UNIVERSITY of SOUTHAMPTON

FACULTY of BUSINESS and LAW

SOUTHAMPTON BUSINESS SCHOOL

Optimal Inventory Policies with Postponed Demand by Price Discounts

by

Muzaffer Alım

Thesis for the degree of Doctor of Philosophy in Management Science

October 2018

Declaration of Authorship

I, Muzaffer Alım, declare that this thesis titled, 'Optimal Inventory Policies with Postponed Demand by Price Discounts' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: ______________________________

Date: ______________________________

UNIVERSITY of SOUTHAMPTON

Abstract

FACULTY of BUSINESS and LAW
SOUTHAMPTON BUSINESS SCHOOL

Doctor of Philosophy in Management Science

Optimal Inventory Policies with Postponed Demand by Price Discounts

by Muzaffer Alım

This thesis introduces a demand postponement policy in order to improve the performance of inventory management under batch ordering, advance demand information, capacitated/uncapacitated and periodic/continuous review inventory systems. The main aim of this study is to find integrated demand postponement and inventory policies. The structure of the thesis consists of five main chapters which starts with an introduction in Chapter 1 which summarizes the main objectives of the study with a background information, followed by a Chapter 2 presenting an overview of the relevant literature and the methodology. Chapter 3 as the first research paper, an inventory problem with stochastic demand and batch ordering and lost sales based on a real case is introduced and a demand postponement policy applied on this system to convert some of the lost sales to advance demand. A Markov Decision Process model is proposed and it is solved through Linear Programming (LP). The dual of the primal model is used to reduce the computational effort and it is tested with several numerical data sets. The optimal inventory policy and discount policy for different batches are shown for managerial insights. In Chapter 4, the same problem without batch ordering is formulated by Markov Decision Process (MDP) solved by Backward induction algorithm. In addition, the demand pattern is changed to Advance Demand Information (ADI) which combines both stochastic and deterministic demand. The properties of optimal inventory and postponement policy parameters are analyzed and the numerical experiments are carried out under the uncapacitated and capacitated systems to show the impact of the postponement policy. The comparison of policy parameters with the literature shows that the demand postponement policy is highly effective for the efficient use of capacity. In Chapter 5, the extension of the problem to a continuous review inventory system with distribution strategies is studied by an Net Present Value (NPV) approach. The effectiveness of demand postponement under different financial settings are examined and an extensive numerical experiments are presented.

The thesis ends with a conclusion in Chapter 6 including the summary of the results, limitations of the study and further research directions.

Keywords. Operational research, Supply Chain management, Logistic, Inventory management, Advance demand information, net present value, Periodic/continuous review inventory models, demand postponement, price discount

UNIVERSITY of SOUTHAMPTON

Özet

FACULTY of BUSINESS and LAW
SOUTHAMPTON BUSINESS SCHOOL

Doctor of Philosophy in Management Science

Fiyat Indirimleri ile Ertelenen Taleple Optimal Stok Politikalari

by Muzaffer Alım

Bu çalışmada, fiyat indirimleri ile sağlanan talep erteleme politikasının çeşitli şartlardaki stok sistemlerinin performansını iyileştirmede nasıl katkı yapacağı incelenmiştir. Asıl amaç, stok sistemini en iyileyecek bütünleşmiş bir stok ve talep erteleme politikası bulmaktır. Çalışma beş ana kısımdan oluşmaktadır. Birinci bölümde, problem ile ilgili genel bir bilgi verilmekte ve çalışmanın amaçları belirtilmektedir. Ardından ikinci bölümde stok problemleri ile ilgili literatür taraması ve ilgili metodoloji genel hatları ile sunulmuştur. İlk makale olan üçüncü bölümde ise stokastik talep ve parti halinde sipariş verilebilen stok sistemlerinin talep erteleme ile iyileştirilmesi çalışması yapılmıştır. Problem, Markov Karar Değişkeni Süreci yöntemi ile formülüze edilmiş ve beklenen toplam kar optimize edilmiştir. Modelin çözümünde Lineer Model kullanılmış ve çeşitli parametreler altındaki en iyi stok ve indirim politikaları gösterilmiştir. Dördüncü bölümde, bir önceki problemdeki parti halindeki sipariş kısıtı kaldırılmış ve problem Markov Karar Süreci ile modellenmiştir. Talep, bünyesinde deterministik ve stokastik talep bulunduran erken talep bilgisi ile değiştirilmiştir. Formülasyon, dinamik programlama ile çözümlenmiş olup en uygun stok ve erteleme politikalarının özellikleri analiz edilmiştir. Model, farklı problem parametreleri ile kapasiteli ve kapasitesiz stok modelleri için test edilmiş ve erteleme politikasının etkisi gözlemlenmiştir. Literatür ile karşılaştırma yapıldığında fiyat indirimi ile talep erteleme politikasının, stok kapasitesinin verimli kullanımında oldukça etkili olduğu görülmüştür. Besinci bölümde, problem sürekli kontrol altında ve dağıtım stratejilerini de içeren stok sistemleri için incelenmiştir. Net bugünkü değer yaklaşımı kullanılarak talep erteleme politikasının farklı finansal parametreler altındaki performansı ölçülmüş ve geniş bir numerik sonuç sunulmuştur. Bu çalışmanın son bölümünde elde edilen sonuçlar özetlenmiş ve çalışmanın sınırları ve gelecek çalışmalar için öneriler sunulmuştur.

Keywords. Yoneylem arastirmasi, tedarik zinciri yonetimi, lojistik, stok yonetimi, erken talep bilgisi, net bugunku deger yontemi, periyodik ve devamli stok sistemleri, talep erteleme, fiyat indirimi

Acknowledgements

I would like to express my sincere gratitude to a number of people as my PhD journey comes to an end. Without their guidance and support, I would not be able to complete my thesis.

Firstly, I would like to thank my main supervisor Dr. Patrick Beullens for his continuous guidance during my PhD and related research, for his patience, motivation and deep knowledge in the area. His guidance helped me in all the time that I am in need of support. I sincerely thank him for all his efforts on me and for seeing me as one of his colleagues instead of a student. I hope I would continue to have the opportunity of making research with him. I owe special thanks to Prof. Julia Bennels for her support as my second supervisor.

Second, I am grateful to the people in CORMSIS (Center for Operational Research, Management Sciences and Information Systems) at Southampton Business School for creating such a great academic environment with regular academic talks from all over the world. They were valuable experience for me and I have been benefited from those talks. I am also thankful for the students and the staff of the University of Southampton for providing a such work place.

Last but not the least, I would like to thank my wife for her continuous moral support during my stressful times and to my family for supporting me spiritually throughout writing this thesis and my life in general.

Contents

Declaration of Authorship i

Abstract ii

Ozet iv

Acknowledgements vi

Contents vii

List of Figures x

List of Tables xi

Abbreviations xii

1 Introduction 1
- 1.1 Context of the Research Problems . . . 1
- 1.2 Context of Methodology . . . 3
- 1.3 Research Aims and Objectives . . . 3
- 1.4 General Outline of the Thesis . . . 5

2 Literature Review 7
- 2.1 Overview of the Literature . . . 7
- 2.2 Lot Sizing Problems . . . 9
  - 2.2.1 Economic Order Quantity Models . . . 9
  - 2.2.2 Dynamic Lot Sizing Models . . . 10
- 2.3 Inventory Models by Net Present Value (NPV) Approach . . . 11
- 2.4 Methodology . . . 12
  - 2.4.1 Mathematical Programming . . . 12
  - 2.4.2 Markov Decision Process . . . 14
- 2.5 Conclusion . . . 15

3 Converting lost sales to advance demand with promised delivery date 17
- 3.1 Introduction . . . 18
- 3.2 Overview of Literature . . . 20
- 3.3 Model Development . . . 22
  - 3.3.1 Problem Description . . . 22
  - 3.3.2 Mathematical Formulation . . . 23

3.3.3 Linear Programming Model . . . 26
3.4 Computational Experiments . . . 27
3.4.1 The effects of Batch Sizes . . . 28
3.4.2 The effect of Demand Postponement . . . 30
3.5 Conclusion . . . 33

4 Selective state-dependent purchasing of buyers' willingness to wait 34
4.1 Introduction . . . 35
4.2 Literature Review . . . 37
4.3 Problem Description . . . 40
4.3.1 Demand Structure . . . 40
4.3.2 Demand Postponement . . . 43
4.4 Model Formulation . . . 44
4.4.1 State and Decision Variables . . . 44
4.4.2 State Transitions . . . 46
4.4.3 Cost Function . . . 47
4.4.4 Recursive Equation . . . 47
4.5 Structure of the Optimal Policy . . . 48
4.6 Analysis & Discussion . . . 50
4.6.1 Action Elimination . . . 50
4.6.2 Solution Algorithm . . . 51
4.7 Computational Results . . . 51
4.7.1 Uncapacitated Inventory . . . 53
4.7.1.1 The value of ADI . . . 53
4.7.1.2 The effect of Postponement on Inventory Policy . . . 54
4.7.2 Capacitated Inventory . . . 56
4.7.2.1 Sensitivity Analysis on Discount Amount . . . 58
4.8 Conclusions . . . 60

5 Joint inventory and distribution strategy for online sales with a flexible delivery option 61
5.1 Introduction . . . 62
5.2 Related research . . . 64
5.3 Description of the System . . . 65
5.3.1 General characteristics . . . 65
5.3.2 Inventory system . . . 67
5.3.3 Delivery system . . . 68
5.3.4 Customer's acceptance of flexible delivery . . . 71
5.4 Model Development . . . 71
5.4.1 Inventory level and order quantity . . . 72
5.4.2 Annuity streams of cash-flows . . . 73
5.5 Properties and algorithm . . . 76
5.5.1 Special Case: Inventory profit only . . . 76
5.5.2 Determining the distribution strategy . . . 77
5.5.3 General Case: Solution Algorithm . . . 78
5.6 Numerical Experiments . . . 80
5.6.1 Special case: Inventory profits only . . . 81

5.6.2 General case: Inventory-Distribution system . . . 83
5.6.3 Impact flexible delivery on distance travelled . . . 86
5.7 Conclusions . . . 87

6 Conclusions **90**
6.1 Overview . . . 90
6.2 Summary of the Main Contributions . . . 90
6.2.1 Chapter 3. NPV Analysis of a Periodic Review Inventory Model with Price Discount . . . 91
6.2.2 Chapter 4. Inventory Decisions under Advance Demand Information Driven by Price Discount . . . 91
6.2.3 Chapter 5. Analysis of Replenishment and Discount Policies with Distribution by NPV Analysis . . . 91
6.3 Limitations of the thesis and directions for future research . . . 92

A Supplement to Chapter 4 **93**
A.1 Proof of Lemma 4.1 . . . 93
A.2 Proof of Theorem 4.3 . . . 94

B Supplement to Chapter 5 **95**
B.1 Proof of Lemma 5.1 . . . 95
B.2 Proof of Prop 1 . . . 95
B.3 Proof of Lemma 5.2 . . . 96
B.4 Proof of Lemma 5.3 . . . 97

Bibliography **98**

List of Figures

2.1 Classification of Lot Sizing Problem 9

3.1 Sequence of Events 23
3.2 Optimal Order Quantity by Batch Size 28
3.3 Reorder Level by Batch Size 29
3.4 Deviation on Profit by Batch Sizes and Setup Cost 30
3.5 Inventory and Postponement Policy 32

4.1 Demand Pattern 43
4.2 Discount Decision 44
4.3 Order & Discount Policy 53
4.4 Sensitivity of Discount $d_{t1}, d_{t2} = d_{t1} + 0.5$, $K = 100, h = 1, p = 9, P = 0.5, C = 15$ 59
4.5 Analysis of d_{t2} when $d_{t1} = 4, \lambda = (3, 1, 2)$ 59
4.6 Analysis of d_{t1} when $d_{t2} = 4, \lambda = (3, 1, 2)$ 59

5.1 Process flow of order placement and delivery. 63
5.2 Inventory Flow with Constant β 68
5.3 $n(k)$ values within an inventory cycle 69
5.4 Cash-flows of revenue streams and inventory system costs 73
5.5 U=3 weeks 82
5.6 U=6 weeks 82
5.7 Impact of flexible delivery on total AS profits ($p = 1.3w$, $g = p - r$, $S = 80$, $c_u = 0.05w$, $\alpha = 0.2$, $U = 3$ weeks) 84
5.8 Effects of Discount by $\beta(r)$ ($p = 1.3w$, $w = 10$, $S = 80$, $g = p - r$, $U = 3$ weeks) 85

List of Tables

3.1 Notation . . . 24
3.2 Optimal Policy, $K = 10, h = 1, \lambda = 8, O_{t,t} \in \{0, ..., 6\}$ . . . 31
3.3 Sensitivity analysis of postponement, $d = 1, u = 20$ . . . 32

4.1 Notation . . . 41
4.2 $K = 100, h = 1, p = 9, D_{t-1,t+1} \in 0, ...10, T = 12$ . . . 54
4.3 $K = 100, h = 1, p = 9, d = 1, P = 0, \lambda = 4, 1, 1, D_{t-1,t+1} \in 0, ...10, T = 12$ 55
4.4 $K = 100, h = 1, p = 9, d = 1, P = 1, \lambda = 4, 1, 1, D_{t-1,t+1} \in 0, ...10, T = 12$ 56
4.5 $K = 100, h = 1, p = 9, d = (1, 1.5), P = 0.5, \lambda = 4, 1, 1, D_{t-1,t+1} \in \{0, ...10\}, Q_t \in \{0, ..., 2\}, C = 20$ . . . 57
4.6 Impact of ADI and Postponement, $T = 12, K = 100, h = 1, p = 9, d_{ti} = (1, 1.5)$ . . . 58

5.1 Impact of flexible delivery ($p = 1.3w$, $w = 10$,$r = 0.01p$, $g = 0$) . . . 81
5.2 Impact of deposit value ($S = 80$, $\alpha = 0.2$, $p = 1.3w$, $w = 10$,$r = 0.01p$) . . 82
5.3 Impact of flexible delivery on distribution AS costs ($S = 80$,$\alpha = 0.2$, $w = 10$, $r = 0.01p$, $\beta = 0.5$, $g = p - r$) . . . 84
5.4 Average travelled distances ($c_u = 0.05w$,$p = 1.3w = 13$, $g = p - r$, $r = 0.01p$, $A = 50$ miles^2, $\gamma = 0.30$) . . . 86

Abbreviations

SCM	Supply Chain Management
EOQ	Economic Order Quantity
EPQ	Economic Production Quantity
DLSP	Dynamic Lot Sizing Problem
MP	Mathematical Programming
LP	Linear Programming
MILP	Mixed Integer Linear Programming
IP	Integer Programming
MDP	Markov Decision Process
DP	Dynamic Programming
ADI	Advance Demand Information
NPV	Net Present Value
DLT	Demand Lead Time
SLT	Supplier Lead Time
TC	Total Cost
ASTP	Annuity Stream of Total Profit

Chapter 1

Introduction

This chapter consists of four sections. Section 1.1 provides a overview of the research problem and Section 1.2 introduces the relevant methodology used in research chapters. In Section 1.3, the main research aims and objectives are discussed and in Section 1.4 presents a detailed outline of the thesis.

1.1 Context of the Research Problems

Supply Chain Management (SCM) is the controlling and planning of all supply chain activities, starting from supplying raw material and including production, stocking, and distribution of the finished products to right locations and ending with the delivery to end customers. The main objective of SCM is to establish the integration between the partners such as suppliers, manufacturers, warehouses and retailers in order to minimise the total cost while answering customers' needs. For key literature on SCM and future research directions, see Power (2005) and Burgess and Koroglu (2006).

Supply chain systems which are designed for a stable environment might be vulnerable to uncertainties on both the supply and the demand side. Inaccurate demand forecasting, variable/stochastic lead times, price changes in the market, disruptions due to the natural and human disasters and uncompleted shipments create uncertainties in the supply chain (Tang 2006b). Companies deal with these by having inventory as a preparation to avoid delays on their services. While holding high amounts of inventories might be a solution to deal with uncertainties, keeping higher levels of inventories could be too costly. Inventory management then plays a key role through the process of SCM due to its direct impact on both cost and customer service. The competitive markets require a strong need for inventory management to determine the right amount of inventory while considering the balance between the service level and cost of having it (Russell and Taylor 2011).

Inventory management is defined as the controlling of product kept in stock over time. The main purpose is to answer the questions of "When to order?" and "How much to order?" under various problem environments. The first model for inventory management is introduced by Harris (1913) and known as the Economic Order Quantity (EOQ) model. In spite of its deterministic and simple problem assumptions, the EOQ model has been a valuable tool and is still considered to be a fundamental method in the inventory literature (Cárdenas-Barrón et al. 2014). The constant demand assumption has been changed to a demand varying over time in the Dynamic Lot Sizing (DLS) problem and a Dynamic Programming (DP) approach is proposed by Wagner and Whitin (1958) to solve it. The literature has been growing in order to deal with more complicated problems such as stochastic demand, non zero lead times, backorder/lost sales, quantity discounts, restricted orders, capacitated inventory systems, perishable products etc. For an extensive literature review, we refer to Brahimi et al. (2006).

Inventory models are often categorized by the stochastic or deterministic nature of the model parameters. For the stochastic cases, the uncertainties on both supply and the demand make flexible inventory management crucial. The traditional inventory approach focuses on finding the optimal ordering policy for meeting a given demand pattern. On the other hand, the modern approach works to change the demand pattern if applicable and to simultaneously set an inventory policy which leads to less inventory costs and higher profits. Well-known examples are the class of quantity discount models which focus on changing the demand by offering discounts. Offering discounts may reduce the uncertainty on demand or increase the demand rate and is by now a widely applied method (Shin and Benton 2007, Wee 1999, Weng 1993).

Although the forecasting and planning for stochastic demand is getting more sophisticated, there still exists a high risk of being stock out (Yang and Burns 2003). Recent improvements of information technology and increased usage of online channels have equipped inventory managers with more accurate information from customers (Özer

2011). This includes the demand information which is placed by customers in advance of their due date, and which is called Advance Demand Information (ADI). Real-life examples of this demand type can be observed in many areas such as room booking of hotels, flight reservations for airline companies, pre-order of new technological products or computer games etc.

1.2 Context of Methodology

The first methodological approach is to study the inventory problem with batch ordering and demand postponement formulated by Markov Decision Process (MDP) with an NPV objective function. The Linear Programming (LP) model and its dual have been used to carry out the numerical experiments for finite time horizon. The second approach is to study the extension of the first problem to a case which has advance demand information. The problem is formulated by MDP and we solve it by Dynamic Programming DP for capacitated and uncapacitated inventory systems under finite time horizon. DP is one of the most common methods in literature on dynamic lot sizing problems. To reduce the computational effort, additional valid inequalities are presented. The third methodology used is calculating the annuity stream of an inventory cash flow, i.e. the Net Present Value (NPV) approach. Although classical models are commonly used in the literature, the postponement of payments and order decisions force us to consider the time value of money in order to obtain accurate inventory solutions. In addition, we include the distribution of the goods by using average travelled distance model (Daganzo and Newell 1985). The problem then is solved by exhaustive search algorithm for numerical experiments.

1.3 Research Aims and Objectives

The main purpose of this thesis is to analyse several variants of inventory problems and to improve their efficiency by using price discounts as a mechanism for manipulating the demand. The main aims of each chapter are summarized as follows.

The first paper focuses on periodic review inventory systems with stochastic demand, batch ordering and improving its performance by postponing demand using price discounts with NPV objective function.

The main research aim of the second paper is to develop an optimal inventory and discount policy for a periodic review inventory system with stochastic Advance Demand Information (ADI) by modeling the problem as a Markov Decision Process for capacitated and uncapacitated cases.

In the third paper, we address a continuous review inventory system and investigate the importance of payment structures on inventory policies by using a Net Present Value (NPV) approach and a distribution strategy based on the average travelled distance is integrated into the inventory model.

In more detail, the research objectives are summarized as follows:

The objectives of the first paper will be:

- to identify the inventory models for batch ordering system in literature;
- to formulate the periodic review inventory system with batch ordering and discount decision as an MDP and solve it by LP model;
- to analyse the relation between batch ordering and demand postponement;
- to perform computational experiments under several parameters settings.

The objectives of the second paper are:

- to review the relevant literature on advance demand information and include the discount decision to the inventory systems with ADI;
- to investigate the conditions under which the price discount mechanism performs better and brings benefit to the system;
- to develop an MDP formulation and its solution by Dynamic Programming;
- to introduce additional valid inequalities to improve the computational performance of the solution techniques;

- to perform extensive computational experiments with several parameters under capacitated and uncapacitated cases.

The objectives of the third paper are:

- to research the area of continuous review inventory systems with discounts in the literature;
- to formulate the problem in second paper to a continuous review case with constant demand rate;
- to test the impact of financial terms (postponed payment, deposit, interest rate) on inventory system;
- to see the impact of outsourcing distribution or self distribution by considering the average travelled distance Daganzo and Newell (1985) with the inventory model.

1.4 General Outline of the Thesis

The remainder of this transfer thesis is organized as follows. Chapters 2 presents an overview to the literature of inventory problems and review the various solution methods for inventory problems. These methods include Mathematical Programming, Dynamic Programming, heuristics and Markov Decision Process.

Chapter 3 introduces a batch ordering inventory problem with stochastic demand and lost sales. The demand postponement policy is applied to convert some lost sales into advance demand information and an MDP formulation is proposed. The model is solved through Linear Programming method and a numerical experiments are carried out to obtain some managerial insights.

Chapter 4 extends the problem in the first research paper to a backordering case without batch ordering. The inventory system already contains some advance demand and postponement policy is used to buy more advance demand when it is needed and profitable. The problem is solved by Dynamic Programming and the structure of the optimal policy is discussed. The model is numerically tested for uncapacitated and capacitated inventory systems and an extensive numerical results are presented.

Chapter 5 investigates the impact of such demand postponement policy on continuous review inventory models. The NPV value of objective function is considered to test the system with different financial settings. A distribution strategy is included into the problem which is to choose whether outsource the delivery or make it locally. An exhaustive search algorithm is used to find the integrated inventory and distribution plan under various system parameters.

Chapter 6 concludes the thesis with a conclusion including the discussion of the limitation of the study and further research directions.

Chapter 2

Literature Review

This chapter begins with an overview to the supply chain, its improvements and shifting demand across time. Next, the evolution of the lot sizing problems which are Economic Order Quantity and Dynamic Lot Sizing models is reviewed. Then, the implications of NPV into lot sizing problems are discussed. The various solutions methods including mathematical programming, dynamic programming, heuristics and Markov Decision Process are discussed and finally, there is a conclusion to compare the methodologies and connect it to the research chapters.

2.1 Overview of the Literature

A Supply Chain (SC) is the entire process to deliver the product from the supplier to the end user. Mainly this procedure can be examined under two classes: (1) Production Planning and Inventory Control and (2) the distribution and Logistic Process (Beamon 1998). Supply Chain Management (SCM) is characterized as the controlling, arranging and integration of these process. Earlier it was basic with a stream of raw material to manufacturer and then to the markets. However, with the shorter product lifecycles, uncertainties on increasing demand, competitive market environment, off-shoring and outsourcing strategies make the SCM more challenging (Tang and Nurmaya Musa 2011). Failure of a partner in a SC creates disruption for all partners both upstream and downstream (Yang and Yang 2010). Thus, there has been an increasing need on effective and flexible SCM to deal with these challenges.

There has been a growing attention on making SCM more flexible and robust. Tang (2006a) propose nine strategic ways including postponement and revenue management for this purpose. They classified the postponement into the classes. First one is the manufacturing postponement which is to postpone the customization, final assembly or packing of the product until the order information received from the customers Yang and Burns (2003). The second on logistic postponement on the other hand is to postpone the changing on inventory locations to the latest point possible (Pagh and Cooper 1998). These studies use the postponement to manage the supply side. Yang and Yang (2010) review the postponement strategies on supply and discuss the complexity of application of such strategies.

The flexibility on supply side is very limited since there is a fixed capacity restriction on supply in most situations. In such cases, the firms focus on demand management. Tang (2006b) summarizes the strategies for demand management as shifting demand across time, markets and products.

Shifting demand across time is getting more popular with the effective usage in industries such as airlines, hotels, utilities as increase the usage of online channels Xu et al. (2017). Customers are offered some price incentives to shift their demand to off-peak periods. Similarly pre-order incentives are applied to gather the demand information from the customers prior to the release of the product. Some customers prefer advance bookings due to uncertainty on the future availability of the product Seref et al. (2016). By getting this information, the firm can overcome the lack of forecasting (Tang et al. 2004) or could increase the effective usage of capacity (Zhuang et al. 2017). These should not be compared with the price incentives to increase the total demand. The main focus on shifting demand is to change the timing of the demand regardless of aiming to change its size. Another study on this area is Iyer et al. (2003) which analyses the demand postponement strategy when the demand exceeds the short term supply capacity. The main idea they define behind demand postponement is to preempt stock-outs or shortages to reduce the expected costs. The shifting demand over times by price incentives are considered under the context of Revenue Management(RM). Quante et al. (2009), Talluri and Hillier (2005) offer a detailed review of revenue management strategies to manage the demand.

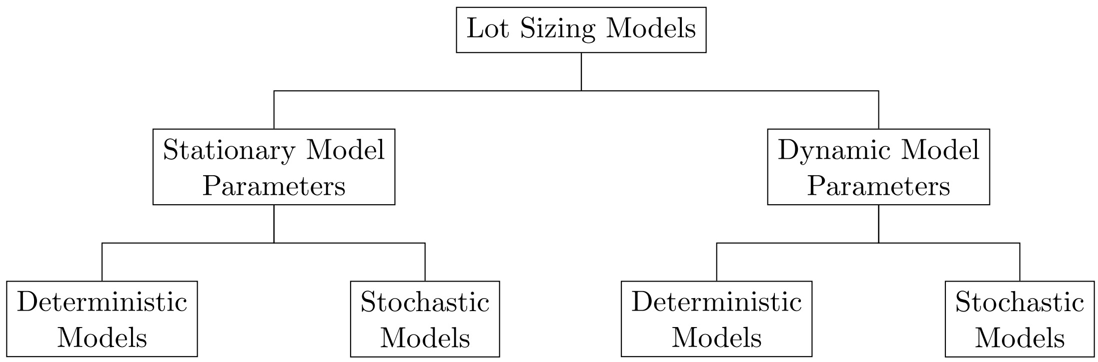

FIGURE 2.1: Classification of Lot Sizing Problem

2.2 Lot Sizing Problems

The lot sizing problem has received much attention from both academia and practice since the first publication in 1913. The reason of this interest is the direct impact of inventory management on both customer service and cost. The lot sizing problems have been classified into different categories. One of the main classification by Glock et al. (2014) is based on the technical structure of the problem as in Fig. 2.1. The more detailed version of theirs has been presented by Aggarwal (1974).

Fig. 2.1 considers the classification of lot sizing problems by the nature of the problem parameters which could be changing over time (stationary-dynamic) and uncertain (stochastic-deterministic). Specifically, based on the nature of the demand the inventory models are named as Economic Order Quantity (EOQ) for constant demand rate or Dynamic Lot Sizing (DLS) for time varying demand. These two are the main models of lot sizing and they have been extended to multiple cases. In the next sections, we present an overview to these models with their extensions in literature.

2.2.1 Economic Order Quantity Models

The first lot sizing model is introduced by Harris in 1913 and called Economic Order Quantity(EOQ) model. The model is developed for a continuous review inventory system on an infinite time horizon with constant demand rate. It provides the optimal order size which is satisfying the trade-off between the ordering and holding costs under stationary parameters and deterministic demand. This model has been extended for different problem environments. Hax and Candea (1984) analyses several extensions of EOQ including backorders and lost sales models. Bakker et al. (2012) review the EOQ models with deterioration in which the items on the stock has a limited lifetime on stock. The EOQ model is called Economic Production Quantity (EPQ) when it considers producing item with a finite production rate (Holmbom and Segerstedt 2014). The review of EOQ and EPQ models with deterministic parameters and partial backordering has been studied by Pentico and Drake (2011). The EOQ model with stochastic parameters has also been interested in literature. Yano and Lee (1995) provides a review of stochastic cases where yield and demand are random for continuous and discrete time models. We refer readers to an early review of EOQ literature (Erlenkotter 1990), as recent study Holmbom and Segerstedt (2014) and Glock et al. (2014) which provide extensive review on Economic lot sizing problems.

Up to this point, all reviewed studies are introduced based on the nature of the problem parameters. There are studies focusing on studying on the content of the problem. Under various parameters settings, the two stage and multi stage cases of the lot sizing problem have been reviewed by Goyal and Gunasekaran (1990), Cárdenas-Barrón (2007) and Bakker et al. (2012). The extensions of the problem with consideration of scheduling and incentives have also been studied in the literature (Glock et al. 2014). Lot sizing and scheduling decisions are closely related to each other so they are often studied in combination. This has been named as Economic Lot Scheduling Problem (ELSP) in which the sequences of the products in production need to be determined in addition to the inventory decisions (Elmaghraby 1978). An early review of ELSP is presented by Graves (1981) and Winands et al. (2011) for the stochastic problem. Incentives are another content issue in literature. Pricing strategies as incentives are the most common way for demand management in inventory models and Elmaghraby and Keskinocak (2003) presents an extensive review of the relation between dynamic pricing and demand. Quantity discounts as an application of incentives are offered if the order amount is higher (Pentico and Drake 2011, Shin and Benton 2007). There are also studies to offer delay in payment instead of quantity discounts (Chung et al. 2005, Pal and Chandra 2014).

2.2.2 Dynamic Lot Sizing Models

The stationary lot sizing problem is updated to a case where the demand is varying over time and it is called Dynamic Lot Sizing Problem (DLSP). It has been received considerable attention in inventory literature especially for periodic review inventory systems. The first paper is Wagner and Whitin (1958) which studies DLSP for a single item and single supplier case. They propose a dynamic programming approach to solve the DLSP with deterministic time varying demand and fixed ordering and linear holding costs. Another early study on DLSP is Zangwill (1966) which includes the backlogging into the DLSP. Aksen et al. (2003) analyse the DLSP with lost sales and propose an efficient algorithm for solution. Lee et al. (2001) introduce time windows for the demand in DLSP and it is required to satisfy the demand within that time window. For the detailed review of DLSP, we refer readers to the studies of Holmbom and Segerstedt (2014) and Brahimi et al. (2017) and Karimi et al. (2003) for capacitated DLSP.

The DLSP has also been considered for two and multi stages and the reader may be referred to Gupta and Keung (1990), Aggarwal (1974) and Brahimi et al. (2006) for a detailed review. Production scheduling has a significant impact on stock especially when multiple items are on production line. Thus, the integrated problem of DLSP with scheduling has been taken in analysis by Beck et al. (2015), Robinson et al. (2009) and Staggemeier and Clark (2001). Incentives are also used often in DLSP problems. Chung (1987) and Federgruen and Lee (1990) consider the quantity discounts in an DLSP. Furthermore, Mazdeh et al. (2015) add multiple suppliers into the problem where each suppliers have different quantity discounts. Thus, the model needs to determine the supplier and the optimal lot size. The same problem with backlogging has been studied by Ghaniabadi and Mazinani (2017).

2.3 Inventory Models by Net Present Value (NPV) Approach

In classic EOQ models, the main objective is to minimise the Average Cost (AC). The payment structure does not have an impact on the objective. Some studies focus on this issue and use Net Present Value (NPV) approach in the lot sizing problem to consider the time value of money. Hadley (1964) is the first study to use NPV in a lot sizing problem and they compare the results with average cost model. Sun and Queyranne (2002) investigate the NPV approach on multiproduct and multistage production and inventory model and show that AC is a good approximation of NPV for a deterministic demand. Chao (1992) investigate the NPV for an inventory system with stochastic demand. Grubbström (2007) presents the transform methodologies for using NPV in a stochastic environment.

All the studies show that NPV model is more accurate than AC model. However, the usage of NPV in literature is very limited. This is due to the complexity of NPV model and the close results of AC to NPV. Van der Laan (2003) point out that AC is a nice approximation for fast moving products on stock, low interest rates and not changing payment structure with the inventory policy. However, when there is a payoff of long term investments, the time value of money has to be considered (Marchi et al. 2016). Hsieh et al. (2008) investigate the lot sizing problem with two warehouses and deterioration with an NPV objective value and they claim that the reorder interval of AC objective must be longer than NPV. There are some other interesting studies on NPV and we refer readers to Grubbström (1998), Giri and Dohi (2004), Beullens and Janssens (2014) and Ghiami and Beullens (2016).

2.4 Methodology

In this section, we introduce the most commonly used methodology in inventory literature by their applications in literature. Both exact methods and heuristics are applied to inventory problem in literature. Among these methods, we present a background information on Mathematical Programming (MP) including Linear Programming, Integer Programming and Mixed Integer Programming (MIP) and Markov Decision Process (MDP). We refer readers to Karimi et al. (2003) and Buschkühl et al. (2008) for a review of solution approaches for capacitated lot sizing problem and Tempelmeier (2013), Buschkühl et al. (2008) and Aloulou et al. (2013) for stochastic DLSP.

2.4.1 Mathematical Programming

The Mathematical Programming (MP) includes the Linear Programming (LP), Integer Programming (IP) and Mixed Integer Programming (MIP). Linear programming is an Operations Research (OR) technique to optimize (either maximise or minimise) the models with linear objectives and constraints Taha (2007). The graphical solution method of LP is enumerating all the basic solutions (corner points). But, the simplex method only focuses on a few points on these solutions which makes this method more effective. Integer programming is a kind of LP with a condition that all the variables are restricted to be integer. MIP is an integer programming but some of the variables can take on real number values. For the theory behind the LP and its applications are presented by Lewis (2008).

There have been extensive real life applications of mathematical programming methods including scheduling, facility location, transportation, production and inventory etc. The nature of lot sizing inventory problem mostly requires to have both continuous and integer decision variables which lead to MIP. Implications of MIP are very common in literature. Pochet and Wolsey (2011) contain papers on alternative applications of MIP on production and inventory systems with a background knowledge on MIP.

Tempelmeier (2013) model the single item uncapacitated lot sizing problem as an MIP for fixed and variable replenishment periods. He also updates the model for a stochastic case. For the capacitated case, Drexl and Kimms (1997) present the MIP models for both discrete and continuous lot sizing and scheduling problem. Gao et al. (2008) compare the MIP model for coordinated DLSP with the LP relaxation. They use the result of LP relaxation as a lower bound for MIP and they obtain (0.022%) as a worst optimality gap. Zhao and Guan (2014) take the dual of LP model incorporating Bellman equations for stochastic uncapacitated lot sizing problem. Lulli and Sen (2004) use branch and price algorithm for multistage stochastic problem and compare it with MIP. On the other hand, Guan et al. (2005) define some valid inequalities to the same problem and use branch and cut algorithm. Lee et al. (2013) construct a MIP model for LSP with multiple suppliers and quantity discounts.

In most cases, it has been shown that the MIP models are NP hard for lot sizing problems and therefore many researchers focus on heuristic methods (Drexl and Kimms 1997). Mazdeh et al. (2015) develop a heuristic method based on Fordyce - Webster Algorithm for the DLSP with supplier selection and quantity discounts. Lee et al. (2013) study the same problem with MIP model and compare it with the Genetic Algorithm. For the complicated problem or NP hard case, the Genetic Algorithm provide nearly optimal results in a short computational time. Gaafar (2006) develop a Genetic Algorithm for batch ordering inventory models. Baciarello et al. (2013) compare the several heuristic methods for the solution of uncapacitated single item LSP in terms of their performance. Meta-heuristics are also widely applied in DLSP and Jans and Degraeve (2007) review them and make a comparison among these heuristics. Beck et al. (2015) review the Wagner-Whitin algorithm and provide an extension for dynamic lot sizing heuristics including Least unit cost, Groff's rule, Leinz-Bossert-Habenicht.

2.4.2 Markov Decision Process

Markov Decision Process (MDP) provides a mathematical framework to formulate the decision sequence where the system can move stochastic or deterministic to another state (Puterman 1994). The main property of Markov is that the state on next stage only depends on the current state and given decision at the current stage. Except the previous state, Markov property requires states to be history independent. In fact, the theory behind the MDP lies upon the recursive Bellman Equations. Main components of MDP are state variables from a state space, decisions, reward/cost function and transition function and probability. MDP is mainly categorized into two class, as Finite Horizon and Infinite Horizon problems.

Puterman (1994) applies the MDP into different cases including stochastic inventory, shortest route and critical path problem and discrete time queuing systems. The implication of MDP into an inventory problem with batch ordering and lost sales has been presented by Woensel, van et al. (2013). Fianu and Davis (2018) study the equitably distribution of donated and uncertain supplies and formulate it as a discrete time Markov Decision Process. Ahiska et al. (2013) use the MDP approach for a stochastic inventory problem with two suppliers including one that is unreliable with better price and other one is perfectly reliable. Qiu and Loulou (1995) consider the MDP for an inventory system with random demand and multiple products. White (1985) summarizes the problems defined by MDP in literature and identifies the MDP models which are implemented into real cases.

The main solution algorithms for MDP are value iteration, policy iteration and Linear programming. Value iteration is the most common method used in dynamic programming and for the finite time horizon it is identical to the backward induction algorithm (Powell 2007). The application of dynamic programming on inventory problems starts with the model proposed by Wagner and Whitin (1958). After this seminal work, there has been a lot of studies considering the dynamic programming for the inventory problems. Yano and Lee (1995) use the dynamic programming approach to solve the stochastic DLSP. Similarly, Gallego and Özer (2001) consider the backward induction algorithm for the solution of periodic review inventory problem with stochastic advance demand. The policy iteration, on the other hand starts with an initial policy and then calculates the value of that policy. Then a new policy is chosen and the value is getting improved by the changes on policy. Ye (2011) proves that policy iteration is strongly polynomial for the fixed discounted MDPs. Another solution method is to use the Linear Programming model. Hernández-Lerma O. (1999) state that LP approach allows to identify the optimal policy on a subset of state space. The superiority of the LP model on value iteration only arises for the relatively small state and action spaces. While the LP model with 50000 constraints is assumed to be a large problem, dynamic programs with the same size of constraints are considered relatively small problems (Powell 2007). Similarly, Abbasi-Yadkori et al. (2014) state that LP is not practical for the large size of variables and constraints and study an approximation method for MDP with large scale state spaces.

2.5 Conclusion

In this chapter, we review the lot sizing problems and the applications of various operational research methods into the lost sizing problems. The MDP formulation of lot sizing problems can be solved by dynamic programming, policy iteration and linear programming. Although the most common method is dynamic programming, the literature shows that LP models can perform better for small state sizes. Also there are recent improvements on LP solvers which can solve large problem instances in an acceptable time. These will lead us to consider the LP model as a solution method for lost sales problem in Chapter 3.

For the finite time horizon lot sizing problem, the research studies in literature prefer dynamic programming. Considering backordering increases the state size since the inventory level can be negative. Thus, in Chapter 4, we use dynamic programming and MDP formulation helps us to state the structure of the optimal policies. Additional state space reduction is introduced and the computational time is reduced. Finally, in the review of NPV approach, we see that the AC model cannot be a good approximation of NPV when the payment structure changes by the inventory policy. The postponement decision cause customers to pay a deposit when they place the order and pay the remaining on the delivery. Therefore, the payment structure depends on discount and inventory policy. These motive us to consider the NPV of the objective function in Chapter 5. We test the system under different financial settings by the help of NPV approach.

Chapter 3

Converting lost sales to advance demand with promised delivery date

Abstract

This paper looks into the periodic review inventory problem with batch ordering and postponed demand by price discounts. The demand pattern is stochastic and unsatisfied demand on time is lost. Beside the classical inventory decisions of order time and amount, there are also postponement decisions of how many demand to be discounts and for how long. To entice customers for postponement, we offer price discount for only some part of the customers at some times. Customers are accepting the postponement with a rate. We formulate the problem with Markov Decision Process (MDP) and solve it through Linear Programming (LP) model. Dual of the primal model is taken into account as to reduce the computational time. We present extensive numerical experiments to show how the optimal policy is affected by changing batch size and postponement policy.

3.1 Introduction

Matching supply and demand in an efficient way is the main purpose of inventory management. Traditionally, solution strategies were focused on determining an optimal supply policy to meet given demand characteristics. More and more, however, inventory theory also considers options to actively influence demand patterns, and seeking for jointly optimized supply and demand management policies.

On the supply side, replenishment policies are often based on particular conditions set out by suppliers. In the classical EOQ model, it is assumed that the order size can take any value. In many production/inventory systems, however, there can be various restrictions. A common restriction is the specification of a minimum order quantity (Zhou et al. 2007, Kesen et al. 2010), either imposed as a hard constraint or a soft constraint in the form of a penalty cost for smaller orders, or a quantity discount if the order exceeds this threshold. Next to this being a strategy for the supplier to stimulate demand, the underlying reasons typically relate to the supplier's cost structure in its production or transportation system, making it not sufficiently economical to supply small quantities. Another form of restriction is known as *batch ordering* (Veinott 1965, Broekmeulen and van Donselaar 2009). Here, the order size has to be an integer number of a base batch quantity. Application of batch ordering is e.g. common in the retail industry (groceries and other products). It is often the consequence of the way individual products are being packaged in larger boxes at production sites and/or fit onto pallets for distribution to wholesalers or retailers, where again costs savings in production, distribution, and inventory handling may be an underlying reason why dealing in larger units is adopted in the supply chain. This practise can furthermore have other benefits, such as minimising the risk of damage to products. Selling in larger quantities also induces the buyers to keep on average more stock, which results in financial benefits for the supplier of receiving profits earlier, referred to as the supplier's reward in Beullens and Janssens (2014).

We study a inventory system with stochastic unit-sized demand but with batch ordering from a supplier. The motivation for this study originates from a project undertaken for an organization supplying and maintaining information technology equipment to the University of Southampton. The company provides laptops, monitors and other computer equipment and peripherals to the university's staff and postgraduate students.

The focus of study was on the supply of laptops and desktops. Staff and students, as customers and users of these computers, can request these via an online ordering system. This demand is difficult to predict and is to be regarded as stochastic. The university expects that these requests can be fulfilled quickly. As some time is needed to install hardware and software components individualized to the user, the company effectively needs to be able to fulfill these orders from stock on hand. To help achieve this, as well as achieve sharp prices from the producers, and further reduce on installation and maintenance costs, the computer types available to the users is restricted to a selection of standardized types. In addition, the orders from the producers of each of these types have to occur in batches (pallets) with base batch quantities ranging around 10 to 20, depending on the computer type.

The company felt that this batch ordering was perhaps restrictive. They thus desired insight into the impact of different base batch quantity levels on financial (holding) costs and service levels, so that they could get an understanding of the commercial value of reducing base quantity levels in potential re-negotiations with suppliers.

Another issue the firm wanted to investigate relates to the demand side. It was quite important to the company to ensure customer satisfaction and part of the performance measures they needed to track was the amount of customers orders completed according to the customer desired delivery time. While customers were encouraged to submit their requests ahead of their needs, this almost never occurred, which means that products always needed to be kept on hand, and delivered within three to five days. One way in which customers can be induced to order in advance is to offer price incentives linked to requested delivery time. However, offering these schemes to all customers can also becomes expensive. A good fraction of demand, however, was related to the upgrade of computers still working. The company thought that it would be possible to convert some of this demand to later delivery times by offering a small financial incentive. This only needed to occur at times when stock was running short. Next to improving service performance, such a strategy will also affect the optimal ordering policy from suppliers. Examining the impact of such a demand postponement strategy has thus to be considered jointly with the first question about the impact of batch ordering.

In this paper, we develop a stylized model that has most of these characteristics and can be used to develop insights into these questions. We consider a batch ordering periodic review inventory problem with stochastic demand and demand postponement enticed by price discounts. The supplier lead-time is shorter than the length of a review period. Without a postponement option, demand arriving within a period has to be met in that period or else is lost. While in the above described example most of this demand would result in backorders instead, the assumption of lost sales is more pertinent for computer stores facing competition. The products are supplied from an external supplier but the order size has to be in integer numbers of a base quantity.

The demand postponement with a price discount is used to persuade some of the customers to accept later delivery within a given maximum delay time. By this, we aim to transform some of the lost sales into advance demand information, which avoids the firm from losing profit. We would also expect this postponement strategy to reduce the negative impact of the batch ordering constraint. The postponement option is only offered to customers in periods when it is needed, and not all customers giving this option would accept it. We consider the maximisation of the total profit and model the problem as a Markov Decision Process (MDP). Our main objective is to investigate the system performance with the postponement policy and increase insight into the conditions under which postponement is beneficial under different levels of base batch quantities.

The remaining parts of this article are organized as follows. In Section 3.2, we provide a review of relevant literature on batch ordering, demand postponement and lost sales. In Section 3.3, we describe the problem environment with relevant notation and assumptions and develop the model using a MDP. We conduct the computational results in Section 3.4. Finally, conclusions and areas for further research are given in Section 4.8.

3.2 Overview of Literature

Shifting demand across time is an important aspect in the studied problem. There are two approaches in the literature. The first strategy is to obtain the demand information from the customers at earlier times. Li and Zhang (2013) examine the benefits of offering price discounts to customers who pre-order prior to release date of a (new) product. Huang and Van Mieghem (2013) focus on increasing insight from data mining on the behaviour of customers in an online sales setting. The second strategy is to convince customers to accept later delivery, often referred to as *demand postponement*, which might be helpful in particular when the system is not able to satisfy earlier delivery. Kremer and Van Wassenhove (2014) and Yang and Burns (2003) discuss the potential benefits of postponement as a means of sharing supply chain risk with the customers. Although there have been studies on strategies and benefits of the postponement, and while it is often encountered in other areas (e.g. booking systems for airlines), few applications of postponement are available in the inventory literature. Yang and Yang (2010) state that this may be in part a result from the increased complexity of the analysis and management of inventory systems with demand postponement options.

Another aspect in this study is batch ordering. This restriction refers to the case where the products are to be supplied in multiples of fixed supply lots (e.g., carton, pallet, container, full truckload). Veinott (1965) is one of the earliest works on batch ordering for dynamic lot sizing problems with a constant lead time. Zheng and Chen (1992) study the quantized ordering by considering the (r, nQ) inventory policies, where Q and r are decision variables. The comparison of this policy with the (s, S) policy shows very little change under the assumption that both s and S are a multiple of Q. Although a simple form of optimal policy does not exist, (r, nQ) policies are commonly applied in the batch ordering models, where Q is the base batch quantity. For the multi-echelon case with exogenously determined batch size, Chen (2000) shows the optimality of the (r_i, n_iQ) policy, which consists of ordering n_i batches when the inventory position is at or below the reorder level r_i at stage i. Lagodimos et al. (2012) show that the batch ordering restriction, for situations with a fixed review interval T, can introduce a significant cost increase when applying a (r, nQ, T) policy. Fixed review periods in multi-echelon situations are also studied in Chao and Zhou (2009) and Shang and Zhou (2010). While dynamic programming is a frequently applied solution approach, Gaafar (2006) apply a genetic algorithm approach to solve the batch order dynamic lot sizing problem with backorders allowed.

A variation known as partial batch ordering occurs when the order is allowed to be in any size, but where a total set-up cost is charged as an integer multiple of a fixed batch size. For example, if the batch size is 100 and the order quantity is 330, then four times the unit set-up costs is being charged. Tanrikulu et al. (2009) reports that full batch ordering is favored in inventory systems with relatively high backordering cost. Alp et al. (2014) study the problem based on a case where the order size is a partially or fully loaded truck. They formulate the infinite time horizon problem as a MDP and propose two heuristic policies based on the analysis of the MDP formulation. They find that the option of postponing demand would help achieve full truckloads. According to our knowledge, this seems to be the only study in which selective demand postponement is considered in the context of stochastic inventory theory.

In comparison to backorders, stochastic lost sales models are less commonly studied and are harder to solve. Woensel, van et al. (2013) seem first to study the batch ordering inventory system with lost sales and handling costs. The model in this paper has similar characteristics to their model, however also includes the demand postponement strategy. Instead of solving the MDP formulation by backward induction as in Woensel, van et al. (2013), we demonstrate how a reformulation based on dual linear programming can solve instances relatively quickly. Despite these computational advantages, the approach seems relatively unused when compared with other methods in the literature on stochastic inventory theory, see also Powell (2007).

3.3 Model Development

This section introduces the problem with relevant notation and the mathematical formulation of the MDP for an inventory system with stochastic demand, demand postponement and batch ordering. We consider this problem for a single item under a periodic review inventory system.

3.3.1 Problem Description

The products are supplied from an external supplier and batch size a_t placed at the beginning of period t to be delivered at the beginning of period $t + L$, where L is the supplier lead time. L is assumed to be less than the review period. The order size has to be in multiples of batch and each batches contains q items.

The sequence of the events is illustrated in Figure 3.1.

We improve the inventory system with a postponement policy. At some periods, the customers are offered a price discount in exchange of postponing their due dates. Only some of the customers accept the late deliveries. The discount decision z_{ti} refers the

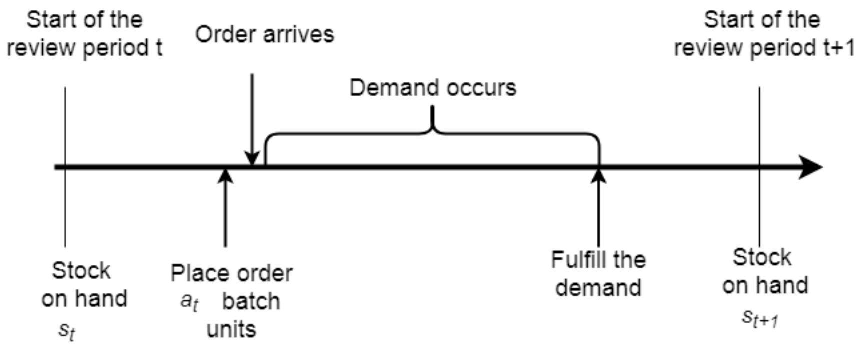

FIGURE 3.1: Sequence of Events

demand size at period t to be decided to postpone to period $i \in \{t+1, ..., t+w\}$ where w is the maximum waiting time for customers. The customers acceptance rate is β_t.

In detail, at the beginning of period t, the manager reviews the on hand inventory level s_t and places an order of a_t with a fixed ordering cost plus unit costs. The order arrives in the same review cycle before the fulfillment of the demand. The order a_t has to be a non-negative integer and multiple of a fixed batch size q, i.e. $a_t \in \{0, q, 2q, ...\}$. By the arrival of the order, the observed demand via postponement is satisfied. To avoid the postponed demand from being lost, the total of on hand stock and placed order is restricted to be greater than or equal to the postponed demand. Then the stochastic demand is observed and postponement decision z_{ti} is made to postpone some of the demand to period i. Next, the remaining demand is fulfilled by on hand stock and unsatisfied demand is lost.

The main notation used in this study is given in Table 3.1.

3.3.2 Mathematical Formulation

We formulate the problem by Markov Decision Process (MDP) for maximising the value of total profit for a finite time horizon. In this section, we introduce the components of MDP which are state and decision variables, transition function and probability, reward function and recursive equation.

The state variable for the inventory problem is mostly the inventory level on hand (Puterman 1994, Woensel, van et al. 2013). The inventory level is a non negative variable due to the lost sales. Including the postponement policy into a classic inventory problem

Table 3.1: Notation

	Parameters
s_t	On hand inventory level at the beginning of period t
D_t	Stochastic demand at period t
$O_{t,i}$	Observed demand consists of postponed demand $i \in t, ..., t+w-1$
K	Setup cost per order (£/order)
u	Purchase cost per unit (£/product)
h	Holding cost per unit (£/product/period)
y_t	Selling price (£/product)
q	Size of a single batch ($units/pallet$)
L	Supplier lead time
T	Time horizon
β_t	Rate of customers who accept to be postponed at period t
w	The maximum postponement length
d_{ti}	Discount amount for postponement from period t to i(£/product)
$\delta(a_t)$	The binary decision variable for order decision $\delta(a_t) \in \{0,1\}$
Q_t	Demand which is willing to accept the late delivery
x_t	The inventory level after the decisions and fulfillment of observed demand
	Variables
a_t	Number of batches ordered in period t
$z_{t,i}$	Postponed demand at period t delayed to period $i \in \{t+1, ..., t+w\}$

makes us to consider the observed demand vector which includes the sum of previous postponement demand. With these, the state variables at the beginning of period t are (s_t, O_t), where s_t is on hand inventory level and O_t is the observed demand vector as in (3.1).

$$O_t = (O_{t,t}, ..., O_{t,t+w-1}) \quad \text{where } O_{t,f} = \sum_{i=f-w}^{t-1} z_{i,f} \tag{3.1}$$

O_t vector has $O_{t,f}$ for $f = t, ..., t+w-1$ which is the cumulative postponed demand at period t for the next periods. These postponement decisions are made in previous periods.

At the beginning of period t, the manager reviews the state variables s_t and O_t and places the order decision a_t. The placed order decision plus on hand inventory must be large enough to cover the observed demand for that period. After the order decision, the stochastic demand which is assumed as based on a known distribution with a probability $p_k = P[D = k]$, $k = 0, 1, ...$ is revealed and the postponement decision $z_{t,f}$ placed. For the simplicity we show the total postponed demand at period t by z_t in (3.2). Then the demand is fulfilled and any unsatisfied demand is lost. The state variables (s_{t+1}, O_{t+1}) for the next period are updated in (3.3), (3.4).

$$z_t = \sum_{f=t+1}^{t+w} z_{t,f} \tag{3.2}$$

$$s_{t+1} = max_{\substack{s_t+a_t \geq O_{t,t} \\ a_t \in \{0,q,2q,...\} \\ z_t \leq \beta_t D_t}} \Big\{ s_t + a_t - O_{t,t} - D_t + z_t, 0 \Big\} \tag{3.3}$$

$$O_{t+1} = (O_{t,t+1} + z_{t,t+1}, ..., O_{t,t+w-1} + z_{t,t+w-1}, z_{t,t+w}) \tag{3.4}$$

The inventory cannot be negative since backlogging is not permitted. In transition (3.3), we add a constraint to guarantee that the on hand stock plus order arrival is sufficient enough to cover the observed demand at that period. This will avoid to lost the postponed demand and guarantee the delivery of postponed demand. We update the observed demand with the postponement decision placed at period t by (3.4). The inventory level after the order decision and satisfaction of observed demand is also updated as in (3.5).

$$x_t = s_t + a_t - O_{t,t} + z_t \tag{3.5}$$

The probability of a transition from period $i \equiv (s_t, O_t)$ to $j \equiv (s_{t+1}, O_{t+1})$ is shown by $p_{ij}(a_t, z_t)$ in (3.6).

$$p_{ij}(a_i, z_i) = \begin{cases} 0 & \text{if } s_{t+1} > x_t \\ P[D = i] & \text{if } x_t \geq s_{t+1} > 0 \text{ and } z_t \leq \beta_f D \quad \forall i = 0, ..., x_t - 1 \\ P[D > x_t] & \text{if } s_{t+1} = 0 \text{ and } z_t \leq \beta_f D \end{cases} \tag{3.6}$$

After the decisions are made, the revenue function (i.e., one-period transition profit) of the current period is calculated based on the ending inventory level as in 3.7.

$$r(s_t, O_t, a_t, z_t) = \sum_{j=0}^{x_t-1} P[D_t = j](jy_t - (x_t - j)h) + \sum_{j=x_t}^{\infty} P[D_t = j]x_t y_t \tag{3.7}$$

For a finite time horizon, the value function is calculated based on a recursive formulation which terminates at $t = T$. (3.8) is to calculate the value of being in state $i_t \equiv (s_t, O_t)$ at period t. When the time horizon is completed, (3.9) provides termination condition which considers only the reward function at the current period.

$$v_t(i_t) = max_{a \in A_{i_t}} \Big\{ r(i_t, a_t, z_t) + \sum_{f \in D} p_{ij}(a_t, z_t) V_{t+1}(j_{t+1}) \Big\} \quad \forall t \in \{0, 1, ..., T-1\} \tag{3.8}$$

$$V_T^*(i_T) = r(i_T, a_T, z_T) \quad \forall a_T \in A_{i_T} \tag{3.9}$$

3.3.3 Linear Programming Model

Puterman (1994) states that linear programming formulations are useful in the study of solving MDPs because of its elegant theory, the fact that constraint inclusions are easy, and offer the advantage of sensitivity analysis. For these reasons and also in order to investigate the computational difference between LP model and backward algorithm, we formulate the MDP for finite horizon as LP model. The LP formulation is;

$$Min \sum_{t \in T} \sum_{s \in S} \alpha_s V_{t,s} \tag{3.10}$$

$$V_{t,s} - \lambda \sum_{j \in S} p(j|s, a, z) V_{t+1,s'} \geq r(s, a, z) \quad \forall a \in A_s, s \in S \forall t \in \{0, 1, ..., T-1\} \tag{3.11}$$

$$V_{T,s} \geq r(s, 0, 0) \tag{3.12}$$

The primal model has $|S|$ decision variables and inequality constraints of dimension $|S|\times|A|$. Since the duality of the model will be computationally better, the dual model is presented in (3.13),(3.14).

$$Max \sum_{t\in T}\sum_{s\in S}\sum_{a\in A_s} r(s,a,z)x(t,s,a,z) \tag{3.13}$$

$$\sum_{a\in A_j} x(t+1,j,a,z) - \lambda\sum_{s\in S}\sum_{a\in A_s} p(j|s,a,z)x(t,s,a,z) = \alpha_{s'}\forall t \in \{0,1,...,T-1\} \tag{3.14}$$

$$x(0,s,a,z) = \alpha_s \tag{3.15}$$

where $x(t,s,a,z)$ refers to the total discounted joint probability that the system occupies state s with an order decision of a and postponement decision z by an initial probability distribution over states $\alpha(s)$. Unless $x(t,s,a,z)$ becomes greater than zero, an optimal decision has not been found yet.

3.4 Computational Experiments

In Section 3.3.2, we develop the mathematical model to formulate the stochastic demand and demand postponement on a batch ordering inventory problem. The unsatisfied demand is lost and we introduce a demand postponement policy which can turn the lost sales into advance demand information. Additionally, we would expect that postponement could also help to deal with the negative impact of batch ordering.

In this section, we investigate how the system behaves with different system parameters through computational experiments. These experiments are carried out in two sections. At first, we did not apply our demand postponement policy to analyse the status quo. And it is observed that how the policy is changing and batch ordering makes the biggest impact under what conditions. We use probabilistic transition to get more accurate results. Next, we apply our postponement strategy and examine the optimal joint inventory and postponement policy. However, due to computational difficulties we assume that transition is deterministic with expected demand. Yet, we calculate the expected reward by probabilities of demand. The comparison with the status quo (with deterministic transition) gives us insight about the value of demand postponement.

The numerical values of parameters are set based on the data obtained from ISolutions case. The demand is assumed to be stochastic based on a Poisson distribution with a mean of λ. The lead time is less than the review period and the demand is satisfied at the end of the period. The other parameters used in numerical tests are set at the following values, $\lambda = 8$, $K = 0, 10, 20$, $y = 1.5u, 1.7u$, $h = 1, 3$.

3.4.1 The effects of Batch Sizes

There is not know simple form of inventory policy for the problem introduced in Section 3.3.1. Only when the base order quantity is $q = 1$, the optimal policy is known to be an (s, S) policy which is order up to S when the inventory level drops to s or below (Woensel, van et al. 2013). We conducted a number of experiments to illustrate the structure of the optimal policy for batch ordering. As an illustration, Figure 3.2 depicts the optimal order quantity by inventory level on hand for different base batch quantities with $K = 10$, $y = 1.5u$, $u = 20, h = 1$.

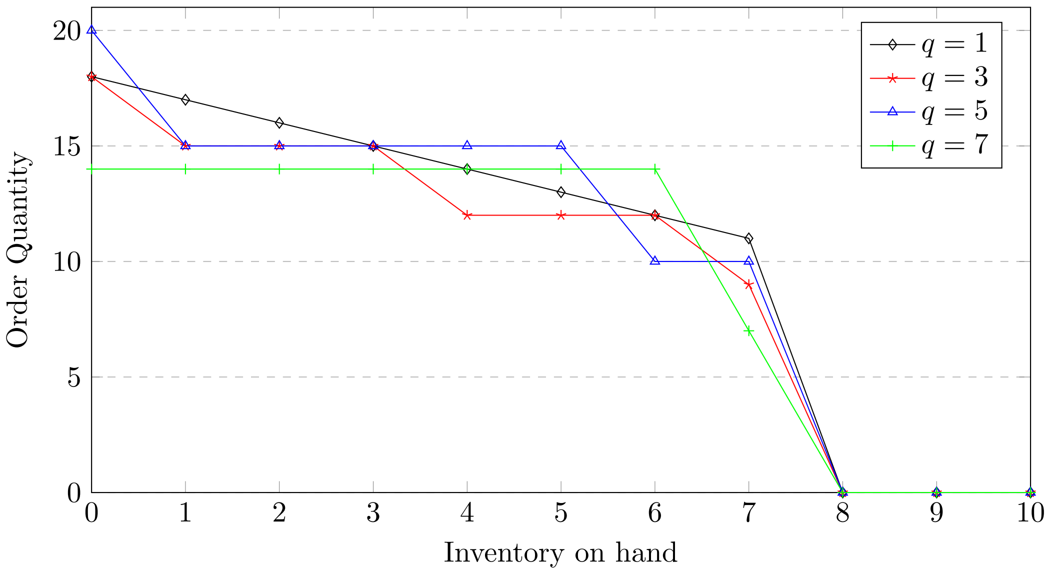

FIGURE 3.2: Optimal Order Quantity by Batch Size

When $q = 1$, the numerical results indicate that the optimal policy is (s, S) policy as in Woensel, van et al. (2013). Changing the base batch quantity results in changes to the optimal policy. The optimal order quantity becomes a step-wise decreasing function of the on hand inventory. The step size is increasing with higher values of q. Although the order amount is varying by batch sizes, the reorder level shows no difference in Figure 3.2 but we further investigate the impact of batch ordering on reorder level in 3.3 by setup cost cost and holding cost. The reason to test with different holding cost is that higher holding cost forces the system to keep less stock on hand. But ordering in batches does not allow the system to choose the 'right' amount to keep on hand. Therefore we expect the higher holding cost to increase the negative impact of batch ordering.

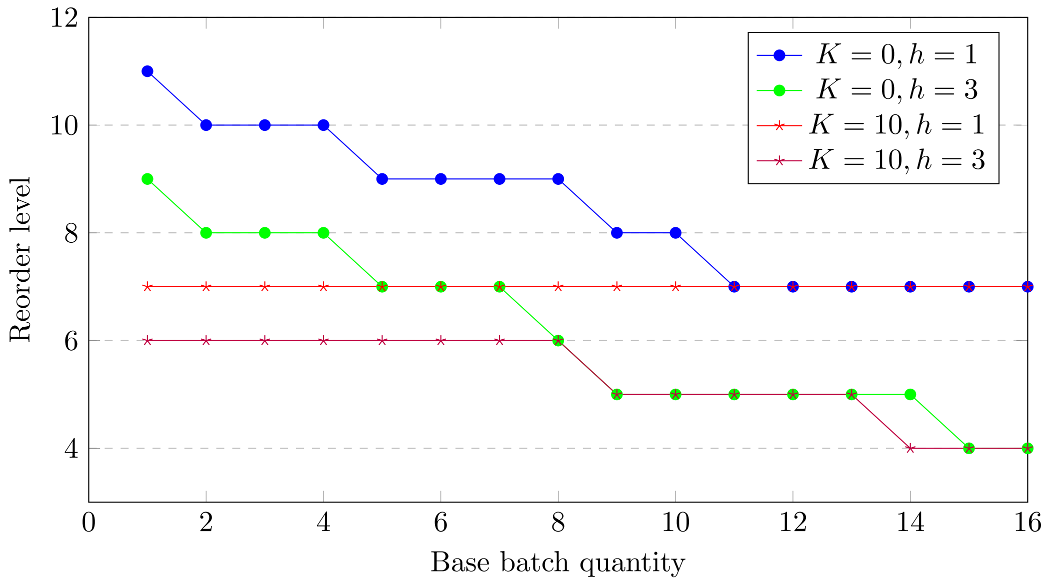

FIGURE 3.3: Reorder Level by Batch Size

Although the graph in Figure 3.3 shows no clear pattern, it does offer insight into the following. Normally the reorder level is decreasing with higher setup costs. However, in the Figure 3.3, after a certain batch size, the setup cost becomes ineffective on the reorder level. This could mean that impact of batch ordering is superior to the setup cost and batch ordering is the dominant parameter that drives reorder level. The holding cost is also an important parameter that affects reorder level, the latter reducing with higher holding cost. In addition, the reorder level does not change by batch size when $K = 10, h = 1$. On the other hand it makes significant changes when holding cost is increased to $h = 3$. This leads us to think that the inventory system with higher holding cost reacts more to batch ordering restriction.

Due to the changes on optimal policy by base batch quantity, we also observe the variation on total expected profit. The impact of batch sizes on profit is calculated by $\delta_m = 100(Profit_1 - Profit_q)/Profit_1$ which represents the percentage deviation on profit by the batch size q while the base profit is for $q = 1$. Figure 3.4 depicts the profit change by batch size and for various setup/ordering costs

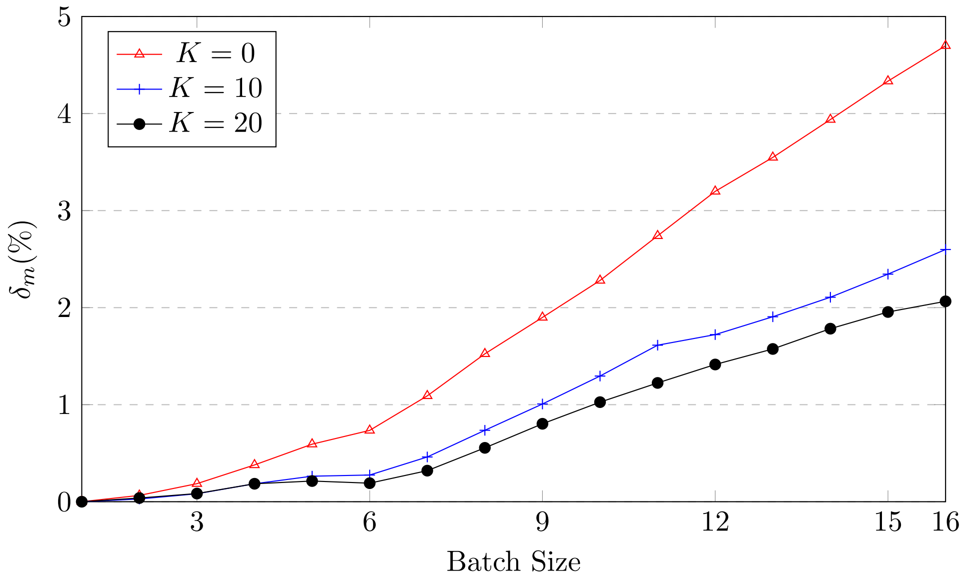

FIGURE 3.4: Deviation on Profit by Batch Sizes and Setup Cost

The profit deviation is getting higher by increasing batch size and lower value of ordering cost, K. For the small batch sizes, the difference is relatively small compared to the higher batch sizes. This clearly shows that the batch ordering makes more impact for the inventory system with lower ordering costs. This is an expected solution since the higher the frequency of order is the lower the value of K. Increasing the number of orders will increase the impact of batch ordering.

3.4.2 The effect of Demand Postponement

In Section 3.4.1 the impact of batch ordering is analysed and the conditions at which batch restriction has more impact are presented. In this section, we apply our demand postponement strategy which aims to convert lost sale into observed advance demand. We test the value of postponement under different parameter settings. To reduce the computational effort, the expected value of demand is used in calculating transition probabilities. The demands that accept the later deliveries are postponed to the next period ($w = 1$). By doing so, the observed demand vector becomes one dimensional and the policy depends on observed demand $O_{t,t}$. The discount policy is shown to follow a (w, W) policy: to offer postponement up to W if the inventory on hand is within $[w, W)$.

We introduce this policy based on the numerical results and show the optimal inventory and discount policy in (3.2).

TABLE 3.2: Optimal Policy, $K = 10, h = 1$, $\lambda = 8$, $O_{t,t} \in \{0, ..., 6\}$

Parameters		0	1	2	3	4	5	6	Avg. Profit	$\delta(\%)$
$\beta = 0$	$S(.)$	20	21	22	23	24	25	26	685.799	0
	$s(.)$	9	10	11	12	13	14	15		
$\beta = 0.25$	$S(.)$	20	21	22	23	24	25	26	714.750	4.22
	$s(.)$	8	9	10	11	12	13	14		
	$W(.)$	12	13	14	15	16	17	18		
	$w(.)$	9	10	11	12	13	14	15		
$\beta = 0.5$	$S(.)$	20	21	22	23	24	25	26	742.140	8.22
	$s(.)$	6	7	8	9	10	11	12		
	$W(.)$	12	13	14	15	16	17	18		
	$w(.)$	7	8	9	10	11	12	13		
$\beta = 0.75$	$S(.)$	22	23	24	25	26	27	28	768.598	12.07
	$s(.)$	5	6	7	8	9	10	11		
	$W(.)$	12	13	14	15	16	17	18		
	$w(.)$	6	7	8	9	10	11	12		

The increasing customer acceptance reduces the reorder level as expected and shows more saving on average profit which is within $4-12\%$. When the β is 0.25 or 0.5, the order up to level remains same with no postponement case. However, making $\beta = 0.75$ also affects the order up to level and increase it. The results indicate that the postponement policy can improve the system performance especially if the customer reaction against postponement is positive. The manager should focus on convincing customers to get more benefit. The results in (3.2) are valid when $q = 1$. But the question still remains how the postponement performs under batch ordering restriction. Thus, we continue the numerical test with the inventory and discount policy by different values of batch sizes. The tests are made with $\beta = 0.5$ and the results are presented in (3.5).

In (3.5), the order quantity levels below 5 depict the number of postponed demand. The discount policy remains the same for all sizes of batches. The step size of batch ordering is reducing due to the postponement policy. The reorder level for $q = 5$ when there is no postponement is 9 while it reduces to 6 when the $\beta = 0.5$. This is a nice reduction but we keep discount amount $d = £1$. For higher discount amount, the difference will

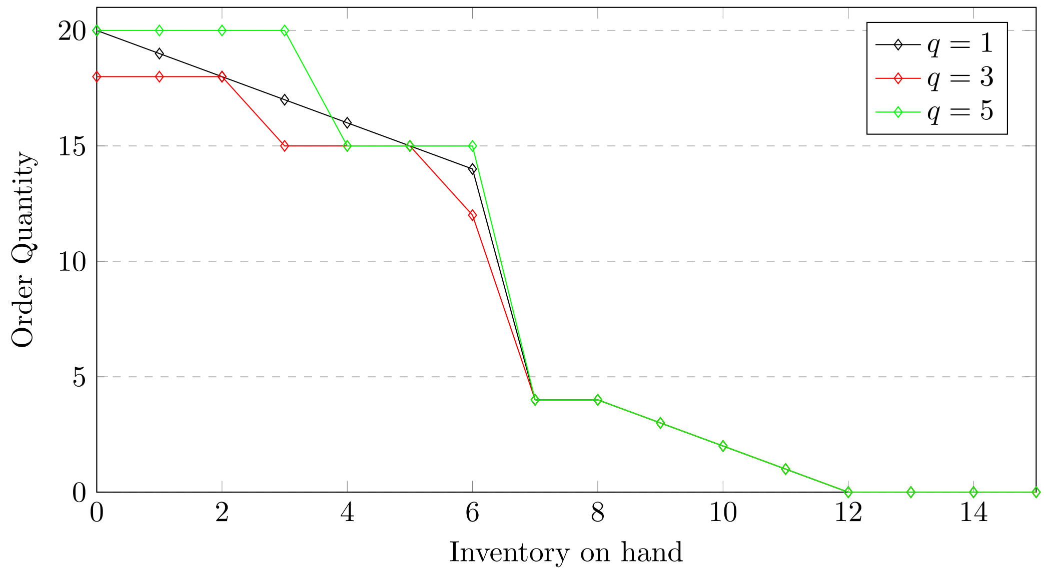

FIGURE 3.5: Inventory and Postponement Policy

be getting closer. Thus, we make a sensitivity analysis of parameters on system profit and show the results in (3.3).

TABLE 3.3: Sensitivity analysis of postponement, $d = 1, u = 20$

		Setup Cost K=10			Setup Cost K=20		
Parameters	β	1	3	5	1	3	5
h=1, y=1.5u	0.25	714.75	713.70	711.31	700.81	699.64	697.33
	0.50	742.14	741.02	738.75	729.67	728.47	726.23
	0.75	768.60	767.33	763.16	757.09	755.62	751.64
h=3, y=1.5u	0.25	624.81	622.32	618.71	602.77	601.36	598.53
	0.50	651.68	621.70	647.24	634.86	633.12	630.43
	0.75	679.26	649.79	674.35	664.54	661.99	659.61
h=1, y=1.7u	0.25	885.34	884.07	881.20	871.02	869.56	866.73
	0.50	912.56	911.31	908.64	899.81	898.46	895.82
	0.75	939.17	937.88	933.93	927.49	926.13	921.99

The gap between the profit of $q = 1$ and $q = 5$ is expected to reduce by more postponement. The gap is 0.46% when $\beta = 0.5$ but it is 0.71% for $\beta = 0.75$. Based on the results in (3.3) the gap is always increasing when the β increases from 0.5 to 0.75 but reducing when β moves from 0.25 to 0.5. This gives us the insight that more usage of the postponement strategy does not always reduce the negative impact of batch ordering. In another words, the postponement makes bigger impact when $q = 1$. The gap between batch sizes is increasing by higher holding cost and setup cost.

3.5 Conclusion

In this chapter, we have studied and analysed a batch ordering problem with lost sales. We propose a postponement strategy which avoids losing profit due to lost sales. If the company can convince some lost sales to convert to advance demand and by offering customers some discount, mutual benefit can occur.

The findings in numerical experiments state the impact of batch ordering and demand postponement under different system conditions. The impact of batch ordering is relatively higher for higher holding and lower setup costs. But the postponement policy makes the impact of setup cost reverse. In case of offering postponement under high setup costs, the negative impact of batch ordering is higher. Normally, it would be higher for lower setup cost. This indicates that postponement policy performs better for lower setup costs.

One of the limitation of this study is the curse of dimensionality due to the requirement to keep track of observed demand. The numerical results, however, indicate structure in the optimal policy, which can be used in further research on the construction of heuristic methods. Continued improvements in the technological capability and increasing availability of better solvers makes the use of LP models an attractive avenue for solving MDPs.

Chapter 4

Selective state-dependent purchasing of buyers' willingness to wait

Abstract

We consider a single item, periodic review inventory system with advance demand. Regular customers place their orders with a heterogeneous time ahead of their needs within a planning time horizon. The focus in the literature has been on how to stimulate customers towards advance demand. Predicting how demand will shift can be problematic, however, and backorders may still occur. We focus on how a firm can address backorders under a given advance demand pattern by a mechanism of compensation from which both the firm and the customers will be benefited. We consider that the firm may offer a price discount to customers for accepting later deliveries. Discounts are only offered in some periods and to some customers when there is a benefit for the firm to postpone some of the demand. Customers may decline the offer, but then face the probability of a backorder. If they agree, they get a promised delivery date and financial compensation. In each period, the firm has to decide whether to offer a price discount and for how long, and whether to order from its supplier and how much. We formulate the problem as a Markov Decision Process and solve it by a backward induction algorithm. We establish the structure of optimal policies with respect to stationary costs. Numerical examples

illustrate the properties of the state-dependent optimal policies obtained for the both uncapacitated and capacitated inventory systems.

4.1 Introduction

Inventory control is concerned with matching the supply of a product to its demand. Supplier lead-time (SLT), i.e. the time it takes to receive the products of an order placed at the supplier, forms an essential component of many inventory models. In general, however, also the time when customers place an order may differ from the time the order is arranged to be delivered. In the inventory literature, this time-lag is defined as the customer lead-time or 'demand lead-time' (DLT), and the techniques to control and exploit knowledge of DLTs as inventory control under Advance Demand Information (ADI).

Matching the supply to uncertain demand is a challenging task if SLT is non-zero, i.e. when supply cannot occur instantaneously. If supply exceeds demand, holding costs are incurred, and problems with warehouse capacity may arise. In the other case, backorders or lost sales occur. Improved performance may be reached by either reducing the SLT or increasing the DLT (Kremer and Van Wassenhove 2014). Typically, the first is thought of improving the responsiveness of the system, while the latter improving its anticipatory power. Reducing SLT can be challenging if there are hard constraints imposed by production and transportation times etc. In some contexts, there might be a higher potential to increase the DLT. However, this usually requires renegotiating the terms offered to customers.

Changing the DLT is considered in various businesses in the context of revenue management (Talluri and Hillier 2005). Increasing the DLT can be realized in two different ways. The first is known as advance booking or pre-ordering, whereby customers inform the company of their orders sooner, but without changing their due date. This strategy is used by Apple, Amazon and Playstation in particular for new products, not yet released. For movies, games, or electronic products, firms may use it to induce customers to commit using these services at a later scheduled date, who in return receive a guarantee on their availability at a discount (Li and Zhang 2013). This strategy locks in demand at earlier times, and enables the firm to better anticipate how to meet these future commitments with current decisions. It can indeed be said to improve the system's anticipatory power. A second strategy to increase DLT is known as demand postponement, and consists of offering customers incentives to postpone the date of delivery. It is a strategy applied in the airline industry as a means to ameliorate overbooking (Tang 2006b), whereby customers are offered financial compensation if they would accept later flights. This strategy is rather different to pre-ordering because it is only offered selectively and when needed. Rather than being anticipatory, this approach to increasing DLT is reactive, and only selectively applied depending on the state of the system.

The literature review (Section 4.2) shows that the majority of studies on ADI in the context of production-inventory systems focus on getting more DLT from customers by offering them pre-announced DLT-dependent price schemes. Because these schemes are open to all customers, the level of uptake may highly affect the profitability of the firm, and it may be difficult to predict how the customers base will react to changes to the scheme. In this paper, we examine the option to use demand postponement as a means of better matching supply and demand in these systems. Provided that at least a fair proportion of customers are open to demand postponement, the financial risk involved in this approach seems less, as the firm can control the selective application of the policy. In this paper we do not model risk explicitly, but investigate the impact on the firm's expected profits.

In particular, we examine the potential of demand postponement under various DLT order patterns in a periodic-review inventory system. These patterns may range from very short DLTs (no ADI), to various heterogeneous DLT patterns (ADI). The demand postponement option is offered at some time periods only and to some particular customers only. We assume that the firm engages in such negotiations at the time when customers place their orders. From a customer service perspective, this approach is better than when such negotiations would take place afterward and thus closer to an initially set due-date. It also signals that the firm is knowledgeable about its own supply chain constraints, and desires to make promises it can keep. As an incentive, the firm will offer a price discount on the order if a customer accepts the later delivery date. Like in other ADI models, the fact of having a policy to affect DLTs will also affect the order policy of the firm for orders placed at its supplier. The inventory control problem for the firm is hence to find the best joint supply ordering and ADI synchronization policy with respect to the trade-off between the loss due to awarded financial incentives and the gain in operational and backorders cost.

4.2 Literature Review

The study of ADI in inventory control is relatively recent. Hariharan and Zipkin (1995) introduce the demand lead-time (DLT) concept and address the case of a continuous review system where all customers have the same DLT. They show that increasing DLT is equivalent to reducing SLT as both reduce uncertainty on future demand in the same manner. When DLT grows and becomes equal to SLT, the optimal inventory policy changes from make-to-stock to make-to-order.

The more general case of customers being heterogeneous in their DLT is addressed in Gallego and Özer (2001). They consider a periodic inventory problem and model the ADI by the vector of observed demand during period t, $D_t = (D_{t,t}, ..., D_{t,t+N})$ and where $D_{t,f}$ is the demand placed in period t to be delivered at period $f(f \geq t)$, with a DLT of $(f-t)$. In the case of zero set-up costs, they show that a state dependent base stock policy is optimal, and that demand information about due dates that fall behind the SLT plus a review period (which is called the *protection period*), has no operational value. For strictly positive set-up costs, the optimal policy structure becomes a state dependent (s, S) policy. For a detailed review of literature with ADI, we refer readers to Gallego and Özer (2002) and Özer (2011).

The capacitated inventory system situations of the above two models are presented in Wijngaard and Karaesmen (2007) and Özer and Wei (2004), respectively. For positive set-up costs, the latter authors prove the optimality of a threshold policy. This policy involves ordering at full capacity when the inventory level at the end of the protection period (which is called the *modified inventory level*) falls below a threshold, and otherwise order nothing. The papers illustrate that ADI increases the effective usage of capacity. Multi-stage inventory systems with ADI are examined in Gallego and Özer (2003), Özer (2003) and Sarkar and Shewchuk (2013). Lu et al. (2003) study a multiproduct assemble-to-order system with ADI and stochastic lead times.

Other research in ADI systems investigates the benefits of allowing the firm to ship products to customers earlier than the due date. This approach is referred to as flexible delivery or early fulfillment. The exploitation of early delivery can significantly reduce inventory costs, see Karaesmen et al. (2004), Sarkar and Shewchuk (2016), Xu et al. (2017). Wang and Toktay (2008) investigate the cases with homogeneous and heterogeneous DLTs, and prove that increasing DLT is preferred over reducing SLT for the homogeneous DLT case. A variation is examined in Sarkar and Shewchuk (2013), in which there are two customer classes with equal priority, one class having customers with zero DLT, and the other class with a positive constant DLT. Xu et al. (2017) study the flexible delivery on ADI with homogeneous DLT and a penalty cost if the delivery is made later than its due date.

Most studies assume that ADI, once obtained, is fixed. In studies with imperfect ADI, such as in Tan et al. (2007), Gayon et al. (2009) and Benjaafar et al. (2011), it is however allowed that customers may cancel placed orders or make amendments to their order. The division of customers into different classes is sometimes used so that not only customers may have different DLTs but also different probabilities about cancellations or amendments. Gayon et al. (2009) exploit this in a Markov Decision Process having different transition probabilities for different demand classes.

Since the results in above studies have identified benefits to the inventory system owner of having more customers adopting longer DLTs, some studies focus on methods to establish such transitions towards more ADI. Most studies have focused on offering financial incentives and evaluating the trade-off with the additional gains from increased ADI in lowering inventory or other system costs. Unlike in classic price discount models, where the overall demand may be affected, the studies in ADI typically keep total demand constant, but focus on establishing how price discounts may affect receiving future demand information earlier, or thus increasing DLT through advance booking. As customers might differ in their willingness to wait for their orders, Chen (2001) offers a price schedule where unit prices are non-increasing function of DLT, and customers can view the price schedule in advance and self-select according to their preference. Kunnumkal and Topaloglu (2008) use price discounts to reduce the standard deviation of demand to a level which minimises the total relevant costs.

It should be noted that increasing DLT by financial incentives does not always lead to net gains. This is supported by the study of Karaesmen et al. (2004), who model the price discount as a function of DLT for the case of homogeneous customers in a production/inventory system in order to identify the optimal ADI level. Li and Zhang (2013) consider a model offering a price discount or price guarantee to customers in exchange of adopting a larger DLT through pre-ordering, but find that this policy actually may reduce the seller's profit in situations of low demand seasonality. Buhayenko and van Eikenhorst (2015) study the coordination of a supplier production schedule with deterministic demand through a price discount as an incentive for changing customers due dates and when no backorders are allowed.

In Chen (2001) or Karaesmen et al. (2004), the aim is to obtain ADI through an incentive scheme offered to all customers in advance. The objective of this study is to take a given ADI setting as a starting point for seeking further improvements. That is, we assume that some level of ADI is present in the demand pattern, including the case when there is no ADI at all (DLT $= 0$). We focus on establishing when the seller would benefit from offering to 'buy' additional ADI by deciding when to make selective offers for demand postponement to some customers at the time they place their order.

While demand postponement is a strategy already widely applied in the airline industry, inventory systems are very different in terms of structure. In an inventory system, for example, there are no fixed flights schedules. An inventory system may or may not have capacity constraints, and both products and orders generally differ in when they enter and leave the system. Customers may also reject the offer of demand postponement and then expect to have the products at the desired time, which the firm may or may not be able to meet. Latter case will result in a backorder.

We study the value of demand postponement in a system with stochastic demand that is similar in set-up to that of Gallego and Özer (2001). We formulate the problem of finding the best inventory and offer synchronization strategy as a Markov Decision Process and solve it by backward induction. We examine the potential of this strategy in uncapacitated and capacitated systems, in cases of zero and positive set-up costs, and at various levels of ADI present in the system.

4.3 Problem Description

This section provides the formal description of the problem and introduces in particular the modelling of the demand structure in the context of DLT and postponement.

A periodic review inventory problem for a single product type is considered. Table 4.1 summarizes the notation. The firm supplies its stock from an external supplier. A supply order placed at the beginning of period t is delivered at the beginning of period $t + L$, where L is the supplier lead time. With each supply order placed the firm incurs a fixed setup cost K_t and pays for the items based on a purchase cost per unit item u_t. These cost parameters can be dependent on the time period t in which the order is placed.

The firm has a large customer base. For each customer, it has agreed on a contract which specifies an (customer-specific) agreed DLT that applies to all orders the customer will place. The maximum possible DLT is called the *information horizon* and denoted by N, where we assume $N \geq L + 2$ (without loss of generality, see Section 4.4.1). Customers can place orders at random time periods. Most of the time, an order placed in period t is accepted with the aim to deliver the products in period $t + DLT$, i.e. according to the agreed DLT. In some periods, the firm will consider making a demand postponement offer to some customers in order for them to arrange delivery beyond the agreed DLT, who either accept or reject this offer.

The sequence of events in period t can be summarized as follows: (1) Order from supplier if placed in period $t - L$ is received; (2) Decision is made to either place a supply order and the quantity, or not to place an order; (3) During period t, customer orders are received, if relevant then postponement offers are made, and the responses from customers to accept or reject the offer are received; (4) At the end of the period, customer orders needing delivery are fulfilled or backordered.

4.3.1 Demand Structure

We consider the advance demand pattern structure as modelled in Gallego and Özer (2001). For the sake of clarity, we shortly review its main components. We do not yet consider the impact of the postponement policy, which will be introduced in Section 4.3.2.

During a period, the firm gathers the demand information from customers, each of whom may have a specific contract DLT. If $D_{t,s}$ denotes the total demand placed at period t to be delivered according to the contract DLT at period $s \in \{t, .., t+N\}$, then at the end of period t, the firm has obtained a demand vector $D_t = (D_{t,t}, ..., D_{t,t+N})$.

A similar process has occurred during periods prior to t. Therefore, at the beginning of period t and prior to obtaining D_t, the cumulative observed demand to be delivered at period $s \in \{t, .., t+N\}$ according to the contract DLT, will consist of demand orders placed in periods no earlier than period $s-N$ and no later than $t-1$. If none of the demand is postponed, the cumulative observed demand at the start of period t to be

Table 4.1: Notation

	Parameters
T	Time Horizon
N	Information Horizon
L	Supplier Lead Time
C	Capacity Limit
α	Opportunity cost of capital rate per period
P_t	Fraction of demand which is accepting the discount at period t
I_t	Inventory on hand at the beginning of period t
B_t	Backorder level at the beginning of period t
x_t	Modified Inventory position at period t before the decisions
y_t	Inventory position at period t after decisions made at period t
	Demand
$O_{t,s}$	Cumulative observed demand for period s at the beginning of period t
O_t	Observed demand beyond protection period at the beginning of period t
O_t^L	Observed protection period demand at the beginning of period t
$D_{t,s}$	Demand placed at period t to be delivered at the end of period s
D_t	Demand placed at the end of period t, $(D_{t,t}, ..., D_{t,t+N})$
Q_t	Demand willing to accept the price discount at period t
	Costs
K_t	Setup cost per order in period t (£/order)
u_t	Purchase cost per unit in period t (£/product)
h_t	Holding cost per unit in period t (£/product/period)
p_t	Shortage cost per unit in period t(£/product/period)
d_{ti}	Discount amount for a product to postpone to period i from period t(£/product)
$c_t(s_t; a_t)$	The expected cost including inventory and decision costs.
	Variables
z_t	Order amount placed in period t at supplier to be delivered at $t+L$
q_{ti}	Amount of discounted demand at period t delayed to $i \in \{t+L+1, ..., t+N\}$

delivered in s thus consists of:

$$O_{t,s} = \sum_{r=s-N}^{t-1} D_{r,s}. \tag{4.1}$$

At the beginning of any period t, the observed quantity $O_{t,s}$ can be expected to be only part of the demand destined to be delivered in s, since at any period $k \in \{t, .., s\}$ additional demand orders will arrive. Let $U_{t,s}$ denote the quantity of yet unobserved demand at the start of t for delivery in s, then:

$$U_{t,s} = \sum_{r=t}^{s} D_{r,s}, \tag{4.2}$$

where these $D_{r,s}$ quantities are still under determined random variables. The sum of $O_{t,s}$ and $U_{t,s}$, were the latter quantity known, would thus be the total quantity to be delivered in s at the beginning of t, prior to making any future offers to postpone.

The *protection period* refers to the period covering the next $L+1$ periods; any customer orders to be delivered during this time can only be fulfilled from on hand stock and planned supply order arrivals. At the beginning of t, the total observed demand to be delivered in periods that fall within the protection period is thus:

$$O_t^L = \sum_{s=t}^{t+L} O_{t,s}, \tag{4.3}$$

while the total unobserved demand within this protection period equals:

$$U_t^L = \sum_{s=t}^{t+L} U_{t,s} \tag{4.4}$$

Figure 4.1 illustrates the demand pattern structure available in period t. Observed and unobserved demand at the beginning of this period are displayed above the x-asis, where the components of D_t obtained during the period are displayed below.

Example 4.3.1. Let us assume that we have a case of $N = 2, L = 0$, and no postponement is offered. Then at the beginning of period t, we have $O_{t,t}$ and $O_{t,t+1}$ as observed cumulative demand from previous periods. Also we have unobserved demand $U_{t,t} = D_{t,t}$ and $U_{t,t+1} = D_{t,t+1} + D_{t+1,t+1}$. During the period t, the demand vector $D_t = (D_{t,t}, D_{t,t+1}, D_{t,t+2})$ is observed. We deliver $O_{t,t}+D_{t,t}$. To move towards the next period,

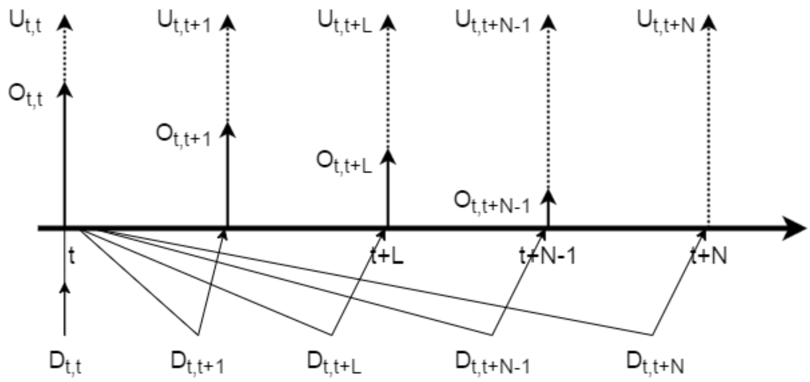

FIGURE 4.1: Demand Pattern

we update the observed demand as $O_{t+1,t+1} = O_{t,t+1} + D_{t,t+1}$ and $O_{t+1,t+2} = D_{t,t+2}$. The unobserved demand is also updated, $U_{t+1,t+1} = D_{t+1,t+1}$ and $U_{t+1,t+2} = D_{t+1,t+2} + D_{t+2,t+2}$.

4.3.2 Demand Postponement

Customers are offered a price discount to postpone the due date for their demand. The discount decision is only available for the protection period demand at only some times. The reason to consider offering a discount only to these demands is to protect the system against the lead time demand. The postponement period is a decision variable. Each customer is treated separately and they might have different delay periods.

After observing the D_t, the manager decides whether to offer price discount. The protection period demand placed at t consists of $(D_{t,t}, ..., D_{t,t+L+1})$ and the discounted demand among these are postponed to a period $i \in \{t + L + 2, ..., t + N + 1\}$. Postponement period is also a decision variable in our problem. Only a specific rate of the customers P_t accepts the discount and the demand willing to accept the postponement is shown by 4.5.The discount decision offered to the demand placed at period t is illustrated in Figure 4.2.

$$Q_{t+1} = P_t \sum_{f=t}^{t+L+1} D_{t,f} \tag{4.5}$$

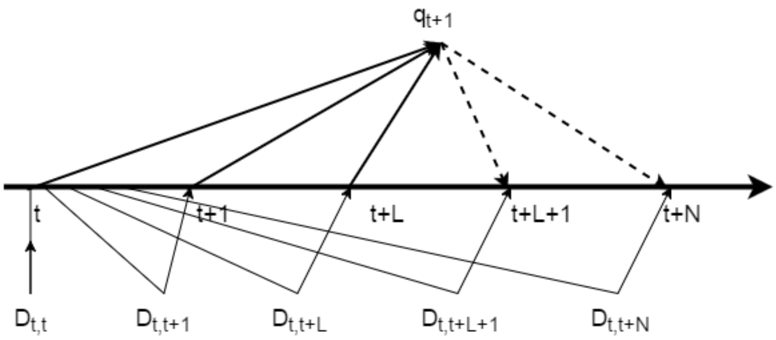

Figure 4.2: Discount Decision

4.4 Model Formulation

The mathematical model and its components are presented in this section. We formulate the problem as a discrete-time Markov Decision Process (MDP). The main assumptions of the problem are as follows.

1. Supplier lead time is deterministic and constant.
2. Discounts are not announced beforehand. After all customer order placements are collected during a period, then the decision is made whether or not to offer the price discount postponement offer.
3. The amount of discount per product offered can be a function of the number of periods delivery will be postponed relative to the DLT of the demand.
4. The unobserved part of the demand is not dependent on the observed part.
5. We assume that stationary costs, action and transition functions.
6. All unsatisfied demands are backordered.
7. Early delivery of a demand is not allowed.

The components of a MDP are state space, decision space, transition function/probability and reward function.

4.4.1 State and Decision Variables

At each decision epoch, the system occupies a state from a state space. Gallego and Özer (2001) show that the state space for advance demand information has a dimension of $1+(N-L-1)^+$. The state variables are modified inventory position x_t and observed demand beyond protection period O_t. Adding postponement decision into the system obligates us to keep track of the information of available demand Q_t for postponement. So, the state space becomes $2+(N-L-1)^+$ dimensional. Notice that when $N < L+2$, the system becomes two dimensional. However the discount decision will not take place since there is no point affecting demand beyond the protection period. Therefore the problem becomes one dimensional.

The state of the MDP at the beginning of period t is shown by $s_t = (x_t, O_t, Q_t)$ where;

$$x_t \equiv I_t + \sum_{s=t-L}^{t-1} z_s - B_t - O_t^L \tag{4.6}$$

$$O_t \equiv (O_{t,t+L+1}, ..., O_{t,t+N-1}) \quad \textit{where;} \quad O_{t,s} = \sum_{r=s-N}^{t-1} (D_{r,s} + q_{r,s}) \tag{4.7}$$

$$Q_t = P_{t-1} \sum_{f=t-1}^{t+L} D_{t-1,f} \tag{4.8}$$

x_t is the planned ending inventory at the end of period $t+L$ by the planned order arrivals including on hand inventory and minus observed demand for protection period as in (4.6). O_t is the cumulative observed demand vector beyond protection period composed of placed demand and delayed demand at previous periods. This is the demand information needed to be considered for future periods. The last state variable is Q_t which is the number of demand willing to accept the postponement.

Based on the state variables, order and discount decisions $a_t = (z_t, q_{ti})$ are placed. where z_t is the order decision and q_{ti} is the postponed demand placed at period t to be delivered at period i. The sum of postponed demand is shown as q_t by (4.9). The number of demand available for postponement is limited by Q_t, the limitation constraint is given in (4.10).

$$q_t = \sum_{i=t+L+1}^{t+N} q_{t,i} \tag{4.9}$$

$$\sum_{i=t+L+1}^{t+N} q_{t,i} \leq Q_t \tag{4.10}$$

The cost placing the decision set $a_t = (z_t, q_{ti})$ is notated as $\beta(a_t)$ and calculated by 4.11.

$$\beta(a_t) = K\delta(z_t) + z_t u + \sum_{i=t+L+1}^{t+N} (d_i q_{ti}) \tag{4.11}$$

After the decisions are made, the modified inventory position is updated to y_t in 4.12. Unless there is a capacity constraint, the order amount is unbounded. For the capacitated systems considered in this paper, y_t is capacitated by a limit C_t as in (4.13).

$$y_t = x_t + z_t + q_t \tag{4.12}$$

$$y_t \leq C_t \tag{4.13}$$

The discount decision increases the ending inventory as the ordering decision in (4.12). But also, the demand beyond protection period increases by price discount decisions, since the part of the current demand is postponed to those periods.

4.4.2 State Transitions

The system state at the next decision stage depends on the state and the decisions made at the current period. The transition between one state to another could be either deterministic or stochastic. However, considering the stochastic transition increases the computational complexity of the model exponentially. Therefore, in our study, we focus on deterministic transition like most ADI literature and the function of moving to $s_{t+1} = (x_{t+1}, O_{t+1}, Q_{t+1})$ are shown in (4.14), (4.15), (4.16).

$$x_{t+1} = y_t - D_{t,t} - \sum_{s=t+1}^{t+L+1} D_{t,s} - O_{t,t+L+1} - q_{t,t+L+1} \tag{4.14}$$

$$O_{t+1} = (O_{t+1,t+L+2}, ..., O_{t+1,t+N}) \tag{4.15}$$

$$Q_{t+1} = P_t \sum_{f=t}^{t+L+1} D_{t,f} \tag{4.16}$$

4.4.3 Cost Function

After the decisions, the net inventory level at the end of period $t + L$ becomes $y_t - U_t^L$ where U_t^L is the unobserved demand for protection period. The expected inventory costs at that period is calculated based on y_t by (4.17). The expected value is due to the uncertainty on demand placed at period t.

$$\hat{G}_t(y_t) \equiv \frac{\gamma^{t+L}}{\gamma^t} Eg(y_t - U_t^L) \tag{4.17}$$

where γ is the discount factor and $g(x)$ is total expected holding and backorder cost based on the inventory level x. We assume that it is a convex function and coercive since $\lim_{|x|\to\infty} g(x) = \infty$. These are the common assumptions in inventory literature and they are valid when the holding and backorder costs are linear.

The total expected cost including inventory and decisions cost is then shown by c_t and calculated by (4.18).

$$c_t(s_t; a_t) = \hat{G}_t(y_t) + \beta(a_t) \tag{4.18}$$

4.4.4 Recursive Equation

With all the given components, the recursive equation also known as Bellman Equations are formulated as in (4.19) and (4.20).

$$V_T^*(s_T) = c_T(s_T, a_T) \quad \forall s_T \in S, \forall a_T \in A_{s_T} \tag{4.19}$$

$$V_t^*(s_t) = min_{a_N \in A_{s_N}} \Big\{ c_t(s_t; a_t) + \alpha E V_{t+1}(s_{t+1}) \Big\} \quad \forall t \in \{0, 1, ..., T-1\} \tag{4.20}$$

$V_t^*(s_t)$ is the optimal expected cost when the system in state s_t at time t. 4.19 presents the termination condition when time horizon is completed.

4.5 Structure of the Optimal Policy

The recursive function is given in 4.21 and now we will further analyse the system to derive the structure of the optimal policy. For simplification, the function $P_t(y_t, O_t, Q_t)$ denotes the expected cost of moving from a state at t to another state in $t+1$ with ignoring the decision costs as in (4.22).

$$V_t^*(s_t) = min_{z_t \geq 0, q_t \leq Q_t} \Big\{ \hat{G}_t(y_t) + K\delta(z_t) + z_t u + \sum_{i=t+L+1}^{t+N} (d_i q_{ti}) + \alpha E V_{t+1}(s_{t+1}) \Big\} \quad (4.21)$$

$$P_t(y_t, O_t, Q_t) = \hat{G}_t(y_t) + \alpha E V_{t+1}(x_{t+1}, O_{t+1}, Q_{t+1}) \quad (4.22)$$

Next, we define $H_t(x_t, O_t, Q_t)$ as the cost of moving to another state when the order decision is placed. Thus, it is expected that when $H_t(x_t, O_t, Q_t) \leq 0$ then it is optimal to order since the cost of moving to the next state with an order decision is negative.

$$H_t(x_t, O_t, Q_t) = K_t + min_{y_t \geq x_t} \{P_t(y_t, O_t, Q_t) - P_t(x_t, O_t, Q_t)\} \quad (4.23)$$

Lemma 4.1. *When the $g_t(x)$ is convex, then for any given (O_t, Q_t), the following statements are;*

i $H_t(., O_t, Q_t)$ is convex.

ii $V_t(x, O_t, Q_t)$ is nondecreasing convex in x.

When the Lemma 4.1 is satisfied, then the reorder level s_t is the maximum inventory level in which the H_t function becomes positive. The order up to level S_t is the minimum value of inventory after the order decision in which expected cost of new inventory is less than the cost of current inventory as shown in (4.24).

$$s_t(O_t, Q_t) = max\{x : H_t(x, O_t, Q_t) \leq 0\}$$
$$S_t(O_t, Q_t) = min\{y : P_t(y, O_t, Q_t) \leq P_t(x, O_t, Q_t) \text{ for all } x\} \tag{4.24}$$

After the structure of replenishment policy, the discount policy is presented now. Intuitively, it is expected that order decision and discount decision cannot take place together since they both have the same purpose. In fact, one of the aim of using discount is to postpone the order decision.

For the case $N = L+2$, we have two dimensional discount decision. The discount policy needs to determine when to offer the discount and postpone to either period $t+L+1$ or $t+L+2 = t+N$. The benefit function when the discount decision with delay time i is placed defined as in 4.25. The delay period affects the expected future cost differently due to their different impact on transition functions which is previously defined.

$$J_t^i(x_t, O_t, Q_t) = min_{q_{ti} \leq Q_t}\{d_i q_{ti} + P_t(y_t, O_t, Q_t) - P_t(x_t, O_t, Q_t)\} \tag{4.25}$$

Definition 4.2. Let us define the inventory values of $w_t^1(O_t, Q_t)$ and $w_t^2(O_t, Q_t)$ where as the minimum inventory position at which postponing is better than ordering. Thus, postponement starts from $w_t^i()$ up to the level $W_t^i(O_t, Q_t)$.

$$w_t^i(O_t, Q_t) = min\{x : J_t^i(x, O_t, Q_t) \leq min\{0, H_t(x, O_t, Q_t)\}\}$$
$$W_t^i(O_t, Q_t) = min_{q_{ti} \leq Q_t}\{y = x + q_{ti} : P_t(y, O_t, Q_t) \leq P_t(x, O_t, Q_t)\} \tag{4.26}$$

Then the optimal expected cost value at time t with the decision is presented as in (4.27).

$$V_t^*(x_t, O_t, Q_t) = P_t(x_t, O_t, Q_t) + min\{0, H_t(x_t, O_t, Q_t), min\{J_t^i(x_t, O_t, Q_t)\}\} \tag{4.27}$$

4.6 Analysis & Discussion

4.6.1 Action Elimination

Increasing number of state space cause to increase computational complexity. Reducing action space will help to narrow the searching area at each iteration. In this part, we focus on reducing the action sets in order to improve the computational complexity.

Theorem 4.3. *The following statements hold;*

1. *If* $K_2 > K_1 > 0$ *then* $s_t(o|K_2) \leq s_t(o|K_1)$
2. $s(o_2) \geq s(o_1)$ *if* $o_2 \leq o_1$
3. $s_{t-1}(o) \geq s_t(o)$

Statement 1 indicates that for the systems with higher setup cost will have lower reorder level. While the observed demand beyond protection period is increasing, the order level is decreasing and shown in Statement 2. These statements clearly show that the reorder level is decreasing by increasing setup cost and increasing observed demand. For finite-horizon with a discount rate, the reorder level is reducing by time since the system is getting closer to the terminal point as in Statement 3.

These will lead us to accept that reorder level for zero setup cost case is an upper bound for positive setup cost cases. Gallego and Özer (2001) prove that the optimum policy is base stock policy when the setup cost is zero. Thus, for stationary problems, base stock parameter of myopic policy which only minimises the cost of current stage is optimal for finite time horizon. This parameter is calculated by 4.28.

$$y^m_{min} = min\{y : G_t(y) = min_x G_t(x)\} \tag{4.28}$$

$$y^m_{min} = max\{y : G_t(y) = min_x G_t(x)\} \tag{4.29}$$

Now, when we take the statements in Theorem 4.3, it is concluded that the reorder level for zero setup cost is an upper bound for reorder level of positive setup cost case. Calculating the policy parameter for $K = 0$ is as easy as shown in 4.28 and this could be used to reduce the action space on a case with $K > 0$. We assume that $s_t(K_0)$ be the optimal base stock policy parameter for the zero setup cost case. Then, we can make an additional constraint as (4.30)

$$z_t = 0 \quad if \quad x_t \geq s_{K_0} \tag{4.30}$$

Practically, (4.30) will reduce the searching area. Because when the modified inventory level, x_t is greater than the reorder level, s_{K_0} found by (4.28), then it is known that the optimal decision is not to order.

4.6.2 Solution Algorithm

Solving the Bellman equations is often performed by backward induction. The backward algorithm starts to calculate the value function at the last period for each of the states. It then steps back in time by one period and calculates the values by adding it to the next periods value function. The algorithm is shown in Algorithm 1.

Algorithm 1: Backward Algorithm

Step 1. Initialize the terminal values. Set $t = N$;
$V_N(s_N) = c_N(s_N) \quad \forall\ s_N \in S$;

Step 2. Substitute $t-1$ for t and compute $V_t(s_t)\ \forall\ s_t \in S$ by ;
$V_t^*(s_t) = min_{a \in A_{s_t}} \left\{ c_t(s_t, a) + \alpha \sum_{s' \in S} p_t(s'|s, a) V_{t+1}(s'_{t+1}) \right\}$;
Set ;
$a_{t,s_t} \equiv argmin_{a \in A_{s_t}} \left\{ c_t(s_t, a) + \alpha \sum_{s' \in S} p_t(s'|s, a) V_{t+1}(s'_{t+1}) \right\}$;

Step 3. If t=1 stop; Otherwise return to step 2 ;

4.7 Computational Results

In this section, we carry out a number of numerical experiments to analyse the behaviour of optimal policies under advance demand and the impact of postponement on these policies. The most inclusive and simple case on our problem without loss of generality is to test the case where $N = L + 2$ in which the observed demand beyond protection period has become one dimensional. We follow Gallego and Özer (2001) for the demand pattern and numerical values of parameters.

Advance demand $D_t = (D_{t,t}, D_{t,t+1}, D_{t,t+2})$ is modelled as Poisson distribution with means $\lambda = (\lambda_0, \lambda_1, \lambda_2)$ when $N = 2$. To see the impact of more ADI available we test the model with different values of λ. While doing this, we keep the total demand rate constant so that we are able to see the effect of the policy which shifts the urgent demands to advance demands. The value of ADI in the uncapacitated and capacitated cases are analysed and compared.

The effect of the discount is also tested under different parameters. Since we have two dimensional discount decision, we offer higher discount amount to those whose postponement period is longer. The customers reaction to postponement is chosen as $P = 0, 0.5, 1$. When $P = 0$, then only at the current period, we can offer postponement if there is any demand available. For future periods, there is no demand available for postponement. This can be seen as the worst case scenario. $P = 0.5$ means that only half of the customers welcome the postponement for all periods. All customers accept the discount when $P = 1$ which could be stated as the best case scenario in terms of customers acceptance of postponement.

Gallego and Özer (2001) prove that the optimal inventory policy for zero setup cost case with advance demand is the base stock policy. When the inventory level drops below the base stock level, then it is optimal to order up to the base stock level. Since the cost is calculated based on the inventory level at the end of the protection period, the ordering decision can increase the inventory level to the base stock level. As the order decision takes place free of charge when the setup cost is zero, we do not need to consider the demand postponement. Thus, our discount policy does not make any difference. So we focus on cases with positive setup costs.

For the positive setup cost without discount decision, the optimal policy is a state dependent $(s(o), S(o))$ policy which is order up to $S(o)$ if inventory level is at $s(o)$ or below. Adding discount decisions turns the optimal policy to $(s(o,Q), S(o,Q), w^i(o,Q), W^i(o,Q))$ where $i \in \{L+1, .., N\}$ refers to the delay period. The order policy is similar to non discount case but the parameters are also dependent on Q.

The discount policy is to offer discount up to $W^i(o,Q)$ if the inventory level is greater than $w^i(o,Q)$ with a delay period i. The lower bound for $w^1(o,Q)$ is the reorder level while its upper bound is $w^2(o,Q)$. Due to the availability of demand for discount, the $W^i(o,Q)$ is limited by $w^i(o,Q) + Q$.

Figure 4.3 illustrates the optimal decisions based on the modified inventory positions.

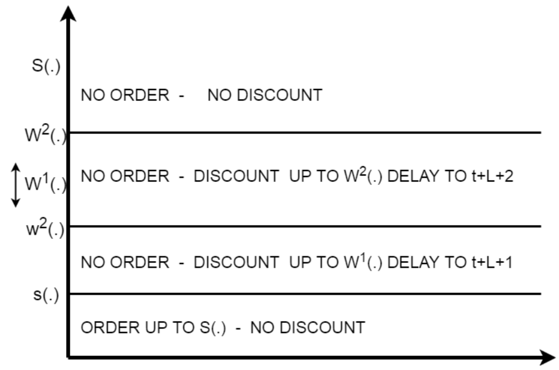

FIGURE 4.3: Order & Discount Policy

In the numerical experiments, we aim to show the value of ADI and effect of demand postponement in capacitated and uncapacitated inventory systems. Then we would like to compare these two policies to derive insight into the value of obtaining more ADI or using postponement policy. The performance of both policies are presented under different conditions.

4.7.1 Uncapacitated Inventory

4.7.1.1 The value of ADI

In this section, we test the value of ADI and policy parameters for an uncapacitated inventory system. Initially we ignore the demand postponement to see the status quo of ADI. We test the no discount model with different means of demand starting from no advance demand case ($\lambda = (6, 0, 0)$) to fully advance demand ($\lambda = (0, 0, 6)$). The percentage saving on cost is calculated by $\delta = 100(Cost_{noADI} - Cost_{\lambda})/Cost_{noADI}$. The results are presented in Table 4.2.

Increasing the ADI present in the system has a reduction on cost up to 6.79%. This value also give the idea to the manager on how much to invest for obtaining ADI from customers. As expected, the reorder level and order up to level is decreasing by more advance demand since the uncertainty on demand is getting disappeared. Although we use shortage cost relatively higher than the holding cost, one may use service level constraint which limits the number of backorders as well.

TABLE 4.2: $K = 100, h = 1, p = 9,\ D_{t-1,t+1} \in 0, ...10,\ T = 12$

λ		0	1	2	3	4	5	6	7	8	9	10	Avg.Cost	$\delta(\%)$
6,0,0	$S(.)$	35	36	37	38	39	40	41	42	43	44	45	289.085	0
	$s(.)$	2	1	1	1	1	1	1	1	1	0	0		
4,1,1	$S(.)$	32	33	34	35	36	37	38	39	40	41	42	286.896	0.76
	$s(.)$	0	0	0	-1	-1	-1	-1	-1	-1	-1	-1		
3,1,2	$S(.)$	30	31	32	33	34	35	36	37	38	39	40	283.335	1.99
	$s(.)$	-1	-1	-1	-1	-2	-2	-2	-2	-2	-2	-2		
2,1,3	$S(.)$	28	29	30	31	32	33	34	35	36	37	38	279.632	3.27
	$s(.)$	-2	-2	-2	-2	-2	-3	-3	-3	-3	-3	-3		
1,1,4	$S(.)$	26	27	28	29	30	31	32	33	34	35	36	275.789	4.60
	$s(.)$	-3	-3	-3	-3	-3	-3	-4	-4	-4	-4	-4		
0,0,6	$S(.)$	24	25	26	27	28	29	30	31	32	33	34	269.454	6.79
	$s(.)$	-4	-4	-4	-4	-4	-4	-4	-4	-5	-5	-5		

4.7.1.2 The effect of Postponement on Inventory Policy

In this section, we evaluate the impact of demand postponement on inventory policy and inventory cost. We test the model with $P = 0, 1$ to show how the policy is changing between the worst and the best case scenarios. In this section, we offer the same discount amount to all customers regardless of their postponement period. The costs of each case is compared with non discounted case and the percentage deviation is calculated. The results on Table 4.3 reveal the optimal ordering and discount policies when the customers accepting probability, $P = 0$. This means that we have a chance to use the advantage of price discount only at the current period. Further periods customers do not accept the discount. The results for $P = 1$ is illustrated on Table 4.4. This time all the customers accept the discount when offered. Therefore more improvement on costs can be seen and when compared with $P = 0$ case, the order up to level is also decreasing with reduction on reorder level.

Even with $P = 0$, there still exists a slight improvement on average inventory costs up to 1% based on the demand available to accept the postponement at current period.

TABLE 4.3: $K = 100, h = 1, p = 9, d = 1, P = 0, \lambda = 4, 1, 1,\ D_{t-1,t+1} \in 0, ...10,\ T = 12$

C		0	1	2	3	4	5	6	7	8	9	10	Avg. Cost	$\delta(\%)$
0	$S(.)$	32	33	34	35	36	37	38	39	40	41	42	286.896	0
	$s(.)$	0	0	0	-1	-1	-1	-1	-1	-1	-1	-1		
	$S(.)$	32	33	34	35	36	37	38	39	40	41	42		
	$s(.)$	-1	-1	-1	-1	-2	-2	-2	-2	-2	-2	-2		
1	$W_2(.)$	10	11	12	13	14	15	16	17	18	19	20	286.288	0.21
	$w_2(.)$	5	6	7	8	9	10	11	12	13	14	15		
	$W_1(.)$	5	6	6	6	6	6	6	6	6	6	6		
	$S(.)$	32	33	34	35	36	37	38	39	40	41	42		
	$s(.)$	-2	-2	-2	-2	-2	-3	-3	-3	-3	-3	-3		
2	$W_2(.)$	10	11	12	13	14	15	16	17	18	19	20	285.711	0.41
	$w_2(.)$	4	5	6	7	8	9	10	11	12	13	14		
	$W_1(.)$	5	6	6	6	6	6	6	6	6	6	6		
	$S(.)$	32	33	34	35	36	37	38	39	40	41	42		
	$s(.)$	-3	-3	-3	-3	-3	-3	-4	-4	-4	-4	-4		
3	$W_2(.)$	10	11	12	13	14	15	16	17	18	19	20	285.163	0.60
	$w_2(.)$	3	4	5	6	7	8	9	10	11	12	13		
	$W_1(.)$	5	6	6	6	6	6	6	6	6	6	6		
	$S(.)$	32	33	34	35	36	37	38	39	40	41	42		
	$s(.)$	-4	-4	-4	-4	-4	-4	-4	-5	-5	-5	-5		
4	$W_2(.)$	10	11	12	13	14	15	16	17	18	19	20	284.634	0.79
	$w_2(.)$	2	3	4	5	6	8	9	10	11	12	13		
	$W_1(.)$	5	6	6	6	6	6	6	6	6	6	6		
	$S(.)$	32	33	34	35	36	37	38	39	40	41	42		
	$s(.)$	-5	-5	-5	-5	-5	-5	-5	-5	-6	-6	-6		
5	$W_2(.)$	10	11	12	13	14	15	16	17	18	19	20	284.124	0.97
	$w_2(.)$	1	2	3	4	5	7	8	9	10	11	12		
	$W_1(.)$	5	6	6	6	6	6	6	6	6	6	6		

Order up to level, S is not changing on this case while the reorder level s is decreasing when more demand is available for discount. Similarly when we consider the $P = 1$, the reorder level is also decreasing but also order up to level is decreasing. This means that additional to ordering later, we also order less which reduces the holding cost as well. There has been observed a significant reduction on total expected cost which is around 17%. One may expect that the benefit will be higher when holding and setup costs are higher.

It is clear that more ADI increases the system performance as seen in Table 4.2. However, when we compare it with the results on Tables 4.3 and 4.4, better potential can be seen when using demand postponement. Our intuition supports this because ADI is available

TABLE 4.4: $K = 100, h = 1, p = 9, d = 1, P = 1, \lambda = 4, 1, 1,\ D_{t-1,t+1} \in 0, ...10,\ T = 12$

C		0	1	2	3	4	5	6	7	8	9	10	Avg. Cost	$\delta(\%)$
0	$S(.)$	24	25	26	27	28	29	30	31	32	33	34	239.296	16.59
	$s(.)$	1	0	0	0	0	0	0	0	0	0	0		
1	$S(.)$	24	25	26	27	28	29	30	31	32	33	34	238.876	16.74
	$s(.)$	0	0	-1	-1	-1	-1	-1	-1	-1	-1	-1		
	$W_2(.)$	6	7	7	8	9	10	11	12	13	14	15		
	$w_2(.)$	1	2	3	4	5	6	7	8	10	11	12		
	$W_1(.)$		2	3	4	5	6	6	6	6	6	6		
2	$S(.)$	24	25	26	27	28	29	30	31	32	33	34	238.468	16.88
	$s(.)$	-1	-1	-1	-2	-2	-2	-2	-2	-2	-2	-2		
	$W_2(.)$	6	7	7	8	9	10	11	12	13	14	15		
	$w_2(.)$	0	1	2	3	4	5	7	8	9	10	11		
	$W_1(.)$		2	3	4	5	6	6	6	6	6	6		
3	$S(.)$	24	25	26	27	28	29	30	31	32	33	34	238.081	17.02
	$s(.)$	-2	-2	-2	-2	-3	-3	-3	-3	-3	-3	-3		
	$W_2(.)$	6	7	7	8	9	10	11	12	13	14	15		
	$w_2(.)$	-1	0	1	2	3	4	6	7	8	9	10		
	$W_1(.)$		2	3	4	5	6	6	6	6	6	6		
4	$S(.)$	24	25	26	27	28	29	30	31	32	33	34	237.713	17.14
	$s(.)$	-3	-3	-3	-3	-3	-4	-4	-4	-4	-4	-4		
	$W_2(.)$	6	7	7	8	9	10	11	12	13	14	15		
	$w_2(.)$	-2	-1	0	1	2	3	5	6	7	8	9		
	$W_1(.)$		2	3	4	5	6	6	6	6	6	6		
5	$S(.)$	24	25	26	27	28	29	30	31	32	33	34	236.066	17.72
	$s(.)$	-4	-4	-4	-4	-4	-4	-5	-5	-5	-5	-5		
	$W_2(.)$	6	7	7	8	9	10	11	12	13	14	15		
	$w_2(.)$	-3	-2	-1	0	1	2	4	5	6	7	8		
	$W_1(.)$		2	3	4	5	6	6	6	6	6	6		

for all customers. On the other hand, with the discount policy, we only buy the ADI when we need it. Especially for higher setup costs, the system already needs to keep large stock most of the time. So, the uncertainty on demand does not effect the system much. But at the times when the stock is at critical levels, obtaining ADI at that time will be greatly beneficial. This can be achieved by our postponement policy.

4.7.2 Capacitated Inventory

In this section, we test ADI and postponement policy on capacitated inventory systems. The literature on ADI mentions that ADI increases the efficient use of capacity. Similar to this, we expect that our postponement policy is going to increase the efficient usage of capacity. However, there is a challenge on usage the ADI and postponement policy. We only offer postponement to protection period demand. However, increasing presence of ADI causes to have less demand available for discount. With the numerical experiments, we aim to solve this issue to decide under which circumstances ADI is better.

In Table 4.5, we report the optimal replenishment and discount policy for a capacitated case. There supposed to be a reduction on reorder level when O_t is higher. However, we observe an increase on reorder level when the O_t closes to 10. This would suggest that if higher demand is waiting for the next period, it is better to order earlier. Also when compared to the uncapacitated case, the discount usage starts at higher inventory positions. This indicates that postponement is more needed on capacitated case.

Özer and Wei (2004) show that a threshold policy which is to order full capacity if the inventory position drops to a threshold level is optimal for capacitated case. Our findings do not verify this policy when the discount policy is in use. Our intuition also supports the idea that if more demand is available for discount, then the order up to level will be lower than the capacity level.

TABLE 4.5: $K = 100, h = 1, p = 9, d = (1, 1.5), P = 0.5, \lambda = 4, 1, 1,\ D_{t-1,t+1} \in \{0, ...10\}, Q_t \in \{0, ..., 2\}, C = 20$

Q	Parameter	0	1	2	3	4	5	6	7	8	9	10
0	$S(.)$	17	18	19	20	20	20	20	20	20	20	20
	$s(.)$	1	1	1	0	0	0	-1	-2	-2	-2	-1
1	$S(.)$	17	18	19	20	20	20	20	20	20	20	20
	$s(.)$	0	0	0	0	-1	-1	-2	-3	-3	-2	-1
	$W_2(.)$	8	9	10	11	12	13	14	15	16	17	18
	$w_2(.)$	2	3	4	5	6	7	8	9	10	11	12
	$W_1(.)$	2	3	4	5	6	6	6	6	6	6	6
2	$S(.)$	17	18	19	20	20	20	20	20	20	20	20
	$s(.)$	-1	-1	-1	-1	-2	-2	-3	-4	-3	-2	-1
	$W_2(.)$	8	9	10	11	12	13	14	15	16	17	18
	$w_2(.)$	1	2	3	4	5	6	7	8	9	10	11
	$W_1(.)$	2	3	4	5	6	6	6	6	6	6	6

Next, we address the benefits of ADI and demand postponement and compare them for different cases. The computational results are presented in Table 4.6.

The cost is increasing with the capacity constraint as expected. But the surprising result exists when we use our demand postponement strategy. When $P = 1$, our policy results in relatively small costs compared to the uncapacitated case with no discount policy.

Table 4.6: Impact of ADI and Postponement, $T = 12, K = 100$, $h = 1, p = 9, d_{ti} = (1, 1.5)$

Demand	Capacity=15			Capacity=20			No Capacity
	$P = 0$	$P = 0.5$	$P = 1$	$P = 0$	$P = 0.5$	$P = 1$	$P = 0$
$6, 0, 0$	449.7	368.0	310.0	378.5	318.8	286.5	289.1
$4, 1, 1$	407.0	360.6	300.6	353.3	314.8	281.1	286.9
$3, 1, 2$	384.9	338.9	299.2	335.7	301.6	278.8	283.3
$2, 1, 3$	366.9	345.7	299.5	321.0	305.3	278.3	279.6
$1, 1, 4$	352.0	325.4	300.3	308.2	294.4	279.6	275.8
$0, 0, 6$		328.2			292.8		269.5

This clearly shows that if a company can convince customers to accept the postponement, then they can completely remove the disadvantage of the capacity constraint.

ADI is also reducing the negative impact of capacity. However, when the demand pattern reaches a certain ADI, more ADI might not bring more benefit when having a postponement policy in place. In Table 4.6, when $P = 0.5$ and $C = 20$, the demand pattern $(3, 1, 2)$ has better cost than $(2, 1, 3)$. Thus, we are now able to solve the contradiction between having more ADI or increasing P. In addtion, the relative benefits of postponing demand can further increase when considering that having more ADI would (when offered to all customers) incurs extra costs from incentives. This aspect has not been considered in our computational experiments.

4.7.2.1 Sensitivity Analysis on Discount Amount

In this section, we observe the system improvements under different discount values. Since the postponement could bring significant improvements to the system, then the customers need to be encouraged to accept it. In reality, it is expected that the customers willingness increases by more discount. We test the sensitivity of discount amount with different advance demand structures under capacitated case. Since there are two postponement periods available, we have two different discounts d_{t1} and d_{t2}. We assume that d_{t2} is always equal to $d_{t1} + 0.5$. The tests are taken with $P = 0.5$ and the results are illustrated in Figure 4.4.

Although in previous experiments we set discount to $(1, 1.5)$, the results show that it can be higher based on the advance demand in the system. The less advance demand results in more customers being offered postponement. If there is no advance demand which is $(6, 0, 0)$, the discount amount could reach up to 14. Starting from a larger ADI

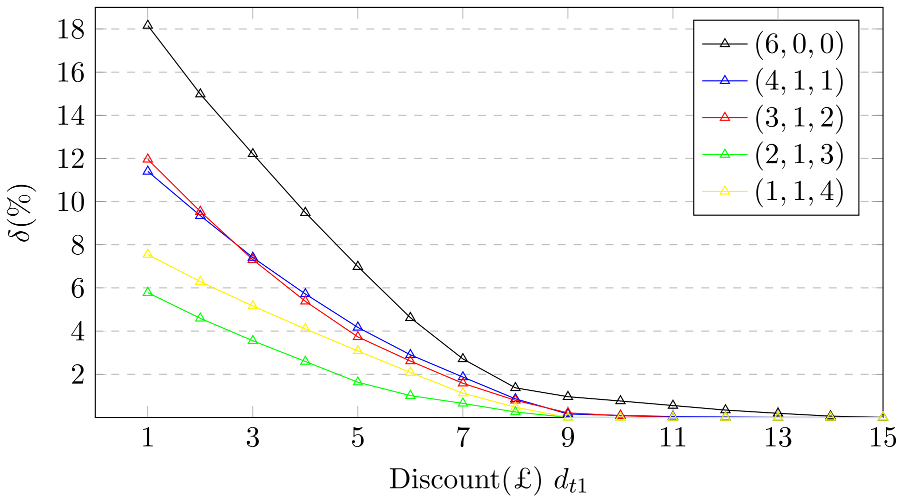

FIGURE 4.4: Sensitivity of Discount $d_{t1}, d_{t2} = d_{t1} + 0.5,\ K = 100, h = 1, p = 9, P = 0.5, C = 15$

in system, the usage and benefit from postponement reduces. When the company thinks of offering the discount, the balance between the cost of obtaining ADI and discount amount has to be carefully considered.

Next, we analyse the effect of different prices for different postponement periods. Therefore, we set d_{t1} to £4 first, and observe the changes on improvement by different values of d_{t2} with $P = 0.5, 1$. Next, the same process is repeated to analyse the effect of d_{t1}. The results are presented in Figure 4.5 and 4.6.

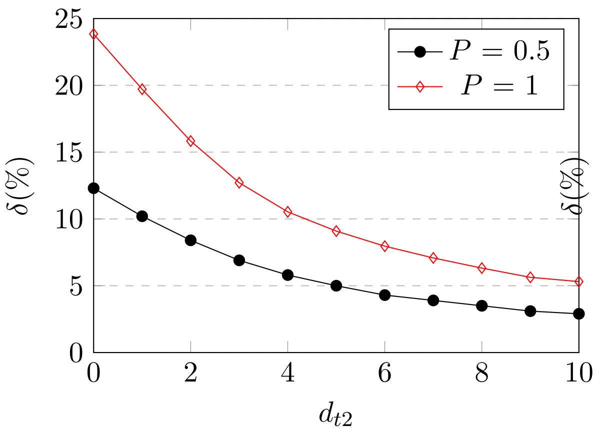

FIGURE 4.5: Analysis of d_{t2} when $d_{t1} = 4, \lambda = (3, 1, 2)$

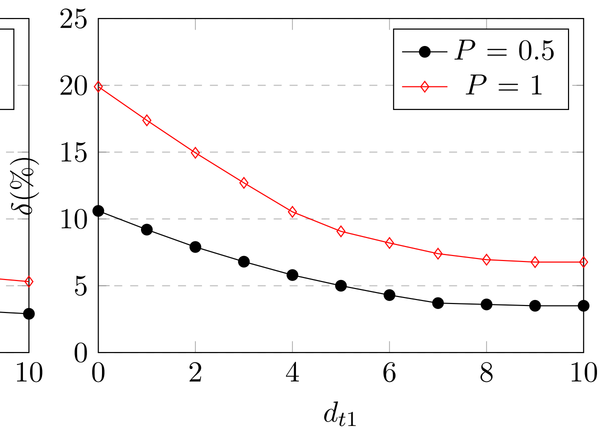

FIGURE 4.6: Analysis of d_{t1} when $d_{t2} = 4, \lambda = (3, 1, 2)$

When we are calculating the cost reduction percentage δ, we take the cost without postponement as base cost. For $P = 1$ case, the cost reduction is between $25 - 5\%$. When we fix the d_{t1}, then it becomes larger. This clearly shows that longer postponement periods lead to more reduction in costs. Thus, the customers needs to be convinced for longer postponement period. With the increasing discount amount, the difference between different probabilities of acceptance is getting lower.

4.8 Conclusions

In this paper, we establish the structure of optimal inventory and discount policies for an inventory model with advance demand information and price discounts. In return of discounts, customers are expected to wait longer for their demand.

For the zero setup cost case, a simple base stock policy is optimal and there is no need to use discounts. When the setup cost is positive, state dependent policies are optimal. The parameters of the policy depend on the observed demand beyond the protection period and the available demand for the postponement/discount decision. Numerical examples indicate that the system cost is reducing when more advance demand and demand for discounts are present. Increasing customers acceptance of discounts results in reducing reorder and order up to levels.

To our knowledge, this is the first study to develop and analyse a selective state-dependent strategy to purchasing buyer's willingness to postpone their delivery in a variety of ADI settings. This strategy seems, in particular, to be of value in capacitated inventory situations with a modest level of ADI. It is argued that this strategy can result in a more profitable system in comparison to a strategies that aim to introduce more ADI in the system.

The practical drawback from this approach is the computational complexity of the model and inventory/postponement policy. Further research may thus be devoted to faster heuristic methods, and examining the efficiency of policies with a simpler structure.

Chapter 5

Joint inventory and distribution strategy for online sales with a flexible delivery option

Abstract

This paper develops a strategy to jointly optimise the inventory and distribution for an online sales firm. The firm has to decide how to distribute the products from its warehouse to customers: this can either be done by using a company-owned vehicle, or by outsourcing to a third-party transportation company. The online sales environment includes a flexible delivery option that gives a discount to customers in return. This option is offered when the inventory level in the warehouse is lower than a threshold level. Customers accepting flexible delivery pay a deposit at the time they place the order and pay the remaining reduced price at the time of delivery. By offering the flexible delivery option, the firm aims to reduce the cost of distribution to the customers as well as postpone the timing of paying an outside supplier for stock replenishment. As the timing of cash-flows are dependent on the customer behaviour and the inventory and distribution strategy, the profit function is the Net Present Value of future cash-flows. We analyse the benefit of flexible delivery to the firm and perform sensitivity analysis with respect to various parameters. The profitability of flexible delivery depends on price setting and customer behaviour. Flexible delivery, in this model, has great potential to reduce transport distances and emissions when firms use their own vehicles.

5.1 Introduction

Online sales has been pioneered in the retail sector of products that can be transported as small parcels by postal or courier services. The strong competition this has generated has led to closures of many traditional stores in the high street. Online sales channels allow customers to save on the time and effort needed to select and order desirable products, but do introduce a time lag between order placement and delivery. Can this disadvantage turned into an advantage for both customers and online firm?

The difference between customers' willingness to wait is recognized by several online retailers by offering customers a choice between delivery options, where price discounts are given when customers accept increased (uncertainty on) delivery time. Delivery options range from fast, say within 24 or 48 hours, to delivery at any time within a given time period of e.g. 7 days. We refer to the latter as a *flexible delivery option*. If greater flexibility to plan the delivery allows the online retailer to make costs savings, providing a financial incentive to customers giving the firm this increased planning flexibility could thus make economic sense.

The use of online sales channels has become increasingly important for businesses who sell larger-sized products. In the UK, for example, there now exist a fair number of online sales companies offering products and materials for home improvement projects. This may range from kitchen and bathroom units and appliances to materials for plumbing, electrics, and joinery projects, servicing both professional tradesmen and DIY enthusiasts. The bulkier or heavier products require transport in specialized (company-owned) vehicles for the delivery to customer locations. The model developed and analysed in this paper is particularly aimed at achieving a better understanding of the impact of flexible delivery for these types of businesses.

The time lags between ordering and delivery will vary with the sector. In the online sales of books, delivery options are typically restricted to hours or days. In the home improvement sector, fast delivery may be within one week, and a flexible delivery may be any time within several weeks. Another difference is that customers ordering home improvement products may, instead of having to pay the full price when ordering, only be required to pay a deposit. The residual payment due is then settled on the day of delivery. Given that this payment may occur several weeks later, it is thus of interest to model the impact of the timing of payments on the profitability of the online business.

The costs that arise from meeting online sales naturally depend on how the supply chain is organized. In this paper, we assume that the products are stored at the online firm's single depot, from which 'last-mile' transportation needs to be organized to customer locations within a sales region surrounding the depot[1]. See also Figure 5.1. The online retailer can make use of its own vehicle to make deliveries, or alternatively outsource to a third party for a fixed price per product (or unit of volume or weight). Next to delivery costs, costs also arise from the inventory stored at the depot and order costs for the replenishment from an upstream supplier. In this paper, we aim to get more insight into how inventory and transportation costs can be minimised in this setting, and how this depends on other system characteristics such as the proportion of customers that can be induced to select flexible delivery.

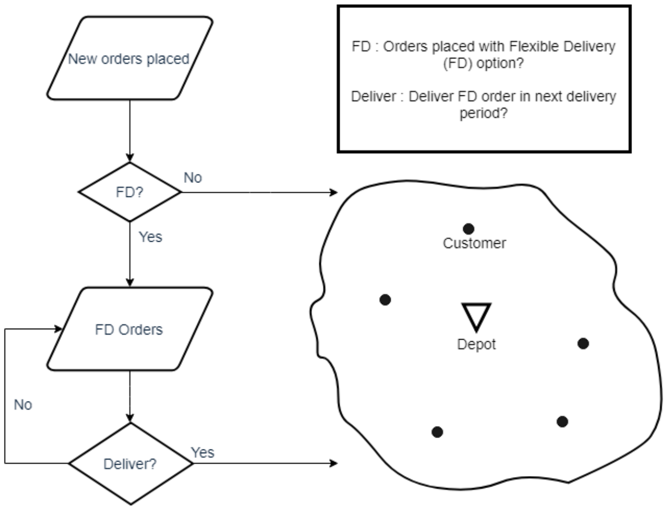

Figure 5.1: Process flow of order placement and delivery.

As the value and timing of both revenue and cost cash-flows will be affected by the choices made by customers and the firm, the minimisation of average costs only is not an appropriate objective. We therefore develop the model with the objective of maximising the Net Present Value (NPV) of the firm's future cash-flows. The usefulness of this approach for the study of production-inventory systems was first demonstrated in Grubbström (1980), see also Beullens and Janssens (2011).

[1]This situation can also represent the case of a firm operating one single online sales website, but adopting a decentralized organization of its supply chain into disjoint service regions, where each sales region is serviced by a single depot.

5.2 Related research

Considering the different ingredients discussed above, this study is of relevance to the research on supply chain inventory and distribution management, and in particular to the use of price incentives for shifting demand across time.

Pricing strategies are widely applied in supply chain literature for various purposes. In the area of revenue management, this approach has been studied in great detail, see e.g. Quante et al. (2009), Agatz et al. (2013). This area of research is primarily concerned with maximising the revenue extractable from a capacitated resource with often a fixed cost structure. This area of research finds application in booking of (airline, rail, ferry) travel, hotel rooms, and car rentals. Pricing strategies will affect both total demand as well as the timing of when demand is announced.

Many papers in the inventory literature consider pricing strategies as a means to increase the total demand or speed up the demand rate, rather than shift demand, as in the studies on price-dependent demand by Papachristos and Skouri (2003), Wee (2001), Dye et al. (2007) and quantity discounts by Shin and Benton (2007), Wee (1999), Weng (1993). Dynamic pricing in the presence of inventory considerations is another area of application extensively discussed in Elmaghraby and Keskinocak (2003).

Using financial incentives to cause demand to shift across time is an area of research that has been mainly analysed within the contexts of advance booking (or pre-orders) and demand postponement, respectively. Advance booking programs entice customers to place their demand earlier by offering them (financial) incentives, see e.g. Tang et al. (2004) and Li and Zhang (2013). According to Xu et al. (2017), the success of online retail channels has made it easier for companies to obtain such advance bookings. The benefits of using advance booking systems include that it could greatly improve the demand forecast for new products, as in Tang et al. (2004), or could lead to a better match between demand and capacity availability, as in Zhuang et al. (2017).

The purpose of demand postponement is to convince customers to shift delivery to a later date. The concept was first introduced in Iyer et al. (2003), who studied the value of postponement of a fraction of demand as to allow the firm to procure additional capacity and reduce overall stockouts. Wu and Wu (2015) use demand postponement to create a capacity buffer for urgent demand, leading to a reduced expected risk of shortages on urgent demand in addition to an overall increased effectiveness of capacity usage. Tang (2006b) use the strategy as a means to mitigate supply chain disruptions.

The effect of customers choosing for the flexible delivery option in our model is comparable to demand postponement by enticing customers to switch from the fast delivery option towards a delivery at a later date. This study, however, differs from the previous literature as decisions need to be made as to when to deliver to which customer based on the joint consideration of inventory and distribution costs.

Our model and analysis is also different from previous literature on demand postponement in that it considers the impact of the logistics strategy on the timing of payments in an NPV framework. Permissible delay in payment when a certain amount is ordered (Chung et al. 2005), selling price- dependent demand (Dye et al. 2007) and backorders with a deposit (Ghiami and Beullens 2016) are examples in which payment structures are shown to affect inventory decisions.

The mechanism of price discounts in combination with deposit values we adopt in this study was first introduced in Ghiami and Beullens (2016), who used this mechanism in a production-inventory system with backorders. The impact of deposit value in that situation was limited. In our inventory-distribution model with the flexible delivery option, we find that its impact is more pronounced.

The paper is further organized as follows. Section 5.3 introduces the description of the system. In Section 5.4, the optimisation problem is formulated. Section 5.5 develops properties and presents the algorithm. Section 5.6 provides insights derived from a variety of numerical experiments. Conclusions and further research areas are presented in Section 5.7.

5.3 Description of the System

5.3.1 General characteristics

We consider a single product ordered by customers via an online store at a given constant demand rate y. Upon placing their order, customers may be given a choice between the fast (or normal) delivery option and a flexible delivery option. In the fast delivery option, the customer pays the full sales price p when placing the order. In the flexible delivery option, the customer receives a small discount r on the full sales price, but pays upon placing the order a pre-agreed deposit g, while paying the remainder $p - r - g$ upon the time of actual delivery. We note that in most online systems we would have $g = p - r$. However, in applications were customers' financial trustworthiness can be verified, a reduced deposit value $g < p - r$ may help in convincing more customers accepting the flexible delivery option, in particular in cases where the possible waiting time may become larger. Furthermore, we assume that more customers will choose the flexible delivery option the larger the discount value r. Part of the aims of the study is to investigate how system performance is affected under various different choices for r and g values.

The performance of the system is further affected by decisions about how to place replenishment orders from an upstream supplier. In line with EOQ-type model characteristics, we consider that in steady-state a constant order quantity yT will be placed at equidistant moments in time, where T denotes the inventory cycle time measuring the time between two consecutive orders arriving in the retailer's warehouse. In addition, system performance is further determined by the choices made about how the demand orders are delivered to customer locations. The choice here is between performing the distribution by a company-owned vehicle, and the option of outsourcing the delivery to a third party. The company can also choose to outsource the delivery of only part of the demand, and make this variable over time. The model constructed will maximise the Net Present Value of this system by choosing the optimal joint replenishment and distribution strategy.

Parameters

y	Constant demand rate per unit time $y > 0$
p	Sales price per unit of the product $p > w$
w	Cost price of a unit product
r	Discount amount per unit of product, $0 \leq r < p - w$
g	Deposit paid per unit of product, $g < p - r$
S	Setup cost per replenishment
$I(t)$	Inventory level at time t
U	Maximum waiting time promised to customers
Δt	Time of a delivery period
A	Size of the total distribution region
ρ	Average distance from a customer in distribution region to the depot
$n(k)$	Number of customers need to be served in the k^{th} delivery period
$d(k)$	Average distance travelled in the k^{th} delivery period
$c(k)$	Total distribution cost of the k^{th} delivery period
$c_u(k)$	Distribution cost per unit product in the k^{th} delivery period when delivered by the company-owned vehicle
c_o	Transport cost per unit product charged by third-party
β	Ratio of demand accepting the flexible delivery option, $0 < \beta < 1$
α	Opportunity cost of capital rate
ζ	Euclidean metric factor
γ	Vehicle operation cost per mile

Decision Variables

T	Inventory cycle time, $T = K\Delta t$ with K integer, $\Delta t << T$, $T > 0$
FT	Time period during which flexible delivery option is not available, $0 \leq F \leq 1$
$J(k)$	Binary variable to decide whether to use own vehicle (1) or not (0)

5.3.2 Inventory system

An important feature not yet discussed is that the flexible delivery option is not necessarily always made available to customers. That is, the flexible delivery option is only made available within a time period $(1-F)T$ of a replenishment inventory cycle. The reason for this is to allow the replenishment cycle T, should this be optimal, to become (much) larger than the maximum time U that customers may have to wait when choosing the flexible delivery option. To minimise waiting time for customers, it is best to have the flexible delivery option made available at the end of each inventory cycle[2].

Figure 5.2 illustrates the inventory level as a function of time for the first inventory cycle. It is assumed that the system starts at time 0 with no outstanding demand orders. At time FT, the flexible delivery is introduced for a time $(1-F)T$. During this period, it is assumed that customers choose the flexible option at a rate of βy (further discussed in Section 5.3.4). Delivery of these orders will be postponed until after the next replenishment order has arrived at time T. Note that the first replenishment order only needs to cover the demand during the first cycle T for those demands to be delivered with the normal delivery option. From the second cycle and onwards, the replenishment order also needs to cover the postponed demand, but is otherwise identical to the first cycle. The system then repeats itself at infinitum.

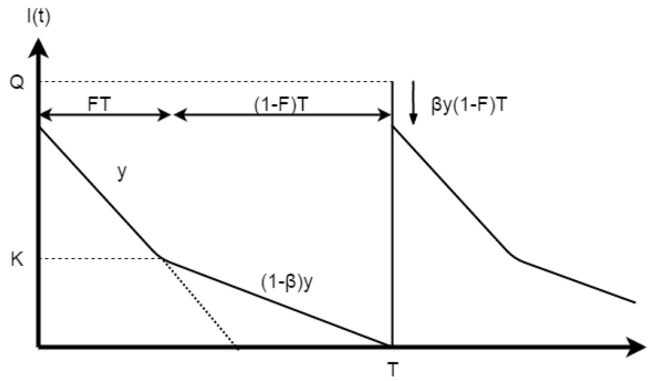

FIGURE 5.2: Inventory Flow with Constant β

5.3.3 Delivery system

Distribution to customer locations is organized at discrete moments in time (Δt, $2\Delta t$ ..., $k\Delta t$, ...). Let $n(k)$ denote the number of customers to be served in delivery period

[2]This will also maximise the NPV of revenues and holding costs. The NPV of distribution costs alone (see Figure 5.3), however, would be maximised by starting each cycle with the flexible delivery time window instead. Distribution costs are typically small, however, so overall the adopted model in this paper will be financially better. For high values $g = p - r \approx p$, it may be better to have the first inventory cycle different, i.e. being of length $T_1 \neq T$ in which the flexible delivery option is offered throughout. The second and all future inventory cycles would then be of length T and follow the structure as in Figure 5.3. We do not pursue this possibility in this paper.

k and $n(1)$ is the number of customers at the first delivery. We then have (with mod indicating the modulo operator):

$$n(1) = \begin{cases} y\Delta t & \Delta t \leq FT \\ y\Delta t - y\beta(\Delta t - FT) & FT < \Delta t < FT + \Delta t \\ (1-\beta)y\Delta t & FT + \Delta t \leq \Delta t \leq T \end{cases} \tag{5.1}$$

$$n(k) = \begin{cases} y\Delta t & \Delta t < k\Delta t \mod T \leq FT \\ y\Delta t - y\beta(k\Delta t \mod T - FT) & FT < k\Delta t \mod T < FT + \Delta t \\ (1-\beta)y\Delta t & FT + \Delta t \leq k\Delta t \mod T \leq T \\ n(1) + (1-F)Ty\beta & k\Delta t \mod T = \Delta t \text{ and } k > 1 \end{cases} \tag{5.2}$$

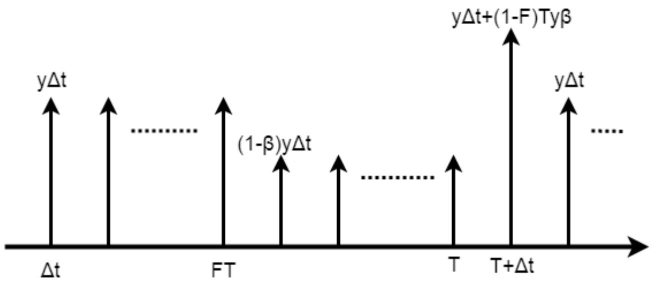

FIGURE 5.3: $n(k)$ values within an inventory cycle

Fig. 5.3 illustrates the resulting distribution pattern. When the delivery occurs at a time when we do not offer the flexible delivery option, the number of demands to be delivered equals $y\Delta t$. If the delivery occurs in the time interval during which flexible delivery is offered, the number of deliveries is $(1-\beta)y\Delta t$. When it is in the interval of $(FT, FT+\Delta t)$, then both types of demand are to be distributed. Also, at the first delivery moment after any replenishment order has arrived (except for the first cycle shown by $n(1)$), the delivery of postponed demand from the previous cycle has to occur. Therefore the initial delivery after the order arrival serves usual customers plus the postponed customers.

For the delivery of $n(k)$, the company can choose between using a company-owned vehicle, or outsourcing to a third party. When the company makes its own delivery, the distribution cost includes fuel, driver wages, vehicle depreciation and (un)loading costs. Among these, we focus on the costs directly related to the travelled distance. The average distance travelled in the k^{th} delivery is approximated in (5.3). This model is derived from the continuous approximation method developed in Daganzo and Newell (1985), where $\zeta(n(k)) \approx 0.57$ and γ is the vehicle operation cost per mile including fuel costs and driver's wages (at average vehicle speed). The main assumption underlying (5.3) is that the $n(k)$ customer locations are independently drawn from a uniform distribution over the service region of area A. The first term measures twice the average distance ρ between the warehouse and a customer location. The second term measures the shortest distance to visit all $n(k)$ customer locations. The corresponding unit distribution cost is $c_u(k)$ as in (5.4).

$$d(k) \approx 2\rho + \zeta(n(k))\sqrt{An(k)} \tag{5.3}$$

$$c_u(k) \approx \frac{\gamma d(k)}{n(k)} \tag{5.4}$$

If the warehouse is located inside the service region, Danganzo's formula can be simplified to $d(k) \approx \zeta\sqrt{A(n(k)+1)}$, where $\zeta \approx 0.71$. Without much loss of generality, we will adopt this setting here.

An external distribution partner is offering a cost of c_o per unit item for delivery. The unit delivery cost of the company's own vehicle depends on the demand intensity across the service area. If the unit cost is less than the outsourcing cost then the company delivers the product itself. The decision variable $J(k)$ is thus as given by (5.5), and the total distribution cost of the k^{th} delivery as in (5.6).

$$J(k) = \begin{cases} 1 & c_u(k) < c_o \\ 0 & otherwise \end{cases} \tag{5.5}$$

$$c(k) = n(k)\Big[J(k)c_u(k) + (1 - J(k))c_o\Big] \tag{5.6}$$

5.3.4 Customer's acceptance of flexible delivery

It is in line of expectation that more customers will be induced to accept the flexible delivery option the higher the discount and the lower the maximum waiting time and deposit value. We will examine the case that the fraction β is influenced by the discount amount r, maximum waiting time U, and deposit value g according to the following relationship:

$$\beta(r, U, g) = e^{-\frac{\theta(U, \frac{g}{p-r})}{r}}, \tag{5.7}$$

where $\theta(U, g/(p-r)$ is a factor that measures the *customer resistance* to the discount offer, which increases with U and $g/(p-r)$. According to (5.7), β grows faster and faster towards 1 with the increase in r, and approaches 1 in the limit for large r. Note, however, that r will always remain finite as $r < p - g$. For given U, g and r values, $\beta(r, U, g) = \beta$ is thus assumed to be a constant.

5.4 Model Development

As the value and the timing of revenues and cost cash-flows are affected by the choices made by customers and the firm, we develop the model with the objective to maximise the Net Present Value of the firm's future cash-flows. We summarize the main modelling assumptions:

- Demand occurs at a constant rate and individual customer orders are unit sized.
- Demands which accept the discounts are postponed to the first delivery after the replenishment order from the supplier is received.
- Customers who choose fast delivery pay the full price upon placing the order online.
- Customers who accept flexible delivery pay a deposit when placing the order, but pay the discounted remainder at the delivery time epoch in which the delivery is made.
- Initial inventory is zero and first replenishment occurs at $t = 0$.

- There is no constraint on the replenishment order size and no quantity discounts are offered by the external supplier.
- The replenishment lead-time is zero (or any non-zero constant value).
- The supplier is paid upon the arrival of the replenishment order in the firm's depot.
- Out-of-pocket holding cost are assumed small and will be ignored but we consider the opportunity cost of capital invested in stock.
- The retailer depot is located within the customer service region.
- When the third-party transportation company is used to delivery in a period, the cost incurred is paid out at the end of that delivery period.
- When using the company-owned vehicle for transport in a delivery period, costs are incurred at the end of the delivery period.
- Continuous compounding is used for the net present value analysis.
- The time horizon is infinite.

5.4.1 Inventory level and order quantity

For $\Delta t << T$, the inventory level $I(t)$ as a function of time over one inventory cycle follows with sufficient accuracy the pattern as in Figure 5.2. The change of inventory level over any cycle can thus be described by the following differential equation:

$$\frac{dI(t)}{d(t)} = \begin{cases} -y & , 0 \leq t \leq FT \\ -(1-\beta)y & , FT \leq t \leq T \end{cases} \tag{5.8}$$

the solution of which is given by:

$$I(t) = \begin{cases} I(0) - yt & , 0 \leq t \leq FT \\ I(FT) - (1-\beta)y(t-FT) & , FT \leq t \leq T \end{cases} \tag{5.9}$$

Since backorders are not allowed, we must have $I(t) \geq 0$. As it is unnecessary to keep stock that is not used, $I(T) = 0$ for each inventory cycle. In the first cycle, the order quantity $Q_1 = I_1(0)$ needs to cover the demand that needs to be fulfilled in that cycle:

$$Q_1 = yFT + (1-\beta)y(1-F)T. \tag{5.10}$$

The first term is the demand in the period $(0, FT)$, and the second term is the demand in period (FT, T). In all subsequent cycles, the order quantity $Q = I(0)$ covers the amount as in (5.10) plus the postponed demand from the previous cycle $\beta y(1-F)T$, which gives:

$$Q = Q_1 + \beta y(1-F)T = yT. \tag{5.11}$$

5.4.2 Annuity streams of cash-flows

Since the time horizon is infinite, the appropriate objective function is given by the Annuity Stream (AS) of the relevent incoming and outgoing cash-flows. The relevant cash-flows for the revenues and inventory costs is shown in Fig.5.4. (The cash-flows of the distribution system are not shown in this figure.)

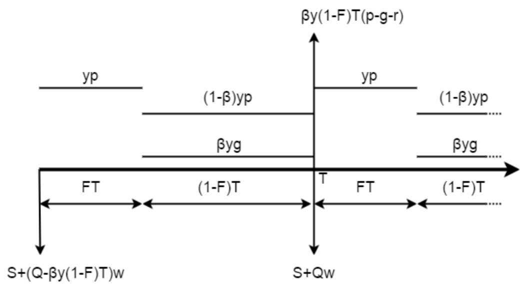

FIGURE 5.4: Cash-flows of revenue streams and inventory system costs

At the start of each inventory cycle, the firm incurs a set-up cost S independent of order size. In all but the first order cycle, the firm has to pay $wQ = wyT$ for the products replenished, see (5.11). The total replenishment cost is then $C = S + wyT$. For the first order cycle, since $Q_1 = Q - \beta y(1-F)T$, the total replenishment cost can be written as:

$$C_1 = S + w(Q - \beta y(1-F)T) = C - C_0^-. \tag{5.12}$$

The annuity stream of replenishment costs are thus given by:

$$\begin{aligned} ASC &= \sum_{i=0}^{\infty}(\alpha Ce^{-i\alpha T}) - \alpha C_0^- = \alpha C\frac{1}{1-e^{-\alpha T}} - \alpha C_0^- \\ &\approx C\left(\frac{1}{T}+\frac{\alpha}{2}\right) - \alpha C_0^- \end{aligned} \tag{5.13}$$

The approximated value on the right-hand-side is derived from the Maclaurin expansion of the exponential terms in αT, where second and higher orders of α are ignored.

The postponed demand from previous periods is satisfied by on hand inventory and these customers pay the discounted and deposit-subtracted price as given in (5.14). Not having postponed demand at $t = 0$, this revenue stream starts at $t = T$, thus the annuity stream value is as given in (5.15).

$$R_1 = \beta y(1-F)T(p-g-r) \tag{5.14}$$

$$\begin{aligned} ASR_1 &= e^{-\alpha T}\left(\sum_{i=0}^{\infty}\alpha R_1 e^{-i\alpha T}\right) = \alpha R_1\frac{e^{-\alpha T}}{1-e^{-\alpha T}} \\ &\approx R_1\left(\frac{1}{T}-\frac{\alpha}{2}\right) \end{aligned} \tag{5.15}$$

Between 0 and FT in each inventory cycle, sales occur at demand rate y for a price of p. The present value of this revenue for the first cycle is given in (5.16), and the equivalent annuity stream (from all cycles) in (5.17).

$$R_2 = \int_0^{FT} pye^{-\alpha t}dt = \frac{py}{\alpha}\left(1-e^{-\alpha FT}\right) \tag{5.16}$$

$$\begin{aligned} ASR_2 &= \sum_{i=0}^{\infty}\alpha R_2 e^{-i\alpha T} = py\left(\frac{1-e^{-\alpha TF}}{1-e^{-\alpha T}}\right) \\ &\approx py\left(F - \frac{\alpha TF^2}{2} + \frac{\alpha TF}{2}\right) \end{aligned} \tag{5.17}$$

At time FT, the flexible delivery option becomes available. The demand rate requiring the fast delivery reduces to $(1-\beta)y$ and generates a revenue stream at the price p. The demand rate βy corresponding to the flexible delivery customers generates a deposit g

in this interval. Present value of this revenue between FT and T is given in (5.18), and corresponding annuity stream function in (5.19).

$$\begin{aligned} R_3 &= \int_{FT}^{T} ((1-\beta)yp + \beta yg)e^{-\alpha t}dt \\ &= \frac{((1-\beta)yp + \beta yg)}{\alpha}\left(e^{-\alpha FT} - e^{-\alpha T}\right) \end{aligned} \tag{5.18}$$

$$\begin{aligned} ASR_3 = \sum_{i=0}^{\infty} \alpha R_3 e^{-i\alpha T} &= ((1-\beta)yp + \beta yg)\left(\frac{e^{-\alpha TF} - e^{-\alpha T}}{1 - e^{-\alpha T}}\right) \\ &\approx ((1-\beta)yp + \beta yg)\left(1 - F + \frac{\alpha TF^2}{2} - \frac{\alpha TF}{2}\right) \end{aligned} \tag{5.19}$$

The distribution cost over time depends on the number of customers served at each delivery $n(k)$, and on the choice of distribution strategy, producing different possible values for $c(k)$ as given by (5.6). The structure of the cash-flows of the distribution system, are given by the reflection about the y-axis of Figure 5.3, and replacing the $n(k)$ by $c(k)$ values. The present value of the distribution cost in every cycle but the first is thus given by (5.20). In the first cycle, we do not have postponed demand from previous periods. So instead of delivering a postponed demand, we only need to consider $c(1)$. This correction is used in the annuity stream function as given by (5.21).

$$C_{dist} = c(\frac{T}{\Delta t} + 1)e^{-\Delta t} + \sum_{k=2}^{T/\Delta t} c(k)e^{-k\Delta t} \tag{5.20}$$

$$\begin{aligned} ASC_{dist} &= \sum_{i=0}^{\infty} \alpha C_{dist} e^{-i\alpha T} - \alpha\left(c(\frac{T}{\Delta t} + 1) - c(1)\right)e^{-\Delta t} \\ &\approx C_{dist}\left(\frac{1}{T} + \frac{\alpha}{2}\right) - \alpha c(\frac{T}{\Delta t} + 1)e^{-\Delta t} + \alpha c(1)e^{-\Delta t} \end{aligned} \tag{5.21}$$

The annuity stream of total profit over an infinite time horizon is the sum of all revenues minus the setup, unit and distribution costs.

$$ASTP(F,T) = ASR_1 + ASR_2 + ASR_3 - ASC - ASC_{dist}$$

Our objective is to maximise the annuity stream of the total profit while keeping maximum waiting time under a certain limit.

$$\begin{aligned} \text{maximise} \quad & ASTP(F,T) \\ \text{subject to} \quad & (1-F)T \leq U \end{aligned}$$

5.5 Properties and algorithm

To facilitate further analysis as well as algorithm design, we present some properties related to the impact of flexible delivery on the replenishment pattern and distribution strategy, respectively.

5.5.1 Special Case: Inventory profit only

In this section, we only consider the inventory system. This situation may arise in the case where the firm has outsourced the delivery to a third party for an agreed total price based on an agreed annual demand volume, and paid out according to a fixed payment structure that is independent of the actual delivery schedules. Flexible delivery will then only affect the replenishment strategy and we can focus on only the inventory system related profits arising from the cash-flows depicted in Figure 5.4. The relevant profit function is given by $ASTP$ as above but where ASC_{dist} is a constant and can thus be ignored.

Lemma 5.1. *The objective function $ASTP(F,T)$ is a concave function under positive setup cost.*

Based on Lemma 5.1, we conclude that the F and T values which make the partial derivatives zero are optimal.

Proposition 1. Let the $T_{F\leq 1}$ and $T_{F=1}$ be the optimal T values in the NPV model where the flexible delivery option is offered and not offered, respectively. Then:

$$T_{F\leq 1} \geq T_{F=1}$$

Proposition 1 shows that the cycle time for the model without a flexible delivery being offered is a lower bound for the model with the flexible delivery option.

Lemma 5.2. *For the flexible delivery option to be profitable, the discount amount should satisfy the condition:*

$$r \leq \frac{\sqrt{\frac{8S\alpha w}{y}}}{2 - \sqrt{\frac{2S\alpha}{yw}}} \tag{5.22}$$

Lemma 5.2 calculates the value of r which makes $F = 1$. This indicates that when the discount amount reaches this level, it is no longer profitable to offer the flexible delivery option.

Note that the above results are only valid when the distribution cost is ignored.

5.5.2 Determining the distribution strategy

In principle, the distribution strategy needs to decide on the optimal choice of $c(k)$ for $k = 1, 2, 3, \ldots$. For any values of T and F, however, the sum of distribution costs as given by (5.20) within an inventory cycle is of a structure in which we recognise up to four different demand intensity situations, see (5.2) and Figure 5.3. The unit distribution cost for the company when it would execute the delivery by itself is decreasing by the demand intensity on the service area. We can see this from (5.4): the numerator increases with the square root of $n(k)$, while the denominator increases by $n(k)$).

Let c_u^1 and c_u^2 be the minimum and maximum value of the unit distribution cost when using the companies' own vehicle, for any values of T and F. Given the above, c_u^1 is thus the unit distribution cost that would arise from servicing the highest possible demand intensity. As seen from Figure 5.3, the highest demand intensity would arise in the period just after replenishments:

$$y\Delta t + (1-F)Ty\beta \leq y\Delta t + Uy\beta, \tag{5.23}$$

where the upper bound follows from constraint $(1-F)T \leq U$. Likewise, c_u^2 would arise in the period when flexible delivery is offered and when the demand level is:

$$(1-\beta)y\Delta t. \tag{5.24}$$

Lemma 5.3. *For any values of T and F:*

1. *Distribution is fully outsourced if*

$$c_o \leq c_u^1 = \gamma\zeta\sqrt{\frac{A}{y\Delta t + Uy\beta}} \tag{5.25}$$

2. *Distribution is always conducted by the company if*

$$c_o \geq c_u^2 = \gamma\zeta\sqrt{\frac{A}{(1-\beta)y\Delta t}} \tag{5.26}$$

3. *Distribution is partially outsourced if*

$$c_u^1 \leq c_o \leq c_u^2, \tag{5.27}$$

where $J(k)$ is decided according to (5.5) for every k in an inventory cycle with given F and T values.

Lemma 3 will be used in the solution algorithm.

5.5.3 General Case: Solution Algorithm

The decision variables are T and F, and $J(k)$ as given by (5.5). The main Algorithm 3 performs an exhaustive search over T and F, and calls Algorithm 2 which decides on the optimal distribution strategy $J(k)$ over an inventory cycle of length T and given F and such that it maximises the objective function ASTP.

The algorithm assumes that T is restricted to an integer multiple of Δt. If we have the fast delivery distribution every x days, then $\Delta t = x/365$, and a maximum waiting time of three weeks corresponds to $U = 21/365$, for example. We restrict the search for an optimal T value to a maximum of $1/\Delta t$, or one year, while search for optimal F

values over 100 possible fractions of T. These settings can be easily adjusted if a more refined search is desired. However, the algorithm implementation in C++ on a normal PC solves all numerical examples we investigated at a sufficient level of accuracy while keeping computational time at less than 1 second for each instance.

If the model only needs to consider the inventory system, then:

$$T_{EOQ} = \lfloor \frac{T_{F=1}}{\Delta t} \rfloor,$$

due to Proposition 1, and the call to $ATSP(F,T)$ in Algorithm 3 simply needs to evaluate this function. In case the inventory-distribution system is to be optimised, $T_{EOQ} = 1$ and Algorithm 2 is used to evaluate $ATSP(F,T)$.

Algorithm 2: Calculation of Profit Function

return $ASTP(F,T)$
if $c_o \le c_u^1$ **then**
  $J(k) = 0 \quad \forall k \in [1, \frac{1}{\Delta t}]$
else if $c_o \ge c_u^2$ **then**
  $J(k) = 1 \quad \forall k \in [1, \frac{1}{\Delta t}]$
else
  for $k = 1$ **to** $\frac{1}{\Delta t}$ **do**
    if $c_o \le c_u(k)$ **then**
      $J(k) = 0$
    else
      $J(k) = 1$
    end if
  end for
end if
Calculate $ASTP(F,T)$ with given F, T and $J(k)$ values

Algorithm 3: Exhaustive Search Algorithm for NPV Model

---

return $T^*, F^*, maxprofit$

initialize; $maxprofit, T_{EOQ}, U$

for $k \geq \frac{T_{EOQ}}{\Delta t}$ **to** $\frac{1}{\Delta t}$ **do**

  for $j = 0$ **to** 100 **do**

    $T = k\Delta t$

    $F = 0.01j$

    if $(1-F)T \leq U$ **then**

      Call $ASTP(F,T)$

      if $ASTP(F,T) \geq maxprofit$ **then**

        $T^* = T$

        $F^* = F$

        $maxprofit = ASTP(F,T)$

      end if

    end if

  end for

end for

5.6 Numerical Experiments

We examine the impact of the flexible delivery option and choice of distribution strategy on the performance of the inventory-distribution system for a set of instances. The numerical values of the parameters are similar to the study of Ghiami and Beullens (2016). However, to clearly show the impact of delivery, we increase the demand rate to 1000 instead of 100 which is used by Ghiami and Beullens (2016). Thus, parameters are set at the following value ranges: $y = 1,000$ per year; $S = 50$ or 80; $p = 2w, 1.5w$, or $1.3w$; $\alpha = 0.2$ or 0.1 (per year); $\Delta t = 4/365$ (fast delivery within 4 days); $U = 3$ or 6 weeks (maximum waiting time for flexible delivery option). Other parameters are set as further specified.

First, we examine the special case of only considering inventory profits. We then consider the joint inventory and distribution system in which we optimally decide when to outsource the distribution. We end with the situation in which outsourcing is not available.

5.6.1 Special case: Inventory profits only

We consider here the special case introduced in Section 5.5.1. We compare the performance of a system offering flexible delivery ($F \leq 1$) to a system in which this option is not available ($F = 1$). In each case, we use the model to determine the optimal inventory system strategy that maximises the AS profits $ASTP$ for the firm. The percentage difference is calculated as:

$$\delta = 100(ASTP_{F\leq 1} - ASTP_{F=1})/ASTP_{F=1} \tag{5.28}$$

The higher value of δ, the higher the benefit obtained from using the flexible delivery option. The difference is tested with various customers reactions and setup costs. Although we consider $r = 1\%p$, its maximum value r_{max} beyond which the flexible delivery option cannot be profitable is also calculated (Lemma 2). A summary of results is given in Table 5.1.

TABLE 5.1: Impact of flexible delivery ($p = 1.3w$, $w = 10$,$r = 0.01p$, $g = 0$)

	$F = 1$		$F \leq 1$ (U=3 weeks)			$F \leq 1$ (U=6 weeks)			
Parameters	T	β	T	F	$\delta(\%)$	T	F	$\delta(\%)$	$r_{max}(\%p)$
$\alpha = 0.1$, $S = 50$	0.32	0.1	0.32	0.82	0.10	0.32	0.63	0.16	
		0.5	0.33	0.82	0.54	0.35	0.66	0.86	2.47
		0.9	0.35	0.83	0.99	0.40	0.70	1.63	
$\alpha = 0.1$, $S = 80$	0.40	0.1	0.41	0.86	0.13	0.40	0.70	0.23	
		0.5	0.42	0.86	0.67	0.44	0.73	1.16	3.13
		0.9	0.42	0.86	1.22	0.46	0.74	2.15	
$\alpha = 0.2$, $S = 50$	0.22	0.1	0.23	0.74	0.26	0.23	0.48	0.36	
		0.5	0.23	0.74	1.30	0.26	0.54	1.95	3.51
		0.9	0.26	0.77	2.37	0.29	0.59	3.74	
$\alpha = 0.2$, $S = 80$	0.28	0.1	0.28	0.79	0.31	0.29	0.59	0.50	
		0.5	0.30	0.80	1.60	0.31	0.62	2.57	4.47
		0.9	0.30	0.80	2.90	0.33	0.64	4.80	

As can be observed, the customers' acceptance rate β greatly affects the benefits of flexible delivery. Although customers do no pay deposits ($g = 0$) and thus delay paying until delivery takes place, increased profits still result from adopting a larger maximum delivery window U. Profits increase at higher values of the opportunity cost of capital α. With higher set-up costs, cycle times T increase but the waiting time constraint U will limit the period in which the flexible delivery is offered.

Fig. 5.5 and Fig. 5.6 show the impact on the relative performance of the flexible delivery option system at different marginal profit values, when $\alpha = 0.2$, $S = 80$, $w = 10$, $g = 0$, and $r = 0.01p$. The results clearly show that the flexible delivery option is more effective for low profit margins and is then also more sensitive to the maximum waiting time.

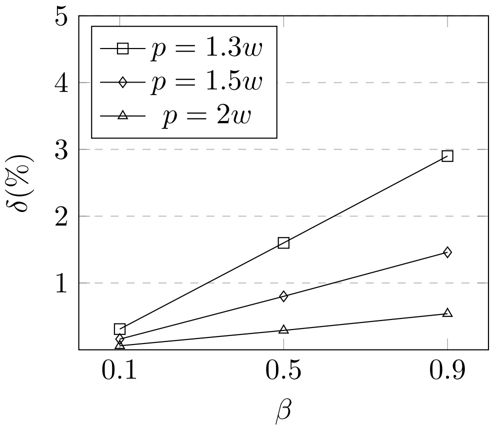

FIGURE 5.5: U=3 weeks

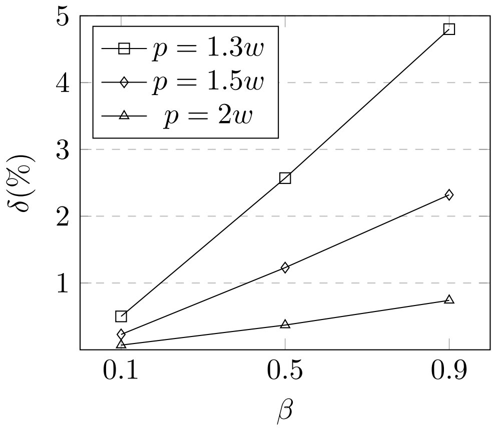

FIGURE 5.6: U=6 weeks

Another financial setting that can be examined in an NPV model is the impact of the deposit amount g. In many online sales environments, it is reasonable to assume that customers pay the full discounted price at the time of placing the order, which corresponds in our model with a deposit value $g = p - r$. It is intuitively clear that profits will improve with higher deposit values and in particular for larger U values. Table 5.2 presents results which show the impact of deposit value on system performance that confirm this.

TABLE 5.2: Impact of deposit value ($S = 80$, $\alpha = 0.2$, $p = 1.3w$, $w = 10$,$r = 0.01p$)

		$F \leq 1$ (U=3 weeks)			$F \leq 1$(U=10 weeks)		
Parameters	β	T	F	$\delta(\%)$	T	F	$\delta(\%)$
$g = 0$	0.1	0.28	0.79	0.31	0.29	0.40	0.56
	0.5	0.30	0.80	1.60	0.34	0.45	3.13
	0.9	0.30	0.80	2.90	0.37	0.49	6.15
$g = 0.5(p - r)$	0.1	0.28	0.79	0.34	0.29	0.32	0.90
	0.5	0.30	0.80	1.75	0.32	0.38	4.73
	0.9	0.30	0.80	3.18	0.35	0.43	8.85
$g = p - r$	0.1	0.28	0.79	0.38	0.29	0.35	1.20
	0.5	0.30	0.80	1.91	0.31	0.39	6.08
	0.9	0.30	0.80	3.47	0.31	0.39	11.03

5.6.2 General case: Inventory-Distribution system

In this section, we consider the general inventory-distribution system in which the company also determines the optimal distribution strategy. We calculate the impact of the flexible delivery option on the distribution costs as follows:

$$\delta_{dist} = 100(ASC_{dist}(F=1) - ASC_{dist}(F \leq 1)/ASC_{dist}(F=1), \tag{5.29}$$

where $F = 1$ corresponds to the case in which the flexible delivery option is not offered. In each case, we use the model to find the joint optimal inventory-distribution strategy that maximises the $ASTP$ value, and use the corresponding values obtained as given by (5.21) in the evaluation of (5.29).

Table 5.3 shows how flexible delivery affects distribution costs on the chosen instance at six different levels of the distribution cost c_o charged by the third party.

When $c_o = c_u^1$, the distribution is always outsourced (see Lemma 3). At first sight, one may perhaps not expect any improvements in the distribution costs from flexible delivery in that case. However, since the flexible delivery option affects the timing when we pay the third party for its distribution tasks, the NPV model does still show a fair benefit occurring. For example, with $U = 3$ weeks, we would still expect a benefit of 2.5% from the flexible delivery option.

When $c_o = c_u^2$, distribution is never outsourced (Lemma 3). The benefits from flexible delivery now become larger. This can be expected since we will not only have a benefit from postponing distribution costs (as in the case of $c_o = c_u^1$), but also will reduce the actual average unit distribution cost per demand delivery over a cycle.

In the case of $c_o = c_u(k)$, we assume that c_o equals the company's unit delivery cost when delivering to fast delivery customers in periods in which the flexible delivery option is not offered. In this situation, we expect that only the deliveries including postponed demand are fulfilled by the company, while the delivery of normal demand customers in other periods can be outsourced. The results in Table 5.3 now show that flexible delivery will the highest potential to improve the firm's distribution costs, with values ranging between 8% and 22%.

TABLE 5.3: Impact of flexible delivery on distribution AS costs ($S = 80$,$\alpha = 0.2$, $w = 10$, $r = 0.01p$, $\beta = 0.5$, $g = p - r$)

Exp.	Unit Cost	Dist. Strategy	U=3 weeks T	U=3 weeks F	U=3 weeks δ_{dist}	U=6 weeks T	U=6 weeks F	U=6 weeks δ_{dist}
1	$c_o = c_u^1 = 0.05w$	Outsource	0.30	0.80	2.5	0.29	0.59	5.0
	$c_o = c_u^1 = 0.10w$		0.30	0.80	2.5	0.29	0.59	5.0
2	$c_o = c_u^2 = 0.05w$	Local delivery	0.30	0.80	3.3	0.29	0.59	7.6
	$c_o = c_u^2 = 0.10w$		0.27	0.78	3.5	0.26	0.54	8.4
3	$c_o = c_u(k) = 0.05w$	Partial Outsource	0.26	0.77	8.3	0.26	0.54	17.8
	$c_o = c_u(k) = 0.10w$		0.26	0.77	8.3	0.21	0.43	21.9

We now consider the impact of flexible delivery on the total system AS profits. The relative improvement from flexible delivery is calculated as in (5.28), but now including the variable distribution cost. Figure 5.7 illustrates the percentage improvement as a function of r for a setting that corresponds to the setting in the bottom row of Table 5.2, while the distribution cost parameters are set as the bottom row in Table 5.3, case $c_o = 0.05w$. Flexible delivery in a jointly optimised inventory-distribution system performs much better than when only accounting for inventory profits. At $r = 1\%p$ and $\beta = 0.5$, for example, flexible delivery could improve the inventory profits by 1.91% (see Table 5.2), while in the jointly optimised inventory-distribution system this becomes 4.47%. Figure 5.7 also shows that, while a higher β brings more benefit, the maximum profitable discount value r increases only modestly with β.

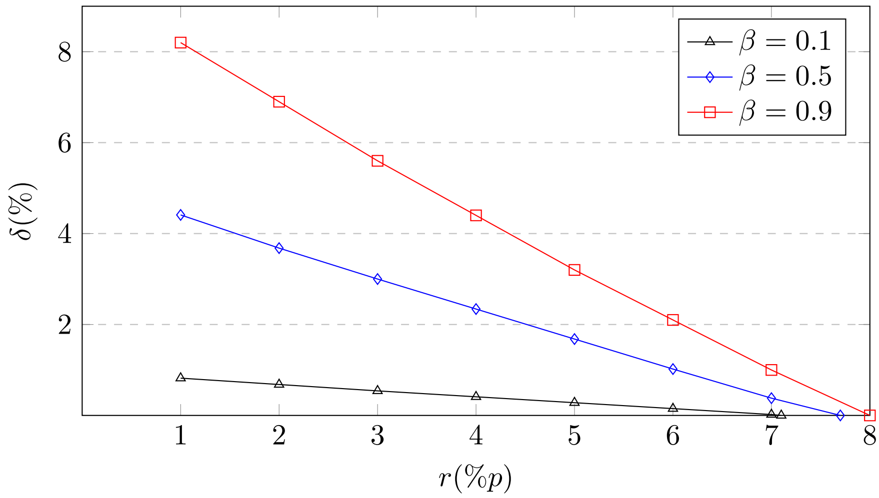

FIGURE 5.7: Impact of flexible delivery on total AS profits ($p = 1.3w$, $g = p - r$, $S = 80$, $c_u = 0.05w$, $\alpha = 0.2$, $U = 3$ weeks)

Figure 5.7 also illustrates that is may not always be worthwhile to offer more discount in order to get more customer acceptance of the flexible delivery option. Instead of offering 5% discount to get $\beta = 0.9$, for example, the $\beta = 0.5$ case with 2% discount would achieve a higher overall profit for the firm. Firms wanting to apply the flexible delivery policy should thus consider the trade-off between customer acceptance and incentive structure. A low flexible delivery acceptance with low discount amount can sometimes be better than a higher acceptance with higher discounts.

In order to illustrate this trade-off more clearly, we end this section with an experiment in which we make use of the customer acceptance function defined in Section 5.3.4. We use the same parameter settings as used in Figure 5.7, but now let β be a function of r for three different values of the resistance function $\theta = 0.1, 0.2, 0.5$ (for $U = 3$ weeks and $g = p - r$). We can now use the algorithm repeatedly over a range of r values to produce the results as illustrated in Fig. 5.8. As shown, the optimal discount would then be $2\%, 3\%, 4\%$ when $\theta = 0.1, 0.2, 0.5$, respectively. At these optimal discount rates, customer acceptance is at the level $\beta = 0.68, 0.60, 0.38$, respectively (evaluate (5.7)). The higher the customer resistance, the higher the optimal discount value but the lower the customer acceptance rate and the lower the total profit increase achievable from flexible delivery. When applying the flexible delivery option, firms should thus also aim to determine the customer acceptance function.

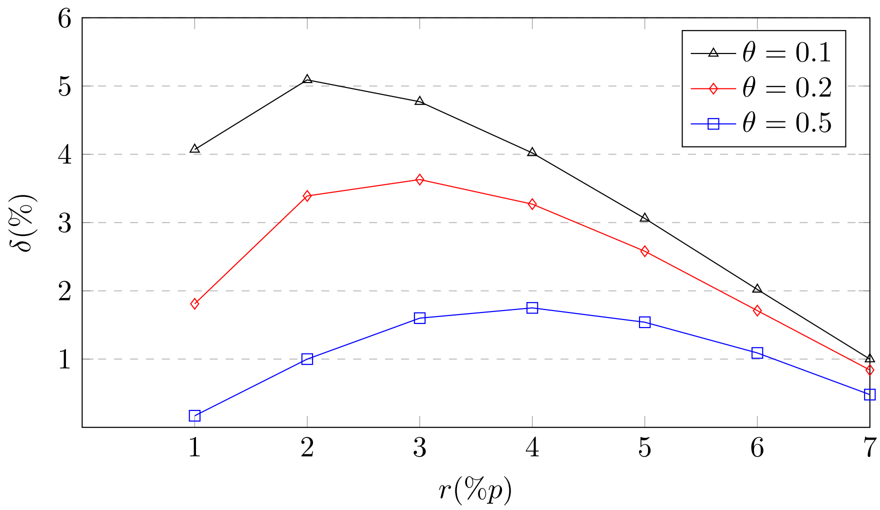

FIGURE 5.8: Effects of Discount by $\beta(r)$ ($p = 1.3w$, $w = 10$, $S = 80$, $g = p - r$, $U = 3$ weeks)

To summarize, we have shown that the flexible delivery option can improve system performance. The joint optimisation of the inventory and distribution strategy shows larger benefits than only considering the inventory strategy, in particular when the company has the option to distribute the delivery between a third party and its own vehicle. The company is able to offer larger discounts to customers accepting flexible delivery when performing part of the distribution by itself. To ensure optimal overall performance of the system with flexible delivery, firms need to investigate how customers respond to the offer, in particular how acceptance of flexible delivery is a function of the discount offer, the maximum waiting time, and deposit value.

5.6.3 Impact flexible delivery on distance travelled

In this part, we illustrate the potential impact of flexible delivery on the average travelled distance when the distribution is solely executed by the firm. This measure may be of interest to firms to assess their impact on emissions from transportation. The performance measure of interest is now:

$$\delta_{AVD} = 100(AVD_{F=1} - AVD_{F\leq 1})/AVD_{F=1}, \tag{5.30}$$

where $AVD_{F=1}$ is the average distance travelled in the optimal solution that maximises the AS profits of the joint inventory-distribution system in the case that no flexible delivery is offered, while $AVD_{F=1}$ is the average distance travelled in the optimal solution when the flexible delivery option is offered.

The numerical experiments are summarized in Table 5.4. As expected, travelled distance reduces for higher maximum waiting times and customer acceptance levels. Flexible delivery can thus not only help the firm to increase its profitability, but may also help to reduce the impact on the environment. When customers are sensitive to this topic, it may help the firm to make customers aware of this positive environmental benefit and thereby increase customers' willingness to accept flexible delivery.

TABLE 5.4: Average travelled distances ($c_u = 0.05w$,$p = 1.3w = 13$, $g = p - r$, $r = 0.01p$, $A = 50$ miles^2, $\gamma = 0.30$)

Max Waiting	$F = 1$			$F \leq 1$				
Time	T	F	Dist	β	T	F	Dist	$\delta_{avgdist}(\%)$
3 weeks	0.29	1	2276.84	0.1	0.29	0.8	2267.81	0.396
				0.5	0.3	0.8	2196.59	3.525
				0.9	0.26	0.77	2018.86	11.330
6 weeks	0.29	1	2276.84	0.1	0.29	0.59	2255.68	0.930
				0.5	0.29	0.59	2095.03	7.985
				0.9	0.26	0.54	1720.36	24.441

5.7 Conclusions

Shifting the demand over time has become an increasingly studied strategy in supply chain management. Most studies have analysed demand postponement in a context of production resource capacity constraints. We have developed a model to study the value of flexible delivery in a continuous review inventory and distribution system. The majority of studies have focused on static postponement schemes available to all customers at all times. In our case we only offer the postponement at a time when the inventory level reaches a critical value, and incorporated not only the discount as a means to induce customers to accept flexible delivery, but also considered the deposit value and maximum waiting time as parameters that can affect customer acceptance. In order to see the impact of changes in the timing of cash-flows as a result of offering the flexible delivery option, we use a NPV model that is able to consider these financial aspects of the system. Numerical experiments under different parameters settings were presented to analyse the benefits of flexible delivery in comparison to a system in which this option is not offered.

The findings presented show that demand postponement through flexible delivery can bring significant benefits. With respect to inventory costs, more benefits can be expected when the marginal profit on a product remains modest, as then the optimisation of the inventory costs is relatively more important in determining the firm's profits. At marginal value $p = 1.3w$, we find that profits from optimising the inventory system may be boosted with a few percentages by offering a discount on the sales price of a smaller to equal magnitude, unless the firm's replenishment set-up cost and opportunity cost of capital are very small. We have further shown how profitability depends on customer acceptance rate (β), discount offered (r), maximum waiting time (U), and deposit value (g).

Larger benefits from flexible delivery are observed when jointly optimising the inventory-distribution system. The benefit of flexible delivery to the distribution system depends on the price of outsourcing distribution relative to the firm's own delivery cost structure; the largest benefits in the distribution system can be expected when the firm finds it optimal to partially outsource delivery in delivery periods with low customer density. Total profits in the inventory-distribution system analysed can, ceteris paribus, more than double in comparison to the inventory system (i.e. with fixed distribution cost structure).

For given r, U and g, benefits of the flexible delivery option generally increase with β, but customer acceptance can be expected to increase with r but decrease with U and g relative to $p - r$. We have shown that the model can be used to determine the optimal discount value r when knowing the function $\beta(r, U, g)$ for different customer resistance function values $\theta(U, g)$. Higher customer resistance values lead to higher optimal discount values, lower acceptance rates, and lower overall profits achievable. Extensions of this approach can be readily adopted to identify the optimal combination of values for all three parameters r, U, and g. To ensure optimal overall performance of the system with flexible delivery, knowledge of $\beta(r, U, g)$ is thus important. Further research into how firms can gather sufficient information to determine how customers respond to the flexible delivery option is warranted. When the firm only uses its own vehicle for transport, we have shown that flexible delivery can help to reduce transport distances and thus save on emissions.

The NPV methodology adopted in this paper also illustrates its benefits in showing how changes in the timing of payments can affect system performance. While Ghiami and Beullens (2016) found that deposit values g do not greatly affect performance in their production-inventory NPV model, they only tested values of deposits up to $20\%p$, which is reasonable in the context of backorders. In our context of online ordering, it is possible and in fact quite common that customers pay in full when placing the order. In our inventory-distribution system, the impact of g is thus found to be more significant. We also found that accounting for the timing of transportation costs is of importance in our system. Even when fully outsourced, i.e. when the total outgoing payments to the third party transportation company are unaffected by the distribution strategy, the change in the timing of these costs arising from flexible delivery decreased AS distribution costs by 2.5% for 3 weeks and 5% for a 6 weeks maximum waiting time. Models not based on the NPV methodology would not be able to identify such impacts of flexible delivery on system performance.

Further research in this area can investigate how system performance is affected by the customer acceptance function, i.e. by the shape of the function $\beta(r, g, U)$. Different possible flexible delivery schemes can also be investigated, as well as different types of distribution systems or cost structures. Since flexible delivery tends to increase replenishment order sizes, its value may further increase in situations where suppliers would offer quantity discounts, or offer a delay in payment if the order is higher than a threshold level, as in e.g. Chung et al. (2005) and Pal and Chandra (2014). Additional benefits of larger shipments may also lead to reductions in emissions from transportation between the firm and its supplier, as argued in Van Hoek (1999).

Chapter 6

Conclusions

6.1 Overview

The main focus of this thesis is to make the inventory system more flexible to the uncertainties through a demand postponement policy. In Chapters 1-2 an overview introduction including the research aims and the literature review of lot sizing problems and relevant methodology to identify and address the gaps and address the gaps and methodology. Chapters 3-5 are the main research papers which present three different mathematical models.

In this chapter, we present an summary of contributions and the conclusions, limitations and the further research directions of research problems.

6.2 Summary of the Main Contributions

This thesis has studied a demand postponement policy on several capacitated and uncapacitated inventory systems. We model the problem and propose appropriate methodology for the solution. The contributions of the research papers are summarized in following sections.

6.2.1 Chapter 3. NPV Analysis of a Periodic Review Inventory Model with Price Discount

The research content in this paper is based on a real case inventory problem which has batch ordering and lost sales. Using the demand postponement policy helps us to convert some lost sales into observed advance demand at some inventory levels. Agreeing with the customers for later deliveries prevents lost sales. We model the problem by a Markov Decision Process (MDP) formulation and solve it through the LP model. The structure of MDP could give some insight which can be used to develop a heuristic method. Although this research looks similar to the partial backordering problems, agreeing with the customers makes out case more realistic. Moreover, based on the customers willingness to wait, more delays can be offered which, is not applicable for backordering.

6.2.2 Chapter 4. Inventory Decisions under Advance Demand Information Driven by Price Discount

In this research, an inventory system with stochastic advance demand information (ADI) is studied for capacitated/uncapacitated inventory systems. We model the problem with demand postponement by Markov Decision Process (MDP) and provide a backward induction algorithm as a solution technique for finite time horizon. Due to the complexity of the model, we develop an action elimination method to bound the reorder level and order up to level. The structure of the optimal integrated replenishment and postponement policy is presented. The numerical findings indicate that the benefit of demand postponement can have a potential to deal with the negative impact of capacity limit.

6.2.3 Chapter 5. Analysis of Replenishment and Discount Policies with Distribution by NPV Analysis

An update of the problem in Chapter 4 to a continuous review case with distribution strategies has been studied in this section. Beside, having a replenishment policy, the firm also needs to decide whether to outsource the delivery process or not. The continuous approximation of travelled distance (Daganzo and Newell 1985) has been used to calculate the total distance. Since the postponement policy affects the number demand delivered at each delivery, it has a big impact on delivery process. The analysis of delivery costs by postponement policy is investigated. An extensive numerical experiments are presented to show the value of demand postponement on inventory and distribution policies with different financial settings.

6.3 Limitations of the thesis and directions for future research

For the real life applications of the models developed in this thesis, the limitations of the research models need to be highlighted and the some further research avenues are explored.

Firstly, one of the biggest limitations is the computational complexity of the MDP problems. This limitation only affects the Chapter 3 and Chapter 4. In Chapter 3, the postponement period is limited to one period in numerical experiments and the information horizon is limited to $N = L + 2$ to reduce computational effort. However, the MDP formulation gives us some structural insights to develop an efficient heuristic method which could be a good future research direction.

The assumption of having deterministic demand in Chapter 5 is also another limitation. A more general model should be developed to deal with the stochastic demand and stochastic lead time. This could a fruitful research idea for future and it is expected that the demand postponement will make bigger impact for stochastic cases. Considering the customer acceptance rate dependent on multiple parameters (discount amount, waiting time, deposit etc.) is also another interesting research idea.

One of the consequences of considering demand postponement into inventory systems is to make order quantities tend to increase. While the models in this thesis have not considered the benefits of larger order sizes, many papers in literature focus on this issue by offering quantity discounts for larger order size (Chung et al. 2005) or a delay in payment if order size is greater than a threshold level (Pal and Chandra 2014). Another benefit of larger order size is to reduce the number of orders which lead to reduction in emissions due to transportation (Van Hoek 1999). These could attract researchers as future studies on benefits of delivery postponement on supply chain.

Appendix A

Supplement to Chapter 4

A.1 Proof of Lemma 4.1

The proof can be obtained by using the induction argument on recursive function. When $t = T$, the problem becomes a single period inventory problem. Then $V_T(., O_T, Q_T) = G_T(.) + \beta(a_t)$ and $G_T(.)$ is a convex and coercive function based on the standard assumptions of inventory literature and $\beta(a_t)$ is also an increasing function. So that makes $V_T(s_T)$ convex. By using the induction for $t = T - 1$, we can show the convexity of V_t and P_t functions for other periods.

$$V_t(s_t) = min\{c_t(s_t, a_t) + G_T(s_{t+1})\}$$

$$P_t(y_t, O_t, Q_t) = G_t(y_t) + \lambda EV_T(x_T, O_T, Q_T)$$

With the convexity of $V_T(s_T)$, we conclude that $P_t(y_t, O_t, Q_t)$ is a convex function due to the sum of two convex functions and $\lim_{|y|\to\infty} P(y, O, Q) = \infty$. This properties show that there is a point $x < S_t(O_T, Q_T)$ which satisfies the following function since the P goes to infinity by $|x|\to\infty$.

$$P(x, O_t, Q_t) > K_t + P(S_t(O_t, Q_t), O_t, Q_t)$$

By considering the equation of $H_t(.)$ at (4.23), it can be derived that $H_t(.)$ has a sign change from - to + due to the properties of $P(y, O, Q)$.

A.2 Proof of Theorem 4.3

The proof of first statement based on the definition of $s_t(O_t, Q_t)$. Since it is the maximum value of x at which $H_t(x, O_t, Q_t)$ is non-positive. The higher setup cost needs to have higher difference between $P(y, O, Q)$ and $P(x, O, Q)$. Since the higher value of $P(x, O, Q)$ can be obtained by lower x. Thus, the optimal value of $s_t()$ reduces.

For the proof of Statement 2, we refer reader Özer and Wei (2004) for a detailed proof.

The similar analysis is applied for the Statement 3. We show that $P_t(x, O, Q)$ is decreasing by t and it is true that $P_{t-1} \geq P_t$ due to the induction with discount rate. The $H(.)$ function will be higher by increasing $P(.)$ values. Thus, it will be negative for lower values of $s_t(.)$.

Appendix B

Supplement to Chapter 5

B.1 Proof of Lemma 5.1

Taking the second partial derivatives of $ASTP(F,T)$ with respect to F and T yields

$$\frac{\partial^2 ASTP}{\partial F^2}(F,T) = -\alpha\beta(p-g)Ty < 0,$$
$$\frac{\partial^2 ASTP}{\partial T^2}(F,T) = \frac{-2S}{T^3} < 0,$$

The $p-g$ is always positive from the definition, then $ASTP(F,T)$ is concave while the S positive.

B.2 Proof of Prop 1

Recall that the Lemma 5.1 shows that the objective function is concave. Then, the variables obtained by partial derivatives are optimal. The optimal values of T are;

$$T = \frac{\partial ASTP}{\partial T}(F,T) = \sqrt{\frac{2S}{\alpha y(w-\beta(1-F)(2w+r+(1-F)(p-g))}}$$

$$T_{F\leq 1} \geq T_{F=1}$$

$$\sqrt{\frac{2S}{\alpha y(w-\beta(1-F)(2w+r+(1-F)(p-g))}} \geq \sqrt{\frac{2S}{\alpha wy}}$$

$$2w+r+(1-F)(p-g) \geq 0$$

w, r and $p-g$ are always non-negative and $F \leq 1$ which left hand side of the equation is always greater than or equal to 0.

B.3 Proof of Lemma 5.2

The proof is based on the $ASTP(F,T)$ function when $F=1$. The r value can be easily derived from the partial derivatives of $ASTP(F,T)$.

$$F = \frac{\partial ASTP}{\partial F}(F,T) = -\frac{2\alpha Tw + (\alpha r - 2\alpha p + 2\alpha g)\,T - 2r}{(2\alpha p - 2\alpha g)\,T} = 1$$

$$-\frac{w}{p-g} - \frac{(\alpha T-2)r}{2\alpha(p-g)T} = 0$$

$$r = \frac{2\alpha wT}{2-\alpha T}$$

$$T = \frac{\partial ASTP}{\partial T}(F,T) = \sqrt{\frac{2S}{\alpha y(w-\beta(1-F)(2w+r+(1-F)(p-g))}}$$

$$T = \sqrt{\frac{2S}{\alpha yw}} \text{when F=1}$$

When we solve the two equations of r and T then the discount amount which makes $F=1$ is;

$$r = \frac{\sqrt{\frac{8S\alpha w}{y}}}{2-\sqrt{\frac{2S\alpha}{yw}}}$$

B.4 Proof of Lemma 5.3

The findings on Lemma 5.3 are obtained based on the number of demand $n(k)$ and unit distribution cost c_u functions (described in Section 5.3.3). The unit distribution cost with the maximum number of demand is taken as the minimum cost while it is maximum for the minimum number of demand.

Bibliography

Abbasi-Yadkori, Y., Bartlett, P., Malek, A., 2014. Linear programming for large-scale Markov decision problems. In: Proceedings of the 31st International Conference on Machine Learning.

Agatz, N., Campbell, A. M., Fleischmann, M., van Nunen, J., Savelsbergh, M., 2013. Revenue management opportunities for internet retailers. Journal of Revenue and Pricing Management 12 (2), 128–138.

Aggarwal, S. C., 1974. A review of current inventory theory and its applications. International Journal of Production Research 12 (4), 443–482.

Ahiska, S. S., Appaji, S. R., King, R. E., Warsing, D. P., 2013. A markov decision process-based policy characterization approach for a stochastic inventory control problem with unreliable sourcing. International Journal of Production Economics 144 (2), 485–496.

Aksen, D., Altinkemer, K., Chand, S., 2003. The single-item lot-sizing problem with immediate lost sales. European Journal of Operational Research 147 (3), 558–566.

Aloulou, M. A., Dolgui, A., Kovalyov, M. Y., 2013. A bibliography of non-deterministic lot-sizing models. International Journal of Production Research 52 (8), 2293–2310.

Alp, O., Tim Huh, W., Tan, T., 2014. Inventory control with multiple setup costs. Manufacturing & Service Operations Management 16 (1), 89–103.

Baciarello, L., D'Avino, M., Onori, R., Schiraldi, M. M., 2013. Lot sizing heuristics performance. International Journal of Engineering Business Management 5, 5.

Bakker, M., Riezebos, J., Teunter, R. H., 2012. Review of inventory systems with deterioration since 2001. European Journal of Operational Research 221 (2), 275–284.

Beamon, B. M., 1998. Supply chain design and analysis:: Models and methods. International Journal of Production Economics 55 (3), 281 – 294.
URL http://www.sciencedirect.com/science/article/pii/S0925527398000796

Beck, F. G., Grosse, E. H., Teßmann, R., 2015. An extension for dynamic lot-sizing heuristics. Production & Manufacturing Research 3 (1), 20–35.

Benjaafar, S., Cooper, W., Mardan, S., 2011. Production-inventory systems with imperfect advance demand information and updating. Naval Research Logistics 58 (2), 88–106.

Beullens, P., Janssens, G., 2014. Adapting inventory models for handling various payment structures using net present value equivalence analysis. International Journal of Production Economics 157, 190–200.

Beullens, P., Janssens, G. K., 2011. Holding costs under push or pull conditions–the impact of the anchor point. European Journal of Operational Research 215 (1), 115–125.

Brahimi, N., Absi, N., Dauzère-Pérès, S., Nordli, A., 2017. Single-item dynamic lot-sizing problems: An updated survey. European Journal of Operational Research 263 (3), 838–863.

Brahimi, N., Dauzere-Peres, S., Najid, N. M., Nordli, A., 2006. Single item lot sizing problems. European Journal of Operational Research 168 (1), 1–16.

Broekmeulen, R. A., van Donselaar, K. H., 2009. A heuristic to manage perishable inventory with batch ordering, positive lead-times, and time-varying demand. Computers & Operations Research 36 (11), 3013–3018.

Buhayenko, V., van Eikenhorst, E., 2015. Synchronisation of customers' orders using discounts. International Journal of Systems Science: Operations & Logistics, 1–15.

Burgess, K., S. P., Koroglu, R., 2006. Supply chain management: a structured literature review and implications for future research. International Journal of Operations & Production Management 26 (7), 703–729.

Buschkühl, L., Sahling, F., Helber, S., Tempelmeier, H., 2008. Dynamic capacitated lot-sizing problems: a classification and review of solution approaches. OR Spectrum 32 (2), 231–261.

Chao, H.-p., 1992. The eqq model with stochastic demand and discounting. European Journal of Operational Research 59 (3), 434–443.

Chao, X., Zhou, S. X., 2009. Optimal policy for a multiechelon inventory system with batch ordering and fixed replenishment intervals. Operations Research 57 (2), 377–390.

Chen, F., 2000. Optimal policies for multi-echelon inventory problems with batch ordering. Operations Research 48 (3), 376–389.

Chen, F., 2001. Market segmentation, advance demand information, and supply chain performance. Manufacturing and Service Operations Management 34 (1), 482–499.

Chung, C., 1987. An optimal algorithm for the quantity discount problem. Journal of Operations Management 7 (1-2), 165–177.

Chung, K., Goyal, S., Huang, Y., 2005. The optimal inventory policies under permissible delay in payments depending on the ordering quantity. International Journal of Production Economics 95 (2), 203–213.

Cárdenas-Barrón, L. E., 2007. Optimizing inventory decisions in a multi-stage multi-customer supply chain: A note. Transportation Research Part E: Logistics and Transportation Review 43 (5), 647–654.

Cárdenas-Barrón, L. E., Chung, K.-J., Treviño-Garza, G., 2014. Celebrating a century of the economic order quantity model in honor of ford whitman harris. International Journal of Production Economics 155, 1–7.

Daganzo, C. F., Newell, G. F., 1985. Physical distribution from a warehouse: Vehicle coverage and inventory levels. Transportation Research Part B: Methodological 19 (5), 397–407.

Drexl, A., Kimms, A., 1997. Lot sizing and scheduling — survey and extensions. European Journal of Operational Research 99 (2), 221–235.

Dye, C.-Y., Ouyang, L.-Y., Hsieh, T.-P., 2007. Inventory and pricing strategies for deteriorating items with shortages: A discounted cash flow approach. Computers & Industrial Engineering 52 (1), 29–40.

Elmaghraby, S. E., 1978. The economic lot scheduling problem (elsp): Review and extensions. Management Science 24 (6), 587–598.

Elmaghraby, W., Keskinocak, P., 2003. Dynamic pricing in the presence of inventory considerations: research overview, current practices, and future directions. IEEE Engineering Management Review 31 (4), 47–47.

Erlenkotter, D., 1990. Ford whitman harris and the economic order quantity model. Operations Research 38 (6), 937–946.

Federgruen, A., Lee, C.-Y., 1990. The dynamic lot size model with quantity discount. Naval Research Logistics 37 (5), 707–713.

Fianu, S., Davis, L. B., 2018. A markov decision process model for equitable distribution of supplies under uncertainty. European Journal of Operational Research 264 (3), 1101–1115.

Gaafar, L., 2006. Applying genetic algorithms to dynamic lot sizing with batch ordering. Computers & Industrial Engineering 51 (3), 433–444.

Gallego, G., Özer, Ö.., 2001. Integrating replenishment decisions with advance demand information. Management Science 47 (10), 1344–1360.

Gallego, G., Özer, Ö., 2002. Supply Chain Structures; Coordination, Information and Optimization. Kluwer Academic Publisher, Ch. Optimal Use of Advance Demand Information in Supply Chain Management, pp. 119–160.

Gallego, G., Özer, Ö.., 2003. Optimal replenishment policies for multiechelon inventory problems under advance demand information. Manufacturing & Service Operations Management, 157–175.

Gao, L.-L., Altay, N., Robinson, E. P., 2008. A comparative study of modeling and solution approaches for the coordinated lot-size problem with dynamic demand. Mathematical and Computer Modelling 47 (11-12), 1254–1263.

Gayon, J., Benjaafar, S., de Vericourt, F., 2009. Using imperfect advance demand information in production-inventory systems with multiple customer classes. Manufacturing & Service Operations Management 11 (1), 128–143.

Ghaniabadi, M., Mazinani, A., 2017. Dynamic lot sizing with multiple suppliers, backlogging and quantity discounts. Computers & Industrial Engineering 110, 67–74.

Ghiami, Y., Beullens, P., 2016. Planning for shortages? net present value analysis for a deteriorating item with partial backlogging. International Journal of Production Economics 178, 1–11.

Giri, B., Dohi, T., 2004. Optimal lot sizing for an unreliable production system based on net present value approach. International Journal of Production Economics 92 (2), 157–167.

Glock, C. H., Grosse, E. H., Ries, J. M., 2014. The lot sizing problem: A tertiary study. International Journal of Production Economics 155, 39–51.

Goyal, S., Gunasekaran, A., 1990. Multi-stage production-inventory systems. European Journal of Operational Research 46 (1), 1–20.

Graves, S. C., 1981. A review of production scheduling. Operations Research 29 (4), 646–675.

Grubbström, R. W., 1980. A principle for determining the correct capital costs of work-in-progress and inventory. International Journal of Production Research 18 (2), 259–271.

Grubbström, R. W., 1998. A net present value approach to safety stocks in planned production. International Journal of Production Economics 56-57, 213–229.

Grubbström, R. W., 2007. Transform methodology applied to some inventory problems. Zeitschrift für Betriebswirtschaft 77 (3), 297–324.

Guan, Y., Ahmed, S., Nemhauser, G. L., Miller, A. J., 2005. A branch-and-cut algorithm for the stochastic uncapacitated lot-sizing problem. Mathematical Programming 105 (1), 55–84.

Gupta, Y. P., Keung, Y., 1990. A review of multi-stage lot-sizing models. International Journal of Operations & Production Management 10 (9), 57–73.

Hadley, G., 1964. A comparison of order quantities computed using the average annual cost and the discounted cost. Management Science 10 (3), 472–476.

Hariharan, R., Zipkin, P., 1995. Customer-order information, leadtimes, and inventories. Management Science 41 (10), 1599–1607.

Harris, F. W., 1913. How many parts to make at once. Factor, The Magazine of Management 10 (2), 135–136.

Hax, A., Candea, D., 1984. Production and Inventory Management. Prentice-Hall. URL https://books.google.co.uk/books?id=Wq9vQgAACAAJ

Hernández-Lerma O., L. J., 1999. The Linear Programming Approach. Vol. 42 of In: Further Topics on Discrete-Time Markov Control Processes. Applications of Mathematics (Stochastic Modelling and Applied Probability). Springer, New York, NY.

Holmbom, M., Segerstedt, A., 2014. Economic order quantities in production: From harris to economic lot scheduling problems. International Journal of Production Economics 155, 82–90.

Hsieh, T.-P., Dye, C.-Y., Ouyang, L.-Y., 2008. Determining optimal lot size for a two-warehouse system with deterioration and shortages using net present value. European Journal of Operational Research 191 (1), 182–192.

Huang, T., Van Mieghem, J. A., 2013. Clickstream data and inventory management: Model and empirical analysis. Production and Operations Management 23 (3), 333–347.

Iyer, A. V., Deshpande, V., Wu, Z., 2003. A postponement model for demand management. Management Science 49 (8), 983–1002.

Jans, R., Degraeve, Z., 2007. Meta-heuristics for dynamic lot sizing: A review and comparison of solution approaches. European Journal of Operational Research 177 (3), 1855–1875.

Karaesmen, F., Liberopoulos, G., Dallery, Y., 2004. The value of advance demand information in production/inventory systems. Annals of Operations Research 126 (1-4), 135–157.

Karimi, B., Fatemi Ghomi, S., Wilson, J., 2003. The capacitated lot sizing problem: a review of models and algorithms. Omega 31 (5), 365–378.

Kesen, S. E., Kanchanapiboon, A., Das, S. K., 2010. Evaluating supply chain flexibility with order quantity constraints and lost sales. International Journal of Production Economics 126 (2), 181–188.

Kremer, M., Van Wassenhove, L. N., 2014. Willingness to pay for shifting inventory risk: The role of contractual form. Production and Operations Management 23 (2), 239–252.

Kunnumkal, S., Topaloglu, H., 2008. Price discounts in exchange for reduced customer demand variability and applications to advance demand information acquisition. International Journal of Production Economics 111 (2), 543–561.

Lagodimos, A., Christou, I., Skouri, K., 2012. Optimal (r,nq,t) batch ordering with quantized supplies. Computers & Operations Research 39 (2), 259–268.

Lee, A. H., Kang, H.-Y., Lai, C.-M., Hong, W.-Y., 2013. An integrated model for lot sizing with supplier selection and quantity discounts. Applied Mathematical Modelling 37 (7), 4733–4746.

Lee, C.-Y., Eetinkaya, S., Wagelmans, A. P. M., 2001. A dynamic lot-sizing model with demand time windows. Management Science 47 (10), 1384–1395.

Lewis, C., 2008. Linear programming: Theory and applications. Whitman College.

Li, C., Zhang, F., 2013. Advance demand information, price discrimination, and preorder strategies. Manufacturing & Service Operations Management 15 (1), 57–71.

Lu, Y., Song, J., Yao, D. D., 2003. Order fill rate, leadtime variability, and advance demand information in an assemble-to-order system. Operations Research 51 (2), 292–308.

Lulli, G., Sen, S., 2004. A branch-and-price algorithm for multistage stochastic integer programming with application to stochastic batch-sizing problems. Management Science 50 (6), 786–796.

Marchi, B., Ries, J., Zanoni, S., Glock, C., 2016. A joint economic lot size model with financial collaboration and uncertain investment opportunity. International Journal of Production Economics 176, 170–182.

Mazdeh, M. M., Emadikhiav, M., Parsa, I., 2015. A heuristic to solve the dynamic lot sizing problem with supplier selection and quantity discounts. Computers & Industrial Engineering 85, 33–43.

Özer, Ö., 2003. Replenishment strategies for distribution systems under advance demand information. Management Science 49 (3), 255–272.

Özer, Ö.., 2011. Inventory Management: Information, Coordination, and Rationality. Springer, pp. 321–365.

Özer, Ö., Wei, W., 2004. Inventory control with limited capacity and advance demand information. Operations Research 52 (6), 988–1000.

Pagh, J., Cooper, M., 1998. Supply chain postponement and speculation strategies: How to choose the right strategy. Journal of Business Logistics 19 (2), 13.

Pal, M., Chandra, S., 2014. A periodic review inventory model with stock dependent demand, permissible delay in payment and price discount on backorders. Yugoslav Journal of Operations Research 24 (1), 99–110.

Papachristos, S., Skouri, K., 2003. An inventory model with deteriorating items, quantity discount, pricing and time-dependent partial backlogging. International Journal of Production Economics 83 (3), 247–256.

Pentico, D. W., Drake, M. J., 2011. A survey of deterministic models for the eoq and epq with partial backordering. European Journal of Operational Research 214 (2), 179–198.

Pochet, Y., Wolsey, L. A., 2011. Production planning by mixed integer programming. Springer.

Powell, W. B., 2007. Introduction to Markov Decision Processes. Wiley Blackwell, Ch. 3, pp. 47–90.
URL https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470182963.ch3

Power, D., 2005. Supply chain management integration and implementation: a literature review. Supply Chain Management: An International Journal 10 (4), 252–263.

Puterman, M. L., 1994. Markov decision processes. Wiley.

Qiu, J., Loulou, R., 1995. Multiproduct production/inventory control under random demands. IEEE Transactions on Automatic Control 40 (2), 350–356.

Quante, R., Meyr, H., Fleischmann, M., Jan 2009. Revenue management and demand fulfillment: matching applications, models, and software. OR Spectrum 31 (1), 31–62.
URL https://doi.org/10.1007/s00291-008-0125-8

Robinson, P., Narayanan, A., Sahin, F., 2009. Coordinated deterministic dynamic demand lot-sizing problem: A review of models and algorithms? Omega 37 (1), 3–15.

Russell, R., Taylor, B., 2011. Operations management, 1st Edition. Wiley.

Sarkar, S., Shewchuk, J., 2013. Use of advance demand information in multi-stage production-inventory systems with multiple demand classes. International Journal of Production Research 51 (1), 57–68.

Sarkar, S., Shewchuk, J. P., 2016. Advance demand information and early fulfillment in inventory management with two demand classes. International Journal of Logistics Systems and Management 25 (2), 175.

Seref, M. M. H., Seref, O., Alptekinoglu, A., Erengüç, S. S., 2016. Advance selling to strategic consumers. Computational Management Science 13 (4), 597–626.

Shang, K. H., Zhou, S. X., 2010. Optimal and heuristic echelon (r, nq, t) policies in serial inventory systems with fixed costs. Operations Research 58 (2), 414–427.

Shin, H., Benton, W., 2007. A quantity discount approach to supply chain coordination. European Journal of Operational Research 180 (2), 601–616.

Staggemeier, A. T., Clark, A. R., 2001. A survey of lot-sizing and scheduling models. In: 23rd Annual Symposium of the Brazilian Operational Research Society (SOBRAPO). Campos do Jordão, Brazil, p. 938–947.

Sun, D., Queyranne, M., 2002. Production and inventory model using net present value. Operations Research 50(3), 528–537.

Taha, H. A., 2007. Operation Research an Introduction, 8th Edition. Prentice Hall and India Pvt Ltd.

Talluri, K. T., Hillier, F. S., 2005. Theory and Practice of Revenue Management. Springer US.

Tan, T., Güllü, R., Erkip, N., 2007. Modelling imperfect advance demand information and analysis of optimal inventory policies. European Journal of Operational Research 177 (2), 897–923.

Tang, C., 2006a. Robust strategies for mitigating supply chain disruptions. International Journal of Logistics 9 (1), 33–45.

Tang, C., Rajaram, K., Alptekinoglu, A., Ou, J., 2004. The benefits of advance booking discount programs: Model and analysis. Management Science 50 (4), 465–478.

Tang, C. S., 2006b. Perspectives in supply chain risk management. International Journal of Production Economics 103 (2), 451–488.

Tang, O., Nurmaya Musa, S., 2011. Identifying risk issues and research advancements in supply chain risk management. International Journal of Production Economics 133 (1), 25–34.

Tanrikulu, M., Sen, A., Alp, O., 2009. A joint replenishment policy with individual control and constant size orders. International Journal of Production Research 48 (14), 4253–4271.

Tempelmeier, H., 2013. Stochastic Lot Sizing Problems. Springer New York, New York, NY, pp. 313–344.
URL https://doi.org/10.1007/978-1-4614-6777-9_10

Van der Laan, E., 2003. An npv and ac analysis of a stochastic inventory system with joint manufacturing and remanufacturing. International Journal of Production Economics 81-82, 317–331.

Van Hoek, R. I., 1999. From reversed logistics to green supply chains. Supply Chain Management: An International Journal 4 (3), 129–135.

Veinott, A. F., 1965. The optimal inventory policy for batch ordering. Operations Research 13 (3), 424–432.

Wagner, H., Whitin, T., 1958. Dynamic version of the economic lot size model. Management Science 5 (1), 89–96.

Wang, T., Toktay, B., 2008. Inventory management with advance demand information and flexible delivery. Management Science 54 (4), 716–732.

Wee, H. M., 1999. Deteriorating inventory model with quantity discount, pricing and partial backordering. International Journal of Production Economics 59 (1-3), 511–518.

Wee, H., L. S., 2001. Replenishment and pricing policy for deteriorating items taking into account the time-value of money. International Journal of Production Economics 71 (1-3), 213–220.

Weng, Z. KevinWong, R. T., 1993. General models for the supplier's all-unit quantity discount policy. Naval Research Logistics 40 (7), 971–991.

White, D. J., 1985. Real applications of markov decision processes. Interfaces 15 (6), 73–83.

Wijngaard, J., Karaesmen, F., 2007. Advance demand information and a restricted production capacity: on the optimality of order base-stock policies. OR Spectrum 29(4), 643–660.

Winands, E., Adan, I., van Houtum, G., 2011. The stochastic economic lot scheduling problem: A survey. European Journal of Operational Research 210 (1), 1–9.

Woensel, van, T., Erkip, N., Curseu - Stefanut, A., Fransoo, J., 2013. Lost sales inventory models with batch ordering and handling costs. BETA publicatie : working papers. Technische Universiteit Eindhoven.

Wu, Z., Wu, J., 2015. Price discount and capacity planning under demand postponement with opaque selling. Decision Support Systems 76, 24–34.

Xu, H., Gong, Y. Y., Chu, C., Zhang, J., 2017. Dynamic lot-sizing models for retailers with online channels. International Journal of Production Economics 183, 171–184.

Yang, B., Burns, N., 2003. Implications of postponement for the supply chain. International Journal of Production Research 41 (9), 2075–2090.

Yang, B., Yang, Y., 2010. Postponement in supply chain risk management: a complexity perspective. International Journal of Production Research 48 (7), 1901–1912.

Yano, C. A., Lee, H. L., 1995. Lot sizing with random yields: A review. Operations Research 43 (2), 311–334.

Ye, Y., 2011. The simplex and policy-iteration methods are strongly polynomial for the markov decision problem with a fixed discount rate. Mathematics of Operations Research 36 (4), 593–603.

Zangwill, W. I., 1966. Production smoothing of economic lot sizes with non-decreasing requirements. Management Science 13 (3), 191–209.

Zhao, C., Guan, Y., 2014. Extended formulations for stochastic lot-sizing problems. Operations Research Letters 42 (4), 278–283.

Zheng, Y.-S., Chen, F., 1992. Inventory policies with quantized ordering. Naval Research Logistics 39 (3), 285–305.

Zhou, B., Zhao, Y., Katehakis, M. N., 2007. Effective control policies for stochastic inventory systems with a minimum order quantity and linear costs. International Journal of Production Economics 106 (2), 523–531.

Zhuang, W., Chen, J., Fu, X., 2017. Joint dynamic pricing and capacity control for hotels and rentals with advanced demand information. Operations Research Letters 45 (5), 397–402.