage (Tayur et al. 1999).

In the early 80s, most manufacturers believed that low cost and high quality were the most fundamental sources of competitive advantage. Interestingly enough, quality is now no more a discriminating market factor as in the past; rather it has become a "order qualifying characteristic" (Monczka and Trent 1995). Success in many businesses now depends largely on time-based competition. This concept was made popular by the classic works of Stalk and Hout (1990) and Blackburn (1991) who showed how firms could gain advantage by being faster than competitors in different aspects of their operations. In recent times a huge volume of academic as well as popular literature has been published on this issue (e.g., So and Song 1998; Suri 1998).

Two other related developments have increased the importance of time-based strategies. The first is the growth of service organisations. The service sector now accounts for 55% of the United States Gross Domestic Product (Center for Retailing Education and Research 2001) and 68% of the Canadian Gross Domestic Product (Industry Canada 2001). In service industries, customers regard total service time as a key concern - the shorter the sojourn time in the facility, the better (Stevenson 1999). The second development is the growth in the use of the internet as a robust channel for commerce. Research projections for firms that do business both on and offline indicate that by the end of 2000, 25% of their revenue will come from the Web. This number is forecast to rise to 34% in 2001 and 50% by 2002 (Levy 2000). Internet shoppers tend to be driven by one of two rewards - either the best price on a readily-available item or finding something special faster than possible by any other means (E-commerce News 2000; Smith, Bailey and Brynjolfsson 2000). A recent survey of overall performance at 110 organisations in five major manufacturing sectors by Performance Management Group, a subsidiary of high-tech management consultants PRTM, indicates that the best in the class performers focus their attention on achieving breakthroughs in costs and speed (Geary and Zonnenberg 2000).

As speed became a driver of business success, lead time reduction emerged as a dominant issue in manufacturing strategy (van Beek and van Putten 1987; Suri 1998; Hopp and Spearman 2000). Lead times in a supply chain can have a number of elements including product development time, supply lead time, set-up time, manufacturing/service time, waiting time and delivery time. For our research we assume that the product or service has already been developed and the firm is in regular operation. We concentrate only on those lead time elements that are related to producing and delivering the product or service to the customer.

There are many advantages of reducing lead times. Some of them include lowering WIP (Work-in-Process), better scheduling, better quality, reduction of bullwhip effect, better service and lower cost (Karmarkar 1993; Simchi-Levi et al. 2000). It is not only that the *mean* lead time is important - high *variability* in lead times makes planning very difficult. It has been known for some time that reducing the variability of supply/manufacturing lead time usually causes lower levels of raw materials/finished goods safety stock and hence lower costs. Lower variability also may cause lower "safety times" and fewer difficulties in co-ordination and scheduling. For some more advantages of reducing lead times refer to Suri (1998), Hopp and Spearman (2000) and to Chapter 2 of this thesis.

Another effect of lead time reduction is becoming apparent - its effect on final customer demand and price. Though the concept is intuitively appealing, the economics literature does not normally deal with the relationship between demand and lead time; the focus has been mainly on the effects of price on demand. However, empirical studies by Sterling and Lambert (1989), Blackburn et al. (1992), Maltz and Maltz (1998), Smith et al. (2000), etc. suggest that the length of the waiting/delivery time can have a significant effect on customer demand. In industrial markets, a 5% decrease in delivery time can result in almost 24% drop in purchases by the existing customer base (Ballou 1998). Recent operations management literature has begun to recognise this relation by modelling demand as a function of both price and delivery time (So and Song 1998; Palaka et al. 1998; So 2000 and the references therein). Firms also realise that lower delivery times can bring in a price premium. For example, Federal Express can charge almost 50% more for guaranteed next day 8am delivery than for guaranteed next day 5pm delivery. More anecdotal evidences of price premium for shorter delivery times can be seen in Magretta (1998), Blackburn et al. (1992) and Ballou (1998).

Most times, lead time reduction can only be realised by investment (Zipkin 1991). Reductions in lead time might be in the form of investment in better communication, newer machines, improved process design, set-up time reduction, better modes of transportation or possibly standardisation of processes. Several applications of lead time reduction techniques and their effects on market share, internal efficiency and customer satisfaction can be found in Garg and Lee (1999), Suri (1998), a study by Helsinki University of Technology (http://130.233.88.250/hyperlogi) and Hopp and Spearman (2000).

It is typical for investments to yield diminishing returns to the scale of investment. This naturally raises the question of how much to invest in lead time reduction so as to obtain the maximum benefit, measured by an appropriate objective function. It is then necessary to have models that weigh the advantages gained by shorter lead times against the associated costs. The models we develop in this thesis address both the costs and benefits related to lead time reduction issues.

1.2.2 Relation between Product Characteristics and Lead Time Reduction in Supply Chains

The previous sub-section explained the reason behind the recent surge of interest in lead time reduction for supply chains and the costs and benefits associated with such reductions. However, the focus of lead time reduction depends to a large extent on the competitive environment of the supply chain, especially on its product characteristics (Fisher 1997; Ramdas and Spekman 2000; Chopra and Meindl 2001).

What is meant here by product characteristics? We can group most products into the following two categories:

a) Products that are standard or "functional" in nature and hence can be produced before receipt of a customer order are called make-to-stock products. The production system creates goods in anticipation of demand, customer orders are typically filled from existing stock and production orders are used to replenish those stocks. Examples include basic food products like baking soda, standard electrical components like resistors, lumber, diapers and light bulbs (Zipkin 2000; Hopp and Spearman 2000; Fisher 1997);

b) Products that are customised in nature are produced in response to customer demand where each customer waits until his/her order is completed and are called make-to-order products. Examples include custom furniture, courier services, hair-cutting and custom machine tools (Zipkin 2000). Indeed, almost all services are make-to-order in nature.

If we examine the demand characteristics of these two types of products, they are quite different (refer also to Fisher 1997 and Chopra and Meindl 2001 for more details). Normally, the waiting time for make-to-stock products is almost nil (either they are there or not), and customers are primarily price-sensitive. Hence, such firms aim for a cost leadership strategy to attract customers. In the case of make-to-order products, customers wait for the product or service and so they really "feel" the delivery time. For such products, customers are not only price but also delivery time sensitive. Hence, make-to-order firms must differentiate themselves strategically based on price and delivery time they offer to customers (for more on competitive strategy refer to Porter 1998). The nature of make-to-stock products also makes their demand much more predictable than for make-to-order products. Another difference is that the profit margin is typically lower for make-to-stock products than for make-to-order products (Fisher 1997).

It is important to recognise that the supply chain's design must complement the nature of the demand for the product and the competitive strategy (Fisher 1997; Ramdas and

Spekman 2000; Hill and Khosla 1992; Chopra and Meindl 2001). For make-to-stock products, the supply chain should supply predictable demand efficiently at the lowest possible cost, and lead time reduction initiatives to be undertaken should be focussed on cost reduction. On the other hand, make-to-order supply chains should invest in decreasing the delivery time so that it can respond quickly to unpredictable demand, without increasing price "too much".

- **Supply Chain Design for Make-to-Stock Firms**

If some make-to-stock firm uses mark-up pricing based on their operating costs (Wang and Zhao 2000; Hay and Morris 1991), it may wish to invest in improving their processes to reduce their operating costs and gain a larger share of price-sensitive customers. For make-to-stock firms, inventory costs can be a significant portion of operating costs. For example, in the case of retail industry, inventory cost can be as high as 80% of the total operating cost and even for make-to-stock manufacturing firms it is as much as 65% of the total operating cost (Ballou 1998; CAPS Research 2000). Hence it is natural that many such firms have targeted inventory cost reduction as a means to achieve their goal of reducing operating costs (Fisher 1997).

As far as inventory cost reduction is concerned, one of the ways it can be realised is through lead time reduction. The two elements of lead time that have been the greatest targets for cost reduction are - i) supply lead time, and ii) set-up time (or cost) (Chopra and Meindl 2001). The growth of supply chains means that the supply lead time between the elements of the chain is now a competitive priority (Australian National Audit Office Report 1997-98; Chopra and Meindl 2001). Many buyers are ready to pay for the investments undertaken by their suppliers in supply lead time reduction. Even with that cost, the buyer seems to be better off as far as total inventory cost is concerned (Purchasing Online 1998). Similarly, investments in set-up time (cost) reduction can also reduce the total inventory costs for make-to-stock firms (Porteus 1985; Hopp and Spearman 2000). However, the increased demand resulting from cost reductions, in turn, will affect the original investment decision. Hence, make-to-stock supply chains must take this "circularity" into account while deciding on the optimal lead time.

- **Supply Chain Design for Make-to-Order Firms**

Make-to-order firms try to attract customers by catering to their lead time sensitivity. Such firms (or supply chains) use three main strategies to utilise speed to attract customers - i) serving customers as fast as possible, ii) encouraging potential customers to obtain a delivery lead time "quote" prior to ordering, and iii) guaranteeing a "uniform" delivery lead time for all potential customers (for examples of each type refer to So and Song 1998). However, make-to-order firms focus not only on the length of the delivery time but also on its reliability. The issue of delivery time reliability is especially important to retain customers and for repeat business.

For make-to-order firms, one of the ways to achieve shorter delivery times is to invest in increasing their capacity. Since services cannot be inventoried, optimal capacity design is especially important in service sector (Stevenson 1999). Make-to-order firms must also keep in mind two other issues - i) customers may be ready to pay a price premium for early delivery, and ii) if lower delivery times can attract more customers, it may lead to economies of scale for the firm (i.e., lower operating costs). Hence, make-to-order firms must account for all these costs and benefits while determining their optimal delivery time.

1.2.3 Research Agenda

From the above discussion we can conclude that:

a) Lead time reduction is a key concern for supply chains, has many potential benefits, but requires investment;

b) The focus of lead time reduction appears to be different for make-to-stock and make-to-order firms.

Since the mid-1980s, the strategic benefits of models and tools from Operations Research to analyse the consequences of integration and the use of new technologies or processes before their introduction have become well known (Maloni and Benton 1997). Given the widespread recognition of the need for effective supply chain management, and the importance of speed as a competitive prerogative, we decided to focus our thesis on analytical models of lead time management issues in supply chains.

Specifically we wanted to address the following topics:

a) Develop delivery lead time management models for make-to-order supply chains;
b) Develop lead time management models for make-to-stock supply chains focusing on the lead time elements of supply time and set-up time.

While some lead time management issues for both make-to-order and make-to-stock supply chains have been investigated before in the literature, we have not seen any model that takes into account all the elements that we include in ours.

1.3 Modelling Strategy

For our research, we assume a three party supply chain - a profit-maximising firm, its supplier and its final customers. The firm is dealing in a single end product. To address the specific research issues, identified in Section 1.2, we have developed three separate analytical models.

- ***Delivery Lead time Management Model for Make-to-Order Firms***

For this model, we focus on the firm and its customers assuming that the firm is dealing in a make-to-order product/service. The firm announces a uniform delivery lead time for all customers within which they guarantee to satisfy each customer order. The customers are ready to pay a price premium for shorter delivery times and the firm knows that it can obtain economies of scale by attracting more customers. The firm has to invest in increasing capacity so that it will be able to keep the delivery time low even when it is attracting a lot of customers (refer to Figure 1.3(a)). This part of the thesis will develop an analytical model to help a make-to-order firm optimally determine its lead time by trading off the benefits of reduced delivery time against the costs of investment (Model A).

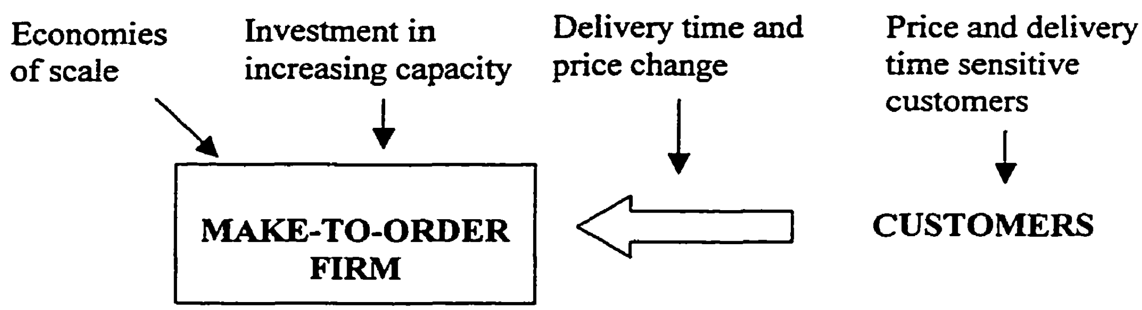


Figure 1.3(a): Supply Chain for Model A

- ***Lead time Management Model for Make-to-Stock Firms***

For this model, we assume that the firm is dealing in a make-to-stock product, which it either produces or buys from a supplier and sells directly to the customers. When the product is procured from a supplier, supply lead time might be the most important component of overall lead time and needs to be properly managed. Hence we first focus on supply lead time management models for make-to-stock firms.

Supply Lead Time Management Model

In this model, we focus on the firm and its supplier. The firm procures material from the supplier in batches following a (Q, r) policy and stores it in its warehouse. The replenishment lead time is stochastic. The customer demand is constant, occurs one unit at a time and is fulfilled directly from the warehouse. The firm wishes to reduce the supply lead time and it is ready to pay for any investment that will be done by the supplier for this reduction. The lead time reduction can be attained through different types of investments. While such reductions will lower the inventory costs for the firm, the cost of the investment must also be accounted for (refer to Figure 1.3(b)). This part of the thesis will model the inventory cost for the make-to-stock firm incorporating both the costs and the benefits of reduced supply time, and help the firm determine the cost-minimising supply lead time, investment and inventory policy values (Model B).

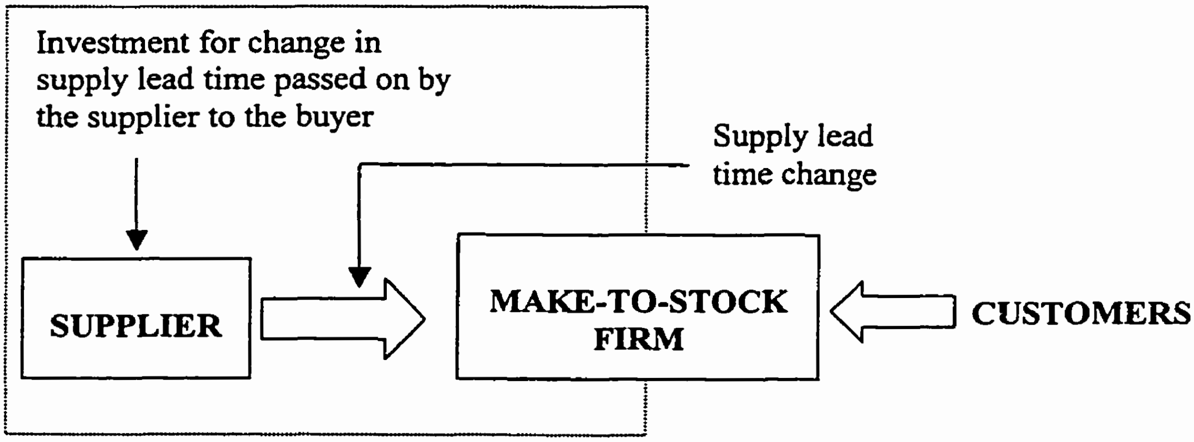


Figure 1.3(b): Supply Chain for Model B

If we assume that the make-to-stock firm uses mark-up pricing over operating costs, proper investments in lead time reduction might bring down the total inventory costs and hence the price of the product for such firms. This implies that the demand will go up since customers are price sensitive. However, the increased demand will have an effect on the extent of the lead time reduction decision itself (Figure 1.3(c)). Hence, in this setting, both price and demand are functions of the operating variables. This part of the thesis will develop an integrated analytical production-marketing model to help the make-to-stock firm maximise its profit by proper selection of lead time (Model C). Specifically, we will focus on set-up time and supply lead time elements of overall lead time.

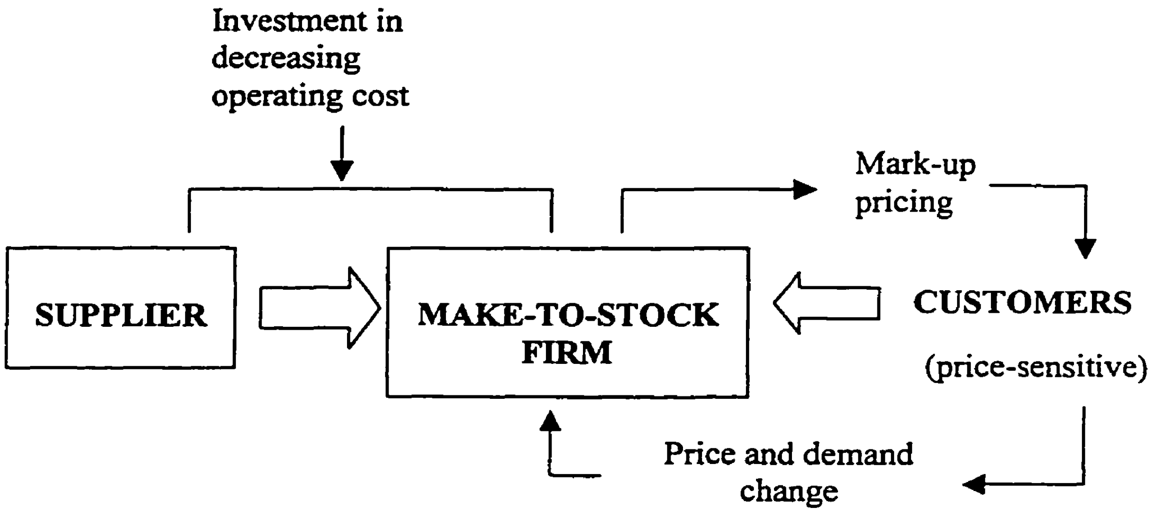


Figure 1.3(c): Supply Chain for Model C

1.4 Organisation of Remainder of the Thesis

To satisfy our research objectives the remainder of this thesis will be organised as follows: Chapter 2 will present a review of the literature relevant to the issues outlined in this chapter. It will also show how components of some previous models can be used in our research and where our research fits into the related body of knowledge. Chapter 3 will deal with delivery lead time management for make-to-order firms taking into account the demand and price characteristics for such products (Model A). Chapter 4 will tackle the issue of supply lead time management and how investment in supply lead time reduction can minimise inventory costs for make-to-stock firms (Model B). In Chapter 5 we will present a model of how make-to-stock firms can determine their set-up times and supply lead times to maximise their profits (Model C). An overall summary of the results and recommendations for future research will be presented in Chapter 6.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

In the previous chapter we dealt with the motivation behind this research and its relevance to modern business environments. While lead time reduction is beneficial for all supply chains, it requires investments whose goals vary depending on the competitive environment of the supply chain. Specifically, we discussed make-to-order and make-to-stock supply chains and how the different demand characteristics of these two require different supply chain design and lead time reduction foci. In this thesis, our aim is to develop models that trade-off the costs and benefits of lead time reduction to determine optimal lead time for make-to-order and make-to-stock firms. Though no previous research captures all the issues that we address, we draw on work done in related areas. This chapter will examine related models and explain their significance to our research.

In this chapter we will address the literature related to the following:

a) Growth of supply chains and the buyer-supplier interface;
b) Importance of lead time, in general, in modern enterprises and specifically the relation between lead time and inventory costs;
c) (Q, r) continuous review stochastic inventory models;
d) Process improving investments, especially investments in lead time reduction;
e) Price and lead time sensitive demand;
f) Integrated Production-Marketing models;
g) Mark-up pricing models;
h) Forms of Investment and Demand functions.

2.2 Growth of Supply Chains

From the late 1980s, the corporate world increasingly used outsourcing for products and services once made in-house. This led to the growth of supply chains (Bryne 1996). More and more firms are now focusing on better management of their supply chains. The goal of effective supply chain management is efficient integration of all parties of the chain in order to minimise system wide costs while satisfying the service level requirements (Simchi-Levi et al. 2000). This integration can be difficult because of the decentralised nature of the chain, short product life cycles and increased customer expectations; however, properly integrated supply chains have a bottom line advantage. According to recent estimates there is a difference of about six to eight per cent savings in yearly revenue between the performance of an average and first-class supply chain. That amounts to approximately \$60 to \$80 million in savings for a company with a billion dollars in annual revenue (Gort 2000).

One of the basic issues in proper integration of supply chains is a close buyer-supplier relationship. The success of the Japanese manufacturers in utilising this relationship has been highly publicised. Frazier et al. (1988) pointed out that such relationships sometimes require specialised or "idiosyncratic" investments by the supplier. This can lead to relatively "high risk" from both the buyer's and the supplier's point of view and in some cases costs of these investments are likely to outweigh its benefits. Newman (1989) and Kalwani and Narayandas (1995) have also expressed similar concerns and pointed out that this might lead to problems in the relationship. In their research on idiosyncratic investments, Fazel et al. (1998) and Levi (1999) point out that the supplier might pass some of the relation-specific investment back to the buyer that requires the investment.

In summary, we can conclude that:

- The recent emphasis on effective supply chain management shows that our research is timely.

- Articles related to the buyer-supplier interface support our modelling assumption that if there is a substantial relation-specific investment required by a supplier, the supplier could pass a part or whole of the investment to the buyer.

2.3 Importance of Lead Time in Modern Business

Seminal studies by Stalk and Hout (1990) and Blackburn (1991) led firms to realise the importance of speed in modern business. This was further strengthened by the results from empirical studies by Jackson et al. (1986), Marr (1994) and several others (refer to Ballou, Chapter 4, 1998). Firms started to focus on reduction of time throughout their operations like supply time, set-up time, manufacturing time and delivery time. It became clear that "time-based competition" is a survival strategy for firms. To effectively compete, firms need to differentiate themselves based on price and length and reliability of the lead time offered to customers (i2's White Paper on E-Business 2000; Monczka and Morgan 2000; Chopra and Meindl 2001).

There is a huge volume of research pointing out the advantages of lead time reduction, some of which were already discussed in Chapter 1. With longer lead times, schedules must be frozen over a longer horizon. This increases the chance of incorrect demand forecast. Longer and more variable lead times are also usually associated with higher lead time demand variability. In general, safety stocks are related to the variability of demand over lead time, which grow larger as lead time increases leading to "excess" safety stocks. Hence, lead time reduction can result in reduction of inventory costs. In production for assembly, variability causes difficulty in co-ordinating parts and requires large intermediate and finished goods safety stocks (Karmarkar et al. 1985). Other benefits of lead time reduction include the ability to quickly fill customer orders that cannot be filled from stock, and reduction in the bullwhip effect (Simchi-Levi et al. 2000).

In more recent times, shorter delivery lead times have been also associated with larger demand and a price premium, especially for make-to-order products (Ballou 1998; So and Song 1998; Weng 1999). Recent research (Garg and Lee 1999) has indicated the potential of lead time reduction strategies in handling product variety. There have been several other recent studies focusing extensively on the techniques of lead time reduction and its effects (Karmarkar 1993; Groenevelt 1993; Song and Zipkin 1996; Suri 1998; Garg and Lee 1999; Hopp and Spearman 2000).

In most of the above research, the emphasis is mainly on the benefits of lead time reduction and the methods of achieving it, ignoring the costs associated with such reductions.

2.3.1 Lead Time and Inventory Costs

Many of the analytical models related to lead time reduction have focussed exclusively on the effect of such reductions on inventory costs. The reason for this focus is the important role played by inventory in the modern economy. As we indicated in Chapter 1, inventory costs comprise the major portion of the total operating cost for a make-to-stock firm. In more general terms, as of March 1999, businesses in the United States excluding government and not-for-profit firms held about $1.1 trillion worth of inventories which is much more than their total monthly sales (Zipkin 2000). Many of the success stories of the last few decades, like Japanese manufacturing companies, Wal-Mart or Dell can be attributed to their ability to operate with substantially lower inventories than their counterparts (Zipkin 2000). One of the ways in which these organisations achieve a lower level of inventory is by lead time reduction throughout their supply chains (Chopra and Meindl 2001).

It is during any replenishment lead time that the likelihood of a stock-out is highest. Hence, the most important element in any inventory related research is the lead time demand (LTD). LTD is composed of two parts: lead time duration and demand per unit time. Inventory models are normally categorised based on their LTD characteristics - deterministic or stochastic, and review frequency - periodic, i.e., inventory level is known only at certain points in time, or continuous, i.e., inventory level is known at all times. There are two basic issues in any inventory model - when to order (reorder point) and how much to order (batch size). With the increased use of information technology for inventory tracking, inventory levels are more or less known at all times. Hence, the most popular inventory control model used in practice is the continuous review, order quantity/reorder point type (Q, r) policy (Chopra and Meindl 2001). With this policy, an order for quantity Q is placed as soon as the inventory position (= inventory on hand + inventory on orders – backorders) drops to a fixed reorder point, r (Zheng 1992).

Economic Order Quantity (EOQ) models and their extensions are used for models with deterministic LTD. For such models, calculation of a reorder point is straightforward and the primary decision variable is the batch size (Q). EOQ models determine the optimal batch size that will minimise the relevant inventory costs per unit time by trading off the set-up cost and holding cost. For a detailed review of how EOQ models and its extensions can help firms in reducing their inventory costs refer to Lee and Nahmias (1993).

If either or both the constituents of LTD are random, the LTD will also be stochastic. For stochastic inventory models, both the reorder point (r) and the batch size (Q) are decision variables and the objective is normally the minimisation of expected inventory costs per unit time. These are called as (Q, r) models. For stochastic LTD, stock-outs are possible. Unmet demands are either backordered or lost. Stochastic models deal simultaneously with two trade-offs - i) between set-up and holding costs, and ii) between backordering and holding costs.

There are certain models that primarily focus on the trade-off between the holding and backordering costs. These models can be either single-period (newsvendor problem) or multi-period (base stock policy). Gerchak and He (2000) used a mean preserving transformation, which changes the demand variability while keeping the mean constant, to show the effect of randomness of demand for an arbitrary demand distribution on optimal inventory cost and order quantity in a newsvendor problem. While the optimal inventory cost always increases with variability, the optimal order quantity does not necessarily follow the same pattern. Gerchak and Mossman (1992) point out the conditions under which the optimal order quantity also increases with variability. Song (1994a) used stochastic ordering for a model where demands form a compound Poisson process and lead times are stochastic. The author investigated the effect of both stochastically larger and more variable supply lead times when base stock policy is optimal. While the optimal base-stock level is higher for larger lead times, it does not necessarily lead to higher optimal long-run average cost. More variable lead time always leads to a higher optimal average cost but the optimal base stock level depends on the cost structure and will be higher if and only if the penalty cost is high relative to the holding cost rate. In a related article (Song 1994b), a similar model was investigated with the performance measure being infinite horizon expected total discounted cost. In this case it is not always true that larger lead-time demand will have larger optimal base stock level. For a detailed review of inventory models, in general, refer to Graves et al. (1993) and Zipkin (2000). For stochastic inventory models, particularly single-period or multi-period problems refer to Porteus (1990). Since our primary interest is in continuous review stochastic inventory models with backordering, we will consider them in greater detail.

2.4 Continuous Review (Q, r) Stochastic Inventory Models with Backordering

Interest in continuous review, stochastic inventory models with backordering started with the classic work of Hadley and Whitin (1963) who developed such an inventory cost model with backordering cost per unit (no time dimension) and the assumptions of one order outstanding and a positive reorder point. They proved the joint convexity in Q and r of the cost function under some restrictive conditions and showed how to determine the cost-minimising batch size and reorder point. For the case of one order outstanding and backordering cost per unit per unit time they developed a model for Poisson demand, but could not prove the joint convexity of the cost function. Other studies that analysed and compared the effects of lead time demand on inventory levels and costs used different techniques like simulation (Gross and Soriano 1969) and numerical computation (Vinson 1972; Naddor 1978).

Research on stochastic (Q, r) inventory models has shown that a key issue is the variability of LTD, especially the standard deviation, whether it originates from demand or lead time duration. The variability of LTD requires some safety stock, the stock in excess of mean lead time demand, for the situations when LTD is high. An analytical study by Das (1975) showed that in a Hadley-Whitin type of (Q, r) inventory model, the average inventory cost is dependent on the variance of the lead time demand and not on the mean. Bagchi et al. (1986) use a case study and numerical examples to show the importance of incorporating the variability of lead time in determining the distribution of demand during lead time and safety stock levels. They recommend that a compound distribution of demand during lead time, or a good approximation of it, be used to calculate safety stocks. In a similar vein, Eppen and Martin (1988) also considered setting safety stock levels in the presence of stochastic lead times for cases when both lead time and demand are random variables. Their research concerns cases when the parameters of the distributions are known as well as the case where they are unknown. For some more early research in this area refer to Zheng (1992) and Lee and Nahmias (1993).

Up to the mid 1980s most of the research in (Q, r) models was based on approximate models following the traditional Hadley-Whitin framework of time-independent backordering costs and a single order outstanding. The assumption of single order outstanding allowed variability of the LTD to be modelled due to variability of both demand and lead time duration. For models with multiple orders outstanding, lead times were assumed to be constant with the variability of LTD coming only from demand. Without this assumption, models with stochastic lead times and multiple orders outstanding created the problem of order crossing.

The (Q, r) model got a major boost from the seminal works of Zipkin in mid 1980s. Zipkin (1986a) showed that if the supply system is assumed to be exogenous and sequential then we can develop (Q, r) inventory models with stochastic lead times even for more than one order outstanding (for more details refer to Zipkin 2000). In the same paper, Zipkin also showed that the relation between the limiting values of the random variables – Inventory level (*IL*), Inventory position (*IP*) and lead time demand (*LTD*),

$IL = IP - LTD$, is valid under very general conditions and that IP and LTD are independent. Zipkin (1986b) showed that popular approximations of (Q, r) models with backordering cost per unit backorder per unit time, where the backordering cost term is simplified assuming single order outstanding, may perform poorly when mean lead time demand is large compared to Q or when lead time demand is highly variable. He went on to prove that the exact expression for backorders, with more than one order outstanding and possible negative reorder point, is a jointly convex function of the control variables, Q and r. The implication of these two papers is that it was possible to develop "exact" (Q, r) models without any need for assumptions like one order outstanding, positive reorder point or constant lead time. In a recent article, Zhang (1998) significantly simplified Zipkin's (1986b) proof of the joint convexity of the inventory cost expression in terms of Q and r.

The understanding of the exact continuous review stochastic (Q, r) inventory models was further augmented by Zheng (1992) who compared such models to corresponding deterministic EOQ models. The backordering cost considered was per backorder per unit time. The research was an extension of Federgruen and Zheng (1988) where a discrete model was considered and a simple and efficient algorithm for cost minimisation and calculation of control variables was developed. The main results from the paper (1992) show that average inventory costs and optimal order quantity in stochastic models are larger than their deterministic counterparts. An interesting result is that the relative increase in the costs incurred by using the quantity determined by EOQ instead of that from the stochastic model is no more than 1/8 and vanishes when the ordering costs are significant relative to other costs. Gallego (1998) extended Zheng's research by capturing the distributional information about lead time demand into its mean and variance and solving the resulting problem against the worst possible distribution. This is sometimes called a maximal approximation. Gallego obtained bounds on optimal long run average inventory costs and batch size for exact (Q, r) models using the maximal approximation. Gallego showed that with $\sqrt{2}$ EOQ batch size, the cost penalty would be no more than 6.07% of the overall optimal.

The absence of closed form expressions for (Q, r) models, even for approximate models, means that analytical comparative statics for such models is relatively difficult. Gerchak (1990) showed the direction of change in the reorder point as shortage penalty, expected demand and holding costs are changed in the (Q, r) model under the assumptions of single order outstanding and backordering cost independent of time. Bookbinder and Çakanyildirim (1999) have also performed first order analytical comparative statics with the assumptions of single order outstanding and positive reorder point. Recently De Groote and Zheng (1997) and Zipkin (2000) have shown how the optimal reorder point and batch size will vary with changes in set-up cost, holding cost and backordering cost for an "exact" (Q, r) model. They have also shown how the inventory cost increases with the standard deviation of the LTD and developed limits to the optimal cost and batch size in terms of standard deviation of LTD.

Based on the above discussion we can conclude that:

- The importance of lead time in modern business is consistent with our emphasis on models for effective lead time management in supply chains.
- The continuous review stochastic inventory models are concerned mainly with understanding the stochastic inventory systems more clearly and relating them to deterministic models. While some researchers show that an increase in variability of lead time demand can adversely affect the cost, none of them addresses the issue that there might be a price to be paid to reduce the variability. In our research we will account for both the costs and benefits associated with lead time reduction to determine the "optimal" lead time. The "exact" (Q, r) model, with proper modifications, will form the basis of our model.

2.5 Process-Improving Investments

The success of Japanese firms with their policy of continuous process improvement created the impetus for a large number of firms to invest in process improvements. This also had an effect on the research paradigm in the production/inventory area by changing the nature of certain parameters from exogenously given to endogenously determined by proper investments (Gerchak and Parlar 1991). A substantial amount of research has been done in the area of process-improving investments. Some examples include, research on - i) investments in reduction of yield randomness in an EOQ model (Gerchak and Parlar 1990), ii) investments in process quality improvement (Porteus 1986b), and iii) how investments by one entity of a supply chain affects the other parties of the chain (Gilbert and Cvsa 2000). For detailed review of this literature refer to Nye (1997) and Ray, Gerchak and Jewkes (2000). However, in our research the focus is on investments in lead time reductions, specifically on set-up time and supply lead time reductions.

2.5.1 Investments in Reduction of Set-up Time

Most of the research dealing with investments in set-up time reduction assumes set-up cost to be a surrogate for set-up time, i.e., investments in set-up time reduction will also reduce the set-up cost. One of the earliest proponents of the research on investments in set-up cost reduction was Porteus (1985). In the first part of his paper, Porteus showed that such investments in a traditional average inventory cost per unit time EOQ model makes sense solely on the basis of benefits obtained in the form of reduced inventory costs. The basic setting in Porteus' research is a classical undiscounted EOQ model with the option of investing in reducing set-up cost. The EOQ model is optimised for batch size and set-up cost. The total cost function per unit time includes both the inventory costs (set-up costs + holding costs) and an opportunity cost for the investment. For two special cases of investment functions, Logarithmic and Power – both decreasing convex in set-up cost, Porteus was also able to show that the objective function will be strictly concave-convex in the relevant region and has a unique local minimum. Porteus subsequently extended his work to the cases of discounted EOQ models (1986a) and simultaneous investments in set-up cost reduction and process quality improvement (1986b). Leschke and Weiss (1997) extended the work of Porteus (1985) to help managers decide how to allocate investments in set-up cost reduction programs in a multi-product environment. They use a transformation of Porteus' model to show that it is better to standardise set-ups across several products in stages than to focus on a single set-up and reduce it as much as possible before proceeding to the next set-up.

Nasri et al. (1990) dealt with the issue of investing in reduction of set-up costs in a model with stochastic lead times and time-independent backorders. Normally, set-up cost reduction models assumed a static cost reduction approach, i.e., the decision to invest in reduction is made only at the initial set-up. In Hong et al. (1996), the authors examined different production policies where a decision to reduce set-up costs could be made at the beginning of each planning cycle.

Nye (1997) considered the question of interdependence of investments for improvement in manufacturing processes. The specific improvement practices considered are setup time reduction and quality improvement, both of which require investment. This research considers not only the traditional EOQ model for showing the interdependence but also the effect of congestion in the form of WIP costs. In Hariga (2000), the author investigates the "approximate" (Q, r) model of Hadley-Whitin with normally distributed lead time demand where lead time is a deterministic function of batch size and set-up time and investments can be made to reduce the set-up time. For a detailed review of literature on set-up time (cost) reduction refer to Nye (1997).

2.5.2 Investments in Reduction of Supply Lead Time

In EOQ models the only way we can benefit from lead time reduction is through set-up time reduction that also reduces set-up cost. In a deterministic demand scenario, any change in supply lead time duration will only change the reorder point but not the optimal batch size or the cost. Hence, almost all the models related to replenishment lead time reduction deal with stochastic LTD in a (Q, r) framework. These models can be divided into two groups - i) those where variability of lead time demand (LTD) is due only to demand variation while lead time duration is deterministic, and ii) those where variability of LTD is due only to variability of lead time duration and demand is constant.

The earliest research where variability of LTD is from stochastic demand and deterministic lead time duration seems to be that of van Beek and van Putten (1987). They showed how in the classical Hadley-Whitin (Q, r) model, in addition to batch size and reorder point, lead time duration can be controlled by investment. Hill and Khosla (1992) extended that model by assuming an investment cost more general than van Beek and van Putten (1987). Liao and Shyu (1991a) consider Poisson demand in a (Q, r) setting and the objective is to minimise the expected inventory cost per unit time by determining the optimal reorder point and lead time pair while batch size is assumed to be known. In a related paper, Liao and Shyu (1991b) consider the same model with normal demand, and where lead time is decomposed into components each having different piecewise linear crashing cost for reduction. Ben-Daya and Raouf (1994) extend the Liao and Shyu (1991b) normal demand model to also include order quantity as a decision variable. They consider the crashing cost to be a continuous function of lead time. Li et al. (1997) consider a (Q, r) model with backordering where there is a cost per unit backordered only. Investment can be made to reduce the decision variables - set up cost, lead time and variance of demand forecast error simultaneously. Hariga and Ben-Daya (1999) consider models with partial backordering and lost sales where there is a crashing cost associated with reducing lead time. The model is solved for both complete and partial information about the lead time demand distribution.

The earliest research where variability of LTD is due only to variability of lead time duration and demand is constant seems to be that of Gerchak and Parlar (1991). The authors dealt with the classical Hadley-Whitin continuous review (Q, r) inventory model with backordering cost per unit backordered. This is one of the few papers that use the mean preserving transformation to capture the effect of investments without assuming any particular LTD distribution. The decision variables are reorder point, batch size and variability of LTD. Paknejad et al. (1992) analyses the options of investment in reduction of lead time variance only or reduction of lead time variance and set-up cost simultaneously. The basic model is a finite range stochastic lead time inventory model with backordering cost per unit per unit time. Numerical results indicate that simultaneous investments result in lower cost and batch size than separate investments, pointing towards meaningful interaction between reduction in lead time variance and set-up cost. For Choi (1994) the variables considered are variance of lead time duration and quality level. A service level criterion is used rather than explicit backordering costs in a stochastic, continuous review (Q, r) model. A recent paper by Bookbinder and Çakanyildirim (1999) also considers a (Q, r) model with backordering cost per unit per unit time. Their model assumes single order outstanding and compare the model where lead time is endogenous to the case where it is exogenous. For a detailed review of literature on supply lead time reduction refer to Ray, Gerchak and Jewkes (2000).

These papers show that costs and benefits associated with lead time reduction make it imperative to balance the two to arrive at the optimal solution. They also show that intelligent investments in lead time reduction can lead to inventory cost reduction. We will follow this approach.

However, most models dealing with investments in set-up time/supply lead time reduction focus on cost minimisation assuming demand to be constant or stochastic with mean demand rate being constant and do not capture the effect of reduced cost on market demand. The recent emphasis on "integrated" supply chain models implies that it is important to consider not only the effect of investments on efficiency but also their ultimate effect on price and demand. We will refer to some "integrated" supply chain models in Section 2.7 and indicate why it is especially necessary to develop such models to investigate process-improving investments.

Also there are two issues that have not yet been fully addressed in the existing research on investments in supply lead time reduction in (Q, r) models:

a) Most focus on models with assumptions like backordering cost per unit, disregarding the duration of the shortage, or one order outstanding. However, most recent research on (Q, r) models assumes a backordering cost per unit per time, allows more than one order outstanding and negative reorder point. Our work deals with investing in lead time reduction in the latter framework;

b) Most models assume that the investment in lead time reduction is a one-time investment. However, several other investment strategies might be used for such reductions. We will show that ignoring the nature and frequency of the type of investments in lead time reduction when deciding on the "optimal" strategy may result in sub-optimal decisions.

2.6 Price and Lead Time Dependent Demand

The relation between price and demand is one of the best known relations in microeconomics. As we indicated in the previous section, for a long time, operations management (OM) did not emphasise the demand side of any supply chain. In the few instances where demand was not assumed to be totally exogenous, it was assumed to be sensitive to price only, but unaffected by operational variables. There is some literature in OM that deals with price-sensitive demand (refer to Porteus 1990; Eliashberg and Steinberg 1993; Petruzzi and Dada 1999).

Recent OM literature recognises that long customer waiting times might have an adverse effect on the demand rate, especially for make-to-order products. Hence, for such products demand rate should not only be dependent on price but also on delivery/waiting time. Customers might be even ready to pay a price premium for shorter delivery times for such products. However, we must keep in mind that firms may need to invest in increasing capacity to shorten their delivery times. Also, models have to account for the congestion that might be caused by the increase of demand. Congestion can lead to increased WIP cost and/or waiting time for customers.

Research in lead-time-dependent-demand models typically focuses on internal pricing and capacity selection issues for service facilities by taking into account user's delay costs and capacity costs (Dewan and Mendelson 1990; Stidham 1992). The consumer's choice depends on price and on the waiting time (full price = price charged + cost of waiting). As an increase in demand might increase congestion and thus the waiting time (leading to decrease in demand), the firm's profit will depend on scheduling, outsourcing and pricing decisions.

Many of the researchers have used a game-theoretic framework for investigation of pricing and capacity selection issues. Li (1992) explored the role of inventory in response time competition by examining the behaviour of customers and competing firms. Lederer and Li (1997) studied the issue of competition between firms serving delay-sensitive customers and the resultant effect on price, production rate and scheduling policies. In a recent paper, Ha (1998) has extended the research of Dewan and Mendelson (1990) by deriving incentive-compatible pricing schemes that can achieve optimal arrival rates and induce delivery-time-sensitive customers to choose optimal service rates when service is jointly produced by the customers and the facility. For some other related research based on game-theoretic frameworks, refer to So and Song (1998).

There exist various other streams of literature that investigates lead-time dependent demand and/or price. This includes the use of quoted customer lead times to explore the impact of due-date setting on demand and profitability (Duenyas and Hopp 1995; Weng 1999). Hill and Khosla (1992) constructed a model where demand is a function of actual delivery time and price and the firm's objective is to maximise profit by optimal selection of price and lead time. But their model is totally deterministic. On the other hand, Buss et al. (1994) determine the best production capacity where demand is stochastic but do not consider the impact of lead times on demand. Weng (1996) models price premium for shorter lead times but does not consider the effect of lead time and/or price on demand rate or any investment in increasing capacity.

While all these lines of research are important, as So and Song (1998, pg 30) point out, they are basically different from the recently popular strategy of committing to a "uniform" delivery time guarantee for all customers, the focus of our research. In the case of a delivery time guarantee, firms advertise a uniform delivery lead time for all customers within which they guarantee to satisfy each customer order. The length of the delivery time guarantee is a decision variable that directly affects overall demand. In practice, usually it is very difficult to quantify user's delay costs, which is used by most research. Therefore, it makes sense to use a reliability constraint to ensure a satisfactory service level once the uniform delivery time guarantee is selected. The strategy of committing to a uniform delivery lead time has been investigated by So and Song (1998), Palaka, Erlebacher and Kropp (1998), So (2000) and Rao et al. (2000b). The basic setting is that the demand rate is a function of price and/or length of the uniform guaranteed delivery time. Some investment needs to be made in increasing capacity so that the delivery times can be reduced. There is also a service level constraint and/or WIP holding and penalty costs. The main objective is to find the optimal price and/or guaranteed delivery time that maximises profit per unit time. While So and Song (1998) and Palaka et al. (1998) deal with a single firm, So (2000) extended So and Song's work by analysing the impact of using delivery time guarantees in the presence of competition. Rao et al. (2000b) integrate uniform delivery time guarantee strategy with production planning. For more detailed analysis of each of the four papers refer to Chapter 3.

Based on our above discussion we can conclude that these papers clearly show the recent trend of assuming demand to be a function of both price and lead time. Some of these papers directly address the issue of a uniform delivery time guarantee for all customers, which is also the focus of one part of our research.

Our model will be significantly different from the existing research on uniform delivery time guarantee in two regards:

a) We will address the issue of a price premium for shorter delivery times;
b) We will explicitly model the economies of scale that may be realised through increased demand by committing to a shorter delivery time.

2.7 Integrated Production-Marketing Models in the Presence of Investment

Marketing often has incentives based on revenue while production has incentives based on cost. However, actions that maximise revenues or minimise costs may not maximise profits. Hence, it is necessary to develop models that take into account the production-marketing interactions to attain the goal of maximising profitability of a firm (Chopra and Meindl 2001). Eliashberg and Steinberg (1993) also give a useful argument for why joint production-marketing decision-making is important, and provide a comprehensive review of such integrated models up to the late 1980s. Such integrated modelling attains even more importance when firms invest in reducing their costs to increase demand.

Though the volume of integrated production-marketing literature is not very large, its history is quite long. Note that here we will mainly discuss continuous time concave-cost models with static pricing (for literature review on dynamic pricing models refer to Deng and Yano 2000). The first integrated production-marketing model of this kind was formulated by Whitin (1955) who incorporated pricing into the traditional framework of the EOQ model through a linear price-demand relation where demand is price-sensitive. The objective was to determine the price the firm should charge in order to maximise its profits. This problem was later explicitly solved by Porteus (1985, Section 6).

Porteus (1985, Section 7) was the first that approached the problem of joint production-marketing decisions when an investment is made in changing some operating parameter. Eliashberg and Steinberg (1993) pointed out that incorporation of investment costs associated with changing the set-up cost makes Porteus' model much more realistic than Whitin's model. Portues modelled demand as price-sensitive and examined the situation where the firm can invest in reducing the set-up cost in a traditional EOQ framework. The objective of the firm is to maximise its profit (revenue - production cost - holding cost - set-up cost - investment cost) by optimal selection of demand rate (or price) and set-up cost. The complex nature of the problem means that the problem could be solved only for some special demand and investment functions. Though Porteus' paper's first part, where he proves the inventory-cost reducing effect of investments in set-up cost reduction, might be one of the most widely cited OM articles of recent times, the last part, despite being more general, seems to have been largely ignored.

More recent models following Whitin's or Porteus' framework includes Cheng (1990), Sajjad and Sarker (1991), Min and Chen (1995) and Weng (1995). These models assume price to be a decision variable, independent of operating costs. Zipkin (1992) also formulated a model of an integrated production-marketing system through a queue that takes into account the congestion effect for higher demand. However, price is an independent decision variable and the objective is to maximise the firm's profit with respect to price, batch size and reorder point.

Ladany and Sternlieb (1974) consider price to be a fixed mark-up over production costs in a profit-maximising EOQ model where the demand rate is decreasing in price. Lee and Nahmias (1993) point out that explicitly relating operating costs and demand makes the objective function much more complex and simplifying assumptions about the demand and production cost functions are needed to obtain closed form expressions. Lee and Rosenblatt (1986) also consider a model similar to that of Ladany and Sternlieb but include advertising investments and quality problems.

Other streams of literature that have modelled integrated production-marketing issues include the economics of queues (refer to So and Song 1998 and Section 2.6) and single-period stochastic inventory models with pricing (refer to Petruzzi and Dada 1999). In most of this literature, price is treated as an independent decision variable and either no investment are considered or the investments are in increasing capacity.

All the above models imply that since process-improving investments have the prospect of decreasing costs and increasing market demand, it is necessary to develop integrated production-marketing models to investigate such investments. This issue will be one of the cornerstones of our research also.

However, our models differ from the existing literature in two major aspects:

a) The models that consider price as an independent decision variable do not explicitly account for the effect of operating costs on price and demand. When the product is such that the profit margin is very low, it is quite natural to determine price as a mark-up over the total operating cost. In those situations it is more realistic to assume that price and demand are both functions of the operating variables. However, we acknowledge that for high-profit goods it might be more natural to assume price to be an independent decision variable;

b) The models that consider price to be a mark-up (e.g., Ladany and Sternlieb 1974), it is a mark-up over the production costs only and not the entire operating costs. In addition, these models do not take into account the investment required to effect process improvements.

As we will show later, both of these issues will have a significant effect on modelling.

2.8 Mark-Up Pricing

Two common methods of pricing referred to in the literature are cost-based pricing and market-based pricing. One of the most popular cost-based methods is mark-up pricing where prices are established based on an estimated total cost plus a percentage mark-up. The mark-up rate depends on the product line, tradition, competition, and other market factors (US Department of Defense Contract Pricing Reference Guides 2000). This type of pricing is frequently used for make-to-stock products in the manufacturing sector, in the apparel industry and in the retail industry (refer to Chapter 5).

Hall and Hitch (1939) and later empirical studies by Eckstein and Fromm (1968) and Coutts, Godley and Nordhaus (1978) found that many firms set prices relative to some notion of average cost and a reasonable mark-up to cover profits. According to Hay and Morris (1991) almost 75% of firms use some variant of mark-up pricing. Bloch and Olive (1997) noted that cost changes play a dominant role in determining prices. Their model also seems to suggest that the change in price and cost for manufacturing companies are proportional, suggesting a mark-up model. There are several models involving mark-up pricing in the operations management literature (Ladany and Sternlieb 1974; Lee and Rosenblatt 1986; Wang and Zhao 2000 to name a few). For more details on mark-up pricing refer to Hay and Morris (1991) and Lohr and Park (2000). The relation between mark-up pricing and the profit-maximisation pricing advocated in the economics literature will be discussed in Chapter 5.

Based on the ample empirical evidence regarding the use of mark-up pricing for make-to-stock products, we will assume that as the pricing technique for our make-to-stock supply chain.

2.9 Forms of Investment and Demand Functions

There are a variety of functional forms used for process-improving investments in the literature - general convex, exponential, logarithmic, power, piecewise linear. It is very difficult to say which type of investment function mirrors the practical situation most closely. A comprehensive list of different types of investment functions used in the literature has been given in Nye (1997).

There are two main types of demand functions found in OM literature: linear (Porteus 1985; Palaka et al. 1998) and constant elasticity (Hill and Khosla 1992; Weng 1995; So and Song 1998). As the name implies, in case of constant elasticity the price elasticity of demand (or lead time elasticity, as the case may be) remains constant while the slope of the curve changes. For the linear demand case, demand elasticity increase with price while the slope remains constant. Many researchers have used linear demand functions because of analytical simplicity.

In our research, the demand and investment functions will be selected to strike a balance between analytical tractability and reality.

2.10 Summary

From the literature review we can conclude that though there are significant relevant research contributions to lead time management issues for make-to-stock and make-to-order products/services, there are some gaps in them:

- The existing research dealing with determining optimal delivery lead time for make-to-order product/services does not account for the price premium from lower delivery times and the economies of scale from higher demand.
- The relevant research on the effects of investments in supply lead time reduction on inventory costs in a make-to-stock scenario does not take into account all possible types of such investments and the resultant effects and none of them models the effect of lower costs on customer demand.
- None of the research investigating integrated production-marketing models for make-to-stock products with process-improving investments in set-up time/supply lead time is based on mark-up pricing over the entire operating cost and hence does not explicitly account for the effects of all the operating variables on price and demand.

Our research, investigating lead time management issues in make-to-stock and make-to-order supply chains, will address all the above "gaps" so as to develop more comprehensive models. Our research aims to make a significant contribution to the operations management literature on "time-based" competition.

CHAPTER 3

DELIVERY LEAD TIME MANAGEMENT FOR MAKE-TO-ORDER FIRMS

3.1 Introduction

Firms specialising in make-to-order products or services often use a time-based competitive strategy, since customers are not only sensitive to the price they are paying but also to the length and reliability of the delivery lead time. For example, Japanese machine tool exports to the United States (US) surged from $22.1 million in 1973 to $687.5 million in 1981. Much of this increase had to do with the shorter and more reliable delivery lead times offered by the Japanese manufacturers. The traditional practice of order backlog management used by US machine tool manufacturers implicitly assumed that the customers would wait for the make-to-order machine tools. However, by the late 1970s and early 1980s many foreign firms, especially Japanese ones, started to offer fast delivery of quality machines to US customers. Very quickly, many of the US customers changed their allegiance to the foreign firms (National Research Council Report 1983). Accordingly, this chapter will focus on how make-to-order firms (or supply chains) can manage their delivery lead times to maximise their profits (Model A of Chapter 1).

While there have been a number of strategies used by make-to-order firms to use speed to attract customers, guaranteeing a "uniform" delivery lead time for all potential customers has become quite popular recently. Many companies are adopting the strategy of advertising a uniform delivery lead time for all customers within which they guarantee to satisfy each customer order (So and Song 1998). This includes manufacturing firms like Titliest/Foot-Joy, a leading manufacturer of customised golf balls, LeatherTech, a manufacturer of customised leather furniture (Rao, Swaminathan and Zhang 2000b) and also service facilities like Wells Fargo Bank, Lucky supermarket and Federal Express (So and Song 1998).

While the strategy of offering a uniform delivery lead time guarantee may attract many customers, there is a risk if the firm announces a very short delivery lead time that attracts a lot of customers. The demand may then exceed the companies' capacity to respond. In such situations, the waiting time for customers may be greater than the guaranteed delivery lead time offered by the firm. This can lead to a penalty cost for the manufacturer and/or it might lead to decrease in repeat business. With this strategy, it is important to have some internal mechanism to ensure that the promised delivery lead times are feasible and reliably met.

The traditional economics literature deals primarily with the effect of price on customer demand, but not with the effect of lead times. Since the late 1980's, a large volume of operations management literature started to recognise that customer demand increases both with shorter lead times and lower prices (Hill and Khosla 1992; Duenyas and Hopp 1995; So and Song 1998; Ballou 1998). Increased demand, in turn, can bring down unit operating costs through economies of scale (Scherer 1980). Recent studies seem to suggest that the effect of delivery lead time is more than just on the demand rate. Karmarkar (1993) pointed out that lead times are most probably inversely related to market share or price premiums or both. Ballou (1998) also noted that shorter delivery lead times could result in a price premium. For example, shipping costs from Amazon.com are more than double when the delivery lead time guarantee is around two days than when it is around one week. While there has been pressure on firms to reduce their delivery lead times, this pressure has opened up new opportunities for companies that are able to satisfy this requirement. Some customers are ready to pay a price premium for shorter and more reliable delivery lead times. Many cutting-edge supply chains are aware of this added incentive to reduce delivery lead times. One of the ways that firms can reliably satisfy a guaranteed delivery lead time is by investing in increasing capacity (So and Song 1998; Palaka et al. 1998). The firms then must trade-off the potential for increased demand and price against the costs of investment.

In this chapter, we model a supply chain of a make-to-order firm and its customers where the firm is using the strategy of announcing a "uniform" delivery lead time guarantee for all its customers. Customer demand is random and the mean demand rate is a function of both price and guaranteed delivery lead time, and the market price is determined by the length of the guaranteed delivery lead time. More specifically, the first part of this chapter presents an analytical approach for a make-to-order firm to maximise its profit by optimal selection of a guaranteed delivery lead time. Mean demand is modelled as a decreasing function of price and guaranteed delivery lead time while price itself is a decreasing function of the guaranteed delivery lead time. The model takes into account that - i) reducing lead time by increasing capacity will require investment, and ii) the company must be able to satisfy the guaranteed delivery lead time according to a pre-specified reliability level. In the second part of this chapter we expand our initial model to incorporate the economies of scale by assuming that higher demand can reduce unit operating costs.

There are several papers that assume that the mean arrival rate to any service/manufacturing facility depends on guaranteed delivery lead time and/or price, i.e., demand increases when the price and/or the length of the guaranteed delivery lead time decreases. In one of the seminal papers in this area, So and Song (1998) model the firm as a queuing system where the mean customer demand has a log-linear relationship with price and guaranteed delivery lead time. The objective is to maximise the profit per unit time by suitable selection of the decision variable values - length of the guaranteed delivery lead time, price and capacity. The revenue function is the product of mean demand per unit time and profit per unit, while cost consists of capacity cost. The constraints are in the form of delivery reliability, non-negativity and queue stability. Some analytical comparative statics of the parameters are performed for the optimal decision variables. With the help of numerical examples they are also able to show that direct operating and capacity costs will have a significant effect on the optimal decision variable values.

In Palaka, Erlebacher and Kropp (1998) the mean demand rate decreases linearly with price and delivery lead time. Their objective is to maximise profit per unit time. While the capacity costs and the decision variables for Palaka et al. are similar to that of So and

Song (1998), Palaka et al. also explicitly take into consideration WIP costs and penalty costs. The constraints are delivery reliability, non-negativity and queue stability. Their analysis suggests that there is a critical service level that affects the problem solution.

So (2000) extended So and Song's work by analysing the impact of using delivery lead time guarantees in the presence of competition. While the basic setting remains the same, the focus is on investigating how firms select the best price and guaranteed delivery lead time in the presence of multiple-firm competition. The first part of the paper analyses the optimisation problem and its solution in a multiple-firm setting. In the latter part, the author illustrates how different firm and market characteristics would affect the optimal strategies.

Rao et al. (2000b) integrate a uniform delivery lead time guarantee strategy with production planning for a make-to-order firm. Demand depends on delivery lead time, but price is an exogenous parameter. Unlike other research, Rao et al. do not use WIP holding costs or a delivery reliability constraint. Though the firm has an in-house capacity restriction, it can buy from an infinite capacity supplier, albeit at a higher per unit outsourcing cost than in-house production cost. The discrete production schedule for the firm is synchronised with the guaranteed delivery lead time and the firm optimises on the delivery lead time to maximise the long-run average expected profit per period. They also provide some analytical comparative statics with respect to outsourcing cost, selling price and production capacity.

So and Song (1998), Palaka et al. (1998), So (2000) and Rao et al. (2000b) assume demand per unit time to be dependent on price and/or guaranteed delivery lead time. However, they do not consider the relationship between price and guaranteed delivery lead time. We extend previous research by explicitly taking into account the fact that customers may be willing to pay a price premium for shorter delivery lead times. The numerical examples in the previous papers also show that operating costs play an important role for firms. However, none of the papers analytically model the effect of demand on operating cost. We include economies of scale by modelling unit operating cost as a decreasing function of the mean demand rate.

3.2 Overview and Assumptions of the Physical System

We consider a supply chain in which a firm that is dealing in a make-to-order product/service announces a uniform guaranteed delivery lead time, L, within which it promises to satisfy each customer order. Orders arrive for processing/service according to a Poisson process with mean rate λ. There is a single server in the facility and the processing times of the orders are independent and exponentially distributed with a mean processing rate of μ. The assumption of exponential service times, though simplistic, makes the problem tractable without significant loss of accuracy (refer to So and Song 1998 and Palaka et al. 1998). Customers are served on a first-come-first-served (FCFS) basis. The mean customer demand rate depends on the price, p, and the stated delivery lead time guarantee, L. We assume that customers prefer shorter delivery lead times and lower prices. We also assume that the firm has performed some market research and is aware of how much of a price premium it can obtain from the market by guaranteeing a shorter delivery lead time. For example, UPS or FedEx perform market research to learn how much of a price premium customers are willing to pay for shorter delivery lead times. Hence, the price, p, is higher for a guarantee of shorter delivery lead time, L. We assume that raw material is available whenever required. If the firm is a service facility, there will be no holding cost for raw material and the customers' waiting cost will be indirectly taken care of by the effect of delivery lead time on demand. For manufacturing facilities we do not model holding costs in Sections 3.4 and 3.5.

The firm has established an internal target delivery lead time reliability level, s^R ($0 \leq s^R \leq 1$), which is the probability that a random customer will have a waiting time of L or less in the facility. We assume the target reliability level to be set by management as an internal performance measure not announced to the customers. As failure to satisfy an arriving customer within L might have an adverse impact on repeat business, s^R will be close to 1. Hence, the occurrences of actual waiting time being greater than guaranteed delivery lead time will be rare. Under this assumption, we will not explicitly account for penalty cost incurred by failing to meet delivery lead time guarantee in the models in Sections 3.4 and 3.5. However, as we will show later, our formulation using the service level constraint is consistent with a setting where a firm has to pay a penalty cost, as long as the penalty is independent of the length of the delay.

The firm can invest in increasing the processing rate, μ, through, for example, hiring extra workers or acquiring improved equipment. Successive investments in increasing μ by the same amount will cost more; thus it is reasonable to assume that the investment function for increasing μ, $M(\mu)$, is increasing and convex. The objective of the firm is to maximise its profit per unit time subject to satisfying the delivery reliability constraint. The entire supply chain system is shown in Figure 3.2.1.

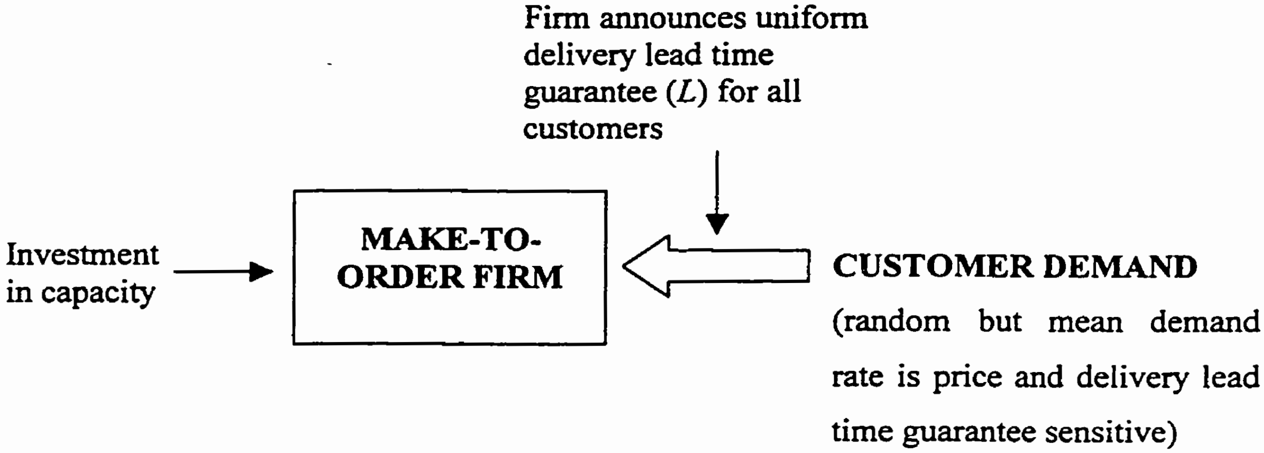


Figure 3.2.1: Supply Chain System for Make-to-Order Firms

3.3 Notation

The following notation will be used for this chapter:

λ = mean demand rate (units/unit time)

μ = mean processing rate (units/unit time)

p = unit market price of the product/service (\$/unit)

m = unit operating cost ($/unit)

L = uniform guaranteed delivery lead time announced to the customers (time)

W = steady state actual waiting time in the facility, i.e., sojourn time (time)

s = actual delivery reliability level, $P(W < L)$

s^R = minimum desired delivery reliability level (i.e., $s \geq s^R$)

$M(\mu)$ = investment cost per unit time to achieve the processing rate μ ($/unit time)

3.4 Analytical Model of the System

We assume that the mean demand rate, λ, depends linearly on L and p, i.e.,

$$\lambda(p, L) = a - b_1 p - b_2 L, \tag{3.4.1}$$

where:

a represents the mean demand rate when both p and L are zero and b_1 and b_2 represent the price and delivery lead time sensitivities of the mean demand rate, respectively (a, b_1, $b_2 > 0$). A higher value of a represents a higher overall potential for demand.

The linear demand function has the desirable properties that price and delivery lead time elasticity of demand are higher at higher prices and guaranteed delivery lead times that even the more popular Cobb-Douglas function does not have (Palaka et al. 1998). These properties are desirable since we would expect that the customers would be more sensitive to long delivery lead times when they are paying more and also sensitive to high prices when they have long waiting times. A linear demand function will also help us to obtain basic qualitative insights without much analytical complexity.

We explicitly model price premiums for shorter delivery lead times by assuming that a guarantee of a shorter delivery lead time can command a higher market price. We assume that the firm has done market research and knows what price premium it will be able to charge for committing to a shorter delivery lead time. Furthermore, we assume that the firm can approximate the relation between p and L by a linear relationship - for a guaranteed delivery lead time of L, the market price, p, will be given by (for a particular demand rate):

$$p = d - e'L, \tag{3.4.2}$$

where:
d = price when $L = 0$, i.e., the maximum price the market is willing to pay,
and
e' = delivery lead time sensitivity of price ($d, e' > 0$).

Combining (3.4.1) and (3.4.2) we can express λ in terms of L as:

$$\lambda = (a - b_1 d) - (b_2 - b_1 e')L = a' - b'L, \tag{3.4.3}$$

where:
$a' = a - b_1 d$,
and
$b' = b_2 - b_1 e'$.

Note that both a' and b' can, in theory, be unrestricted in sign. However, we will assume that $a' > 0$, since otherwise when b' is positive, λ will be negative for all L. As L increases, both λ and p decreases and any decrease in p increases demand.

If $b' > 0$, λ decreases with L - this is the case to which most recent operations management literature refers. This represents the situation where customers are "more lead-time-sensitive than price-sensitive" (i.e., $b_2 > b_1 e'$). Then the decrease of demand rate due to increase in L will be more than the increase of demand rate due to corresponding decrease of p (p decreases since L increases). Some thought shows that $b' < 0$ (i.e., $b_1 e' > b_2$) also makes sense when customers are ready to wait longer to pay a lower price. In this case, λ increases with L. Customers are "more price-sensitive than lead-time-sensitive", i.e., the decrease of demand rate due to increase in L will be less than the increase of demand rate due to corresponding decrease of p (p decreases since L increases). For $b' = 0$, λ is constant (= a') for any L - the customers are "equally sensitive towards price and lead-time", i.e., $b_2 = b_1 e'$ and so the decrease of demand rate due to increase in L will be equal to the increase of demand rate due to corresponding decrease of p.

This type of customer price and lead time sensitivity has been referred to in the literature. Blackburn et al. (1992) pointed out that there are both "price-sensitive" and "time-sensitive" customers in the market. The former segment always chooses a lower price even with longer delivery times while the latter segment is ready to pay a price premium for shorter delivery times. A recent paper by Smith et al. (2000) noted a similar phenomenon in the electronic marketplace. Internet retailers try to create price discrimination among customers using this difference in price and value of time.

Since the firm wishes to maximise expected profit per unit time, π, its goal can be written as:

$$\text{(P3.1)} \quad \underset{\mu, L}{\text{Maximise}}\ \pi(\mu, L) = (p - m)\lambda - M(\mu), \qquad (3.4.4)$$

subject to:

$s = P(W < L) = 1 - e^{-(\mu-\lambda)L} \geq s^R$ (delivery reliability constraint),

$\mu > \lambda$ (system stability constraint),

$p \geq m \geq 0$, $L > 0$, $\lambda \geq 0$ (non-negativity constraints),

where λ is given by (3.4.3), p by (3.4.2) and from our assumption about the investment function, $M(\mu)$ is an increasing convex function in μ. For this section we assume the unit operating cost, m, to be constant; in the next section we will introduce economies of scale.

The form of the delivery reliability constraint is based on the fact that for an $M/M/1$ queue the waiting time has an exponential distribution. Note that at high service levels the waiting time distribution is well approximated by the exponential distribution even for a $G/G/s$ queue (So and Song 1998).

We can now present the following propositions:

Proposition 3.4.1: $\pi(\mu, L)$ in P3.1 is (i) decreasing concave in μ, (ii) decreasing convex in L for $b' \geq 0$ and (iii) concave in L for $b' < 0$.

Proof: Differentiating $\pi(\mu, L)$ with respect to μ we have[1]:

$$\frac{\partial\pi(\mu,L)}{\partial\mu} = -M_\mu; \; \frac{\partial^2\pi(\mu,L)}{\partial\mu^2} = -M_{\mu\mu}. \tag{3.4.5}$$

With the assumption that $M(\mu)$ is increasing convex in μ, it is clear that $\pi(\mu, L)$ is decreasing concave in μ. Differentiating $\pi(\mu, L)$ with respect to L we have:

$$\frac{\partial\pi(\mu,L)}{\partial L} = \lambda_L(p - m) + \lambda p_L; \; \frac{\partial^2\pi(\mu,L)}{\partial L^2} = \lambda_{LL}(p - m) + 2\lambda_L p_L + \lambda p_{LL}. \tag{3.4.6}$$

As both p and λ are linear in L, λ_{LL} and p_{LL} are both equal to zero. Thus, from (3.4.2), (3.4.3) and (3.4.6) we can say that if $b' \geq 0$, then $\pi(\mu, L)$ is decreasing convex in L and if $b' < 0$, then $\pi(\mu, L)$ is concave in L. ■

Proposition 3.4.2: The service level is increasing concave in μ for all b' but increasing concave in L for only $b' = 0$.

Proof: Differentiating the service level, s, with respect to L we have:

$$\frac{\partial s}{\partial\mu} = Le^{-(\mu-\lambda)L}; \; \frac{\partial^2 s}{\partial\mu^2} = -L^2e^{-(\mu-\lambda)L}. \tag{3.4.7}$$

From (3.4.7) we can say that s is increasing concave in μ. Similarly, differentiating s with respect to L and from the stability condition, we can easily prove that s is also increasing concave in L for $b' = 0$. ■

---

[1] For this thesis, Z_Q will represent the first derivative of function Z with respect to Q and Z_{QQ} will represent the second derivative.

Proposition 3.4.1 gives some understanding of the response surface of $\pi(\mu, L)$ in μ and L. From Proposition 3.4.2, it is clear that there is a one-to-one relation between s and μ. As μ approaches λ, s tends towards 0, while as μ approaches ∞, s approaches 1. This is quite intuitive since when μ is very high (for a fixed L) we will be able to satisfy all demand reliably, while when μ is almost equal to λ then there will be heavy congestion and we will not be able to satisfy demand as reliably.

Propositions 3.4.1 and 3.4.2 lead to our third proposition (Palaka et al., 1998 and So and Song, 1998 also obtain a similar result).

Proposition 3.4.3: At optimality, the delivery reliability constraint is binding.
Proof: From the delivery reliability constraint we know that the service level must be at least s^R (i.e., $s \geq s^R$). This implies that the processing rate must be at least the minimum required to achieve the minimum desired service level, (i.e., $\mu \geq \mu^R(L)$ where $1 - e^{-(\mu^R(L)-\lambda)L} = s^R$). We can illustrate the above by Figure 3.4.1 (recall that stability constraint requires $\mu > \lambda$).

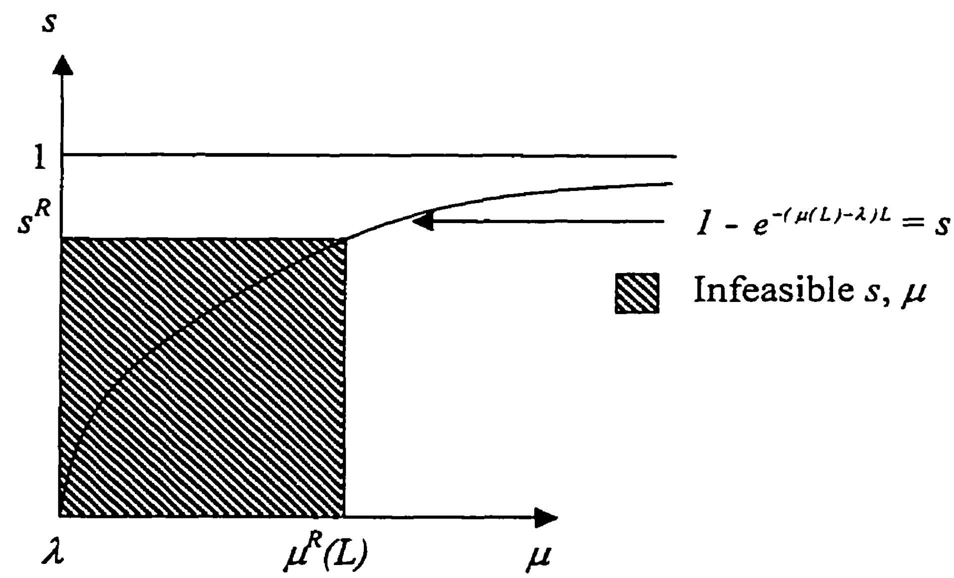


Figure 3.4.1: Plot of Service Level (s) versus Processing Rate (μ)

From Proposition 3.4.1 we know that $\pi(\mu, L)$ is decreasing concave in μ. It is obvious that for a profit maximising firm, for a fixed L, the objective would be to make μ as small as possible while maintaining feasibility. This implies that at optimality, μ should be equal to $\mu^R(L)$ where:

$$\mu^R(L) = \frac{-ln(1-s^R)}{L} + \lambda, \tag{3.4.8}$$

is the processing rate required to achieve the minimum desired service level, s^R. Therefore, at optimality $s = s^R$. ■

The optimal processing rate, μ^*, will therefore lie on the curve $\mu^R(L)$. If we can find the optimal L, L^*, we can substitute it in (3.4.8) to determine the optimal processing rate, μ^*, and in (3.4.2) to determine the optimal price, p^*. Note that $\mu^R(L)$ clearly shows that $\mu^* > \lambda$ (since $[-ln(1-s^R)]$ and $L > 0$), i.e., at optimality, the stability condition will be satisfied. Since the reliability constraint is tight at optimality, to explicitly model a penalty cost for failing to meet the time guarantee, we can just add a term $\chi(1 - s^R)$ to the unit operating cost representing the expected penalty, where χ is the penalty cost for each instance of failure to meet the guarantee independent of the length of the delay (So 2000).

Proposition 3.4.4: $\mu^R(L)$ is a decreasing convex function of L for $b' \geq 0$ and convex for $b' < 0$.

Proof: If we differentiate $\mu^R(L)$ with respect to L we have:

$$\frac{\partial \mu^R}{\partial L} = -\left(\frac{n}{L^2} + b'\right), \ \frac{\partial^2 \mu^R}{\partial L^2} = \left(\frac{2n}{L^3}\right) \geq 0,$$

$$\frac{\partial^3 \mu^R}{\partial L^3} = -\frac{6n}{L^4} \leq 0 \text{ and } \frac{\partial^4 \mu^R}{\partial L^4} = \frac{24n}{L^5} \geq 0, \tag{3.4.9}$$

where $n = [-ln(1-s^R)]$ $(n \geq 0)$.

From (3.4.9) we can tell that if $b' \geq 0$, $\mu^R(L)$ is a decreasing convex function of L (So and Song 1998), and if $b' < 0$, $\mu^R(L)$ is convex in L. ■

Figure 3.4.2 illustrates $\mu^R(L)$ for $b' \geq 0$ and $b' < 0$. Note that for any b', as $L \to 0$, $\mu^R(L) \to \infty$. For $b' \geq 0$, as L increases, $\mu^R(L)$ decreases in a convex, monotone fashion. For $b' < 0$, as L increases, $\mu^R(L)$ initially decreases reaching its minimum at $L = \sqrt{\frac{n}{(-b')}}$ and then increases.

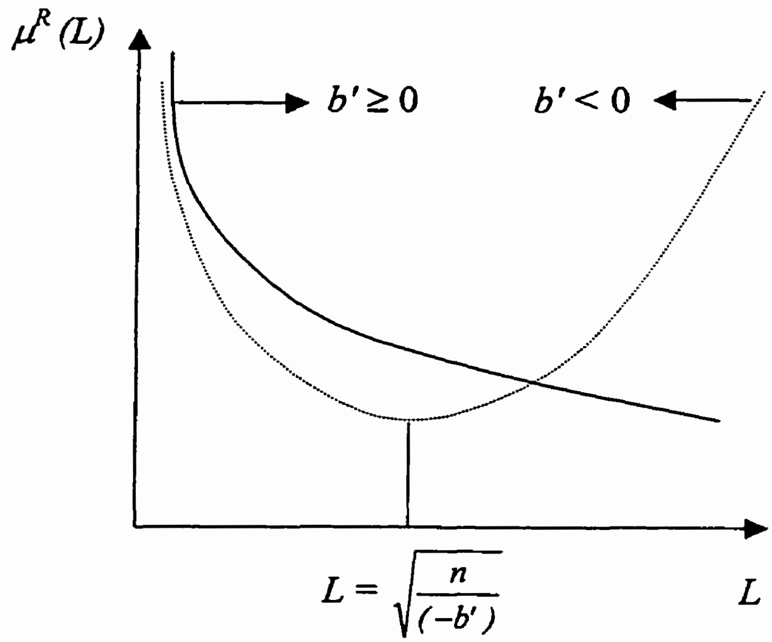


Figure 3.4.2: Plot of $\mu^R(L)$ versus L

Having illustrated $\mu^R(L)$, we can transform the problem P3.1 into a new problem P3.2 in terms of a single variable, L:

$$\text{(P3.2)} \quad \underset{L}{\text{Maximise}}\ \pi(L) = \lambda(p - m) - M(\mu^R(L)), \tag{3.4.10}$$

subject to:

$\lambda \geq 0, L > 0, p \geq m \geq 0,$

where:

$\lambda = a' - b'L, p = d - e'L,$

and

$$\mu^R(L) = \frac{-ln(1-s^R)}{L} + \lambda(L).$$

Now the problem is simply to find L^*, the value of L that maximises $\pi(L)$ for P3.2. For the rest of this section we will suppress the argument of $\pi(L)$ of (3.4.10) unless otherwise stated.

First let us consider the feasible range for L. If $b' \geq 0$, then $\lambda \geq 0$, i.e., $L \leq (a'/b')$. The condition $p \geq m$ implies that $(d - e'L) \geq m$, i.e., $L \leq \dfrac{d-m}{e'}$. Then for $b' \geq 0$, the feasible region for L is (0, min $\left(\dfrac{a'}{b'}, \dfrac{d-m}{e'}\right)$). If $b' < 0$ then the condition $\lambda \geq 0$ will be satisfied for any $L > 0$ and the feasible region for L is (0, $\left(\dfrac{d-m}{e'}\right)$). Note that from (3.4.10) and assumptions about $M(\mu)$ we can conclude that as L tends towards the feasible limits, the profits are negative - infinite at lower limit and finite at the upper limit. Differentiating π with respect to L we have:

$$\pi_L = \lambda_L (p - m) + \lambda p_L - M_\mu(\mu^R(L)) \frac{\partial \mu^R}{\partial L}, \tag{3.4.11}$$

and

$$\pi_{LL} = 2\lambda_L p_L - M_{\mu\mu}(\mu^R(L)) (\frac{\partial \mu^R}{\partial L})^2 - M_\mu(\mu^R(L)) \frac{\partial^2 \mu^R}{\partial L^2}. \tag{3.4.12}$$

<u>**Case 1: $b' \leq 0$**</u>

Proposition 3.4.5: For $b' \leq 0$, π in P3.2 is concave in L.

Proof: If $b' \leq 0$, we know $\lambda_L \geq 0$ and $p_L < 0$ implying that $2\lambda_L\, p_L \leq 0$. Then from (3.4.9) and the assumption that M is increasing convex in μ, (3.4.12) will be negative and hence π in P3.2 is concave in L. ■

Case 2: $b' > 0$

Proposition 3.4.6: For $b' > 0$ and $M_{\mu\mu\mu} \geq 0$, $\pi_L = 0$ can have zero, one or two feasible solutions.

Proof: When $b' > 0$, then $\lambda_L < 0$ and $p_L < 0$, thus $2\lambda_L\, p_L > 0$ and the sign of π_{LL} is unrestricted. Rearranging $\pi_L = 0$ from (3.4.11), we have:

$$\lambda_L(p - m) + \lambda p_L = M_\mu(\mu^R(L))\, \frac{\partial \mu^R}{\partial L}. \tag{3.4.13}$$

Differentiating both sides of (3.4.13) with respect to L we have:

$$\frac{\partial}{\partial L}(LHS) = 2\lambda_L p_L = 2b'e';\ \frac{\partial^2}{\partial L^2}(LHS) = 0. \tag{3.4.14}$$

$$\frac{\partial}{\partial L}(RHS) = M_{\mu\mu}(\mu^R(L))\,(\frac{\partial \mu^R}{\partial L})^2 + M_\mu(\mu^R(L))\,\frac{\partial^2 \mu^R}{\partial L^2}; \tag{3.4.15}$$

$$\frac{\partial^2}{\partial L^2}(RHS) = M_{\mu\mu\mu}(\mu^R(L))\,(\frac{\partial \mu^R}{\partial L})^3 + 3\,M_{\mu\mu}(\mu^R(L))\,(\frac{\partial \mu^R}{\partial L})\frac{\partial^2 \mu^R}{\partial L^2} + M_\mu(\mu^R(L))\frac{\partial^3 \mu^R}{\partial L^3}. \tag{3.4.16}$$

Taking into account the constraints of P3.2, it is easy to see that the LHS of (3.4.13) is always negative and linearly increasing. It will start from a finite negative value and increase linearly to a finite negative value at the upper feasible limit of L. If we assume that $M_{\mu\mu\mu} \geq 0$ (i.e., not only do successive investments cost more, the rate also increases as μ increases) and since M_μ and $M_{\mu\mu}$ are positive, from (3.4.9) we can deduce that (3.4.16) is negative for $b' > 0$. The RHS of (3.4.13) will always be negative for feasible L, will tend to - ∞ as L tend towards 0 and increase in a concave manner to a finite negative value at the upper feasible limit of L. With the above forms of LHS and RHS of (3.4.13), we can convince ourselves that there can be either zero, one or two feasible solutions to $\pi_L = 0$ (refer to Figure 3.4.3). ■

From these observations, we can conclude that when demand is decreasing in L, π may not be unimodal for feasible L. Let us now consider the three possible outcomes:

a) No solution to $\pi_L = 0$: In this case the RHS is always below the LHS for feasible L and hence π will always be increasing (Figure 3.4.3(i));

b) One solution to $\pi_L = 0$: In this case there might be two situations. It might be that RHS < LHS for small L and as L increases the RHS will intersect the LHS from below. In this case π will be increasing up to the solution for $\pi_L = 0$ and then decreasing (Figure 3.4.3(ii)). It might also be that while RHS < LHS for small L, as L increases the RHS rather than intersecting is tangent to LHS from below. In this case also the RHS is always below the LHS for feasible L and hence π will always be increasing for feasible L and the solution to $\pi_L = 0$ will be the inflection point;

c) Two solutions to $\pi_L = 0$: In this case the RHS will first intersect the LHS from below and then it will intersect from above. The profit, π, is initially increasing, then decreasing and then again increasing in L (Figure 3.4.3(iii)).

Note the only thing we can say about π for $b' > 0$ is the range where π will be increasing or decreasing and not whether it will be concave or convex.

What can we then say about L^* for Case 1 and Case 2? First, since π_L is increasing as $L \to 0$ (the lower feasible limit of L), L^* will be given by either the feasible solution(s) to $\pi_L = 0$ or the upper limit for the feasible limit of L. As there are a finite number of possible alternatives we can easily compare the profit at those alternatives to find L^*. We can reduce the possible alternatives further since from Figure 3.4.3(iii) it is clear that for multiple feasible solutions to $\pi_L = 0$ for $b' > 0$, the larger solution can never be L^*. So, it is relatively simple to determine L^*.

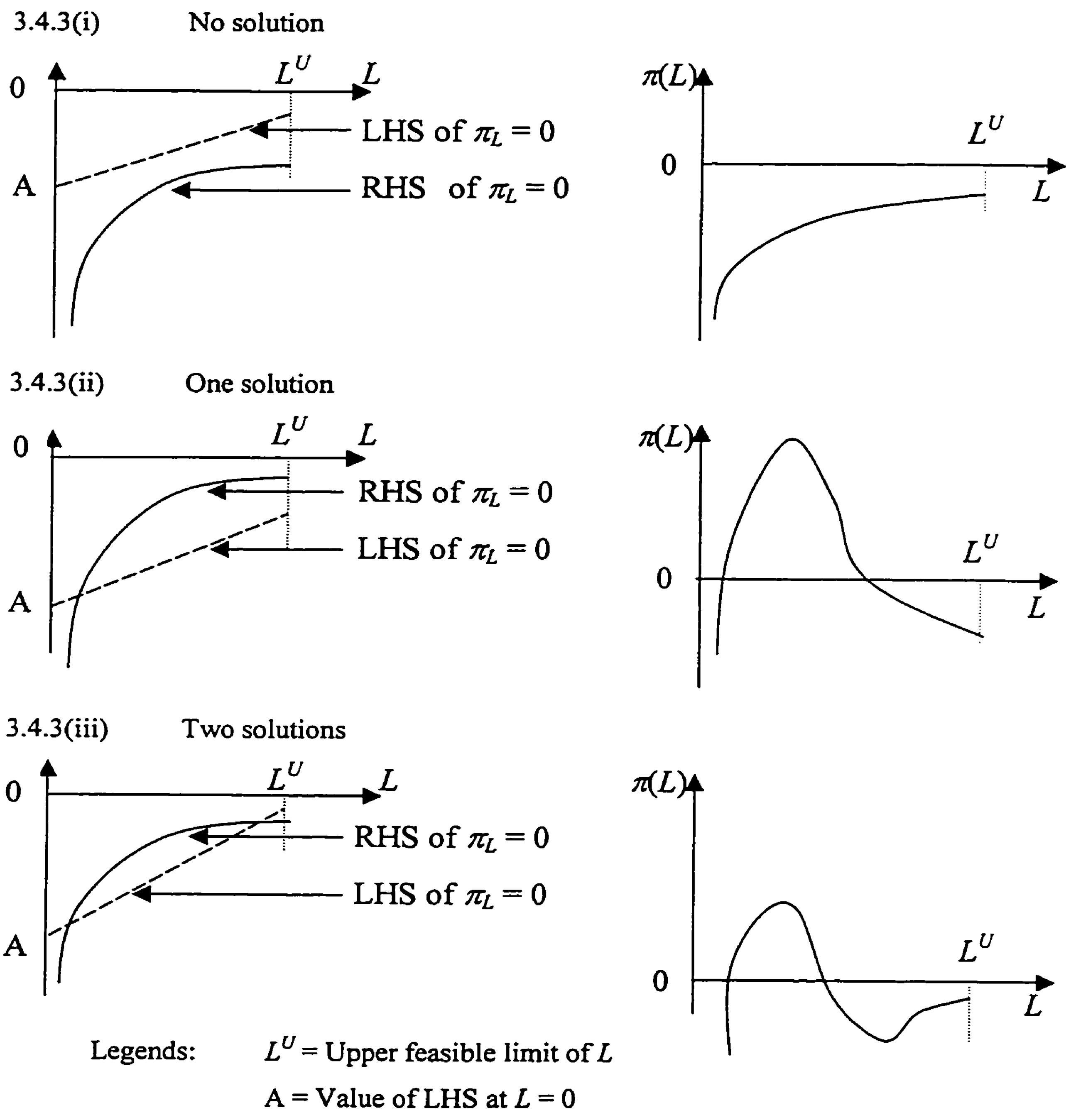


Figure 3.4.3: Three Possible Forms of $\pi(L)$ for $b' > 0$

It is not necessary that $\pi(L^*)$ will be positive. If $\pi(L^*)$ is negative (e.g., if L^* = upper feasible limit of L) it implies that the firm cannot make a profit even when acting optimally and hence should not be in the business at all. If there is no feasible solution to $\pi_L = 0$ (Figure 3.4.3(i)), the objective function will then be negative (increasing) for the entire feasible range of L. If, however, $\pi(L^*) > 0$, the firm can announce a guaranteed delivery lead time of L^* and set its processing rate by substituting L^* into (3.4.8) to satisfy the service level. The market price (p^*) will be determined by (3.4.2) and the mean demand rate will then be given by (3.4.3).

3.4.1 Numerical Examples

Example 3.4.1: First let us consider an example with the following parameters:
$a = 100$, $b_1 = 5$, $b_2 = 8$, $d = 10$, $e' = 1$, $m = 1$, $M(\mu) = A\mu^3$ where $A = 0.0001$ and $s^R = 0.98$. We have $a' = a - b_1 d = 50$, $b' = b_2 - b_1 e' = 3$, $\lambda = 50 - 3L$ and $p - m = 9 - L$.

In this example, customers are more lead-time-sensitive than price-sensitive (since $b' > 0$), $L \in (0, 9)$, $\lambda \in (23, 50)$ (λ is decreasing in L) and $p \in (1, 10)$. The solution for the constrained problem P3.2 in this example is given in Table 3.4.1. Note that π is unimodal but not concave (similar to Figure 3.4.3(ii)).

Different parameter values can give rise to cases similar to Figures 3.4.3(i) and 3.4.3(iii).

Example 3.4.2: Suppose $b_2 = 2$ in Example 3.4.1, so that $b' = -3$ and $\lambda = 50 + 3L$. While the feasible region of L and the range of p remain the same as Example 3.4.1, in this example $\lambda \in (50, 87)$ and is increasing in L - customers are more price-sensitive than lead-time-sensitive. Since $b' < 0$, π is known to be concave in L.

Example 3.4.3: In this example, $a = 73$ and $b_2 = 2$ while all the other parameters are the same as in Example 3.4.1. Then $a' = 23$ and $b' = -3$ and $\lambda = 23 + 3L$. For this example, customers are more price-sensitive than lead-time-sensitive and λ is increasing in L. Note that for this example the feasible ranges of L, p and λ are all similar to Example 3.4.1. Since $b' < 0$, π is known to be concave in L.

We compare the values of the optimal guaranteed delivery lead time, processing rate, price, demand and profit for the above examples in Table 3.4.1.

Table 3.4.1: Comparison of L^*, μ^*, p^*, λ^* and π^* for Examples 3.4.1 - 3.4.3

	b'	L^*	$\mu(L^*)$	$p(L^*)$	$\lambda(L^*)$	$\pi(L^*)$
Example 3.4.1	3	0.26	64.27	9.74	49.22	403.64
Example 3.4.2	-3	0.39	61.18	9.61	51.17	417.65
Example 3.4.3	-3	0.81	30.27	9.19	25.42	205.50

These examples show that the sign of b' has a significant impact on the ptimal decision variable values. Recall that for $b' < 0$, customers are ready to wait longer if they can pay a lower price. Comparison of Examples 3.4.1 and 3.4.2 show the effect of the change in sign of b' on the optimal decision variable values, demand and profit. For $b' < 0$, L^* is larger and p^* and μ^* are smaller than for $b' > 0$. Increased revenue from higher demand and reduced investment cost more than offset the loss of revenue from lower price.

Comparison of Examples 3.4.1 and 3.4.3 shows that even when the ranges of p and λ are similar, a negative b' forces L^* to be higher and μ^* to be lower (leading to less investment) so that p^* can be reduced to cater to more price-sensitive customers. Comparison of Examples 3.4.2 and 3.4.3 shows that lower a leads to higher L^* and lower π^*. While the comparative statics of So & Song (1998) and Palaka et al. (1998) show a similar phenomenon, our model not only takes into account the price and lead-time sensitivity of demand rate but also explicitly accounts for the relationship between price and the length of the guaranteed delivery lead time.

3.4.2 Effect of Treating p as a Decision Variable

We have chosen to explicitly model the relationship between p and L rather than to assume that p and L are independent decision variables. While this reduces the number of decision variables, it captures for managers a relationship which exists in practice, and which could lead to a decision error if ignored. In this section, our aim is to see the effect on profit if managers ignore the dependence of price on the length of the delivery lead time guarantee and assume p to be a decision variable, independent of L.

With both p and L as decision variables, $\lambda(p, L) = a - b_1 p - b_2 L$ $(b_1, b_2 > 0)$ and the optimal μ depends on both p and L, i.e., $\mu^R(p, L) = \dfrac{-\ln(1-s^R)}{L} + \lambda(p,L)$. The problem is now to maximise $\pi(p, L) = \lambda(p, L)(p - m) - M(\mu^R(p, L))$.

Example 3.4.4: Let us consider an example with the following parameters:
$a = 73$, $b_1 = 5$, $b_2 = 2$, $m = 1$, $d = 12$, $e' = 0.9$, $M(\mu) = A\mu^3$ where $A = 0.0001$ and $s^R = 0.98$.

With these parameters, the solution with p as a decision variable will be:
$L_p^* = 0.40$, $p^* = 7.98$ and $\mu(p^*, L_p^*) = 42.03$.

If we use our model (p dependent on L), the solution will be:
$L^* = 3.44$, $\mu(L^*) = 22.73$, $p(L^*) = 8.91$, $\lambda(L^*) = 21.59$ and $\pi(L^*) = 169.55$.

The differences in the optimal values are expected since the models are dissimilar. Table 3.4.2 shows how large the effect on profit will be if in reality p is related to L, as in our model, but the manager ignores it and determines L^* assuming p to be an independent decision variable (i.e., guarantees $L = 0.4$ rather than 3.44 in our model).

Table 3.4.2: Comparison of μ^*, p^*, λ^* and π^* for Different L for Example 3.4.4

	$\mu^*(L)$	$p^*(L)$	$\lambda^*(L)$	$\pi^*(L)$
Model assuming p and L to be related as $p = d - eL$ ($L^* = 3.44$)	22.73	8.91	21.59	169.55
Model assuming p and L to be decision variables ($L = 0.40$)	23.76	11.64	14.00	147.63

In this example, customers are more price-sensitive than lead-time-sensitive. However, if we ignore the relation between p and L when it actually exists, we will be guaranteeing a much shorter delivery lead time (L = 0.4) compared to the optimal (L = 3.44). Though the price will be higher the demand will be much less. Also, we have to invest more to satisfy the small L with the desired reliability. The net result is that profit will be about 13% lower! We can also show examples when the opposite will happen, i.e., the firm will be offering a "higher" L for more lead-time-sensitive than price-sensitive customers, and will be loosing profit.

There will be cases when assuming p and L to be independent decision variables will not have much impact on profits. For example, in Example 3.4.1 assuming p to be a decision variable would have given L_p^* as 0.25 and profit would be quite close to the optimal profit with L^* = 0.26. However, in general, it is clear that when p and L are related we should model the relationship explicitly.

We performed numerical experiments to observe the effect of s^R on L. For both $b' \geq 0$ and $b' < 0$, L^* is increasing convex in s^R. Note that this effect is similar to that of service level on inventory cost implying that the length of the delivery lead time guarantee can be thought of as an inventory of time and more "time inventory" is required as the service level increases.

In this section, we were able to develop a model for a profit-maximising make-to-order firm whose demand rate depends on price and a uniform guaranteed delivery lead time while price itself is determined by the length of the guaranteed delivery lead time. The firm has to satisfy some delivery reliability constraint but can invest in reducing its processing time. The resulting problem is different from previous literature as it explicitly accounts for the relationship between price and delivery lead time. This model can capture two distinct consumer preferences: i) where the customers are willing to pay more for faster delivery, and ii) where the customers are ready to wait longer to pay less. By using a reliability constraint and a relation between price and delivery lead time, we were able to express the problem in terms of a single variable, guaranteed delivery lead time.

We were able to show that the form of the solution will be rather simple, and through some numerical examples that it is important that the customer preferences are taken into account while deciding upon the optimal policy for the firm. The optimal policies (including investment decisions) for firms whose customers are more sensitive towards price than delivery lead time will be quite different from firms whose customers wants shorter delivery lead time. We also showed how our model gives rise to different decisions than models that assume price and delivery lead time to be independent decision variables. Ignoring the relation between price and the guaranteed delivery lead time can lead to investing in lead time reduction to guarantee a shorter delivery lead time when customers want lower prices and are willing to wait longer or not providing short enough delivery lead times when the market is willing to pay a price premium for shorter delivery lead times. This is consistent with the empirical findings of Sterling and Lambert (1989) that management often subjectively sets customer service levels that are not consistent with customer preferences, not realising that customers might have different needs than the seller.

3.5 Analytical Model Incorporating Economies of Scale

Companies may be able to achieve economies of scale by spreading fixed costs over a large production volume. For such operations, it is reasonable to assume that the unit operating costs is a decreasing function of the demand rate, at least within a certain volume range (Scherer 1980). Such economies of scale are present in almost all types of firms - manufacturing or service. Numerical examples of So and Song (1998), Palaka et al. (1998) and So (2000) show that operating costs may have a significant impact on the optimal operating characteristics of a firm. In this section, we analytically explore the implications of scale economies on the basic model introduced in Section 3.4.

While the exact nature of the scale economies will depend on many factors, here we explore the case when the unit operating cost, $m = u\lambda^{(-v)}$, is decreasing convex with respect to the mean demand rate. In the relation, v (> 0) is the sensitivity of unit operating costs with respect to the mean demand rate and u is a finite constant. We assume that the economies of scale can come from any number of sources. The relationship between λ and L as well as between p and L and the characteristics of the investment function, M, remain the same as in Section 3.4.

The problem of maximising expected profit per unit time can now be written as:

$$\text{(P3.3)} \quad \underset{\mu, L}{\text{Maximise}}\ \pi(\mu, L) = (p - m)\lambda - M(\mu), \tag{3.5.1}$$

subject to:

$P(W < L) \geq s^R$, i.e., $1 - e^{-(\mu-\lambda)L} \geq s^R$ (delivery reliability constraint),

$\mu > \lambda$ (system stability constraint),

$p \geq m \geq 0,\ L > 0,\ \lambda \geq 0$ (non-negativity constraints).

In the above expression, $p = d - e'L$, $\lambda = a' - b'L$ and hence, $m = u(a' - b'L)^{(-v)}$. From the definition, m is decreasing convex with respect to λ. When $b' < 0$ (demand increases with increase in L), λ_L is positive and m is decreasing convex in L and when $b' > 0$ (demand decreases with increase in L), λ_L is negative and m is increasing convex in L.

Proposition 3.5.1: $\pi(\mu, L)$ of P3.3 is (i) decreasing concave with respect to μ, and (ii) decreasing with respect to L for $b' \geq 0$.

Proof: Part (i) is similar to the proof for Proposition 3.4.1. Differentiating (3.5.1) with respect to L we have:

$$\pi_L(\mu, L) = \lambda_L(p - m) + \lambda p_L - \lambda m_L;\ \pi_{LL}(\mu, L) = 2\lambda_L p_L - 2\lambda_L m_L - \lambda m_{LL}. \qquad (3.5.2)$$

For $b' \geq 0$, $\pi_L \leq 0$ and thus $\pi(\mu, L)$ is decreasing with respect to L. While for $b' = 0$, $\pi(\mu, L)$ is linear in L, for $b' > 0$, it can be either convex or concave. ■

Note that for $b' < 0$, we cannot tell whether π is increasing/decreasing or convex/concave.

The reasoning in Section 3.4 showing that the optimal μ will be along $\mu^R(L)$ given by (3.4.8), is still valid as the expression for $\mu^R(L)$ is independent of m and so remains unchanged. We can now express problem P3.3 in terms of the single decision variable L as:

$$\text{(P3.4)}\quad \underset{L}{\text{Maximise}}\ \pi(L) = (p - m)\lambda - M(\mu^R(L)), \qquad (3.5.3)$$

subject to:

$p \geq m \geq 0,\ L,\ \lambda \geq 0$ (non-negativity constraints).

In (3.5.3), $p = d - e'L$, $\lambda = a' - b'L$, $m = u(a' - b'L)^{(-v)}$, $\mu^R(L) = (n/L) + \lambda$ and $n = -\ln(1 - s^R) \geq 0$. The feasible region for L will be given by the constraints $p \geq m \geq 0$, $L > 0$ and $\lambda \geq 0$. As in Section 3.4, as $L \to 0$, $\pi(L) \to -\infty$ but can be positive or negative at the feasible limits of L. Our objective is to find L^* that maximises $\pi(L)$ in P3.4. The value of L^* will determine the optimal values of the other decision variables and also the optimal profit. For the rest of this section we suppress L as the argument of $\pi(L)$ in (3.5.3) unless otherwise stated.

Differentiating (3.5.3) yields (recall that λ_{LL} and $p_{LL} = 0$):

$$\pi_L \quad = \lambda_L(p - m) + \lambda p_L - \lambda m_L - M_L(\mu^R(L))$$

$$= (-b')(p - m) - e'\lambda + (-b')uv\lambda^{(-v)} - M_\mu(\mu^R(L))\frac{\partial \mu^R}{\partial L}, \tag{3.5.4}$$

and

$$\pi_{LL} = 2\lambda_L p_L - 2\lambda_L m_L - \lambda m_{LL} - M_{\mu\mu}(\mu^R(L))\left(\frac{\partial \mu^R}{\partial L}\right)^2 - M_\mu(\mu^R(L))\frac{\partial^2 \mu^R}{\partial L^2}. \tag{3.5.5}$$

We follow the same method as we did in Section 3.4 to find L^*.

On rearranging the terms of $\pi_L = 0$ we have:

$$\lambda_L(p - m) + \lambda p_L - \lambda m_L = M_\mu(\mu^R(L))\frac{\partial \mu^R}{\partial L}. \tag{3.5.6}$$

Differentiating both sides of (3.5.6) with respect to L we have:

$$\frac{\partial}{\partial L}(LHS) = 2\lambda_L p_L + (uv)(\lambda_L)^2(\lambda^{-(v+1)})(1 - v); \tag{3.5.7}$$

$$\frac{\partial^2}{\partial L^2}(LHS) = -(uv)(1 + v)(1 - v)(\lambda^{-(v+2)})(\lambda_L)^3. \tag{3.5.8}$$

$$\frac{\partial}{\partial L}(RHS) = M_{\mu\mu}(\mu^R)\left(\frac{\partial \mu^R(L)}{\partial L}\right)^2 + M_\mu(\mu^R)\left(\frac{\partial^2 \mu^R(L)}{\partial L^2}\right). \tag{3.5.9}$$

From (3.4.9) and characteristics of M, the RHS of (3.5.6) is increasing in L for any b'. To solve the constrained problem P3.4, we will have to consider three cases: i) $b' = 0$, ii) $b' < 0$ and iii) $b' > 0$.

Case 1: $b' = 0$

In this case $\lambda_L = 0$, $p_L < 0$, $m_L = 0$ and $m_{LL} = 0$.

Proposition 3.5.2: For $b' = 0$, π is concave in L.

Proof: From (3.4.9) and remembering that $\lambda_L = 0, p_L < 0, m_L = 0$ and $m_{LL} = 0$ when $b' = 0$, it is clear that (3.5.5) is negative and π is concave in L for $b' = 0$. ■

<u>**Case 2: $b' < 0$**</u>

In this case $\lambda_L > 0, p_L < 0, m_L < 0$ and $m_{LL} > 0$.

Proposition 3.5.3: For $b' < 0$, $M_{\mu\mu\mu} \geq 0$ and $M_{\mu\mu\mu\mu} \geq 0$, the RHS of (3.5.6) will be increasing concave up to the unique solution for $\frac{\partial^2(RHS)}{\partial L^2} = 0$ (say, L^s) and then it will be increasing convex and the value of the RHS at L^s will be positive.

Proof: Refer to Appendix 3.1. ■

Example 3.5.1: For better understanding, let us consider a numerical example with the following parameters:

$a = 73$, $b_1 = 5$, $b_2 = 2$, $d = 10$, $e' = 1$ ($\Rightarrow b' = -3$), $A = 0.001$ and $s^R = 0.98$. The feasible range of L is (0, 9.54). In Figure 3.5.1, we show the RHS for $\pi_L = 0$ in (3.5.6).

The RHS is increasing throughout the feasible region of L, is negative till $L = 1.14$ $\left(=\sqrt{\frac{n}{(-b')}}\right)$ and then positive, and concave till $L = 4.07$ (=L^s, unique solution of $\frac{\partial^2(RHS)}{\partial L^2} = 0$, refer to Appendix 3.1) and then convex.

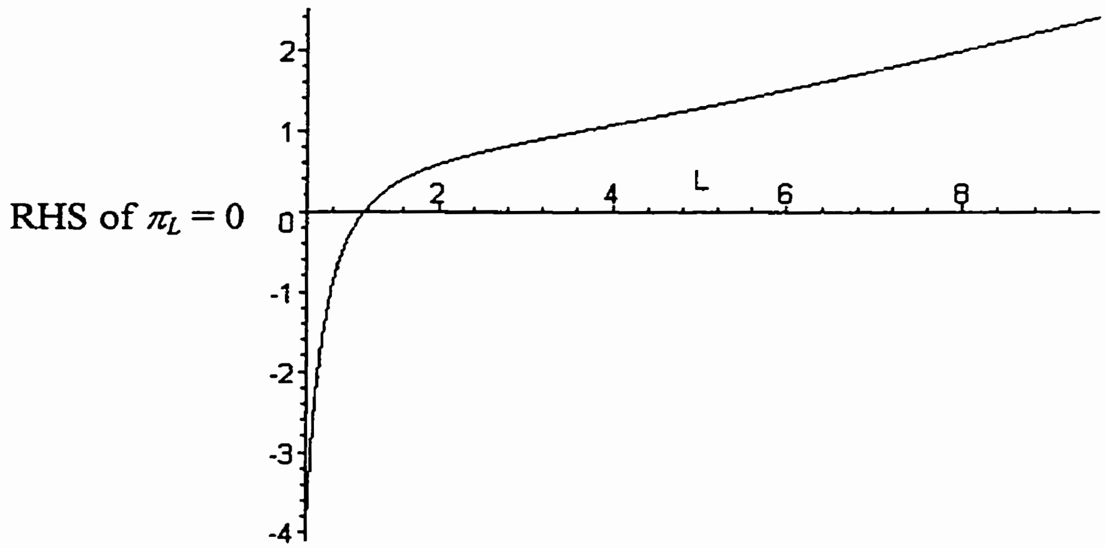


Figure 3.5.1: Plot of RHS of $\pi_L = 0$ versus L for $b' < 0$

Now we can develop a proposition regarding the structure of π_L.

Proposition 3.5.4:, If $v \geq 1$, $M_{\mu\mu\mu} \geq 0$ and $M_{\mu\mu\mu\mu} \geq 0$, for $b' < 0$, $\pi_L = 0$ will have zero or one feasible solution.

Proof: The LHS of (3.5.6) in this case will be unrestricted in sign, finite as $L \to 0$ and decreasing convex in L. The RHS of (3.5.6) will be always increasing, negative for $L < \sqrt{\dfrac{n}{(-b')}}$ (RHS $\to -\infty$ as $L \to 0$), positive for $L \geq \sqrt{\dfrac{n}{(-b')}}$, concave for $L \leq L^s$ and convex for $L > L^s$. Hence, we can convince ourselves that there can be zero or one feasible solution to $\pi_L = 0$. ■

Case 3: $b' > 0$

In this case $\lambda_L < 0$, $p_L < 0$, $m_L > 0$ and $m_{LL} > 0$.

Proposition 3.5.5: If $\nu \geq 1$ and $M_{\mu\mu\mu} \geq 0$, for $b' > 0$, $\pi_L = 0$ can have zero, one, two or three feasible solutions.

Proof: The LHS of (3.5.6) in this case will be always negative and concave in L (from (3.5.7) and (3.5.8)). The RHS of (3.5.6) will be always negative, increasing and concave (from 3.4.15 and 3.4.16). As $L \to 0$, the RHS $\to -\infty$ but the LHS is finite and negative. We can now convince ourselves that $\pi_L = 0$ can have zero, one, two or three feasible solutions (refer to Figure 3.5.2). ■

Since the lower feasible limit of L for this section is not necessarily zero, the possible cases are now more complex than in Section 3.4. However if we assume the lower feasible limit to be equal to zero then the four possible outcomes of Proposition 3.5.5 will look like Figure 3.5.2. As $L \to 0$, $\pi \to -\infty$ and $\pi_L \to +\infty$, i.e., π is increasing at an infinite rate. At the upper limit of L, both π and π_L can be positive or negative.

a) No solution to $\pi_L = 0$: In this case the RHS is always below the LHS for feasible L and hence π will always be increasing (Figure 3.5.2(i));

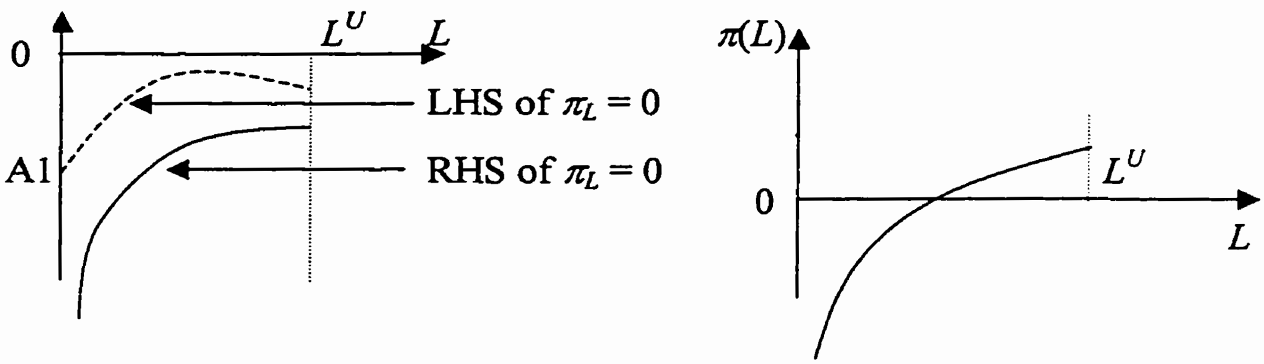
b) One solution to $\pi_L = 0$: In this case, π will be increasing up to the solution for $\pi_L = 0$ and then decreasing (Figure 3.5.2(ii));

c) Two solutions to $\pi_L = 0$: In this case, the RHS will first intersect the LHS from below and then it will intersect from above. The profit function, π, is initially increasing, then decreasing and then again increasing in L (Figure 3.5.2(iii));

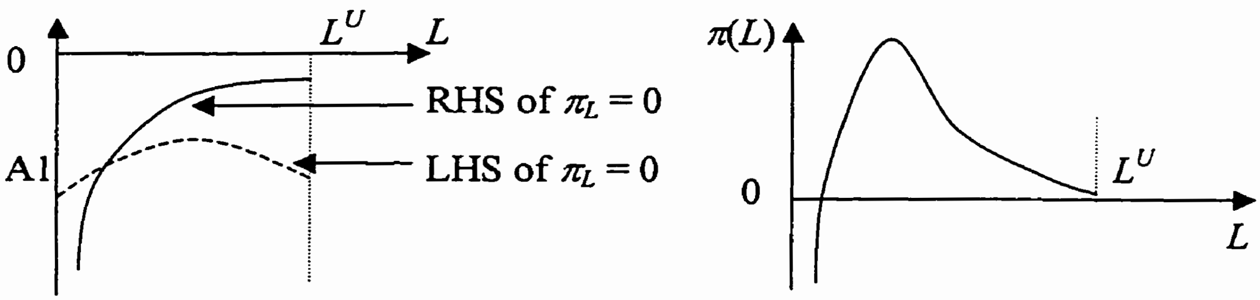
d) Three solutions to $\pi_L = 0$: The nature of intersections and the resultant structure of π are shown in Figure 3.5.2(iv).

What can we now say about L^* for Cases 1, 2 and 3? When the lower limit of feasible L is not zero, then L^* will be given by either the feasible solution(s) to $\pi_L = 0$ or one of the limits for feasible range of L. Since there are a finite number of possible alternatives we can compare the profit at those alternatives to obtain L^*. As in Section 3.4, there may be cases when the RHS and the LHS may not intersect, but one might be tangent to another. It is quite straightforward to find the shape of π in those cases also. We can also reduce the possible alternatives further by reasoning (e.g., for $b' = 0$ or $b' < 0$, if there is a feasible solution to $\pi_L = 0$, then it will be L^*; for $b' > 0$, if there are three feasible solutions to $\pi_L = 0$, the middle solution can never be L^*). So, it is again relatively simple to determine L^*.

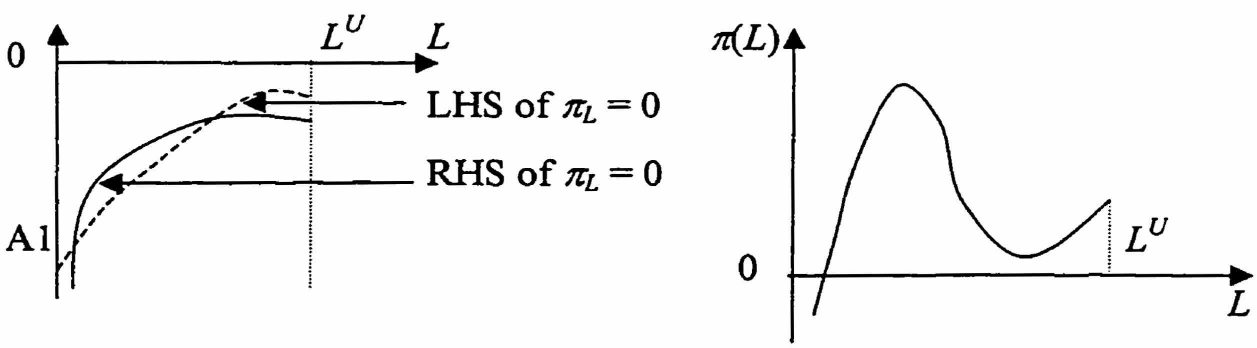
3.5.2(i) No solution



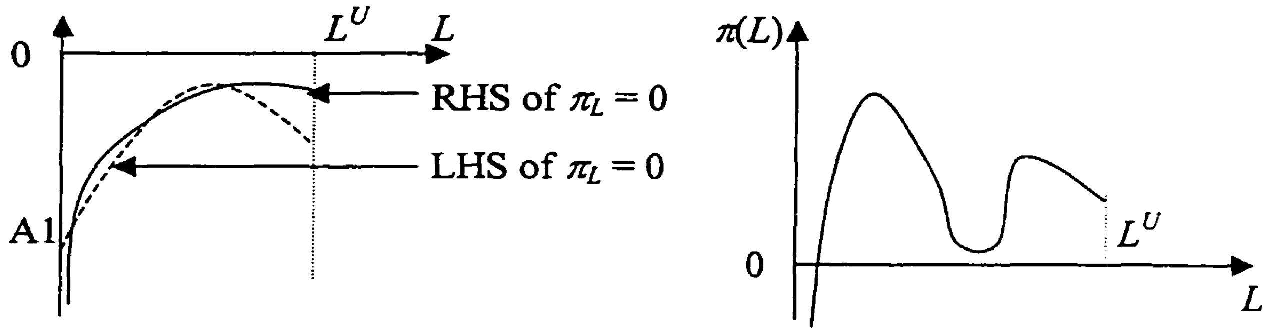
3.5.2(ii) One solution



3.5.2(iii) Two solutions



3.5.2(iv) Three solutions



Legends: L^U = Upper feasible limit of L

A1 = Value of LHS at $L = 0$

Figure 3.5.2: Four Possible Forms of $\pi(L)$ for $b' > 0$

As in Section 3.4, $\pi(L^*)$ may not be positive. If $\pi(L^*)$ is negative, it implies that the firm cannot make a profit even when acting optimally and hence should not be in the business at all. If $\pi(L^*) > 0$, then the firm can announce a guaranteed delivery lead time of L^* and set its processing rate by substituting L^* into (3.4.8) to satisfy the service level. The market price (p^*) will be determined by (3.4.2) and this combination of p^* and L^* will give the mean demand rate from (3.4.3). The mean demand rate will induce the operating cost, $m(L^*)$ ($=u\lambda^{*(-v)}$), and the firm's profit will be maximised.

3.5.1 Numerical Examples

The following examples show how our decision-making will be affected when we consider economies of scale.

Example 3.5.2: Let us assume the same parameters as in Example 3.4.1:
$a = 100$, $b_1 = 5$, $b_2 = 8$, $d = 10$, $e' = 1$, $M(\mu) = A\mu^3$ where $A = 0.0001$ and $s^R = 0.98$. Also, $u = 185$, $v = 1.5$, so that $m = 185(50 - 3L)^{-1.5}$.

Here $a' = 50$, $b' = 3$ (i.e., customers are more lead-time-sensitive than price-sensitive) and $\lambda = 50 - 3L$. For this example, $L \in (0, 9.54)$, $\lambda \in (24.56, 50)$ and $m \in (0.52, 1.52)$. In Table 3.5.1 we compare the optimal values of L^*, μ^*, p^*, m^*, λ^* and π^* for Example 3.4.1 (without economies of scale, i.e., $m = 1$) to that of Example 3.5.2 (with economies of scale).

Table 3.5.1: Comparison of L^*, μ^*, p^*, m^*, λ^* and π^*

for Examples 3.4.1 and 3.5.2

	L^*	$\mu(L^*)$	$p(L^*)$	$m(L^*)$	$\lambda(L^*)$	$\pi(L^*)$
Example 3.4.1	0.26	64.27	9.74	1	49.22	403.64
Example 3.5.2	0.25	64.60	9.75	0.535	49.25	426.50

Note that while π is unimodal within the feasible range of L, it is not concave for Example 3.5.2. Though we used the same parameter values, L^* is lower when economies of scale exist. Since for this example, $b' > 0$, lower L^* leads to higher μ^* and p^*. This is intuitive; we have more incentive to guarantee a shorter delivery lead time so as to attract more customers and thereby decrease m. A guarantee of a shorter delivery lead time will also command a higher price. However, we have to invest in increasing the processing rate so that we can satisfy the reliability constraint even at the higher demand level.

Similarly we can easily show examples that firms having different economies of scale (i.e, only the value of parameter v is different) will guarantee different L^* to maximise their profits. The direction of change in L^* as v changes will depend on the sign of b', i.e., whether customers are more price-sensitive or lead-time-sensitive. Though there is not much difference in the L^* for Examples 3.4.1 and 3.5.2, this is not always the case, as we show in the next example.

Example 3.5.3: Let us consider an example similar to Example 3.4.4 with $b' < 0$ (i.e., customers are more price-sensitive than lead-time-sensitive). The parameters are: $a = 73$, $b_1 = 5$, $b_2 = 2$, $d = 12$, $e' = 0.9$, $u = 175$, $v = 1.45$, $M(\mu) = A\mu^3$ where $A = 0.0001$ and $s^R = 0.98$. Then $\lambda = 13 + 2.5L$ and $m = 175(13 + 2.5L)^{-1.45}$.

We already solved this problem assuming that no economies of scale exist (i.e., $m = 1$) with p explicitly related to L ($L^* = 3.44$) as well as with p as a decision variable independent of L ($L_p^* = 0.4$). Solving this problem assuming p is explicitly related to L and that economies of scale exist, we obtain $L^* = 4.41$. Table 3.5.2 shows what the effect on profit will be, if in reality p is related to L and economies of scale exist as in our model, but the manager ignores either economies of scale or both economies of scale and the relation between p and L (i.e., guarantees $L = 3.44$ or 0.4 rather than 4.41 in our model).

Table 3.5.2: Comparison of μ^*, p^*, m^*, λ^* and π^* for Different L for Example 3.5.3

	$\mu^*(L)$	$p^*(L)$	$m^*(L)$	$\lambda^*(L)$	$\pi^*(L)$
Model with economies of scale and p and L related by $p = d - eL$ ($L^* = 4.41$)	24.91	8.03	1.74	24.02	149.54
Model without economies of scale and p and L related by $p = d - eL$ ($L = 3.44$)	22.73	8.91	2.03	21.59	147.23
Model without economies of scale and p and L as independent decision variables ($L = 0.40$)	23.76	11.64	3.81	14.00	108.27

The illustrative example clearly shows the importance of not only explicitly modelling the relation between p and L, but also the effect of demand on the operating cost. If a manager recognises the relation between p and L but ignores the effect of demand on operating cost when it really exists, the firm will be loosing a profit of about 1.5% (i.e., 149.54 vs 147.24). For large companies, this percentage difference can be substantial in dollar terms.

If the manager also ignores the explicit relation between p and L as in (3.4.2), then the firm stands to loose profit of about 38% (i.e., 149.54 vs 108.25)! We noted in Example 3.4.4 that for "more price-sensitive than lead-time-sensitive" customers, accounting for the dependence of price on the length of the guaranteed delivery lead time will lead to a comparatively higher L to bring down the price and increase demand. If in addition, economies of scale exist for unit operating costs, taking that into account will result in an even higher delivery lead time guarantee L to increase demand further and by that decrease operating cost. Note that our numerical examples are only illustrative; the exact amount of profit loss from using the "wrong" model will depend on the parameter values and can be larger or smaller than the results in Table 3.5.2.

In summary, this section presented a model of a profit-maximising make-to-order firm whose demand depends on guaranteed delivery lead time and price (which itself depends on the length of the delivery lead time guarantee) and economies of scale exist for operating costs. The firm must satisfy a delivery reliability constraint and can invest in increasing its processing rate. We were able to determine the optimal value of the guaranteed delivery lead time that will maximise the firm's profit. We can say that for practising managers it is not only important to know the customer preferences (as in Section 3.4), but also to take into account the effect of economies of scale when they are present.

3.6 Analytical Model with Holding and Backordering Costs

In Sections 3.4 and 3.5 we assumed that there is no holding cost for the raw materials work-in-process (WIP) and that the penalty cost is time-independent. Though for the physical system discussed in Section 3.2 these assumptions are reasonable, there might be firms for which it is necessary to develop models relaxing those assumptions. Suppose pre-processing requirements make it necessary for some firms to "take out" the raw material as soon as the order is received rather than wait for the order to reach the server before doing so. Then the raw material will be waiting in front of the server and there will be a WIP holding cost. If the facility is congested, then this holding cost can become significant. Similarly, if the penalty cost depends on the length of the customer waiting time, our model should be modified to explicitly account for such penalty costs.

In the model of this section all the other conditions remain the same as in Section 3.4, i.e., raw material is still available whenever required, demand depends on price and guaranteed delivery lead time while price itself is determined by the length of the guaranteed delivery lead time. In this section, we assume the unit operating cost to be constant (= m). The expected profit per unit time can now be written as:

$\pi(\mu, L)$ = Expected revenue per unit time - expected WIP holding cost per unit time - expected penalty cost per unit time - investment cost per unit time

$$= \lambda (p - m) - \frac{h'\lambda}{(\mu-\lambda)} - \frac{\xi\lambda}{(\mu-\lambda)} e^{-(\mu-\lambda)L} - M(\mu), \qquad (3.6.1)$$

where:

$\lambda = a' - b'L$,

$p = d - e'L$,

h' = WIP holding cost per unit per unit time (\$/unit/unit time),

and

ξ = penalty cost per time per unit lateness (\$/unit/unit time).

The first term of the right hand side of (3.6.1) represents expected revenue per unit time while the last term represents the investment costs per unit time. The second term is the WIP holding cost that can be calculated from Little's law. The third term represents the penalty cost for late jobs. Recall that for an *M/M/1* queue, the probability that a job/service is late is given by $P(W > L) = e^{-(\mu-\lambda)L}$ and the expected lateness given that a job is late is given by $[1/(\mu-\lambda)]$ (refer to Palaka et al. 1998).

The problem for the firm is now:

$$\text{(P3.5)}\ \underset{\mu, L}{\text{Maximise}}\ \pi(\mu, L) = \lambda (p - m) - \frac{h'\lambda}{(\mu-\lambda)} - \frac{\xi\lambda}{(\mu-\lambda)} e^{-(\mu-\lambda)L} - M(\mu), \qquad (3.6.2)$$

subject to:

$P(W < L) \geq s^R$, i.e., $1 - e^{-(\mu-\lambda)L} \geq s^R$ (delivery reliability constraint),

$\mu > \lambda$ (system stability constraint),

$p \geq m \geq 0$, $L > 0$, $\lambda \geq 0$ (non-negativity constraints).

We feel that even with an explicit penalty cost, the delivery reliability constraint should be there. In this case, we can define s^R to be a service level below which the demand rate falls drastically. So, the firm cannot afford to allow its service level to dip below s^R.

Proposition 3.6.1: $\pi(\mu, L)$ of P3.5 is concave with respect to μ.

Proof: Differentiating (3.6.2) with respect to μ we have:

$$\pi_\mu(\mu, L) = \frac{h'\lambda}{(\mu-\lambda)^2} + \frac{\xi\lambda}{(\mu-\lambda)^2} e^{-(\mu-\lambda)L} + \frac{\xi\lambda L}{(\mu-\lambda)} e^{-(\mu-\lambda)L} - M_\mu; \tag{3.6.3}$$

$$\pi_{\mu\mu}(\mu, L) = -\frac{2h'\lambda}{(\mu-\lambda)^3} - \frac{2\xi\lambda}{(\mu-\lambda)^3} e^{-(\mu-\lambda)L} - \frac{2\xi\lambda L}{(\mu-\lambda)^2} e^{-(\mu-\lambda)L} - \frac{\xi\lambda L^2}{(\mu-\lambda)} e^{-(\mu-\lambda)L} - M_{\mu\mu}. \tag{3.6.4}$$

Under the conditions $\mu > \lambda \geq 0$, $L > 0$ and $M_{\mu\mu} \geq 0$, it is clear that $\pi_{\mu\mu}(\mu, L) \leq 0$, i.e., π of (3.6.2) is concave in μ. ■

Differentiating (3.6.2) with respect to L we have:

$$\pi_L(\mu, L) = -e'\lambda - b'(p-m) + \frac{h'b'\mu}{(\mu-\lambda)^2} + \frac{\xi b'\mu}{(\mu-\lambda)^2} e^{-(\mu-\lambda)L} + \frac{\xi\lambda(\mu-\lambda+b'L)}{(\mu-\lambda)} e^{-(\mu-\lambda)L}; \tag{3.6.5}$$

$$\pi_{LL}(\mu, L) = 2b'e' - \frac{2h'(b')^2\mu}{(\mu-\lambda)^3} - \frac{2\xi(b')^2\mu}{(\mu-\lambda)^3} e^{-(\mu-\lambda)L} + 2\xi b' e^{-(\mu-\lambda)L} - \frac{2\xi(b')^2 L}{(\mu-\lambda)} e^{-(\mu-\lambda)L}$$
$$- \frac{2\xi(b')^2 L}{(\mu-\lambda)^2} e^{-(\mu-\lambda)L} - \frac{\xi\lambda(\mu-\lambda+b'L)^2}{(\mu-\lambda)} e^{-(\mu-\lambda)L}. \tag{3.6.6}$$

We cannot determine the sign of either (3.6.5) or (3.6.6). For $b' > 0$, the first two terms of π_L are negative while others are positive while for π_{LL} only the first term is positive and the others are negative. For $b' < 0$, the 2nd term of π_L is positive and the last term can be positive. For π_{LL} only the 4th term is positive. Hence π can be either convex or concave with respect to L.

If we assume L to be a given constant, then the mean demand rate and price of equations (3.4.2) and (3.4.3) will be constant and there is only one decision variable, μ. The constraint $1 - e^{-(\mu-\lambda)L} \geq s^R$ now implies that $\mu > [(-ln(1-s^R)/L) + \lambda] = \mu^R(L)$ (as in 3.4.8). Note that in this case, $\mu^R(L)$ is a constant. Since π is concave with respect to μ, we can solve $\pi_\mu = 0$ and let μ that solves $\pi_\mu = 0$ be $\mu_{opt}(L)$. If $\mu_{opt}(L) > \mu^R(L)$, then we would want μ^* to be $= \mu_{opt}(L)$. If $\mu_{opt}(L) \leq \mu^R(L)$, then we would want μ^* to be equal to $\mu^R(L)$. This implies that $\mu^* = \max(\mu^R(L), \mu_{opt}(L))$. Figure 3.6.1 illustrates the concept.

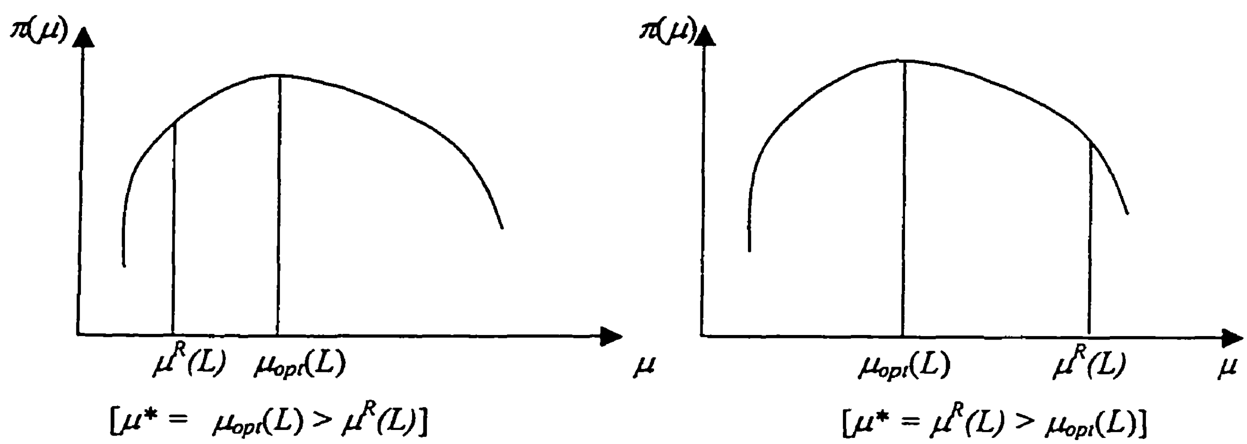


Figure 3.6.1: Optimal μ when L is Constant

In general, the solution of P3.5 will be analytically complex. However, using the concept of μ^* we can have some understanding of its solution.

Proposition 3.6.2: The optimal service level for P3.5 will be given by $s^* = Max\ (s^c, s^R)$ and the optimal processing rate by $\mu^*(L) = \frac{1}{L}[-ln(1 - Max(s^c, s^R))] + \lambda$.

Proof: From (3.4.7), we know that s is an increasing concave function of μ. This implies that there is a one-to-one relation between s and μ. Now let us define a critical service level, s^c, given by $s^c = 1 - e^{-(\mu_{opt}(L)-\lambda)L}$, i.e., s^c is the service level given by $\mu_{opt}(L)$, the processing rate maximising π for fixed L. From (3.4.8), we already know that $\mu^R(L)$ is the processing rate required to achieve the minimum desired service level s^R. Since for a fixed L, $\mu^* = \max\ (\mu^R(L), \mu_{opt}(L))$, the optimal service level will be given by:

$$s^* = Max\ (s^c, s^R). \tag{3.6.7}$$

We can now say that μ^* will satisfy $1 - e^{-(\mu^*-\lambda)L} = Max\ (s^c, s^R)$ from which:

$$\mu^*(L) = \frac{1}{L}[-ln(1 - Max(s^c, s^R))] + \lambda. \tag{3.6.8}$$

■

Note in (3.6.8), s^R is a known parameter fixed by management while s^c is a function of L. Only if p is independent of L would s^c be a constant the value of which will depend on the problem parameters (refer to Palaka et al., 1998).

Replacing the expression for $\mu^*(L)$ into the original expression of π in (3.6.2) we have the transformed problem P3.6 in terms of single variable, L.

$$\text{(P3.6)}\ \underset{L}{\text{Maximise}}\ \pi(L) = \lambda\,(p - m) - \frac{h'\lambda L}{x(L)} - \frac{\xi\lambda L}{x(L)}e^{-x(L)} - M(\mu^*(L)), \tag{3.6.9}$$

subject to:

$\lambda \geq 0, L > 0, p \geq m \geq 0,$

where:

$x(L) = -ln[1 - \text{Max}\ (s^c, s^R)]$, $\lambda = a' - b'L$, $p = d - e'L$ and $\mu^*(L)$ is given by (3.6.8).

Problem P3.6 is a function of only one variable, L. If we can determine the L^* that maximises (3.6.9), we can then substitute it in (3.6.8) to obtain μ^* and in (3.4.2) to obtain p^*. However, since s^c is also a function of L, the expression for π_L for P3.6 will be analytically difficult to handle and we leave this for future research. However, we can solve one special case of the general problem.

3.6.1 Special Case

Let us consider the special case of $b' = 0$ ($b_2 = b_1 e'$), i.e., while both p and L affects λ and also L affects p, the direct effect of L (b_2) is equal to the indirect effect of L through p ($b_1 e'$).

In this case π can be written as:

$$\pi(\mu, L) = a'(d - m - e'L) - \frac{h'a'}{(\mu - a')} - \frac{\xi a'}{(\mu - a')} e^{-(\mu - a')L} - M(\mu). \qquad (3.6.10)$$

We can show that π is concave with respect to L. Note that $b_2 - b_1 e' = 0 \Rightarrow e' = \frac{b_2}{b_1}$. Let us define the critical service level, $s^c = 1 - \left(\frac{e'}{\xi}\right)$ (e' must be $\leq \xi$) and $L_{opt}(\mu)$ as the solution to $\pi_L = 0$. From $\pi_L = 0$, we can show that $1 - \left(\frac{e'}{\xi}\right) = 1 - e^{-(\mu - a')L_{opt}(\mu)}$. The service level, s, is an increasing concave function of L. Following previous logic (in this case with respect to L rather than μ), we can say that $s^* = Max\ (s^c, s^R)$. Note that s^c is now a constant dependent on the parameter values e' and ξ. The expression for L^* will be given by:

$$L^*(\mu) = \frac{-\ln(1 - s^*)}{(\mu - a')}. \qquad (3.6.11)$$

As s^* is between zero and 1 and the constraint on μ is $\mu > a'$, so L^* will be positive. Replacing $L^*(\mu)$ in (3.6.10) we now have an objective function (to be maximised) in terms of a single variable μ with the constraints that $\mu > a'$ and $p \geq m \geq 0$ (as $s^* = Max\ (s^c, s^R)$, the delivery reliability constraint will always be satisfied). Since $\pi(\mu, L^*(\mu))$ is concave with respect to μ, the solution to $\pi_\mu = 0$ can give us μ^*. Replacing μ^* into (3.6.11) gives us L^* and then from (3.4.2) we can determine p^*. This combination of (μ^*, L^*, p^*) will maximise the constrained problem. This critical service level $1 - \left(\frac{e'}{\xi}\right)$ is in agreement with that of Palaka et al. (1998) for the special case $b' = 0$ (in their paper $s_c = 1 - \frac{b_2}{b_1\xi}$).

Example 3.6.1: Let us consider a numerical example with the following parameters: $a = 100$, $b_1 = 5$, $b_2 = 5$, $e' = 1$ (so that $b' = 0$), $d = 10$, $m = 1$, $M(\mu) = A\mu^3$ where $A = 0.0001$, $s^R = 0.98$, $h' = 1$ and $\xi = 10$. Then $s^c = 0.1$ implying that $s^* = s^R = 0.98$. In this case we can easily determine that $L^* = 0.2727$ and $\mu^* = 64.345$. The value of L^* will determine the price and the demand rate.

In this section we introduced WIP holding costs and penalty costs to the profit maximising problem of Section 3.4. The present literature that deals with holding and penalty costs do not take into account the relation between p and L. As we have seen in Section 3.4, it is important to take into consideration the relation between p and L when they exist. Otherwise, the model may give us substantially sub-optimal solutions. We could formulate the profit-maximising problem taking holding and penalty costs into account and assuming price to be dependent on the guaranteed delivery lead time, L. We were also able to express the objective function in terms of only the decision variable L. Though we could solve a special case, the structure of the general problem makes analytical solution difficult. We leave this for future research.

3.7 Conclusions and Future Research Opportunities

In this chapter, we modelled a make-to-order supply chain consisting of a firm and its customers where the mean demand rate is a function of price and guaranteed delivery lead time and the market price is determined by the length of the guaranteed delivery lead time. We then extended our model by incorporating economies of scale where the unit operating cost is a decreasing convex function of the mean demand rate. The firm can invest in increasing capacity to guarantee a shorter delivery lead time but must be able to satisfy the guaranteed delivery lead time according to a specified reliability level.

Our models explicitly accounted for "price-sensitive" and "lead-time-sensitive" customers. We showed how the firm could select the optimum length of the guaranteed delivery lead time to maximise its profits by a relatively simple procedure. Our numerical examples clearly showed that ignoring the dependence of market price on the guaranteed delivery lead time and the economies of scale, when they really exist, can lead to potentially large profit losses for the firm. It is also important for firms to take note of the inherent preference of the customers for price or delivery lead time when making decisions. We also extended our model by explicitly accounting for holding cost and penalty cost for the firm. However, in that case we were only able to formulate the problem without deriving many insights.

We feel that the models in this chapter captured some of the most salient characteristics of make-to-order firms and can help such firms to better manage their delivery times to gain competitive advantage in a time-based-competition environment.

Future research opportunities to extend our model include:

a) As So and Song (1998) point out, customers may be sensitive to the service level delivered. Hence, the mean demand rate may be a function of the service level in addition to that of price and delivery lead time;

b) Analytically solve the case where our model explicitly accounts for the WIP holding costs and penalty costs (i.e., Section 3.6);

c) Though we put forth our reasons for assumption of linear demand (Section 3.4), extension of our model with non-linear demand of the form used by So and Song (1998) and in our Chapter 5 might give new insights.

CHAPTER 4

SUPPLY LEAD TIME MANAGEMENT FOR MAKE-TO-STOCK FIRMS

4.1 Introduction

The growth of supply chains and emergence of speed as a key competitive priority means that supply lead time is one of the main performance drivers of the supply chain performance. There seems to be a growing feeling that supply lead time reduction opportunities need to be identified and adopted more widely (Australian National Audit Office Report 1997-98; Chopra and Meindl 2001). In this chapter we develop models to investigate the effects of investment in supply lead time reduction on inventory costs for a make-to-stock firm (Model B of Chapter 1).

Any type of lead time reduction requires careful planning and sometimes substantial investments (Zipkin 1991). For example, if faster processing or testing reduces lead times, capital intensive equipment acquisitions may be required. Some investments might not even be in the form of capital expenditure, but may amount to holding extra inventory. Clearly, the size and type of investment required to reduce lead times will depend on the type of process involved and also on the extent of change desired.

In the context of supply chains, management of lead times between a supplier and a buyer are crucial and often negotiated. A buyer may wish a supplier to shorten replenishment lead time to reduce the buyer's inventory costs or to reduce customer response time. This may involve the supplier, for example, investing in installing a new machine or a better information system or a new warehouse (e.g., some companies like General Motors and British Aerospace have "convinced" their suppliers to build warehouses near their assembly plant, so that supply lead time will be reduced). From the supplier's point of view, this investment might benefit a number of buyers, or it maybe primarily for a specific buyer (relation-specific or idiosyncratic investment; refer to Levi 1999, for more details of such investments). As several authors have pointed out, many small companies simply cannot afford improvements owing to the costs involved (Zipkin 1991; Groenevelt 1993; Bensaou 1999) and they seek to pass on some of the investment to the buyer (Chopra and Meindl 2001). This type of recouping of investment costs will happen if and only if there is no major disparity in the power between the two parties.

In recent times, we have seen that buyers are ready to pay for investments made by suppliers in lead time reduction since they believe that their "total cost" will still be reduced (Anderson et al. 1997; Chopra and Meindl 2001). For example, companies using more expensive transport services are finding that savings in inventory costs more than compensate for increased transportation costs (Australian National Audit Office Report 1997-98).

The focus of this chapter will be to investigate the effectiveness of investment in supply lead time reduction when the supplier will make the investment but will pass on the cost of the investment, partly or fully, to the buyer. We will develop analytical models that can capture the costs and benefits of such lead time reductions from the buyer's view point and can assist the buyer in deciding how much of lead time reduction to pursue.

Though there is a significant literature on investment in lead time reduction within the (Q, r) modelling framework, as we indicated in Chapter 2, there are two issues that have not yet been properly addressed:

a) Most research focuses on models that use a backordering cost per unit, disregarding the duration of the shortage or assume one order outstanding. However, state-of-the-art (Q, r) models provide a more "exact" representation that assumes a backordering cost per unit per unit time, allows more than one order outstanding and a negative reorder point. Our work deals with investing in lead time reduction in the latter framework;

b) Most research does not differentiate between the frequency and the nature of the investment and the effect this has on modelling. This chapter will address the issue of different types of investment in lead time reduction in the (Q, r) framework and show that it is very important for the buyer to consider such issues in decision making.

4.2 Overview of the Physical System

We consider a single firm (buyer) who procures a make-to-stock product in batches from a supplier according to a continuous review (Q, r) control system. Customer demand occurs one unit at a time and is satisfied directly from the warehouse. Since it is a make-to-stock product, it is reasonable to assume constant customer demand. The procurement lead time is stochastic which results in a stochastic lead time demand for the buyer. All unmet demands are backordered. The buyer wants the supplier to improve replenishment lead time by reducing its mean and/or variability. A schematic representation of the two-party supply chain is shown in Figure 4.2.1.

Suppose the supplier, in order to respond to the buyer, must make a substantial investment to reduce the supply lead time. The investment might be relation-specific if the improvements will not help the supplier vis-a-vis its other customers or it might be for a number of customers. We assume that the supplier can pass on at least part of the investment cost to the buyer (i.e., there is no major disparity in market power between the buyer and the supplier). The buyer will be the ultimate decision maker and has to take both the costs and benefits of lead time reduction into consideration before deciding how much of a reduction in lead time to request, if any. Hence, we assume that the supplier will first inform the buyer of the cost consequences of investment in lead time reduction but will make the investment only after the buyer has made a decision.

Though we assume that the buyer and the supplier are two separate entities, they may also be parts of the same organisation. Since the environment we are considering is make-to-stock, any reduction in the procurement lead time (either external or internal) has no impact on the delivery time to the ultimate customer. This lead time reduction will only affect the buyer's inventory costs (Hill and Khosla 1992).

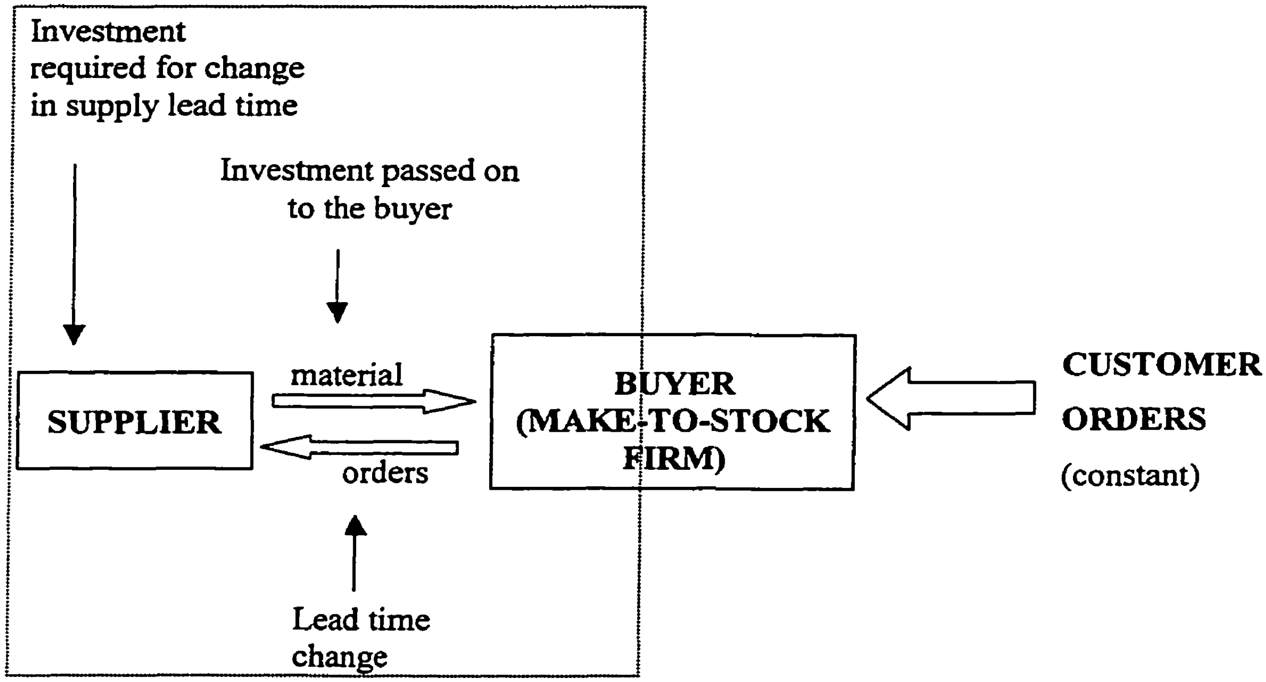


Figure 4.2.1: Supply Chain System for Chapter 4

In the presence of stochastic lead times, to obtain proper analytical results in a (Q, r) framework, we must make sure that: (i) deliveries of orders cannot cross in time, and (ii) the supply lead time is independent of the number and size of the outstanding orders (Porteus 1990). These conditions will hold in the above supply chain under the assumptions that the supply system is exogenous and sequential (Zipkin 2000). An exogenous system means that while the supplier's overall workload may fluctuate over time, the buyer's orders contribute little to these fluctuations. This implicitly assumes a reasonably large supplier. An exogenous system ensures that the supply lead times for the orders will be independent of the number and size of the outstanding orders. A sequential system is one for which the supplier processes the buyer's orders in a FIFO (First-In-First-Out) manner. This implies that there will be no order crossing even if there is more than one order outstanding at any time. For examples demonstrating the generality of these assumptions refer to Zipkin (2000).

4.3 Notation

The following notation will be used in this chapter:

$I(t)$	= inventory on hand at time t (units)
$B(t)$	= backorders at time t (units)
$IL(t)$	= inventory level at time $t = I(t) - B(t)$ (units)
$IP(t)$	= inventory position at time t
	$= IL(t)$ + orders outstanding at time t (units)
$E(I)$	= expected on hand inventory (units)
$E(B)$	= expected backorders (units)
K	= fixed set-up cost incurred for each order ($/order)
c	= purchase price/unit for the buyer from the supplier ($/unit)
i	= period interest rate (while our model can handle different interest rates for the buyer and the supplier, for notational simplicity we assume it to be equal for both parties) ($/$/unit time)
h	= holding cost/unit/time = ic ($/unit/unit time)
b	= backordering penalty cost/unit backordered/time ($/unit/unit time)
ζ	= backordering penalty cost/unit backordered ($/unit)
r	= reorder point (units)
Q	= batch size (units)
λ	= demand rate (units/unit time)
C	= expected total inventory cost per unit time ($/unit time)

X will represent the continuous random variable denoting lead time demand (LTD). It will have cumulative distribution function (CDF) F, density f and mean μ.

4.4 The (Q, r) Model

For demands occurring one at a time it is known that (Q, r) policy is optimal (Zipkin 2000). With (Q, r) inventory control, the buyer orders Q units as soon as the *IP* reaches the reorder point, r. Normally, the planning horizon is infinite and the objective is to minimise long-run expected total inventory costs per unit time. The classical inventory model of this type was developed by Hadley and Whitin (1963). That model includes (average) ordering costs, inventory holding costs and backorder penalty costs. Purchase cost is not considered, as for their model it does not depend on the decision variables - Q (batch size) and r (reorder point). The most often used expression for expected long-run total inventory cost per unit time is (Nahmias 1997):

$$C(Q, r) = \frac{K\lambda}{Q} + h\{r + \frac{Q}{2} - \mu\} + \frac{1}{Q}[\zeta\lambda \int_{x=r}^{\infty}(x-r)f(x)dx]. \tag{4.4.1}$$

The above model is an approximation of the exact cost expression. It is adequate for some situations, assuming arrival process is Poisson, backordering is negligible and time dependent backordering costs is not present. For details about the conditions when the approximation might not work well, refer to Zipkin (1986b) and Chapter 2.

Because of the importance of Zheng's (1992) research on continuous review (Q, r) modelling to this thesis, it will be helpful to review his basic model. The author models a fully backordered, single-item, continuous review inventory system where demands arrive at rate λ. Demand is random and the relevant costs are ordering, holding and backordering penalty. The backordering cost is taken to be per unit per unit time, b, and the objective is minimisation of long-run expected inventory costs per unit time. Zheng combined the backordering and holding costs together and termed it as inventory costs. If $G(y)$ is the rate of accumulation of expected inventory costs at time $(t + L^T)$ when the inventory position at time t equals y and X is the random variable for lead time demand, i.e., total demand during the time interval from t to $(t + L^T)$ where L^T is the constant procurement lead time, we have:

$$G(y) = E\,[h(y - X)^+ + b(X - y)^+], \tag{4.4.2}$$

where $(x)^+ = \max(x, 0)$.

With the above definition of $G(y)$, the total cost function can be written as:

$$C(Q, r) = \frac{K\lambda}{Q} + \frac{1}{Q}\left[\int_{y=r}^{r+Q} G(y)dy\right]. \tag{4.4.3}$$

The above model holds under the conditions that the inventory position in steady state is uniformly distributed on the interval $(r, r+Q)$ and is independent of lead time demand. These conditions are met when a non-decreasing stochastic process with stationary increments and continuous sample paths can model cumulative demand. Hence, this model does not require the assumption of Poisson arrivals like Hadley and Whitin's model. Zipkin (1986a, 1986b) shows that these conditions hold even for quite general stochastic lead time distributions. The above cost equation assumes inventory position, lead time demand and the decision variables, Q and r, are continuous. This assumption is reasonable and is quite good as long as the order quantity is not too small (Zipkin 1986a, 1986b). If the demands and inventory positions are assumed to be discrete then the cost function is:

$$C(Q, r) = \frac{K\lambda}{Q} + \frac{1}{Q}\left[\sum_{y=r+1}^{r+Q} G(y)\right]. \tag{4.4.4}$$

In the above case, the steady state inventory position is uniform in the interval $\{r+1, r+2, \ldots, r+Q\}$. Zheng studies the continuous cost function because of ease of comparison to EOQ model (with backordering) where Q and r are typically assumed to be continuous. It can be shown that under a given (Q, r) policy, $E(I) = (1/Q)\int_{y=r}^{r+Q} E(y - X)^+ dy$ and $E(B) = (1/Q)\int_{y=r}^{r+Q} E(X - y)^+ dy$, and it can also be verified that $E(I) = (\frac{Q}{2} + r - L^T\lambda) + E(B)$.

Using this relation, (4.4.3) can be written as:

$$C(Q, r) = \frac{K\lambda}{Q} + hE(I) + bE(B) = \frac{K\lambda}{Q} + h(\frac{Q}{2} + r - L^T\lambda) + (h+b)E(B). \qquad (4.4.5)$$

Zipkin (1986b) proved the joint convexity of $E(B)$ in Q and r for (4.4.5). All other terms in (4.4.5) are clearly convex; so $C(Q, r)$ is also jointly convex in Q and r. Note that the term $G(y)$ of (4.4.3) can also be written as:

$$G(y) = (h+b) \int_{t=0}^{y} F(t)dt + b(L^T\lambda - y), \qquad (4.4.6)$$

which implies:

$$C(Q, r) = \frac{K\lambda}{Q} + \frac{1}{Q} \int_{y=r}^{r+Q} [(h+b)\int_{t=0}^{y} F(t)dt + b(L^T\lambda - y)]dy, \qquad (4.4.7)$$

where $L^T\lambda = \mu$, is the mean lead time demand.

The joint convexity in Q and r makes possible the sequential minimisation of the above cost function – first with respect to r and then with respect to to Q. The first-order condition with respect to r (with Q fixed) for the optimal reorder point, r^*, is then, $r = r^*(Q)$ (for any $Q > 0$) if and only if, $G(r) = G(r+Q)$ with $G(y)$ defined as in (4.4.6). Then the problem becomes that of minimising $C(Q, r^*(Q))$. Defining $H(Q) = G(r^*(Q))$, it can be shown that the cost function becomes:

$$C(Q) = \frac{Kd}{Q} + \frac{1}{Q} \int_{y=0}^{Q} H(y)dy. \qquad (4.4.8)$$

The convexity of the cost function gives the optimality condition for optimal batch size, Q^*, as $H(Q^*) = C(Q^*)$ and, finally, (Q^*, r^*) is optimal for $C(Q, r)$ if and only if $C(Q^*, r^*) = G(r^*) = G(r^* + Q^*)$.

Even with these simultaneous equations it might be difficult to solve for the optimal values. While for a given Q, the optimal r is independent of K, assuming r^* is selected properly, Q^* will depend on both ordering costs and inventory costs. Results comparing the stochastic model with an EOQ model with backordering have already been discussed in Chapter 2.

In the above model, the main random variable of interest is the lead time demand (LTD). Though Zheng's model assumed that the variability in LTD stems from stochastic demand (assuming supply lead time to be deterministic), all the results hold even if the variability in the LTD stems from the lead time duration (with demand constant), or even if both lead time duration and demand are stochastic (Zipkin 2000). For stochastic lead times, it is necessary to assume an exogenous and sequential system, like ours, for the results to hold.

4.5 The (Q, r) Model with Reduced Lead Time

Suppose, at the request of the buyer, the supplier decides to reduce the lead time duration. In this section, we develop a (Q, r) model with reduced supply lead time. In the next section we will incorporate the investment costs required to reduce the supply lead time. Any reduction in lead time duration will change the LTD. We assume α to be the decision variable that signifies the reduction of LTD due to a change in lead time duration. Let the new "reduced" LTD random variable be $\hat{X}$. A particular form of transformation that allows us to reduce mean and variability simultaneously is simply:

$$\hat{X} = \alpha X. \tag{4.5.1}$$

In this case, $E(\hat{X}) = \alpha\mu$ and $\mathrm{Var}(\hat{X}) = \alpha^2\mathrm{Var}(X)$, both decrease when α declines. Note that the transformation is not general in the sense that it assumes that the standard deviation (std) and mean of the LTD decline by the same fraction, i.e., the CV (co-efficient of variation) remains constant. This modelling approach is reasonable given the literature that indicates that as the absolute mean lead time is reduced, it will also reduce the absolute lead time variation and the CV will remain almost constant (Ballou 1998).

The transformation in (4.5.1) does not assume any particular lead time demand distribution, unlike models in the literature which are generally based on normal LTD distribution. Since in our case the investment cost required to reduce α will also be quite general, the transformation in (4.5.1) is not as restrictive as it may seem.

Proposition 4.5.1: For only non-negative values, $\hat{X}$ of the form of (4.5.1) is stochastically non-decreasing in α.

Proof: $\frac{d}{d\alpha}E[\psi(\hat{X})] = E[\frac{d}{d\alpha}[\psi(\hat{X})]] = E[\frac{d}{d\alpha}[\psi(\alpha X)]] = E[\psi_{\alpha}(\alpha X)X]$. As long as ψ is increasing and X is non-negative, it is clear that $\hat{X}$ is stochastically non-decreasing in α (Ross 1983). For a related proof also refer to Bookbinder and Çakanyildirim (1999).■

This transformation was used by Gupta and Gerchak (1995), Bookbinder and Çakanyildirim (1999) and Gerchak and He (2000). While ideally the only restriction on α should be $\alpha \geq 0$, for an existing system it makes sense to have $0 \leq \alpha \leq 1$ with $\alpha = 1$ as the status-quo. When $\alpha = 0$, both the mean and the variance of the lead time duration, and hence the LTD, are also equal to zero.

We can now develop the cost model from the buyer's viewpoint. It is known that the limiting distributions of the random variables IL, IP and lead time demand (X) are related by $IL = IP - X$. Since, IP and X are independent, and if IP, X, Q and r are taken as continuous, then IP is uniformly distributed between r and $(r + Q)$ (Zipkin 1986a). IL is related to I and B by $IL = I - B$. The expected value of the difference of two random variables is the difference of their expected values. Thus:

$$E(I) = E(B) + E(IP) - E(X). \qquad (4.5.2)$$

Replacing X by $\hat{X}$ (from 4.5.1), we have:

$$E(I) = E(B) + E(IP) - E(\hat{X}), \qquad (4.5.3)$$

and

$$hE(I) + bE(B) = h\{E(B) + E(IP) - E(\hat{X})\} + bE(B)$$

$$= (h + b)E(B) + h\{E(IP) - E(\hat{X})\}. \tag{4.5.4}$$

Hence, with the transformation, we can write the expected long-run total inventory cost per unit time as:

Expected total inventory cost / unit time = $C(Q, r, \alpha)$

$$\begin{aligned}
&= \text{Set-up/Ordering cost} + \text{Holding Cost} + \text{Backordering Cost} + \text{Purchase Cost} \\
&= \frac{K\lambda}{Q} + hE(I) + bE(B) + c\lambda \\
&= \frac{K\lambda}{Q} + (h + b)E(B) + h\{E(IP) - E(\hat{X})\} + c\lambda. \qquad (4.5.5)
\end{aligned}$$

As IP is uniformly distributed between r and $(r+Q)$, by conditioning and unconditioning on IP we obtain the following expression:

$$E(IP) = \frac{1}{Q}\int_{s=r}^{r+Q} E(IP = s)\, ds, \tag{4.5.6}$$

and

$$E(B) = \frac{1}{Q}\int_{s=r}^{r+Q} E(B|IP = s) ds. \tag{4.5.7}$$

Note that $(B|IP = s)$ represents the backorder level given that the inventory position at which the order is placed is s. When $IP = s$ and the lead time demand variable is $\hat{X}$, $(B|IP = s)$ is defined as follows:

$$\begin{aligned}
& (B|IP = s) && = 0 && \text{if} && \hat{x} < s \\
& && = (\hat{x} - s) && \text{if} && \hat{x} \geq s \\
\Rightarrow\ & (B|IP = s) && = 0 && \text{if} && \alpha x < s \text{ (from 4.5.1)} \\
& && = (\alpha x - s) && \text{if} && \alpha x \geq s \\
\Rightarrow\ & (B|IP = s) && = 0 && \text{if} && x < s/\alpha \\
& && = (\alpha x - s) && \text{if} && x \geq s/\alpha.
\end{aligned}$$

The above is quite straightforward. If the LTD is more than the *IP* at which the order is placed, then there will be backordering; otherwise there will be no backordering.

Therefore,
$$E(B|IP = s) = \int_{(s/\alpha)}^{\infty} (\hat{x} - s) f(x) dx$$
$$= \int_{(s/\alpha)}^{\infty} (\alpha x - s) f(x) dx$$
$$= \alpha \int_{(s/\alpha)}^{\infty} (x - \frac{s}{\alpha}) f(x) dx.$$

From (4.5.7) and $E(B|IP = s)$ we obtain:

$$E(B) = \frac{1}{Q} \int_{r}^{r+Q} E(B|IP = s) = \frac{\alpha}{Q} \int_{r}^{r+Q} \{ \int_{(s/\alpha)}^{\infty} (x - \frac{s}{\alpha}) f(x) dx \} ds. \tag{4.5.8}$$

From (4.5.6), $E(IP)$ can be easily seen to be:

$$E(IP) = \frac{1}{Q} \int_{r}^{r+Q} E(IP = s) ds = \frac{1}{Q} \left(\frac{(r+Q)^2}{2} - \frac{r^2}{2} \right)$$
$$= r + (Q/2). \tag{4.5.9}$$

From (4.5.5), (4.5.8) and (4.5.9) and recalling that $E(\hat{X}) = \alpha\mu$, we have:

$C(Q, r, \alpha)$

$$= \frac{K\lambda}{Q} + h\{r + \frac{Q}{2} - \alpha\mu\} + \frac{(h+b)\alpha}{Q} \int_{r}^{r+Q} \{ \int_{(s/\alpha)}^{\infty} (x - \frac{s}{\alpha}) f(x) dx \} ds + c\lambda. \tag{4.5.10}$$

Simplifying,
$$\int_{s=r}^{\infty} \{ \int_{x=(s/\alpha)}^{\infty} (x - \frac{s}{\alpha}) f(x) dx \} ds$$
$$= \int_{s=r}^{\infty} \{ \int_{x=(s/\alpha)}^{\infty} \overline{F}(x) dx \} ds$$
$$= \int_{x=(r/\alpha)}^{\infty} \overline{F}(x) dx \int_{s=r}^{\alpha x} ds$$

$$= \quad \alpha \int_{x=(r/\alpha)}^{\infty} (x - \frac{r}{\alpha})\, \overline{F}(x)dx, \qquad (4.5.11)$$

where $\overline{F} = 1 - F$.

Similarly, $\int_{s=r+Q}^{\infty} \{ \int_{x=(s/\alpha)}^{\infty} (x - \frac{s}{\alpha}) f(x)dx$

$$= \quad \alpha \int_{x=(r+Q)/\alpha}^{\infty} (x - \frac{r+Q}{\alpha}) \overline{F}(x)dx. \qquad (4.5.12)$$

Now (4.5.10) can be written as:

$$C(Q, r, \alpha) = \frac{K\lambda}{Q} + h\{r + \frac{Q}{2} - \alpha\mu\} + \frac{(h+b)\alpha^2}{Q}[N(\frac{r}{\alpha}) - N(\frac{r+Q}{\alpha})] + c\lambda, \qquad (4.5.13)$$

where $N(y/\alpha) = \int_{x=(y/\alpha)}^{\infty} [x - (y/\alpha)]\overline{F}(x)dx$. Note that F still represents the distribution of the original LTD.

Another way of representing $C(Q, r, \alpha)$ is to simplify the term $hE(I) + bE(B)$. Proceeding as in (4.5.8) we can easily show that:

$$E(I) = \frac{\alpha}{Q} \int_{r}^{r+Q} \{ \int_{0}^{(s/\alpha)} (\frac{s}{\alpha} - x)\, f(x)dx\} ds. \qquad (4.5.14)$$

Replacing (4.5.8) and (4.5.14) and simplifying we obtain:

$hE(I) + bE(B)$

$$= \quad \frac{1}{Q} \int_{r}^{r+Q} \{(h\,\alpha(\int_{0}^{(s/\alpha)} (\frac{s}{\alpha} - x) f(x)dx)) \;+\; (b\,\alpha(\int_{(s/\alpha)}^{\infty} (x - \frac{s}{\alpha})\, f(x)dx))\} ds$$

$$= \quad \frac{1}{Q} \int_{r}^{r+Q} \{h\,\alpha(\frac{s}{\alpha})F(\frac{s}{\alpha}) - h\alpha \int_{0}^{(s/\alpha)} x f(x)dx$$

$$+ b\alpha[\mu - \int_{0}^{(s/\alpha)} x\, f(x)dx] - b\alpha(\frac{s}{\alpha})[1 - F(\frac{s}{\alpha})]\} ds$$

$$= \qquad \frac{1}{Q}\int_{r}^{r+Q} \{(h+b)\,\alpha \int_{0}^{(s/\alpha)} F(x)dx + b\alpha(\mu - \frac{s}{\alpha})\}ds, \qquad (4.5.15)$$

and

$$C(Q, r, \alpha)$$

$$= \frac{K\lambda}{Q} + hE(I) + bE(B) + c\lambda$$

$$= \frac{K\lambda}{Q} + \frac{\alpha}{Q}\int_{r}^{r+Q} G(\alpha, y)dy + c\lambda, \qquad (4.5.16)$$

where $G(\alpha, y) = (h+b)\int_{0}^{(y/\alpha)} F(x)dx + b(\mu - \frac{y}{\alpha})$.

While both (4.5.13) and (4.5.16) represent the same expected long-run total inventory costs per unit time for the buyer, depending on the circumstances use of one might be preferable than the other. This cost model is "exact" under the assumptions that cumulative demand can be modelled by a non-decreasing stochastic process with stationary increments and continuous sample paths and the supply system is exogenous and sequential. Under these conditions, it is applicable for stochastic lead time demand and more than one order outstanding, regardless of whether the variability arises from demand or supply lead time. Remember that in our case demand is constant and variability is due to lead time duration.

4.5.1 Mean Preserving Transformation

Another transformation that can be used keeps the mean LTD constant, while variability decreases with α (hence called a Mean Preserving Transformation). In that case, the "reduced" lead time demand random variable $\hat{X}$ will be given by:

$$\hat{X} = \alpha X + (1 - \alpha)\mu. \qquad (4.5.17)$$

Note that, $E(\hat{X}) = \alpha E(X)+(1-\alpha)\mu = \alpha\mu+(1-\alpha)\mu=\mu$ (for all α) and $\mathrm{Var}(\hat{X}) = \alpha^2\mathrm{Var}(X)$. Not only does the variance of the transformed variable increase in α, the transformed random variable become more variable in the general notion of variability (for more details refer to Gerchak and He 2000). Since the mean of the transformed variable remains the same for all α and $\hat{X}$ is non-negative for all α, to prove the assertion about variability we have to show that for $\hat{X}$ of the form of (4.5.17), $E[\psi(\hat{X})]$ is non-decreasing in α where ψ is a convex function (Ross 1983).

Proposition 4.5.2: $\hat{X}$ of the form of (4.5.17) becomes more variable as α increases.

Proof: $\frac{d}{d\alpha}E[\psi(\hat{X})]= E[\frac{d}{d\alpha}[\psi(\hat{X})]] = E[\frac{d}{d\alpha}[\psi(\alpha X+(1-\alpha)\mu)]]$

$= E[\psi_\alpha[(X-\mu)\alpha+\mu](X-\mu)]$. Replacing the random variable Y for $\alpha(X - \mu)$ (note that $E(Y) = 0$) we have, $E[\psi_\alpha[(X-\mu)\alpha+\mu](X-\mu)] = \frac{1}{\alpha}E[\psi_\alpha(Y+\mu)Y]$.

However, $E[\psi_\alpha(Y+\mu)Y] = \int_{-\infty}^{0} y\psi_\alpha(y+\mu)f(y)dy + \int_{0}^{\infty} y\psi_\alpha(y+\mu)f(y)dy$. Since ψ is convex, it implies that ψ_α is increasing and hence,

$\int_{-\infty}^{0} y\psi_\alpha(y+\mu)f(y)dy + \int_{0}^{\infty} y\psi_\alpha(y+\mu)f(y)dy$

$\geq \int_{-\infty}^{0} y\psi_\alpha(\mu)f(y)dy + \int_{0}^{\infty} y\psi_\alpha(\mu)f(y)dy.$

Since, $\int_{-\infty}^{0} y\psi_\alpha(\mu)f(y)dy + \int_{0}^{\infty} y\psi_\alpha(\mu)f(y)dy = \psi_\alpha(\mu)E(Y)=0$ and $\alpha \geq 0$, it is clear that $\frac{d}{d\alpha}E[\psi(\hat{X})] \geq 0$ and hence, $E[\psi(\hat{X})]$ is non-decreasing in α. Hence proved.■

With the transformation of (4.5.17),

$$E(B) = \frac{\alpha}{Q}\int_{r}^{r+Q}\{\int_{x=s/\alpha}^{\infty}(x-\frac{s-(1-\alpha)\mu}{\alpha})f(x)dx\}ds,$$

and $$E(IP) = r+(Q/2).$$

So, the long-run expected inventory costs per unit time for the buyer will be given by (after some simplification):

$$C(Q, r, \alpha) = \frac{K\lambda}{Q} + h\{r + \frac{Q}{2} - \mu\} +$$

$$\frac{(h+b)\alpha^2}{Q}[N(\frac{r-(1-\alpha)\mu}{\alpha}) - N(\frac{(r+Q)-(1-\alpha)\mu}{\alpha})] + c\lambda . \quad (4.5.18)$$

Though the transformation (4.5.17) was used by Gerchak and Parlar (1991) and Gerchak and Mossman (1992), some reflection on (Q, r) models leads to the realisation that, in general, changes only in the mean lead time demand confer no cost benefits, do not change Q, and change r by the same amount as change in the mean. Thus the transformations (4.5.1) and (4.5.17) are essentially equivalent here. This is formally proved in Appendix 4.1. In particular, for constant demand, as in our model, reduction only in mean lead time duration reduces only the mean lead time demand and hence has no effect on Q or C. Since the transformation in (4.5.1) is simpler to work with analytically, we will adopt it as our model. This result implies that in the domain of the (Q, r) model, MPT as a technique has virtually no added significance.

4.6 The (Q, r) Model with Investments in Lead Time Reduction

In Section 4.5 we developed (Q, r) cost models with reduced lead time duration. However, both cost equations (4.5.13) and (4.5.16) have ignored the costs needed to achieve a reduction in lead time. Using the transformation, the parameter we invest in reducing is α (assuming demand is constant). If there was no cost in reducing α, the inventory policy we will be following will be the EOQ model with backordering (which is a first-order approximation of the stochastic model). Zheng (1992, pgs 94-95) shows that the cost of the optimal EOQ model with backordering will always be less than that of the optimal stochastic model cost with backordering. If the cost of changing α is nil, the best option for the buyer is obviously $\alpha = 0$. But, as we have already mentioned, most of the time there will be a cost associated with reducing α.

While the functional form of the investment needed to reduce α will be case specific, it is reasonable to assume that such costs will be decreasing convex in α, i.e., successive reductions in α will require larger and larger investments per unit reduction. Let the investment function to reduce α be denoted by $M(\alpha)$. A reduction of α means a reduction in both the mean and variance of the lead time duration, and therefore a reduction in the mean and variance of the lead time demand. Though, at this point, we do not specify any particular form for $M(\alpha)$, we assume that there is a finite cost to maintain the status-quo lead time duration distribution (i.e., $M(1)$ is finite) and that the cost to reduce the mean and variance of the lead time to zero {i.e., $M(0)$} is essentially infinite. From this preamble, it is reasonable to assume that $M(\alpha)$ has the following characteristics:

$$M(\alpha) \geq 0,\ M_{\alpha}(\alpha) \leq 0,\ M_{\alpha\alpha}(\alpha) \geq 0. \tag{4.6.1}$$

From an "engineering" perspective, $M(\alpha)$ might be very different from investment to investment.

To get further insights into how companies invest in reducing lead time in practice, we contacted several Canadian companies. Their experience seems to suggest that investment in reducing α by the supplier can be of two main types: i) Recurring investment, where $M(\alpha)$ represents each instalment, and ii) One-time investment, where $M(\alpha)$ represents that investment.

Recurring investments can be incurred: i) Per unit (e.g., use of more skilled labour that reduces processing time but increases per unit variable cost), or ii) Per cycle (e.g., costlier but faster transportation each cycle), or iii) Per unit time (e.g., special maintenance checks each period by the supplier to decrease machine downtime).

The one-time investment might also be of three types depending on whether the life of the investment depends on: i) Number of units produced by the investment (e.g., a test equipment or machine that can be used for certain number of parts), or ii) Number of cycles it can be used (e.g., an apparatus that can reduce set-up time but can be used only certain number of cycles), or iii) Lifetime of its use (e.g., lease of an IT facility valid for certain amount of time).

Let us explain the concept in some more detail. When the buyer requests the supplier to reduce the lead time duration, the supplier examines its own facility and notes that the greatest opportunity to reduce supply lead time lies in reducing transportation time. Suppose at status-quo, the supplier is using rail for transportation. To reduce its lead time, it decides to use trucks. While it will reduce both the mean and variability of lead time duration (and so the mean and variability of lead time demand), the supplier has to incur an extra cost per cycle for transportation (Ballou 1998) which it will pass on to the buyer. The supplier can reduce the mean and variability of lead time duration further by using air transport which will require even more cost per cycle on the part of the buyer to be paid to the supplier. As α is getting reduced from rail to truck to air (i.e., both mean and variability of lead time duration and lead time demand), the buyer is incurring an extra cost; however, this extra cost is a recurring cost per cycle.

On the other hand, the most opportunity to reduce the mean and variability of supply lead time may be in using more skilled labourers that will reduce the processing time. Such an action will increase the direct labour cost for the supplier, which it will pass on to the buyer. In this case, the buyer will be paying a price premium every unit for early delivery.

Though both types of investment lead to a reduction of α, the "physical" type of investment done by the supplier and the effect of the extra cost passed on to the buyer is very different. In practice, the opportunities for improvement most likely will be finite and hence $M(\alpha)$ will have a finite number of discrete values; but for analytical simplicity, we assume $M(\alpha)$ to be continuous.

By focusing on the different nature of investments and amortisation schemes we have been able to make our model general enough to be used in diverse models of lead time reduction related to transportation, maintenance and capital expenditure, as is evident from the examples given above. Most previous articles seem to focus on recurring investments per unit time or one-time investment where the life of the investment depends on lifetime of its use.

Depending on the nature of the relationship, only a fraction, say θ ($0 < \theta \leq 1$), rather than all of the investment might be passed on by the supplier to the buyer. One of the cases where $\theta = 1$ is for an in-house supplier and then both the buyer and the supplier are parts of the same organisation. Though we assumed that the investment would be relation-specific, in reality, there might be cases where the investment in supply lead time reduction can benefit a number of buyers. Then it might be reasonable to allocate some part of the investment for the particular buyer and the fraction of the investment to be allocated (θ) can be decided by negotiation (for some related ideas about this type of investment allocation refer to Gerchak and Gupta 1991). However, some types of investments, e.g., changing the transportation mode that is used solely for one particular buyer, will require no investment allocation. In that case, whether θ should be equal to 1 or less than 1 will be dependent on the nature of the buyer-supplier relationship.

Thus, the nature of the investment will result in 6 different types of models. While the basic inventory model has been already discussed, we will now illustrate how the nature of the investment will affect modelling. For analytical simplicity we will assume that the one time investment will have a long life, so that as an approximation we can assume infinite life (this approximation works well; refer to literature review of Nye 1997).

We will use the following additional notation to signify the different investment functions and amortisation schemes: $M^{AB}(\alpha)$ where the first superscript indicates whether the investment is one-time (1) or recurring (R). The second superscript signifies whether the investment is done per unit (U), per cycle (C) or per period (T) for recurring investment or whether the life of the investment depends on the number of units produced (U), cycles (C) or periods (T) it will be used for one-time investment. For example, $M^{RC}(\alpha)$ represents each instalment of a recurring investment done each cycle and $M^{IU}(\alpha)$ represents a one-time investment where its life depends on the number of units produced. If only a fraction θ of the investment costs is passed on, then Table 4.6.1 shows the effects on $C(Q, r, \alpha)$ in (4.5.13), depending on the nature of the investment. Note that here we are assuming that the supplier can reduce the lead time by making any one of the investments, i.e., the investments are mutually exclusive. In reality, there may be cases where more than one type of investment can be done simultaneously. It will only result in more complex cost functions and we leave it for future research.

Table 4.6.1: Cost Models with Different Investment Functions

Type of Investment	Investment Incurred	Changes in $C(Q, r, \alpha)$ of (4.5.13)
Recurring	1A) Per unit	c replaced by $c(\alpha) = c + \theta M^{RU}(\alpha)$ and h by $h(\alpha) = ic(\alpha)$
	1B) Per cycle	[2] K replaced by $K(\alpha) = K + \theta M^{RC}(\alpha)$
	1C) Per unit time	Extra term $\theta M^{RT}(\alpha)$
One time	2A) Life depends on the number of units produced	[3] c replaced by $c(\alpha) = c + [\theta M^{IU}(\alpha)(i/\lambda)]$ and h by $h(\alpha) = ic(\alpha)$
	2B) Life depends on the number of cycles used	Extra term $[\{\theta M^{IC}(\alpha)i\}/(\lambda/Q)](\lambda/Q) = \theta M^{IC}(\alpha)i$
	2C) Life depends on the length of use	Extra term $\theta M^{IT}(\alpha)i$

[2] In this case, we can view the situation as if the cost per set-up has increased from K to $[K + \theta M^{RC}(\alpha)]$ while length of the set-up time (and hence lead time) has decreased. This model is a mixture of two cases: i) increased set-up cost, and ii) reduced lead time.

[3] Note that while Gerchak and Parlar (1991) referred to this type of investment, they did not take into account the effect on holding costs. For Models 1A and 2A, both the purchase and the holding cost will be affected. Though less safety stock may be required, each unit will cost more to procure and hold.

The developments of the recurring investment models are relatively straightforward.

Model 1A:

If there is a recurring unit investment of $M^{RU}(\alpha)$ per unit and a fraction θ of it is decided to be passed on to the buyer, then the buyer has to pay a purchase price of $c(\alpha) = c + \theta M^{RU}(\alpha)$ per unit and the holding cost will be $h(\alpha) = ic(\alpha)$ (as α decreases both the purchase price and the holding cost goes up). Hence, the purchase cost, c, in (4.5.13) or (4.5.16) will be replaced by $c(\alpha)$ and the holding cost per unit per time, h, by $h(\alpha)$.

Model 1B:

If there is a recurring per cycle investment of $M^{RC}(\alpha)$ each cycle and a fraction θ of it is passed on, then though there will be no effect on the purchase or the holding cost, the effective set-up cost per cycle, $K(\alpha) = K + \theta M^{RC}(\alpha)$, will change with α. As α decreases, the buyer has to pay an increased effective set-up cost.

Model 1C:

If there is a recurring investment of $M^{RT}(\alpha)$ each unit of time and a fraction θ of it is passed on, then there will be an extra cost, $\theta M^{RT}(\alpha)$, to be paid lumpsum by the buyer every unit of time (e.g., a lumpsum annual amount) and this amount will increase as α decreases.

The development of the one-time investment models particularly those for Models 2A and 2B, are more involved.

Model 2A:

Suppose that a one-time investment $M^{IU}(\alpha)$ is made by the supplier and the life of the investment depends on the number of units produced. Assuming that the investment lasts for a large number of units, i.e., it has an infinitely large life, and a fraction θ of the investment is passed on:

Amortised investment / year (assuming the unit of time to be a year) = $\theta M^{IU}(\alpha)i$.

The price increase per unit (since λ is the annual demand) = $\dfrac{\theta M^{IU}(\alpha)\, i}{\lambda}$.

After the investment, the unit purchase price will be: $c(\alpha\ ) = c + \dfrac{\theta M^{IU}(\alpha)\, i}{\lambda}$.

As α decreases, the buyer has to pay a price premium for early delivery. The holding cost/unit/time will now be: $h(\alpha) = i(c + \dfrac{\theta M^{IU}(\alpha)\, i}{\lambda})$. As c, i and λ are constants, $h(\alpha)$ has the same properties as $M^{IU}(\alpha)$, i.e., $h(\alpha) \geq 0,\quad h_{\alpha}(\alpha) \leq 0,\quad h_{\alpha\alpha}(\alpha) \geq 0$.

Hence, after the investment, the new long-run average total inventory cost for the buyer will be (from (4.5.13)):

$$C(Q, r, \alpha) = \frac{K\lambda}{Q} + h(\alpha)\{r + \frac{Q}{2} - \alpha\mu\} + \frac{(h(\alpha)+b)\alpha^2}{Q}[N(\frac{r}{\alpha}) - N(\frac{r+Q}{\alpha})] + \frac{h(\alpha)}{i}\lambda, \tag{4.6.2}$$

where $h(\alpha) = i(c + \dfrac{\theta M^{IU}(\alpha)\, i}{\lambda})$.

Note that for Models 1A and 2A, both the purchase cost and the holding cost are functions of the decision variable α. Hence, purchase cost must be a part of our "total inventory cost" unlike the traditional (Q, r) models where purchase cost is independent of decision variables and hence can be ignored for cost minimisation purpose.

Model 2B:

Suppose that the supplier has to make a one-time investment of $M^{IC}(\alpha)$ that will have a lifetime of T^C order cycles. Since each time period has (λ/Q) order cycles, the investment will have a lifetime of $N^C = T^C/(\lambda/Q)$ time periods. If the supplier wishes to pass on a cost of $\theta M^{IC}(\alpha)$, the equivalent cost per period will be (Fraser et al. 2000):

$$\frac{\theta M^{1C}(\alpha)i(1+i)^{N^C}}{(1+i)^{N^C}-1}. \tag{4.6.3}$$

If there are (λ/Q) > 1 cycles per period, then the cost per cycle for the supplier that will be passed on to the buyer will be:

$$\frac{[\theta M^{1C}(\alpha)i(1+i)^{N^C}]/[(1+i)^{N^C}-1]}{\lambda/Q}. \tag{4.6.4}$$

If N^C is large, then (4.6.4) becomes approximately $[\theta M^{IC}(\alpha)i]/(\lambda/Q)$ (Fraser et al. 2000). However, since there will be (λ/Q) cycles per unit time for the buyer, the "approximate" cost expression for the buyer will only have an extra term of $\theta M^{IC}(\alpha)i$.

If the approximation that N^C is large is not applicable, then the "exact" cost function of the buyer will have an extra term $[\theta M^{1C}(\alpha)i(1+i)^{N^C}]/[(1+i)^{N^C}-1]$ which itself will contain both the decision variables, α and Q. Note that we only show the "approximate" total cost expression in Table 4.6.1 and it will be taken as the cost expression for Model 2B unless otherwise indicated.

Model 2C:

The development of Model 2C is relatively straightforward. Suppose that a one-time investment $M^{IT}(\alpha)$ is made by the supplier and the life of the investment depends on the number of units produced. Assuming the investment has an infinitely large life and a fraction θ of the investment is passed on, the amortised investment / year (assuming unit of time to be year) = $\theta M^{IT}(\alpha)i$, and so the only effect on the cost function will be in the form of an extra term.

We note that in Models 1A and 2A, the additional cost factored into the purchase price affects the holding costs. Some might argue that in the remaining models the additional costs should also be factored into the purchase price and influence holding costs. However, for decision-making purposes it is necessary to emphasise the "engineering" aspects of the investment, and it will not be proper to roll the additional cycle or period costs into the effective purchase price (or acquisition cost) of the product.

4.7 Convexity Analysis

The buyer's objective is to minimise the cost function C in Table 4.6.1 by proper selection of the values of the three decision variables - Q, r and α. The optimal decision variable values for the buyer will be different depending on the type of the investment done by the supplier and how it is passed on to the buyer. Hence, it is important for the buyer to consider that issue before deciding on its optimal strategy.

In this section, we concentrate on investigating the convexity of the cost function. If proven so, first-order conditions (FOCs) will be both necessary and sufficient to determine the optimum decision variable values. Let us first analyse the cost equation (4.5.5):

$$C(Q, r, \alpha) = \frac{K\lambda}{Q} + (h+b)E(B) + h\{\frac{Q}{2} + r - \alpha\mu\} + c\lambda. \qquad (4.7.1)$$

Proposition 4.7.1: The cost function of (4.7.1) is jointly convex in Q, r and α.

Proof: In a (Q, r) policy, an order is placed whenever the IP reaches r. A backorder will occur only when the lead time demand exceeds the inventory position at which the order was placed and if it is less, then there will be no backorder. Hence, we can write, $E(B) = E_{\hat{X},IP}(max(\hat{X} - IP), 0)$.

For our case, $\hat{X} = \alpha X$ and IP is uniformly distributed over (r, $r + Q$). So we can substitute $IP = r + QU$ where U is uniform on (0, 1] and $\hat{X} = \alpha X$. To prove the joint convexity of $E(B)$ it is sufficient to show that max $\{(\alpha X - r - QU), 0\}$ is jointly convex in Q, r and α for fixed values of (X, U) (Zhang 1998). It is obvious that $(\alpha X - r - QU)$ is jointly convex in the three decision variables Q, r and α for fixed values of X and U. As max $\{g(.), 0\}$ is convex for any convex function $g(.)$, $E(B)$ is also jointly convex in Q, r and α. As the other terms of (4.7.1) are clearly convex, we can say that the cost function of (4.7.1) is jointly convex in Q, r and α. ■

While the basic inventory cost model with transformed LTD is jointly convex in Q, r and α, it is not clear what will happen when the investment cost is also taken into account (i.e., models of Table 4.6.1).

Proposition 4.7.2: Models 1C, 2C and 2B of Table 4.6.1 are jointly convex in the three decision variables, Q, r and α.

Proof: For Models 1C, 2C and 2B, the only difference in the cost function from (4.7.1) is an extra term. Since the investment function is by assumption convex in α, it is clear that the extra term, which is independent of Q and r, will be convex in α for all the three models. So, the cost equations for these three models are jointly convex in Q, r and α even after taking the investment cost into account. ■

For the "exact" Model 2B, we can show that the sufficient condition for joint convexity will be that $(M^{IC}_{\alpha\alpha})(M^{IC}) \geq (M^{IC}_{\alpha})^2$ (Appendix 4.2). This condition is satisfied by decreasing convex investment functions like power ($A\alpha^{-a}$, A, $a > 0$) and logarithmic ($A(1\text{-}\ln(\alpha))$, $A > 0$) used by Porteus (1985).

Proposition 4.7.3: A sufficient condition for joint convexity of Model 1B in Table 4.6.1 is $2(M^{RC}{}_{\alpha\alpha})(M^{RC}) \geq (M^{RC}{}_{\alpha})^2$.

Proof: Refer to Appendix 4.3. ■

Obviously this condition is also satisfied by decreasing convex investment functions like power ($A\alpha^{-a}$, A, $a > 0$) and logarithmic ($A(1\text{-}\ln(\alpha))$, $A > 0$).

Proposition 4.7.4: Models 1A and 2A of Table 4.6.1 are jointly convex in Q and r for a fixed α and also convex in α for fixed Q and r.

Proof: For joint convexity of Models 1A and 2A with respect to Q and r for a fixed α refer to Zheng (1992). For convexity with respect to α for fixed Q and r refer to Appendix 4.4. ■

Models 1A and 2A are not, in general, jointly convex in Q, r and α. While $E(I)$ is convex, and so is $h(\alpha) = ic(\alpha)$, the product of two convex functions may not be convex[4]. A sufficient condition for these mo dels to be convex is the joint convexity in Q, r and α of $h(\alpha)E(I)$ (since $E(B)$ is jointly convex in the three variables and $c(\alpha)$ is by assumption convex in α), i.e., joint convexity in Q, r and α of:

$$\frac{h(\alpha)}{Q}\int_r^{r+Q}\left[\int_0^{(s/\alpha)}(s-\alpha x)f(x)dx\right]ds. \tag{4.7.2}$$

To obtain more insights into the behaviour of the cost functions, we did extensive numerical experiments with different types of LTD distributions, e.g., exponential, logistic, and gamma, and different types of investment functions, e.g., power and logarithmic (details given in Section 4.9). From our numerical experiments, the cost

[4] Example: $f = -\log x$, $g = x^2$; while f and g are individually convex for all positive values of x, (fg) is not convex for all positive values of x. A sufficient condition for product of two convex functions to be convex is that both of them are monotone of the same sign. We cannot guarantee this in our case.

functions, while not necessarily convex, seems to be "univalleyed"[5] in the relevant region of the decision variables, i.e.,

$$\frac{1}{Q^*(\alpha)}\int_{r^*(\alpha)}^{r^*(\alpha)+Q^*(\alpha)}[\frac{d}{d\alpha}(\alpha G(\alpha,y))]dy+\frac{h_a\lambda}{i}=0, \tag{4.7.3}$$

where $Q^*(\alpha)$ and $r^*(\alpha)$ denote optimal Q and r for a given α, has a unique solution in the relevant region of α, $0 \leq \alpha \leq 1$ (Appendix 4.5). If C is "univalleyed", FOCs are both necessary and sufficient to determine the optimum values of the decision variables. If there is no solution, it implies that $\alpha^* = 1$ (as α tends towards 0, the cost will tend towards $\propto$, but is finite for $\alpha = 1$).

In the previous sections, the assumption was that the demand is constant and the variability in lead time demand is solely due to lead time duration variability. In that case if we know the demand rate (λ), and the mean (μ) and variance (σ^2) of the status-quo lead time demand ($\alpha = 1$), then from the optimal value of α, α^*, we can deduce the optimal mean and variability of lead time duration from the buyer's viewpoint (Appendix 4.6). However, all our previous analytical results and the insights in the following sections are based on the lead time demand distribution and not on the lead time duration distribution. Hence, even if the demand is random, we can proceed as before and determine the optimal value of α. When the demand is random, the variability in lead time demand comes from both demand and lead time duration and then the mean and the variance of the lead time demand is given by (Nahmias 1997):

$$\mu = \lambda\mu_L, \tag{4.7.4}$$

and

$$\sigma^2 = (\sigma_\lambda)^2\mu_L + \lambda^2(\sigma_L)^2, \tag{4.7.5}$$

[5] We use the term univalleyed in a cost minimisation in the same sense as unimodality is used for profit maximisation.

where μ and σ^2 are the respective mean and variance of the lead time demand distribution, μ and $(\sigma_L)^2$ are the respective mean and variance of the lead time duration distribution and λ and $(\sigma_\lambda)^2$ are the respective mean and variance of the demand distribution. As long as the reduction in lead time demand is *solely* due to lead time duration reduction and the mean and variance of the demand remains the same for all lead time, we can still deduce the optimal mean and variance of lead time duration (Appendix 4.6). Though the analysis in this chapter, including the numerical examples, is based on a constant demand, in fact, the model is much more general and can be used even with random demand. The exact values in the numerical examples might vary but the insights, we believe, will remain the same. However, note that when demand is random a change in only the mean lead time duration will change both the variance, as well as the mean, of the LTD.

4.8 Analytical Comparative Statics of Q^* and r^* with respect to α

A careful examination of the 6 models of Table 4.6.1 reveals that for constant i and λ, Models 1A and 2A are structurally similar (with $M^{RU}(\alpha)$ in Model 1A replaced by $\frac{M^{IU}(\alpha)i}{\lambda}$ in Model 2A), Models 1C and 2C are structurally similar (with $M^{RT}(\alpha)$ in Model 1C replaced by $M^{IT}(\alpha)i$ in Model 2C) and Models 2B and 2C are structurally similar ($M^{IT}(\alpha)$ in Model 2C replaced by $M^{IC}(\alpha)$ in Model 2B). Hence, we can restrict our detailed analysis to 3 basic models: 1A, 1B and 1C, which from now on we will refer to as "unit", "cycle" and "time" models respectively.

In this section we will perform the comparative statics of Q^* and r^* with respect to α for the unit, cycle and time models. In other words, we examine the effect of decreasing the supply lead time duration on the optimal reorder point and batch size. This issue is important from the buyer's perspective since any change in Q^* and r^* affects the buyer's operations directly, e.g., warehouse size, unloading dock design, manpower planning.

Proposition 4.8.1: For time and unit models, r^* and $(r^* + Q^*)$ decreases as α decreases.

Proof: Refer to Appendix 4.7. ■

It is known that any decrease in mean lead time demand only decreases r^*, but has no effect on Q^* and any decrease in variability of lead time demand reduces Q^* (De Groote and Zheng 1997). Based on these two effects and Proposition 4.8.1, we can conclude *that for the time and unit models, Q^* will also decrease as α decreases*. The effect of α on r^* and Q^* is intuitive since if α gets smaller, we would expect that the buyer has to hold less safety stock (since lead time demand will be less variable) and order less in each batch (since it will not take long for the supplier to deliver, there is no point of ordering more).

Proposition 4.8.2: For the cycle model, r^* decreases as α decreases but $(r^* + Q^*)$ may increase as α decreases.

Proof: Refer to Appendix 4.7. ■

For the cycle model, we are able to analytically show that while r^* decreases as α decreases, $(r^* + Q^*)$ may increase as α decreases. This implies that Q^* may increase as α decreases. This also makes sense if we think carefully about the cycle model: as α decreases, the effective set-up cost, $K(\alpha) = K + \theta M^{RC}(\alpha)$, increases. There are two opposite effects produced: a decrease of α will reduce Q^* (DeGroote and Zheng 1997) while an increase in $K(\alpha)$ will increase Q^*. Depending on which effect is stronger, Q^* may increase as α decreases. The effect of increased effective set-up cost can be so high that even $(r^* + Q^*)$ might increase as α decreases, though r^* itself will always decrease as α decreases.

Example 4.8.1: Suppose for the cycle model, the LTD has an exponential (β) distribution and the investment functions is of the power form. The parameters have the following values: $\theta = 1$, $K = 100$, $b = 18.75$, $c = 0.75$, $i = 0.05$, $A = 87$, $\lambda = 500$, $\beta = 0.001$. The values of $r^*(\alpha)$, $Q^*(\alpha)$ and $r^*(\alpha) + Q^*(\alpha)$ are plotted versus α in Figures 4.8.1(a), 4.8.2(b) and 4.8.3(c) respectively.

It is clear that while $r^*(\alpha)$ decreases with α, both $Q^*(\alpha)$ and $r^*(\alpha) + Q^*(\alpha)$ initially decrease and then increase with α.

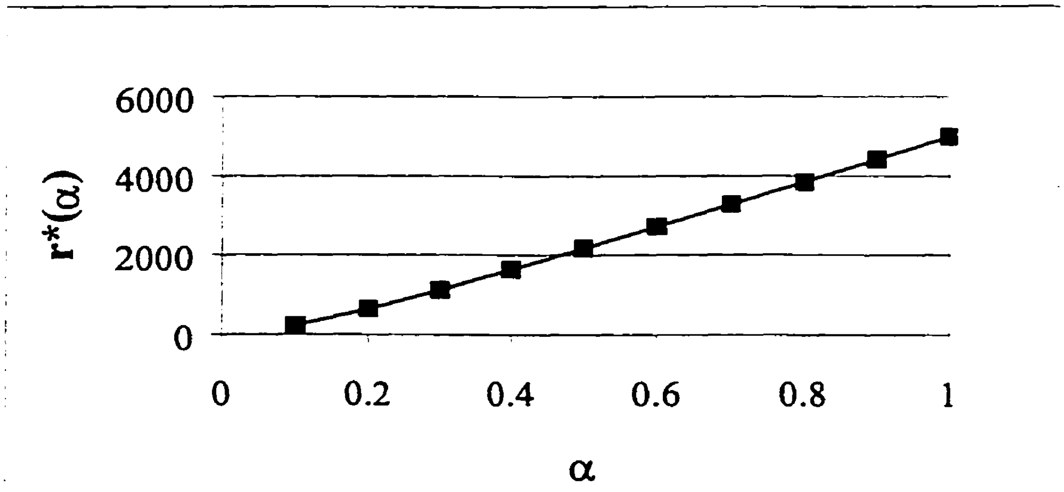


Figure 4.8.1(a): $r^*(\alpha)$ versus α for the Cycle Model

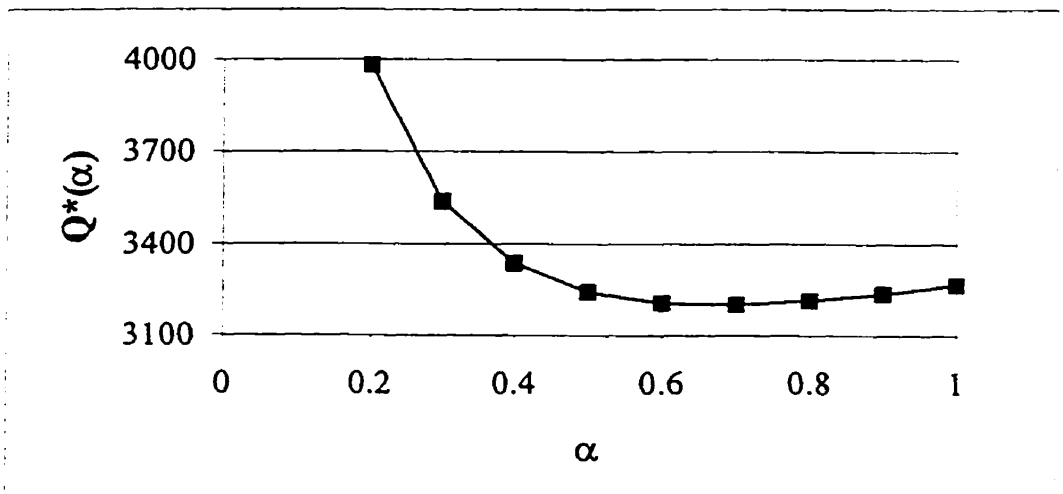


Figure 4.8.1(b): $Q^*(\alpha)$ versus α for the Cycle Model

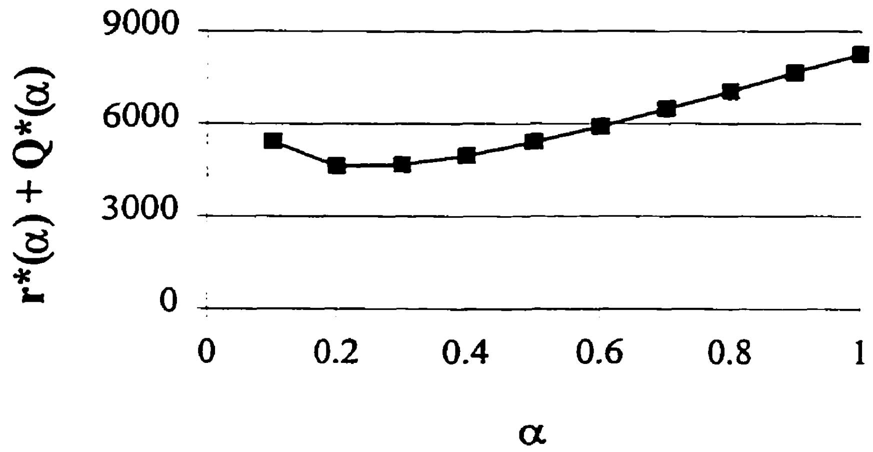


Figure 4.8.1(c): $r^*(\alpha) + Q^*(\alpha)$ versus α for the Cycle Model

This apparently counter-intuitive result is due to the fact that in our model the trade-off is between a reduction in lead time duration and an increase in set-up cost. As lead time (set-up time) is reduced, we can use that free time (capacity) to increase batch size and partly compensate for the increase in set-up cost. As is evident from Appendix 4.7, the effect of α for the cycle model will depend on the value of $\theta M^{RC}{}_{\alpha}(\alpha)$.

4.9 Numerical Examples

This section will report numerical examples explaining some of our previous assertions. Bagchi et al. (1986) suggests that the most common LT distributions are gamma, exponential and normal. We use the logistic distribution in place of normal as it approximates normal distribution quite accurately while the CDF and right-tail distribution can be obtained in closed form (refer to Gerchak and Parlar 1991 and references therein for more details). We primarily worked with 3 LTD distributions - exponential (β), logistic (μ, β) and gamma (2, β). We used two different investment functions for $M(\alpha)$ - Power, A/α, and Logarithmic, $A[1 - \ln(\alpha)]$ (for more details on the investment functions refer to Porteus 1985).

- **Cost Function of the Unit Model**

We proved the unconditional convexity of cycle and time models, but for the unit model we have conditions for the cost function to be convex or univalleyed. One of our first goals for the numerical experiments was to assess the shape of the cost function for the unit model over a wide variety of parameter settings and all combinations of LTD distributions and investment functions. For each value of α we calculated $Q^*(\alpha)$, $r^*(\alpha)$ and the corresponding $C(\alpha)$. Recall that C is jointly convex in Q and r for a fixed α even for the unit model. Then we plotted the cost function $C(Q^*(\alpha), r^*(\alpha), \alpha)$ versus α. In the following example, we plot $C(Q^*(\alpha), r^*(\alpha), \alpha)$ versus α for two combinations of LTD distribution and investment functions.

Example 4.9.1: In Figure 4.9.1(a) we plot $C(Q^*(\alpha), r^*(\alpha), \alpha)$ as a function of α for the unit model with an exponential (0.00667) LTD distribution, logarithmic investment function and the following parameter values: $\theta = 1$, $K = 15$, $b = 15$, $c = 1$, $i = 0.00085$, $A = 0.1$, $\lambda = 7$.

In Figure 4.9.1(b), the LTD for the unit model has a logistic (150, 20.7) distribution, the investment function is of the power form and the parameter values are as follows: $\theta = 1$, $K = 15$, $b = 22$, $c = 1$, $i = 0.025$, $A = 0.075$, $\lambda = 10$.

In both the above examples, the cost function is clearly univalleyed. Similarly, in all our numerical experiments, the cost function was univalleyed implying that there is an unique combination of Q, r and α that minimises C even for the unit model.

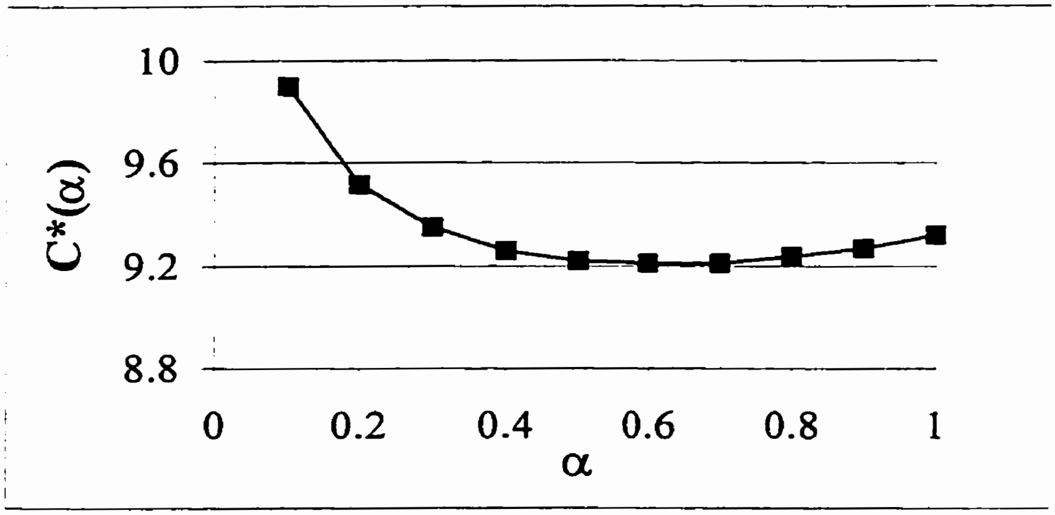


Figure 4.9.1(a): $C^*(\alpha)$ versus α for the Unit Model (a)
(Exponential LTD Distribution, Logarithmic Investment Function)

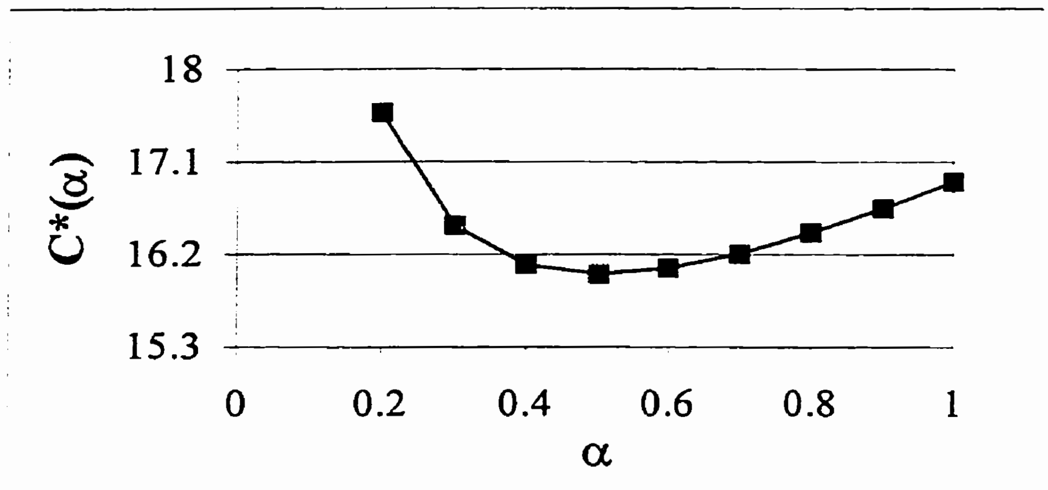


Figure 4.9.1(b): $C^*(\alpha)$ versus α for the Unit Model (b)
(Logistic LTD Distribution, Power Investment Function)

- **An Illustrative Example**

If the trivariate cost function $C(Q, r, \alpha)$ is convex or univalleyed, then the optimal values of the decision variables can be determined from the first order conditions (FOCs). For example, for the unit model, the cost function (4.5.16) can be written as:

$$C(Q, r, \alpha) = \frac{K\lambda}{Q} + \frac{\alpha}{Q}\int_{r}^{r+Q} G(\alpha, y)dy + c(\alpha)\lambda, \qquad (4.9.1)$$

where:

$G(\alpha, y) = (h(\alpha) + b)\int_{0}^{(y/\alpha)} F(x)dx + b\{\mu - (y/\alpha)\}$,

$h(\alpha) = ic(\alpha)$,

and

$c(\alpha) = c + \theta M^{RU}(\alpha)$.

For the cycle model, (4.5.16) will be:

$$C(Q, r, \alpha) = \frac{\{K + \theta M^{RC}(\alpha)\}\lambda}{Q} + \frac{\alpha}{Q}\int_{r}^{r+Q} G'(\alpha, y)dy + c\lambda, \qquad (4.9.2)$$

where:

$G'(\alpha, y) = (h + b)\int_{0}^{(y/\alpha)} F(x)dx + b\{\mu - (y/\alpha)\}$,

and

$h = ic$.

For the time model, (4.5.16) will be:

$$C(Q, r, \alpha) = \frac{K\lambda}{Q} + \frac{\alpha}{Q}\int_{r}^{r+Q} G'(\alpha, y)dy + c\lambda + \theta M^{RT}(\alpha). \qquad (4.9.3)$$

For the unit model the three FOCs are (by differentiating 4.9.1):

$$C_Q(Q, r, \alpha) = 0 \qquad \Rightarrow -\frac{K\lambda}{Q^2} - \frac{\alpha}{Q^2}\int_r^{r+Q} G(\alpha, y)dy + \frac{\alpha}{Q}G(\alpha, r+Q) = 0$$

$$\Rightarrow C(Q, r, \alpha) - c(\alpha)\lambda = \alpha G(\alpha, r+Q); \qquad (4.9.4)$$

$$C_r(Q, r, \alpha) = 0 \qquad \Rightarrow G(\alpha, r) = G(\alpha, r+Q); \qquad (4.9.5)$$

$$C_\alpha(Q, r, \alpha) = 0 \qquad \Rightarrow \frac{1}{Q}\int_r^{r+Q} G(\alpha, y)dy + \frac{\alpha}{Q}\int_r^{r+Q} G_\alpha(\alpha, y)dy + c_\alpha(\alpha)\lambda = 0.$$

(4.9.6)

Similarly we can also find the FOCs for cycle and time models by differentiating (4.9.2) and (4.9.3) respectively with respect to the decision variables - Q, r and α.

We used Maple to solve the three FOCs to obtain Q^*, r^* and α^*. We might also use some iterative technique by fixing the values of two decision variables at a time and solving for the third one until some convergence criterion is achieved (refer to Gerchak and Parlar 1991). As long as C is univalleyed, convergence is guaranteed.

Example 4.9.2: Let us take the case of unit model with the following parameter values at the status quo: $\theta = 1$, $K = 11$, $b = 16.125$, $c = 1$, $i = 0.00096$, $A = 0.075$, $\lambda = 178$.

Suppose also that at status-quo (i.e., $\alpha = 1$) the LTD is exponentially (β) distributed with a mean of 6250 units (i.e., $\beta = 0.00016$). If we assume the unit of time to be days, then the demand per day is 178 units (assumed constant) and the mean and standard deviation of lead time duration is approximately 35 days. The firm is paying \$1.075 per unit to the supplier, including the cost of maintaining the lead time duration at status quo. The optimal strategy for the firm at status quo will be: Q^* = 5241 and r^* = 57,916, i.e., whenever the inventory position at the firm's warehouse reaches 57,916 units, the firm will order 5241 units from the supplier. By following this inventory policy the firm will be incurring a cost of C^* = \$254.20 per day. However, the firm is not satisfied with the status quo. It has to hold almost 10.5 months of inventory as well as order 30 days of inventory at a time. Hence, the firm (buyer) wishes to reduce the supply lead time duration.

At the request of the buyer, the supplier looks at its operation and decides that it can reduce the lead time duration only by making some type of recurring per unit investment (e.g., extra labour to process each unit faster). Also the investment will be of logarithmic type, i.e., it will cost a fixed amount to reduce α by a fixed percentage (Porteus 1985). The supplier specifies to the buyer how much extra per unit the buyer has to pay to reduce lead time for various values of α. In this case, the buyer has to pay extra 4.83% per unit to bring α to 0.5, extra 9.67% per unit to bring α to 0.25 and so on. The buyer's manager takes that into account and uses our unit model to make a decision. The optimal decision variable values will be as follows: $Q^* = 3094$, $r^* = 11{,}412$ and $\alpha^* = 0.21$.

The manager of the firm should instruct the supplier to invest in lead time reduction and to reduce the mean and standard deviation of the lead time duration from 35 days to 7.35 days (more accurately the distribution itself will change; however, from managerial viewpoint this is easier to understand). The buyer is ready to pay a price premium of close to 11% per unit for this reduction. The buyer's optimal reorder point will be reduced to 11,412 units, i.e., approximately 64 days inventory, and the optimal order quantity will be approximately 17.5 days demand! Why is the buyer ready to pay the price premium? Even after paying the premium the buyer's optimal costs will be C^* = \$227.50 per day, almost 10.5% less than its status-quo optimal cost. Note that there is no point of urging the supplier to achieve "perfect delivery", at least as far as inventory cost is concerned. For example, a mean lead time duration of 3.5 days (with standard deviation of also 3.5 days, i.e., $\alpha = 0.1$) will require a price premium of almost 16%. Though the optimal reorder point will then be around 28.5 days stock and the optimal order quantity will be 14 days demand, the optimal cost will be almost 1.3% more than the overall optimal.

In the following table (Table 4.9.1) we show the individual cost elements (purchase, set up and holding + backordering), batch size, reorder point and cost for the unit model for this example at status-quo and at the optimal solution.

Table 4.9.1: Cost Elements for Example 4.9.2

	Status-quo ($\alpha = 1$)	Optimal ($\alpha = 0.21$)
Batch size (units)	5241	3094
Reorder point (units)	57,916	11,412
Purchase costs ($/day)	191.35	212.06
Set up costs ($/day)	0.37	0.63
(Holding + Backordering) costs ($/day)	62.48	14.81
Total cost ($/day)	254.20	227.50

At the status-quo, the purchase cost accounted for almost 75% of the total cost and the remaining 25% was from holding and backordering costs. In the example, we deliberately kept the set-up cost quite small. There are two reasons for this: (i) If set-up cost is high, then it is known that EOQ model will work quite well (Zheng 1992), and (ii) In recent times in retail environments if the transportation cost is taken to be a part of the purchase cost itself (i.e., transportation cost is paid for by the supplier), then the so called ordering cost has become very low. In the optimal solution, set-up cost has gone up, though as a percentage it is still quite small, and the purchase cost accounts for almost 93% of the total cost. Due to reduction in supply lead time duration, the backordering and holding costs now account for only about 7% of the total cost. We can easily develop "extreme" examples where the advantage from investing in supply lead time reduction can be much more significant.

What would have been the result if the buyer did not take into account that the investment done by the supplier is of the unit type and instead assumed that it is as in cycle or time models? With the same parameter values the optimal decision variable values for the *cycle* model will be: $Q^* = 2084$, $r^* = 5$ and $\alpha^* = 0.0003$, and the optimal decision variable values for the *time* model will be: $Q^* = 2033$, $r^* = 64$ and $\alpha^* = 0.0022$.

Note that the optimal values will be very different than those for the unit model. If the buyer follows any of these policies, the cost penalty will also be significant. For example, following the optimal values of the decision variables from the time model will make the cost almost 21% more than the optimal and following the cycle model will make it almost 33% more than the optimal. In practice, most probably the investment costs for the unit model will be very different than that of cycle or time models and hence chances for such mistakes will be rare. However, we want to point out here that it is important to consider the type of investment done by the supplier in lead time reduction to arrive at the optimal decision.

We can also develop similar examples for cycle and time models. An example of a situation where the cycle model is appropriate is when the buyer pays for the transportation cost separately to the supplier and the supplier feels that the only way to reduce supply lead time is by reducing the transportation time. Then the buyer might have to pay extra transportation cost per cycle, i.e., increased effective set-up cost. Similarly, an example of a situation where the time model is appropriate is when the supplier feels that it can reduce supply lead time by reducing the downtime of its machines. The buyer might pay an extra lumpsum amount per year to the supplier to have a better maintenance program but even after paying it, the buyer maybe better off.

Basically, there are two important insights that come out of the example:

a) If used properly, there can be substantial cost reduction from investment in supply lead time reduction. Even after paying a price premium (for the unit model) or a premium on the set-up cost (for the cycle model) or a lumpsum amount (for the time model), the buyer firm can decrease its inventory costs. However, the buyer must be careful before taking any decision. Depending on the parameter values, the amount of cost reduction might vary. If the cost of reducing the supply lead time is high, i.e., if A is high, then it might be that the status-quo is optimal. Similarly there might also be other situations where investment in supply lead time reduction might not make much sense. However, all these will come out of the analytical model we have built. Our model will help managers in deciding when to go for supply lead time reduction and if lead time reduction is necessary then what should be the optimal supply lead time, and associated optimal investment and inventory levels;

b) Our models also show that the managers must be very careful about correctly modelling the supplier's investments in lead time reduction and which model to use (i.e., unit or cycle or time) while deciding on the optimal decision variable values.

Note that these models are very much in tune with the recent phenomenon of focusing on "total" inventory cost rather than just purchase cost (Australian National Audit Office Report 1997-98; Purchasing Online 1998). Our models are also consistent with the JIT philosophy since we reduce the non-value added supply lead time (Stevenson 1999).

- **Effect of Investment Functions**

We investigated the effects of the investment functions on the model. We used logarithmic ($A[1 - \ln(\alpha)]$) and power (A/α) investment functions for all three LTD distributions. For both time and unit models it is seen that for all values of α, $C(Q^*(\alpha), r^*(\alpha), \alpha)$ is lower for the logarithmic investment than for the power and Q^*, r^*, α^* and C^* for the power investment are greater than that for the logarithmic investment (for the same A).

Example 4.9.3: In Figure 4.9.2 we plot $C(Q^*(\alpha), r^*(\alpha), \alpha)$ versus α for the unit model for exponential (0.00016) LTD distribution and the following parameter values for both the investment functions: $\theta = 1, K = 15, b = 15, c = 1, i = 0.002, A = 0.1, \lambda = 250$.

The optimal decision variable values and costs for the two investments are as follows:

Logarithmic: $Q^* = 2901, r^* = 9773, \alpha^* = 0.2, C^* = 344.6$.

Power: $Q^* = 3843, r^* = 24{,}231, \alpha^* = 0.48, C^* = 366.86$.

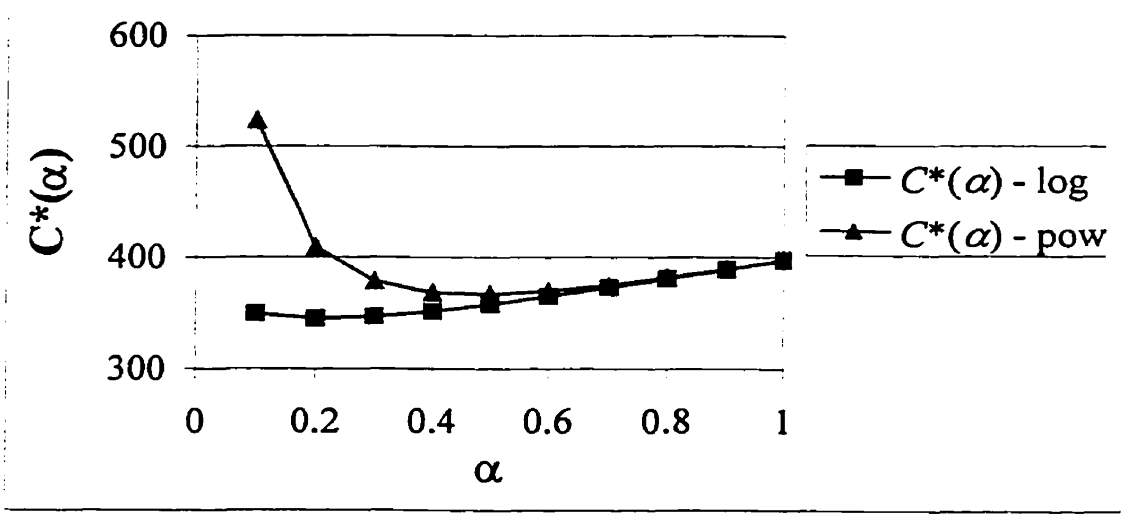


Figure 4.9.2: $C^*(\alpha)$ versus α for Different Investment Functions (Exponential LTD Distribution, Unit Model)

The above result is quite intuitive since for all values of α the investment for logarithmic is less than the investment for power, except at $\alpha = 1$ when they are equal. This leads to lower α^* and hence lower r^*, Q^* and C^*. For the same reason we noted that for the cycle model, while r^*, α^* and C^* follow a similar pattern, Q^* for the power investment function might be less than Q^* for the logarithmic investment function. Recall that lower α^* might lead to higher Q^* for the cycle model.

Example 4.9.4: For the cycle model, suppose that the LTD has an exponential (0.00016) distribution with the following parameter values: $\theta = 1$, $K = 15$, $b = 15$, $c = 1$, $i = 0.002$, $A = 10000$, $\lambda = 250$.

The optimal decision variable values and costs for the two investments are then as follows:

Logarithmic: $Q^* = 70{,}839$, $r^* = 14{,}988$, $\alpha^* = 0.43$, $C^* = 416.34$.

Power: $Q^* = 66{,}212$, $r^* = 24{,}944$, $\alpha^* = 0.65$, $C^* = 424.20$.

In this case while r^*, α^* and C^* are lower for the logarithmic investment than the power, Q^* is higher.

Example 4.9.5: Consider another example for the cycle model where the LTD has a gamma distribution (2, 3125) with the following parameter values: $\theta = 1$, $K = 15$, $b = 15$, $c = 1$, $i = 0.002$, $A = 500$, $\lambda = 250$.

The optimal decision variable values and costs for the two investments are then as follows:

Logarithmic: $Q^* = 19{,}972$, $r^* = 3039$, $\alpha^* = 0.13$, $C^* = 294.35$.

Power: $Q^* = 20{,}494$, $r^* = 8787$, $\alpha^* = 0.34$, $C^* = 304.32$.

In this case all Q^*, r^*, α^* and C^* are lower for the logarithmic investment than the power.

- **Effect of Neglecting α for the Holding Cost in the Unit Model**

As we noted before, while time and cycle models are convex, the unit model is difficult to be proven convex or univalleyed. Convexity does hold for a unit model if we assume that while purchase cost will increase as α decreases, there will be no effect on the holding cost (as in Gerchak and Parlar 1991 and Bookbinder and Çakanyildirim 1999). However, we feel that it is important to take into consideration the effect on holding cost, as otherwise the model would be underestimating the inventory cost and lead to wrong operational decisions. Sometimes this underestimation can be quite high.

Example 4.9.6: Suppose for the unit model the LTD has an exponential (0.00016) distribution and the investment is of the logarithmic form with the following parameter values: $\theta = 1$, $K = 15$, $b = 15$, $c = 1$, $i = 0.004$, $A = 0.75$, $\lambda = 250$.

The optimal decision variable values and costs are then as follows:

$Q^* = 3033$, $r^* = 34{,}866$, $\alpha^* = 0.77$, $C^* = 771.72$.

However, if we solve the model using the same parameters but ignoring the investment's effect on holding cost, the optimal decision variable values and costs will then be as follows:

$Q^* = 2834$, $r^* = 25{,}557$, $\alpha^* = 0.56$, $C^* = 736.13$.

For this example the underestimation in cost is almost 4.5% (we can easily construct more "extreme" examples). Note that when the effect on holding costs is not taken into account, then more investment is done, and both the optimal batch size and reorder point will be lower. In this case batch size is almost 6.5% lower and reorder point is almost 26% lower. Ignoring the effect on holding cost can also lead to the decision of instructing the supplier to invest in lead time reduction, when really the status-quo is optimal. Hence, it is very important to take into consideration the effect on the holding cost along with that on the purchase cost for the unit model.

- **Approximations**

Even approximations like (i) $E(IL) \cong E(I)$ (assuming backordering time to be negligible as has been used by Nahmias 1997), or (ii) one order outstanding (like Bookbinder and Çakanyildirim 1999) for the unit model might not produce a convex cost function. Note that the first approximation will always underestimate the "exact" model's cost while the second approximation will always overestimate it. The only way we can prove those approximations to be jointly convex in the three variables easily is by assuming the holding cost to be independent of α. However, as we have already indicated, under the assumption of holding cost being independent of α we can prove the convexity of even the "exact" model. Hence, these approximations would not be of much use here.

The one possible approximation that can be employed is to use the optimal decision variable values from the model assuming that holding cost is independent of α but the purchase cost is a function of α, which is provably convex, in unit models. Though we noted before that such model by itself underestimates the true cost, using the decision variable values from that model in the "true" unit model can act as an approximation. In Example 4.9.6 such approximation will result result in a cost of 776.87, only about 0.66% higher than the optimal cost. From our numerical experiments, again with all combinations of the three LTD distributions and two investment functions, we can say that this approximation will work fairly well unless A or i are very high. This is intuitive, since when A or i are high the effect of holding costs will be significant. In passing, we would also like to mention that the approximation of one order outstanding performs poorly for highly variable LTD distributions and low K. This is in line with Zipkin's (1986b) assertions for a standard (Q, r) model without α. In that sense, examples with exponential LTD with one order outstanding assumption as in Bookbinder and Çakanyildirim (1999) are not advisable. We also noted that one-order-outstanding approximation performs poorly for low values of b. This is intuitive since high values of b will naturally lead to one order being outstanding and so low values of b exposes the model to problems.

4.10 Numerical Comparative Statics

FOCs for the exact (Q, r) model, even without α, are very involved and it is not possible to obtain closed form solutions for the optimal decision variables. But De Groote and Zheng (1997) and Zipkin (200) have performed comparative statics of an exact (Q, r) model. While it might be possible to do some analytical comparative statics for our trivariate cost model, it would be quite complex. Therefore, we report only numerical comparative statics for the parameters.

For our numerical comparative statics, we also used three LTD distributions (logistic, gamma and exponential) and two investment functions (power and logarithmic). The focus is on the effect on optimum decision variable values rather than cost. Following are our results for the different parameters involved (in the following ↑ denotes increases, ↓ denotes decreases and ↑\↓ denotes that it might increase or decrease):

- **System Cost Parameters**

- ***Set-up Cost (K)***

As K ↑, for unit and time models $\rightarrow Q^*$ ↑, r^* ↑\ ↓, α^* ↑;

for the cycle model $\rightarrow Q^*$ ↑, r^* ↓, α^* ↓.

As K increases, as expected Q^* increases for time and unit models. At the same time we reduce our investment in α to partly compensate for possible increase in holding cost (so α^* increases). This effect on α is quite intuitive. As Q^* increases, the number of orders will decrease and so the frequency of the inventory level reaching the reorder point will decrease. This will create an incentive for reduced investment in α. In the traditional (Q, r) model, it can be shown that as K ↑, while Q^* ↑, r^* ↓ (Zheng 1992). However, in our model, under certain circumstances, r^* may increase as K increases.

Example 4.10.1: Suppose for the unit model the LTD is exponentially distributed (0.00016), the investment function is logarithmic and the parameter values are as follows: $\theta = 1, b = 15, c = 1, i = 0.002, A = 0.1, \lambda = 250$.

For $K = 15$, the optimal decision variable values are: $Q^* = 2901$, $r^* = 9773$, $\alpha^* = 0.20$.
For $K = 40$, the optimal decision variable values are: $Q^* = 4149$, $r^* = 9651$, $\alpha^* = 0.21$.

Example 4.10.2: Suppose for the unit model the LTD has a logistic (1000, 150) distribution, the investment function is logarithmic and the parameter values are as follows: $\theta = 1, b = 50, c = 5, i = 0.00125, A = 1, \lambda = 5$.

For $K = 15$, the optimal decision variable values are: $Q^* = 207$, $r^* = 1053$, $\alpha^* = 0.49$.
For $K = 40$, the optimal decision variable values are: $Q^* = 300$, $r^* = 1068$, $\alpha^* = 0.51$.

While for Example 4.10.1, r^* decreases with increase in K, for Example 4.10.2, r^* increases with increase in K.

The reason for this apparently counter-intuitive behaviour of r^* is that in some cases, the increase in Q^* alone is not enough to counterbalance the increase in α^* and so r^* also needs to increase to have more safety stock. This is contrary to what Bookbinder and Çakanyildirim (1999) reported on the effect of K on r^*, i.e., they reported that even in the presence of α, as K increases, r^* will always decrease.

The effect for the cycle model is also intuitive; increasing K leads to an increase in Q^* and if Q^* is increasing due to K, we would expect α^* to decrease. The decrease of α^* will just have an added impact on Q^* and we can use the reduced α^* to decrease r^*.

Note that as K increases, α^* increases for the time and unit models. This implies that there will be some threshold value of K which will make $\alpha^* = 1$ and for any K larger than the threshold value, the status-quo will be optimal. We think that this result is important since it shows that when K is sufficiently large then there is no point in reducing supply lead time; rather the focus should be on set-up time (cost) reduction (refer to Chapter 5). As both Zheng (1992) and Zipkin (2000) pointed out, when the ordering cost is large and variability is low (i.e., EOQ type models) the more important trade-off is between the set-up cost and the holding cost. In that case, the focus should be on reducing the set-up cost.

- ***Backordering cost/unit/unit time (b)***

As $b\uparrow$,	for unit and time models	$\rightarrow Q^*\downarrow, r^*\uparrow/\downarrow, \alpha^*\downarrow$;
	for the cycle model	$\rightarrow Q^*\uparrow, r^*\uparrow/\downarrow, \alpha^*\downarrow$.

It can be shown that for a traditional (Q, r) model as b increases, r^* and $(r^* + Q^*)$ must increase (De Groote and Zheng 1997), while the effect on Q^* alone is not obvious.

However, if we think carefully, the apparently counter-intuitive effect of b on r^* we report is not so strange. If b increases, there are three ways of compensating - reduce α^*, increase r^*, or both. While in most cases both occur, if the investment cost is very small and/or holding cost is high, it is better to reduce α^* only than to increase r^*. Our model will first try to see whether the cost of changing α is small or not. If it is inexpensive then it will try to lower α^* as much as possible which can lead to reduced r^* and Q^* and hence savings in holding costs. Obviously, sometimes it may be necessary to reduce α^* and increase r^* simultaneously. In all cases, α^* will decrease and the model will compensate by reducing Q^* to save on holding costs.

That this apparently counter-intuitive result makes sense can also be seen from the following argument: Let us assume that at status-quo the backordering cost is significant. This will imply that r^* will be greater than the mean lead time demand, i.e, there will be some safety stock. Now suppose the backordering cost increases, but is finite and the cost of changing α is zero. In that case, the optimal strategy will be the EOQ model with backordering. For the EOQ model with backordering, it is well known that r^* is always less than the mean lead time demand as long as the backordering cost is finite (Zipkin 2000). Hence, it is clear that if the cost of changing α is zero, then even as b increases, r^* will decrease. Only when the cost of changing α is greater than some threshold value will r^* start increasing with b.

Example 4.10.3: Suppose the LTD distribution is gamma (2, 3125) and the investment function is of the power form for the unit model and the parameter values are as follows: $\theta = 1, K = 15, c = 1, i = 0.001, A = 0.1, \lambda = 250$.

For $b = 5$, the optimal decision variable values are: $Q^* = 5058, r^* = 28{,}317, \alpha^* = 0.90$.
For $b = 40$, the optimal decision variable values are: $Q^* = 4862, r^* = 31{,}188, \alpha^* = 0.81$.

Example 4.10.4: Suppose the LTD distribution is exponential (0.00016) and the investment function is logarithmic for the unit model and the parameter values are as follows: $\theta = 1, K = 15, c = 1, i = 0.002, A = 0.1, \lambda = 250$.

For $b = 15$, the optimal decision variable values are: $Q^* = 2901$, $r^* = 9773$, $\alpha^* = 0.20$.
For $b = 50$, the optimal decision variable values are: $Q^* = 2775$, $r^* = 9743$, $\alpha^* = 0.17$.
While for Example 4.10.3, r^* increases with increase in b, for Example 4.10.4, r^* decreases with increase in b!

For the cycle model, as b increases, α^* decreases. Decreased α^* increases the effective set-up cost for this model and so Q^* increases. Since Q^* is increasing, so there might not be any further need to raise r^* too! Only if the increase of Q^* is not enough to compensate for the increased b, it will be required to increase r^* also. Hence, in this case also, as b increase, r^* can either increase or decrease.

Example 4.10.5: Suppose the LTD distribution is gamma (2, 3125) and the investment function is of the power form for the cycle model and the parameter values are as follows: $\theta = 1, K = 15, c = 1, i = 0.002, A = 500, \lambda = 250$.

For $b = 5$, the optimal decision variable values are: $Q^* = 19{,}835$, $r^* = 8295$, $\alpha^* = 0.37$.
For $b = 40$, the optimal decision variable values are: $Q^* = 21{,}057$, $r^* = 9185$, $\alpha^* = 0.32$.

Example 4.10.6: Suppose the LTD distribution is exponential (0.00016) and the investment function is of the logarithmic form for the cycle model and the parameter values are as follows: $\theta = 1, K = 15, c = 1, i = 0.002, A = 10000, \lambda = 250$.

For $b = 5$, the optimal decision variable values are: $Q^* = 67{,}961$, $r^* = 15{,}203$, $\alpha^* = 0.51$.
For $b = 40$, the optimal decision variable values are: $Q^* = 73{,}084$, $r^* = 14{,}806$, $\alpha^* = 0.37$.

While for Example 4.10.5, r^* increases with increase in b, for Example 4.10.6, r^* decreases with increase in b!

- ***Holding cost/unit/unit time (h)***

  As we have already defined $h = ic$, the effect of h can be decomposed into two parts: i) effect of i, and ii) effect of c.

  As $i / c \uparrow$, for unit, cycle and time models $\rightarrow Q^* \downarrow, r^* \downarrow, \alpha^* \downarrow$;

The effect of i and c on the models is intuitive. If i or c is higher, holding cost per unit will increase. Hence we would want Q^* and r^* to be lower, to save on the holding costs. This would require the model to reduce α^*.

However, in the case of Models 2A and 2C, the effect of i and c can be quite different. This is clear from the expression of the cost functions for those two models. In those models, if i increases, we would want the investment to be lower (implying α^* will be higher) and also we would want to hold less stock to reduce inventory holding costs. This would lead to reduced Q^* and also have a downward effect on r^*. However note that increased α^* itself will try to increase r^*. Hence, depending on which effect is stronger, r^* will either increase or decrease. On the contrary, when c increases there is no need to reduce investment and so then α^* will reduce to lower Q^* and r^* so that holding costs can be reduced.

- **Investment Cost Parameters**

- ***Investment Cost (A)***

  As $A \uparrow$, for unit and time models $\rightarrow Q^* \uparrow/\downarrow, r^* \uparrow, \alpha^* \uparrow$;

  for the cycle model $\rightarrow Q^* \uparrow, r^* \uparrow, \alpha^* \uparrow$.

Increase of A implies that more investment will be needed to change α. Hence, we would not want to decrease α much and so α^* increases for unit and time models. This increase in α^* will cause an increase in r^* also in these models. For time and unit models, while in most cases Q^* will increase with A to take care of the higher α^*, there are instances (e.g., exponential LTD distribution and logarithmic investment function) when Q^* decreases as A increases. This happens since the increase in r^* itself is enough to "handle" the excess supply lead time and it makes further increase in Q^* superfluous.

For the cycle model, an increase in A will also lead to an increase in α^* to reduce the amount of investment. Increased α^* will lead to increase in r^*, as in the unit and time models. For the cycle model, any increase in α^* will decrease the effective set-up cost that will try to decrease Q^*. On the other hand, increased α^* will try to increase Q^*. From our numerical experiments it seems that the latter effect will always dominate the former and so for the cycle model as A increases, Q^* will tend to increase.

- ***Fraction of Investment cost passed on*** **(θ)**

As $\theta\uparrow$,	for unit and time models	$\rightarrow Q^* \uparrow/\downarrow$, $r^* \uparrow$, $\alpha^* \uparrow$;
	for the cycle model	$\rightarrow Q^* \uparrow$, $r^*\uparrow$, $\alpha^* \uparrow$.

As θ increases, the supplier passes on more of the investment to the buyer. Hence, as expected, the effect of an increase in θ on the optimal decision variable values is similar to that of increase in A for all the models.

- **Demand Rate (λ)**

As $\lambda\uparrow$,	for unit, cycle and time models	$\rightarrow Q^* \uparrow$, $r^* \uparrow$, $\alpha^*\downarrow$.

This effect is more complex to explain than the other parameters. As λ increases, both μ and σ for LTD also increases (i.e., the distribution of the LTD itself changes). The relation is given by: $\mu = \lambda\mu_L$ and $\sigma^2 = \lambda^2\sigma_L^2$ as shown in (4.7.4) and (4.7.5) respectively. In this case, Q^* and r^* increases to take care of the increased demand while α^* decreases to at least partly "control" the increase of Q^* and r^*. In our case we are also trying to reduce α.

So there are two opposite effects acting on the LTD distribution simultaneously - an increase of λ and decrease of α. Though from our numerical experiments we noticed the above effects on Q^* and r^*, we would expect that the ultimate effect on Q^* and r^* will depend on the overall effect of increasing λ and decreasing α on the LTD. A detailed investigation of the effect of λ on the decision variables for a traditional (Q, r) model can also be found in Zipkin (2000) and De Groote and Zheng (1997).

The decrease of α^* with λ implies that there is a threshold value of λ below which the staus-quo will be optimal for any demand rate (i.e., $\alpha^* = 1$). This is intuitive, since if there is not enough demand, then there is no point in investment in supply lead time reduction. Such investments make sense only above a certain demand level. This result is also in line with the assertion of Porteus (1985) regarding investment in set-up cost reduction where he shows that such investments make sense only above a certain critical demand rate.

Example 4.10.7: Suppose the LTD distribution is exponential and the investment function is logarithmic for the unit model and the parameter values are as follows (remember in this case as λ will change, the LTD distribution will also change): $\theta = 1$, $K = 15$, $b = 5$, $c = 1$, $i = 0.001$, $A = 0.2$, $\mu_L = 25$ and $\sigma_L = 25$.

In this case for $\lambda \leq 7.25$ (approximately), $\alpha^* =1$, i.e., the demand rate must be more than 7.25 units per unit time for any investment in supply lead time reduction to make sense.

Example 4.10.8: Suppose the LTD distribution is exponential and the investment function is logarithmic for the cycle model and the parameter values are as follows: $\theta = 1$, $K = 15$, $b = 10$, $c = 1$, $i = 0.001$, $A = 10000$, $\mu_L = 25$ and $\sigma_L = 25$.

In this case for $\lambda \leq 130$ (approximately), $\alpha^* =1$, i.e., the demand rate must be more than 130 units per unit time for any investment to make sense.

While the effect of λ on the decision variables is important, we feel that a more interesting issue to investigate is how the optimal cost per unit (C^*/λ) changes with λ. It is not surprising that a bigger company (i.e., higher demand) has higher total cost or total profit. Though total cost or profit can be a measure of the size of the company, it cannot be a measure of efficiency (cost per unit) or profitability (profit per unit). While this has been recognised in ex-post financial analysis of firm's performance (e.g., Return on Investment, Earnings per share), we have not been able to locate any significant research on this in ex-ante stochastic operations management models. Most of the models have total cost/profit as objective rather than cost/profit per unit. However, recently Gerchak et al. (2000) have shown that in a newsvendor framework the optimal decision variable values for a ratio objective can be very different from absolute objectives and using one for the other can result in significant losses.

We plotted (C^*/λ) versus λ for our different models. In all cases, we found that there is a decreasing convex relation between (C^*/λ) and λ. We have plotted two examples in Figures 4.10.1 and 4.10.2 - one for an exponential LTD distribution and logarithmic investment function for the unit model (Figure 4.10.1) and another for exponential LTD distribution and logarithmic investment function for the cycle model (Figure 4.10.2). The decreasing convex relation is clear in both cases. This type of relation is present for all models for all LTD distribution and investment function combinations signifying that there are decreasing economies to scale in inventory costs. Zipkin (2000) notes that for a traditional (Q, r) model, C^* is increasing in σ, becoming nearly linear for large σ (i.e., for large λ, since in our model λ is proportional to σ). But for us α is a decision variable and we are investing in reducing it and reduction of α will also reduce σ. Hence, we would expect that it will require a larger λ than that in traditional (Q, r) model for C^* to be linear. In Figure 4.10.2 we plot (C^*/λ) at $\alpha = 1$ for exponential LTD distribution and logarithmic investment function for the cycle model and compare it to the overall optimal model. It is clear that when we invest in reducing α, the economies of scale are "more" than from traditional (Q, r) model alone. But the decreasing "convex" optimal cost per unit signifies that even here economies of scale will disappear for "high"(er) λ.

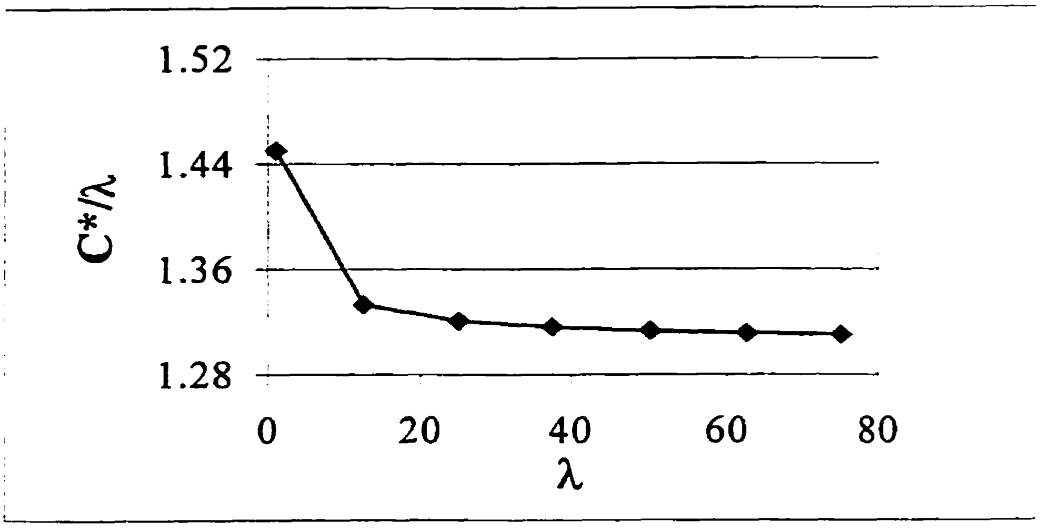


Figure 4.10.1: ***C*/λ*** **versus** ***λ*** **for the Unit Model**

(Exponential LTD Distribution, Logarithmic Investment Function)

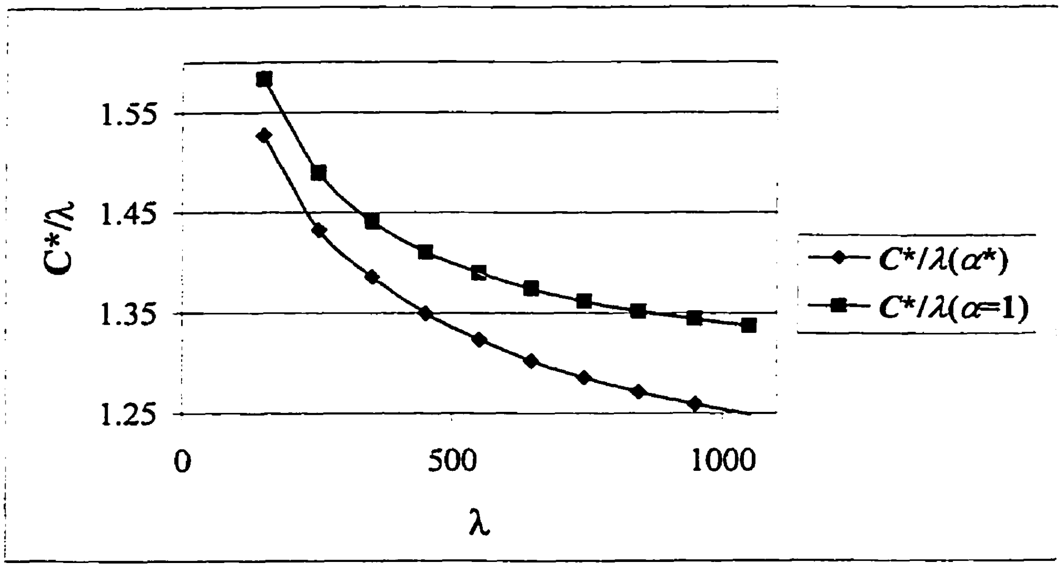


Figure 4.10.2: ***C*/λ*** **versus** ***λ*** **for the Cycle Model at Optimal** ***α*** **and at** ***α*** **= 1**

(Exponential LTD Distribution, Logarithmic Investment Function)

Even with its inherent limitations regarding the number and range of experiments, the above numerical comparative statics make one thing clear - our intuitive reasoning and even some results pertaining to traditional (Q, r) models might fail when the additional variable α is introduced. As is evident, the cost of changing α, the type of the model and the LTD distribution plays a big role. Since, in our trivariate model, all the three decision variables are inter-related, depending on the cost of changing α, the model will adjust the optimal decision variable values and this adjustment for Q^* and r^* can be very different from traditional models. But this interaction of three decision variables will also render analytical comparative statics even more difficult than for existing (Q, r) models.

4.11 Conclusions and Future Research Opportunities

In this part of the research, we showed analytically and numerically the effects of investments in supply lead time reduction in a two-party make-to-stock supply chain. We used a continuous review (Q, r) model and incorporated the effects of reduction in lead time duration on the model through a single variable, α. We also took into account the cost (in the form of some investment) associated with this lead time reduction. In practice, investments can be one-time or recurring. It might also vary depending on the nature of the reduction needed (e.g., per unit or per cycle or per unit time). Our overall trivariate model captured both the costs and benefits of lead time reduction for the different types of possible investment.

The cost models were analysed in detail as to their convexity. Numerical experiments with different LTD distributions and investment functions were performed to obtain a better feel of the behaviour of the model. The numerical experiments clearly show not only the benefits of investment in supply lead time reduction but also the importance of taking into consideration the type of the investment done by the supplier and how costs are passed on to the buyer before deciding on the optimal strategy. The complexity of the cost models render analytical comparative statics difficult. We performed extensive numerical comparative statics, which showed that the three decision variables are very much interdependent. Some of the results are quite counter-intuitive.

These new models have both technical and managerial contributions. From a technical standpoint, we incorporated a new variable which represents the supply lead time in the traditional (Q, r) model. There is a cost of reducing the supply lead time that depends on the type of the investment done by the supplier. This created six analytical models and we were able to perform the convexity and sensitivity analysis of those models. From our analysis we were able to show that not only investments in supply lead time reduction can result in significant savings but also that these models are "different" from traditional (Q, r) models. Our analysis showed results that are seemingly counter-intuitive but makes perfect sense in the new model. We were able to explain the reasons behind such counter-intuitive results.

From a managerial perspective, we were able to show that it makes sense to invest in supply lead time reduction. Even when the buyer is "paying" for supply lead time reduction it can give substantial benefits. However, managers should not blindly go for supply lead time reduction. Our models will help them in deciding when investment in supply lead time reduction makes sense and when not to invest in such reductions. Our models will also help decide their "optimal strategy" both in terms of inventory (reorder point and batch size) and time (supply lead time) that will minimise their inventory costs. In this age of large-scale outsourcing, we feel that our models can help supply chain managers reduce inventory costs and give them a competitive advantage in the marketplace.

In terms of scope for future research, it would be nice to prove analytically that the unit model is univalleyed. Similarly, analytical comparative statics will be better than numerical results. But a trivariate model without closed form solutions for any of the decision variables (as ours is) can make such analysis extremely difficult.

One of the comparative statics that we think is particularly interesting is the consequence of varying λ. We noted that (C^*/λ) is decreasing convex with respect to λ. This opens up an interesting avenue of research regarding the effect of inventory costs on market demand. Higher demand leads to reduced unit costs. In case of make-to-stock products, the unit cost is often related to unit price by mark-up pricing. If investment in supply lead time reduction brings down the unit cost, it will also lead to reduced unit price. If the buyer is willing to pass on the reduction in cost due to lead time reduction to the final customer (refer to Figure 4.2.1), then demand for make-to-stock products will increase (since price-sensitive). So we can expect that a chance for increased demand will provide more incentive for investment. We feel that an inventory model with demand being a function of cost by its effect on price can be very interesting and give a much better picture of the entire supply chain. One of the main goals of any cost reduction initiative is to achieve a larger market share by increasing demand. Hence, there is a need to couple the inventory models and the relevant marketing models to establish market effectiveness of increased speed. In the next chapter we will address this issue.

CHAPTER 5

LEAD TIME MANAGEMENT FOR MAKE-TO-STOCK FIRMS

5.1 Introduction

In this chapter we will develop an integrated production-marketing framework which will allow profit-maximising make-to-stock firms to determine their optimal lead time taking into account the costs and benefits associated with lead time reduction (Model C of Chapter 1). In the first part of the chapter, we will focus on a general modelling framework that forms the basis of the latter part of the chapter where we develop specific models.

Recent studies in such diverse fields as health care (Connor et al. 1998) and grocery industries (McGoldrick 1993) have shown that cost reduction initiatives by firms have led to price reductions and, in turn, increased market share due to customers' price-sensitivity. Growth in the Japanese share of the US market in the automobile industry (Nanto, Cooper and Bass, 1995), and in electronic goods, long thought to be the domain of US manufacturers (Zipkin 1991), can be partly attributed to the lower Japanese prices. The price advantage stemmed largely from lower production costs (Rao 2000a). For example, in the early 1980s, production costs of Japanese firms were about 25% lower than their US counterparts. This allowed the Japanese firms to charge a lower selling price and capture a larger market share (National Research Council Report 1983).

Fisher (1997) contends that for any "functional" product, one that has a reasonable life cycle and fairly stable demand, competition is typically very high, i.e., the profit margin is low and customer's price sensitivity is high. As we indicated in Chapter 1, these types of market characteristics are normally seen in make-to-stock products. For such products, customers are primarily price-sensitive; hence, the firms aim for cost leadership focusing on maximising efficiency and minimising costs in the supply chain (also refer to Chopra and Meindl 2001).

Encouraged by the Japanese experience, many make-to-stock companies have invested in process improvements to reduce costs, the most successful improvement strategies involving capital investments (Zipkin 1991). Cachon and Fisher (1999) find that the benefits from cost-reducing process improvements can be significantly more valuable than those derived from information sharing.

Since inventory costs comprise a large portion of the total operating costs of many make-to-stock firms, they have focussed specifically on inventory cost reduction as a primary means of reducing operating costs. It is well known that reduction of either internal or supply lead time can result in reduced inventory costs (Karmarkar 1993; Zipkin 2000; Chapter 4 of this thesis). There are several anecdotal examples in recent literature that show that one of the most popular process improvement techniques used by firms to reduce their inventory cost is lead time reduction (Suri 1998; Simchi-Levi et al. 2000). Note that while lead time reduction can yield many benefits, here we are focusing mainly on the effects of lead time reduction on inventory costs.

Mark-up pricing is a strategy used for make-to-stock products in the manufacturing sector (Bloch and Oliver 1997), internet pricing (Wilson 2000), apparel industry (San Francisco Fashion Industry Report 2000) and retail industry (Wang and Zhao 2000). In mark-up pricing, price is based on the unit operating cost plus a constant percentage mark-up (or a constant amount) which depends on factors such as the industry and the product type (US Department of Defense Contract Pricing Reference Guides 2000).

Traditional economists contend that mark-up pricing is not consistent with market-based profit-maximisation pricing, i.e., marginal analysis approach. However, the simplicity of cost-based pricing makes it a very appealing alternative for many firms. Empirical studies have revealed that very seldom do managers use the concept of equating marginal cost to that of marginal revenue in setting prices. Rather, most managers work in terms of markups or profit margins as their basis for pricing (see number of references in Hay and Morris 1991).

Despite this apparent inconsistency, it is not very difficult to relate the concept of mark-up pricing and profit-maximising pricing (Hay and Morris 1991). In profit-maximising pricing, we equate marginal revenue (MR) to marginal cost (MC). MR is given by:

$$MR = price(1 + \frac{1}{demand\ elasticity}). \tag{5.1.1}$$

If MC is taken to be equal to a constant average cost, then the profit-maximising equality, MR = MC can be simplified to show that the optimal mark-up should be inversely proportional to the demand elasticity. This also makes intuitive sense. If there are close substitutes existing in the market, then firms cannot charge a high mark-up. However, for price-inelastic products the firm can extract a large price premium from customers. There is ample evidence that firms vary the mark-up inversely with the price elasticity for the product's demand (Bliss 1988; Hay and Morris 1991). Hence, the firms that use mark-up pricing implicitly strive for profit maximisation pricing. However, this implicit profit maximisation is done based on empirical evidence of the product's price elasticity and not explicitly through a marginal analysis approach.

In environments where some make-to-stock firm has a cost advantage and knows that customers are price-sensitive, it might use mark-up pricing. Furthermore, the popularity of mark-up pricing is likely to be sustained because of the recent trend of cost transparency. The cost of products is becoming more "transparent" nowadays due to widely available information on the internet. This implies that firms now have less opportunity to extract a price premium. Sinha (2000) postulates that, under such circumstances, customers will pay the seller's actual costs and a "reasonable" premium, i.e., that the price should be based on the cost and a "reasonable" mark-up.

Keeping in mind its widespread use as a simple and effective approximation to profit-maximising behaviour, we will assume in our research that a percentage mark-up is the pricing technique used by make-to-stock firms. For a particular product, the mark-up percentage will be fixed. However, the manager will decide on the constant percentage based on her/his experience pertaining to the product's price-elasticity.

Hay and Morris (1991) highlighted the need to link profit margins to demand conditions facing the firm. According to their work, firms which employ mark-up pricing estimate a unit cost based on normal ranges of production, independent of actual output (i.e., for pricing purposes they think of their average total cost curve to be horizontal). Ignoring the demand curve might generate substantial overestimation or underestimation of sales, which then might make the calculated average cost incorrect. But still this approximation is done largely to avoid the circularity that would otherwise develop of having to estimate demand (in order to determine output and associated unit costs) before the price derived from mark-up pricing is known. In our research, we show that even while using mark-up pricing it is possible to tackle the issue of "circularity". We develop models that, rather than assuming the operating cost to be constant, explicitly take into account how the demand affects average costs.

If some make-to-stock firm uses mark-up pricing, it may try to reduce operating costs through some process-improving investments. The firm then has a choice of either to reduce price to gain a greater market share or to keep the price constant and let the increased profits flow right to the bottom line. We feel that this is a question of firm's strategy. For many make-to-stock firms, the focus of pricing is to improve market share rather than to maximise short-run profit. If a firm has a cost advantage and its customers are price-sensitive, it makes sense for the firm to cut price to gain market share (Rao et al. 2000; Hay and Morris 1991) even if such cost advantage is short-lived (since the competition will eventually catch up). The temporary advantage can be sustained within an overall strategy of "continuous improvement".

As Likierman (1981) points out, in competitive markets, increasing market share is imperative for both cost competitiveness as well as market power. For example, in the 1980's Japanese firms used their lower costs to under-price North American manufacturers. This created a perception among the customers that "Detroit" had been overcharging them and this perception of price unfairness is hurting North American manufacturers even to this day (Sinha 2000). Our setting is a profit-maximising firm that deals in a make-to-stock product and has price-sensitive customers. The firm's strategy is to attain a cost advantage in the market through process-improving investments. Our assumption that the firm passes on its savings from process improvements to the customers as a price reduction in order to improve its market share is thus reasonable.

5.2 General Model

Let us consider a price-setting retailer/manufacturer (henceforth termed firm) selling a single make-to-stock product for which customers are price-sensitive. The demand rate for the product, λ (units/unit time), depends on the unit selling price of the product, p (\$/unit):

$$\lambda = f_1(p). \tag{5.2.1}$$

The firm may be a monopolist, or may be one of several competing firms that offer similar products. However, we do not explicitly model the competition between firms except to assume that demand is decreasing in price (Deng and Yano, 2000).

The firm sets its price based on its total unit operating cost, m (\$/unit):

$$p = f_2(m), \tag{5.2.2}$$

so that:

$$\lambda = f_1(f_2(m)). \tag{5.2.3}$$

Note that mark-up pricing is a special case of (5.2.2) where $f_2(m) = \eta m$ ($\eta > 1$) and (η - 1)*100 percent is the desired per unit contribution margin. Recall that for make-to-stock firms operating costs consist mainly of relevant inventory costs (Chapter 1). Hence for us, the operating costs per unit will include set-up costs, purchase and/or production costs and inventory costs (holding and backordering). The firm can invest in projects that will reduce operating costs, and so we also include an investment cost per unit time in the total operating cost per unit:

m = [(set-up cost + production and/or purchase cost + inventory holding cost + backordering cost + investment cost) per unit time] / [demand per unit time]. (5.2.4)

We note that demand depends on price, which depends on unit operating costs, which in turn depends on demand. The firm's objective is to maximise its profit per unit time, π:

(P5.1) Maximise $\pi = (p - m)\lambda$,

subject to:

$p \geq m \geq 0$ and $\lambda \geq 0$,

where m is given by (5.2.4), p by (5.2.2) and λ by (5.2.1).

Theoretically, the following general procedure can be used to solve the optimisation problem:

Step 1: Determine the relationship between λ and m (as in (5.2.3)).

Step 2: Substitute (5.2.3) into (5.2.4) and obtain an explicit expression for m in terms of the relevant decision variables, i.e.,

$$m = f_3\,(\mathrm{DV}). \qquad (5.2.5)$$

Note that m may, at least initially, appear in both the numerator and denominator of the right-hand-side of the expression in (5.2.4), and thus writing (5.2.5) explicitly may or may not be possible.

Step 3: Substitute (5.2.5) into (5.2.3) to express λ in terms of the decision variables:

$$\lambda = f_1(f_2(f_3(\text{DV}))). \tag{5.2.6}$$

Step 4: Solve the following maximisation problem for the firm:

(P5.2) Maximise $\pi = (p - m)\lambda$,

subject to:

$p \geq m \geq 0$ and $\lambda \geq 0$,

where p, m and λ are given by (5.2.2), (5.2.5) and (5.2.6) respectively.

In general, the above model can be used in situations where operating costs are reduced by investments in changing some operating parameters, demand depends solely on price and price is a known, deterministic function of the operating costs. Obviously the complexity of the problem will vary from case to case. In our research we will be concerned with using this model to guide decisions on investments in lead time reduction. Several examples where our model can be used to help firms in making such decisions are:

(a) A deterministic customer demand setting where an Economic Order Quantity (EOQ) policy is used. The firm can then invest in reducing set-up time. If we assume the set-up time to be proportional to set-up cost, this is equivalent to investment in set-up cost reduction (Porteus 1985). As indicated in Chapter 2, Porteus (1985) showed that investments in reducing set-up costs can lead to decrease in relevant inventory costs per unit time. If the firm is using mark-up pricing, then reduction in inventory cost can reduce price and the reduced price will lead to increased demand from price-sensitive customers. However, the demand rate is also a parameter for the original cost minimisation problem. This implies that operating costs will be a function of the operating variables, batch size (Q) and set-up cost (K), as will be both price and demand. Now the firm's profit maximisation problem will have an objective function with only batch size and set-up cost as decision variables. In Section 5.3 we develop models to show how to determine the optimal set-up cost (time) for such investments;

(b) A stochastic lead time demand setting where a (Q, r) policy is applied. The firm may decide to invest in reducing the mean and/or variability of the lead time demand by reducing the procurement lead time duration. In Chapter 4, we showed that such investments could also lead to a decrease in expected long-term inventory cost per unit time for the buyer, which in turn would affect the price and demand rate. In Section 5.4 we formulate models to help firms in making optimal supply lead time decisions where the operating costs are a function of the operating variables, reorder point (r), batch size (Q) and lead time duration variability (α). Both price and demand, and hence the profit, will depend on the decision variables - batch size, reorder point and lead time duration (or demand) variability.

Our literature review shows that integrated inventory-marketing models like the ones we develop in this chapter have not been addressed thoroughly in the traditional operations management literature. Based on our review of previous research, we also observe that:

a) Models that consider price as an independent decision variable do not explicitly account for the effect of operating costs on price and demand though price and demand are related by a demand function;
b) Models that consider mark-up pricing, employ a mark-up over the production cost only and not the entire operating cost. In addition, these models do not take into account the investment required to affect process improvements.

Our models in this chapter will address the above two issues and show that they will have considerable effect on the optimal decision of firms.

Since we are dealing with make-to-stock products, we assume that customer demand is deterministic but price-sensitive (refer to Chapter 1). Also note that in the tradition of previous researchers (e.g., Porteus 1985; Hariga 2000 and references therein) we will use set-up cost as a surrogate for set-up time.

5.3 Investments in Set-up Cost Reduction

In Section 5.3.1 we develop models where no investment in reducing set-up cost is possible. These models will set the stage for Section 5.3.2 where we will deal with models in which investments can be done in reducing the set-up cost.

5.3.1 EOQ Model with Price-Sensitive Demand and Mark-up Pricing

The basic model setting in this section is similar to that of Section 5.2, i.e., a firm buying/producing a single make-to-stock item and selling it directly to price-sensitive customers. We make the usual EOQ assumptions (refer to Lee and Nahmias, 1993, pg 9) except that the demand rate, λ, is a decreasing function of retail price per unit, p. We assume that the firm uses a constant percentage mark-up over the total operating cost per unit, m, to determine the price. Operating costs include a set-up cost of K per order, a holding cost of h per unit per unit time and a purchase/production cost per unit, c. The holding cost is assumed to consist primarily of the cost of capital invested in the inventory. Each order is for a batch of Q units, and c is a decreasing function of Q, i.e., the purchase/production cost exhibits economies of scale based on the batch size. The scale economies might be due to a supplier quantity discount or economies of scale in transportation (Lee and Rosenblatt 1986; Chopra and Meindl 2001). Since the holding cost per unit, h, equals ic, where i is the carrying cost per unit per unit time, it too will be a function of Q. The firm's ordering or set-up cost per unit time is $(K\lambda / Q)$; the holding cost per unit time is $(icQ / 2)$ and the purchase/production cost per unit time is $(c\lambda)$. From (5.2.4), the firm's total operating cost per unit, m, will be given by:

$$m = \frac{\frac{K\lambda}{Q} + \frac{hQ}{2} + c\lambda}{\lambda}, \tag{5.3.1}$$

where c is a function of Q and $h = ic$.

Since price depends on the operating costs and demand is a function of price, demand is also a function of the operating costs. Note that demand itself affects the unit operating costs. The objective of the firm is to maximise its profit per unit time, as in problem P5.2 in Section 5.2, where the only decision variable is the batch size, Q. The operating cost per unit, m, will be obtained by solving (5.3.1) in terms of Q. A schematic representation of the proposed system is shown in Figure 5.3.1.

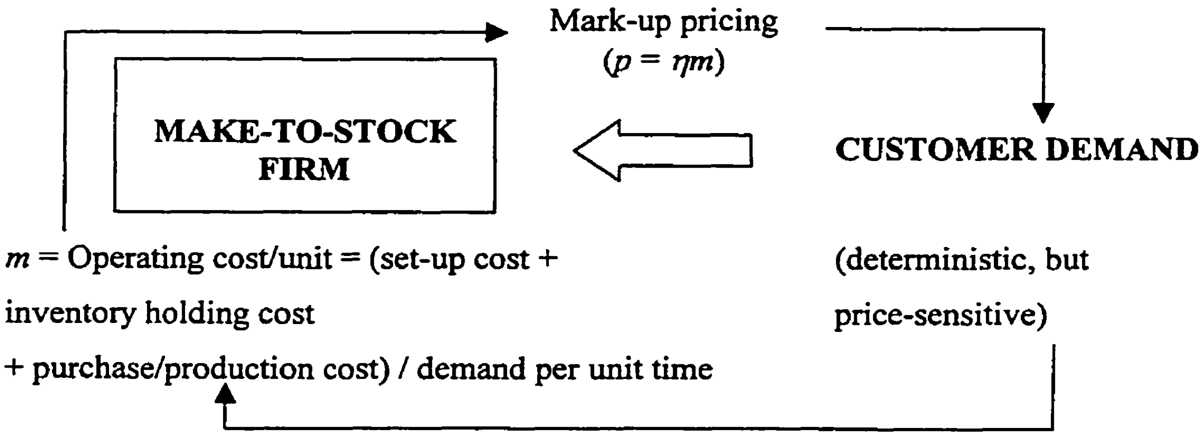


Figure 5.3.1: Supply Chain System for Section 5.3.1

5.3.1.1 Log-linear demand function, mark-up pricing and general non-increasing unit purchase cost

Before analysing a general form of demand function, in this section we will assume a particular log-linear demand function so that we are able to get better insights. We initially choose a demand function of the form:

$$\lambda = a(p^{-2}), \qquad (5.3.2)$$

where a higher value of a represents higher overall potential for demand. This function while having the desirable properties of constant demand elasticity is also analytically tractable. Since the price is a fixed mark-up over the total operating cost per unit, m, we have:

$$p = \eta m. \tag{5.3.3}$$

We can now express λ in terms of m as:

$$\lambda = (a\eta^{-2})m^{-2}. \tag{5.3.4}$$

The per unit purchase cost function, $c(Q)$, is assumed to be a general, non-increasing function of Q. The profit function for the firm will then be:

$$\pi = (p - m)\lambda = [a\eta^{(-2)}(\eta - 1)]m^{(-1)}. \tag{5.3.5}$$

Let

$$u = \mathrm{c}(Q)Q\frac{i}{2a\eta^{-2}}, \tag{5.3.6}$$

and

$$v = \frac{K}{Q} + c(Q). \tag{5.3.7}$$

Substituting (5.3.4) into (5.3.1) and solving for m we obtain:

$$m = \frac{1 \pm \sqrt{1 - 4uv}}{2u}. \tag{5.3.8}$$

As long as the discriminant of (5.3.8) is positive, both roots will be real and positive. However, if we substitute the two roots of m in the profit function in (5.3.5), we can show that the root corresponding to the minus sign will always give a higher profit than the root corresponding to the plus sign (Appendix 5.1). Hence, we can ignore the root corresponding to the plus sign for further analysis.

For each value of Q there will thus be a single relevant m (and hence π) and we want to find out the Q for which π will be maximised. Substituting the root corresponding to the minus sign of (5.3.8) into (5.3.5) we have:

$$\pi = \frac{2a\eta^{-2}(\eta-1)u}{1-\sqrt{1-4uv}}. \tag{5.3.9}$$

The firm's problem can be written as:

(P5.3) $\underset{Q}{\text{Maximise}}\,\pi = \dfrac{2a\eta^{-2}(\eta-1)u}{1-\sqrt{1-4uv}},$

subject to:

$0 < 4uv \le 1.$

We can show that under certain conditions the profit function will be semi-strictly quasiconcave in Q for feasible Q (Appendix 5.2). When π is semi-strictly quasiconcave, a local maximum will be the global maximum (Schaible 1981) and the optimum Q^* will be given by the solution to the equation $\pi_Q = 0$, and the optimal profit (π^*) will be given by substituting Q^* into (5.3.9).

5.3.1.2 Log-linear demand function, mark-up pricing and constant unit purchase cost

In this section, we will analyse the case when unit purchase cost is constant, i.e., $c(Q) = c$ $\forall$ Q, and all other conditions remain the same as before. With this assumption, the condition $0 < 4uv \le 1$ implies that $0 < K \le [(a\eta^{(-2)}/2ic) - Qc]$ and we can prove an even stronger result about the profit function than just semi-strict quasiconcavity.

Proposition 5.3.1: For constant unit purchase/production cost, the profit function (π) is concave in Q for feasible Q.

Proof: Refer to Appendix 5.3. ■

The equation for $\pi_Q = 0$ will be a quadratic equation in Q - one solution will be negative and the other positive. Since batch size should be positive, the optimum Q will be given by:

$$Q^*(K) = \frac{1}{c}\left[\sqrt{\frac{2Ka\eta^{-2}}{ic}} - 2K\right]. \tag{5.3.10}$$

Note that (5.3.10) requires only that $K \leq (a\eta^{(-2)}/2ic)$ which is less restrictive than the condition for $1 - 4uv \geq 0$ and so will be satisfied by all feasible K. The explicit expression for Q^* in (5.3.10) also allows us to investigate analytically the nature of the optimal batch size.

Proposition 5.3.2: Q^* is concave in K and reaches its maximum at $K = a\eta^{(-2)}/8ic$.
Proof: Differentiating $Q^*(K)$ of (5.3.10) twice with respect to K we can easily show that Q^* is concave in K and by solving $\partial Q^*(K)/\partial K = 0$ for K, we can show that $Q^*(K)$ reaches its maximum at $K = a\eta^{(-2)}/8ic$. ■

The implication of Proposition 5.3.2 is interesting. In almost all types of cost minimisation inventory models, including stochastic ones, the optimal batch size is always monotone increasing concave in K. Our model produces different results. Though Q^* is still concave in K, it is not monotone increasing. This implies that for the type of firms we are modelling, managers must be careful about reducing batch size when set-up cost is decreased. The explanation lies in the inter-relationship between demand and the operating variable, batch size, itself (we will come back to this issue later).

There might be a tendency on the part of many firms to set price as a mark-up over only the production/purchase cost. Ladany and Sternlieb (1974) analysed the profit-maximising batch size of such a firm where price is taken to be a mark-up over purchase cost only (i.e., $p = \eta c(Q)$). In that case, the effects of set-up cost and holding cost on the price are not accounted for explicitly. When $c(Q) = c$ (> 0) $\forall$ Q, then Ladany and Sternlieb's model will be equivalent to the traditional EOQ model where λ is the demand corresponding to the price, ηc (note that all other conditions remain the same as in our model). With this demand rate the profit maximising batch size in their model is given by:

$$Q^*_{LS} = \frac{1}{c}\sqrt{\frac{2Ka\eta^{-2}}{ic}}. \tag{5.3.11}$$

Comparing (5.3.10) and (5.3.11), it is clear that omitting set-up and holding cost from operating costs when determining price will lead to a different optimal batch size.

Proposition 5.3.3: $Q^*_{LS} > Q^*$ for positive K and the difference will increase as K increases and/or c decreases.

Proof: $Q^*_{LS} - Q^* = \dfrac{2K}{c}$. Hence, Q^*_{LS} will always be greater than Q^* for any $K > 0$. The difference is linear increasing in K and convex decreasing in c. ∎

Since, in our model, demand and price are both functions of batch size, we would expect that there would be such a difference. However, from a managerial standpoint, it is important to note that if K is large and/or c is small, the difference in the optimal batch sizes might be substantial. As K increases we would expect the set-up cost to have a larger effect compared to purchase/production cost on the total operating cost and hence in our model the price and demand will be affected more, resulting in larger difference in optimal batch sizes. On the other hand, as c increases the effect of purchase/production cost on the total unit operating cost will increase and hence the difference in the value of optimal batch sizes will decrease. Also note that Q^*_{LS}, as expected, is monotone increasing concave in K.

It is important to investigate the effect on the firm's profit if Q^*_{LS} is used in our model instead of Q^*. Since in our model demand itself is a function of the decision variable, Q, any type of comparison should be based on a profit function including the purchase cost (this is unlike in the traditional EOQ model where purchase cost is independent of Q and hence should be ignored). Comparing the profit functions we obtain:

$$\frac{\pi(Q^*)}{\pi(Q^*_{LS})} = \left(\frac{u(Q^*)}{u(Q^*_{LS})}\right) * \left(\frac{1-\sqrt{1-4u(Q^*_{LS})v(Q^*_{LS})}}{1-\sqrt{1-4u(Q^*)v(Q^*)}}\right). \tag{5.3.12}$$

By definition, $\pi(Q^*)/\pi(Q^*_{LS}) \geq 1$. In fact, this ratio can be quite large. If we plot $\pi(Q^*)/\pi(Q^*_{LS})$ versus K, it appears to be increasing convex in K (refer to Figure 5.3.2 in page 147 for a plot of this ratio for Example 5.3.1.1) with the ratio $\rightarrow 1$ as $K \rightarrow 0$. This reason behind this intuitive result is same as we indicated before regarding the difference between Q^*_{LS} and Q^*. As K increases, the effect of set-up costs on the total operating cost will increase and hence it will be more harmful to use Q^*_{LS} as the optimal batch size. However, as $K \rightarrow 0$, purchase/production cost will be the defining element in the operating costs and hence it will be natural to base the price just on that cost. In the following numerical example we will show the detailed effect of using Q^*_{LS} instead of Q^*.

Example 5.3.1.1: The parameter values are as follows: K = 1100, i = 0.3, η = 1.2, a = 6000, and $c = 1\ \forall\ Q$. With our model Q^* = 3327.71. In the Ladany and Sternlieb model, p = 1.2 (1.2*1), so λ = 4166.67 and thus Q^*_{LS} = 5527.71. Table 5.3.1 shows the detailed cost and revenue elements with Q^* and Q^*_{LS} in our model.

Table 5.3.1: Cost and Revenue Elements with Q^* and Q^*_{LS} for Example 5.3.1.1

	$Q^* = 3327.71$	$Q^*_{LS} = 5527.71$
A) Set-up cost / time	499.16	212.36
B) Holding cost / time	499.16	829.16
C) Purchase cost / time	1510.04	1067.13
D) Total Operating Cost / time (= A + B + C)	2508.35	2108.65
E) Demand rate ($\lambda = ap^{-2}$)	1510.04	1067.13
F) Operating Cost per unit (m = D/E)	1.66	1.98
G) Price / unit ($p = \eta m$)	1.99	2.37
H) Revenue ($= p\lambda$)	3010.03	2530.37
I) Profit ($\pi = [p - m]\lambda$)	501.67	421.73

For this numerical experiment, Q^*_{LS} is approximately 40% greater than Q^* and the optimal operating cost per unit is larger if Q^*_{LS} is used in our model. The total operating costs are lower with Q^*_{LS}; however, λ is also smaller with Q^*_{LS} resulting in higher unit operating cost, m. The net result is that the profits are about 19% lower using Q^*_{LS} in our model instead of Q^*. With $K = 1250$, the difference in profit can be as large as 30%.

In our model, the inventory costs (set-up + holding) depend on λ which itself depends on the inventory costs (through m). This explicit circular dependence makes this model much more complex and realistic than the normal EOQ model. The optimal batch size attempts to minimise the unit operating cost rather than the absolute operating cost which is one of the reasons for the "apparently counter-intuitive" behaviour of Q^* (for a somewhat related idea refer to Gerchak, Hassini and Ray 2000).

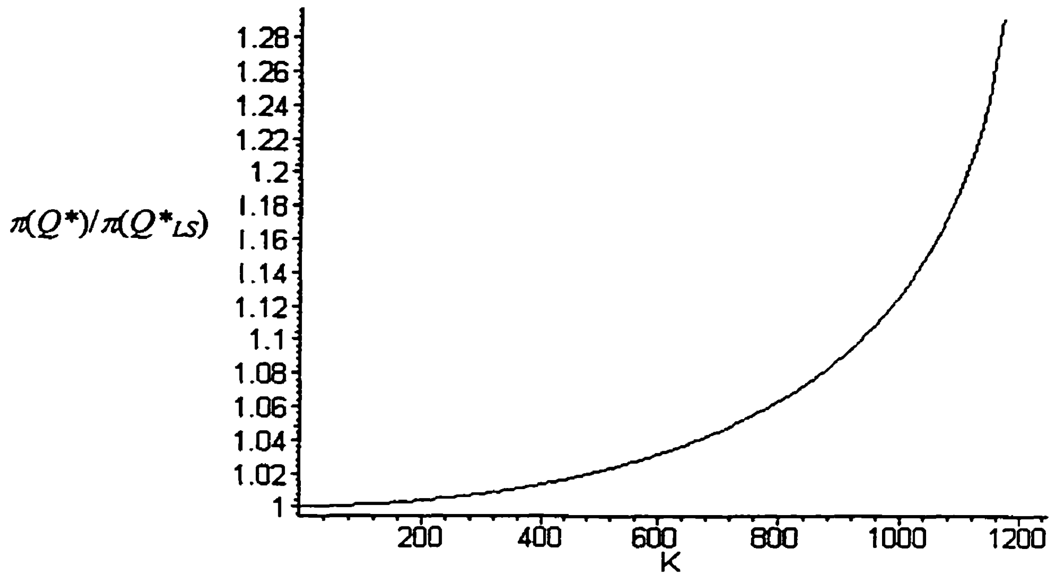


Figure 5.3.2: Plot of $\pi(Q^*)/\pi(Q^*_{LS})$ versus K for $c(Q) = c\ \forall Q$
($i = 0.3$, $\eta = 1.2$, $c = 1$, $a = 6000$)

If we vary K, then Q^* is increasing up to $K = 1736.11$ ($= a\eta^{-2}/8ic$) and then decreasing. If due to some reason K decreases from 1800 to 1750, the firm should not blindly decrease the batch size. Note that the Q^* in our model can be thought of as the "modified" optimal EOQ ($= \sqrt{\dfrac{2K\lambda(Q^*)}{ic}}$), but $\lambda(Q^*)$ itself will be a complex function of Q^*. In the following sub-section we will extend the model of this section by assuming a more general form of decreasing unit purchase cost function.

5.3.1.3 Log-linear demand function, mark-up pricing and non-increasing unit purchase cost function of power form

Let us assume a more general form of unit purchase cost: $c(Q) = c + (d/Q)$ ($c, d > 0$), but the form of log-linear demand function remains the same as before. This type of unit purchase cost function has been used in the literature (Ladany and Sternlieb 1974; Lee and Rosenblatt 1986). With these assumptions, the profit function will be of the same form as (5.3.9), but with:

$$u = Q\,(c + \frac{d}{Q})\frac{i}{2a\eta^{-2}}\,, \tag{5.3.13}$$

and

$$v = \frac{K+d}{Q} + c. \tag{5.3.14}$$

The optimisation problem for the firm will be similar to P5.3.

We can now have the following proposition:

Proposition 5.3.4: The profit function (π) is semi-strictly quasiconcave in Q for feasible Q.

Proof: From (5.3.13) and (5.3.14) it is clear that u is linear increasing in Q while v is decreasing convex in Q. Also uv is convex in Q. The rest of the proof is similar to Appendix 5.2. ■

Note that Ladany and Sternlieb (1974) did not prove the unimodality of their profit function even for this particular functional form of $c(Q)$. Though we prove that $\pi(Q)$ is unimodal, it is not possible to obtain an explicit solution for Q^* from $\pi_Q = 0$ assuming $c(Q) = c + (d/Q)$; however $\pi_Q = 0$ can easily be solved using any standard mathematical package. The solution to $1 - 4uv = 0$ (a quadratic, concave function in Q) will give the limits of feasible Q. At the smaller root, the discriminant will be increasing and at the larger root it will be decreasing. Since at both limits the profit function will be positive, then for feasible Q the profit will be positive. This is intuitive since our price is a mark-up over the total operating cost. We will again use a numerical example to show the effect on π if Q^*_{LS} (i.e., the optimal batch size when the price is a mark-up over only $c(Q)$) is used in place of Q^* for this model.

Example 5.3.1.2: Let the parameters be: $K = 1400$, $i = 0.1$, $\eta = 1.2$, $a = 6000$, $c = 2$ and $d = 100$. The optimal batch size in our model, $Q^* = 2437.46$ and $\pi^* = 256.93$. For the Ladany and Sternlieb model, $Q^*_{LS} = 3745.72$. In Table 5.3.2 we show the detailed effect of using Q^*_{LS} in our model.

Table 5.3.2: Cost and Revenue Elements with Q^* and Q^*_{LS} in Example 5.3.1.2

	c	m	p	λ	π
$Q^* = 2437.46$	2.04	3.24	3.89	396.08	256.93
$Q^*_{LS} = 3745.72$	2.03	3.55	4.25	331.41	235.02

Note that when price is a mark-up over just the purchase costs, unless the mark-up is "large" enough, profits may actually become negative. As in the last section, Q^*_{LS} is much greater than Q^* (about 35%) and profit is much lower (about 9.5%). The reasoning for the low profit will be similar to Example 5.3.1.1. However, when c is a function of Q, Q^* might not be $\sqrt{\dfrac{2K\lambda(Q^*)}{ic(Q^*)}}$ as it was in Section 5.3.1.2.

From our numerical experiments we observe that $Q^*_{LS} > Q^*$. However, the generality of the observation is difficult to prove analytically because of the complex nature of the first order condition:

$$\pi_Q = (u_Q)\sqrt{1-4uv} + 2uv(u_Q) - u_Q + 2u^2(v_Q) = 0, \tag{5.3.15}$$

where u and v are given by (5.3.13) and (5.3.14) respectively.

Even for the specific demand and unit purchase cost function, equation (5.3.15) will be quite complex. This should not be surprising since it is also difficult to prove the unimodality of Ladany-Sternlieb model. From our numerical experiments it also appears that $Q^*_{LS} - Q^* \neq (2K/c(Q^*))$, in general.

For this section, there is no explicit solution for $Q^*(K)$, though we can determine an explicit expression for the lower bound on its value. This lower bound will help to reduce the search space for Q^*. Let us define Q_l as the solution to $T_Q = 0$ where $T = 1 - \sqrt{1-4uv}$.

Proposition 5.3.5: Q_l is a lower bound on Q^*.

Proof: Q^* is derived from the solution to $\pi_Q = u_Q T - T_Q u = 0$. Since u_Q is non-negative, so Q^* requires that $T_Q \geq 0$. We can show that T is convex in Q. Therefore, the lower bound on Q^* will be given by Q_l. ■

For this section, the expression for Q_l is:

$$Q_l = \frac{1}{c}\sqrt{d(K+d)}. \tag{5.3.16}$$

To determine the behaviour of Q^* we have to resort to total differentiation. Total differentiation of $\pi_Q = 0$ with respect to K gives us the following expression for $\partial Q^*(K)/\partial K$ (refer to Appendix 5.4):

$$\frac{\partial Q^*(K)}{\partial K} = \frac{u_{QK}T + u_Q T_K - uT_{QK} - u_K T_Q}{T_{QQ}u - u_{QQ}T}. \tag{5.3.17}$$

We have, $u_{QK} = 0$, $u_{QQ} = 0$ and $u_K = 0$. The expression in (5.3.17) simplifies to:

$$\frac{\partial Q^*(K)}{\partial K} = \frac{u_Q T_K - uT_{QK}}{T_{QQ}u}. \tag{5.3.18}$$

It can be easily shown that, $T_K \geq 0$ and $u_Q \geq 0$ and the expression for T_{QK} will be:

$$T_{QK} = 2\frac{(uv)_{QK}\sqrt{1-4uv} + \dfrac{2(uv)_Q(uv)_K}{\sqrt{1-4uv}}}{1-4uv}. \tag{5.3.19}$$

Since $(uv)_{QK} \leq 0$, $(uv)_Q \geq 0$ and $(uv)_K \geq 0$, the sign of T_{QK} and (5.3.18) can be either positive or negative. Our numerical experiments confirm that Q^* will not be necessarily monotone in K; as in Section 5.3.1.2, it will be concave in K. In Table 5.3.3, we show the values of $Q^*(K)$ for different values of d. In all cases, Q^* is concave in K. To prove the concavity of Q^* with respect to K we will have to determine the sign of $\partial^2 Q^*(K)/\partial K^2$. This expression is quite complex (refer to Appendix 5.4). We leave the analysis for future research.

From Table 5.3.3 we also note that while Q^* might be monotone in d, the direction is not clear. It is increasing for smaller K while decreasing for larger K. The behaviour of Q^* with respect to d is intuitive. For smaller values of K, Q^* increases with d to take advantage of lower production/purchase cost. But as K increases and the effect of d becomes less significant, Q^* reduces to decrease the holding cost. Managers must thus be very careful when deciding about the optimal batch size even when the economies of scale vary. For a profit-maximisation model like ours, the explicit relation between the demand and the operating variables can give results that run counter to most traditional cost minimisation models.

We can also have the following proposition for the model of this section.

Proposition 5.3.6: The condition $\sqrt{1-4uv} \leq (1/3)$ is sufficient for the profit function (π) to be concave in Q for feasible Q.

Proof: Refer to Appendix 5.5. ■

The expression $\sqrt{1-4uv}$ is positive concave in Q and at the feasible limits of Q it will be equal to 0 (increasing at lower limit while decreasing at upper limit). So, we are sure that for Q sufficiently close to the feasible limits, π is concave in Q. Also for any $Q_l \leq Q \leq Q^*$, π is also concave in Q. However, it is difficult to analytically prove concavity of π for all feasible Q.

Table 5.3.3: $Q^*(K)$ for Different Values of K and d ($i = 0.3$, $\eta = 1.2$, $c = 1$, $a = 6000$)

K	$d = 0$	$d = 10$	$d = 20$	$d = 30$	$d = 40$	$d = 50$
900	3200.00	3204.86	3209.52	3213.97	3218.22	3222.27
1100	3327.71	3329.44	3331.01	3332.41	3333.64	3334.7
1300	3409.25	3408.48	3407.55	3406.48	3405.26	3403.89
1500	3454.97	3452.08	3449.05	3445.88	3442.58	3439.14
1700	3471.84	3467.09	3462.21	3457.2	3452.06	3446.78
1900	3464.83	3458.4	3451.84	3445.15	3438.34	3431.39

5.3.1.4 General price-sensitive, decreasing demand function, mark-up pricing and general non-increasing unit purchase cost

In Sections 5.3.1.1 - 5.3.1.3 we showed that it is possible to determine the profit-maximising batch size when demand is price sensitive and price is assumed to be a percentage mark-up over the operating costs. However, in all the three previous sub-sections we assumed specific form of the demand and/or the unit production/purchase cost to obtain closed form solutions. In this section we will analyse the make-to-stock firm's maximisation problem assuming a more general demand function.

Suppose we assume a general log-linear demand function of the form:

$$\lambda = a(p^{\varphi}), \tag{5.3.20}$$

where a higher value of a represents higher overall potential for demand and φ (< 0) represents the constant price elasticity (So and Song 1998; Ladany and Sternlieb 1974; Lee and Rosenblatt 1986) while p is equal to ηm.

Substituting λ into (5.3.1) and solving to obtain $m = g(Q)$ (all others parameters are assumed constant), the firm's profit can be written as:

$$\pi(Q) = a\eta^{\varphi}(\eta\text{-}1)\, g(Q)^{(\varphi+1)}. \tag{5.3.21}$$

Differentiating (5.3.21) with respect to Q it is possible to arrive at the following conclusions:

Observation 1:

(i) For $\varphi > -1$, concavity of g in Q is sufficient (not necessary) for concavity of π in Q while convexity of g in Q is necessary (not sufficient) for convexity of π in Q and π_Q will have the same sign as g_Q.

(ii) For $\varphi < -1$, convexity of g in Q is necessary (not sufficient) for concavity of π in Q while concavity of g in Q is sufficient (not necessary) for convexity of π in Q and π_Q will have the opposite sign to g_Q.

For $\varphi > -1$, the profit will be increasing in Q if the operating cost is increasing in Q, i.e., as φ tends towards zero, we can increase m which will increase p, but the relative inelasticity of demand will result in demand being almost fixed and hence π will increase with m. For the special case of constant demand (i.e., $\varphi = 0$) and $c(Q) = c\ \forall\ Q$, π will be convex in Q implying that the optimal Q from our model is at either of the feasible limits. For $\varphi = -1$, the profit function will be constant ($=a\eta^{-1}[\eta\text{-}1]$) for any Q. This behaviour is not unexpected. As is well known in microeconomics, when demand is inelastic in price (i.e., $0 > \varphi \geq -1$), a price increase will lead to non-decrease in the firm's revenue (So and Song 1998; Pappas and Brigham 1979). Since our basic setting is an environment where demand is price-sensitive, assuming $\varphi = -2$ (Sections 5.3.1.1 - 5.3.1.3), i.e., the demand is "sufficiently price-sensitive", is justified. However, with this general log-linear form of the demand function we do not obtain many managerial insights into the firm's optimal action.

In Sections 5.3.1.1 to 5.3.1.4, we were able to develop a profit maximising model for a make-to-stock firm that determines its price as a constant percentage mark-up over the unit operating cost and sells its products to price-sensitive customers. We were able to prove the unimodality of the profit function and determine easily-computable explicit expressions for optimal batch size or its bounds for practising managers. In these types of firms it is important for managers to note that one of the most basic tenets of inventory models which specifies that the optimal batch size will always increase with set-up cost, does not hold. When operating cost, price and demand are explicitly related, then for profit-maximisation models, the optimal batch size can behave quite differently. Also we show that using an EOQ batch size, or optimal batch size assuming price to be a mark-up over just the production cost, can result in substantial profit loss. Managers should be especially careful in choosing the batch size when the set-up cost is a major portion of the operating cost.

5.3.2 EOQ Model with Price-Sensitive Demand, Mark-up Pricing and Investments in Set-up Cost Reduction

Though the models in Section 5.3.1 had no investment in lead time reduction, they are important since they provide us valuable insights and form the basis of the models of this section where we incorporate investments in set-up cost reduction. We consider a firm in the same setting as in Section 5.3.1 except that now the firm has the option of investing in reducing its set-up cost, K. As indicated in the literature review, both practical experience and academic research (Porteus 1985; Cachon and Fisher 1999) has clearly proven the effectiveness of investment in set-up cost reduction. Following Porteus (1985, Section 2) and other research in this area, we assume that the investment function $a(K)$ denotes the cost of fixing the set-up cost at level K. We peg the cost for fixing the set-up cost at a particular level rather than to a change in it. An opportunity cost of $ia(K)$ is charged per unit time as part of the operating cost for the investment. Like in Porteus (1985), the investment cost can be thought of as either one-time irreversible investment cost or as a revocable lease that specifies a fee to be paid per unit time to maintain that set-up cost level. A schematic representation of the proposed physical and conceptual system will now look like as shown in Figure 5.3.3.

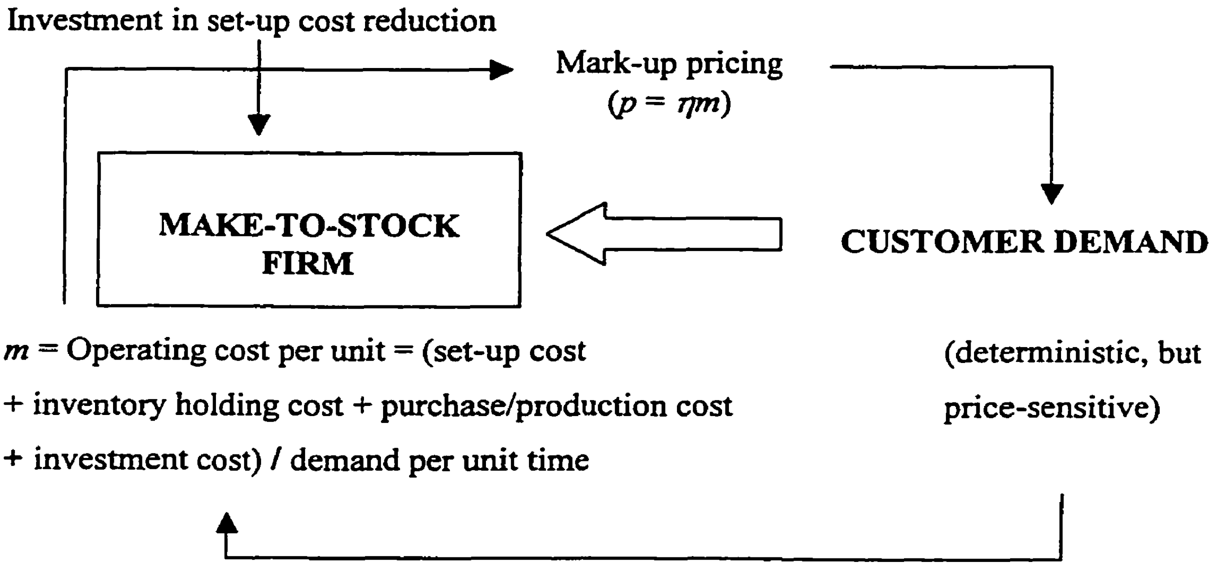


Figure 5.3.3: Supply Chain System for Section 5.3.2

With the investment cost as a part of the operating cost, expression (5.3.1) take the form:

$$m = \frac{\frac{K\lambda}{Q} + \frac{hQ}{2} + c\lambda + ia(K)}{\lambda}, \qquad (5.3.22)$$

where, as in Section 5.3.1, the unit production cost, c, can be a constant or a function of Q, and $h = ic$.

Now the profit maximisation model will have two explicit decision variables, K and Q. The basic model structure and the solution method of this section will be the same as in Section 5.3.1. However the investment cost will affect the operating cost and hence demand, which in turn will influence the operating cost and investment decision.

The profit maximisation problem for the firm will now be:

(P5.4) $\underset{Q,K}{\text{Maximise}}\ \pi$,

subject to:

$$0 < 4wz \leq 1,$$

where:

$$\pi = \frac{2a\eta^{-2}w}{1-\sqrt{1-4wz}}, \tag{5.3.23}$$

$$w = \frac{i}{a\eta^{-2}}\{\frac{Qc(Q)}{2} + a(K)\}, \tag{5.3.24}$$

and

$$z = \frac{K}{Q} + c(Q). \tag{5.3.25}$$

Note that in this section π, w and z are all functions of Q and K.

5.3.2.1 Profit maximisation with respect to Q

In this section we investigate the maximisation problem P5.4 with respect to the decision variable, Q, assuming K to be constant.

We will concentrate on two cases: i) $c(Q) = c\ \forall\ Q$, and ii) $c(Q) = c + (d/Q)$.

- For both cases, $T = 1 - \sqrt{1-wz}$ will be convex in Q and the profit function will be semi-strictly quasiconcave in Q implying that the solution to $\pi_Q = 0$ will give the optimal Q, Q^*. However, we cannot obtain closed form results solving for Q in $\pi_Q = 0$ even for $c(Q) = c\ \forall\ Q$.
- For both cases we can use the concept of Proposition 5.3.6 to prove that $\sqrt{1-4wz} \leq (1/3)$ is sufficient for concavity of π.
- The sign of $\partial Q^*(K)/\partial K$ is not obvious. However, from numerical examples we can deduce that $Q^*(K)$ is not monotone in K but concave. In Table 5.3.4 we show $Q^*(K)$ as a function of K for different types of investment functions, $a(K)$: i) Power (a(K) =

b/K where b is a positive constant), and ii) Logarithmic ($a(K) = j - b\ln(K)$ where both j and b are positive constants) and for both $c(Q) = c \ \forall\ Q$ and $c(Q) = c + (d/Q)$. Note that since we would expect that successive reductions in K will require larger and larger investments per unit reduction it is plausible that $a(K)$ should be decreasing convex in K. Both the forms of $a(K)$ satisfy the condition and are the most frequently used types of investment functions for this type of analysis (Porteus 1985; Nye 1997). In all cases $Q^*(K)$ is not monotone, but concave in K.

Table 5.3.4: $Q^*(K)$ versus K for Different $a(K)$ ($i = 0.1$, $\eta = 1.5$, $c = 2$, $a = 6000$)

Power Investment $a(K) = b/K$, $b = 100, d = 0$		**Power Investment** $a(K) = b/K$, $b = 100, d = 10$		**Logarithmic Investment** $a(K) = j - b\ln(K)$, $j = 100, b = 10, d = 0$	
K	$Q^*(K)$	K	$Q^*(K)$	K	$Q^*(K)$
1000	2227.46	1000	2232.43	1000	2220.52
1500	2452.83	1500	2454.44	1500	2444.61
2000	2564.34	2000	2563.77	2000	2554.96
2500	2603.08	2500	2600.87	2500	2592.61
3000	2590.15	3000	2586.55	3000	2578.57
3500	2538.05	3500	2533.21	3500	2525.33

5.3.2.2 Profit maximisation with respect to K

In this section, we investigate the maximisation problem P5.4 with respect to the set-up cost, K, assuming the batch size to be a given parameter. Note that this investigation might be worthwhile in itself, if for some reason the firm has to fix the batch size at some particular value (e.g., a material handling constraint) and hence the only option available to the firm is to invest in reducing the set-up cost.

In this section, we will again focus on power ($a(K) = b/K$, $b > 0$) and logarithmic ($a(K) = j - b\ln(K)$, $j, b > 0$) investment functions. Before going into the details of the decreasing convex investment functions, we would like to examine what will happen for a linear investment function (say, $b - jK$, $b, j > 0$).

For linear investment of the form $b - jK$ ($K \leq b/j$) we can show that wz is concave in K. Therefore, the range of feasible K might be continuous or discontinuous. Assuming that the range is continuous (i.e., the maximum value of $4wj$ is ≤ 1), it can be shown that for $c(Q) = c + (d/Q)$ or $c(Q) = c \; \forall \; Q$, the solution to the first order condition ($w_K T - T_K w = 0$ with $T = 1 - \sqrt{1 - wz}$) will give us $K^* = (Qc(Q)/2j) + (b/j) \geq (b/j)$. Hence, there is no optimal solution within the feasible range of K. The first differentiation of the profit function is either positive or negative within the feasible range. The profit function is thus monotone and can be either increasing or decreasing. This implies that either of the limits will be the optimum value.

For the power investment function, $a(K) > 0$ is always satisfied. For the logarithmic investment $a(K) > 0$ requires that $K < e^{(j/b)}$. For both investment functions and cost functions of the form $c(Q) = c \; \forall \; Q$ and $c(Q) = c + (d/Q)$, wz will be convex in K. For logarithmic investment function, convexity of (wz) requires that $K \leq Qc(Q)$ and we will assume it to be true in the rest of the section. So we can tell that T will also be convex in K. However, w is not concave in K; rather it is decreasing convex in K, and hence it is not necessary that π will be semi-strictly quasiconcave in K.

For both types of investment functions, as $K \rightarrow 0$, $wz \rightarrow \infty$. So, the lower limit for feasible K, K_L, must be strictly positive. If we assume that the minimum of $4wz$ does not exceed 1 and for the logarithmic investment function at the upper feasible limit of K (K_U), $4wz > 1$ and $(wz)_K > 0$, then there must be some feasible range of K and wz and T cannot be monotone within it. It is easy to show examples where π will not be concave in K for both power and logarithmic investment functions (Figure 5.3.4). So, we have to try to prove the unimodality of π with respect to K.

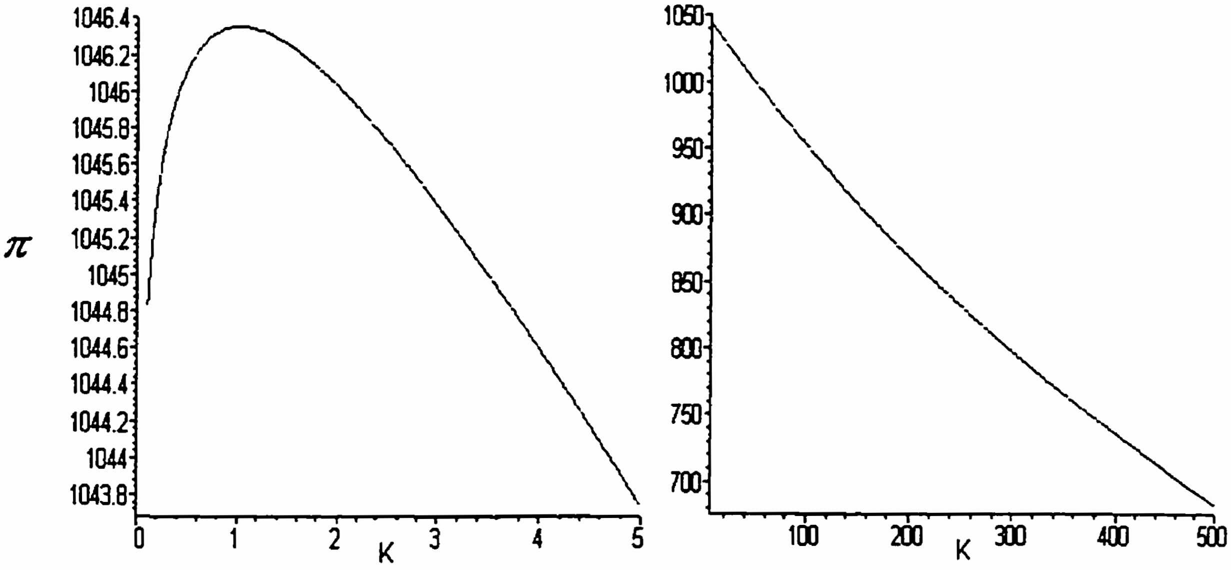


Figure 5.3.4(a): Plot of π as a Function of K for $a(K) = j - b\ln(K)$ and $c(Q) = c\ \forall Q$ ($i = 0.2$, $\eta = 1.5$, $c = 1$, $a = 5000$, $j = 100$, $b = 10$, $Q = 1000$)

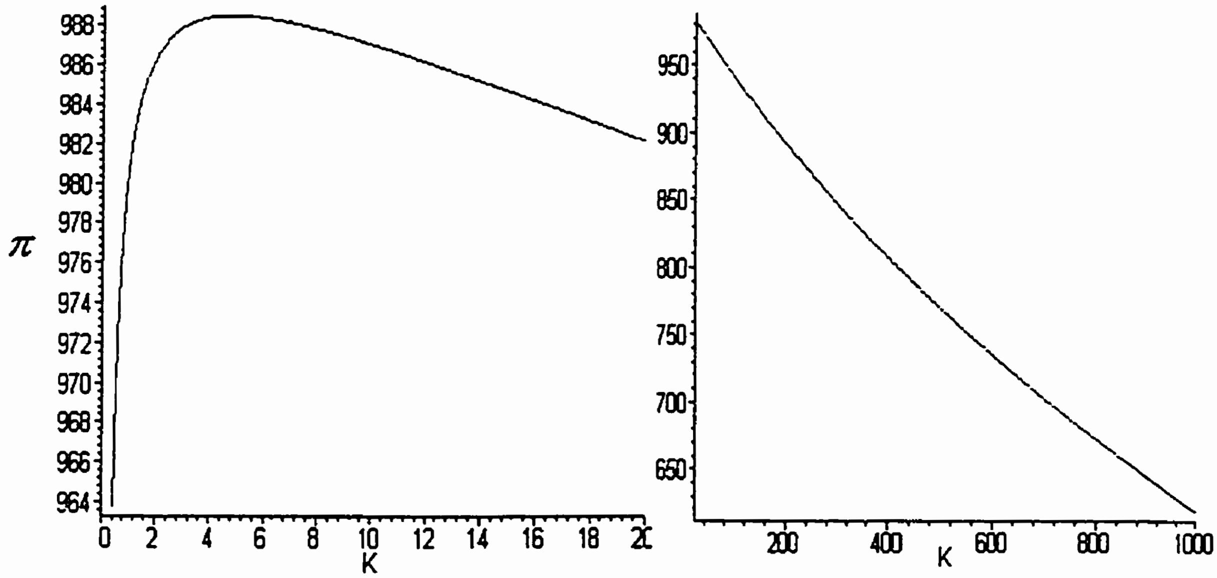


Figure 5.3.4(b): Plot of π as a Function of K for $a(K) = b/K$ and $c(Q) = c + (d/Q)$ ($i = 0.2$, $\eta = 1.2$, $c = 1$, $a = 5000$, $b = 50$, $d = 10$, $Q = 2000$)

The first order condition (FOC) for π with respect to K (π_K) will be of the form $w_K T = T_K w$. Differentiating both sides of the FOC with respect to K we have:

$$\frac{\partial(LHS)}{\partial K} = w_{KK}T + w_K T_K ; \ \frac{\partial^2(LHS)}{\partial K^2} = 2w_{KK}T_K + w_K T_{KK} + w_{KKK}T. \tag{5.3.26}$$

$$\frac{\partial(RHS)}{\partial K} = wT_{KK} + w_K T_K ; \ \frac{\partial^2(RHS)}{\partial K^2} = wT_{KKK} + 2w_K T_{KK} + w_{KK}T_K . \tag{5.3.27}$$

Differentiation of T yields:

$$T_K = \frac{2(wz)_K}{\sqrt{1-wz}}; \quad T_{KK} = \frac{2(wz)_{KK}\sqrt{1-wz} + 4\dfrac{[(wz)_K]^2}{\sqrt{1-wz}}}{1-wz}. \tag{5.3.28}$$

Proposition 5.3.7: $\pi_K = 0$ will have either one or three solutions.

Proof: For both types of investment functions, as $K \rightarrow K_L$, $T_K \rightarrow -\infty$. Since w_K is negative and T is positive, the LHS of the FOC will always be negative. Though T_K is unrestricted in sign (recall that T is convex in K), we are only interested in $T_K < 0$ (for $T_K \geq 0$, π_K will be negative). Since $w \geq 0$, $w_K \leq 0$, $w_{KK} \geq 0$, $w_{KKK} \leq 0$ and $T_{KK} \geq 0$ (for both investment functions), so for $T_K < 0$ we can show that the LHS will always be negative (finite negative as $K \rightarrow K_L$), increasing and concave. For $T_K < 0$ it is possible to prove that $T_{KKK} \leq 0$ (Appendix 5.6). Then the RHS will also be increasing and concave. However, as K tends to its lower feasible limit, the RHS will tend to $-\infty$, increase in a concave manner up to $T_K = 0$ and then become positive. If T is not monotone, then for this type of RHS and LHS, it follows that $\pi_K = 0$ will have either one or three solutions. ■

As $K \rightarrow K_L$, π will be increasing at infinite rate and for $T_K > 0$, π will be decreasing in K. If $\pi_K = 0$ has only one solution then π must be increasing up to that point and then decreasing, and so the unique solution to FOC will give us the profit-maximising K^*. If $\pi_K = 0$ has three solutions, then the profit function will be increasing from K_L up to the first solution, then decreasing up to the second solution, again increasing up to the third solution and then finally decreasing up to K_U. It is obvious that the maximum value of π will be given by either the first or the third solution to the FOC and so K^* can be determined easily by simple comparison of the value of π at the first and the third solution.

Though Proposition 5.3.7 shows that $\pi_K = 0$ can have three solutions, in all our numerical experiments with both types of investments, the profit function was always unimodal in K. A sufficient condition that $\pi_K = 0$ will have an unique solution is to prove that the slope of the RHS of the FOC is greater than the slope of the LHS of the FOC for $T_K \leq 0$, i.e., $(\text{RHS})_K > (\text{LHS})_K$ for $T_K \leq 0$. However, we can show by numerical examples that it is not, in general, true.

We can show that the optimal set-up cost, K^*, must be $\leq K_I$ (the solution to $T_K = 0$) and so K_I can be an upper bound on K^*. For the power investment function we can even show that the upper bound is independent of Q (for logarithmic investment, K_I will be a function of Q). The upper bound will help us reduce the search region for K^*, especially for the power investment case when the upper bound is very easily computable.

Proposition 5.3.8: In general $K^* \leq K_I$ and, for a power investment function, $a(K) = b/K$, $K^* \leq \sqrt{2b}$.

Proof: The solution of the FOC requires that $T_K < 0$. As we have already said, T is convex in K, implying that K^* must be $\leq K_I$. From (5.3.28), the solution to $T_K = 0$ will be given by the solution to $(wz)_K = 0$. For the power investment function the solution to $(wz)_K = 0$ will simplify to $(1/2) - (b/K_I^2) = 0$, implying that $K^* \leq \sqrt{2b}$. ■

It is somewhat interesting that for the power investment function, the upper bound of K^* is independent of Q of all parameters except b. We would normally expect it to be dependent on parameters other than b as well.

5.3.2.3 *Joint Profit maximisation with respect to Q and K*

In this section, we will investigate the maximisation of P5.4 jointly with respect to the decision variables, Q and K. We have already shown in Figure 5.3.4 that π is not necessarily concave in K. However, the profit function can still be unimodal with respect to the two decision variables. One of the ways to prove unimodality in this case is to find $Q^*(K)$ and then try to prove the unimodality of $\pi(Q^*(K), K)$ with respect to K. Since we have already proved that π will be semi-strictly quasiconcave in Q, the solution to $w_Q T - wT_Q = 0$ will give us the unique $Q^*(K)$. If we replace Q in (5.3.23) by this $Q^*(K)$ we will obtain $\pi(Q^*(K), K)$. Now we have to prove the unimodality of $\pi(Q^*(K), K)$ with respect to K, i.e., that

$$(w_Q \frac{\partial Q^*(K)}{\partial K} + w_K)T - (T_Q \frac{\partial Q^*(K)}{\partial K} + T_K)w = 0 \tag{5.3.29}$$

has a unique solution within the feasible range of K (in (5.3.29) T represents $1 - \sqrt{1 - w(Q^*(K),K)z(Q^*(K),K)}$). Since we are concerned with $Q^*(K)$ and we know that $w_Q T - T_Q w = 0$, thus we have to prove that $w_K T - T_K w$ has a unique solution for feasible K. Before going further we state the following proposition.

Proposition 5.3.9: For the model of Section 5.3.1.2, $\pi(Q^*(K), K)$ will be decreasing in K.
Proof: For the model of Section 5.3.1.2, we can show that T_K will be positive for feasible K and $w_K = 0$ implying that $\pi_K(Q^*(K), K) \leq 0$ for feasible K. Hence, π is decreasing in K.

■

The above proposition is rather intuitive since if there were no cost to reduce K, we would want it to be as low as possible. However, from our previous results, we know that T is convex in K and T_K can be both positive and negative. Since w_K is always negative, we are only concerned with $T_K \leq 0$ (for $T_K \geq 0$, π will be decreasing). Also as $K \to K_L$, $T_K \to -\infty$ but w_K will be finite negative. So, as $K \to K_L$, we can say that π will be increasing at infinite rate. But analytically it is difficult to prove that there will be unique solution to

$w_K T - T_K w = 0$ for feasible K. However, our extensive numerical experiments with power and logarithmic investment functions always resulted in a unique solution. Based on the fact that we know that $\pi(Q^*(K), K)$ will be increasing near the lower feasible limit and will be decreasing for any K for which $T_K \geq 0$ and from our numerical experiments we are confident that, in general, $\pi(Q^*(K), K)$ will be unimodal in K.

We performed numerical experiments to compare the optimal values of the decision variables and the profit from our model with two possible alternatives that a firm might employ:

a) Alternative I: If the firm does not explicitly take into account that lower operating costs from investments in set-up cost reduction can be passed on to the customers so as to increase demand. That is, it selects its price as a mark-up over a constant production/purchase cost, estimate demand based on that price and then determines the optimal batch size and set-up cost based on the estimated demand.

b) Alternative II: If the firm decides not to explicitly take into account the effect of lower operating costs on demand and chooses to use the decision variables resulting from solving a model assuming p and K to be independent decision variables.

Our numerical experiments show that, as expected, the optimal batch size and set-up cost resulting from our model can be very different from those of either of the above alternatives. However, if the firm uses values of decision variables derived from the alternatives within our model, the "loss in profit" can be significant.

Example 5.3.2.1: Let us assume that $a(K) = b/K$. With the parameters: $i = 0.35$, $\eta = 1.2$, $a = 350$, $b = 2600$ and $c = 0.5$ $\forall$ Q, the optimal decision variable values in our model are as follows: $Q^* = 386$ and $K^* = 26.97$. The optimal profit, $\pi(Q^*, K^*) = 68.48$. For Alternative I, $p = 0.6$ and $\lambda = 972.22$. The optimal decision variable values in this case will be $K_1^* = 21.35$ and $Q_1^* = 487$ (refer to Sections 4 and 5 of Porteus 1985). For Alternative II, the optimal decision variable values are: $K_2^* = 34.41$ and $p^* = 1.23$. The induced optimal batch size in this case will be $Q_2^* = 302$ (refer to Sections 6 and 7 of Porteus 1985). In Table 5.3.5 we show the effect on profit, demand, price and operating cost if we use either of the alternative optimal decision variable values in our model in place of Q^* and K^*.

Table 5.3.5: Cost and Revenue Elements with K^*/Q^*, K_1^*/Q_1^* and K_2^*/Q_2^* for Example 5.3.2.1

	m	p	λ	π
$K^* = 26.97$, $Q^* = 386$	0.710	0.852	482.38	68.48
$K_1^* = 21.35$, $Q_1^* = 487$	0.732	0.878	454.24	66.46
$K_2^* = 34.41$, $Q_2^* = 302$	0.730	0.876	456.41	66.61

From the representative numerical experiment it is clear that the optimal batch size and set-up cost of our model are very different from those of Porteus' model. Our optimal set-up cost is about 26% higher than Alternative I and about 22% lower than Alternative II while optimal batch size is about 21% lower than Alternative I and 28% higher than Alternative II. If the optimal decision variable values from either alternative are used in our model, the "loss in profit" can be significant. In our example it is almost 3%, which is quite high for companies in competitive situations. Using the optimal decision variable values from "wrong models" leads to larger operating costs per unit. Larger operating cost leads to higher price and hence lowers demand with the net result being that the profits are lower. If, in addition, there were economies of scale from batch size (e.g., $c(Q) = c + [d/Q]$) and if it is not taken into account (i.e., assuming $c(Q) = c \ \forall \ Q$), it can have an even stronger effect on the optimal decision variable values and hence on profit. Obviously as the economies of scale become more prominent, the detrimental effect on profit of ignoring them becomes more severe.

5.4 Investments in Supply Time Reduction

In Section 5.3 we developed models based on the idea that operating costs can be reduced by investments in set-up time (cost) reduction. In Chapter 4 we showed that investment in supply lead time reduction could also decrease inventory costs. In this section we develop models similar to those of Section 5.3 with the focus now on investment in supplier lead time reduction.

5.4.1 Basic Model

Assume a firm buying a single make-to-stock item from a supplier (internal or external) and selling it directly to price-sensitive customers. However, unlike in Section 5.3, there is a supply lead time from the supplier to the buyer that is random. Though the final customer demand is deterministic, because of the stochastic procurement lead time the lead time customer demand for the buyer will also be stochastic. The buyer is following a (Q, r) policy for its procurement control.

The customer demand rate, λ, is a decreasing function of price per unit charged to the customers, p. The firm sets its price, p, as a constant percentage mark-up over the operating cost per unit, m. Since the lead time demand is random, the firm might loose customers if it does not have sufficient safety stock. Hence, in addition to set-up cost, holding cost and purchase cost, the firm also incurs a backordering penalty cost per unit per unit time of b. The holding cost is assumed to consist primarily of cost of capital invested in inventory. Since the buyer is using a (Q, r) policy, whenever the inventory position of the buyer reaches r, it orders Q units from the supplier. As in Section 5.3, c is a non-increasing function of Q.

The firm's total inventory cost is composed of ordering costs, holding costs, backordering costs and purchase costs per unit time and from Chapter 4 will be given by,

$C(Q, r) = \frac{K\lambda}{Q} + hE(I) + bE(B) + c\lambda$, where $E(I)$ and $E(B)$ have been defined in Chapter 4. The firm's total operating cost per unit, m, will be given by:

$$m = \frac{\frac{K\lambda}{Q} + hE(I) + bE(B) + c\lambda}{\lambda}, \qquad (5.4.1)$$

where c is a function of Q and $h = ic$.

Since customers are price-sensitive, the demand, λ, will be a function of the price, p, and because the firm employs mark-up pricing, $p = \eta m$ where m is given by (5.4.1). The objective of the firm is to maximise its profit per unit time, similar to problem P5.2, where the decision variables are now both the batch size, Q and the reorder point, r. A schematic representation of the proposed system is shown in Figure 5.4.1.

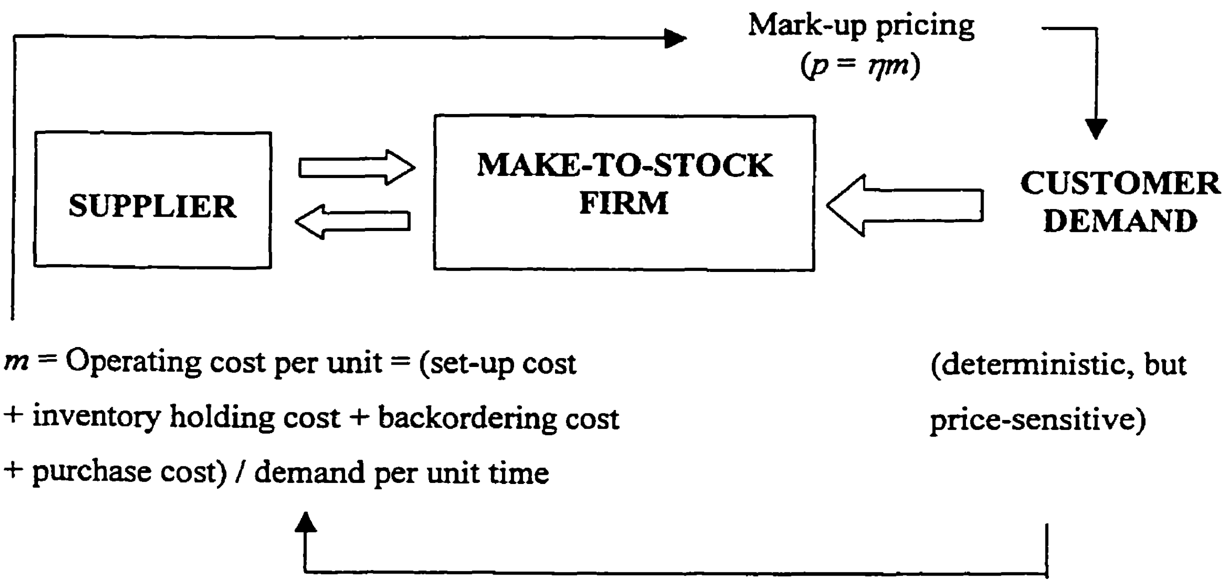


Figure 5.4.1: Supply Chain System for Section 5.4.1

5.4.2 Model with Investments in Supply Lead Time Reduction

In this section we extend the general model developed in Section 5.4.1 assuming that the firm invests in reducing supply lead time. Hence, the system is now similar to those considered in Chapter 4 except that customer demand, λ, is deterministic with a price-sensitive rate. The price will be determined by the firm's operating cost and so both p and λ will be functions of the operating decision variables.

In Chapter 4, we developed six different models based on the nature and frequency of the investment type. For the present model, let us assume that the investment in supply lead time reduction is one-time with the life of the investment depending on the time it is being used. The total cost will be given by:

$$C(Q, r, \alpha) = \frac{K\lambda}{Q} + \frac{\alpha}{Q}\int_r^{(r+Q)} G'(\alpha, y)dy + c\lambda + \theta i M^{1T}(\alpha), \quad (5.4.2)$$

where $G'(\alpha, y) = (h + b)\int_0^{(y/\alpha)} F(x)dx + b(\mu - \{y/\alpha\})$ and $h = ic$.

Note that now there are three operational decision variables - Q (batch size), r (reorder point) and α (refer to Chapter 4). The operating cost will now be given by:

$$m = \frac{\frac{K\lambda}{Q} + \frac{\alpha}{Q}\int_r^{(r+Q)} G'(\alpha, y)dy + c\lambda + \theta i M^{1T}(\alpha)}{\lambda}, \quad (5.4.3)$$

where c is a function of Q and $h = ic$.

The price charged by the firm will be $p = \eta m$ and the demand rate λ will be given by the relation in (5.2.1). However, the demand rate λ itself will affect m and even the investment decision will be affected by the integration of the cost minimisation model with market demand. The problem for the firm can now be written as:

(P5.5) $\underset{Q,r,\alpha}{\text{Maximise}}\ \pi = (p - m)\lambda,$

subject to:

$p \geq m \geq 0$ and $\lambda \geq 0$,

where m is given by (5.4.3), $p = \eta m$ and $\lambda = f_1(p)$ (λ is some decreasing function of p).

While the general procedure to solve this problem will be the same as shown in Section 5.2, the analysis will be much more complex than in Section 5.3. The operating cost affects the price as well as the demand. The demand rate in turn will affect the lead time demand distribution. Hence, the operating variables - Q, r and α - will affect the lead time demand distribution, i.e., $F(x)$. In this case obtaining a closed form solution for m in terms of the decision variables will be difficult. Though we formulate the problem here, we leave the analysis for future research.

5.5 Conclusions and Future Research Opportunities

In this chapter, we set about to model a profit-maximising firm selling a single make-to-stock product to price-sensitive customers. The firm sets its price as a fixed percentage mark-up over its operating costs per unit, and is considering investing in reducing its operating costs by optimal management of its lead time so that it can lower the price to gain a greater market share. For the case of deterministic (but price-sensitive) demand and investment in set-up time (cost) reduction, we were able to formulate a model where price and demand, and hence profit, are functions of the operating variables - batch size and set-up cost - of the firm. We show that when there is explicit dependence of the demand on the operating variables in a profit maximisation model, some well known solution properties from classical inventory management do not hold true anymore. In that case, managers have to be extra careful about choosing optimal operating variables. We were also able to show that investments in increasing efficiency (i.e., decreasing operating cost by some investment in set-up cost reduction) can be passed on as a price decrease to the customers which will increase the demand and the profit for the firm. If the firm does not take this effect into account, and determines the optimal decision variables assuming that the demand is constant or price is an independent decision variable, it will lead to sub-optimal results and the firm would be loosing substantial profit in a competitive market.

We were also able to formulate the problem for the case when the firm is buying the stock from some supplier before selling it and the supply lead time is stochastic. In this case, the lead time demand for the firm will be random and the firm can invest in reducing the variability of the supply lead time either by making the investment itself or by paying the supplier for lead time reduction investments. Again the price and demand (and hence profit) will be functions of the operating variables of the firm. In this case the variables will be - batch size, reorder point and variability of supply lead time. However, the complexity of this problem makes further analysis cumbersome and difficult. We leave this for future research.

We feel that our model adequately captures the main features of make-to-stock firms in a competitive environment and how such firms can use a time-based strategy to reduce costs and increase market share. The main contribution of this research lies in the fact that for the first time in the literature we develop a model where the profit-maximising operational variables for a firm are determined by taking into account how efficiency improvements can explicitly impact the profitability of the firm. We were able to couple the cost reducing operations research based models with relevant microeconomic models to demonstrate how not taking such interactions into account can lead to sub-optimal results.

In terms of scope for future research, further analysis of the model formulated in Section 5.4 with investments in supply lead time reduction would be worthwhile. As we indicated, most of the process-improvement models in operations management (discussed in Sections 2.5 and 2.6) assumed demand to be constant or stochastic with the mean demand rate being constant. It might be useful to use the general framework developed in Section 5.2 to analyse some of those models with price-sensitive demand. Such integrated production-marketing modelling will help in establishing the overall contribution of investments in increasing efficiency.

CHAPTER 6

CONCLUSIONS AND FUTURE RESEARCH

6.1 Summary and Conclusions

In this thesis, we investigated the issue of lead time management in supply chains in different competitive environments. We took into account both the costs and benefits associated with lead time reduction in three analytical models. One model studied delivery lead time management for make-to-order firms, and the other two models analysed supply lead time and set-up time management for make-to-stock firms.

The motivation for this research and some relevant background were presented in Chapter 1. It showed the importance of effective lead time management in modern supply chains and discussed the various costs and benefits that firms need to consider before deciding on the optimal lead time. We also discussed why it is necessary to have a different supply chain design and lead time reduction focus for make-to-order and make-to-stock supply chains.

In Chapter 2 we identified several gaps in the current literature on lead time management in supply chains.

- Models investigating delivery lead time management for make-to-order supply chains did not analytically account for the possibility of a price premium from shorter delivery times or economies of scale from increased demand.
- Previous literature on investments in supply lead time reduction did not employ the state-of-the-art (Q, r) model or capture the different types of possible investments while determining the optimal supply lead time.
- Models investigating set-up time/supply lead time management for make-to-stock supply chains either totally ignored or did not explicitly account for the effect of reduction of operating costs on final customer demand.

In our research, we addressed the above gaps so as to develop more comprehensive models. This has led to many new insights into lead time management in supply chains.

In Chapter 3, we dealt with delivery lead time management issues for make-to-order firms. We modelled a make-to-order supply chain consisting of a firm and its customers where the mean demand rate is a function of price and guaranteed delivery lead time and the market price is determined by the length of the guaranteed delivery lead time. We then extended our model by incorporating economies of scale where the unit operating cost is a decreasing convex function of the mean demand rate. The firm can invest in increasing capacity to guarantee a shorter delivery lead time but must be able to satisfy the customers according to a specified reliability level. Our models explicitly accounted for "price-sensitive" and "lead-time-sensitive" customers. We showed how the firm could select the optimal length of the guaranteed delivery lead time and processing rate to maximise its profits by a relatively simple procedure. Our numerical examples clearly indicated that ignoring the dependence of market price on the lead time offered and economies of scale, when they in fact exist, could lead to potentially large profit losses for the firm. It is also important for firms to take note of the inherent preference of customers for price or lead time when making decisions. We also extended our model by explicitly accounting for WIP holding costs and penalty costs for the firm. In that case we were only able to formulate the problem without providing any analytical solutions. Some of the possible extensions to this model might include:

a) Allowing the mean demand rate to be a function of the service level in addition to price and guaranteed delivery time;
b) Extending our model to a non-linear demand function;
c) Analytical solution of the model developed in Section 3.6 taking holding costs and backordering costs into account.

In Chapter 4, we developed models for investigating investments in supply lead time reduction for make-to-stock products. We used a continuous review "exact" (Q, r) model and captured the effects of investment in lead time duration reduction through investment in changing a single variable, α. In practice, investments can be one-time or recurring. It might also vary depending on the nature of investment (e.g., per unit or per cycle or per unit time). Our six new trivariate models captured both the costs and benefits of lead time reduction for the different types of possible investment. Our analysis of the cost model illustrated the benefits of investments in supply lead time reduction in terms of inventory cost. It also highlighted the importance of taking into consideration the type of the investment made by the supplier and how it is being passed on to the buyer before deciding on the optimal strategy. Our extensive numerical comparative statics showed that the interdependency of the three decision variables could result in some seemingly counter-intuitive results. From this chapter, we can conclude that investments in supply lead time reduction can result in substantial reduction of inventory costs after accounting for all the associated costs and benefits. However, it is important to consider the frequency and nature of the investment while developing models for supply lead time management. Some of the possible extensions to this model might include:

a) Analytical proof that the cost function of the unit model is univalleyed;

b) Analytical comparative statics.

In Chapter 5, we developed models for determining the optimal lead time for make-to-stock firms. In this chapter we considered a profit-maximising firm selling a single make-to-stock product to price-sensitive customers. The firm sets its price as a fixed percentage mark-up over its operating costs per unit and is considering investing in reducing its operating costs by proper management of its lead time. Lower operating costs will allow the firm to reduce price and gain a greater market share. For the case of deterministic demand and investment in set-up time (cost) reduction, we were able to formulate a model where price and demand, and hence profit, are functions of the operating variables - batch size and set-up cost - of the firm. We showed that when there is explicit dependence of the demand on the operating variables in a profit maximisation model, some of the best known solution properties from classical inventory management do not hold true anymore. In that case, managers have to be extra careful about choosing the optimal values of the operating variables. We were also able to show that in this case, if the firm ignores the explicit dependence by either assuming demand to be constant or price to be an independent decision variable, it will lead to sub-optimal results and the firm would be loosing substantial profit. For the case of stochastic lead time demand and investment in supply lead time reduction, we were also able to formulate the problem in terms of the operating variables of the firm - batch size, reorder point and variability of supply lead time. However, the complexity of the problem makes further analysis cumbersome and difficult. Some possible extensions to this model will include:

a) Analysis of the model developed in Section 5.4 where investment is made in supply lead time reduction in a stochastic lead time demand environment;
b) Utilising the general framework developed in Chapter 5 to extend a whole generation of process-improving-investment models that were developed ignoring the price-sensitivity of the customers (refer to Sections 2.5 and 2.6).

From a technical standpoint, the main contribution of this research lies in the fact that we develop new models incorporating issues that were not accounted for in previous models both for make-to-stock as well as make-to-order firms. Our models are able to provide new insights into lead time management issues in supply chains. Specifically, we showed the importance of integrated operations-marketing modelling in making supply chain decisions. From a managerial standpoint, this research can help managers to determine the key issues they must focus upon, depending on their competitive environment, when making their lead time decisions. Our research shows that both make-to-order and make-to-stock firms can gain competitive advantage through lead time management. As e-commerce and outsourcing grows in popularity and customers becomes more demanding, the importance of price and speed as competitive priorities will only increase. Simultaneously, competition will make resources scarcer, necessitating their optimal use. We feel that the integrative nature of our research is a significant addition to the operations management literature on time-based competition.

We would like to add here that as in most analytical research, the ability to use the models developed in this thesis is limited by the number of simplifying assumptions made.

However, we believe that the qualitative insights drawn such as the risks of decision error and importance of integrated production-marketing modelling are applicable for a wide variety of real-life manufacturing and service systems.

6.2 Recommendations for Future Research

The summary and the conclusions presented in the previous section show our current understanding of the lead time management issues for supply chains. However, the models developed in this research have significant potential for being extended and further evolved. We have discussed some of the possible extensions to the models already developed. In this section we will present some ideas for new models.

- **Assemble-to-Order Environment**

In recent years assemble-to-order production has become very popular. This approach combines the effectiveness of make-to-stock and make-to-order environments by producing components to stock and then assembling them as required by customer orders. Normally the result is faster response than the traditional make-to-order approach, with fewer inventories than a make-to-stock approach (Hopp and Spearman 2000). This is the technique used, for example, by Dell for manufacturing its computers. The models developed in Chapters 3, 4 and 5 can be integrated to develop models for an assemble-to-order environment where demand is random with the mean demand rate being sensitive to both price and delivery time.

Assemble-to-order firms have two options to increase demand through investment. One is to reduce the operating cost by investing in supply lead time reduction for suppliers of components (like in Chapters 4 and 5) so that price can be reduced. The other way is to improve the length and accuracy of the delivery time by investing in increasing its own assembly capacity (like in Chapter 3). If there is limited budget available for investment, then the firm has to optimally allocate it to reduce supply lead time and/or delivery time. Though intuitively it appears that the allocation will depend on the sensitivity of demand rate to price and delivery time, the optimal allocation decision is not clear. Development of models to analytically investigate the issue of optimal allocation of constrained resources to improve efficiency and/or responsiveness for an assemble-to-order environment can be a worthwhile endeavour.

- **Lead Time and Price Based on Customer Sensitivity**

In Chapter 3 we assumed that customers are homogeneous - either price-sensitive or lead-time sensitive. However, recent technology has made it possible to collect customer characteristics in more detail (Smith, Bailey and Brynjolfsson 2000). If some firm has such data then it is imperative to use it to make intelligent decisions. The firm then might quote lower prices and higher delivery times to customers who are more price-sensitive than time-sensitive and vice versa. We feel that priority queueing techniques can be used to model the case where, rather than giving "uniform" delivery time guarantee to all customers, the profit-maximising firm can provide different delivery time (and so charge different price) to different customer niches depending on their price and lead time sensitivity.

- **"Active" Supplier**

In our research, the issue of supply lead time reduction has been investigated assuming that the supplier has no information about the final customer demand and so it passes on part or whole of the investment cost to the buyer (i.e., the supplier is relatively passive). However, if the supplier is aware of the demand sensitivity of the final customers, then it might keep its price and delivery times low by its own choice. Such an action will reduce the inventory cost for the buyer and probably result in increased demand. In case of sole supplier (like ours), this means increased order for the supplier also. Development of an integrated supply chain model with a more "active" supplier that takes into account its pricing and lead time decisions' effect on the whole supply chain is an interesting future direction.

Appendix 3.1

Proof of Proposition 3.5.3

From (3.4.9) and (3.5.9), we can say that the RHS of (3.5.6) is always increasing in L, negative up to $L = \sqrt{\frac{n}{(-b')}}$ and then positive. From (3.4.16),

$$\frac{\partial^2 (RHS)}{\partial L^2} = M_{\mu\mu\mu}(\mu^R(L))\,(\frac{\partial \mu^R}{\partial L})^3 + 3\,M_{\mu\mu}(\mu^R(L))\,(\frac{\partial \mu^R}{\partial L})(\frac{\partial^2 \mu^R}{\partial L^2}) + M_{\mu}(\mu^R(L))(\frac{\partial^3 \mu^R}{\partial L^3}).$$

(A3.1.1)

From (3.4.9) and assuming $M_{\mu\mu\mu} \geq 0$, we have (A3.1.1) is negative for $L \leq \sqrt{\frac{n}{(-b')}}$ and so the RHS of (3.5.6) is concave for $L \leq \sqrt{\frac{n}{(-b')}}$. (A3.1.2)

Differentiating (A3.1.1) with respect to L we have:

$$\frac{\partial^3 (RHS)}{\partial L^3} = M_{\mu\mu\mu\mu}(\mu^R(L))\,(\frac{\partial \mu^R}{\partial L})^4 + 6M_{\mu\mu\mu}(\mu^R(L))\,(\frac{\partial \mu^R}{\partial L})^2(\frac{\partial^2 \mu^R}{\partial L^2}) +$$

$$3M_{\mu\mu}(\mu^R(L))(\frac{\partial^2 \mu^R}{\partial L^2})^2 + 4M_{\mu\mu}(\mu^R(L))(\frac{\partial \mu^R}{\partial L})(\frac{\partial^3 \mu^R}{\partial L^3}) + M_{\mu}(\mu^R(L))(\frac{\partial^4 \mu^R}{\partial L^4}).$$

(A3.1.3)

From (3.4.9) and assuming that $M_{\mu\mu\mu} \geq 0$ and $M_{\mu\mu\mu\mu} \geq 0$, we see that only the third term of (A3.1.3) is negative while all others are positive. A sufficient condition for $\frac{\partial^2 (RHS)}{\partial L^2}$ to be increasing is:

$$3M_{\mu\mu}(\mu^R(L))(\frac{\partial\mu^R}{\partial L})(\frac{\partial^3\mu^R}{\partial L^3}) \geq 4M_{\mu\mu}(\mu^R(L))(\frac{\partial\mu^R}{\partial L})(\frac{\partial^3\mu^R}{\partial L^3}). \qquad \text{(A3.1.4)}$$

We can show that condition (A3.1.4) implies that for any $L \geq \sqrt{\frac{n}{(-2b')}}$, $\frac{\partial^2(RHS)}{\partial L^2}$ is increasing.

(A3.1.5)

If (A3.1.1) is positive for higher L (otherwise RHS will be throughout concave), from (A3.1.2) and (A3.1.5) it is clear that there is a unique solution of $\frac{\partial^2(RHS)}{\partial L^2} = 0$ (say L^s) and $\frac{\partial^2(RHS)}{\partial L^2}$ is negative for $L \leq L^s$ and positive for $L > L^s$. Also, $L^s > \sqrt{\frac{n}{(-b')}}$, implying that RHS will be concave till L^s and then it will be convex and at L^s, the RHS will be positive.

Appendix 4.1

Proof of Equivalence of Inventory Cost Function with the Transformations $\hat{X} = \alpha X$ and $\hat{X} = \alpha X + (1-\alpha)\mu$

With $\hat{X} = \alpha X$,

$$C(Q,r,\alpha) = \frac{Kd}{Q} + h(r + \frac{Q}{2} - \alpha\mu) + \frac{(h+b)\alpha}{Q}[J1(r) - J1(r+Q)] + pd\,, \quad \text{(A4.1.1)}$$

where $J1(r) = \int_{s=r}^{\infty} [\int_{x=(s/\alpha)}^{\infty} \overline{F}(x)dx]\, ds.$

With $\hat{X} = \alpha X + (1-\alpha)\mu$,

$$C(Q,r,\alpha) = \frac{Kd}{Q} + h(r + \frac{Q}{2} - \mu) + \frac{(h+b)\alpha}{Q}[J2(r) - J2(r+Q)] + pd\,, \quad \text{(A4.1.2)}$$

where $J2(r) = \int_{s=r}^{\infty} [\int_{x=[s-(1-\alpha)\mu]/\alpha}^{\infty} \overline{F}(x)dx]ds.$

Let us define $Z(x) = \int \overline{F}(x)$ (for this Appendix only).

Then, $J1(r_{mv}) = \int_{s=r_{mv}}^{\infty} [\int_{x=(s/\alpha)}^{\infty} \overline{F}(x)dx]ds$ (r_{mv} is the reorder point for $\hat{X} = \alpha X$)

$$= \int_{s=r_{mv}}^{\infty} [Z(\infty) - Z(\frac{s}{\alpha})]ds\,. \quad \text{(A4.1.3)}$$

With $s = t - (1-\alpha)\mu$,

$$\text{(A4.1.3)} = \int_{t=r_{mv}+(1-\alpha)\mu}^{\infty} [Z(\infty) - Z(\frac{t-(1-\alpha)\mu}{\alpha})]dt$$

$$= \int_{t=r_v}^{\infty} (\int_{x=[t-(1-\alpha)\mu]/\alpha}^{\infty} \overline{F}(x)dx)\, dt$$

$= J2(r_v)$ (r_v is the reorder point for $\hat{X} = \alpha X + (1-\alpha)\mu$)

where $r_v = r_{mv} + (1-\alpha)\mu$.

Similarly, $J1(r_{mv}+Q)=J2(r_v+Q)$ and clearly $h(\frac{Q}{2}+r_{mv}-\alpha\mu)=h(\frac{Q}{2}+r_v-\mu)$.

Therefore, from (A4.1.1) and (A4.1.2) we can say that Q and C are same for $\hat{X}=\alpha X$ as for $\hat{X}=\alpha X+(1-\alpha)\mu$, and $r_v=r_{mv}+(1-\alpha)\mu \ \forall \ \alpha$. This proof is applicable for any distribution, as it does not use any particular property of $\overline{F}(x)$.

Appendix 4.2

Proof of Convexity for Model 2B (exact)

To prove the convexity of "exact" Model 2B we must prove that the extra term (in this Appendix M will represent $M^{IC}(\alpha)$),

$$[\theta M(\alpha)i(1+i)^{N^C}]/[(1+i)^{N^C}-1] = A \text{ (for this Appendix only)},$$

is jointly convex in Q, r and α (all the other terms are jointly convex). Since the extra term is independent of r, we need to prove the joint convexity of A in Q and α.

Let $\theta i = C1$, $(1 + i) = x$, and $T^C/\lambda = C2$ (for this Appendix only). Then we have:

$$A_\alpha = \frac{(C1)(M_\alpha)x^{(C2)Q}}{x^{(C2)Q}-1}, \; A_{\alpha\alpha} = \frac{(C1)(M_{\alpha\alpha})x^{(C2)Q}}{x^{(C2)Q}-1},$$

$$A_Q = (C1)(M)\frac{x^{(C2)Q}(C2)ln(x)}{x^{(C2)Q}-1} - \frac{(x^{(C2)Q})^2(C2)ln(x)}{(x^{(C2)Q}-1)^2},$$

$$A_{QQ} = (C1)(M)(C2)^2[ln(x)]^2\left[\frac{x^{(C2)Q}(x^{(C2)Q}+1)}{(x^{(C2)Q}-1)^3}\right],$$

$$A_{Q\alpha} = -(C1)(M_\alpha)(C2)[ln(x)]\left[\frac{x^{(C2)Q}}{(x^{(C2)Q}-1)^2}\right].$$

It is not difficult to show that $A_{\alpha\alpha} \geq 0$, $A_{QQ} \geq 0$ and $A_{Q\alpha} \geq 0$.

The principal determinant of the hessian matrix ($A_{\alpha\alpha} A_{QQ} - A_{Q\alpha}^2$) on simplification is,

$$(C1)^2(C2)^2 \left[\frac{(x^{(C2)Q})^2}{(x^{(C2)Q}-1)^4}\right] (ln(x))^2[(x^{(C2)Q} + 1)(M_{\alpha\alpha})(M) - (M_{\alpha})^2].$$

Since $(x^{(C2)Q} + 1) \geq 1$, $(M_{\alpha\alpha})(M) - (M_{\alpha})^2 \geq 0$ is sufficient for the hessian to be positive and hence for A to be jointly convex in Q and α.

Appendix 4.3

Proof of Convexity for Model 1B

To prove joint convexity of Model 1B we have to prove the joint convexity of the term,

$E = \dfrac{[K + \theta M]\lambda}{Q}$, in Q and α (for this Appendix M will represent $M^{RC}(\alpha)$).

$$E_Q = -\frac{[K + \theta M]\lambda}{Q},$$

$$E_{QQ} = \frac{2[K + \theta M]\lambda}{Q^3},$$

$$E_\alpha = \frac{\theta(M_\alpha)\lambda}{Q},$$

$$E_{\alpha\alpha} = \frac{\theta(M_{\alpha\alpha})\lambda}{Q},$$

$$E_{Q\alpha} = -\frac{\theta(M_\alpha)\lambda}{Q^2}.$$

It is easy to see that $E_{QQ} \geq 0$, $E_{\alpha\alpha} \geq 0$ and $E_{Q\alpha} \geq 0$.

The principal determinant of the hessian matrix ($E_{\alpha\alpha} E_{QQ}$ - $E_{Q\alpha}^2$) on simplification is,

$$\frac{2K\theta(M_{\alpha\alpha})\lambda}{Q^4} + \frac{2\theta^2\lambda^2(M)(M_{\alpha\alpha})}{Q^4} - \frac{\theta^2(M_\alpha)^2\lambda^2}{Q^4}.$$

It is then easy to show that $2(M_{\alpha\alpha})(M) \geq (M_\alpha)^2$ is a sufficient condition for the determinant to be positive and hence for the term E to be jointly convex in Q and α.

Appendix 4.4

Proof of Convexity for Models 1A and 2A with respect to α for Fixed Q and r

The cost function for Models 1A and 2A (for fixed Q and r) is given by:

$$C(\alpha) = \frac{K\lambda}{Q} + \frac{\alpha}{Q}\int_r^{r+Q} G(\alpha, y)dy + c(\alpha)\lambda, \qquad \text{(A4.4.1)}$$

where $G(\alpha, y) = (h(\alpha) + b)\int_0^{(y/\alpha)} F(x)dx + b(\mu - (y/\alpha))$, $h(\alpha) = ic(\alpha)$.

The only difference in the two models is that for Model 1A, $c(\alpha)$ is given by $[c + \theta M^{RU}(\alpha)]$, while for Model 2A, $c(\alpha)$ is given by $c + [\theta M^{IU}(\alpha)(i/\lambda)]$. However, this difference will not matter since our aim here is to prove the convexity of the cost function with respect to α for fixed Q and r and both i and λ are constant parameters.

On simplification $C(\alpha)$ can be written as:

$$C(\alpha) =$$
$$\frac{K\lambda}{Q} + \frac{\alpha(h(\alpha)+b)}{Q}\int_r^{r+Q}\left[\int_0^{(y/\alpha)} F(x)dx\right]dy + \alpha b\mu - \frac{1}{2Q}\left[(r+Q)^2 - r^2\right] + c(\alpha)\lambda.$$

(A4.4.2)

Since we have assumed that Q and r are constants and $c(\alpha)$ is by assumption convex in α, to prove the convexity of $C(\alpha)$ in α we need to prove that,

$$\frac{\alpha(h(\alpha)+b)}{Q}\int_r^{r+Q}\left[\int_0^{(y/\alpha)} F(x)dx\right]dy = T \text{ (for this Appendix only)},$$

is convex in α, i.e., $T_{\alpha\alpha} \geq 0$.

Then (h represents $h(\alpha)$),

$$T_\alpha = -\frac{(h+b)}{Q}\int_r^{r+Q}\left[\int_0^{(y/\alpha)} xf(x)dx\right]dy + \frac{\alpha h_\alpha}{Q}\int_r^{r+Q}\left[\int_0^{(y/\alpha)} F(x)dx\right]dy,$$

and

$$T_{\alpha\alpha} = \frac{(h+b)}{Q}\int_r^{r+Q}\left[(y/\alpha^3)f(y/\alpha)\right]dy + \frac{\alpha h_{\alpha\alpha}}{Q}\int_r^{r+Q}\left[\int_0^{(y/\alpha)} F(x)dx\right]dy$$

$$-\frac{2h_\alpha}{Q}\int_r^{(r+Q)}\left[\int_0^{(y/\alpha)} xf(x)dx\right]dy.$$

If r is positive, it is easy to see that $T_{\alpha\alpha} \geq 0$ (since $h_\alpha \leq 0$ and $h_{\alpha\alpha} \geq 0$).

Appendix 4.5

Condition under which Cost Functions of Models 1A and 2A will be "Univalleyed"

The cost function for Model 1A is,

$$C(Q,r,\alpha)=\frac{K\lambda}{Q}+\frac{1}{Q}\int_r^{r+Q}[\alpha G(\alpha,y)]\,dy+\frac{h(\alpha)}{i}\lambda,$$

where $G(\alpha,y)=(h(\alpha)+b)\int_0^{(y/\alpha)}F(x)dx+b[\mu-(y/\alpha)]$,

$h(\alpha)=ic(\alpha)$ and $c(\alpha)=c+\theta M^{RU}(\alpha)$.

We know that C is convex in Q and r for fixed α. Let $Q^*(\alpha)$ and $r^*(\alpha)$ be the optimal Q and r for a fixed α. This implies:

$$C(Q^*(\alpha),r^*(\alpha),\alpha)=\frac{K\lambda}{Q^*(\alpha)}+\frac{\alpha}{Q^*(\alpha)}\int_{r^*(\alpha)}^{r^*(\alpha)+Q^*(\alpha)}G(\alpha,y)dy+\frac{h(\alpha)}{i}\lambda.\tag{A4.5.1}$$

In the sequel, we suppress the argument of $Q^*(\alpha)$ and $r^*(\alpha)$.

From (A4.5.1),

$$\begin{aligned}C_\alpha &=\frac{\partial Q^*}{\partial\alpha}\cdot\frac{\partial C}{\partial Q^*}+\frac{\partial r^*}{\partial\alpha}\cdot\frac{\partial C}{\partial r^*}+\frac{1}{Q^*}\int_{r^*}^{r^*+Q^*}G(\alpha,y)dy\\ &\quad+\frac{\alpha}{Q^*}\int_{r^*}^{r^*+Q^*}G_\alpha(\alpha,y)dy+\frac{h_\alpha(\alpha)}{i}\lambda.\end{aligned}$$

By design $\frac{\partial C}{\partial Q^*}$ and $\frac{\partial C}{\partial r^*}$ are both equal to zero, implying that,

$$C_\alpha=\frac{1}{Q^*}\int_{r^*}^{r^*+Q^*}[\frac{d}{d\alpha}(\alpha G(\alpha,y))]dy+\frac{h_\alpha(\alpha)\lambda}{i}.\tag{A4.5.2}$$

To prove that C is univalleyed, it is enough to show that (A4.5.2) = 0 has a unique (or no) solution for $0 \leq \alpha \leq 1$. The condition will be similar for Model 2A with the only difference being that $h(\alpha) = i(c + \frac{\theta M^{1U}(\alpha)i}{\lambda})$.

Appendix 4.6

Optimal Lead Time Duration with Random Demand

It is well known that when demand and lead time duration are both random, the mean and variance of the lead time demand are given by:

$$\mu = \lambda\mu_L \qquad \text{(A4.6.1)}$$

and

$$\sigma^2 = (\sigma_\lambda)^2\mu_L + \lambda^2(\sigma_L)^2, \qquad \text{(A4.6.2)}$$

where μ and σ^2 are the respective mean and variance of the lead time demand distribution, μ and $(\sigma_L)^2$ are the respective mean and variance of the lead time duration distribution and λ and $(\sigma_\lambda)^2$ are the respective mean and variance of the demand distribution.

When the demand is constant (i.e., $(\sigma_\lambda)^2 = 0$ and λ is just the demand rate) and μ and σ^2 at status-quo are known, we can easily find μ_L and $(\sigma_L)^2$ at $\alpha = 1$. The optimal α, α^*, gives us the optimal mean $(\mu_L{}^*)$ and variance $(\sigma_L{}^*)^2$ of the lead time duration:

$$\mu_L{}^* = (\alpha^*)(\mu_L \text{ at } \alpha = 1), \qquad \text{(A4.6.3)}$$

and

$$(\sigma_L{}^*)^2 = (\alpha^*)^2(\sigma_L{}^2 \text{ at } \alpha = 1). \qquad \text{(A4.6.4)}$$

Suppose the demand is random but the reduction in lead time demand comes *solely* from the reduction in lead time duration. The mean and variance of the demand is presumed to be the same for any lead time duration. Let the reduced lead time demand random variable be $\hat{X} = \alpha X$ and there is an investment in reducing α. Like in the constant demand case, we can determine the optimal α^* that minimizes the buyer's cost. Since all our analytical and numerical results are based on LTD distribution and cost associated with changing $\hat{X}$ (i.e., α is our decision variable), all of them will still hold.

If the mean and variance of the lead time demand distribution and demand distribution at status-quo are known, we can determine the mean and variance of the lead time duration at status-quo. For random demand, the optimal mean and variance of the lead time duration then will be given by:

$$\mu_L{}^* \;=\; (\alpha^*)(\mu_L \text{ at } \alpha = 1), \tag{A4.6.5}$$

and $$(\sigma_L{}^*)^2 = \frac{\{(\alpha^*)^2(\sigma^2 \text{ at } \alpha = 1)\} - \{(\sigma_\lambda)^2 \text{ at } \alpha = 1)(\mu_L{}^*)\}}{\lambda^2}. \tag{A4.6.6}$$

From (A4.6.5) and (A4.6.6) it is clear that we can determine the optimal mean and variance of lead time duration for random demand case also.

Appendix 4.7

Analytical Comparative Statics of $r^*(\alpha)$ and $\{r^*(\alpha)+Q^*(\alpha)\}$ with respect to α for Unit, Time and Cycle Models

The FOCs for the time model (recalling that all three models are jointly convex in Q and r for fixed α) can be written as:

$$C(Q,r)-\theta M^{RT}-c\lambda \quad = \quad \alpha G(\alpha,r)=\alpha G(\alpha,r+Q)\ , \qquad \text{(A4.7.1)}$$

where $C(Q,\, r) = \dfrac{K\lambda}{Q}+\dfrac{\alpha}{Q}\displaystyle\int_r^{r+Q} G(\alpha,y)dy + c\lambda + \theta M^{RT}$, with α as parameter.

In the following analysis, r and Q represent $r^*(\alpha)$ and $Q^*(\alpha)$ respectively.

We know from Zheng (1992) that $b \geq (b+h)F(r/\alpha)$. Total differentiation with respect to α of $\alpha G(\alpha,r)=C(Q,r)-\theta M^{RT}-c\lambda$ at r and Q gives (recall that $(\partial C/\partial Q)$ and $(\partial C/\partial r)$ are by design equal to zero):

$$\frac{\partial \alpha G(\alpha,r)}{\partial \alpha}+\frac{\partial \alpha G(\alpha,r)}{\partial r}\frac{\partial r}{\partial \alpha} \quad = \quad \frac{1}{Q}\int_r^{r+Q}[\alpha G_\alpha(\alpha,y)+G(\alpha,y)]\ dy\ . \qquad \text{(A4.7.2)}$$

Simplifying (A4.7.2) we have:

$$\frac{\partial r}{\partial \alpha} \quad = \quad \frac{\dfrac{1}{Q}\displaystyle\int_r^{r+Q} Z(y)dy - Z(r)}{\alpha G_r(\alpha,r)}, \qquad \text{(A4.7.3)}$$

where $Z(y)=\alpha G_\alpha(\alpha,y)+G(\alpha,y)$ and $G_r(\alpha,r)=(b+h)F(r/\alpha)-b$.

$Z(y)$ is a decreasing function of y. We can also show that $Z(r+Q) \geq 0$ if and only if $b/(b+h) \geq [(\int_0^{(r+Q)/\alpha} xf(x)dx)/(\int_0^{\infty} xf(x)dx)]$. This condition will most probably hold for $b >> h$, unless $\alpha \cong 0$, but then $(r + Q)$ will also be $\cong 0$ and so $(r + Q)/\alpha$ will probably still be finite. If we assume that $Z(r+Q) \geq 0$, then both $Z(r)$ and $Z(r + Q)$ are positive and we can show that $\frac{1}{Q}\int_r^{r+Q} Z(y)dy - Z(r) \leq 0$ and $\frac{1}{Q}\int_r^{r+Q} Z(y)dy - Z(r+Q) \geq 0$. So, from (A4.7.3), $(\partial r/\partial\alpha) \geq 0$ since both the numerator and the denominator are negative.

Using total differentiation with respect to α of $\alpha G(\alpha, r+Q) = C(Q,r) - \theta M^{RT} - c\lambda$ and recalling that $G_{r+Q}(\alpha, r+Q) = (b+h)F\left(\frac{r+Q}{\alpha}\right) - b \geq 0$ (Zheng 1992), we can similarly show that $[(\partial r/\partial\alpha) + (\partial Q/\partial\alpha)] \geq 0$. But we cannot analytically determine the sign of $(\partial Q/\partial\alpha)$.

For the unit model, the procedure will remain exactly the same as before with M^{RT} replaced by M^{RU} and h replaced by $h(\alpha) = i[c + \theta M^{RU}]$. In this case, though $Z(y)$ is still a decreasing function, the condition for $Z(r+Q) \geq 0$ will be different. From our numerical experiments we can tell that $Z(r + Q)$ will still be positive in almost all cases. Assuming that $Z(r + Q) \geq 0$ and following the same method as for time model, we can prove that $(\partial r/\partial\alpha) \geq 0$ and $[(\partial r/\partial\alpha) + (\partial Q/\partial\alpha)] \geq 0$.

For the cycle model, $C_Q = 0$ will be:

$$\frac{(K + \theta M^{RC})\lambda}{Q^*(\alpha)} + \frac{1}{Q^*(\alpha)}\int_{r^*(\alpha)}^{r^*(\alpha)+Q^*(\alpha)} [\alpha G(\alpha, y)]dy = \alpha G(\alpha, r^*(\alpha) + Q^*(\alpha)). \tag{A4.7.4}$$

Following the procedure as before (r and Q representing $r^*(\alpha)$ and $Q^*(\alpha)$ respectively) we will have:

$$\frac{\partial r}{\partial \alpha}=\frac{\frac{\lambda}{Q}\theta M_{\alpha}^{RC}}{\alpha G_{r}(\alpha,r)}+\frac{\frac{1}{Q}\int_{r}^{r+Q}Z(y)dy-Z(r)}{\alpha G_{r}(\alpha,r)}, \quad \text{(A4.7.5)}$$

where $Z(y)$ and $G(\alpha, y)$ are as in the time model.

As $\frac{1}{Q}\int_{r}^{r+Q}Z(y)dy-Z(r)$, $\alpha G_{r}(\alpha,r)$ and θM_{α}^{RC} are all negative, we can say that $(\partial r/\partial \alpha)\geq 0$. But now,

$$\frac{\partial r}{\partial \alpha}+\frac{\partial Q}{\partial \alpha}=\frac{\frac{\lambda}{Q}\theta M_{\alpha}^{RC}}{\alpha G_{r+Q}(\alpha,r+Q)}+\frac{\frac{1}{Q}\int_{r}^{r+Q}Z(y)dy-Z(r+Q)}{\alpha G_{r+Q}(\alpha,r+Q)}. \quad \text{(A4.7.6)}$$

While the second part of RHS is positive, the first part is negative (since the numerator is negative while the denominator is positive). So, for the cycle model $[(\partial r/\partial \alpha)+(\partial Q/\partial \alpha)]$ can be positive or negative depending on the value of θM_{α}^{RC}.

Appendix 5.1

Proof that Profit with the Root with Minus Sign is Higher

The profit function is given by:

$$\pi = (p - m)\lambda = (a\eta^{(-2)}(\eta\text{-}1))m^{(-1)}, \tag{A5.1.1}$$

and the two solutions to m are given by:

$$m = \frac{1 \pm \sqrt{1-4uv}}{2u}, \tag{A5.1.2}$$

where:

$$u = c(Q)Q\frac{i}{2a\eta^{-2}}, \tag{A5.1.3}$$

and

$$v = \frac{K}{Q} + c(Q). \tag{A5.1.4}$$

Substituting the two roots in (A5.1.2) in (A5.1.1) we have:

π(negative root of m) - π(positive root of m)

$$= (a\eta^{(-2)}(\eta\text{-}1))\frac{\sqrt{1-4uv}}{v}. \tag{A5.1.5}$$

If $\sqrt{1-4uv} > 0$, then (A5.1.5) is strictly positive implying that the profit function with the negative root of m will always be greater than the profit function with the positive root of m.

Appendix 5.2

Proof of Semi-Strict Quasiconcavity of π

Remembering that c is a function of Q, we can easily show that:

i) $K(c_{QQ}) + Qc(c_{QQ}) + 2Q(c_Q)^2 + 4c(c_Q) \geq 0$ is sufficient for convexity of uv,

and

ii) $2c_Q + Q(c_{QQ}) \leq 0$ is sufficient for concavity of u.

Hence, $(1 - 4uv)$ is concave in Q and since $(1 - 4uv)$ is positive for feasible Q, this implies that $\sqrt{1-4uv}$ is also positive concave. Since the denominator of (5.3.9) is convex in Q and the numerator is concave (both positive), we can conclude that π is semi-strictly quasiconcave in Q (Schaible, 1981).

Both conditions will be satisfied when the unit purchase/production cost is constant, i.e., $c(Q) = c$ (> 0) $\forall$ Q. The condition $2c_Q + Q(c_{QQ}) \leq 0$ is also satisfied by most of the common non-increasing convex unit cost functions we expect to see in the literature - linear, power and logarithmic. These types of functions, but under some additional constraints, will also satisfy the condition $K(c_{QQ}) + Qc(c_{QQ}) + 2Q(c_Q)^2 + 4c(c_Q) \geq 0$.

Appendix 5.3

Proof of Proposition 5.3.1

With 1 - $\sqrt{1-4uv}$ = T, π in (5.3.9) can be written as:

$$\pi = \frac{2a\eta^{-2}(\eta-1)u}{T}. \qquad \text{(A5.3.1)}$$

Differentiating (A5.3.1) twice with respect to Q, we get the following expression for π_{QQ} which is a cubic equation in Q:

$$(u_{QQ})T^2 - uT(T_{QQ}) - 2T(T_Q)(u_Q) + 2u(T_Q)^2. \qquad \text{(A5.3.2)}$$

For constant unit purchase/production cost, $u_{QQ} = 0$ and $T_{QQ} \geq 0$. Of the three solutions to (A5.3.2), two will be complex and one will be negative. So, for any positive Q, (A5.3.2) has the same sign. Note that T is increasing linear in Q, i.e., $T_Q \geq 0 \; \forall \; Q$. It is now easy to prove that (A5.3.2) is negative for any feasible $Q \leq Q^*$ (solution to $u_QT - T_Qu = 0$). Since we know that (A5.3.2) will have the same sign for any positive Q, it must be negative for all feasible Q. Therefore, π is concave for all feasible Q.

Appendix 5.4

Derivation of $\partial Q^*(K)/\partial K$ and $\partial^2 Q^*(K)/\partial K^2$

For Section 5.3.1.3,

$$\pi_Q = 0 \Rightarrow u_Q T - T_Q u = 0. \qquad \text{(A5.4.1)}$$

Differentiating (A5.4.1) with respect to K we have:

$$(u_{QQ}\frac{\partial Q^*(K)}{\partial K} + u_{QK})T + u_Q(T_Q\frac{\partial Q^*(K)}{\partial K} + T_K) - (T_{QQ}\frac{\partial Q^*(K)}{\partial K} + T_{QK})u$$
$$- (u_Q\frac{\partial Q^*(K)}{\partial K} + u_K)T_Q = 0. \qquad \text{(A5.4.2)}$$

Rearranging we have the expression in (5.3.17).

Differentiating (5.3.17) with respect to K we have:

$$\frac{\partial^2 Q^*(K)}{\partial K^2} =$$
$$\frac{1}{(T_{QQ}u)^2}[\{(u_{QQ}\frac{\partial Q^*(K)}{\partial K} + u_{QK})T_K + (T_{QK}\frac{\partial Q^*(K)}{\partial K} + T_{KK})u_Q\}T_{QQ}u - \{(u_Q\frac{\partial Q^*(K)}{\partial K} + u_K)T_{QK}$$
$$+ (T_{QKK} + T_{QKQ}\frac{\partial Q^*(K)}{\partial K})u\}T_{QQ}u - \{(u_Q\frac{\partial Q^*(K)}{\partial K} + u_K)T_{QQ} + (T_{QQK} + T_{QQQ}\frac{\partial Q^*(K)}{\partial K})u\}$$
$$(u_Q T_K - uT_{QK}). \qquad \text{(A5.4.3)}$$

Appendix 5.5

Proof of Proposition 5.3.6

The condition for concavity of π is given in (A5.3.2) of Appendix 5.3. The condition can be written as (recalling that $T = 1 - \sqrt{1-4uv}$):

$$-\frac{u[(uv)_Q]^2}{1-4uv}\{\frac{4}{\sqrt{1-4uv}}-12\}+(2u(uv)_{QQ}+4u^Q(uv)^Q)\{1-\frac{1}{\sqrt{1-4uv}}\} \leq 0. \tag{A5.5.1}$$

Note that $[1 - (1/\sqrt{1-4uv})]$ and $-(u[(uv)_Q]^2)/(1 - 4uv)$ are negative since $0 < 4uv \leq 1$ for feasible Q. Then the sufficient conditions for concavity of π are: i) $\sqrt{1-4uv} \leq (1/3)$, and ii) $2u(uv)_{QQ} + 4u_Q(uv)_Q \geq 0$. It is possible to show that

$$2Q(uv)_{QQ} + 4(uv)_Q \geq 0. \tag{A5.5.2}$$

From (A5.5.2) it is possible to show that the condition $2u(uv)_{QQ} + 4u_Q(uv)_Q \geq 0$ will always be satisfied for feasible Q. So, $\sqrt{1-4uv} \leq (1/3)$ is sufficient for concavity of π.

Appendix 5.6

Proof of $T_{KKK} \leq 0$ for $T_K \leq 0$

Differentiation of T_{KK} in (5.3.28) with respect to K after some simplification yields:

$$T_{KKK} = \frac{2}{X^4}\{((wz)_{KKK}X + \frac{2(wz)_K(wz)_{KK}}{X} + \frac{4[(wz)_K]^2(wz)_K}{X^3})(X^2)$$

$$-(-)4(wz)_K(Z))\}, \tag{A5.6.1}$$

where $X = \sqrt{1-(wz)}$ and Z is the numerator of T_{KK} in (5.3.28) (for this Appendix only).

Noting that feasible K requires $(1 - wz) \geq 0$, i.e., $X \geq 0$, and we are interested in $T_K \leq 0$, from the expression of (5.3.28) we can tell that $(wz)_K \leq 0$. For both investment functions we can show that $(wz)_K \leq 0$, $(wz)_{KK} \geq 0$, $(wz)_{KKK} \leq 0$ and $Z \geq 0$. Then from (A5.6.1) we can prove that $T_{KKK} \leq 0$.

REFERENCES

1. Anderson, D. L., Britt, F. E. and Favre, D. J. (1997). "The Seven Principles of Supply Chain Management", *Supply Chain Management Review*, Spring.
2. Anthony, M. (1998). "Breakthrough Cost Reduction: An Enterprise View", *RONIN Corporation Position Paper Series*, (www.ronin.com).
3. Australian National Audit Office Report. (1997-98). "Performance Management of Defence Inventory", Audit Report No. 5, Part 2, Section 3-4.
4. Bagchi, U., Hayya, J. C. and Chu, C-H. (1986). "The Effect of Lead Time Variability: The Case of Independent Demand", *Journal of Operations Management,* 6, 2, 159 - 177.
5. Ballou, R. H. (1998). *Business Logistics Management*, Prentice Hall, Upper Saddle River, NJ, USA.
6. Barron, K. (1999). "Wal-Mart's Ankle Biters", *Forbes Global*, 18th October.
7. Ben-Daya, M. and Raouf, A. (1994). "Inventory Models Involving Lead Time as a Decision Variable", *Journal of Operational Research Society,* 45, 5, 579-582.
8. Bensaou, M. (1999). "Portfolios of Buyer-Supplier Relationships", *Sloan Management Review*, Summer, 35-44.
9. Blackburn, J. D. (1991). *Time-Based Competition: The Next Battleground in American Manufacturing*, Irwin, Homewood, IL.
10. Blackburn, J. D., Elrod, T., Lindsley, W. B. and Zahorik, A. J. (1992). "The Strategic Value of Response Time and Product Variety", Chapter 13, *Manufacturing Strategy - Process and Content*, C. A. Voss (editor), Chapman and Hall, London.
11. Bliss, C. (1988). "A Theory of Retail Pricing", *Journal of Industrial Economics*, 36, 4, 375-391.
12. Bloch, H. and Olive M. (1997). "Pass-Through Elasticities for Production Costs and Competing Foreign Prices: Evidence from Manufacturing Prices in Seven Countries", Working Paper, Department of Economics, Curtin Business School, Curtin University of Technology, Perth, Western Australia (www.cbs.curtin.edu.au/ef/9708.html).

13. Bookbinder, J. H. and Çakanyildirim, M. (1999). "Random Lead time and Expedited Orders in (Q, r) Inventory Systems", *European Journal of Operational Research*, 115, 300-313.

14. Bryne, J.A. (1996). "Has Outsourcing gone too Far?", *Business Week*, April 1, 26-28.

15. Buss, A. H., Lawrence, S. R. and Kropp, D. H. (1994). "Volume and Capacity Interaction in Facility Design", *IIE Transactions*, 26, 4, 36-49.

16. Cachon, G. P. and Fisher, M. L. (2000). "Supply Chain Inventory Management and the Value of Shared Information", *Management Science*, 46, 8, 1032-1048.

17. CAPS Research (2000). "Cross-Industry Comparison of Standard Benchmarks", Center for Advanced Purchasing Studies, Tempe, AZ, USA.

18. Carbone, J. (2000). "Growth Means More Consolidation, More Services", *Purchasing Online*, 129, 7, 48-57.

19. Center for Retailing Education and Research. (2001). "Refacts from Retailing Management", 8th March (http://www.cba.ufl.edu/CRER/).

20. Cheng, T. C. E. (1990). "An EOQ Model with Pricing Consideration", *Computers and Industrial Engineering*, 18, 4, 529-534.

21. Choi, J-W. (1994). "Investment in the Reduction of Uncertainties in Just-in-Time Purchasing Systems", *Naval Research Logistics*, 41, 257-272.

22. Chopra, S. and Meindl, P. (2001). *Supply Chain Management: Strategy, Planning and Operation*, Prentice Hall, Upper Saddle River, New Jersey.

23. Connor, R. A., Feldman, R. D. and Dowd, B. E. (1998). "The Effects of Market Concentration and Horizontal Mergers on Hospital Costs and Prices", *International Journal of the Economics of Business*, 5, 2, 159-180.

24. Coutts, K., Godley, W. and Nordhaus, W. (1978). *Industrial Pricing in the United Kingdom*, Cambridge University Press, Cambridge, UK.

25. Das, C. (1975). "Effect of Lead Time on Inventory: A Static Analysis", *Operational Research Quarterly*, 26, 2, I, 273-282.

26. De Groote, X. and Zheng, Y. S. (1997). "The Effect of Demand Variability on Inventory Systems", Proceedings of memorial held in INSEAD, France, in honour of Xavier De Groote.

27. Deng, S. and Yano, C. A. (2000). "Joint Production and Pricing Decisions with Set-up Costs and Capacity Constraints", Working Paper, Department of Industrial Engineering and Operations Research, University of California, Berkeley.
28. Dewan, S. and Mendelson, H. (1990). "User Delay Costs and Internal Pricing for a Service Facility", *Management Science*, 36, 12, 1502-1507.
29. Duenyas, I. and Hopp, W. J. (1995). "Quoting Customer Lead Times", *Management Science*, 41, 1, 43-57.
30. Eckstein, O. and Fromm, G. (1968). "The Price Equation", *American Economic Review*, 58, 1158-1183.
31. E-commerce News. (2000). "Consumer Priorities Shifting from Best Merchandise Price to Total Transaction Cost", July 25th.
32. Eliashberg, J. and Steinberg, R. (1993). "Marketing-Production Joint Decision Making", Chapter 18, *Marketing*, Volume 5, *Handbooks in Operations Research and Management Science*, J. Eliashberg and G. L. Lilien (eds.), North-Holland (Elsevier Science Publishers B.V.), Amsterdam, The Netherlands.
33. Eppen, G. D. and Martin, R. K. (1988). "Determining Safety Stock in Presence of Stochastic Lead Times", *Management Science*, 34, 11, 1380-1390.
34. Fazel, F., Fisher, K. P. and Gilbert, E. W. (1998). "JIT Purchasing vs. EOQ with a Price Discount: An Analytical Comparison of Inventory Costs", *International Journal of Production Economics*, 54, 101-109.
35. Federgruen, A. and Zheng, Y-S. (1992). "An Efficient Algorithm for Computing an Optimal (r, Q) policy in Continuous Review Stochastic Inventory Systems", *Operations Research*, 40, 808-813.
36. Fisher, M. L. (1997). "What is the Right Supply Chain for Your Product?", *Harvard Business Review*, March-April, 105-116.
37. Fraser, N. M., Bernhardt, I., Jewkes, E. M. and Tajima, M. (2000). *Engineering Economics in Canada*, 2nd Edition, Prentice-Hall Canada Inc., Scarborough, Ontario.
38. Frazier, G. L., Speakman, R. E. and O' Neal, C. R. (1988). "Just-in-Time Exchange Relationships in Industrial Markets?", *Journal of Marketing*, 52, 4, 52-67.

39. Gallego, G. (1998). "New Bounds and Heuristics for (Q, r) Policies", *Management Science*, 44, 2, 219-233.

40. Garg, A. and Lee, H. L. (1999). "Managing Product Variety: An Operations Perspective", Chapter 17, *Quantitative Models for Supply Chain Management*, S. Tayur, R. Ganeshan and M. J. Magazine (eds.), Kluwer Academic Publishers, Boston.

41. Geary, S. and Zonnenberg, J. P. (2000). "What it Means to be Best in Class?", *Supply Chain Management Review*, 4, 3, 42-49.

42. Gerchak, Y. (1990). "Analytical Comparative Statics for the Continuous Review Inventory Model", *Operations Research Letters*, 9, 215-217.

43. Gerchak, Y. and Gupta, D. (1991). "On Apportioning Costs to Customers in Centralised Continuous Review Inventory Systems", *Journal of Operations Management*, 10, 4, 546-551.

44. Gerchak, Y. and He, Q-M. (2000). "On the Relation between the Benefits of Risk Pooling and the Variability of Demand", Working Paper, Department of Industrial Engineering, Tel-Aviv University, Israel, December.

45. Gerchak, Y. and Mossman, D. (1992). "On the Effect of Demand Randomness on Inventories and Costs", *Operations Research*, 40, 804-807.

46. Gerchak, Y. and Parlar, M. (1990). "Yield Randomness, Cost Trade-offs and Diversification in the EOQ Model", *Naval Research Logistics*, 37, 341-354.

47. Gerchak, Y. and Parlar, M. (1991). "Investing in Reducing Lead Time Randomness in Continuous Review Inventory Models", *Engineering Costs and Production Economics*, 21, 191-197.

48. Gerchak, Y., Hassini, E. and Ray, S. (2000). "Capacity Selection Under Uncertainty with Ratio Objectives", Working Paper, Department of Management Sciences, University of Waterloo, Canada, August.

49. Gilbert, S. M. and Cvsa, V. (2000). "Strategic Commitment to Price to Stimulate Downstream Innovation in a Supply Chain", Working Paper, The University of Texas at Austin, Management Department, CBA, Austin, TX, November.

50. Gort, A. (2000). "Building a Global Unified Supply Chain", Section 3, Briefing Document, The Extended Enterprise, Supply Chain Management and Enterprise Resource Planning (www.mexen.com).

51. Graves, S. C., Rinnoy Kan, A. H. G. and Zipkin, P. H. (eds.). (1993). *Logistics of Production and Inventory*, Volume 4, *Handbooks in Operations Research and Management Science*, S.C.Graves, A.H.G. Rinnoy Kan and P.H.Zipkin (eds.), North-Holland (Elsevier Science Publishers B.V.), Amsterdam, The Netherlands.

52. Groenevelt, H. (1993). "The Just-in-Time System", Chapter 12, *Logistics of Production and Inventory*, Volume 4, *Handbooks in Operations Research and Management Science*, S.C.Graves, A.H.G. Rinnoy Kan and P.H.Zipkin (eds.), North-Holland (Elsevier Science Publishers B.V.), Amsterdam, The Netherlands.

53. Gross, D. and Soriano, A. (1969). "The Effect of Reducing Lead Time on Inventory Levels – Simulation Analysis", *Management Science*, 16, B61-B76.

54. Gupta, D. and Gerchak, Y. (1995). "Joint Product Durability and Lot-Sizing Models", *European Journal of Operational Research*, 84, 371-384.

55. Ha, A. (1998). "Incentive-Compatible Pricing for a Service Facility with Joint Production and Congestion Externalities", *Management Science*, 44, 12 (Part 1 of 2), 1623-1636.

56. Hadley, G. and Whitin, T. (1963). *Analysis of Inventory Systems,* Prentice Hall, NJ.

57. Hall, R. and Hitch, C. (1939). "Price Theory and Business Behaviour", *Oxford Economic Papers*, 2, 12-45.

58. Hariga, M. (2000). "Set-up Cost Reduction in (Q, r) Policy with Lot Size, Set-up Time and Lead Time Interactions", *Journal of The Operational Research Society*, 51, 1340-1345.

59. Hariga, M. and Ben-Daya, M. (1999). "Some Stochastic Inventory Models with Deterministic Variable Lead Time", *European Journal of Operational Research,* 113, 42-51.

60. Hay, D. A. and Morris, D. J. (1991). *Industrial Economics and Organisation: Theory and Evidence*, Oxford University Press, UK.

61. Hayes, R.H., Wheelwright, S. C. and Clark, K. B. (1988). *Dynamic Manufacturing: Creating the Learning Organisation,* The Free Press, New York.

74. Lee, H. L. and Nahmias S. (1993). "Single-Product and Single-Location Models", Chapter 1, *Logistics of Production and Inventory*, Volume 4, *Handbooks in Operations Research and Management Science*, S.C.Graves, A.H.G. Rinnoy Kan and P.H.Zipkin (eds.), North-Holland (Elsevier Science Publishers B.V.), Amsterdam, The Netherlands.

75. Lee, H. L. and Rosenblatt, M. J. (1986). "The Effects of Varying Marketing Policies and Conditions on the Economic Ordering Quantity", *International Journal of Production Research*, 24, 593-598.

76. Leschke, J. P. and Weiss, E. N. (1997). "The Multi-Item Set-up Reduction Investment Allocation Problem with Continuous Investment Cost Functions", *Management Science*, 43, 6, 890-894.

77. Levi, M. (1999). "Supplier Base Management: Selection, Control and Idiosyncratic Investments", Working Paper, SH-DH 1300, The Wharton School, University of Pennsylvania, Philadelphia, USA.

78. Levy, M. (2000). "2000 E-Commerce Recap", 2, 12 (http://ecmgt.com/Dec2000).

79. Li, C-L., Erlebacher, S. J. and Kropp, D. H. (1997). "Investment in Set-up Cost, Lead Time and Demand Predictability Improvement in the EOQ Model", *Production and Operations Management*, 6, 4, Winter, 341-352.

80. Li, L. (1992). "The Role of Inventory in Delivery-Time Competition", *Management Science*, 38, 2, 182-197.

81. Liao, C-J and Shyu, C-H. (1991a). "Stochastic Inventory Model with Controllable Lead Time", *International Journal of Systems Science*, 22, 11, 2347-2354.

82. Liao, C-J and Shyu, C-H. (1991b). "An Analytical Determination of Lead Time with Normal Demand", *International Journal of Operations and Production Management*, 11, 9, 72-78.

83. Likierman, A. (1981). "Pricing Policy in the Texturing Industry", *Journal of Industrial Economics*, 30, 1, 25-38.

84. Lohr, L. and Park, T. (2000). "Integrated Mark-up Rules for Optimal Pricing Decisions in Expanding Markets for Organic Produce", Research Paper, Department of Agricultural and Applied Economics, University of Georgia, Athens, USA.

85. Magretta, J. (1998). "Fast, Global and Entrepreneurial: Supply Chain Management, Hong Kong Style - An Interview with Victor Fung", *Harvard Business Review*, 76, 5, 102-114.
86. Maloni, M. J., and Benton, W. C. (1997). "Supply Chain Partnerships: Opportunities for Operations Research", *European Journal of Operational Research*, 101, 419-429.
87. Maltz, A. and Maltz, E. (1998). "Customer Service in the Distribution Channel: Empirical Findings", *Journal of Business Logistics*, 19, 2, 103-120.
88. Marr, N. E. (1994). "Do Managers Really Know What Service their Customers Require?", *Journal of Physical Distribution and Logistics Management*, 24, 4, 24-31.
89. McGoldrick, P. J. (1993). "Grocery Pricing in the 1990s - War or Peace? ", *Irish Marketing Review*, 101-110.
90. Min, K. J. and Chen, C-K. (1995). "A Competitive Inventory Model with Options to Reduce Set-up and Inventory Holding Costs", *Computers and Operations Research*, 22, 5, 503-514.
91. Monczka, R. M. and Morgan, J. P. (2000). "Competitive Supply Strategies for the 21st Century", *Purchasing Online*, 13th January.
92. Monczka, R. M. and Trent, R. J. (1995). "Purchasing and Sourcing Strategy: Trends and Implications", *Focus Study Executive Summary*, Center for Advanced Purchasing Studies.
93. Naddor, E. (1978). "Sensitivity to Distributions in Inventory Systems", *Management Science*, 24, 1769-1772.
94. Nahmias, S. (1997). *Production and Operations Analysis*, 3rd Edition, Irwin, Homewood, IL.
95. Nanto, D. K., Cooper, W. and Bass G. L. (1995). "Japan - U.S. Automobile and Parts Trade Dispute", *Congressional Research Service: Report for Congress*, 95-725 E, US.
96. Nasri, F., Affisco, J. F and Paknejad, M. J. (1990). "Set-up Cost Reduction in an Inventory Model with Finite Range Stochastic Lead Time", *International Journal of Production Research*, 28, 1, 199-212.

97. National Research Council Report. (1983). "The US Machine Tool Industry and the Defense Industrial Base", *Report of Committee on the Machine Tool Industry*, Manufacturing Studies Board, Commission on Engineering and Technical Systems, National Research Council, National Academy Press, Washington D. C.

98. National Retail Federation. (2000). Report of Mr. Robert Benham, Owner of Balliet's on behalf of the National Retail Federation before the US House Judiciary Committee, 29th June (http://www.house.gov/judiciary/benh0629.htm).

99. Newman, R. G. (1989). "Single Sourcing: Short Term Savings versus Long Term Problems", *Journal of Purchasing and Materials Management*, Vol. 25, No. 2, 20-25.

100. Nye, T. J. (1997). *Simultaneous Improvements in Manufacturing Systems and Effects on Investment Decisions*, Ph.D Thesis, Department of Management Sciences, University of Waterloo, Canada.

101. Paknejad, M. J., Nasri, F. and Affisco, J. F. (1992). "Lead Time Variability Reduction in Stochastic Inventory Models", *European Journal of Operational Research*, 62, 311-322.

102. Palaka, K., Erlebacher, S. and Kropp, D. H. (1998). "Lead Time Setting, Capacity Utilisation, and Pricing Decisions under Lead Time Dependent Demand", *IIE Transactions*, 30, 151-163.

103. Pappas, J. L. and Brigham, E. F. (1979). *Managerial Economics*, 3rd Edition, The Dryden Press, USA.

104. Petruzzi, N. C. and Dada, M. (1999). "Pricing and the Newsvendor Problem: A Review with Extensions", *Operations Research*, 47, 2, 183-194.

105. Porter, M. E. (1998). *Competitive Strategy: Techniques for Analysing Industries and Competitors*, Free Press, New York, USA.

106. Porteus, E. L. (1985). "Investing in Reduced Set-ups in the EOQ Model", *Management Science*, 31, 8, 998-1010.

107. Porteus, E. L. (1986a). "Investing in New Parameter Values in the Discounted EOQ Model", *Naval Research Logistics Quarterly*, 33, 39-48.

108. Porteus, E. L. (1986b). "Optimal Lot Sizing, Process Quality Improvement and Set-up Cost Reduction", *Operations Research*, 34, 1, 137-144.

109. Porteus, E. L. (1990). "Stochastic Inventory Theory", Chapter 12, *Stochastic Models*, Volume 2, *Handbooks in Operations Research and Management Science*, D. P. Heyman and M. J. Sobel (eds.), North-Holland (Elsevier Science Publishers B.V.), Amsterdam, The Netherlands.

110. Purchasing Online. (1998). "How to Help Suppliers Add Value", November 19 (www.manufacturing.net).

111. Ramdas, K. and Spekman, R. E. (2000). "Chain or Shackles: Understanding what Drives Supply Chain Performance", *Interfaces*, 30, 4, 3-21.

112. Rao, A. R., Bergen, M. E. and Davis, S. (2000a). "How to Fight a Price War", *Harvard Business Review*, March-April, 107-116.

113. Rao, U. S., Swaminathan J. M. and Zhang, J. (2000b). "Integrated Demand and Production Management in a Periodic, Make-to-Order Setting with Uniform Guaranteed Lead Time and Outsourcing", Working Paper, GSIA, Carnegie Mellon University, Pittsburgh, September.

114. Ray, S., Gerchak, Y. and Jewkes E. M. (2000). "Lead Time Management in Continuous Review Inventory Systems", Working Paper, Department of Management Sciences, University of Waterloo, Canada, September.

115. Ross, S. M. (1983). *Stochastic Processes*, John Wiley and Sons, USA.

116. Sajjad, Z. and Sarker, R. (1991). "Joint Economic Ordering Policies of Multiple Wholesalers and a Single Manufacturer with Price-Dependent Demand Functions", *The Journal of the Operational Research Society*, 42, 5, 157-164.

117. San Francisco Fashion Industry Report. (2000). "Production", Part III of A Beginning Guide To Apparel Manufacturing at http://www.dnai.com/~sffi/.

118. Schaible, S. (1981). "A Survey of Fractional Programming", in *Generalized Concavity in Optimization and Economics*, S. Schaible and W. T. Ziemba (eds.), Papers presented at the Proceedings of the NATO Advanced Study Institute held at the University of British Columbia, Vancouver, Canada, August 4-15, 1980, Academic Press, NY.

119. Scherer, F. M. (1980). *Industrial Market Structure and Economic Performance*, Houghton Mifflin Company, USA.

120. Simchi-Levi, D., Kaminsky, P. and Simchi-Levi, E. (2000). *Designing and Managing the Supply Chain: Concepts, Strategies and Case Studies*, McGraw-Hill Higher Education, USA.

121. Sinha, I. (2000). "Cost Transparency: The Net's Real Threat to Prices and Brands", *Harvard Business Review*, March-April, 43-50.

122. Smith, M. D., Bailey, J. and Brynjolfsson, E. (2000). "Understanding Digital Markets: Review and Assessment", *Understanding the Digital Economy*, Erik Brynjolfsson and Brian Kahin (eds.), MIT Press, USA.

123. So, K. C. (2000). "Price and Time Competition for Service Delivery", *Manufacturing and Service Operations Management*, 2, 4, 392-409.

124. So, K. C. and Song, J-S. (1998). "Price, Delivery Time Guarantees and Capacity Selection", *European Journal of Operational Research*, 111, 28-49.

125. Song, J-S. (1994a). "The Effect of Lead Time Uncertainty in a Simple Stochastic Inventory Model", *Management Science*, 40, 5, 603-613.

126. Song, J-S. (1994b). "Understanding the Lead Time Effects in Stochastic Inventory Systems with Discounted Costs", *Operations Research Letters*, 15, 85-93.

127. Song, J-S. and Zipkin, P. (1996). "The Joint Effect of Lead Time Variance and Lot Size in a Parallel Processing Environment", *Management Science*, 42, 9, 1352-1363.

128. Stalk, G., Jr. and Hout, T. M. (1990). *Competing Against Time: How Time-Based Competition is Reshaping Global Markets*, The Free Press, New York.

129. Sterling, J. U. and Lambert, D. M. (1989). "Customer Service Research: Past, Present and Future", *International Journal of Physical Distribution and Materials Management*, 19, 2, 17

130. Stevenson, W. J. (1999). *Production and Operations Management*, 6th Edition, Irwin McGraw-Hill, USA.

131. Stidham, S. (1992). "Pricing and Capacity Decisions for a Service Facility: Stability and Multiple Local Optima", *Management Science*, 38, 8, 1121-1139.

132. Suri, R. (1998). *Quick Response Manufacturing: A Companywide Approach to Reducing Lead times*, Productivity Press, USA.

133. Tayur, S., Ganeshan, R. and Magazine, M. J. (1999). "Introduction", *Quantitative Models for Supply Chain Management*, S. Tayur, R. Ganeshan and M. J. Magazine (eds.), Kluwer Academic Publishers, Boston.

134. US Department of Defense Contract Pricing Reference Guides. (2000). "Identifying the Seller's Pricing Objectives and Approaches", Volume 1, Price Analysis, Contract Pricing Reference Guides, Cost, Pricing and Finance Directorate, Defense Procurement, Department of Defense, USA.

135. Van Beek, P. and van Putten, C. (1987). "OR Contributions to Flexibility Improvement in Production/Inventory Systems", *European Journal of Operational Research*, 31, 52-60.

136. Vinson, C. E. (1972). "The Cost of Ignoring Lead Time Unreliability in Inventory Theory", *Decision Sciences*, 3, 87-105.

137. Wang, Y. and Zhao, W. (2000). "Inventory Systems with Decreasing Purchasing Costs", Working Paper, Department of Operations, Weatherhead School of Management, Case Western Reserve University, Cleveland, OH.

138. Weng, Z. K. (1995). "Channel Coordination and Quantity Discounts", *Management Science*, 41, 9, 1509-1522.

139. Weng, Z. K. (1996). "Manufacturing Lead Times, System Utilisation Rates and Lead-time-related Demand", *European Journal of Operational Research*, 89, 259-268.

140. Weng, Z. K. (1999). "Strategies for Integrating Lead Time and Customer-order Decisions", *IIE Transactions*, 31, 161-171.

141. Whitin, T. (1955). "Inventory Control and Price Theory", *Management Science*, 2, 61-68.

142. Wilson, R. F. (2000). "P4: Pricing Strategy as Part of Your Internet Marketing Plan", *Web Marketing Today*, May 9 (www.wilsonweb.com).

143. Zhang, H. (1998). "A Note on the Convexity of Service-Level Measures of the (Q, r) System", *Management Science*, 44, 3, 431-432.

144. Zheng, Y-S. (1992). "On Properties of Stochastic Inventory Systems", *Management Science*, 38, 1, 87-103.

145. Zipkin, P. (1986a). "Stochastic Lead Times in Continuous Time Inventory Models", *Naval Research Logistics Quarterly*, 33, 763-774.

146. Zipkin, P. (1986b). "Inventory Service-Level Measures: Convexity and Approximations", *Management Science*, 32, 8, 975-981.

147. Zipkin, P. (1991). "Does Manufacturing need a JIT Revolution'", *Harvard Business Review*, Jan-Feb, 40-50.

148. Zipkin, P. (1992). "Optimizing the Supply Chain: A Model for Operations-Marketing Planning", Research Working Paper, RWP-90-9, Graduate School of Business, Columbia University, NY

149. Zipkin, P. (2000). *Foundations of Inventory Management*, McGraw-Hill Higher Education, USA.