PRODUCTION PLANNING AND INVENTORY CONTROL OF TWO-PRODUCT RECOVERY SYSTEM IN REVERSE LOGISTICS

PAN JIE

NATIONAL UNIVERSITY OF SINGAPORE

2010

Founded 1905

PRODUCTION PLANNING AND INVENTORY CONTROL OF TWO-PRODUCT RECOVERY SYSTEM IN REVERSE LOGISTICS

PAN JIE

(*M. Eng., Beihang University*)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF INDUSTRIAL AND SYSTEMS ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2010

ACKNOWLEDGEMENTS

First and foremost, I would like to express my profound gratitude to my supervisors, Associate Professor Chew Ek Peng and Associate Professor Lee Loo Hay, who offered numerous suggestions and patient guidance throughout my whole research work. I would also give my thanks to Associate Professor Sum Chee Chuong, Associate Professor Ng Kien Ming and Dr. Huang Boray for their helpful suggestion on amending the thesis.

I greatly acknowledge the support from Department of Industrial and Systems Engineering (ISE) for providing the scholarship and the utilization of the facilities, without which it would be impossible for me to complete the work reported in this dissertation. Specially, I wish to thank the ISE Computing Laboratory technician Mr. Victor Cheo Peng Yim for his kind assistance.

My thanks also go to all my friends in the ISE Department: Han Yongbin, Liu Shudong, Hu Qingpei, Liu Xiao, to name a few, for the joy and encouragement they have brought to me. Specially, I will thank my colleagues in the Computing Lab: Liu Jiying, Aldy, Yao Zhishuang, Long Quan, Yuan Le, Zhang Haiyun, Zhu Zhecheng for the happy hours spent with them.

Finally, I would like to take this opportunity to express my appreciation for my parents. I thank them for suffering with me with their patience and eternal support. It would not have been possible without them.

Table of Contents

SUMMARY

This research focuses on a two-product recovery system in the field of Reverse Logistics. As far as the knowledge about current literature, this research could be regarded as the first study on the multi-product recovery system involving two products and two flows of returned items. Firstly, a periodic review inventory problem is studied on the two-product recovery system in the situation of lost sales over a finite horizon. A dynamic programming model has been developed in order to obtain the optimal policy of production and recovery decisions, which aims to maximize the expected total profit in the finite horizon. However, the model is difficult to be solved efficiently as no nice property could be found. Thus, the special case of the multi-period problem, a single-period problem is investigated.

Secondly, the optimal threshold level policy has been obtained for the system in a single period. For the single-period problem, the usual approach is to use Karush-Kuhn-Tucker (KKT) conditions to find the optimal solution. In this case, the answer is very complex which results in 21 different cases. However, after analyzing these 21 cases, we found out that they can be represented by an optimal multi-level threshold policy. This optimal policy is characterized by 6 order-up-to levels and 3 switching levels. By using the policy, the extension from the two-product situation to a general multi-product situation has been further discussed.

Even though this multi-level threshold policy might not be optimal for the multi-period problem, it is intuitive, easy to use and provides good managerial perspectives. Hence, we apply this policy to the multi-period problem in the situation

of lost sales at first. We have found that different from the single-period problem, the threshold levels will not only depend on the current-period cost parameters, but also on the future cost-to-go function.

Thirdly, we have developed an efficient way to compute these threshold levels:

- Unlike the usual approach which uses a single function (or piecewise function) to represent the cost-to-go function, we just need to estimate the gradient of the cost-to-go function at the points of interest by Monte Carlo simulation. These gradients will be used to compute the threshold level. Hence, the performance of the results will not depend on the function we assume which can be a challenge for most of the approximate dynamic programming approaches.
- We develop an Infinitesimal Perturbation Analysis (IPA) based approach to estimate the gradient. This approach not only uses the least computing resources but also its estimation quality is better.
- The results of the numerical experiments show that the performance of this threshold policy is found to be promising under a wide range of settings.

Finally, we have extended the multi-period problem to the situation of backorder. Furthermore, the lead time effect is investigated based on a simple case, where production lead time and recovery lead time of each product are assumed to be equal to the same nonzero constant. This multi-level threshold policy also shows good performance under a wide range of settings.

List of Tables

List of Figures

List of Symbols

M	length of planning horizon;
i	group index on returned items (i = 1, 2);
j	product index on finished items (j = 1, 2);
$r_{ij}^{(t)}$	quantity of recovering returned item in group i to product j in period t;
$p_j^{(t)}$	quantity of producing product j in period t;
$x_{Sj}^{(t)}$	initial inventory position of product j in period t;
$x_j^{(t)}$	inventory level of product j after production and recovery in period t;
s_j	selling price of product j;
c_{Rij}	unit cost of recovering returned item in group i to product j;
c_{Pj}	production cost of per unit product j;
h_j	inventory holding cost of per unit product j per period;
v_j	penalty cost of per unit shortage of product j per period;
L	lead time of production and recovery processes for each product;
$R_i^{(t)}$	returned items in group i in period t;
$D_j^{(t)}$	demand for product j in period t;
$f(x,\mu,\sigma)$	probability density function w.r.t. x with known parameter (μ,σ);
ER_t	expected revenue in period t;

EC_t	expected cost in period t;
EP_t	expected profit in period t;
$f_t(x_{S1}^{(t)}, x_{S2}^{(t)})$	expected maximum of the expected total profit from period t till final period;
$F(x, \mu, \sigma)$	cumulative distribution function w.r.t. x with known parameter (μ, σ);
$F^{-1}(x, \mu, \sigma)$	inverse function of $F(x, \mu, \sigma)$;
EP	expected profit in the single period;
MEP	maximum expected profit in final period;
MEC	minimum expected cost in final period;
ETP_t	expected total profit from period t till final period;
$\tilde{ETP}_t$	approximation to ETP_t;
ETC_t	expected total cost from period t till final period;
$\tilde{ETC}_t$	approximation to ETC_t;
$ATP_k^{(t)}$	actual profit in period t at sample k of demands and returns;
$ATC_k^{(t)}$	actual cost in period t at sample k of demands and returns;
$u_j^{(t)}$	gradient of the cost-to-go function in period t w.r.t. order-up-to level of product j;
$grad_{j,k}^{(t)}$	sample gradient of $u_j^{(t)}$ at sample k of demands and returns.

Chapter 1 Introduction

1.1 Background

In the recent decades, the management of the flows, opposite to the conventional supply chain flows, is addressed in the emerging field of ‘Reverse Logistics’. The returns flow of products or goods from downstream entity to upstream entity in the supply chain is due to different reasons. Product recovery may initiate the returns flow from users to producers. The returns flow of unsold goods from retailers to manufacturers is another example. Furthermore, the returns flow of defective products or spare parts for repair is also in the field. As for the definition of ‘Reverse Logistics’, there are a few versions, based on different emphases.

According to a White Paper published by the Council of Logistics Management (CLM), Reverse Logistics is introduced as

“[…] the term often used to refer to the role of logistics in recycling, waste disposal, and management of hazardous materials; a broader perspective includes all issues relating to logistics activities carried out in source reduction, recycling, substitution, reuse of materials and disposal”. (Stock, 1992)

As defined by Fleischmann (2001), Reverse Logistics is the process of planning, implementing, and controlling the efficient, effective inbound flow and storage of secondary goods and related information opposite to the traditional supply chain direction for the purpose of recovering value or proper disposal.

According to Dowlatshahi (2005), Reverse Logistics is a $53 billion industry in the US alone. Costs derived from reverse-logistics activities in the US exceed $35 billion per year. The customer returns rate may be as high as 15% of sales, and in sectors such as catalogue sales and e-commerce, it could reach as much as 35%. The following are the most frequently cited reasons for companies to engage in Reverse Logistics (Thierry, Salomon, Van Nunen, & Van Wassenhove, 1995; De Brito & Dekker, 2004; Ravi, Shankar, & Tiwari, 2005):

- Economic reasons, both direct (consumption of raw materials, reduction of disposal costs, recovery of the added value of used products, etc.) and indirect (an environmentally friendly image and compliance with current or future legislation);
- Legal reasons, because current legislation in many countries (including, for example, members of the European Union) holds companies responsible for recovering or properly disposing of the products they put on the market;
- Social reasons, because society is aware of environmental issues and demands that companies behave more respectfully towards the natural environment, especially with regard to issues like emissions and the generation of waste.

The above drivers are closely linked with the available options for recovering value from the products under consideration. Product recovery management may be defined as ‘the management of all used and discarded products, components, and materials for which a manufacturing company is legally, contractually, or otherwise responsible’ (Thierry et al., 1995). According to the re-entry point in the value adding process, there are the following forms of recovery:

- Repair. Products are brought to working order. This implies that typically the quality standard of repaired products is less than those for new products. Usually repair requires minor (dis)assembly, since only the non-working parts are repaired or replaced.
- Refurbishing. Products are upgraded to some pre-specified quality standards. Typically these standards are less than those for new products but higher than those for repaired products.
- Remanufacturing. Used products are recovered such that the quality standards are as strict as those for new products. Necessary disassembly, over-haul, and replacement operations are carried out in the recovery process.
- Cannibalization. This involves selective disassembly of used products and inspection of potentially reusable parts. Parts obtained from cannibalization can be reused in the repair, refurbishing or remanufacturing process.
- Recycling. Materials rather than products are recovered. These materials are reused in the manufacturing of new products.
- Disposal. Products are disposed of in the form of landfilling or incineration.

In the above categorization, the forms of refurbishing and cannibalization are also referred to as reuse. Refurbishing is denoting the reuse at the product level, whereas cannibalization is at the part level. Figure 1.1 describes the Reverse Logistics involving reuse, remanufacturing and recycling.

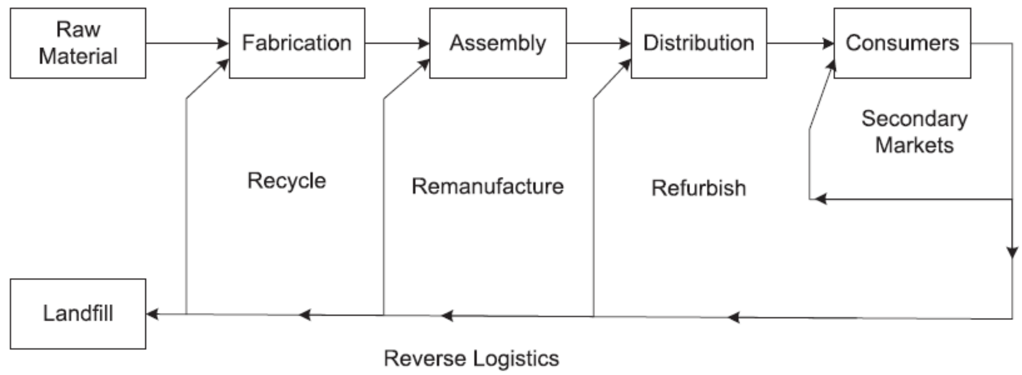


Figure 1.1 Reuse, remanufacturing and recycling in reverse logistics

As the inbound flows of product recovery management, the returns flows are distinguished as follows:

- End-of-use returns. Products are returned when they have reached the end of usage or lease period by customers. Remanufacturing and recycling are the major recovery options for them.
- Commercial returns. Products are returned by the buyer to the original sender for refunding. Reuse, remanufacturing, recycling and disposal are possible recovery options for them.
- Warranty returns. Products failing during use or damaged during delivery, spare parts, and product recalls due to security hazards are included in this category. Repair and disposal are possible recovery options for them.
- Production scrap and by-products. Excess material is reintroduced in the production process. By-products are often transferred to alternative supply chain. Recycling and remanufacturing are possible recovery options for them.
- Packaging. Crates, refillable bottles, pallets, reusable boxes and containers are best known examples in this category. Mostly, reusable

packaging is owned by logistics service providers who take charge of the recollection. Reuse and recycling are possible recovery options for them.

A growing number of industries are now becoming interested in remanufacturing of end-of-use returns. Nowadays, products that can be remanufactured might include machine tools, medical instruments, copiers, automobile parts, computers, office furniture, mass transit, aircraft, tires etc. Table 1.1 lists some large companies within these industries that currently apply product remanufacturing.

Table 1.1 Some companies active in remanufacturing

Company name	**Product**	**References**
Abbott Laboratories	Medical diagnostic instruments	Sivinski and Meegan (1993)
BMW	Car engines, starting motors, alternators	Vandermerwe and Oliff (1991)
De Vlieg-Bullard	Machine tools	Sprow (1992)
Grumman	F-14 aircraft	Kandebo (1990)
Rank Xerox	Copiers	Thierry et al. (1995)
Volkswagen Canada	Car engines	Brayman (1992)

Reverse Logistics has also attracted the attention from academia in recent years (Prahinski & Kocabasoglu, 2006). The research in the field of Reverse Logistics has covered three aspects: design of network structure for collecting the returned products, joint inventory management of recoverable products and serviceable products, operational planning of recovery process and normal production (Fleischmann et al., 1997). Among these aspects, many of the studies published on Reverse Logistics have focused on the inventory management of recoverable products and serviceable products (Rubio, Chamorro, & Miranda, 2008). Some of the most notable works have analyzed the effects of the returns flow on traditional inventory-management models (see, for example, Fleischmann et al., 1997; De Brito & Dekker, 2003; Minner, 2003; Fleischmann & Minner, 2004, for a review). Most of them are carried out on the basis of product recovery system, which undertakes the recovery process of returned products or goods. In many cases, the product recovery system also includes normal production of finished product. In practice, the product recovery system is often implemented as the remanufacturing of end-of-use returns.

According to whether inventory of returned products is allowable, product recovery system is classified into autonomous recovery system and managed recovery system. The autonomous recovery system only contains the inventory of finished product. Once returned products enter the system, they are immediately put into the recovery process. Thus, simple Push-strategy is applicable to this kind of system. However, the managed recovery system contains inventories of both returned product and finished product. Study on this kind of two-echelon inventory system is more complex.

In another aspect, product recovery system is classified according to differentiation of the returns flow or demands flow. In practice, the returned products are categorized according to different criteria, such as quality condition. Thus, the returns flow is divided. On the other hand, the demands flow is divided according to different customer segments, service levels, etc. For different demands flows, different recovery options are taken advantage of. Single-return-flow and single-demand-flow recovery system has been widely studied in the field. There are also few studies on single-return-flow and multi-demand-flow recovery system. A more detailed literature review on product recovery system modeling is given in Chapter 2.

However, to the latest knowledge, multi-return-flow and multi-demand-flow recovery system is almost not investigated.

Production planning and inventory control of the product recovery system has been attracting more research efforts. Many articles have appeared to explore the structure of the optimal policy or propose better heuristic policy for the product recovery system. In particular, we would review some important periodic review models here, which are related to our research. More details could be referred to in Chapter 2.

Simpson (1978) proposes an inventory model based on fixed periodic review of a product recovery system with single product and single flow of returned items, and finds out the optimum solution structure for the multi-period problem. Inderfurth (1997) extends Simpson's model by considering the impact of non-zero lead times both for production and recovery. Kiesmüller and Scherer (2003), present a method for the exact computation of the parameters which determine the optimal periodic policy in Simpson (1978). DeCroix (2006) extends Simpson (1978) and Inderfurth (1997) studies by identifying the structure of the optimal remanufacturing/ordering/disposal policy for a system where used products are returned to a recovery facility. Inderfurth (2001) presents a periodic review model for product recovery in stochastic remanufacturing systems with multiple reuse options, including a disposal option and incorporating uncertainties in returns and demands for the different serviceable options. Teunter (2002) considers a class of ordering policies and proposes EOQ (Economic Order Quantity) formulae (on the basis of the results proposed by Teunter, 2001) that are applicable to inventory systems with discounted costs and with stochastic demand and return. DeCroix et al. (2005) propose a stochastic periodic review model of multistage system with stationary costs and stochastic demand over an infinite horizon. Ahiska and King (2010) discuss inventory optimization of a periodically reviewed single-product stochastic manufacturing/remanufacturing system with two stocking points (recoverable and serviceable inventories) developing a stochastic review period model by using Markov Decision Processes.

From the aforementioned literature, we can find that most work is on single-product recovery system involving a single returns flow and a single demands flow. Only Inderfurth (2001) considers multiple reuse options for multiple demands flows. However, the study on the product recovery involving multiple products and thus multiple demands flows is of practical value.

Many high-tech products, such as personal computers, copiers etc., have very short lifecycle. For their Original Equipment Manufacturers (OEMs) responsible for taking care of the end-of-use returns, well-implemented product recovery system is of much importance to both economical earnings and marketing image of the manufacturers. The product recovery system is required to be capable of dealing with the recovery of multiple products, which belong to the same product family. The returned items of each product can be recovered to finished items of any product at different cost.

In addition, Behret and Korugan (2009) construct a simulation model by using the ARENA simulation program to analyze the effect of quality classification of returned products, and find out that quality-based classification of returned products could result in significant cost savings especially when return rates are high. Therefore, the returned items of all the products are discriminated into multiple groups by different quality conditions or different cost requirements in the recovery process.

1.2 Scope and Purpose of the study

From the aforementioned literature, we can find that most work is on single-product recovery system involving a single returns flow and a single demands flow. Only Inderfurth (2001) considers multiple reuse options for multiple demands flows. However, the study on the product recovery involving multiple products and thus multiple demands flows is of practical value. As one of the multi-product cases, the two-product case is easy to be implemented and could be the basis for the study on a general multi-product case. Therefore, a product recovery system involving two products is selected for this research. In addition, Behret and Korugan (2009) find that quality-based classification of returned products could result in significant cost savings. Thus, in the two-product recovery system studied, we classify the returned items of the two products into two groups by quality in contrast to most work disregarding this classification in the literature.

As far as the knowledge about current literature, this research could be regarded as the first study on the multi-product recovery system involving two products and two flows of returned items. Furthermore, the extension of this research to a general multi-product recovery system is also discussed.

This research aims to obtain the optimal or near-optimal periodic review policy over a finite horizon for the inventory control of a two-product recovery system involving two products and two returns flows.

A dynamic programming model has been developed in order to obtain the optimal policy of production and recovery decisions. However, the model is difficult to be solved efficiently as no nice property could be found. Thus, the special case of the multi-period problem, a single-period problem is investigated. The optimal multi-level threshold policy has been obtained by solving KKT conditions for the single-period problem. Even though this multi-level threshold policy might not be optimal for the multi-period problem, it is intuitive, easy to use and provides good managerial perspectives. Hence, we apply this policy to the multi-period problem. It is further investigated how to compute the threshold levels, which depend not only on the current-period cost parameters but also the future cost-to-go function. We have developed an approximate dynamic programming model to derive the threshold levels in the multi-period situation. The performance of the threshold policy is proved to be good by comparing with the other two heuristic policies from the single-period problem under a wide range of settings.

1.3 Organization

The organization of this thesis is as follows. Chapter 2 reviews the research literature on product recovery system in the field of Reverse Logistics. Chapter 3 describes the two-product recovery system in a finite horizon. A dynamic programming model on this system is developed. Chapter 4 studies the two-product recovery system in a single period. Some good properties on the model of the system are proved. The optimal multi-level threshold policy of production and recovery decisions are obtained by solving KKT conditions. Furthermore, the managerial insights of the policy are provided. In addition, the extension from the two-product situation to a general multi-product situation is discussed. The multi-level threshold policy is assumed to be used for the multi-period problem. Chapter 5 focuses on the study of the two-product recovery system in the situation of lost sales over a finite horizon. An ADP model on the system is developed to help derive the threshold levels. This multi-level threshold policy is compared with two heuristic policies derived from the optimal policy of the single-period problem. In addition, the impact of system parameters is investigated. Chapter 6 and Chapter 7 consider the two-product recovery system in the situation of backorder over a finite horizon. In particular, Chapter 7 investigates the lead time effect of production and recovery processes. Chapter 8 provides a summary of the findings and proposes several possible directions for the future research.

Chapter 2 Literature review

Chapter 2 reviews the previous studies on production and inventory control of product recovery system in the field of Reverse Logistics. Section 2.1 presents a classification table with the objective of intelligibly describing the papers. The studies on production and inventory control of product recovery system with single return flow and single demand flow will be reviewed in Section 2.2. Section 2.3 introduces the studies on production and inventory control of product recovery system with multiple flows of returns or multiple flows of demands or both.

The research in the field of Reverse Logistics has covered three aspects: design of network structure for collecting the returned products, joint inventory management of recoverable products and serviceable products, operational planning of recovery process and normal production (Fleischmann et al., 1997). This branching is due to the stages of reverse logistics activities. From the other perspectives, Reverse Logistics covers green supply chain, closed-loop supply chain etc. Various closed-loop supply chain processes and modeling framework of the closed-loop supply chain are presented (see, for example, Ferguson, M., Souza, G., 2010; Ferguson, M., 2010; Drake, M.J., Ferguson M., 2008, for a review). Paksoy et al. (2011) investigate a number of operational and environmental performance measures, in particular those related to transportation operations, within a closed-loop supply chain.

However, we would focus on production planning and inventory management of product recovery system in the literature review.

2.1 Classification

There are considerable amounts of research work on production planning and inventory management of product recovery system. Hence, it is helpful to provide a classification table, which is used to describe the papers that will be reviewed in the following sections. A general overview of Operations Management problems associated with product recovery is provided in Thierry et al. (1995). A review of quantitative models in the field of reverse logistics is given by Fleischmann et al. (1997). A review of environmentally conscious manufacturing and product recovery is given by Ilgin et al. (2010).

Table 2.1 Legend for classification system

Elements	**Descriptions**
Length of horizon	Single period/Multiple periods/Infinite horizon
Demand type	Deterministic demand/Stochastic demand
Review policy	Periodic review/Continuous review
Sales	Backorder/Lost sale
Products	Single product/Multiple products

2.2 Product recovery system with single return flow and single demand flow

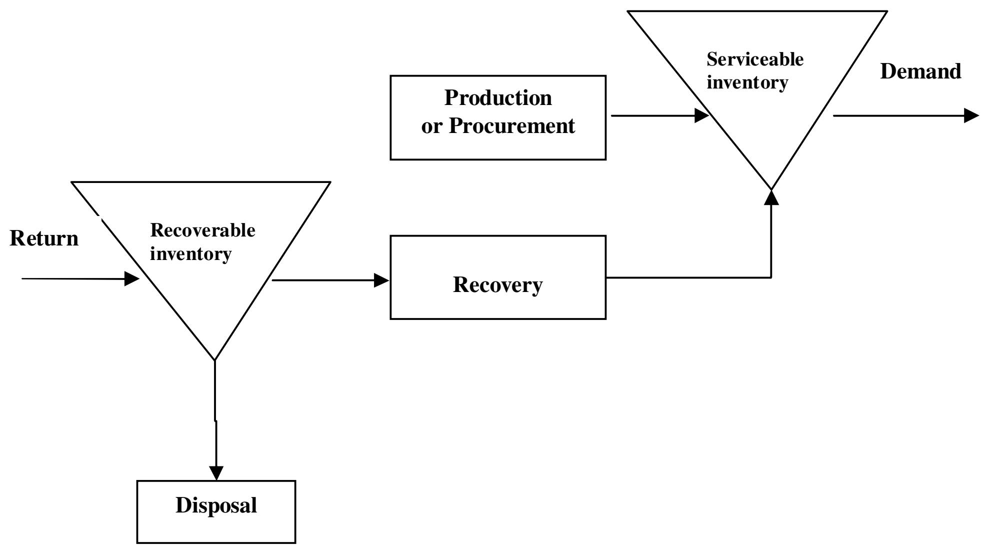


Figure 2.1 Product recovery system with single return flow and single demand flow (adapted from Fleischmann et al., 1997)

In this section we review literature concerning quantitative inventory control models of product recovery system with single return flow and single demand flow, which are independent of each other. From a mathematical inventory theory perspective, deterministic and stochastic models can be distinguished, and the latter can be further subdivided into continuous and periodic review models. We treat each of these groups separately below.

2.2.1 Deterministic models

In deterministic models, the demand flow and return flow are known a priori for the entire planning horizon. Using the taxonomy of inventory theory, Table 2.2

lists deterministic models from literature. For each model, the planning horizon, and the cost criterion of the objective function are indicated. Some models explicitly consider the two types of inventory distinguished in Figure 2.1, whereas others take into account only a single aggregated stock point. Moreover, disposal of excess returns may or may not be allowed. In addition, fixed costs and lead times may or may not be included in the recovery system considered. Table 2.2 has listed some papers with their model characteristics.

Table 2.2 Deterministic inventory models of product recovery system

Literature	Planning horizon	Cost criterion	Number of stock points	Disposal	Fixed costs	Lead times
Schrady (1967)	∞	Avg	2	-	+	+
Mabini et al. (1992)	∞	Avg	2	-	+	+
Richter (1994)	∞	Avg	2	+	+	-
Richter (1996, 1997)	∞	Avg	2	+	+	-
Teunter (2001)	∞	Avg	2	+	+	-
Richter and Sombrutzki (2000)	T	Total	1	-	+	-
Beltran and Krass (2002)	T	Total	1	+	+	-
Minner and Kleber (2001)	T	Total	2	+	-	-

Schrady (1967) first extended the classical economic order quantity (EOQ) model by taking return flow into account. The model is developed on the product recovery system with constant demand rate and return rate, and fixed lead times for production and recovery processes. Disposal is not allowed. The costs considered are fixed setup costs for production and recovery processes and linear inventory holding costs for returned products and finished products. A control policy was proposed with fixed lotsizes for production and recovery where each production batch is followed by n identical recovery batches. The formulae on the optimal value of n and on the optimal lotsizes are derived similar to the classical EOQ model.

Richter (1994) considered Schrady's model for alternating production and recovery batches (i.e. $n = 1$ in the above setting) and analyzed the dependence of the cost function on the return rate. He shows that costs are convex in the return rate if holding costs for recoverables do not exceed serviceable holding costs. Richter (1996, 1997) extends the analysis to the case of multiple consecutive production and recovery batches.

Teunter (2001) considered the same model for a modified disposal policy. The model assumes that all returns occurring during a certain time span are disposed, while all returns thereafter are accepted again. Disposal involves a linear cost per item. Moreover, it assumes different holding costs for recoverable, recovered, and produced items. The formulae on the optimal lotsizes in the policy are derived.

Koh et al. (2002) considered a joint EOQ and EPQ model assuming a proportion of the used products to be returned. They found closed form expressions for the economic order quantity for new products and the optimal inventory level where the recovery process starts. Further they proposed a numerical procedure, which calculates the optimal number of set-ups in both recovery and production processes. Konstantaras and Papachristos (2008) proposed another method to obtain the optimal number of set-ups and proved it to be more computationally efficient.

Besides the above static models, a few dynamic lotsizing models similar to the classical Wagner-Whitin model (Wagner and Whitin, 1958) have been proposed in the field. Most of these models consider a single stock point, which aggregates recoverable inventory and serviceable inventory.

Beltran and Krass (2002) considered dynamic lotsizing for a single stock point facing both demand and returns. This is regarded as the modification of the original Wagner-Whitin model by allowing negative (net) demand. The authors proposed a dynamic programming algorithm, which is of different complexity in the general case and restrictive case.

Richter and Weber (2001) extended the reverse Wagner-Whitin model to the case with additional variable manufacturing and remanufacturing cost. The authors proved the optimality of a policy starting with recovery before switching to production and gave an estimation for the optimal switching point. In addition, the impact of the disposal of excess inventory was investigated on the solution.

Minner and Kleber (2001) proposed an optimal control policy for the product recovery system, where in addition to demand and returns, all actions (production, recovery and disposal) are modeled as non-stationary continuous processes. Results are illustrated in a scenario with seasonality and a fixed time lag between demand and returns. Pontryagin's Maximum Principle is applied to obtain the optimal production and remanufacturing policies for deterministic but dynamic demands and returns when backorders are not allowed.

Kiesmüller (2003b) investigated a one product recovery system for dynamic and deterministic demand and return rates. The optimal production rate, recovery rate and disposal rate are determined for the system under the assumptions of a linear cost structure and zero lead time for production and recovery. Furthermore, the author showed how the results based on zero lead times were used to solve the control problems with positive lead times.

Teunter et al. (2006) study the dynamic lot sizing problem with product returns. The authors propose a model that aims at determining those lot sizes for manufacturing and remanufacturing by minimizing the total cost composed of holding cost for returns and (re)manufactured products and set-up costs.

Konstantaras and Papachristos (2007) propose a single product recovery and a periodic review inventory model with finite horizon and remanufacturing, manufacturing options. Demand is satisfied only by remanufactured or by newly manufactured products. They aim at identifying an optimal policy that specifies the period of switching from remanufacturing to manufacturing, the periods where remanufacturing and manufacturing activities take place and the corresponding lot sizes.

2.2.2 Continuous review stochastic models

Most continuous review models on the product recovery system are stationary and analyze the infinite horizon system behavior. They focus on determining optimal parameter values for predetermined control policies. In almost all cases, demand and returns are modeled as independent Poisson processes. Table 2.3 has listed some papers with their model characteristics.

Table 2.3 Continuous review inventory models of product recovery system

Literature	Planning horizon	Cost criterion	Number of stock points	Disposal	Fixed costs	Lead times
Heyman (1977)	∞	Avg	1	+	-	-
Muckstadt and Iscaac (1981)	∞	Avg	1	-	+	+
Van der Laan et al (1996a,b)	∞	Avg	1	+	+	+
Yuan and Cheung (1998)	∞	Avg	1	-	+	-
Teunter (2002)	∞	NPV	2	-	+	-
Van der Laan et al (1999a,b)	∞	Avg	2	-	+	+
Van der Laan and Salomon (1997)	∞	Avg	2	+	+	+
Inderfurth and van der Laan (1998)	∞	Avg	2	+	+	+

The proposed models can be divided into two groups, with one considering a single aggregated stock point and the other distinguishing recoverable and serviceable inventories. Within the former class, Heyman (1977) analyzed disposal policies to optimize the trade-off between additional inventory holding costs and production cost savings. The demand and returns are modeled as general independent compound renewal processes. He proposed a single parameter disposal level strategy: incoming returns exceeding this level are disposed of. For the case of Poisson distributed demands and returns, he derived an explicit expression for the optimal disposal level. For generally distributed demands and returns, an approximation is given.

Muckstadt and Isaac (1981) considered a similar model where the recovery process is modeled as a multi-server queue. However, disposal decisions are not taken into account. The costs considered comprise serviceable holding costs, backorder costs, and fixed production costs. The production process is controlled by a traditional (*s, Q*)-rule whereas returned products directly enter the recovery queue. Values for *s* and *Q* are determined based on an approximation of the distribution of the net inventory.

Van der Laan et al. (1996a, b) proposed an alternative procedure for determining the control parameters in the above (*s, Q*)-model based on an approximation of the distribution of the net demand during the production lead time. A numerical comparison shows this approach to be more accurate in many cases. Moreover, the model is extended with a disposal option, for which several policies are compared numerically.

Yuan and Cheung (1998) model dependent demand and returns by assuming an exponentially distributed market sojourn time. Moreover, items may eventually be lost with a certain probability. Lead times for both recovery and production are zero and there is no disposal option. The authors proposed an (*s, S*) reorder-order-up-to policy for production based on the sum of items on hand and in the market. The long-run average costs by this policy are calculated based on a two-dimensional Markov process. A numerical search algorithm is proposed for finding optimal control parameter values.

Van der Laan and Teunter (2006) considered a product recovery system including manufacturing and remanufacturing, both of which have equal non-zero lead times. The cost structure consists of setup costs, holding costs, and backorder costs. The system is controlled by certain extensions of (*s, Q*) policy, called push and pull remanufacturing policies. For all policies, the authors presented simple, closed-form formulae for approximating the optimal policy parameters under a cost minimization objective.

Ouyang and Zhu (2008) extended traditional (*s, Q*) model into (s_p, Q, s_d) order-disposal strategy to control the manufacturing/remanufacturing hybrid system assuming demand and returns to be independent Poisson processes. They derived the expression of the system expected total cost per unit time as a function of the control parameters s_p, Q and s_d. They developed heuristic lower and upper bounds for the optimal solution. They compared the disposal strategy with the non-disposal strategy and investigated the robustness of the optimal solution through the numerical examples.

Teunter (2002) distinguished serviceable and recoverable inventory and evaluated an EOQ-based heuristic under assuming demand and returns to be independent Poisson processes. Lotsizes for production and recovery are determined in a deterministic model (see Teunter, 2001, discussed above). Teunter and Vlachos (2002) investigated the impact of a disposal option for a similar situation. They concluded that only under certain circumstances, the disposal option can bring economic benefits.

Van der Laan et al. (1999a, b) analyzed different policies for controlling serviceable and recoverable inventory in the above setting, considering non-zero lead times for production and recovery. In particular, a Push-strategy and a Pull-strategy for recovery are considered while production is controlled by an (*s, S*)-policy concerning the serviceable inventory position (serviceable inventory on hand minus backorders plus outstanding orders). The Pull-strategy-based recovery is also controlled by an (s, S) policy based on the serviceable inventory position. Long-run expected costs for both strategies are computed by evaluating a two-dimensional Markov process. Control parameter values are determined via enumeration. Furthermore, Inderfurth and van der Laan (2001) improved the above model with a modified inventory position used for the case of a large difference between production lead time and recovery lead time. The modification for the inventory position is that only those outstanding orders are considered within a certain time window.

Van der Laan and Salomon (1997) extended the above model to include a disposal option. For the Pull-strategy, the disposal is triggered by an upper bound on the recoverable inventory. However, for the Push-strategy, since the recoverable inventory is limited by the recovery lotsize, the disposal is controlled based on the serviceable inventory position. The authors showed that a disposal option significantly reduces the system costs by avoiding excessive stock in particular for large return rates.

2.2.3 Periodic review stochastic models

The models within this context aim to seek an optimal periodic review policy for production, recovery, and/or disposal decisions. The models can be distinguished by considering an aggregated stock point or both recoverable inventory and serviceable inventory. Within the former class, models differ mainly with respect to the assumptions on the relation between demand and returns. Table 2.4 has listed some papers with their model characteristics.

Table 2.4 Periodic review inventory models of product recovery system

Literature	Planning horizon	Cost criterion	Number of stock points	Disposal	Fixed costs	Lead times
Whisler (1967)	T/∞	Avg	1	+	-	-
Kelle and Silver (1989)	T	Total	1	-	+	-
Toktay et al (1999)	∞	Avg	1	-	-	+
Buchanan and Abad (1998)	T	Total	1	-	-	-
Cohen et al (1980)	T	Avg	1	-	-	-
Inderfurth et al (1998)	T/∞	Avg	N+1	+	-	+

Whisler (1967) analyzed a single stock point receiving issued item returns after a stochastic market sojourn time and constructed a queueing model. The optimal control policy was found to be characterized by two critical numbers $L < U$. Whenever the inventory level at a review epoch lies outside the interval $[L, U]$ it is optimal to produce up to L or dispose down to U, respectively. For intermediate inventory levels the optimal production and disposal decisions depend on additional parameters.

Kelle and Silver (1989) analyzed a similar situation where issued items are returned after a stochastic time lag or are lost eventually. Thus, due to positive average net demand, no disposal option is included. On the other hand, fixed production costs are considered. The authors formulated a chance-constrained integer program, which can be transformed into a dynamic lotsizing model with possibly negative demand, based on an approximation of the cumulative net demand.

Buchanan and Abad (1998) modified the above model by assuming for each period that returns are a stochastic fraction of the number of items in the market. This comes down to an exponentially distributed market sojourn time. Moreover, in each period a fixed fraction of items from the market is lost. Under these conditions the authors derived an optimal production policy depending on two state variables, namely the on-hand inventory and the number of items in the market.

Cohen et al. (1980) considered a similar system assuming a fixed market sojourn time. Moreover, a given fraction of demand in each period will not be returned. In addition, a certain fraction of on-hand inventory is lost due to decay in each period. The authors proposed a heuristic order-up-to policy which is optimal for the case of a market sojourn time of one period.

Simpson (1970) assumed demand and returns to be independent with a positive expected net demand. He proposed a heuristic for computing an order-up-to level under linear costs and a stochastic production lead time when neglecting intermediate backorders cleared by returns.

Mahadevan et al. (2003) modeled a product recovery system in the remanufacturing context assuming demand and returns to be independent Poisson processes. Taking no disposal into account, they applied a Push-strategy to combining production and remanufacturing decisions. They developed several heuristics based on traditional inventory models and investigated the performance of the system as a function of return rates, backorder costs, and lead times of production and remanufacturing. In addition, the lower and upper bounds on the optimal solution were developed.

Kiesmüller and Van der Laan (2001) considered the impact of dependency between demand and returns, which are assumed to be Poisson processes. The returns are correlated to the demand through a constant sojourn time in the market. In addition, two probabilities are introduced: the return probability and the probability that a returned item is in a sufficiently good condition to be recovered. By comparing the performance on the total average costs with the models neglecting the dependency, the authors suggested that it was worth to use the dependency information between demand and returns.

A special class of periodic review models considering product returns is newsboy models. Vlachos and Dekker (2003) assumed a constant fraction of the sold items to be returned and re-sold only once. In Mostard and Teunter (2006), each sold item has a constant probability of being returned and once returned it has a constant probability of being recovered. Returned items can be re-sold more than once. In the above two models the optimal order quantity for the single period was sought.

In addition, within the context of models distinguishing recoverable inventory and serviceable inventory, Simpson (1978) first considered the trade-off between material savings due to reuse of returned products versus additional inventory holding costs. Demand and returns of each period are modeled as generally distributed random variables that are correlated with known information. Optimality of a three parameter (*L, M, U*) policy to control production, recovery, and disposal is shown when neither fixed costs nor lead time are involved. The policy can be interpreted as 'recover while serviceable inventory is below *M*' and then adjust the echelon inventory (i.e. the sum of both recoverable inventory and serviceable inventory) according to Whisler's (*L, U*)-policy. Kiesmüller and Scherer (2003) provided a method for the exact computation of the parameters in the (*L, M, U*) policy. Since the exact computation is quite time consuming, they also provided two different approximations. One is based on an approximation of the value-function in the dynamic programming problem while the other approximation is based on a deterministic model.

Ahiska and King (2010) discuss inventory optimization of a periodically reviewed single product stochastic manufacturing/remanufacturing system with two stocking points (recoverable and serviceable inventories) developing a stochastic review period model by using Markov Decision Processes.

Inderfurth (1997) extended Simpson's model by considering the impact of non-zero lead times both for production and recovery. The difference between both lead times was shown to determine the system complexity. If lead times are equal Simpson's policy can be shown to remain optimal by considering an appropriate inventory position rather than the net stock. In all other cases growing dimensionality of the underlying Markov model prohibits simple optimal control rules. A similar result holds if recoverables cannot be stored but need to be recovered or disposed of immediately. In this case Whisler's (*L, U*)-policy is optimal for equal lead times and for a production lead time excess of one period. All other cases result again in fairly intractable situations.

DeCroix (2006) extends Simpson (1978) and Inderfurth (1997) studies by identifying the structure of the optimal remanufacturing/ordering/disposal policy for a system where used products are returned to a recovery facility. In particular, the author analyzes a multiechelon inventory system with inventory stages arranged in series. DeCroix et al. (2005) propose a stochastic periodic review model of multistage system with stationary costs and stochastic demand over an infinite horizon. Note that in the model, demand can be negative representing returns from customers. The authors also show the optimality of an echelon basestock policy for an infinite-horizon series system where returns go directly to stock.

Kiesmüller (2003a) considered similar situation where production lead time and recovery lead time are different. The recovery system is controlled by (*S, M*)-policy, described by produce-up-to level *S* and remanufacture-up-to level *M*. In contrast with previous models using inventory position as information for decision making, the author defined two variables aggregating related information for production and recovery decisions respectively. The two variables are dependent on the lead time and include all information about outstanding production and recovery orders which will arrive before the new released order. By means of numerical examples, the system performance, measured in average costs per time unit, can be improved substantially especially for large lead time differences.

2.3 Product recovery system with single return flow and multiple demand flows

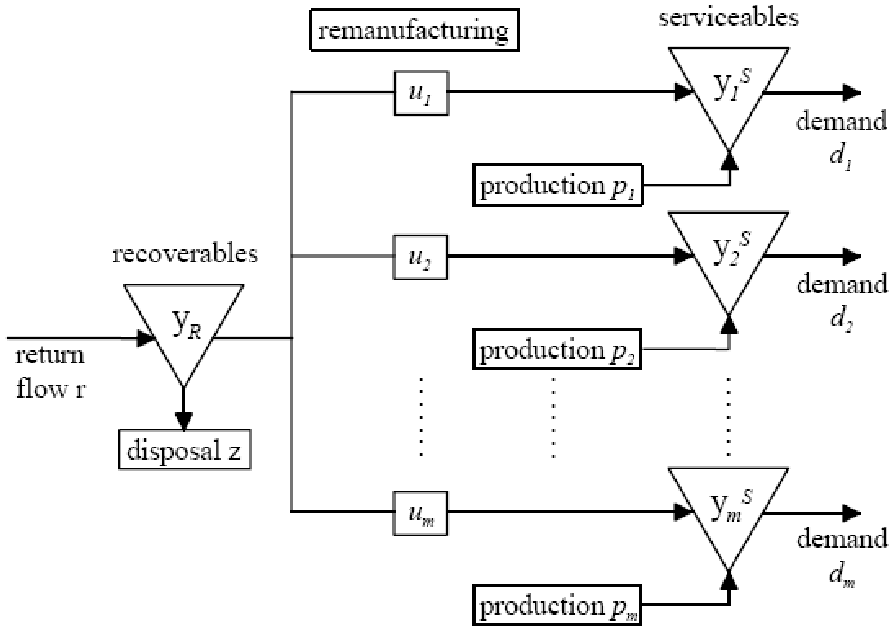


Figure 2.2 Product recovery system with single return flow and multiple demand flows (adapted from Kleber et al., 2002)

In many situations, there are different options of reusing old products. In more details, old products are reused for making new products or spare parts. The new products from reusing old products are for different customer classes having different quality requirements. Therefore, multiple demand flows are possibly included in the product recovery system.

Kleber et al. (2002) presented a continuous model of a product recovery system with returns of a single product and multiple alternating recovery options. Each of the recovery options corresponds to different demand classes, e.g. different product qualities or different markets. In the model, disposal option is included and production option is alternative for each recovery option. The product recovery system is described in Figure 2.2. Demands and returns are assumed to be deterministic but dynamic. The optimal policy of production, recovery and disposal is obtained under linear cost by applying Pontryagin's Maximum Principle.

Inderfurth et al. (2001) investigated a periodic review model of a product recovery system with returns of a single product and multiple alternating recovery options. The system also includes one disposal option. There is no production option as alternation for each recovery option. Backorder is allowed in the system. Demands and returns are assumed to be stochastic with known probability distribution. Taking advantage of a linear allocation rule on product returns, they obtained a fairly simple near-optimal policy, characterized by a dispose-down-to level and a recover-up-to level for each recovery option.

The above-presented two models have considered single returns flow, i.e. returns of a single product. However, in practice, returns of even a single product can vary in quality condition. Then, multiple flows of returns have to be considered, each representing a certain quality class. Behret and Korugan (2009) constructed a simulation model by using the ARENA simulation program to analyze the effect of quality classification of returned products. The analysis denotes that under different cost scenarios quality based classification of returned products have brought significant cost savings, especially when return rates are high.

To recent knowledge of product recovery systems in the literature, multiple flows of returns and multiple flows of demands are not appearing at the same time in these product recovery systems. However, as of the practical value, the product recovery system needs to be studied, which includes both multiple flows of returns and multiple flows of demands. Due to this, the following chapters will focus on the study of two-product recovery system. The returned items of the two products are discriminated into two groups by the required cost in recovery of them rather than by product type. In addition to recovery options using two groups of returned items, the two-product recovery system includes production option to make finished items of two products in order to satisfy customer demands. More details of the two-product recovery system will be introduced in Chapter 3. Obviously, the two-product recovery system includes two flows of returns and two flows of demands.

Chapter 3 The study on two-product recovery system in a finite horizon

Chapter 3 focuses on the study of the two-product recovery system in a finite planning horizon. In the system, the stocking of the two products aims to satisfy stochastic customer demands in each period of the planning horizon. The inventory of the two products can be instantly replenished by production and recovery process as both processes are assumed to have zero lead time. When the system is short of inventory in a certain period, the sale would be either lost forever or allowed to be backordered in future periods. Section 3.1 introduces the details of the two-product recovery system in a finite horizon. In Section 3.2, a dynamic programming model of the system is developed in order to maximize the expected total profit in the finite horizon. Finally, Section 3.3 summarizes the main work in this chapter.

3.1 Introduction

Two products, which belong to the same product family, are provided to customers by an Original Equipment Manufacturer. At the same time, the manufacturer is required to take responsibility of dealing with returned products, which have reached the end of the usage at customers. The manufacturer would take advantage of the returned products in the recovery for finished products, which are assumed to be as good as those from normal production. The manufacturer would build up the two-product recovery system, in which both recovery process and normal production are used to make finished products. As the two products belong to the same product family, returned item of each old product can be recovered to finished item of both new products. Therefore, the returned items of both old products regardless of product identity are discriminated into two groups by quality or the cost requirement in the process of value extraction and recovery. The returned items of each group are assumed to consume the same cost in the recovery for finished items of any certain product. After the discrimination, one group of returned items is always recovered at lower cost than the other. In addition, normal production is more costly than recovery such that normal production would be only used in case of insufficient returned items available for recovery.

In particular, the occurring events in each period of the finite horizon are described here. Firstly, returned items arrive at the recovery system at the beginning of each period. They will be used for recovery in this period. Secondly, after observing on-hand inventories of finished products, the manufacturer would make production and recovery decisions. After that, the inventories of finished products get replenished instantly. The inventories are used to satisfy demands later in the same period. If the demands of current period could not be fully satisfied, the unsatisfied demands would be either lost forever or allowed to be backordered in future periods. Anyway, the penalty cost on the shortage would be incurred. On the other hand, if there are inventories left at the end of the period, the remaining inventories would be carried to subsequent periods and inventory holding cost would be counted in current period. Finally, the remaining returned items are disposed of and the disposal cost is assumed to be negligible. For the recovery system, the revenue is generated from selling finished products. The total cost consists of production cost, recovery cost, inventory holding cost of finished products and penalty cost of shortage.

The objective function is to maximize the expected total profit in a finite horizon. In order to fulfill the aim, we need to find the optimal policy of production planning and inventory control for the system.

3.2 Production and recovery decisions for two products in the multi-period context

3.2.1 Assumptions and notations

Firstly, in order to focus on the interesting aspects of the system and also simplify the modeling, we would make the following assumptions. The relaxation of certain assumptions has been discussed in Chapter 8.

1) Demands for the two products follow independent stationary general distributions;
2) Production and recovery processes of each product have zero lead time;
3) No setup cost is considered for production or recovery process of each product;
4) One unit of returned product is recovered to one unit of finished product;
5) No disposal cost or salvage value is considered for the remaining returned products;
6) No stocking of the remaining returned products is required in each period.

Secondly, in order to simplify the notation, the two products in the system are denoted as product 1 and product 2 respectively. In addition, the returned items of two products are discriminated into two groups, denoted as group 1 and group 2 respectively. Without the loss of generality, returned item in group 2 is assumed to be recovered at lower cost than group 1 for each product. Thus, the related notations are listed as follows (i = 1, 2; j = 1, 2):

M	length of planning horizon;
$r_{ij}^{(t)}$	quantity of recovering returned item in group i to product j in period t;
$p_j^{(t)}$	quantity of producing product j in period t;
$x_{Sj}^{(t)}$	initial inventory position of product j in period t;
s_j	selling price of product j;
c_{Rij}	unit cost of recovering returned item in group i to product j;
c_{Pj}	production cost of per unit product j;
h_j	inventory holding cost of per unit product j per period;
v_j	penalty cost of per unit shortage of product j per period;
$R_i^{(t)}$	returned items in group i in period t;
$D_j^{(t)}$	demand for product j in period t;
$f(x,\mu,\sigma)$	probability density function w.r.t. x with known parameter (μ,σ);
ER_t	expected revenue in period t;
EC_t	expected cost in period t;
EP_t	expected profit in period t;
$f_t(x_{S1}^{(t)}, x_{S2}^{(t)})$	expected maximum of the expected total profit from period t till final period.

The two-product recovery system is described in Figure 3.1. The Figure has shown inbound flows of returned items in group 1 and group 2, and outbound flows of product 1 and product 2 in demand. In particular, the inbound flow of returned items in group 1 and group 2 respectively, are shown on the upper and lower left of this Figure. On the other hand, the outbound flow of product 1 and product 2 respectively, are shown on the upper and lower right of this Figure. In addition, it can be seen from the Figure that there is no stocking of returned items in the system. Once the returned items have been allocated to the recovery for finished products, the remaining returned items would be disposed of. In the Figure, the time index (t) is omitted from the related notations for simplicity. All the notations shown in the Figure are related to the same period.

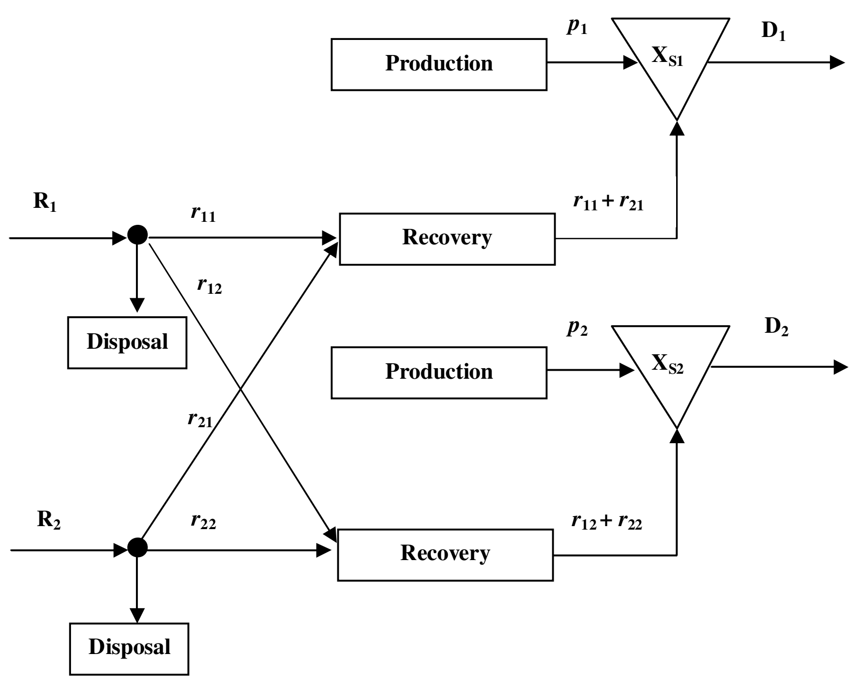


Figure 3.1 The structure of the two-product recovery system

The sequence of the occurring events and cost accounting of a certain period is described in Figure 3.2 as follows (*i*, *j* = 1, 2):

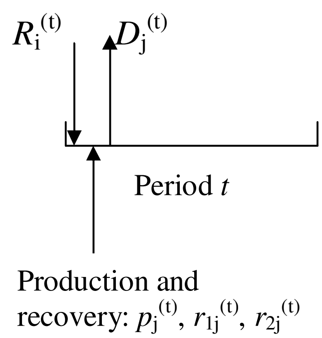


Figure 3.2 The occurring events of the two-product recovery system at period t

In addition, some restrictions on cost parameters are imposed so as to ensure the economical meaningfulness of the study. Firstly, for each product, selling price is higher than production cost, and penalty cost of shortage is higher than the profit from production. Therefore, there exist:

$$s_1 > c_{P1},\ \ s_2 > c_{P2},\ \ v_1 > s_1 - c_{P1},\ \ v_2 > s_2 - c_{P2}.$$

Secondly, for each product, production cost is higher than recovery cost. Otherwise, recovery is unnecessary. Therefore, there exist:

$$c_{P1} > c_{R11},\ \ c_{P1} > c_{R21},\ \ c_{P2} > c_{R12},\ \ c_{P2} > c_{R22}.$$

Finally, the recovery using returned item in group 2 is cheaper than that using returned item in group 1. Therefore, there exist:

$$c_{R21} < c_{R11},\ \ c_{R22} < c_{R12}.$$

3.2.2 Dynamic programming model of the two-product recovery system in the multi-period context

We would develop a dynamic programming model of the two-product recovery system in a finite horizon. The objective is to obtain the optimal policy of production and recovery decisions for the recovery system. Firstly, we consider the calculation of the expected profit in period *t*. At the beginning of the period, the system receives the returned items of two products. They are discriminated into group 1 and group 2, denoted as $R_1^{(t)}$ and $R_2^{(t)}$ respectively. The initial inventory position of product 1 and product 2 are found to be $x_{S1}^{(t)}$ and $x_{S2}^{(t)}$ respectively. Then, we would allocate the returned items to the recovery for finished products. At the same time, production would be used in case of insufficient returned items. Once production and recovery decisions have been made, the inventory of finished products would get replenished instantly. Then, the inventory would be used to satisfy the realization of stochastic demands in this period. The demands for the two products have been assumed to follow the known independent probability distributions. The expected revenue at period *t* is calculated as follows:

$$\begin{aligned}
&ER_t(x_{S1}^{(t)}, x_{S2}^{(t)}, p_1^{(t)}, p_2^{(t)}, r_{11}^{(t)}, r_{12}^{(t)}, r_{21}^{(t)}, r_{22}^{(t)}) \\
&= s_1\mu_1 + s_2\mu_2 - s_1\int_{x_{S1}^{(t)}+p_1^{(t)}+r_{11}^{(t)}+r_{21}^{(t)}}^{\infty} (D_1^{(t)} - x_{S1}^{(t)} - p_1^{(t)} - r_{11}^{(t)} - r_{21}^{(t)}))f(D_1^{(t)}, \mu_1, \sigma_1)dD_1^{(t)} \\
&\quad - s_2\int_{x_{S2}^{(t)}+p_2^{(t)}+r_{12}^{(t)}+r_{22}^{(t)}}^{\infty} (D_2^{(t)} - x_{S2}^{(t)} - p_2^{(t)} - r_{12}^{(t)} - r_{22}^{(t)})f(D_2^{(t)}, \mu_2, \sigma_2)dD_2^{(t)}.
\end{aligned} \tag{3.1}$$

The consumption of returned items in two groups is no more than their availability. Therefore, there are the following constraints:

$$r_{11}^{(t)} + r_{12}^{(t)} \le R_1^{(t)};$$
$$r_{21}^{(t)} + r_{22}^{(t)} \le R_2^{(t)}.$$

At period t, the cost includes production cost, recovery cost, inventory holding cost of finished products and penalty cost of shortage. Therefore, the expected cost of period t is calculated as follows:

$$\begin{aligned}
&EC_t(x_{S1}^{(t)}, x_{S2}^{(t)}, p_1^{(t)}, p_2^{(t)}, r_{11}^{(t)}, r_{12}^{(t)}, r_{21}^{(t)}, r_{22}^{(t)}) \\
&= c_{P1}p_1^{(t)} + c_{P2}p_2^{(t)} + c_{R11}r_{11}^{(t)} + c_{R12}r_{12}^{(t)} + c_{R21}r_{21}^{(t)} + c_{R22}r_{22}^{(t)} \\
&\quad + h_1\int_0^{x_{S1}^{(t)}+p_1^{(t)}+r_{11}^{(t)}+r_{21}^{(t)}} (x_{S1}^{(t)} + p_1^{(t)} + r_{11}^{(t)} + r_{21}^{(t)} - D_1^{(t)}) f(D_1^{(t)}, \mu_1, \sigma_1) dD_1^{(t)} \\
&\quad + v_1\int_{x_{S1}^{(t)}+p_1^{(t)}+r_{11}^{(t)}+r_{21}^{(t)}}^{\infty} (D_1^{(t)} - x_{S1}^{(t)} - p_1^{(t)} - r_{11}^{(t)} - r_{21}^{(t)}) f(D_1^{(t)}, \mu_1, \sigma_1) dD_1^{(t)} \\
&\quad + h_2\int_0^{x_{S2}^{(t)}+p_2^{(t)}+r_{12}^{(t)}+r_{22}^{(t)}} (x_{S2}^{(t)} + p_2^{(t)} + r_{12}^{(t)} + r_{22}^{(t)} - D_2^{(t)}) f(D_2^{(t)}, \mu_2, \sigma_2) dD_2^{(t)} \\
&\quad + v_2\int_{x_{S2}^{(t)}+p_2^{(t)}+r_{12}^{(t)}+r_{22}^{(t)}}^{\infty} (D_2^{(t)} - x_{S2}^{(t)} - p_2^{(t)} - r_{12}^{(t)} - r_{22}^{(t)}) f(D_2^{(t)}, \mu_2, \sigma_2) dD_2^{(t)}.
\end{aligned} \tag{3.2}$$

As $EP_t = ER_t - EC_t$, the expected profit of period t is calculated as follows:

$$\begin{aligned}
&EP_t(x_{S1}^{(t)}, x_{S2}^{(t)}, p_1^{(t)}, p_2^{(t)}, r_{11}^{(t)}, r_{12}^{(t)}, r_{21}^{(t)}, r_{22}^{(t)}) \\
&= s_1\mu_1 + s_2\mu_2 - c_{P1}p_1^{(t)} - c_{P2}p_2^{(t)} - c_{R11}r_{11}^{(t)} - c_{R12}r_{12}^{(t)} - c_{R21}r_{21}^{(t)} - c_{R22}r_{22}^{(t)} \\
&\quad - h_1\int_0^{x_{S1}^{(t)}+p_1^{(t)}+r_{11}^{(t)}+r_{21}^{(t)}} (x_{S1}^{(t)} + p_1^{(t)} + r_{11}^{(t)} + r_{21}^{(t)} - D_1^{(t)}) f(D_1^{(t)}, \mu_1, \sigma_1) dD_1^{(t)} \\
&\quad - (s_1 + v_1)\int_{x_{S1}^{(t)}+p_1^{(t)}+r_{11}^{(t)}+r_{21}^{(t)}}^{\infty} (D_1^{(t)} - x_{S1}^{(t)} - p_1^{(t)} - r_{11}^{(t)} - r_{21}^{(t)}) f(D_1^{(t)}, \mu_1, \sigma_1) dD_1^{(t)} \\
&\quad - h_2\int_0^{x_{S2}^{(t)}+p_2^{(t)}+r_{12}^{(t)}+r_{22}^{(t)}} (x_{S2}^{(t)} + p_2^{(t)} + r_{12}^{(t)} + r_{22}^{(t)} - D_2^{(t)}) f(D_2^{(t)}, \mu_2, \sigma_2) dD_2^{(t)} \\
&\quad - (s_2 + v_2)\int_{x_{S2}^{(t)}+p_2^{(t)}+r_{12}^{(t)}+r_{22}^{(t)}}^{\infty} (D_2^{(t)} - x_{S2}^{(t)} - p_2^{(t)} - r_{12}^{(t)} - r_{22}^{(t)}) f(D_2^{(t)}, \mu_2, \sigma_2) dD_2^{(t)}.
\end{aligned} \tag{3.3}$$

The expected maximum total profit from period t till final period is denoted as $f_t(x_{S1}^{(t)}, x_{S2}^{(t)})$. Assume the expected maximum total profit beyond the planning horizon to be equal to zero, i.e. $f_{M+1}(x_{S1}^{(M+1)}, x_{S2}^{(M+1)}) = 0$. Thus, for the multi-period problem, the Bellman's equation of dynamic programming model can be written as follows (t = 1, 2,..., M):

$$f_t(x_{S1}^{(t)}, x_{S2}^{(t)}) = \underset{R_1^{(t)},R_2^{(t)}}{E}\{\max_{p_1^{(t)},p_2^{(t)},r_{11}^{(t)},r_{12}^{(t)},r_{21}^{(t)},r_{22}^{(t)}}\{EP_t(x_{S1}^{(t)}, x_{S2}^{(t)}, p_1^{(t)}, p_2^{(t)}, r_{11}^{(t)}, r_{12}^{(t)}, r_{21}^{(t)}, r_{22}^{(t)}) + \underset{D_1^{(t)},D_2^{(t)}}{E}[f_{t+1}(x_{S1}^{(t+1)}, x_{S2}^{(t+1)})]\}\}. \tag{3.4}$$

The objective of studying the dynamic programming model is to obtain the optimal policy for the two-product recovery system in a finite horizon. However, due to the curse of dimensionality of dynamic programming, it is intractable to solve dynamic programming problem involving more than two states.

The expected maximum total profit of the final period can be expressed as follows:

$$f_M(x_{S1}^{(M)}, x_{S2}^{(M)}) = \underset{R_1^{(M)},R_2^{(M)}}{E}[\max_{p_1^{(M)},p_2^{(M)},r_{11}^{(M)},r_{12}^{(M)},r_{21}^{(M)},r_{22}^{(M)}} EP_M]. \tag{3.5}$$

As shown in Formula (3.5), the maximization of the expected profit in a single period would be the standing point for solving the dynamic programming model.

3.3 Summary

In this Chapter, firstly, we have introduced the two-product recovery system in a finite horizon. Secondly, we have developed a dynamic programming model of the two-product recovery system. In the following Chapter, we would study the two-product recovery system in a single period, which could be the basis for further study on the recovery system in the multi-period situation.

Chapter 4 The study on two-product recovery system in a single period

Chapter 4 focuses on the two-product recovery system in a single period. Section 4.1 introduces the system. In Section 4.2, the expected profit maximization model of the system is developed. Furthermore, the objective function of the model proves to be concave on production and recovery decisions. Therefore, the optimal solution to the model can be obtained by solving KKT conditions. Based on the optimal solution, the optimal multi-level threshold policy is obtained. The related threshold levels of the policy are discovered. Their managerial insights are further explained. Section 4.3 discusses about the extension from the two-product case to a general multi-product case. Finally, Section 4.4 summarizes the main work in this chapter.

4.1 Introduction

The introduction to the two-product recovery system has been made in Chapter 3. In this chapter, the recovery system is studied in a single period. The occurring events of the two-product recovery system in a single period are similar to those in a certain period of a multi-period horizon as described in Chapter 3. However, the remaining finished products at the end of the single period have to be salvaged. Therefore, we need to consider their salvage value. Hereafter in this chapter, we would assume the salvage value of the remaining finished products to be equal to zero.

The aim of studying the two-product recovery system in a single period is to maximize the expected profit in this period, and to show certain good properties. These properties will help to obtain the optimal policy on production and recovery decisions in the single period.

4.2 Production and recovery decisions for two products in a single period

In this section, we would formulate and analyze the single-period problem on the two-product recovery system. As the related assumptions of the single-period problem are similar to the multi-period problem in Chapter 3, we would not repeat here. With regard to the independency between chapters, we would list the related notations of the single-period problem here. Different from the multi-period problem, the time index will be excluded from these notations.

4.2.1 Notations

The notations of the single-period problem are listed as follows (i = 1, 2; j = 1, 2):

r_{ij} quantity of recovering returned item in group i to product j;

p_j quantity of producing product j;

s_j selling price of product j;

x_{Sj} initial inventory of product j;

x_j replenishment level of product j after production and recovery;

c_{Rij} unit cost of recovering returned item in group i to product j;

c_{Pj} production cost of per unit product j;

h_j	inventory holding cost of per unit product *j*;
v_j	penalty cost of per unit shortage of product *j*;
R_i	returned items in group *i*;
D_j	demand for product *j*;
$f(x,\mu,\sigma)$	probability density function w.r.t. x with known parameter (μ,σ);
$F(x,\mu,\sigma)$	cumulative distribution function w.r.t. x with known parameter (μ,σ);
$F^{-1}(x,\mu,\sigma)$	inverse function of $F(x,\mu,\sigma)$;
EP	expected profit in the single period.

In Chapter 3, some restrictions have been made to cost parameters in order to ensure the economical meaning of studying the two-product recovery system. As different cost structures result in different forms of production and recovery, we would focus on the modeling of the recovery system based on a certain cost structure, which imposes other restrictions on cost parameters. Under the cost structure, we will obtain the optimal policy of production and recovery through solving the model. For the other cost structures, the process of modeling and solving can easily refer to it. The selected cost structure includes the restrictions on cost parameters: $c_{P1}-c_{R11}>c_{P2}-c_{R12}$, $c_{P1}-c_{R21}>c_{P2}-c_{R22}$, and $c_{R11}-c_{R21}>c_{R12}-c_{R22}$.

4.2.2 The expected profit maximization model

At the beginning of the single period, we have known the quantities of returned items and the initial inventory of two products. Then, we would make the optimal production and recovery decisions in order to maximize the expected profit in the single period. As recovery decisions are subject to the availability of returned items, the expected profit maximization model of the single-period problem can be formulated as follows:

$$\max_{p_1\geq 0,\, p_2\geq 0,\, r_{11}\geq 0,\, r_{12}\geq 0,\, r_{21}\geq 0,\, r_{22}\geq 0} EP(p_1, p_2, r_{11}, r_{12}, r_{21}, r_{22})$$

$$\begin{aligned} & s.t. \\ & r_{11}+r_{12}\leq R_1; \\ & r_{21}+r_{22}\leq R_2. \end{aligned} \tag{4.1}$$

In Chapter 3, we have introduced the calculation of the expected profit at period *t*. According to Formula (3.3), the expected profit in a single period can be similarly calculated as follows:

$$\begin{aligned} & EP(p_1, p_2, r_{11}, r_{12}, r_{21}, r_{22}) \\ & = s_1\mu_1 + s_2\mu_2 - c_{P1}p_1 - c_{P2}p_2 - c_{R11}r_{11} - c_{R12}r_{12} - c_{R21}r_{21} - c_{R22}r_{22} \\ & -h_1\int_0^{x_{S1}+p_1+r_{11}+r_{21}} (x_{S1}+p_1+r_{11}+r_{21}-D_1)f(D_1,\mu_1,\sigma_1)dD_1 \\ & -(s_1+v_1)\int_{x_{S1}+p_1+r_{11}+r_{21}}^{\infty} (D_1-x_{S1}-p_1-r_{11}-r_{21})f(D_1,\mu_1,\sigma_1)dD_1 \\ & -h_2\int_0^{x_{S2}+p_2+r_{12}+r_{22}} (x_{S2}+p_2+r_{12}+r_{22}-D_2)f(D_2,\mu_2,\sigma_2)dD_2 \\ & -(s_2+v_2)\int_{x_{S2}+p_2+r_{12}+r_{22}}^{\infty} (D_2-x_{S2}-p_2-r_{12}-r_{22})f(D_2,\mu_2,\sigma_2)dD_2. \end{aligned} \tag{4.2}$$

In order to simplify the expression of Formula (4.2), let $L_1(p_1, r_{11}, r_{21})$ and $L_2(p_2, r_{12}, r_{22})$ denote the accounting items related to stochastic demands for product 1 and product 2 respectively. The two accounting items are expressed as follows:

$$L_1(p_1, r_{11}, r_{21}) = -h_1 \int_0^{x_{S1}+p_1+r_{11}+r_{21}} (x_{S1} + p_1 + r_{11} + r_{21} - D_1) f(D_1, \mu_1, \sigma_1) dD_1 - (s_1 + v_1) \int_{x_{S1}+p_1+r_{11}+r_{21}}^{\infty} (D_1 - x_{S1} - p_1 - r_{11} - r_{21}) f(D_1, \mu_1, \sigma_1) dD_1. \tag{4.3}$$

$$L_2(p_2, r_{12}, r_{22}) = -h_2 \int_0^{x_{S2}+p_2+r_{12}+r_{22}} (x_{S2} + p_2 + r_{12} + r_{22} - D_2) f(D_2, \mu_2, \sigma_2) dD_2 - (s_2 + v_2) \int_{x_{S2}+p_2+r_{12}+r_{22}}^{\infty} (D_2 - x_{S2} - p_2 - r_{12} - r_{22}) f(D_2, \mu_2, \sigma_2) dD_2. \tag{4.4}$$

The inventory level of each product after replenishment is calculated as follows (j = 1, 2):

$$x_j = x_{Sj} + p_j + r_{1j} + r_{2j}. \tag{4.5}$$

As the related decision variables act similarly in the function L_1 (or L_2), the first-order partial derivatives of the function L_1 (or L_2) with respect to them are equal to each other. In details, they are calculated as follows (j = 1, 2):

$$\frac{\delta L_j}{\delta p_j} = \frac{\delta L_j}{\delta r_{1j}} = \frac{\delta L_j}{\delta r_{2j}} = s_j + v_j - (s_j + v_j + h_j) F(x_j, \mu_j, \sigma_j). \tag{4.6}$$

Lemma 4.1**:** *The objective function of the expected profit maximization model is jointly concave on all the decision variables for the single-period two-product recovery system disregarding salvage value of the remaining finished products.*

Proof of Lemma 4.1:

The *Lemma 4.1* is proved if and only if the nonlinear part of the objective function, i.e. the functions L_1 and L_2, could be proved to be concave on all the decision variables. Firstly, we would prove the concavity property of the function L_1.

According to Formula (4.6), the second-order partial derivative of the function L_1 with respect to production decision (p_1) can be calculated as follows (Let a denote its value.):

$$a = \frac{\partial^2 L_1}{\partial {p_1}^2} = -(s_1 + v_1 + h_1) f(x_1, \mu_1, \sigma_1) \leq 0.$$

As the related three decision variables act similarly in the function L_1, all the second-order partial derivatives of the function L_1 with respect to them are equal to each other. Thus, the Hessian matrix of the function L_1 can be expressed as $\begin{bmatrix} a & a & a \\ a & a & a \\ a & a & a \end{bmatrix}$.

Given a random nonzero vector (y_1, y_2, y_3), there exist:

$$(y_1, y_2, y_3). \begin{bmatrix} a & a & a \\ a & a & a \\ a & a & a \end{bmatrix} . \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = a(y_1 + y_2 + y_3)^2 \leq 0 .$$

Therefore, the Hessian matrix of the function L_1 is negative semi-definite. Thus, the function L_1 has been proved to be concave on its related decision variables. Similarly, the concavity of the function L_2 can be proved. Finally, *Lemma 4.1* has been proved.

The concavity of the objective function has been shown in *Lemma 4.1*. Furthermore, we would apply the method of Lagrange Multipliers to find the maximum. For the model, the Lagrangian function (denoted as L) can be expressed as follows:

$$
\begin{aligned}
&L(p_1, p_2, r_{11}, r_{12}, r_{21}, r_{22}, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \lambda_8) \\
&= EP(p_1, p_2, r_{11}, r_{12}, r_{21}, r_{22}) + \lambda_1(r_{11} + r_{12} - R_1) + \lambda_2(r_{21} + r_{22} - R_2) \\
&\quad - \lambda_3 p_1 - \lambda_4 p_2 - \lambda_5 r_{11} - \lambda_6 r_{12} - \lambda_7 r_{21} - \lambda_8 r_{22}.
\end{aligned}
\tag{4.7}
$$

In order to obtain the optimal solution to the model, we will need to consider the KKT conditions for the maximum of the Lagrangian function. As we have known, the necessary conditions are also sufficient for optimality when the objective function is concave and the inequality constraints are linear on decision variables. *Lemma 4.1* has shown the concavity of the objective function on decision variables. In addition, the inequality constraints are linear on decision variables. Therefore, the solution to KKT conditions is also the global maximum of the model. The optimal production and recovery decisions should satisfy all the KKT conditions at the same time. In details, the KKT conditions are listed as follows:

$$
\begin{aligned}
&\frac{\delta EP}{\delta p_1^*} - \lambda_3 = 0; \\
&\frac{\delta EP}{\delta p_2^*} - \lambda_4 = 0; \\
&\frac{\delta EP}{\delta r_{11}^*} + \lambda_1 - \lambda_5 = 0; \\
&\frac{\delta EP}{\delta r_{12}^*} + \lambda_1 - \lambda_6 = 0; \\
&\frac{\delta EP}{\delta r_{21}^*} + \lambda_2 - \lambda_7 = 0; \\
&\frac{\delta EP}{\delta r_{22}^*} + \lambda_2 - \lambda_8 = 0;
\end{aligned}
$$

$$\lambda_1(r_{11}^* + r_{12}^* - R_1) = 0;$$
$$\lambda_2(r_{21}^* + r_{22}^* - R_2) = 0;$$
$$\lambda_3 p_1^* = \lambda_4 p_2^* = \lambda_5 r_{11}^* = \lambda_6 r_{12}^* = \lambda_7 r_{21}^* = \lambda_8 r_{22}^* = 0;$$
$$r_{11}^* + r_{12}^* \le R_1;$$
$$r_{21}^* + r_{22}^* \le R_2;$$
$$p_1^* \ge 0, p_2^* \ge 0, r_{11}^* \ge 0, r_{12}^* \ge 0, r_{21}^* \ge 0, r_{22}^* \ge 0;$$
$$\lambda_i \le 0 \text{ (i = 1,...,8)}.$$

Through solving the above KKT conditions, we can obtain the optimal production and recovery decisions for the two products, which are dependent on the initial inventory of the two products and the availability of returned items. The optimal solution includes 21 cases given in Appendix B. However, after analyzing these 21 cases, we found out that they can be represented by an optimal multi-level threshold policy. This optimal policy is characterized by 6 order-up-to levels and 3 switching levels. Once these threshold levels have been determined, we can use the optimal policy to make the optimal production and recovery decisions. Among the threshold levels, there are three order-up-to levels for each product corresponding to three different replenishment sources: production, recovery using returned items in group 1 and group 2 respectively. These order-up-to levels can be obtained by solving the related KKT conditions as follows:

- **Order-up-to level by production**

  For each product, the order-up-to level by production is defined as the maximum inventory level by the replenishment of production. At the order-up-to level, the marginal profit of further replenishment is equal to zero. With the combination of the related KKT conditions, we can determine the order-up-to level for product 1 as follows:

$$p_1^* > 0 \text{ , } \lambda_3 p_1^* = 0, \text{ and } \left.\frac{\delta EP}{\delta p_1^*}\right|_{x_1^* = AL_0} - \lambda_3 = 0$$

$$\Rightarrow AL_0 = F^{-1}\left(\frac{s_1 + v_1 - c_{P1}}{s_1 + v_1 + h_1}, \mu_1, \sigma_1\right).$$

Similarly, for product 2, the order-up-to level by production can be determined as follows:

$$BL_0 = F^{-1}\left(\frac{s_2 + v_2 - c_{P2}}{s_2 + v_2 + h_2}, \mu_2, \sigma_2\right).$$

- **Order-up-to level by using the returned items in group 1**

For each product, the order-up-to level is defined as the maximum inventory level by the replenishment using the returned items in group 1. If returned items in group 1 are enough for the allocation between the two products, the inventory of the two products will be replenished until the order-up-to level, at which the marginal profit of further replenishment is equal to zero. With the combination of the related KKT conditions, we can determine the order-up-to level of the two products as follows:

$$\frac{\delta EP}{\delta r_{11}^{*}}+\lambda_1-\lambda_5=0; \quad \frac{\delta EP}{\delta r_{12}^{*}}+\lambda_1-\lambda_6=0;$$
$$\lambda_1(r_{11}^{*}+r_{12}^{*}-R_1)=0; \quad \lambda_5 r_{11}^{*}=\lambda_6 r_{12}^{*}=0;$$
$$r_{11}^{*}+r_{12}^{*}<R_1; \quad r_{11}^{*}>0; \quad r_{12}^{*}>0.$$

$$\Rightarrow \left.\frac{\delta EP}{\delta r_{11}^{*}}\right|_{x_1^{*}=AL_1}=\left.\frac{\delta EP}{\delta r_{12}^{*}}\right|_{x_2^{*}=BL_1}=0$$
$$\Rightarrow AL_1=F^{-1}\left(\frac{s_1+v_1-c_{R11}}{s_1+v_1+h_1},\mu_1,\sigma_1\right);$$
$$BL_1=F^{-1}\left(\frac{s_2+v_2-c_{R12}}{s_2+v_2+h_2},\mu_2,\sigma_2\right).$$

- **Order-up-to level by using the returned items in group 2**

For each product, the order-up-to level is defined as the maximum inventory level by the replenishment using the returned items in group 2. If returned items in group 2 are enough for the allocation between the two products, the inventory of the two products will be replenished until the order-up-to level, at which the marginal profit of further replenishment is equal to zero. With the combination of the related KKT conditions, we can determine the order-up-to level of the two products as follows:

$$\frac{\delta EP}{\delta r_{21}^*} + \lambda_2 - \lambda_7 = 0; \quad \frac{\delta EP}{\delta r_{22}^*} + \lambda_2 - \lambda_8 = 0;$$
$$\lambda_2(r_{21}^* + r_{22}^* - R_2) = 0; \quad \lambda_7 r_{21}^* = \lambda_8 r_{22}^* = 0;$$
$$r_{21}^* + r_{22}^* < R_2; \quad r_{21}^* > 0; \quad r_{22}^* > 0.$$

$$\Rightarrow \left.\frac{\delta EP}{\delta r_{21}^*}\right|_{x_1^* = AL_2} = \left.\frac{\delta EP}{\delta r_{22}^*}\right|_{x_2^* = BL_2} = 0$$
$$\Rightarrow AL_2 = F^{-1}\left(\frac{s_1 + v_1 - c_{R21}}{s_1 + v_1 + h_1}, \mu_1, \sigma_1\right);$$
$$BL_2 = F^{-1}\left(\frac{s_2 + v_2 - c_{R22}}{s_2 + v_2 + h_2}, \mu_2, \sigma_2\right).$$

Besides these order-up-to levels, three switching levels, denoted as SW_1, SW_2 and RP, are used to control the interactive allocation of possibly limited returned items between the recovery processes of the two products, which involves the comparison of marginal profits. It can be found from the KKT conditions that the two products have equal marginal profits while being further replenished by a certain recovery source based on the respective inventory level after the optimal replenishment. So the optimal solution by solving the KKT conditions always maintains the equality of marginal profits between the two products. Suppose the final inventory level of product 2 after the optimal replenishment is at a certain order-up-to level, we would need to find out the corresponding inventory level of product 1 such that the two products have equal marginal profits. The details of the corresponding inventory levels, i.e. the switching levels SW_1, SW_2 and RP, are explained as follows:

- **Switching level SW_1**

The switching level SW_1 for product 1 corresponds to the order-up-to level BL_0 for product 2. When the final inventory levels of the two products after the optimal replenishment are at SW_1 and BL_0 respectively, the two products will have equal marginal profits from recovering the returned items in group 1. Thus,

$$\left.\frac{\delta EP}{\delta r_{11}^*}\right|_{x_1^*=SW_1} = \left.\frac{\delta EP}{\delta r_{12}^*}\right|_{x_2^*=BL_0}$$
$$\Rightarrow \left.\frac{\delta EP}{\delta r_{11}^*}\right|_{x_1^*=SW_1} = c_{P2} - c_{R12}$$
$$\Rightarrow SW_1 = F^{-1}\left(\frac{s_1 + v_1 + c_{R12} - c_{R11} - c_{P2}}{s_1 + v_1 + h_1}, \mu_1, \sigma_1\right).$$

- **Switching level SW_2**

The switching level SW_2 for product 1 corresponds to the order-up-to level BL_0 for product 2. When the final inventory levels of the two products after the optimal replenishment are at SW_2 and BL_0 respectively, the two products will have equal marginal profits from recovering the returned items in group 2. Thus,

$$\left.\frac{\delta EP}{\delta r_{21}^*}\right|_{x_1^*=SW_2} = \left.\frac{\delta EP}{\delta r_{22}^*}\right|_{x_2^*=BL_0}$$
$$\Rightarrow \left.\frac{\delta EP}{\delta r_{21}^*}\right|_{x_1^*=SW_2} = c_{P2} - c_{R22}$$
$$\Rightarrow SW_2 = F^{-1}\left(\frac{s_1 + v_1 + c_{R22} - c_{R21} - c_{P2}}{s_1 + v_1 + h_1}, \mu_1, \sigma_1\right).$$

- **Switching level *RP***

The switching level *RP* for product 1 corresponds to the order-up-to level BL_1 for product 2. When the final inventory levels of the two products after the optimal replenishment are at *RP* and BL_1

respectively, the two products will have equal marginal profits from recovering the returned items in group 2. Thus,

$$\left.\frac{\delta EP}{\delta r_{21}^*}\right|_{x_1^*=RP} = \left.\frac{\delta EP}{\delta r_{22}^*}\right|_{x_2^*=BL_1}$$

$$\Rightarrow \left.\frac{\delta EP}{\delta r_{21}^*}\right|_{x_1^*=RP} = c_{R12} - c_{R22}$$

$$\Rightarrow RP = F^{-1}\left(\frac{s_1 + v_1 + c_{R22} - c_{R21} - c_{R12}}{s_1 + v_1 + h_{s1}}, \mu_1, \sigma_1\right).$$

Due to the restrictions on cost parameters mentioned before, we can tell the relative locations of the above-mentioned threshold levels based on their determination formulae in Tables A.1 and A.2 of Appendix A. Among the order-up-to levels of product 1, AL_2 is the highest whereas AL_0 is the lowest, and AL_1 is between them. Furthermore, among the threshold levels of product 1, SW_1 is located between AL_0 and AL_1 whereas the threshold level RP is located between AL_1 and AL_2. The threshold level SW_2 is between SW_1 and RP. In addition, among the order-up-to levels of product 2, BL_2 is the highest whereas BL_0 is the lowest, and BL_1 is between them. The locations of all the threshold levels for the two products can be referred to in Figure 4.1.

In the following section, we would further explain the insights of the above-mentioned threshold levels. At the same time, the managerial insights to the optimal control of two-product recovery system are introduced. With the rules, we would know how to make the optimal production and recovery decisions for the two-product recovery system if the initial inventory of the two products and the availability of returned items are given.

4.2.3 Managerial insights to the optimal control of two-product recovery system in a single period

Firstly, we have described the relative locations of the order-up-to levels and the related threshold levels for the two products in Figure 4.1.

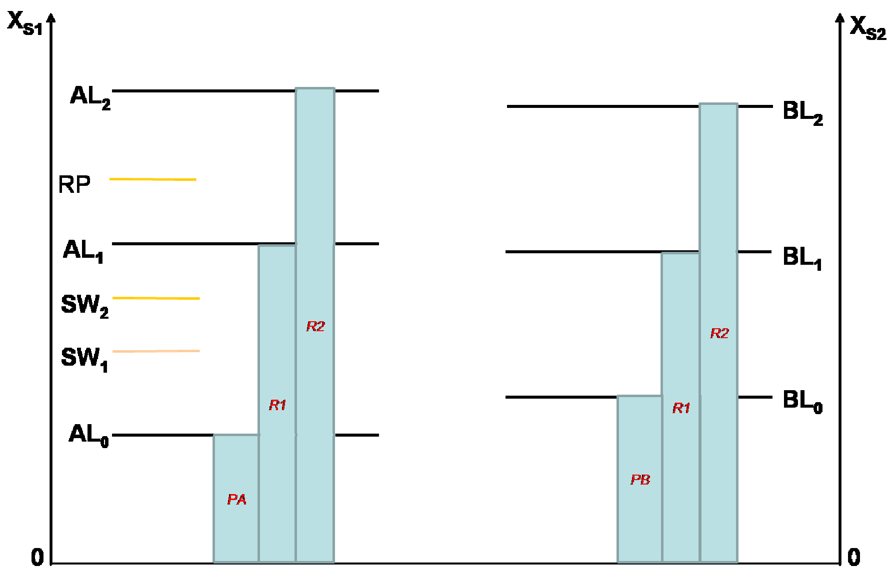


Figure 4.1 The threshold levels for the inventory control of two-product recovery system

In Figure 4.1, *R1* and *R2* denote the returned items incurring high and low recovery cost respectively; *PA* and *PB* denote the production of product 1 and product 2 respectively; X_{S1} and X_{S2} denote the finished item inventory of product 1 and product 2 respectively. Different from traditional inventory problem, the product recovery system involves multiple sources of supplies for each product (either recovery with *R1*, *R2* or doing production). Which sources to be used to replenish the inventory of finished items will depend on the initial inventory levels of finished items, the costs of recovery/production, and the availability of returned items. For example, we would expect the sources with lower recovery/production costs to be used first, and the optimal replenishment level for these sources would be higher than other sources with higher recovery/production costs.

Figure 4.1 has shown the order-up-to levels of *PA*, *R1* and *R2* for product 1, denoted by AL_0, AL_1 and AL_2 respectively; and the order-up-to levels of *PB*, *R1* and *R2* for product 2, denoted by BL_0, BL_1 and BL_2 respectively. Except AL_0 and BL_0, these order-up-to levels are the highest replenishment levels of the finished item inventory for the two products if the respective sources are used, and further replenishment beyond these levels would be unprofitable. For example, the highest replenishment level for product 1 is AL_1 if *R1* is used. The order-up-to levels AL_0 and BL_0 respectively, are the highest replenishment levels for product 1 and product 2 if production is used. In addition, they are also the lowest replenishment levels due to unlimited production capacity.

To explain the insights of the above-mentioned order-up-to levels more clearly, we first assume that the replenishment of the two products are independent, i.e. only one product is available. Taking the replenishment of product 1 for example, we would compare the order-up-to levels of all the replenishment sources (*R1*, *R2* and *PA*). As *R2* is the cheapest, its order-up-to level AL_2 is the highest while *PA* is the most expensive and so its order-up-to level AL_0 is the lowest. Additionally, due to unlimited production capacity, AL_0 is also the lowest level that we would order up to. As for *R1*, its order-up-to level AL_1 is between AL_0 and AL_2. Due to cost difference between the replenishment sources, *R2* would be the first choice of the three replenishment sources, then *R1* if *R2* is used up, and finally *PA* if *R1* is used up. In the meantime, the choice of a particular replenishment source should only be made when this replenishment source is still cost-effective. The final replenishment level would be subject to the availability of returned items.

However, the above-mentioned order-up-to levels are not enough to control the two-product recovery system as there is interaction between the replenishment of the two products on the allocation of possibly limited returned items. In order to fulfill the optimal allocation, we need to refer to the threshold levels SW_1, SW_2 and RP. As shown in Figure 4.1, the threshold level SW_1 is higher than AL_0 but lower than AL_1, RP is higher than AL_1 but lower than AL_2, whereas the threshold level SW_2 is between SW_1 and RP. In Figure 4.1, SW_2 is shown to be lower than AL_1 but it is not always like that because it is not subject to the selected cost structure but the other cost structures. All the three threshold levels are related to the inventory level of product 1. For each threshold level, the determination and insight can be referred to in Appendix A.

As mentioned before, the three threshold levels are defined by comparing the marginal profits of using the recovery with *R1* or *R2* to replenish the two products. By the comparison, *R1* or *R2* would be allocated to the product, which is more profitable to be replenished. Thus, by this kind of allocation, the inventory level of the product with high marginal profit is increased whereas the inventory level of another product remains unchanged. With the inventory level increasing, the product with originally high marginal profit would have its marginal profit decreasing until the two products have equal marginal profits. We would define this kind of allocation rule as 'fair' allocation rule, which aims to balance the marginal profits of the two products being replenished with the recovery. If the allocation is based on the inventory levels, at which the two products have had equal marginal profits already, the allocation would increase the inventory levels of the two products at the same time, and maintain the equality of marginal profits at the final inventory levels of the two products. The details about using the three threshold levels will be introduced later.

Suppose that the initial inventory levels of product 1 and product 2 are lower than their order-up-to levels AL_0 and BL_0 respectively. In this situation, due to the selected cost structure, it is more profitable to replenish product 1 than product 2 by using recovery sources. Therefore, both *R1* and *R2* prefer to replenish product 1 whereas product 2 is replenished by *PB*. Due to unlimited production capacity, product 2 can be always replenished by *PB* to the order-up-to level BL_0. The final level of product 1 after replenishment depends on the availability of returned items. As *R2* is cheaper than *R1*, *R2* will be used at first. Once *R2* is used up and *R1* is still cost-effective to replenish product 1, then *R1* will be used. Based on the situation and different availability of returned items, we would introduce the threshold levels SW_1, SW_2 and *RP* as follows:

- Threshold level SW_1

  The threshold level SW_1 indicates the inventory level of product 1, at which *R1* would switch from the replenishment of product 1 to product 2. If *R2* is not enough to replenish the inventory of product 1 up to the threshold level SW_1, the following allocation of *R1* will be used to replenish product 1 until SW_1 is reached. After that, if there are *R1* left, the remaining *R1* will switch to replenish product 2 in place of *PB*, instead of continuing the replenishment of product 1. If there are *R1* left after replacing all the *PB* at product 2, the remaining *R1* will be allocated to the two products following the 'fair' allocation rule. The 'fair' allocation of *R1* would increase the inventory levels of product 1 and product 2 at the same time until the order-up-to levels AL_1 and BL_1 are reached respectively.

- Threshold level SW_2

  The threshold level SW_2 indicates the inventory level of product 1, at which *R2* would switch from the replenishment of product 1 to product 2. As *R2* is cheaper than *R1*, the threshold level SW_2 is higher than the threshold level SW_1. If *R2* is more than enough to increase the inventory level of product 1 to the threshold level SW_2, *R2* will switch to replenish product 2 in place of *PB* until all the *PB* at product 2 are replaced. After that, if there are still *R2* left, the remaining *R2* will be allocated between the two products following the 'fair' allocation rule. The 'fair' allocation of *R2* would increase the inventory levels of product 1 and product 2 at the same time until the order-up-to levels AL_2 and BL_2 are reached respectively.

- Threshold level *RP*

  The threshold level *RP* indicates the inventory level of product 1, at which product 1 has equal marginal profit from the recovery of *R2*, compared with product 2 at the order-up-to level BL_1. The threshold level *RP* will be involved when the following allocation of *R1* affects the existing allocation of *R2*. Suppose that the 'fair' allocation of *R2* has increased the inventory level of product 1 higher than SW_2 but lower than *RP*, and on the other hand, the inventory level of product 2 has been increased between BL_0 and BL_1. Then, the following allocation of *R1* will replace the existing allocation of *R2* at product 2, and the saved *R2* will be reallocated to product 1. Thus, the inventory levels of the two products will be increased at the same time by the process of replacement and reallocation. By the 'fair' allocation rule, the process would result in the final inventory levels of the two products, at which they have equal marginal profits from the recovery of *R2*.

If *R1* is enough to push the process but does not replace all the *R2* at product 2, *R1* can increase the inventory levels of product 1 and product 2 until the threshold levels *RP* and BL_1 are reached respectively. After that, *R1* will not be cost-effective. Otherwise, if *R1* has replaced all the *R2* at product 2 but does not increase the inventory level of product 1 to the threshold level *RP*. As *RP* is higher than AL_1, the resulting inventory level of product 1 may be either between AL_1 and *RP* or below AL_1. When the resulting inventory level is between AL_1 and *RP* and there are *R1* left, the remaining *R1* will replenish product 2 alone until the order-up-to level BL_1 is reached. When the resulting inventory level is below AL_1 and there are *R1* left, the remaining *R1* will replenish product 2 alone at first until the two products have equal marginal profits from the recovery of *R1*. After that, *R1* will be 'fairly' allocated to the two products until the inventory levels of product 1 and product 2 reach the threshold levels AL_1 and BL_1 respectively.

The allocation of multiple replenishment sources between the two products is more complicated as it depends on the initial inventory levels of the two products and the availability of returned items. It does not make much sense to describe all the allocation situations here. However, the optimal solution has included all the allocation situations, which can be referred to in Appendix B. The optimal solution has been obtained under the selected cost structure. Furthermore, we would describe the replenishment process of the two-product recovery system in Figure 4.2. The replenishment process is implemented in the main algorithm, which calls two sub-algorithms to allocate *R2* and *R1* respectively in sequence.

In the following section, we would discuss about how to extend the results from the two-product case to a general multi-product case.

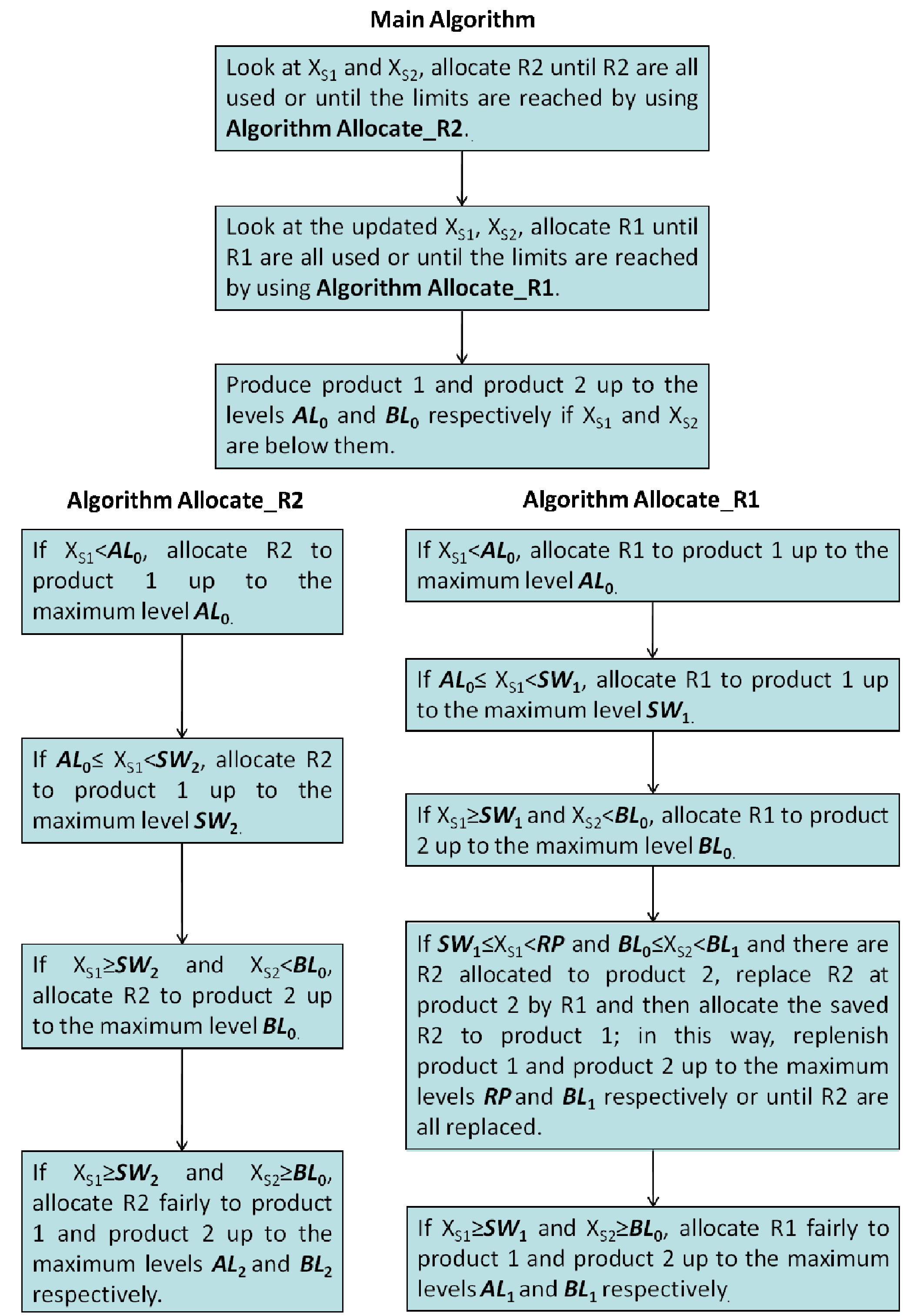


Figure 4.2 The inventory replenishment process of the two-product recovery system in a single period

4.3 The extension to a general multi-product recovery system

The N-product recovery system can be drawn as follows. The demands follow some general distributions.

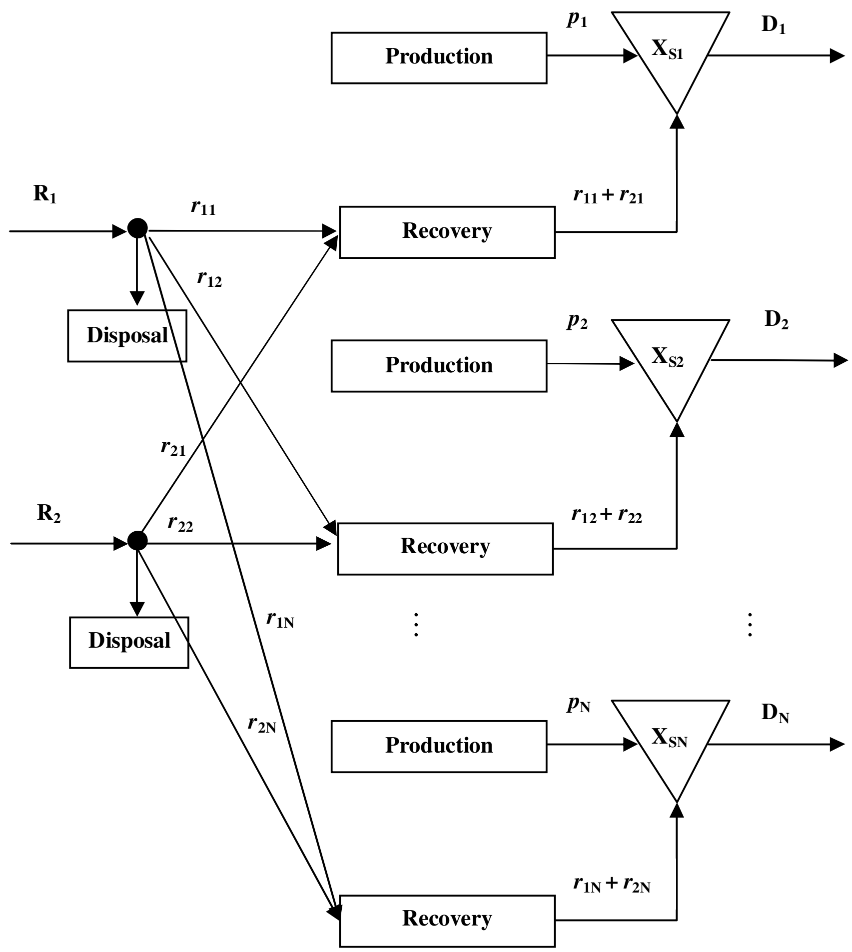


Figure 4.3 The structure of the N-product recovery system

Some restrictions on cost parameters are imposed so as to ensure the economical meaningfulness of the study on the N-product recovery system. Firstly, for each product, selling price is higher than production cost, and penalty cost of shortage is higher than the profit from production. Therefore, there exist: $s_j > c_{Pj}$, $v_j > s_j - c_{Pj}$ (j = 1, 2, …, N). Secondly, for each product, production cost is higher than recovery cost. Otherwise, recovery is unnecessary. Therefore, there exist: $c_{Pj} > c_{R1j}$, $c_{Pj} > c_{R2j}$. Finally, the recovery using returned item in group 2 is cheaper than that using returned item in group 1. Therefore, there exist: $c_{R2j} < c_{R1j}$.

The *N*-product recovery system can be firstly studied in the single period by referring to the two-product recovery system. Similarly, the optimal solution can be found by solving KKT conditions of Lagrangian function. The optimal solution to the single-period problem on the two-product recovery system has been shown in Appendix B based on a choice of cost structure: $c_{P1} - c_{R11} > c_{P2} - c_{R12}$, $c_{P1} - c_{R21} > c_{P2} - c_{R22}$, and $c_{R11} - c_{R21} > c_{R12} - c_{R22}$. For the *N*-product recovery system, the optimal solution can be similarly shown based on the cost structure (j = 1, 2,…, *N*-1): $c_{Pj} - c_{R1j} > c_{P(j+1)} - c_{R1(j+1)}$, $c_{Pj} - c_{R2j} > c_{P(j+1)} - c_{R2(j+1)}$, and $c_{R1j} - c_{R2j} > c_{R1(j+1)} - c_{R2(j+1)}$.

In the *N*-product recovery system, there are three order-up-to levels for each product with respect to three replenishment sources. Based on the selected cost structure, there are three threshold levels for product j to controlling the interactive allocation of recovery replenishment sources between product j and product (j+1). Totally, the number of the threshold levels for the *N*-product recovery system is calculated as 3**N*+3*(*N*-1) = 6**N*-3. Although the solution structure of the optimal solution will increase with the number of products in the recovery system, the *N*-product case can still be formalized as NLP problem and solved by KKT conditions. When the recovery system is extended from the single-period context to the multi-period context, the threshold policy is considered to be used as it is easy to be implemented. For example, for M-period problem of the N-product recovery system, the threshold levels of each period can be determined by referring to the work on the two-product recovery system. However, the computational complexity of the heuristic algorithm would increase with the number of products in the recovery system.

4.4 Summary

For the two-product recovery system involving two groups of returned items based on quality classification, we have obtained the optimal solution to the single-period problem by solving KKT conditions. After analyzing the 21 cases of the optimal solution, we found out that they can be represented by an optimal multi-level threshold policy. Although the policy is a similar threshold policy to many works in the literature, it has more complicated structure due to multiple replenishment sources and multiple products.

This optimal policy is characterized by 6 order-up-to levels and 3 switching levels. For each of the two products, there are 3 order-up-to levels corresponding to different replenishment sources. In addition, there are 3 switching levels to control the allocation of the returned items between the two products. The managerial insights of these threshold levels have been explained. The allocation of returned items between the two products would follow the fair allocation rule. The rule aims to balance the marginal profits from the recovery replenishment source between the two products. Based on the selected cost structure, we have shown the details of the replenishment process. For the other cost structures, the replenishment process and managerial rules can be similarly obtained. In particular, the extension from the two-product situation to a general multi-product situation has also been discussed.

The research results of the two-product recovery system in a single period will be used for further research on the recovery system in the multi-period context. The threshold policy is assumed to be used for the multi-period problem as it is intuitive, easy to use and provides good managerial perspectives even though it might not be optimal.

Chapter 5 The study on two-product recovery system in a finite horizon with lost sale and zero lead time

Chapter 5 focuses on the study of the two-product recovery system in a finite horizon. The stocking of two products in the recovery system aims to satisfy stochastic customer demands in each period of the planning horizon. The inventory of the two products can be instantly replenished by production and recovery processes as both processes are assumed to have zero lead time. If the inventory is in shortage, the recovery system will lose the sales. Section 5.1 introduces the two-product recovery system. In Section 5.2, an ADP model of the system is developed in order to maximize the expected total profit in the finite horizon. The model is used to derive the threshold levels, which are only dependent on the gradient of the cost-to-go function at the points of interest. Section 5.3 provides the details about how to determine the gradient of the cost-to-go function at the points of interest. Section 5.4 gives numerical analysis on the recovery system. Finally, Section 5.5 summarizes the main findings.

5.1 Introduction

The occurring events in each period of the finite horizon are described here. Firstly, returned items arrive at the two-product recovery system at the beginning of each period. Secondly, after observing the on-hand inventories of finished products at the beginning of each period, the manufacturer makes production and recovery decisions. Once the decisions are made, inventories of finished products get replenished instantly. The inventories are used to satisfy demands later in the same period. If the demands cannot be fully satisfied, the sales will be lost and penalty cost of the shortages will be incurred. Otherwise, if there are inventories left at the end of the period, the remaining inventories will be carried to future periods and inventory holding cost will be counted for the period. Finally, the remaining returned items are disposed of and disposal costs are negligible. The revenue is generated from the sale of finished products. The total cost consists of production cost, recovery cost and inventory holding cost of finished products and penalty cost of shortages.

The objective is to maximize the expected total profit of the two-product recovery system in a finite horizon. In order to fulfill the aim, we need to find the optimal policy of production planning and inventory control.

5.2 Approximate Dynamic programming model of the two-product recovery system in the multi-period context

The dynamic programming model of the two-product recovery system in the multi-period context has been introduced in Chapter 3. In this Chapter, the recovery system is assumed to deal with stock shortage in the way of lost sale. The other assumptions and the related notations are referred to as in Chapter 3. Besides, some related notations are listed as follows ($i, j = 1, 2$):

$x_{Sj}^{(t)}$ initial inventory of product j in period t;

$x_j^{(t)}$	inventory level of product j after production and recovery in period t;
$f_t(x_{S1}^{(t)}, x_{S2}^{(t)})$	expected maximum of the expected total profit from period t till final period;
EP_t	expected profit in period t;
MEP	maximum expected profit in final period;
ETP_t	expected total profit from period t till final period;
$\tilde{ETP}_t$	approximation to ETP_t;
$ATP_k^{(t)}$	actual profit in period t at sample k of demands and returns;
$u_j^{(t)}$	gradient of the cost-to-go function in period t w.r.t. order-up-to level of product j;
$grad_{j,k}^{(t)}$	sample gradient of $u_j^{(t)}$ at sample k of demands and returns.

The transition relationship on initial inventory of each product between two subsequent periods can be expressed as follows ($[X]^+ := \max\{X, 0\}$; $j = 1, 2$):

$$x_{Sj}^{(t+1)} = [x_{Sj}^{(t)} + p_j^{(t)} + r_{1j}^{(t)} + r_{2j}^{(t)} - D_j^{(t)}]^+. \tag{5.1}$$

In addition, the order-up-to level of each product at period t, i.e. inventory level after replenishment, is dependent on the initial inventory, which can be expressed as follows:

$$x_j^{(t)} = x_{Sj}^{(t)} + p_j^{(t)} + r_{1j}^{(t)} + r_{2j}^{(t)}. \tag{5.2}$$

As expressed in Chapter 3, the Bellman's equation of dynamic programming is as follows ($t = 1, 2, ..., M$):

$$f_t(x_{S1}^{(t)}, x_{S2}^{(t)}) = \underset{R_1^{(t)}, R_2^{(t)}}{E} \{ \max_{p_1^{(t)}, p_2^{(t)}, r_{11}^{(t)}, r_{12}^{(t)}, r_{21}^{(t)}, r_{22}^{(t)}} \{ EP_t(x_{S1}^{(t)}, x_{S2}^{(t)}, p_1^{(t)}, p_2^{(t)}, r_{11}^{(t)}, r_{12}^{(t)}, r_{21}^{(t)}, r_{22}^{(t)}) + \underset{D_1^{(t)}, D_2^{(t)}}{E} [f_{t+1}(x_{S1}^{(t+1)}, x_{S2}^{(t+1)})] \} \}. \tag{5.3}$$

The objective of studying the dynamic programming model is to obtain the optimal policy for the two-product recovery system in the multi-period context. However, due to the curse of dimensionality of dynamic programming, it is intractable to solve dynamic programming problem involving more than two states. For this kind of dynamic programming problem, suboptimal methods are proposed, which focus on evaluation and approximation of the cost-to-go function. Based on the approximation, the ADP model is proposed to help derive the threshold levels of the threshold policy. Therefore, we can approximate the cost-to-go function at the points of interest by using the gradients as follows:

$$\begin{aligned} &\underset{D_1^{(t)}, D_2^{(t)}}{E} [f_{t+1}(x_{S1}^{(t+1)}, x_{S2}^{(t+1)})] \\ &= \underset{D_1^{(t)}, D_2^{(t)}}{E} [f_{t+1}([x_1^{(t)} - D_1^{(t)}]^+, [x_2^{(t)} - D_2^{(t)}]^+)] \approx u_1^{(t)} x_1^{(t)} + u_2^{(t)} x_2^{(t)}. \end{aligned} \tag{5.4}$$

To find the gradient $u_j^{(t)}$, we can use the first-order derivatives of the cost-to-go function with respect to the inventory level of product j after replenishment, which is expressed as follows ($j = 1, 2$):

$$u_j^{(t)} = \frac{\partial \underset{D_1^{(t)},D_2^{(t)}}{E}[f_{t+1}([x_1^{(t)} - D_1^{(t)}]^+, [x_2^{(t)} - D_2^{(t)}]^+)]}{\partial x_j^{(t)}}. \tag{5.5}$$

The details of computing the gradients will be discussed later. After the above-mentioned approximation, the objective function of the dynamic programming model at period t, denoted as ETP_t, can be expressed as follows:

$$\begin{aligned}
&ETP_t \\
&= EP_t + f_{t+1}(x_1^{(t)}, x_2^{(t)}) \\
&\approx EP_t + u_1^{(t)} x_1^{(t)} + u_2^{(t)} x_2^{(t)} \\
&\approx EP_t + u_1^{(t)}(x_{S1}^{(t)} + p_1^{(t)} + r_{11}^{(t)} + r_{21}^{(t)}) + u_2^{(t)}(x_{S2}^{(t)} + p_2^{(t)} + r_{12}^{(t)} + r_{22}^{(t)}).
\end{aligned} \tag{5.6}$$

As mentioned in Chapter 3, the expected profit function, denoted as EP_t, is expressed as follows:

$$\begin{aligned}
&EP_t(x_{S1}^{(t)}, x_{S2}^{(t)}, p_1^{(t)}, p_2^{(t)}, r_{11}^{(t)}, r_{12}^{(t)}, r_{21}^{(t)}, r_{22}^{(t)}) \\
&= s_1\mu_1 + s_2\mu_2 - (c_{P1}p_1^{(t)} + c_{P2}p_2^{(t)} + c_{R11}r_{11}^{(t)} + c_{R12}r_{12}^{(t)} + c_{R21}r_{21}^{(t)} + c_{R22}r_{22}^{(t)}) \\
&\quad - h_1 \int_0^{x_{S1}^{(t)} + p_1^{(t)} + r_{11}^{(t)} + r_{21}^{(t)}} (x_{S1}^{(t)} + p_1^{(t)} + r_{11}^{(t)} + r_{21}^{(t)} - D_1^{(t)}) f(D_1^{(t)}, \mu_1, \sigma_1) dD_1^{(t)} \\
&\quad - (s_1 + v_1) \int_{x_{S1}^{(t)} + p_1^{(t)} + r_{11}^{(t)} + r_{21}^{(t)}}^{\infty} (D_1^{(t)} - x_{S1}^{(t)} - p_1^{(t)} - r_{11}^{(t)} - r_{21}^{(t)}) f(D_1^{(t)}, \mu_1, \sigma_1) dD_1^{(t)} \\
&\quad - h_2 \int_0^{x_{S2}^{(t)} + p_2^{(t)} + r_{12}^{(t)} + r_{22}^{(t)}} (x_{S2}^{(t)} + p_2^{(t)} + r_{12}^{(t)} + r_{22}^{(t)} - D_2^{(t)}) f(D_2^{(t)}, \mu_2, \sigma_2) dD_2^{(t)} \\
&\quad - (s_2 + v_2) \int_{x_{S2}^{(t)} + p_2^{(t)} + r_{12}^{(t)} + r_{22}^{(t)}}^{\infty} (D_2^{(t)} - x_{S2}^{(t)} - p_2^{(t)} - r_{12}^{(t)} - r_{22}^{(t)}) f(D_2^{(t)}, \mu_2, \sigma_2) dD_2^{(t)}.
\end{aligned}$$

$$(5.7)$$

After substituting Formula (5.7), Formula (5.6) can be further expressed as follows:

$$
\begin{aligned}
&ETP_t(x_{S1}^{(t)}, x_{S2}^{(t)}, p_1^{(t)}, p_2^{(t)}, r_{11}^{(t)}, r_{12}^{(t)}, r_{21}^{(t)}, r_{22}^{(t)}) \\
&\approx \tilde{ETP}_t(x_{S1}^{(t)}, x_{S2}^{(t)}, p_1^{(t)}, p_2^{(t)}, r_{11}^{(t)}, r_{12}^{(t)}, r_{21}^{(t)}, r_{22}^{(t)}) \\
&= (s_1 + u_1^{(t)})\mu_1 + (s_2 + u_2^{(t)})\mu_2 - (c_{P1}p_1^{(t)} + c_{P2}p_2^{(t)} + c_{R11}r_{11}^{(t)} + c_{R12}r_{12}^{(t)} + c_{R21}r_{21}^{(t)} + c_{R22}r_{22}^{(t)}) \\
&\quad - (h_1 - u_1^{(t)})\int_0^{x_{S1}^{(t)}+p_1^{(t)}+r_{11}^{(t)}+r_{21}^{(t)}} (x_{S1}^{(t)} + p_1^{(t)} + r_{11}^{(t)} + r_{21}^{(t)} - D_1^{(t)}) f(D_1^{(t)}, \mu_1, \sigma_1) dD_1^{(t)} \\
&\quad - (s_1 + v_1 + u_1^{(t)})\int_{x_{S1}^{(t)}+p_1^{(t)}+r_{11}^{(t)}+r_{21}^{(t)}}^{\infty} (D_1^{(t)} - x_{S1}^{(t)} - p_1^{(t)} - r_{11}^{(t)} - r_{21}^{(t)}) f(D_1^{(t)}, \mu_1, \sigma_1) dD_1^{(t)} \\
&\quad - (h_2 - u_2^{(t)})\int_0^{x_{S2}^{(t)}+p_2^{(t)}+r_{12}^{(t)}+r_{22}^{(t)}} (x_{S2}^{(t)} + p_2^{(t)} + r_{12}^{(t)} + r_{22}^{(t)} - D_2^{(t)}) f(D_2^{(t)}, \mu_2, \sigma_2) dD_2^{(t)} \\
&\quad - (s_2 + v_2 + u_2^{(t)})\int_{x_{S2}^{(t)}+p_2^{(t)}+r_{12}^{(t)}+r_{22}^{(t)}}^{\infty} (D_2^{(t)} - x_{S2}^{(t)} - p_2^{(t)} - r_{12}^{(t)} - r_{22}^{(t)}) f(D_2^{(t)}, \mu_2, \sigma_2) dD_2^{(t)}.
\end{aligned}
$$

(5.8)

The functions EP_t and $\tilde{ETP}_t$, expressed in Formulae (5.7) and (5.8) respectively, are found to be similar to each other except for some coefficient differences. Therefore, we can prove the concavity of the function $\tilde{ETP}_t$ similar to the function EP_t. Thus, we can find the optimal solution to maximize the function $\tilde{ETP}_t$ by solving KKT conditions. The optimal solution has the same structures as that for the single-period problem in Appendix B. Therefore, the optimal multi-level threshold policy of the single-period problem could be conveniently used for the multi-period problem. However, due to coefficient differences, the threshold levels for the multi-period problem need to be re-computed. For example, the threshold level $AL_0^{(t)}$, can be determined as follows:

$$\left.\frac{\partial \tilde{ETP}_t}{\partial p_1^{(t)}}\right|_{x_1^{(t)}=AL_0^{(t)}} = 0$$

$$\Rightarrow s_1 + u_1^{(t)} + v_1 - c_{P1} - (s_1 + v_1 + h_1)F(AL_0^{(t)}, \mu_1, \sigma_1) = 0 \tag{5.9}$$

$$\Rightarrow AL_0^{(t)} = F^{-1}(\frac{u_1^{(t)} + s_1 + v_1 - c_{P1}}{s_1 + v_1 + h_1}, \mu_1, \sigma_1).$$

Similarly, we can determine the other threshold levels at period *t* for the multi-period problem. In Table 5.1, we have listed the formulae of determining the threshold levels for the single-period problem and the multi-period problem respectively. It could be seen from the formulae that the threshold level for the multi-period problem is only dependent on the gradient of the cost-to-go function at the points of interest.

Unlike the usual approach which uses a single function (or piecewise function) to represent the cost-to-go function across the whole state space, we just need to estimate the gradient of the cost-to-go function at the points of interest. These gradients will be used to compute the threshold level. Hence, the performance of the results will not depend on the function we assume which can be a challenge for most of the approximate dynamic programming approaches.

As the gradients used to compute the threshold level are dependent on the threshold level conversely, we need to employ an iterative algorithm to find the threshold level. In the following, we would explain how to compute the threshold levels, taking the threshold level $AL_0^{(t)}$ as example.

Table 5.1 The formulae of determining the threshold levels for the single-period problem and the multi-period problem

	The single-period problem	The multi-period problem
$AL_0^{(t)}$	$F^{-1}(\frac{s_1+v_1-c_{P1}}{s_1+v_1+h_1},\mu_1,\sigma_1)$	$F^{-1}(\frac{u_1^{(t)}+s_1+v_1-c_{P1}}{s_1+v_1+h_1},\mu_1,\sigma_1)$
$AL_1^{(t)}$	$F^{-1}(\frac{s_1+v_1-c_{R11}}{s_1+v_1+h_1},\mu_1,\sigma_1)$	$F^{-1}(\frac{u_1^{(t)}+s_1+v_1-c_{R11}}{s_1+v_1+h_1},\mu_1,\sigma_1)$
$AL_2^{(t)}$	$F^{-1}(\frac{s_1+v_1-c_{R21}}{s_1+v_1+h_1},\mu_1,\sigma_1)$	$F^{-1}(\frac{u_1^{(t)}+s_1+v_1-c_{R21}}{s_1+v_1+h_1},\mu_1,\sigma_1)$
$SW_1^{(t)}$	$F^{-1}(\frac{s_1+v_1+c_{R12}-c_{R11}-c_{P2}}{s_1+v_1+h_1},\mu_1,\sigma_1)$	$F^{-1}(\frac{u_1^{(t)}+s_1+v_1+c_{R12}-c_{R11}-c_{P2}}{s_1+v_1+h_1},\mu_1,\sigma_1)$
$SW_2^{(t)}$	$F^{-1}(\frac{s_1+v_1+c_{R22}-c_{R21}-c_{P2}}{s_1+v_1+h_1},\mu_1,\sigma_1)$	$F^{-1}(\frac{u_1^{(t)}+s_1+v_1+c_{R22}-c_{R21}-c_{P2}}{s_1+v_1+h_1},\mu_1,\sigma_1)$
$RP^{(t)}$	$F^{-1}(\frac{s_1+v_1+c_{R22}-c_{R21}-c_{R12}}{s_1+v_1+h_1},\mu_1,\sigma_1)$	$F^{-1}(\frac{u_1^{(t)}+s_1+v_1+c_{R22}-c_{R21}-c_{R12}}{s_1+v_1+h_1},\mu_1,\sigma_1)$
$BL_0^{(t)}$	$F^{-1}(\frac{s_2+v_2-c_{P2}}{s_2+v_2+h_2},\mu_2,\sigma_2)$	$F^{-1}(\frac{u_2^{(t)}+s_2+v_2-c_{P2}}{s_2+v_2+h_2},\mu_2,\sigma_2)$
$BL_1^{(t)}$	$F^{-1}(\frac{s_2+v_2-c_{R12}}{s_2+v_2+h_2},\mu_2,\sigma_2)$	$F^{-1}(\frac{u_2^{(t)}+s_2+v_2-c_{R12}}{s_2+v_2+h_2},\mu_2,\sigma_2)$
$BL_2^{(t)}$	$F^{-1}(\frac{s_2+v_2-c_{R22}}{s_2+v_2+h_2},\mu_2,\sigma_2)$	$F^{-1}(\frac{u_2^{(t)}+s_2+v_2-c_{R22}}{s_2+v_2+h_2},\mu_2,\sigma_2)$

When determining the threshold level $AL_0^{(t)}$, the inventory level of product 1 after replenishment, i.e. $x_1^{(t)}$, is set as $AL_0^{(t)}$. As for the $x_2^{(t)}$, we can set it at any of the three order-up-to levels, i.e. $BL_0^{(t)}$, $BL_1^{(t)}$ and $BL_2^{(t)}$. If the cost-to-go function is separable with the inventory levels of the two products, the gradient $u_1^{(t)}$ will be independent of the values of $x_2^{(t)}$. However, the cost-to-go function might not actually be separable, and so we use the average of the gradients at the three points as the approximation of the gradient $u_1^{(t)}$ at $AL_0^{(t)}$. Thus, the gradient $u_1^{(t)}$ can be estimated as follows:

$$\begin{aligned} & u_1^{(t)}(x_1^{(t)}, x_2^{(t)} \mid x_1^{(t)} = AL_0^{(t)}) \\ & \approx \frac{1}{3}\sum_{k=0}^{2} u_1^{(t)}(x_1^{(t)}, x_2^{(t)} \mid x_1^{(t)} = AL_0^{(t)}, x_2^{(t)} = BL_k^{(t)}). \end{aligned} \tag{5.10}$$

To solve (5.10), we need to know $AL_0^{(t)}$ and $BL_k^{(t)}$ (k = 0, 1, 2). However, they can only be determined after having determined the corresponding gradients. Thus, we need to use an iterative approach to search for the threshold level $AL_0^{(t)}$ by using the pre-determined threshold levels of period t+1.

Firstly, the gradient $u_1^{(t)}$ is estimated by using the threshold levels of period t+1 as initial value, i.e. $x_1^{(t)} = AL_0^{(t+1)}, x_2^{(t)} = BL_k^{(t+1)}$. Then, the gradient is used to determine the threshold level $AL_0^{(t)}$ by using Table 5.1, which is further used by smoothing with previous value to obtain the latest value of the threshold level. After that, we re-compute the gradient using the latest threshold level $AL_0^{(t)}$. Then, the gradient is updated with previous value by smoothing, which is used to determine a new value of the threshold level. The computing and updating procedure is repeated until it converges. The algorithm is an iterative learning algorithm. Due to the time-consuming computation, we will stop the iteration when the approximation has satisfied our criterion. In the following algorithm, the stopping criterion can be tuned to suitable values according to the algorithm's performance. The notations of the algorithm are listed as follows:

m	the index for an iteration;
$\overline{u}_1^{(t),m}$	a weighted average of gradient $u_1^{(t)}$ up to iteration m;
$u_1^{(t),m}$	a stochastic gradient obtained at iteration m;
$\overline{AL}_0^{(t),m}$	the threshold level corresponding to a weighted average of gradient at iteration m;
$AL_0^{(t),m}$	a weighted average of the threshold level at iteration m;
$U_1^{(t)}(x_1^{(t)}, x_2^{(t)})$	the set of gradients of the cost-to-go function with respect to $x_1^{(t)}$ by averaging the gradients $u_1^{(t)}(x_1^{(t)}, x_2^{(t)} \mid x_1^{(t)} = AL_0^{(t),m-1}, x_2^{(t)} = BL_k^{(t+1)})$ $(k = 0, 1, 2)$;

The main steps of the algorithm are shown as follows:

Step 1. Set $\overline{u}_1^{(t),0} = u_1^{(t+1)}(x_1^{(t+1)}, x_2^{(t+1)} \mid x_1^{(t+1)} = AL_0^{(t+1)})$ and $AL_0^{(t),0} = AL_0^{(t+1)}$;

Step 2. Set $m = 1$;

Step 3. Obtain $u_1^{(t),m} \in U_1^{(t)}(x_1^{(t)}, x_2^{(t)})$;

Step 4. Update $\overline{u}_1^{(t),m}$ by

$$\overline{u}_1^{(t),m} = \overline{u}_1^{(t),m-1} + \beta_m(u_1^{(t),m} - \overline{u}_1^{(t),m-1});$$

Step 5. Obtain $\overline{AL}_0^{(t),m}$ by referring to the corresponding formula in Table 5.1;

Step 6. Update the solution by

$$AL_0^{(t),m} = AL_0^{(t),m-1} + \alpha_m(\overline{AL0}^{(t),m} - AL_0^{(t),m-1});$$

Step 7. Set $m = m + 1$. If the stopping criterion is not met, go to Step 3.

In general, we stop the algorithm if the total absolute change of the threshold level over a certain number of iterations is small. For example, if $\sum_{i=m-L+1,m>L}^{m} \left\| AL_0^{(t),i} - AL_0^{(t),i-1} \right\| < \delta$ (L: the number of consecutive results used for calculating the absolute change; δ: a small number.), then the algorithm will stop. In addition, because sampling is involved, stochastic gradients can be quite different from iteration to iteration. The instability of stochastic gradients makes the solution obtained in Step 5 fluctuating. Therefore, we need the averaging steps, i.e. Steps 4 and 6, to help stabilize the solution. According to Gupal and Bazhenov (1972), the solution will surely converge to an optimal solution under certain conditions on the stepsizes α_m and β_m, such as

$$\alpha_m \geq 0,\ \ \beta_m > 0,\ \ \alpha_m / \beta_m \to 0,\ \ \sum_{m=1}^{\infty} \alpha_m = \infty,\ \ \sum_{m=1}^{\infty} \alpha_m^2 < \infty.$$

After the algorithm stops, we can obtain the threshold level $AL_0^{(t)}$. Similarly, we can determine the other threshold levels of period t by using the algorithm. However, as mentioned before, we need to have determined the threshold levels of period t+1 at first. As the threshold levels of the last period, i.e. period M, have been determined by solving the single-period problem, we can take advantage of backward way to determine the threshold levels from period M till period t. In addition, the algorithm requires the gradients to be determined at the points of interest. We will introduce it in the following section.

5.3 The determination of the gradient at the points of interest in the multi-period context

As the objective function of the approximate dynamic programming model has similar concave property to the objective function of the single-period problem, we can refer to the optimal policy of the single-period problem for solving the multi-period problem. While solving the multi-period problem, we take advantage of backward induction. Firstly, for the last period of the multi-period horizon, i.e. period *M*, the threshold levels are determined with reference to the single-period problem. Then, for the second last period, the threshold levels are determined by estimating the gradients of the cost-to-go function, i.e. the gradients of the expected maximum expected profit in the last period, which is expressed as follows:

$$\begin{aligned} & u_j^{(M-1)} \\ & = \frac{\partial}{\partial x_j^{(M-1)}} \{ \underset{R_1^{(M)}, R_2^{(M)}, D_1^{(M-1)}, D_2^{(M-1)}}{E} [\max_{p_1^{(M)} \geq 0, p_2^{(M)} \geq 0, r_{11}^{(M)} \geq 0, r_{12}^{(M)} \geq 0, r_{21}^{(M)} \geq 0, r_{22}^{(M)} \geq 0} EP_M ([x_1^{(M-1)} - D_1^{(M-1)}]^+, \\ & [x_2^{(M-1)} - D_2^{(M-1)}]^+, p_1^{(M)}, p_2^{(M)}, r_{11}^{(M)}, r_{12}^{(M)}, r_{21}^{(M)}, r_{22}^{(M)})]\}. \end{aligned} \quad (5.11)$$

In this backward induction method, suppose we are now at period *t* to determine the threshold levels. Up to now, we have determined the threshold levels from period *t*+1 till the last period. Therefore, we have known how to make the optimal replenishment decisions based on the threshold levels and the optimal policy in these periods. In order to determine the threshold levels of period *t*, we need to estimate the gradients of the cost-to-go function, i.e. the gradients of the expected maximum expected total profit earned from period t+1 till the last period, which is expressed as follows:

$$u_j^{(t)} = \frac{\partial}{\partial x_j^{(t)}}\{ \underset{R_1^{(t+1)},R_2^{(t+1)},D_1^{(t)},D_2^{(t)}}{E}[\max_{p_1^{(t+1)}\geq 0, p_2^{(t+1)}\geq 0, r_{11}^{(t+1)}\geq 0, r_{12}^{(t+1)}\geq 0, r_{21}^{(t+1)}\geq 0, r_{22}^{(t+1)}\geq 0} ETP([x_1^{(t)}-D_1^{(t)}]^+,[x_2^{(t)}-D_2^{(t)}]^+, p_1^{(t+1)}, p_2^{(t+1)}, r_{11}^{(t+1)}, r_{12}^{(t+1)}, r_{21}^{(t+1)}, r_{22}^{(t+1)})]\}. \tag{5.12}$$

As there is no closed-form formula to compute the gradient $u_j^{(t)}$, we need to run Monte Carlo simulation, and estimate the gradient based on the simulation results. Firstly, we need to generate N sets of random realization of stochastic returns and demands in each period from period t till period M-1. Among them, sample k is expressed as $\begin{pmatrix} R_{1,k}^{(t+1)} & R_{2,k}^{(t+1)} & D_{1,k}^{(t)} & D_{2,k}^{(t)} \\ .... & & & \\ R_{1,k}^{(M)} & R_{2,k}^{(M)} & D_{1,k}^{(M-1)} & D_{2,k}^{(M-1)} \end{pmatrix}$. The sample value of the cost-to-go function for sample k, is obtained by summing up the profit for the realization from period t+1 till period M-1 and the expected profit at period M after applying the optimal policy for these periods. The $ATP_k^{(\tau)*}$ is used to calculate the profit of period τ ($t < \tau < M$) for sample k. The profit function $ATP_k^{(\tau)*}$ is expressed as follows:

$$\begin{aligned} &ATP_k^{(\tau)*}(x_{S1}^{(\tau)}, x_{S2}^{(\tau)}, p_1^{(\tau)*}, p_2^{(\tau)*}, r_{11}^{(\tau)*}, r_{12}^{(\tau)*}, r_{21}^{(\tau)*}, r_{22}^{(\tau)*}, D_{1,k}^{(\tau)}, D_{2,k}^{(\tau)}) \\ &= s_1(x_{S1}^{(\tau)} + p_1^{(\tau)*} + r_{11}^{(\tau)*} + r_{21}^{(\tau)*}) + s_2(x_{S2}^{(\tau)} + p_2^{(\tau)*} + r_{12}^{(\tau)*} + r_{22}^{(\tau)*}) \\ &-(c_{P1}p_1^{(\tau)*} + c_{P2}p_2^{(\tau)*} + c_{R11}r_{11}^{(\tau)*} + c_{R12}r_{12}^{(\tau)*} + c_{R21}r_{21}^{(\tau)*} + c_{R22}r_{22}^{(\tau)*}) \\ &-(h_{s1}+s_1)[x_{S1}^{(\tau)} + p_1^{(\tau)*} + r_{11}^{(\tau)*} + r_{21}^{(\tau)*} - D_{1,k}^{(\tau)}]^+ - v_1[D_{1,k}^{(\tau)} - x_{S1}^{(\tau)} - p_2^{(\tau)*} - r_{12}^{(\tau)*} - r_{22}^{(\tau)*}]^+ \\ &-(h_{s2}+s_2)[x_{S2}^{(\tau)} + p_2^{(\tau)*} + r_{12}^{(\tau)*} + r_{22}^{(\tau)*} - D_{2,k}^{(\tau)}]^+ - v_2[D_{2,k}^{(\tau)} - x_{S2}^{(\tau)} - p_2^{(\tau)*} - r_{12}^{(\tau)*} - r_{22}^{(\tau)*}]^+. \end{aligned} \tag{5.13}$$

As there is no closed-form formula for the maximum expected profit at final period, i.e. MEP_k , we compute it by maximizing the expected profit $EP_{\mathrm{M,\,k}}$ for sample k, which is calculated as Formula (5.7).

Therefore, the cost-to-go function can be estimated as follows:

$$\begin{aligned}&\underset{D_1^{(t)},D_2^{(t)}}{E}[f_{t+1}(x_{S1}^{(t+1)},x_{S2}^{(t+1)})]\\&\approx\frac{1}{N}\sum_{k=1}^{N}(\sum_{\tau=t+1}^{M-1}ATP_k^{(\tau)*}+MEP_k).\end{aligned}\tag{5.14}$$

As the function $ATP_k^{(\tau)*}$ and the function MEP_k are both continuous functions, it is suitable to approximate the cost-to-go function by Monte Carlo sampling method. Furthermore, the gradient $u_j^{(t)}$ can be approximated by sample average of the gradient over all the realizations. The approximation of the gradient $u_j^{(t)}$ by averaging the sample gradient $grad_{j,k}^{(t)}$ for sample k is expressed as follows:

$$u_j^{(t)}(x_1^{(t)},x_2^{(t)})\approx\frac{1}{N}\sum_{k=1}^{N}grad_{j,k}^{(t)}(x_1^{(t)},x_2^{(t)}).\tag{5.15}$$

We would start from a two-period problem to introduce the determination of the sample gradient. Then, we would extend from the two-period problem to the three-period problem. Finally, we can determine the sample gradient for any multi-period problem by induction. In the determination of the sample gradient, we have taken advantage of an Infinite Perturbation Analysis (IPA) based approach.

5.3.1 The determination of sample gradient in the two-period problem

For the two-period problem, the threshold levels of the last period can be obtained by referring to the single-period problem. The threshold levels of the first period are determined by using the gradients of the cost-to-go function estimated by Monte Carlo simulation. Before that, the sample gradient of the cost-to-go function needs to be determined. The sample k for Monte Carlo sampling is expressed as $(R_{1,k}^{(2)}, R_{2,k}^{(2)}, D_{1,k}^{(1)}, D_{2,k}^{(1)})$. The sample gradient can be calculated as follows (j =1, 2):

$$grad_{j,k}^{(1)}(x_1^{(1)}, x_2^{(1)}) = \frac{\partial MEP_k}{\partial x_j^{(1)}} = \frac{\partial MEP_k}{\partial x_{S1}^{(2)}}\frac{\partial x_{S1}^{(2)}}{\partial x_j^{(1)}} + \frac{\partial MEP_k}{\partial x_{S2}^{(2)}}\frac{\partial x_{S2}^{(2)}}{\partial x_j^{(1)}}. \tag{5.16}$$

As $x_1^{(1)}$ and $x_2^{(1)}$ are assumed to be independent of each other, the perturbation of $x_j^{(1)}$ will be only propagated to $x_{Sj}^{(1)}$ if there is no shortage of product j at the first period. Therefore, Formula (5.16) is further expressed as follows (j = 1, 2):

$$grad_{j,k}^{(1)}(x_1^{(1)}, x_2^{(1)}) = \begin{cases} 0, & \text{if } x_j^{(1)} < D_{j,k}^{(1)}; \\ \dfrac{\partial MEP_k}{\partial x_{Sj}^{(2)}}, & \text{otherwise.} \end{cases} \tag{5.17}$$

The term $\frac{\partial MEP_k}{\partial x_{Sj}^{(2)}}$ in Formula (5.17) can be computed as follows by the related derivatives of the expected profit function $EP_{2,\,k}$, which is expressed in Formula (5.7):

$$
\begin{aligned}
&\frac{\partial MEP_k}{\partial x_{Sj}^{(2)}} \\
&= \frac{\partial EP_{2,k}{}^*}{\partial x_{Sj}^{(2)}} + \left(\frac{\partial EP_{2,k}{}^*}{\partial p_1^{(2)*}} \cdot \frac{\partial p_1^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial EP_{2,k}{}^*}{\partial r_{11}^{(2)*}} \cdot \frac{\partial r_{11}^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial EP_{2,k}{}^*}{\partial r_{21}^{(2)*}} \cdot \frac{\partial r_{21}^{(2)*}}{\partial x_{Sj}^{(2)}}\right) \\
&\qquad + \left(\frac{\partial EP_{2,k}{}^*}{\partial p_2^{(2)*}} \cdot \frac{\partial p_2^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial EP_{2,k}{}^*}{\partial r_{12}^{(2)*}} \cdot \frac{\partial r_{12}^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial EP_{2,k}{}^*}{\partial r_{22}^{(2)*}} \cdot \frac{\partial r_{22}^{(2)*}}{\partial x_{Sj}^{(2)}}\right).
\end{aligned} \tag{5.18}
$$

According to Formula (5.7), the partial derivatives of the function $EP_{2,k}{}^*$ can be determined as follows ($j = 1, 2$):

$$
\begin{aligned}
&\frac{\partial EP_{2,k}{}^*}{\partial x_{Sj}^{(2)}} = s_j + v_j - (s_j + v_j + h_j) F(x_j^{(2)*}, \mu_j, \sigma_j); \\
&\frac{\partial EP_{2,k}{}^*}{\partial p_j^{(2)*}} = \frac{\partial EP_{2,k}{}^*}{\partial x_{Sj}^{(2)}} - c_{Pj}; \\
&\frac{\partial EP_{2,k}{}^*}{\partial r_{1j}^{(2)*}} = \frac{\partial EP_{2,k}{}^*}{\partial x_{Sj}^{(2)}} - c_{R1j}; \\
&\frac{\partial EP_{2,k}{}^*}{\partial r_{2j}^{(2)*}} = \frac{\partial EP_{2,k}{}^*}{\partial x_{Sj}^{(2)}} - c_{R2j}.
\end{aligned} \tag{5.19}
$$

Based on Formulae (5.18) and (5.19), Formula (5.17) is further expressed as follows ($j = 1, 2$):

$$
grad_{j,k}^{(1)}(x_1^{(1)}, x_2^{(1)}) = \begin{cases} 0, \quad \text{if } x_j^{(1)} < D_{j,k}^{(1)}; \\[2em] \dfrac{\partial EP_{2,k}{}^*}{\partial x_{S1}^{(2)}} \dfrac{\partial x_1^{(2)*}}{\partial x_{Sj}^{(2)}} + \dfrac{\partial EP_{2,k}{}^*}{\partial x_{S2}^{(2)}} \dfrac{\partial x_2^{(2)*}}{\partial x_{Sj}^{(2)}} - c_{P1} \dfrac{\partial p_1^{(2)*}}{\partial x_{Sj}^{(2)}} - c_{R11} \dfrac{\partial r_{11}^{(2)*}}{\partial x_{Sj}^{(2)}} \\ \qquad - c_{R21} \dfrac{\partial r_{21}^{(2)*}}{\partial x_{Sj}^{(2)}} - c_{P2} \dfrac{\partial p_2^{(2)*}}{\partial x_{Sj}^{(2)}} - c_{R12} \dfrac{\partial r_{12}^{(2)*}}{\partial x_{Sj}^{(2)}} - c_{R22} \dfrac{\partial r_{22}^{(2)*}}{\partial x_{Sj}^{(2)}}, \quad \text{otherwise.} \end{cases} \tag{5.20}
$$

The above formula involves the partial derivatives of the optimal replenishment decisions with respect to initial inventory. These partial derivatives can be obtained according to the corresponding structure in Appendix B. Suppose that the optimal replenishment decisions for sample *k* match the structure S7 in the Appendix as follows:

S7. $R_{2,k}^{(1)} + x_{S1}^{(2)} < SW_1,\ R_{1,k}^{(1)} + R_{2,k}^{(1)} + x_{S1}^{(2)} > SW_1,\ R_{1,k}^{(1)} + R_{2,k}^{(1)} + x_{S1}^{(2)} + x_{S2}^{(2)} \le SW_1 + BL_0$:

$$p_1^{(2)*} = 0, \qquad r_{11}^{(2)*} = SW_1 - R_{2,k}^{(1)} - x_{S1}^{(2)}, \qquad r_{21}^{(2)*} = R_{2,k}^{(1)};\ (x_1^{(2)*} = SW_1)$$
$$p_2^{(2)*} = SW_1 + BL_0 - R_{1,k}^{(1)} - R_{2,k}^{(1)} - x_{S1}^{(2)} - x_{S2}^{(2)},\ r_{12}^{(2)*} = R_{1,k}^{(1)} + R_{2,k}^{(1)} + x_{S1}^{(2)} - SW_1,\ r_{22}^{(2)*} = 0. \quad (x_2^{(2)*} = BL_0)$$

From the structure S7, we can observe that the inventory levels of the two products after replenishment have reached the threshold levels SW_1 and BL_0 respectively. Therefore, the perturbation on the initial inventory of the two products will not be propagated to the order-up-to level of the two products, i.e. $\frac{\partial x_1^{(2)*}}{\partial x_{S1}^{(2)}} = \frac{\partial x_2^{(2)*}}{\partial x_{S1}^{(2)}} = \frac{\partial x_1^{(2)*}}{\partial x_{S2}^{(2)}} = \frac{\partial x_2^{(2)*}}{\partial x_{S2}^{(2)}} = 0$. However, the perturbation would affect the related replenishment decisions. The impact can be determined as: $\frac{\partial r_{11}^{(2)*}}{\partial x_{S1}^{(2)}} = -1$, $\frac{\partial r_{12}^{(2)*}}{\partial x_{S1}^{(2)}} = 1$, $\frac{\partial p_2^{(2)*}}{\partial x_{S1}^{(2)}} = -1$ and $\frac{\partial p_2^{(2)*}}{\partial x_{S2}^{(2)}} = -1$. Thus, we can conclude that if the initial inventory of product 1 is increased with a small amount Δ, the recovery using returned items in group 1 for product 1 will be saved with the same amount Δ and be reallocated to product 2 in place of production for product 2; on the other hand, if the initial inventory of product 2 is increased with a small amount Δ, the same amount of production for product 2 will be saved. Despite the impact on the related replenishment decisions, the order-up-to level of each product remains unaffected.

According to Formula (5.20), the sample gradient is calculated as follows:

$$grad_{1,k}^{(1)}(x_1^{(1)}, x_2^{(1)}) = c_{R11} + c_{P2} - c_{R12};$$

$$grad_{2,k}^{(1)}(x_1^{(1)}, x_2^{(1)}) = c_{P2}.$$

5.3.2 The determination of sample gradient in the three-period problem

For the three-period problem, the threshold levels of the second and the last period can be obtained by referring to the two-period problem and the single-period problem respectively. Then, the threshold levels of the first period are determined by using the gradients of the cost-to-go function estimated by Monte Carlo simulation. Before that, the sample gradient of the cost-to-go function needs to be determined. The sample k for Monte Carlo sampling is expressed as $\begin{pmatrix} R_{1,k}^{(2)} & R_{2,k}^{(2)} & D_{1,k}^{(1)} & D_{2,k}^{(1)} \\ R_{1,k}^{(3)} & R_{2,k}^{(3)} & D_{1,k}^{(2)} & D_{2,k}^{(2)} \end{pmatrix}$. The sample gradient can be calculated as follows (j = 1, 2):

$$\begin{aligned}
& grad_{j,k}^{(1)}(x_1^{(1)}, x_2^{(1)}) \\
&= \frac{\partial}{\partial x_j^{(1)}} \{ATP_k^{(2)*} + MEP_k\} \\
&= \frac{\partial ATP_k^{(2)*}}{\partial x_j^{(1)}} + \frac{\partial MEP_k}{\partial x_j^{(1)}} \\
&= \frac{\partial ATP_k^{(2)*}}{\partial x_j^{(1)}} + \frac{\partial MEP_k}{\partial x_1^{(2)*}} \frac{\partial x_1^{(2)*}}{\partial x_j^{(1)}} + \frac{\partial MEP_k}{\partial x_2^{(2)*}} \frac{\partial x_2^{(2)*}}{\partial x_j^{(1)}} \\
&= \frac{\partial ATP_k^{(2)*}}{\partial x_j^{(1)}} + \frac{\partial x_1^{(2)*}}{\partial x_j^{(1)}} grad_{1,k}^{(2)}(x_1^{(2)*}, x_2^{(2)*}) + \frac{\partial x_2^{(2)*}}{\partial x_j^{(1)}} grad_{2,k}^{(2)}(x_1^{(2)*}, x_2^{(2)*}).
\end{aligned} \tag{5.21}$$

As the inventory levels of the two products $x_1^{(1)}$ and $x_2^{(1)}$ are assumed to be independent of each other, the perturbation of the $x_j^{(1)}$ will be only propagated to the $x_{Sj}^{(2)}$ if there is no shortage of product *j* at the first period. Therefore, Formula (5.21) is further expressed as follows (j = 1, 2):

$$grad_{j,k}^{(1)}(x_1^{(1)}, x_2^{(1)}) = \begin{cases} 0, & \text{if } x_j^{(1)} < D_{j,k}^{(1)}; \\ \dfrac{\partial ATP_k^{(2)*}}{\partial x_{Sj}^{(2)}} + \dfrac{\partial x_1^{(2)*}}{\partial x_{Sj}^{(2)}} grad_{1,k}^{(2)}(x_1^{(2)*}, x_2^{(2)*}) & \\ \quad + \dfrac{\partial x_2^{(2)*}}{\partial x_{Sj}^{(2)}} grad_{2,k}^{(2)}(x_1^{(2)*}, x_2^{(2)*}), & \text{otherwise.} \end{cases} \tag{5.22}$$

In the above formula, the perturbation of the initial inventory of product *j* at period *2* will be propagated to the final inventory of product *i* at the same period as follows (i, j = 1, 2):

$$\frac{\partial x_i^{(2)*}}{\partial x_{Sj}^{(2)}} = \frac{\partial x_{Si}^{(2)}}{\partial x_{Sj}^{(2)}} + \frac{\partial p_i^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial r_{1i}^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial r_{2i}^{(2)*}}{\partial x_{Sj}^{(2)}}. \tag{5.23}$$

The above formula involves the partial derivatives of the optimal replenishment decisions with respect to the initial inventory of the two products at period *2*. Similar to the two-period problem, these derivatives can be obtained by referring to the corresponding structure of the optimal replenishment decisions in Appendix B. Before that, the threshold levels at period *2* are computed on the basis of the objective function $\tilde{ETP}_2$ considering the final two periods. The two gradients of $\tilde{ETP}_2$ at the point of interest need to be estimated in the two-period context. Suppose that the optimal replenishment decisions at period *2* for sample *k* match the structure S7. With the solution structure mentioned before, the perturbation on the initial inventory of the two products will not affect the order-up-to level of the two products. Therefore, the sample gradient for sample *k* can be calculated as $grad_{j,k}^{(1)}(x_1^{(1)}, x_2^{(1)}) = \frac{\partial ATP_k^{(2)*}}{\partial x_{Sj}^{(2)}}$ according to Formula (5.22).

In Formula (5.22), the partial derivatives of the function $ATP_k^{(2)*}$ can be determined as follows (j = 1, 2):

$$\begin{aligned}&\frac{\partial ATP_k^{(2)*}}{\partial x_{Sj}^{(2)}}\\&=\frac{\partial ATP_k^{(2)*}}{\partial x_{Sj}^{(2)}}+\left(\frac{\partial ATP_k^{(2)*}}{\partial p_1^{(2)*}}\frac{\partial p_1^{(2)*}}{\partial x_{Sj}^{(2)}}+\frac{\partial ATP_k^{(2)*}}{\partial r_{11}^{(2)*}}\frac{\partial r_{11}^{(2)*}}{\partial x_{Sj}^{(2)}}+\frac{\partial ATP_k^{(2)*}}{\partial r_{21}^{(2)*}}\frac{\partial r_{21}^{(2)*}}{\partial x_{Sj}^{(2)}}\right)\\&\qquad+\left(\frac{\partial ATP_k^{(2)*}}{\partial p_2^{(2)*}}\frac{\partial p_2^{(2)*}}{\partial x_{Sj}^{(2)}}+\frac{\partial ATP_k^{(2)*}}{\partial r_{12}^{(2)*}}\frac{\partial r_{12}^{(2)*}}{\partial x_{Sj}^{(2)}}+\frac{\partial ATP_k^{(2)*}}{\partial r_{22}^{(2)*}}\frac{\partial r_{22}^{(2)*}}{\partial x_{Sj}^{(2)}}\right).\end{aligned} \tag{5.24}$$

According to Formula (5.13), the related partial derivatives of the function $ATP_k^{(2)*}$ are listed in Table 5.2 as follows. While calculating these partial derivatives, we have considered all the combinations of demand satisfaction. In order to summarize all the possible expressions, the related index and indicator are excluded from the notations in Table 5.2.

Table 5.2 The partial derivatives of the function $ATP_k^{(\tau)*}$ with respect to initial inventory and replenishment decisions

	$\frac{\partial ATP}{\partial x_{S1}}$	$\frac{\partial ATP}{\partial x_{S2}}$	$\frac{\partial ATP}{\partial p_1}$	$\frac{\partial ATP}{\partial r_{11}}$	$\frac{\partial ATP}{\partial r_{21}}$	$\frac{\partial ATP}{\partial p_2}$	$\frac{\partial ATP}{\partial r_{12}}$	$\frac{\partial ATP}{\partial r_{22}}$
Product 1 underage product 2 underage	s_1+v_1	s_2+v_2	$s_1+v_1-c_{P1}$	$s_1+v_1-c_{R11}$	$s_1+v_1-c_{R21}$	$s_2+v_2-c_{P2}$	$s_2+v_2-c_{R12}$	$s_2+v_2-c_{R22}$
Product 1 underage product 2 overage	s_1+v_1	$-h_2$	$s_1+v_1-c_{P1}$	$s_1+v_1-c_{R11}$	$s_1+v_1-c_{R21}$	$-(h_2+c_{P2})$	$-(h_2+c_{R12})$	$-(h_2+c_{R22})$
Product 1 overage product 2 underage	$-h_1$	s_2+v_2	$-(h_1+c_{P1})$	$-(h_1+c_{R11})$	$-(h_1+c_{R21})$	$s_2+v_2-c_{P2}$	$s_2+v_2-c_{R12}$	$s_2+v_2-c_{R22}$
Product 1 overage product 2 overage	$-h_1$	$-h_2$	$-(h_1+c_{P1})$	$-(h_1+c_{R11})$	$-(h_1+c_{R21})$	$-(h_2+c_{P2})$	$-(h_2+c_{R12})$	$-(h_2+c_{R22})$

With reference to Formula (5.24), the two gradients for sample k are calculated as follows. According to the situation of demand satisfaction in period *2*, we can refer to Table 5.2 to obtain the values of the related partial derivatives in the following formulae.

$$grad_{1,k}^{(1)}(x_1^{(1)}, x_2^{(1)})$$
$$= \frac{\partial ATP_k^{(2)*}}{\partial x_{S1}^{(2)}} + \frac{\partial ATP_k^{(2)*}}{\partial r_{12}^{(2)*}} - \frac{\partial ATP_k^{(2)*}}{\partial r_{11}^{(2)*}} - \frac{\partial ATP_k^{(2)*}}{\partial p_2^{(2)*}} = c_{R11} + c_{P2} - c_{R12};$$

$$grad_{2,k}^{(1)}(x_1^{(1)}, x_2^{(1)})$$
$$= \frac{\partial ATP_k^{(2)*}}{\partial x_{S2}^{(2)}} - \frac{\partial ATP_k^{(2)*}}{\partial p_2^{(2)*}} = c_{P2}.$$

5.3.3 The determination of sample gradient in the *N*-period problem

Since we have learned the induction rule in the determination of the sample gradient from the two-period problem to the three-period problem, we would extend to any *N*-period problem in order to determine the sample gradient at any period *t* of the multi-period horizon for the two-product recovery system. Before considering the *N*-period problem, suppose that we have known how to determine the sample gradient for all the multi-period problems with less than N periods. Therefore, the threshold levels of each period except the first period can be determined for the N-period problem. With these threshold levels, the optimal policy helps to make the optimal replenishment decisions in these periods. In this situation, the sample gradient of the first period for the N-period problem can be calculated as follows (j = 1, 2):

$$\begin{aligned}
& grad_{j,k}^{(1)}(x_1^{(1)}, x_2^{(1)}) \\
&= \frac{\partial}{\partial x_j^{(1)}}\{\sum_{\tau=2}^{N-1} ATP_k^{(\tau)*} + MEP_k\} \\
&= \frac{\partial ATP_k^{(2)*}}{\partial x_j^{(1)}} + \frac{\partial}{\partial x_j^{(1)}}\{\sum_{\tau=3}^{N-1} ATP_k^{(\tau)*} + MEP_k\} \\
&= \frac{\partial ATP_k^{(2)*}}{\partial x_j^{(1)}} + \frac{\partial x_1^{(2)*}}{\partial x_j^{(1)}} grad_{1,k}^{(2)}(x_1^{(2)*}, x_2^{(2)*}) + \frac{\partial x_2^{(2)*}}{\partial x_j^{(1)}} grad_{2,k}^{(2)}(x_1^{(2)*}, x_2^{(2)*}).
\end{aligned} \tag{5.25}$$

The above formula is similar to Formula (5.21) for the sample gradient of the three-period problem, which indicates the same induction rule. Therefore, by induction, we can determine the sample gradient for any N-period problem. For the sample gradient at period t of the M-period horizon, we can take advantage of backward way. In more details, the sample gradient at period *M-1* can be determined by solving the two-period problem considering the final two periods. Then the sample gradient at period *M-2* can be determined by solving the three-period problem considering the final three periods. Finally, in this way of backward induction, the sample gradient for period t can be determined. The process of determining the sample gradient can be referred to in Appendix C.

5.4 Computational results

5.4.1 The convergence of the threshold levels with period

Firstly, we investigate the impact of inventory holding cost rate on the convergence of the threshold levels. A set of system parameters is given as follows:

Cost: $v_1 = 4, v_2 = 6, s_1 = 15, s_2 = 20,$
$c_{P1} = 12, c_{P2} = 15, c_{R11} = 6, c_{R12} = 10, c_{R21} = 2, c_{R22} = 7;$
Demand: $E[D_1] = 200, StDev[D_1] = 60; E[D_2] = 100, StDev[D_2] = 30;$
Return: $E[R_1] = 210, StDev[R_1] = 70; E[R_2] = 45, StDev[R_2] = 15.$

In the following, we have shown the results about the threshold levels when the inventory holding cost rates h_1 and h_2 are both equal to 1, 2 and 3 respectively. Once the difference of all the threshold levels between two consecutive periods is no more than 1%, the convergence is regarded to have been achieved. As shown from the following results, the convergence takes place at the 13th last period, the 7th last period and the 6th last period when the inventory holding cost rates h_1 and h_2 are both equal to 1, 2 and 3 respectively. The results have shown that the threshold levels converge faster if the inventory holding cost rate is higher. When the inventory holding cost rate is high, the trade-off between inventory holding cost and penalty cost of inventory shortage is fulfilled at periods, which are not far from the end of the horizon. Thus, we can refer to the converged threshold levels while making production and recovery decisions for each period of a relatively long finite horizon.

Table 5.3 The threshold levels of each period for the 15-period problem when h_1=h_2=1

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
***M*=1**	176.9	223.1	262.2	184.8	192.5	231.5	93.0	107.0	116.1
***M*=2**	205.9	309.1	445.4	217.4	229.4	351.5	111.6	161.7	202.1
***M*=3**	216.6	370.8	608.5	229.9	244.4	452.3	117.9	202.5	281.9
***M*=4**	221.3	398.5	756.3	235.7	251.6	496.4	119.9	216.7	355.9
***M*=5**	223.2	403.8	883.2	237.8	254.4	521.7	120.6	228.0	422.8
***M*=6**	223.9	415.0	976.8	238.6	255.9	539.5	120.8	245.0	472.2
***M*=7**	224.5	429.8	1034.7	239.5	257.2	556.3	120.7	261.3	502.3
***M*=8**	225.0	441.9	1068.7	240.0	258.0	576.9	120.6	271.0	521.3
***M*=9**	224.9	450.7	1096.2	240.2	258.7	593.0	120.5	276.4	536.6
***M*=10**	225.3	454.1	1124.7	240.4	259.2	603.0	120.5	280.5	552.9
***M*=11**	225.2	457.7	1153.6	240.9	258.9	608.2	120.2	281.8	564.9
***M*=12**	225.0	459.8	1176.4	240.5	259.2	610.4	120.2	282.6	571.6
***M*=13**	224.8	459.7	1191.0	240.4	259.2	612.4	120.3	283.9	573.4
***M*=14**	224.9	459.7	1202.5	240.5	259.3	614.7	120.2	285.5	576.4
***M*=15**	224.8	459.8	1210.3	240.6	259.2	616.4	120.2	285.8	576.3

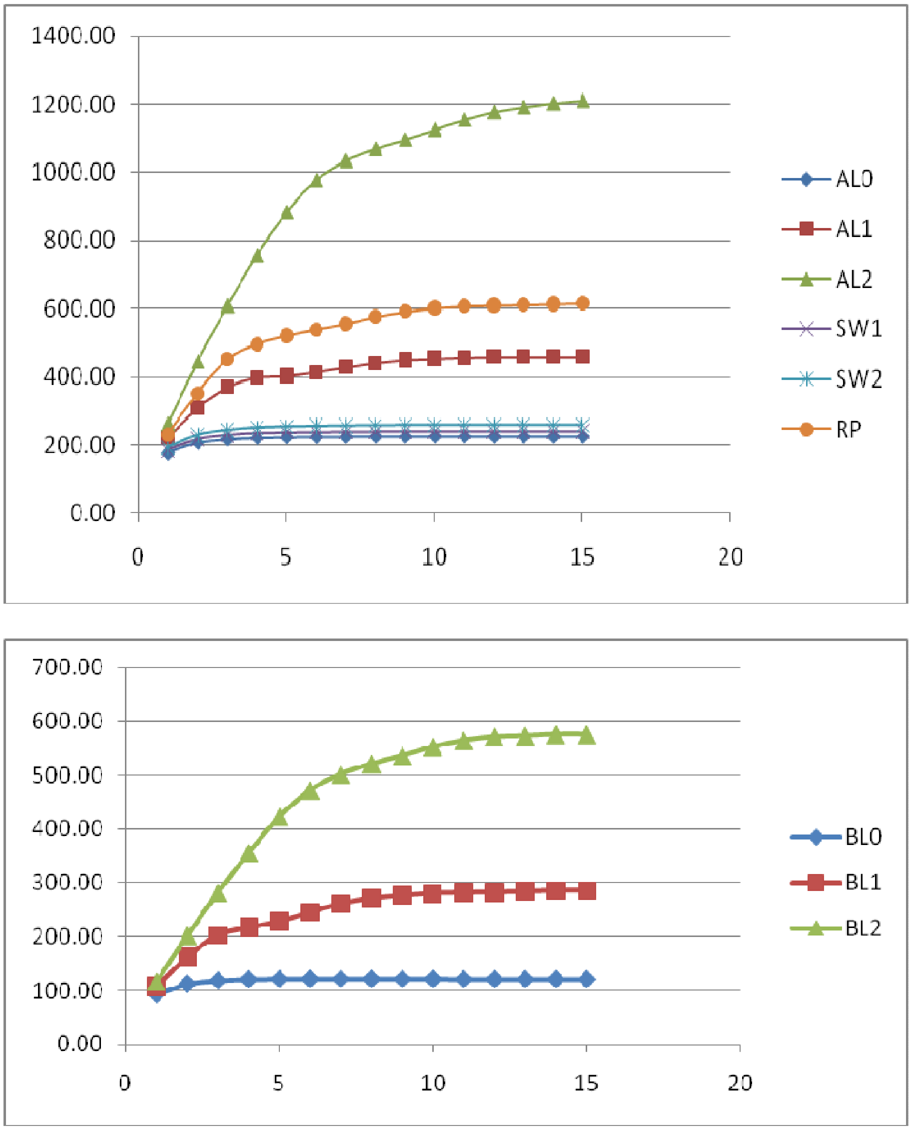
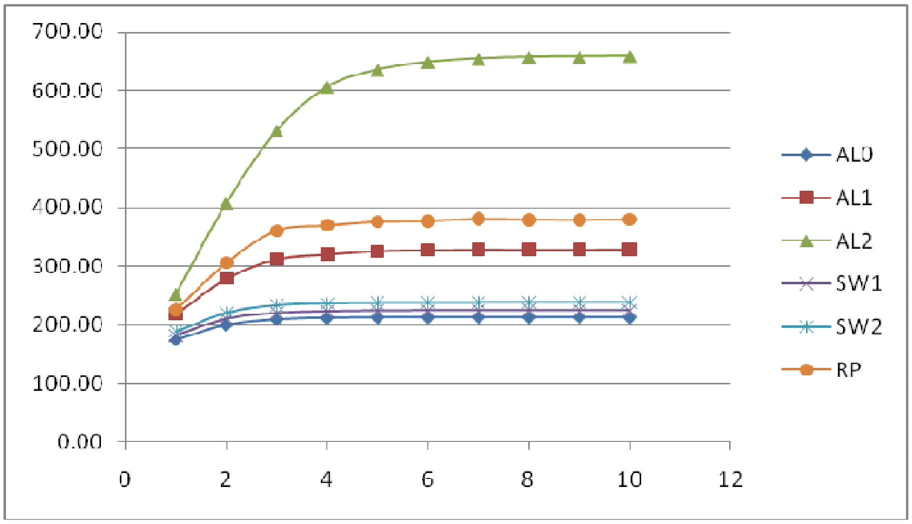


Figure 5.1 The trend of the threshold levels when $h_1=h_2=1$

Table 5.4 The threshold levels of each period for the 10-period problem when $h_1=h_2=2$

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
***M*=1**	174.2	218.2	252.6	181.8	189.2	225.8	91.8	105.4	113.9
***M*=2**	199.8	279.1	408.3	210.4	220.8	306.4	108.5	141.5	189.7
***M*=3**	209.2	312.2	531.6	220.4	232.9	361.3	114.0	171.7	253.8
***M*=4**	211.7	320.6	606.3	223.4	236.3	369.9	115.3	178.8	295.0
***M*=5**	212.6	326.2	635.4	224.5	237.3	376.2	115.9	181.8	309.1
***M*=6**	212.8	327.6	648.6	224.8	237.7	377.6	116.1	183.1	316.1
***M*=7**	212.9	328.6	654.4	225.1	237.8	381.0	116.2	184.3	319.8
***M*=8**	213.0	328.4	657.0	225.2	238.2	379.5	116.1	185.0	320.7
***M*=9**	213.0	328.4	657.8	224.9	238.1	379.3	116.3	185.3	322.0
***M*=10**	212.8	329.0	658.8	224.9	238.0	379.7	116.4	186.0	322.1



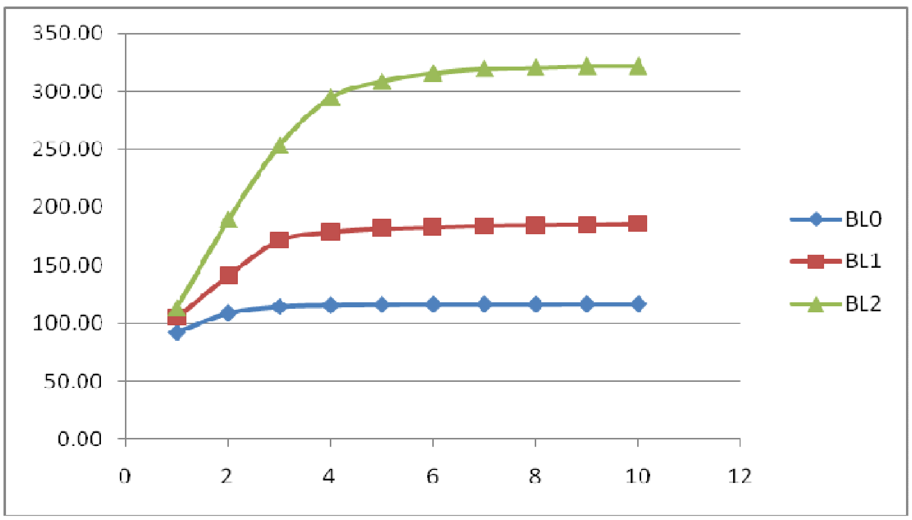


Figure 5.2 The trend of the threshold levels when h_1=h_2=2

Table 5.5 The threshold levels of each period for the 10-period problem when h_1=h_2=3

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
***M*=1**	171.6	213.8	244.9	179.1	186.2	220.9	90.8	103.9	112.0
***M*=2**	194.9	261.5	373.0	204.4	214.1	279.1	105.9	132.1	176.4
***M*=3**	202.5	280.7	460.8	212.8	223.7	307.4	110.7	146.0	223.9
***M*=4**	204.4	285.0	489.9	215.1	226.0	313.6	112.0	151.9	232.7
***M*=5**	205.1	286.8	498.2	215.7	226.5	316.0	112.2	153.3	233.8
***M*=6**	205.2	287.8	502.6	216.1	226.8	317.3	112.2	153.6	234.6
***M*=7**	205.4	287.7	504.6	216.0	227.1	317.3	112.2	155.0	234.8
***M*=8**	205.4	287.7	503.2	215.9	227.0	317.4	112.2	155.1	235.3
***M*=9**	205.6	287.7	503.5	216.0	226.9	316.7	112.2	154.4	235.6
***M*=10**	205.4	287.7	503.9	216.0	227.2	317.1	112.3	154.9	235.8

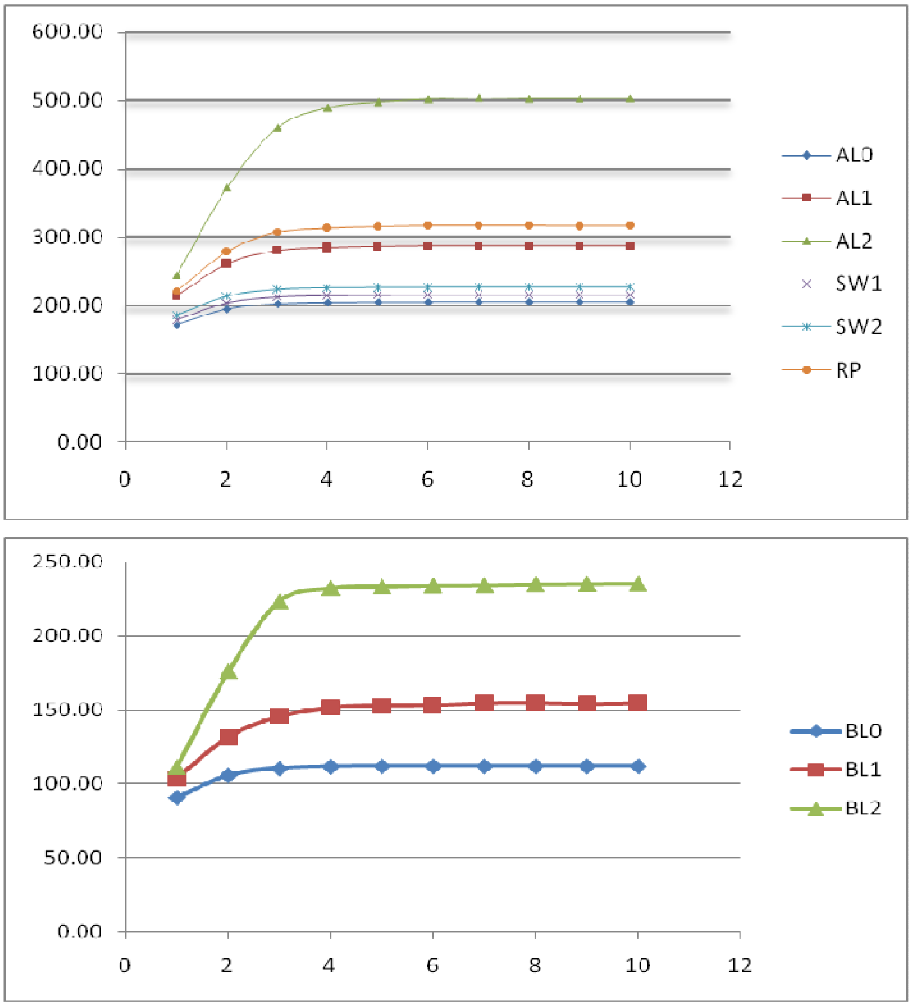


Figure 5.3 The trend of the threshold levels when h_1=h_2=3

5.4.2 The impact of stochastic returns and demands on the threshold levels

Three sets of system parameters are given as follows:

$$\begin{aligned}&S1{:}h_1=3,h_2=3,v_1=4,v_2=6,s_1=15,s_2=20,\\&\quad c_{P1}=12,c_{P2}=15,c_{R11}=6,c_{R12}=10,c_{R21}=2,c_{R22}=7.\end{aligned}$$

$$\begin{aligned}&S2{:}h_1=3,h_2=3,v_1=4,v_2=6,s_1=18,s_2=20,\\&\quad c_{P1}=16,c_{P2}=15,c_{R11}=10,c_{R12}=10,c_{R21}=2,c_{R22}=7.\end{aligned}$$

$$\begin{aligned}&S3{:}h_1=3,h_2=3,v_1=4,v_2=6,s_1=21,s_2=20,\\&\quad c_{P1}=20,c_{P2}=15,c_{R11}=14,c_{R12}=10,c_{R21}=2,c_{R22}=7.\end{aligned}$$

We will investigate the impact of stochastic returns and demands on the threshold levels, which have converged in the multi-period context. Firstly, we investigate how the expected value of returned items affects the threshold levels. Secondly, we investigate the impact of demand variability on the threshold levels.

5.4.2.1 The impact of the expected value of returned items on the threshold levels

We will investigate the impact of the expected value of returned items in two groups on the threshold levels based on the following stochastic demands:

$$E[D_1]=200, StDev[D_1]=60; \quad E[D_2]=100, StDev[D_2]=30.$$

The impact of the expected value of returned items in group 1

Firstly, we investigate the impact of the expected value of returned items in group 1 based on the following scenarios in Table 5.6:

Table 5.6 The scenarios of returned items in group 1

($E[R_2]=45, StDev[R_2]=15$)

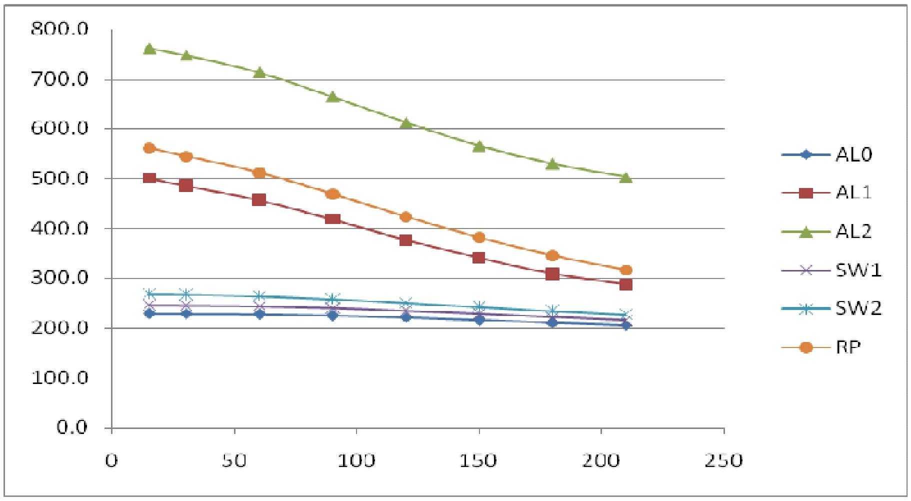
$E[R_1]$	15	30	60	90	120	150	180	210
$StDev[R_1]$	5	10	20	30	40	50	60	70

In the following, the threshold levels of the two products from the approximate dynamic programming model are shown in Table 5.7. Furthermore, the trend of the threshold levels is shown in Figure 5.4. The results have shown that all the threshold levels decrease with the expected value of returned items in group 1. For product 1, the threshold levels AL_1, AL_2 and RP are decreasing faster than its other threshold levels. On the other hand, for product 2, the threshold levels BL_1 and BL_2 are decreasing faster than BL_0. As more returned items are available for the recovery in each period, the threshold levels would be decreased.

As the interactive allocation of the returned items in two groups to the recovery, the expected value of returned items in group 1 would impact the threshold levels of the recovery processes using the returned items in two groups. In addition, the expected value of returned items in group 1 has less impact on the threshold levels related to production and switching. As production process never uses the returned items, it would not be impacted by the expected value of returned items. In addition, the two switching levels related to product 1, i.e. SW_1 and SW_2, are from the comparison of marginal profits of the recovery using returned items in group 1 and group 2 while the inventory level of product 2 is at the threshold level BL_0. Therefore, the expected value of returned items in group 1 has less impact on the two switching levels. However, there is remarkable impact on the threshold level RP, which is from the comparison of marginal profits of the recovery using returned items in group 2 while the inventory level of product 2 is at the threshold level BL_1.

Table 5.7 The threshold levels in different scenarios of returned items in group 1 with parameter set 1

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
$E[R_1]=15$	229.1	500.3	761.4	245.9	267.8	560.7	122.1	228.3	332.1
$E[R_1]=30$	228.6	485.6	747.6	245.3	266.9	544.4	121.5	224.3	328.4
$E[R_1]=60$	227.8	456.7	713.2	243.8	263.4	511.9	120.5	213.1	319.6
$E[R_1]=90$	225.3	418.2	664.4	240.6	257.7	469.0	119.5	203.9	310.5
$E[R_1]=120$	221.1	376.6	613.1	234.8	250.1	423.6	118.1	195.9	299.2
$E[R_1]=150$	216.1	340.8	566.1	228.8	242.1	381.9	116.2	183.6	280.6
$E[R_1]=180$	210.6	308.7	529.8	222.1	234.2	345.7	114.1	169.4	257.2
$E[R_1]=210$	205.4	287.7	503.9	216.0	227.2	317.1	112.3	154.9	235.8



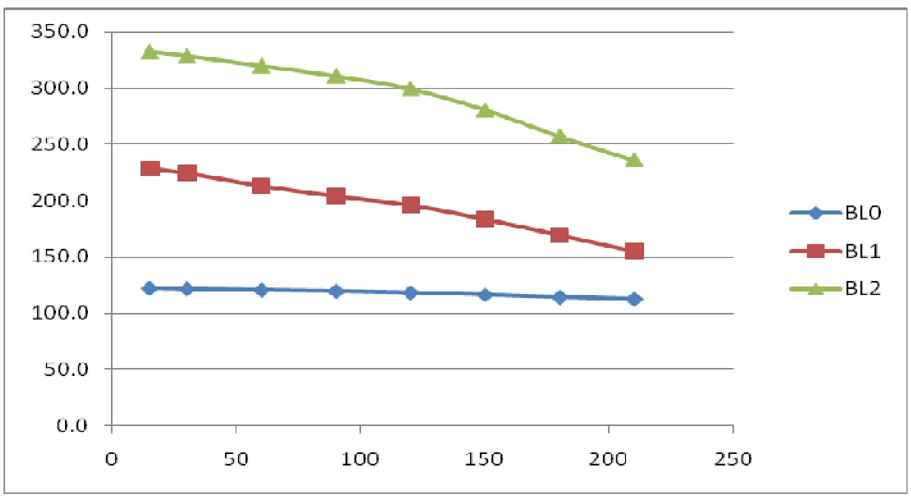


Figure 5.4 The trend of the threshold levels in different scenarios of returned items in group 1 with parameter set 1

Table 5.8 The threshold levels in different scenarios of returned items in group 1 with parameter set 2

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
$E[R_1]=15$	222.9	485.6	960.2	239.8	486.4	766.5	122.2	231.4	331.6
$E[R_1]=30$	222.8	476.6	950.2	239.6	476.0	758.5	121.8	223.5	325.4
$E[R_1]=60$	221.6	451.9	924.5	238.4	451.7	737.2	120.5	210.8	313.4
$E[R_1]=90$	219.3	417.1	883.5	235.1	417.1	700.0	118.9	203.8	305.1
$E[R_1]=120$	215.3	372.2	827.0	229.4	372.1	646.7	116.9	193.2	290.5
$E[R_1]=150$	210.4	337.8	773.9	223.9	337.6	596.0	114.9	182.2	273.5
$E[R_1]=180$	205.1	304.4	734.1	216.8	304.7	557.4	113.6	169.1	251.1
$E[R_1]=210$	200.0	283.2	705.8	211.3	282.5	531.2	112.3	154.6	233.2

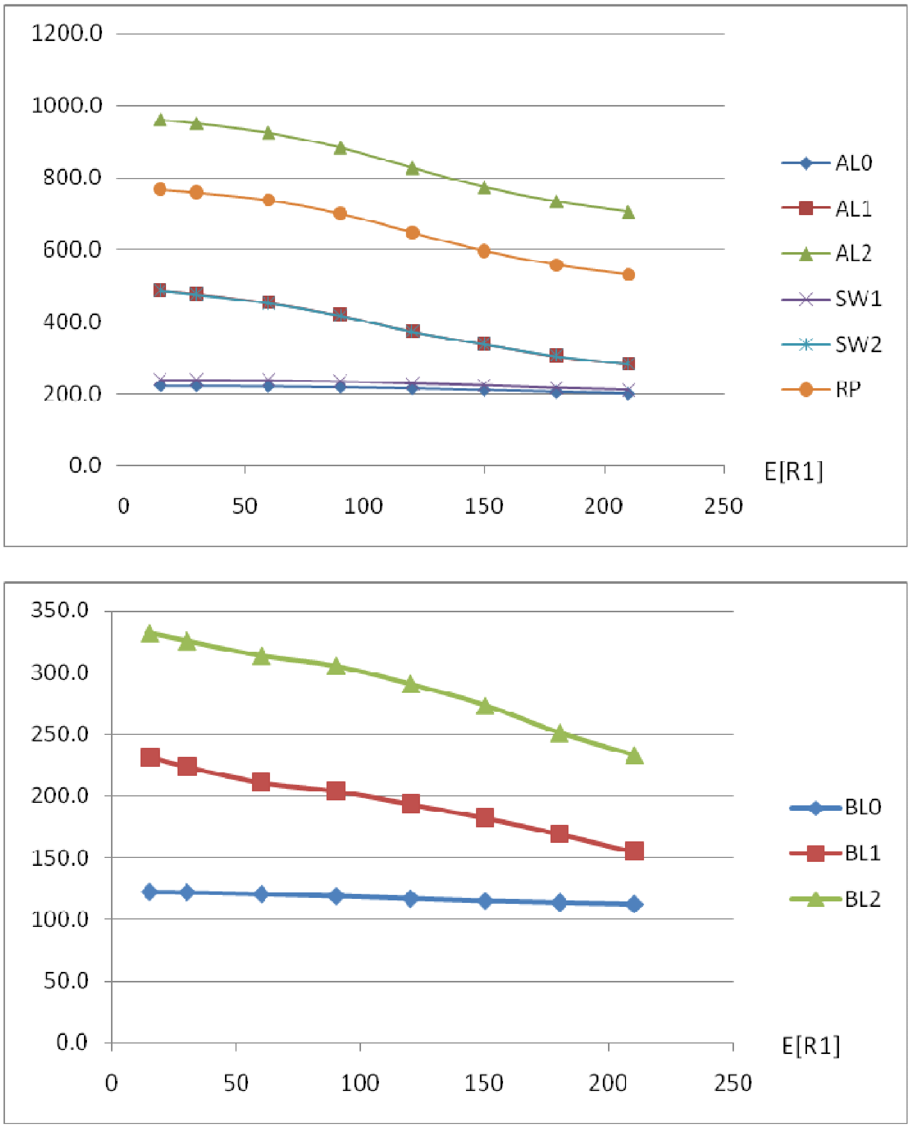


Figure 5.5 The trend of the threshold levels in different scenarios of returned items in group 1 with parameter set 2

Table 5.9 The threshold levels in different scenarios of returned items in group 1 with parameter set 3

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
$E[R_1]=15$	215.8	484.6	1158.7	233.9	686.4	965.7	122.4	230.6	332.8
$E[R_1]=30$	216.3	479.1	1148.7	233.1	680.3	959.4	121.7	222.2	326.2
$E[R_1]=60$	215.1	457.1	1124.1	231.7	665.5	941.4	120.0	207.5	312.5
$E[R_1]=90$	212.5	418.9	1091.6	228.7	642.2	915.8	117.8	204.4	297.3
$E[R_1]=120$	209.2	372.3	1030.6	223.6	592.2	858.5	115.7	193.7	283.4
$E[R_1]=150$	204.2	334.1	973.8	217.7	540.5	801.4	114.7	182.0	267.6
$E[R_1]=180$	198.5	298.5	935.1	211.3	502.2	760.9	113.7	169.6	250.6
$E[R_1]=210$	193.8	277.4	909.4	205.4	475.3	734.2	112.0	155.3	232.5

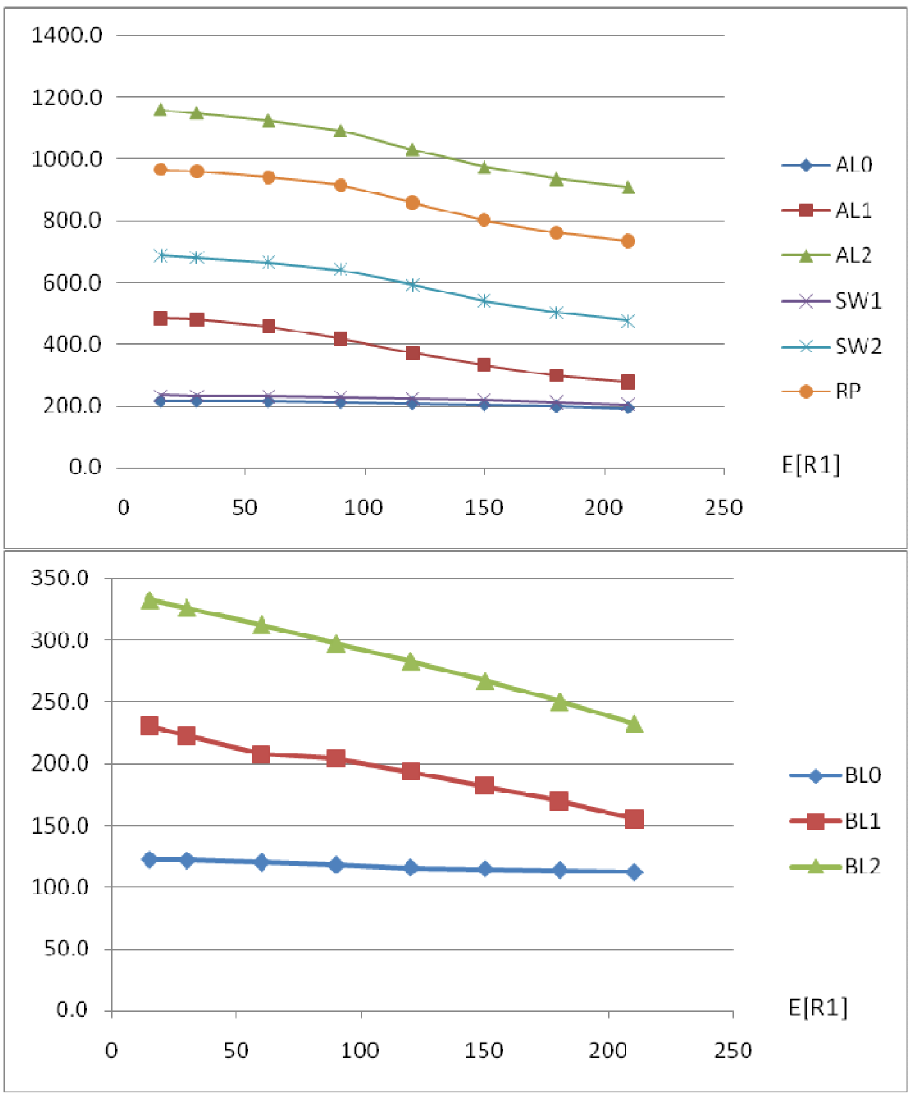


Figure 5.6 The trend of the threshold levels in different scenarios of returned items in group 1 with parameter set 3

The impact of the expected value of returned items in group 2

Secondly, we investigate the impact of the expected value of returned items in group 2 based on the following scenarios in Table 5.10:

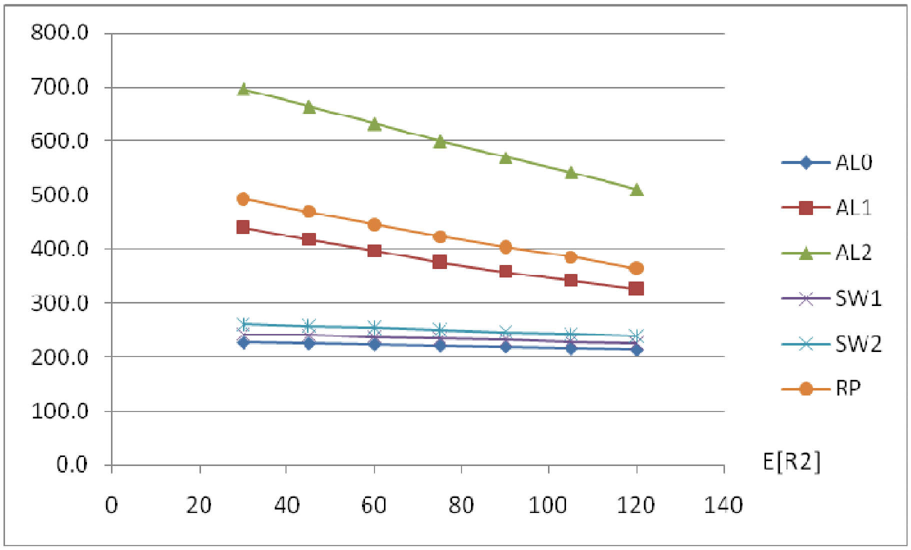
Table 5.10 The scenarios of returned items in group 2 ($E[R_1]=90, StDev[R_1]=30$)

$E[R_2]$	30	45	60	75	90	105	120
$StDev[R_2]$	10	15	20	25	30	35	40

In the following, the threshold levels of the two products from the approximate dynamic programming model are shown in Table 5.11. At the same time, the trend of the threshold levels is shown in Figure 5.7. From the figure, it can be found that all the threshold levels decrease with the expected value of returned items in group 2. For product 1, the threshold levels AL_1, AL_2 and RP are decreasing faster than its other threshold levels. On the other hand, for product 2, the threshold levels BL_1 and BL_2 are decreasing faster than BL_0. The explanation to the results is similar to that on the impact of the expected value of returned items in group 1.

Table 5.11 The threshold levels in different scenarios of returned items in group 2 with parameter set 1

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
$E[R_2]=30$	226.8	440.3	696.5	242.2	260.8	493.3	119.9	206.7	315.5
$E[R_2]=45$	225.3	418.2	664.4	240.6	257.7	469.0	119.5	203.9	310.5
$E[R_2]=60$	223.7	396.2	632.3	237.9	254.2	445.1	118.9	201.2	305.2
$E[R_2]=75$	221.4	376.0	600.4	235.0	250.5	422.9	118.1	198.6	298.5
$E[R_2]=90$	219.0	357.6	571.2	232.2	246.4	403.5	117.2	195.1	291.3
$E[R_2]=105$	216.6	342.0	543.0	228.9	242.6	385.4	116.4	191.1	283.4
$E[R_2]=120$	213.8	326.6	511.2	225.8	239.1	364.7	115.9	185.8	270.7



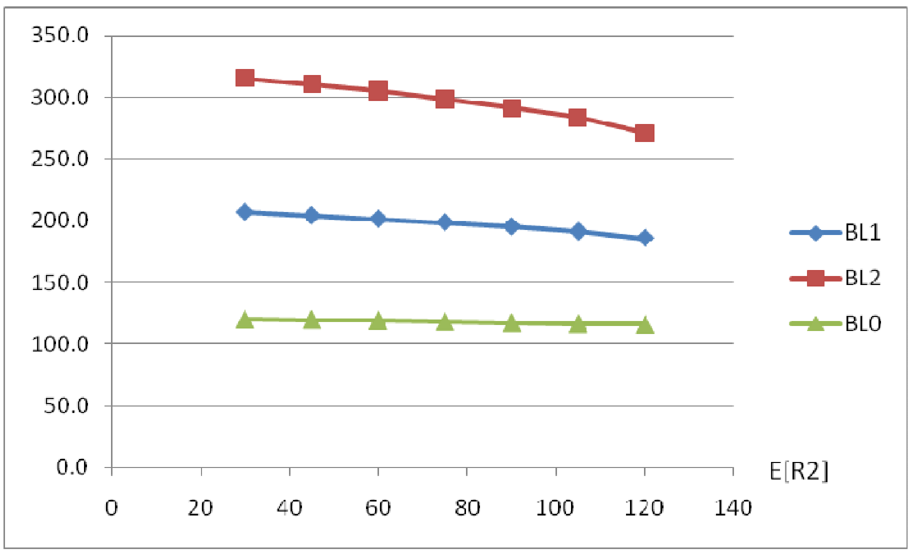
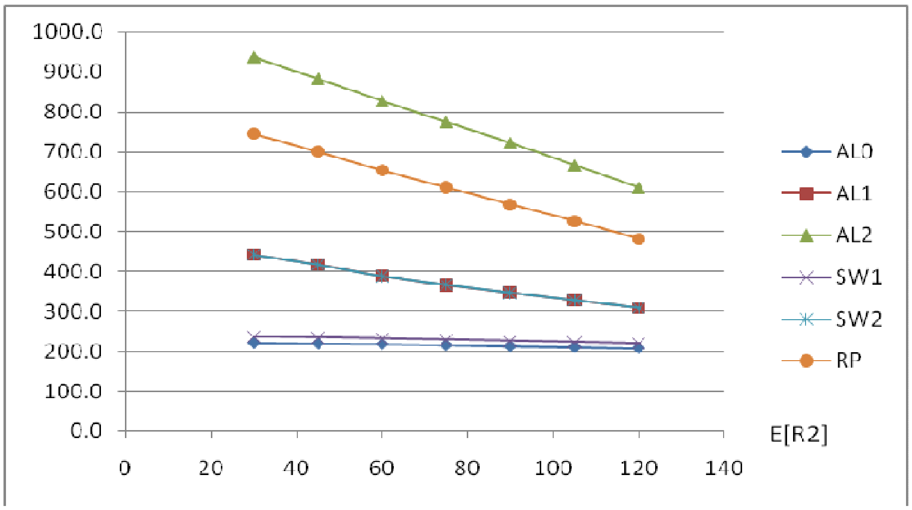


Figure 5.7 The trend of the threshold levels in different scenarios of returned items in group 2 with parameter set 1

Table 5.12 The threshold levels in different scenarios of returned items in group 2 with parameter set 2

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
$E[R_2]=30$	221.2	440.8	937.4	236.8	441.1	744.6	119.1	205.7	307.7
$E[R_2]=45$	219.3	417.1	883.5	235.1	417.1	700.0	118.9	203.8	305.1
$E[R_2]=60$	217.9	387.9	827.9	232.4	388.1	652.9	118.5	200.2	297.1
$E[R_2]=75$	215.5	365.3	775.5	229.6	366.3	610.6	117.4	198.5	291.1
$E[R_2]=90$	212.8	346.4	722.6	226.6	346.5	568.0	116.8	195.4	283.1
$E[R_2]=105$	210.5	327.1	666.9	223.4	327.7	526.5	116.1	191.8	274.5
$E[R_2]=120$	207.9	308.7	611.1	220.1	309.1	481.8	115.7	188.3	263.3



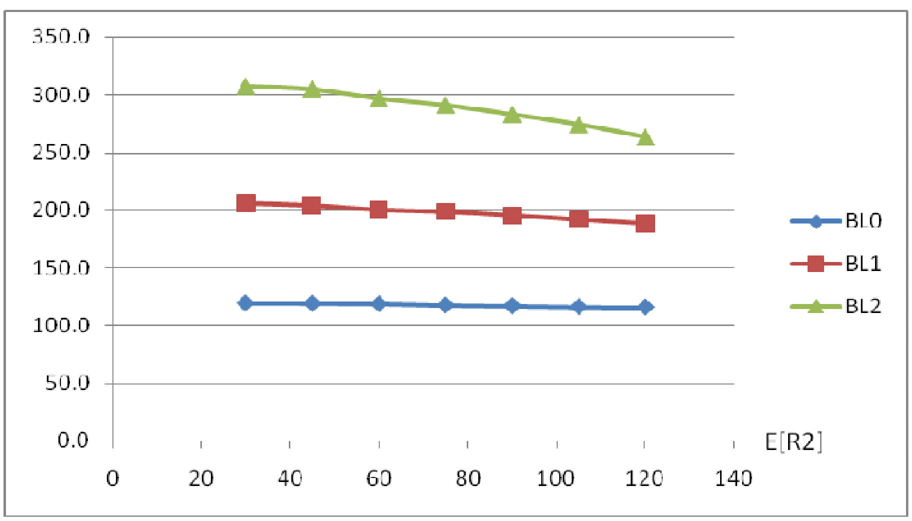


Figure 5.8 The trend of the threshold levels in different scenarios of returned items in group 2 with parameter set 2

Table 5.13 The threshold levels in different scenarios of returned items in group 2 with parameter set 3

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
$E[R_2]=30$	214.4	440.2	1164.9	230.6	681.9	979.9	118.1	205.5	299.7
$E[R_2]=45$	212.5	418.9	1091.6	228.7	642.2	915.8	117.8	204.4	297.3
$E[R_2]=60$	211.0	395.4	1014.0	226.5	602.1	850.2	117.4	202.2	292.6
$E[R_2]=75$	209.1	362.8	933.0	223.5	555.1	778.9	116.8	197.5	285.4
$E[R_2]=90$	206.3	340.6	854.2	220.2	512.4	709.6	116.3	195.2	277.3
$E[R_2]=105$	204.0	320.2	775.7	216.8	471.4	643.8	116.1	192.8	270.3
$E[R_2]=120$	201.0	300.1	700.0	213.6	425.9	577.2	115.7	188.6	261.8

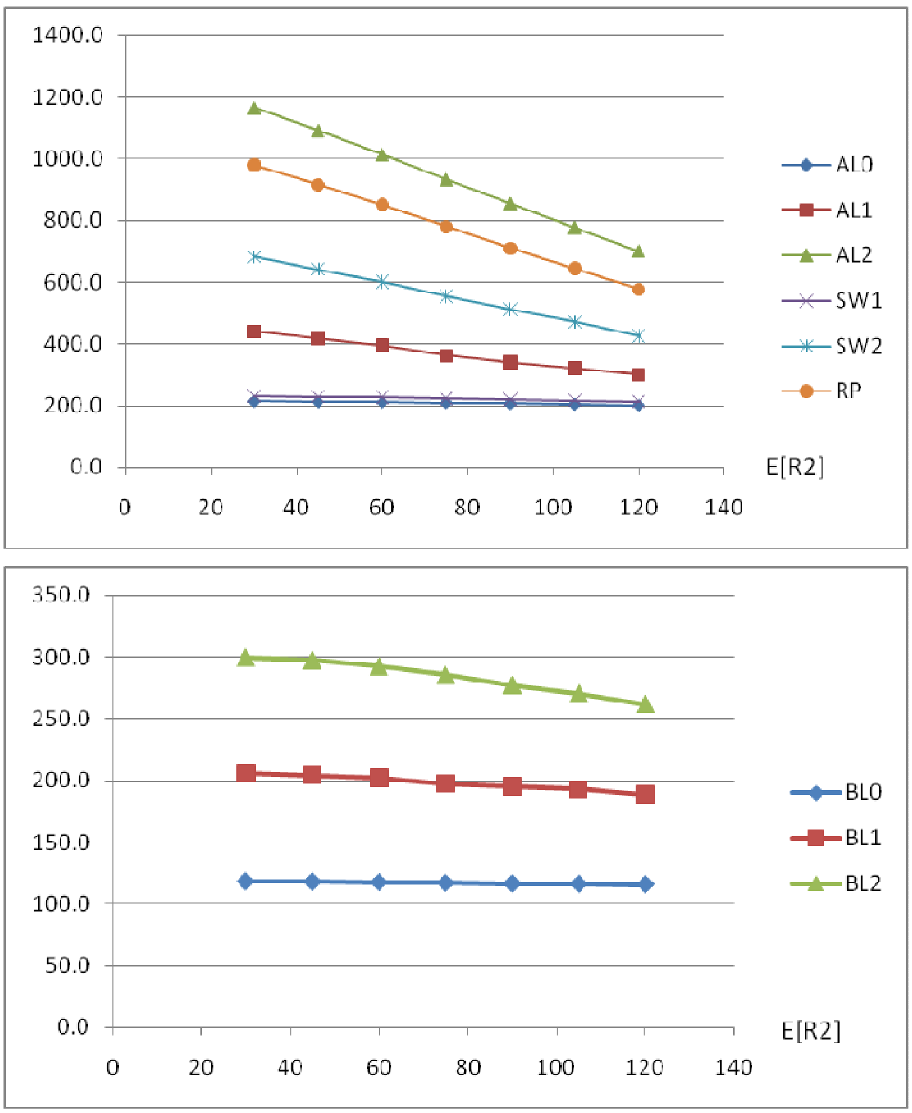


Figure 5.9 The trend of the threshold levels in different scenarios of returned items in group 2 with parameter set 3

5.4.2.2 The impact of demand variability of two products on the threshold levels

We will investigate the impact of demand variability of two products on the threshold levels with the following set of parameters on returned items:

$$E[R_1] = 90, StDev[R_1] = 30; \quad E[R_2] = 45, StDev[R_2] = 15.$$

The impact of demand variability of product 1

Firstly, the impact of demand variability of product 1 is investigated based on the following scenarios in Table 5.14.

Table 5.14 The scenarios of demand for product 1

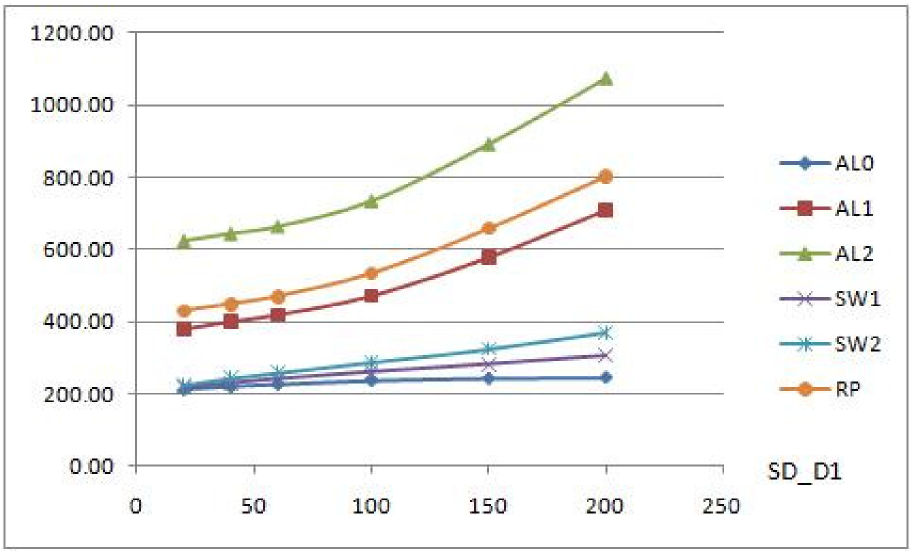
($E[D_1] = 200, E[D_2] = 100, StDev[D_2] = 30$)

***StDev*[*D*$_1$]**	20	40	60	100	150	200
***COV*$_1$**	0.1	0.2	0.3	0.5	0.75	1.0

In the following, the threshold levels of the two products from the approximate dynamic programming model are shown in Table 5.15. In addition, the trend of the threshold levels is shown in Figure 5.10. The results have shown that all the threshold levels related to product 1 increase with the demand variability of product 1 whereas the threshold levels related to product 2 seem unaffected. As the demands for the two products are independent of each other, the impact of demand variability of product 1 would only affect the threshold levels related to product 1. Furthermore, the higher demand variability results in the higher threshold levels to avoid possible stock shortage.

Table 5.15 The threshold levels in different scenarios of demand for product 1 with parameter set 1

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
COV_1=0.1	210.0	378.7	625.5	216.1	223.9	429.1	120.2	203.0	312.9
COV_1=0.2	218.6	399.6	644.8	229.6	242.8	447.7	120.0	205.6	312.2
COV_1=0.3	225.3	418.2	664.4	240.6	257.7	469.0	119.5	203.9	310.5
COV_1=0.5	236.7	470.4	734.9	260.1	287.2	532.6	118.9	202.4	310.2
COV_1=0.75	243.6	577.3	892.2	281.9	324.0	658.0	118.8	206.8	314.6
COV_1=1.0	246.2	708.4	1073.8	305.2	370.5	803.7	118.8	207.0	317.1



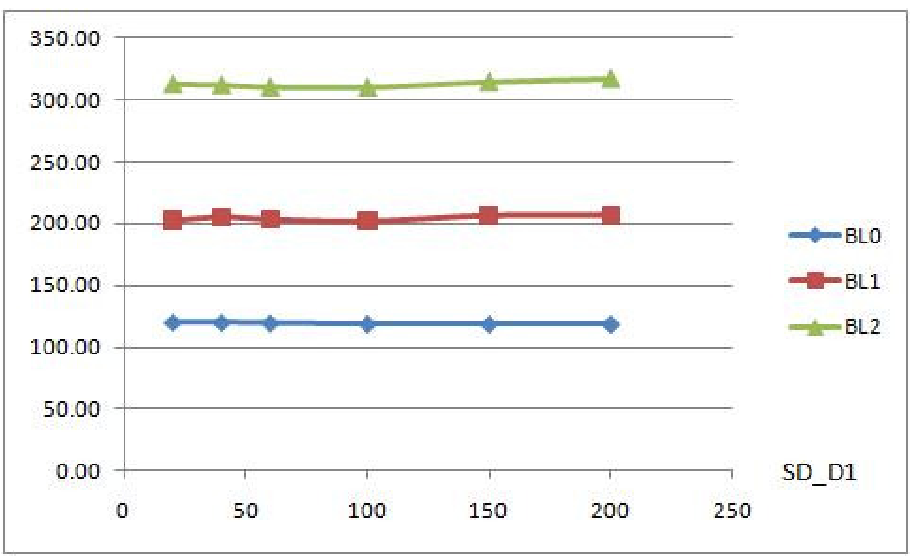
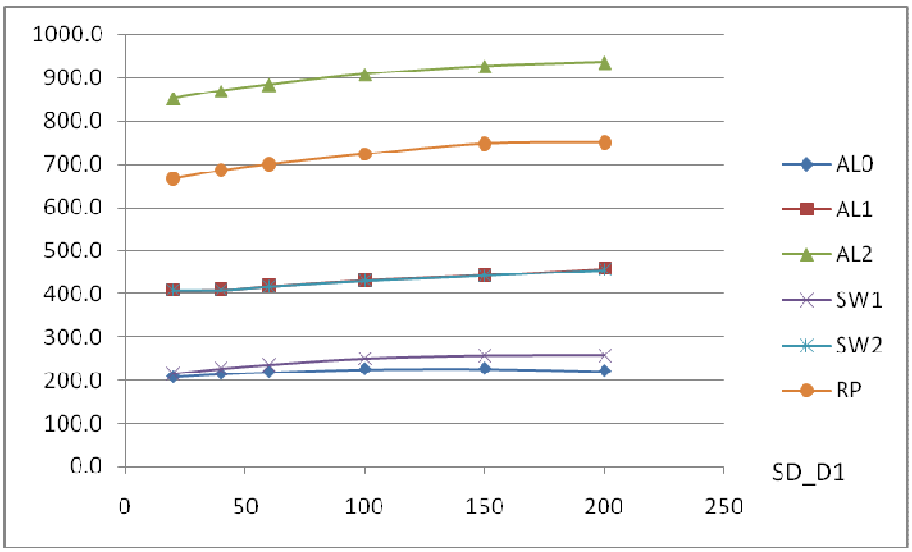


Figure 5.10 The trend of the threshold levels in different scenarios of demand for product 1 with parameter set 1

Table 5.16 The threshold levels in different scenarios of demand for product 1 with parameter set 2

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
COV_1=0.1	208.1	408.0	851.5	214.6	408.3	666.7	119.4	207.3	301.9
COV_1=0.2	214.8	409.0	869.7	226.2	408.9	686.0	119.4	205.2	307.2
COV_1=0.3	219.3	417.1	883.5	235.1	417.1	700.0	118.9	203.8	305.1
COV_1=0.5	225.9	430.7	907.9	249.2	430.3	723.0	118.7	202.1	300.6
COV_1=0.75	227.1	442.4	926.3	256.1	442.5	746.9	119.3	201.0	302.3
COV_1=1.0	221.6	457.0	935.6	257.0	455.9	749.6	120.6	206.4	305.8



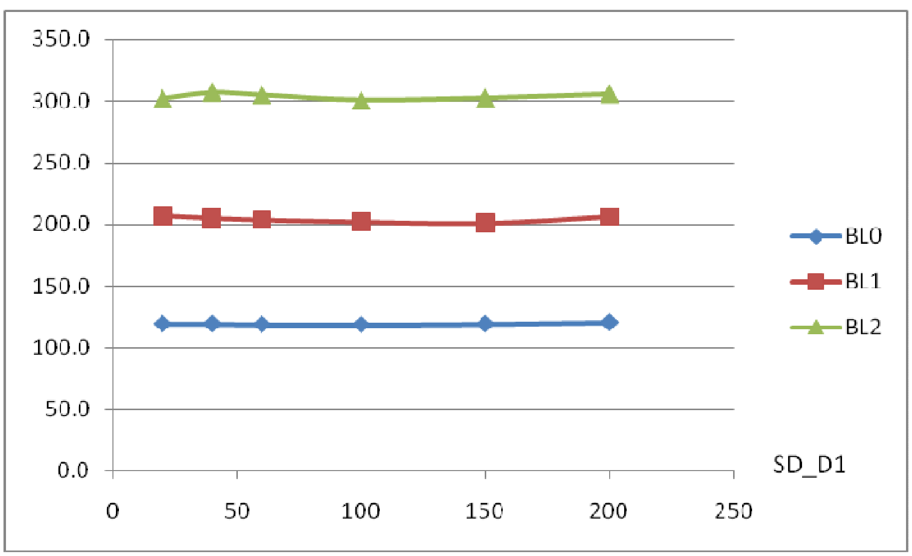


Figure 5.11 The trend of the threshold levels in different scenarios of demand for product 1 with parameter set 2

Table 5.17 The threshold levels in different scenarios of demand for product 1 with parameter set 3

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
COV_1=0.1	205.7	409.5	1060.1	212.3	618.4	881.8	116.8	207.3	293.3
COV_1=0.2	210.4	409.3	1079.9	221.9	630.6	902.3	117.8	204.5	297.2
COV_1=0.3	212.5	418.9	1091.6	228.7	642.2	915.8	117.8	204.4	297.3
COV_1=0.5	215.1	418.4	1089.2	237.9	649.5	918.7	118.3	201.7	296.5
COV_1=0.75	209.1	420.0	1086.0	238.5	654.7	921.6	119.2	200.7	301.8
COV_1=1.0	198.2	424.4	1081.9	231.7	631.5	918.2	120.8	203.7	304.4

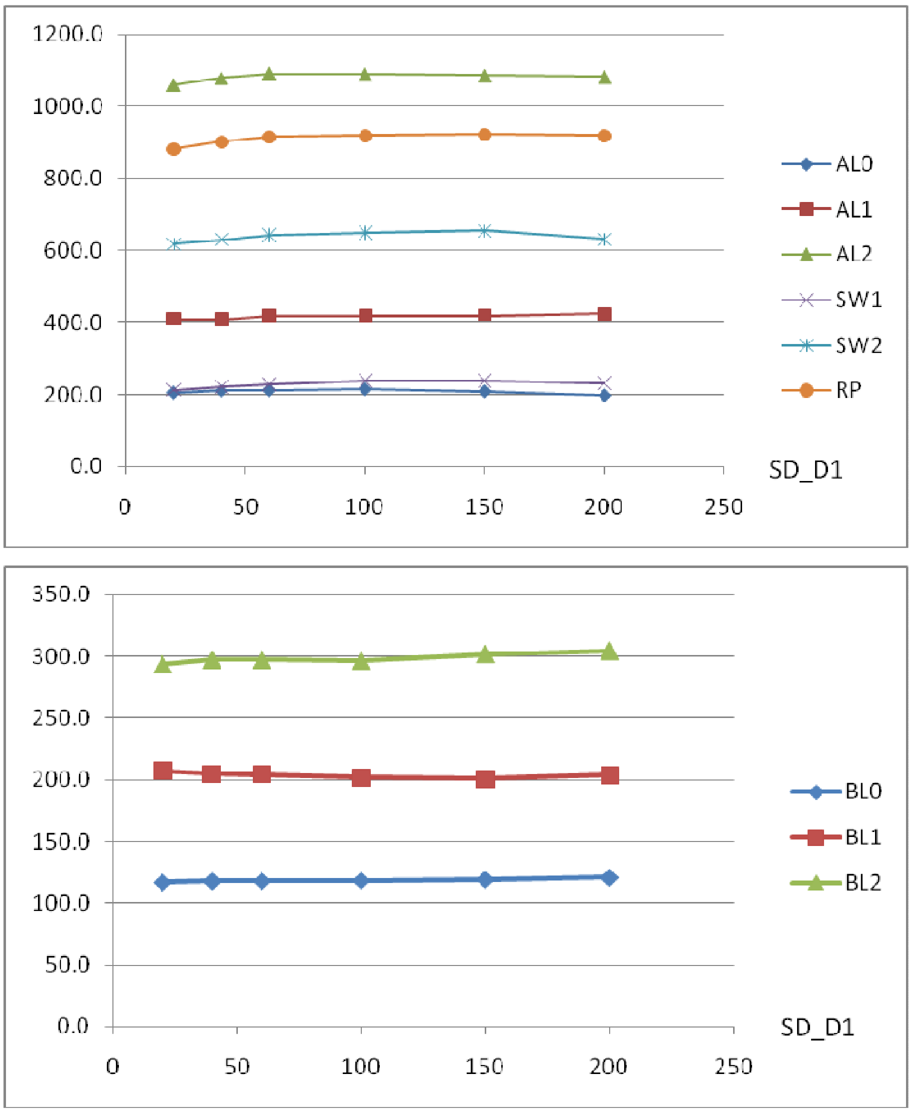


Figure 5.12 The trend of the threshold levels in different scenarios of demand for product 1 with parameter set 3

The impact of demand variability of product 2

Secondly, the impact of demand variability of product 2 is investigated based on the following scenarios listed in Table 5.18:

Table 5.18 The scenarios of demand for product 2

($E[D_2]=100, E[D_1]=200, StDev[D_1]=60$)

***StDev*[*D*₂]**	10	20	30	50	75	100
***COV*₂**	0.1	0.2	0.3	0.5	0.75	1.0

In the following, the threshold levels of the two products from the approximate dynamic programming model are shown in Table 5.19. In addition, the trend of the threshold levels is shown in Figure 5.13. The results have shown that all the threshold levels related to product 2 increase with the demand variability of product 2 whereas the threshold levels related to product 1 seem unaffected. The explanation to the results is similar to that on the impact of demand variability of product 1.

Table 5.19 The threshold levels in different scenarios of demand for product 2 with parameter set 1

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
COV_2=0.1	225.8	425.3	659.2	240.5	258.4	476.4	107.3	187.3	287.1
COV_2=0.2	225.6	422.7	662.9	240.7	258.1	473.2	113.9	194.3	297.8
COV_2=0.3	225.3	418.2	664.4	240.6	257.7	469.0	119.5	203.9	310.5
COV_2=0.5	225.5	419.3	667.4	240.4	257.4	470.9	130.3	229.0	339.2
COV_2=0.75	225.3	425.4	678.0	240.5	257.6	477.8	141.6	274.0	401.1
COV_2=1.0	225.3	432.2	693.4	240.6	258.3	486.6	149.4	339.1	498.8

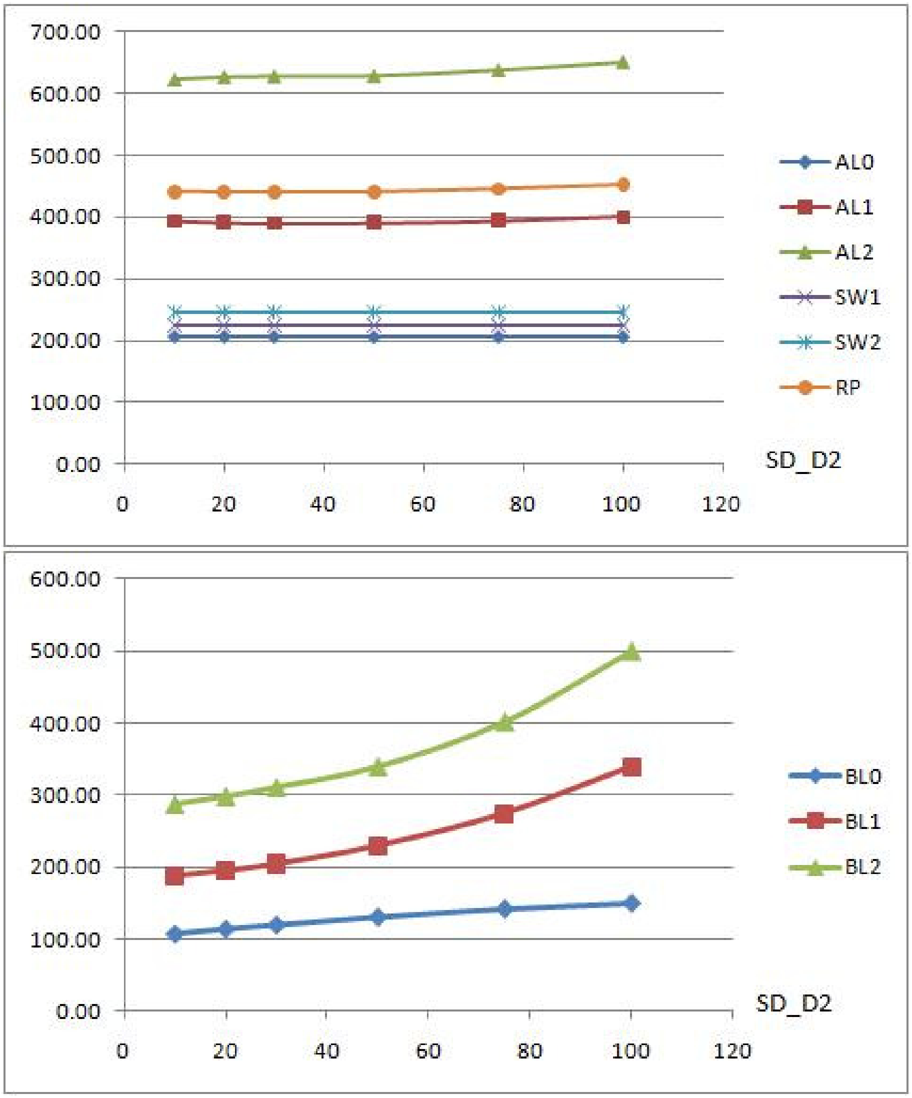


Figure 5.13 The trend of the threshold levels in different scenarios of demand for product 2 with parameter set 1

Table 5.20 The threshold levels in different scenarios of demand for product 2 with parameter set 2

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
COV_2=0.1	220.4	438.8	883.7	236.0	439.2	706.3	106.8	181.4	283.5
COV_2=0.2	219.8	424.2	887.2	235.4	425.7	704.0	113.4	191.9	290.8
COV_2=0.3	219.3	417.1	883.5	235.1	417.1	700.0	118.9	203.8	305.1
COV_2=0.5	219.6	413.1	879.4	234.6	412.3	694.1	130.0	223.6	322.9
COV_2=0.75	219.7	419.5	878.5	235.7	419.2	694.4	141.9	243.4	344.2
COV_2=1.0	219.8	424.3	908.3	235.5	424.0	718.1	148.6	340.1	491.5

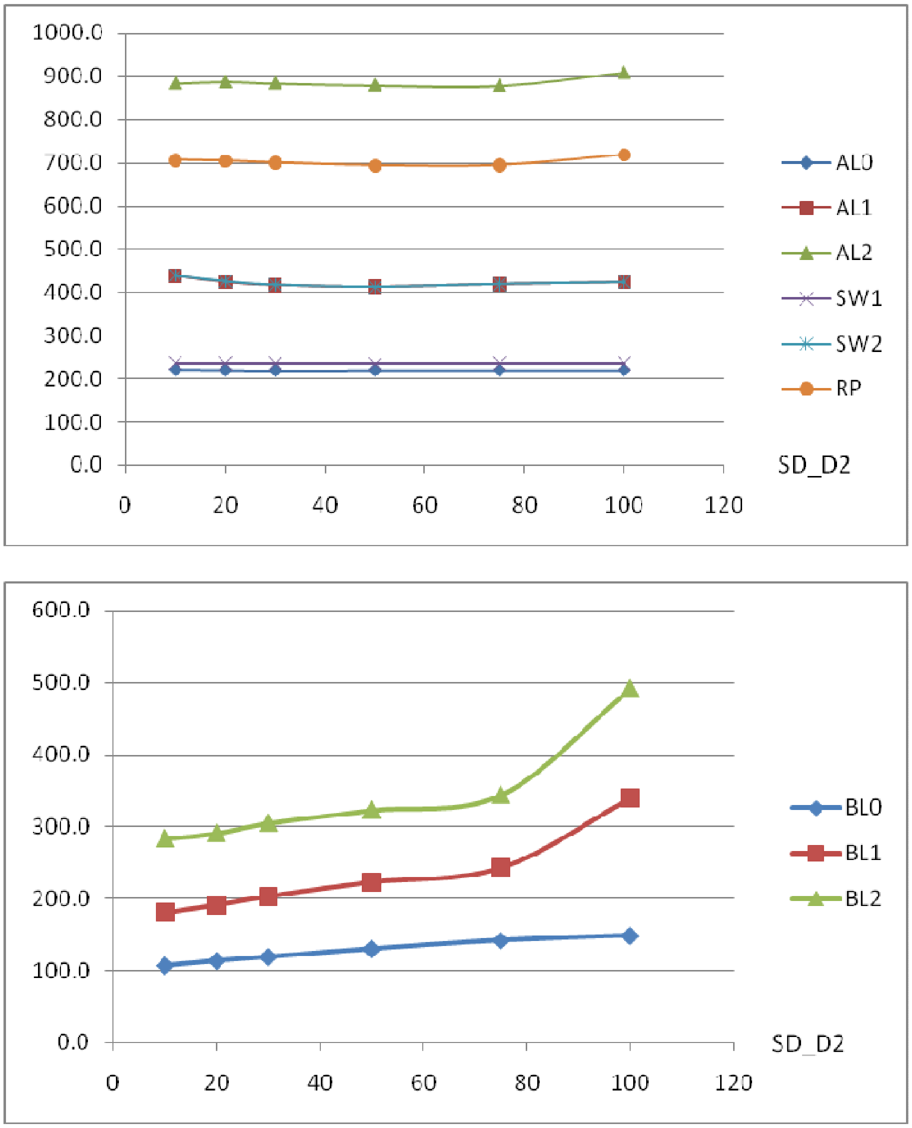


Figure 5.14 The trend of the threshold levels in different scenarios of demand for product 2 with parameter set 2

Table 5.21 The threshold levels in different scenarios of demand for product 2 with parameter set 3

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
COV_2=0.1	212.9	408.4	1096.1	228.2	648.3	921.0	106.3	179.5	264.0
COV_2=0.2	212.6	417.4	1094.6	229.0	646.5	918.6	112.4	192.4	280.2
COV_2=0.3	212.5	418.9	1091.6	228.7	642.2	915.8	117.8	204.4	297.3
COV_2=0.5	213.1	417.0	1084.2	228.3	634.4	907.9	128.9	221.8	319.6
COV_2=0.75	213.4	423.4	1086.9	229.3	633.7	907.4	140.6	242.8	340.9
COV_2=1.0	214.2	440.6	1114.8	230.6	649.6	929.3	149.3	256.2	355.9

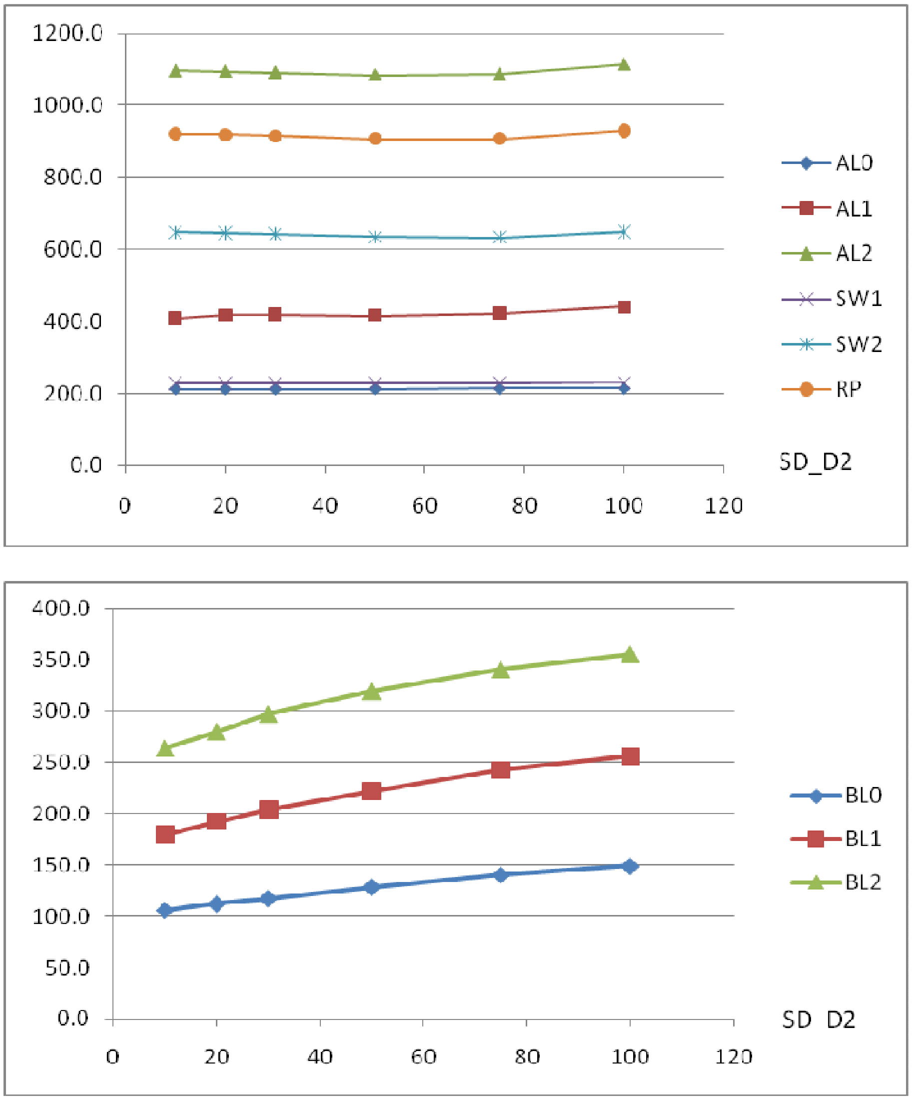


Figure 5.15 The trend of the threshold levels in different scenarios of demand for product 2 with parameter set 3

5.4.3 The comparison of three heuristic policies with respect to the expected average profit

While using the threshold levels to help to make production and recovery decisions in a relatively long horizon, the resulting expected average profit is compared with those values obtained by using two heuristic policies from the single-period problem. The following symbols will be used in the presentation of numerical results:

H1 – Heuristic policy from the single-period problem disregarding scrap values of the remaining finished products;

H2 – Heuristic policy from the single-period problem assuming scrap value of the remaining product 1 and product 2 to be equal to c_{R21} and c_{R22} respectively;

H3 – Heuristic policy from solving the ADP model;

EAP_H1 – Expected average profit calculated while the heuristic policy ***H1*** is used in a relatively long horizon;

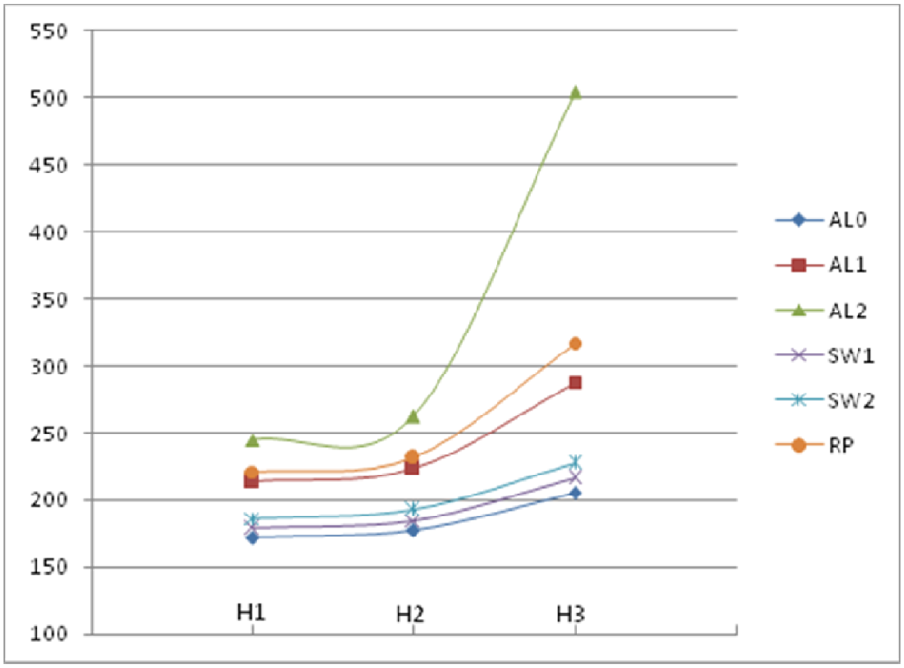
EAP_H2 – Expected average profit calculated while the heuristic policy ***H2*** is used in a relatively long horizon;

EAP_H3 – Expected average profit calculated while the heuristic policy ***H3*** is used in a relatively long horizon.

The optimal policy for the single-period problem is used as heuristic policy for the multi-period problem. Two heuristic policies, denoted as *H1* and *H2* respectively, are derived from solving the single-period problem. The policy *H1* disregards the scrap values of the remaining finished products whereas the policy *H2* assumes the scrap value of product 1 and product 2 to be equal to c_{R21} and c_{R22} respectively. In order to compare the threshold levels of the policies *H1* and *H2* with those of the policy *H3* from solving the approximate dynamic programming model, we have selected the set of threshold levels when $E[R_1] = 210$ in Table 5.7. The threshold levels of the three heuristic policies have been shown in Table 5.22. In Figure 5.16, the threshold levels have been compared between the three heuristic policies. It is found that the corresponding threshold levels of the policy *H3* are highest, whereas those threshold levels of the policy *H1* are lowest. The difference of each corresponding threshold level between the policies *H1* and *H2* is small whereas the difference between the policies *H1* and *H3* is obviously large.

Table 5.22 The threshold levels in three heuristic policies

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
H1	171.6	213.8	244.9	179.1	186.2	220.9	90.8	103.9	112.0
H2	176.9	223.1	262.2	184.8	192.5	231.5	100.0	118.1	132.9
H3	205.4	287.7	503.9	216.0	227.2	317.1	112.3	154.9	235.8



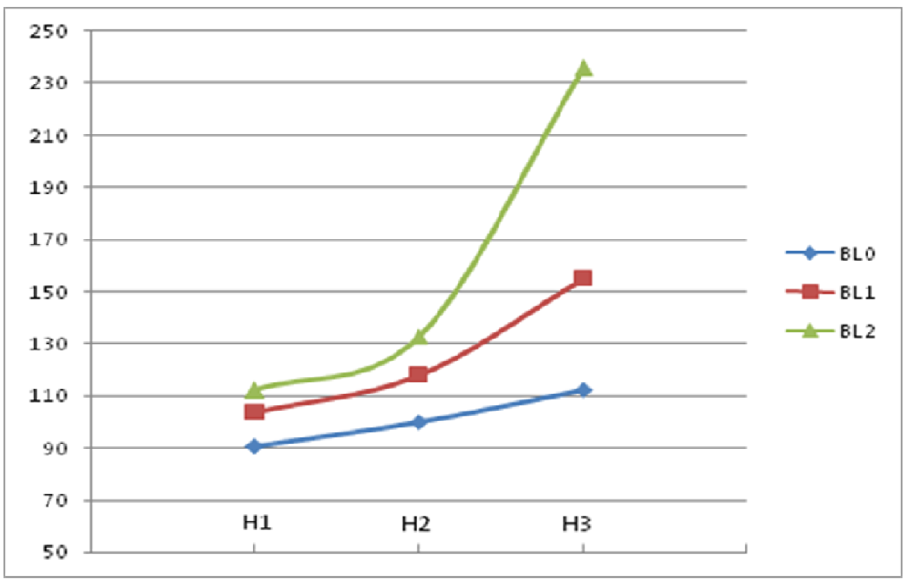


Figure 5.16 The comparison of the threshold levels in different heuristic policies

Using the above three heuristic policies to make production and recovery decisions of the two-product recovery system in a relatively long horizon, the resulting expected average profits are shown in Table 5.23. It can be found that the heuristic policy *H3* performs best, secondly the policy *H2* and finally the policy *H1*. Comparing the expected average profits between the policies *H1* and *H3*, 7.2% increment can be achieved while using the policy *H3* to replace the policy *H1*.

Table 5.23 The expected average profit using different heuristic policies

EAP_H1	2115.3
EAP_H2	2186.2
EAP_H3	2267.4

A bound could be obtained by relaxing some assumptions. However, it might be too loose and become meaningless to be compared with the performance of the results by using our approach. Hence, the multi-level threshold policy by solving the ADP model is compared with the other two heuristic policies, which are derived from the optimal policy for the single-period problem. One of the two heuristic policies assumes scrap value of the remaining finished items to be a nonzero fixed value whereas another assumes scrap value to be zero. By this comparison, the threshold policy by solving the ADP model is found to have the best performance under a wide range of settings. Therefore, to some extent, we have proved that our approach is promising to solve the multi-period problem although the optimal solution is difficult to obtain.

5.5 Summary

In this Chapter, we have developed the ADP model of the two-product recovery system in the situation of lost sale over a finite horizon. The model aims to determine the threshold levels as the multi-level threshold policy from the single-period problem is assumed to be used for the multi-period problem. In the multi-period situation, the threshold levels are found to be only dependent on the gradient of the cost-to-go function at the points of interest.

For a given set of system parameters, we find that the threshold levels of any certain period would converge with the distance of this period from the last period of the planning horizon. In addition, the converging speed of the threshold levels is impacted by the inventory holding cost rate. The higher the inventory holding cost rate, the faster the threshold levels converge. The converging threshold levels are used in the optimal policy, which helps to make production and recovery decisions in the multi-period context. The impact of system parameters on the threshold levels has been investigated. The numerical results have shown that the more returned items from either group in each period would make the threshold levels lower. Among them, the threshold levels AL_1, AL_2, BL_1, BL_2 and RP, related to recovery processes, would obviously decrease with returned items increasing. However, there are small decreases on the threshold levels AL_0 and BL_0, related to production processes, and the threshold levels SW_1 and SW_2, related to switching the allocation of returned items to the recovery processes between the two products. On the other hand, with the increasing demand variability of a certain product, the threshold levels related to this product would increase at the same time whereas the threshold levels related to the other product seem unaffected.

After determining the threshold levels in the multi-period situation, we can use the threshold policy to control the two-product recovery system. The performance of this policy is compared with the two heuristic policies derived from the optimal policy of the single-period problem. Through the comparison of the resulting expected average profit, the policy from solving the ADP model outperforms the other two heuristic policies.

Chapter 6 The study on two-product recovery system in a finite horizon with backorder and zero lead time

Chapter 6 focuses on the two-product recovery system in a finite horizon, in which backorder is allowed. In the recovery system, production and recovery processes are assumed to have zero lead time. Thus, the inventory of the two products can be instantly replenished by production and recovery processes. Section 6.1 introduces the two-product recovery system. In Section 6.2, an ADP model of the recovery system is developed in order to minimize the expected total cost in the finite horizon. The model is used to derive the threshold levels, which are only dependent on the gradient of the cost-to-go function at the points of interest. Section 6.3 provides the details about how to determine the gradient at the points of interest. Section 6.4 gives numerical analysis on the recovery system with respect to the effect of system parameters and provides the comparison with two other heuristic policies. Finally, Section 6.5 summarizes the main findings.

6.1 Introduction

The two-product recovery system in a finite horizon has been introduced in Chapter 3. Furthermore, Chapter 5 focuses on the study of the recovery system dealing with shortages as lost sales. However, Chapter 6 will study the two-product recovery, in which backorder is allowed. Thus, the objective is to minimize the expected total cost of the two-product recovery system in a finite horizon. In order to fulfill the aim, we need to find the optimal policy, which helps to make production and recovery decisions in each period of the finite horizon.

6.2 Approximate dynamic programming model of the two-product recovery system in the multi-period context

The dynamic programming model of the two-product recovery system in the multi-period context has been introduced in Chapter 3. The model aims to maximize the expected total profit in the finite horizon. In this Chapter, the recovery system is assumed to allow unsatisfied demands to be backordered in future periods. Therefore, the dynamic programming model in this Chapter aims to minimize the expected total cost in the finite horizon. The related assumptions and notations are referred to as in Chapter 3. Besides, some related notations are listed as follows (*i, j* = 1, 2):

$x_{Sj}^{(t)}$	initial inventory position of product *j* in period *t*;
$x_j^{(t)}$	inventory position of product *j* after production and recovery in period *t*;
$f_t(x_{S1}^{(t)}, x_{S2}^{(t)})$	expected minimum of the expected total cost from period *t* till final period;
EC_t	expected cost in period *t*;
MEC	minimum expected cost in final period;
ETC_t	expected total cost from period *t* till final period;
$\tilde{ETC}_t$	approximation to ETC_t;
$ATC_k^{(t)}$	actual cost in period *t* for sample *k* of demands and returns.

The transition relationship on initial inventory position of each product between two subsequent periods can be expressed as follows (j = 1, 2):

$$x_{Sj}^{(t+1)} = x_{Sj}^{(t)} + p_j^{(t)} + r_{1j}^{(t)} + r_{2j}^{(t)} - D_j^{(t)}. \tag{6.1}$$

In addition, inventory position of each product after replenishment at period t is dependent on the initial inventory, which can be expressed as follows:

$$x_j^{(t)} = x_{Sj}^{(t)} + p_j^{(t)} + r_{1j}^{(t)} + r_{2j}^{(t)}. \tag{6.2}$$

According to Formula (3.2) in Chapter 3, the expected cost at period t can be calculated as follows:

$$\begin{aligned}
&EC_t(x_{S1}^{(t)}, x_{S2}^{(t)}, p_1^{(t)}, p_2^{(t)}, r_{11}^{(t)}, r_{12}^{(t)}, r_{21}^{(t)}, r_{22}^{(t)}) \\
&= c_{P1}p_1^{(t)} + c_{P2}p_2^{(t)} + c_{R11}r_{11}^{(t)} + c_{R12}r_{12}^{(t)} + c_{R21}r_{21}^{(t)} + c_{R22}r_{22}^{(t)} \\
&\quad + h_1\int_0^{x_{S1}^{(t)}+p_1^{(t)}+r_{11}^{(t)}+r_{21}^{(t)}} (x_{S1}^{(t)} + p_1^{(t)} + r_{11}^{(t)} + r_{21}^{(t)} - D_1^{(t)}) f(D_1^{(t)}, \mu_1, \sigma_1) dD_1^{(t)} \\
&\quad + v_1\int_{x_{S1}^{(t)}+p_1^{(t)}+r_{11}^{(t)}+r_{21}^{(t)}}^{\infty} (D_1^{(t)} - x_{S1}^{(t)} - p_1^{(t)} - r_{11}^{(t)} - r_{21}^{(t)}) f(D_1^{(t)}, \mu_1, \sigma_1) dD_1^{(t)} \\
&\quad + h_2\int_0^{x_{S2}^{(t)}+p_2^{(t)}+r_{12}^{(t)}+r_{22}^{(t)}} (x_{S2}^{(t)} + p_2^{(t)} + r_{12}^{(t)} + r_{22}^{(t)} - D_2^{(t)}) f(D_2^{(t)}, \mu_2, \sigma_2) dD_2^{(t)} \\
&\quad + v_2\int_{x_{S2}^{(t)}+p_2^{(t)}+r_{12}^{(t)}+r_{22}^{(t)}}^{\infty} (D_2^{(t)} - x_{S2}^{(t)} - p_2^{(t)} - r_{12}^{(t)} - r_{22}^{(t)}) f(D_2^{(t)}, \mu_2, \sigma_2) dD_2^{(t)}.
\end{aligned} \tag{6.3}$$

As $f_t(x_{S1}^{(t)}, x_{S2}^{(t)})$ denotes the expected total cost from period t till final period in Chapter 6, we assume $f_{M+1}(x_{S1}^{(M+1)}, x_{S2}^{(M+1)}) = \sum_{i=1}^{2} c_{Pi}[-x_{Si}^{(M+1)}]^+$. The assumption means that normal production would be used to meet the backordered demands, which are not satisfied at final period. The objective of the dynamic programming model is to minimize the expected total cost in the finite horizon. The Bellman's equation of dynamic programming is as follows ($t = 1, 2, ..., M$):

$$f_t(x_{S1}^{(t)}, x_{S2}^{(t)}) = \underset{R_1^{(t)}, R_2^{(t)}}{E}\{ \min_{p_1^{(t)}, p_2^{(t)}, r_{11}^{(t)}, r_{12}^{(t)}, r_{21}^{(t)}, r_{22}^{(t)}} \{EC_t(x_{S1}^{(t)}, x_{S2}^{(t)}, p_1^{(t)}, p_2^{(t)}, r_{11}^{(t)}, r_{12}^{(t)}, r_{21}^{(t)}, r_{22}^{(t)}) + \underset{D_1^{(t)}, D_2^{(t)}}{E}[f_{t+1}(x_{S1}^{(t+1)}, x_{S2}^{(t+1)})]\}\}. \tag{6.4}$$

Under the above-mentioned assumption about the boundary value of Bellman's equation, the objective function of the single-period problem on the final period can be expressed as follows:

$$\begin{aligned}
&EC_M(x_{S1}^{(M)}, x_{S2}^{(M)}, p_1^{(M)}, p_2^{(M)}, r_{11}^{(M)}, r_{12}^{(M)}, r_{21}^{(M)}, r_{22}^{(M)}) \\
&= c_{P1}p_1^{(M)} + c_{P2}p_2^{(M)} + c_{R11}r_{11}^{(M)} + c_{R12}r_{12}^{(M)} + c_{R21}r_{21}^{(M)} + c_{R22}r_{22}^{(M)} \\
&\quad + h_1\int_0^{x_{S1}^{(M)}+p_1^{(M)}+r_{11}^{(M)}+r_{21}^{(M)}} (x_{S1}^{(M)} + p_1^{(M)} + r_{11}^{(M)} + r_{21}^{(M)} - D_1^{(M)}) f(D_1^{(M)}, \mu_1, \sigma_1) dD_1^{(M)} \\
&\quad + (v_1 + c_{P1})\int_{x_{S1}^{(M)}+p_1^{(M)}+r_{11}^{(M)}+r_{21}^{(M)}}^{\infty} (D_1^{(M)} - x_{S1}^{(M)} - p_1^{(M)} - r_{11}^{(M)} - r_{21}^{(M)}) f(D_1^{(M)}, \mu_1, \sigma_1) dD_1^{(M)} \\
&\quad + h_2\int_0^{x_{S2}^{(M)}+p_2^{(M)}+r_{12}^{(M)}+r_{22}^{(M)}} (x_{S2}^{(M)} + p_2^{(M)} + r_{12}^{(M)} + r_{22}^{(M)} - D_2^{(M)}) f(D_2^{(M)}, \mu_2, \sigma_2) dD_2^{(M)} \\
&\quad + (v_2 + c_{P2})\int_{x_{S2}^{(M)}+p_2^{(M)}+r_{12}^{(M)}+r_{22}^{(M)}}^{\infty} (D_2^{(M)} - x_{S2}^{(M)} - p_2^{(M)} - r_{12}^{(M)} - r_{22}^{(M)}) f(D_2^{(M)}, \mu_2, \sigma_2) dD_2^{(M)}.
\end{aligned} \tag{6.5}$$

As the concave property of the objective function of the single-period problem has been proved in Chapter 3, we can obtain the optimal solution to minimize the objective function in Formula (6.5) by solving KKT conditions. The solution structure can be referred to in Appendix B. The formulae of determining the related threshold levels of the optimal policy for the single-period problem have been listed in Table 6.1.

The objective of studying the dynamic programming model is to obtain the optimal policy for the two-product recovery system in a finite horizon. By similar approximation mentioned in Chapter 5, the cost-to-go function of dynamic programming at the points of interest can be represented by using the gradients as follows:

$$f_{t+1}(x_{S1}^{(t+1)}, x_{S2}^{(t+1)}) \approx u_1^{(t)} x_1^{(t)} + u_2^{(t)} x_2^{(t)}. \tag{6.6}$$

In the above formula, $u_j^{(t)}$ (j = 1, 2), which denotes the first-order derivative of the cost-to-go function with respect to inventory level of product j after replenishment, is expressed as follows:

$$u_j^{(t)} = \frac{\partial}{\partial x_j^{(t)}} \{ \underset{D_1^{(t)}, D_2^{(t)}}{E} [f_{t+1}(x_{S1}^{(t+1)}, x_{S2}^{(t+1)})] \}. \tag{6.7}$$

Thus, the objective function of the dynamic programming model, denoted as ETC_t, can be approximated as follows:

$$\begin{aligned} &ETC_t \\ &= EC_t + \underset{D_1^{(t)}, D_2^{(t)}}{E} [f_{t+1}(x_{S1}^{(t+1)}, x_{S2}^{(t+1)})] \\ &\approx EC_t + u_1^{(t)} x_1^{(t)} + u_2^{(t)} x_2^{(t)} \\ &\approx EC_t + u_1^{(t)} (x_{S1}^{(t)} + p_1^{(t)} + r_{11}^{(t)} + r_{21}^{(t)}) + u_2^{(t)} (x_{S2}^{(t)} + p_2^{(t)} + r_{12}^{(t)} + r_{22}^{(t)}). \end{aligned} \tag{6.8}$$

After substituting Formula (6.3), Formula (6.8) can be further expressed as follows:

$$
\begin{aligned}
&ETC_t(x_{S1}^{(t)}, x_{S2}^{(t)}, p_1^{(t)}, p_2^{(t)}, r_{11}^{(t)}, r_{12}^{(t)}, r_{21}^{(t)}, r_{22}^{(t)}) \\
&\approx \tilde{ETC}_t(x_{S1}^{(t)}, x_{S2}^{(t)}, p_1^{(t)}, p_2^{(t)}, r_{11}^{(t)}, r_{12}^{(t)}, r_{21}^{(t)}, r_{22}^{(t)}) \\
&= u_1^{(t)}\mu_1 + u_2^{(t)}\mu_2 + c_{P1}p_1^{(t)} + c_{P2}p_2^{(t)} + c_{R11}r_{11}^{(t)} + c_{R12}r_{12}^{(t)} + c_{R21}r_{21}^{(t)} + c_{R22}r_{22}^{(t)} \\
&\quad +(h_1 + u_1^{(t)})\int_0^{x_{S1}^{(t)}+p_1^{(t)}+r_{11}^{(t)}+r_{21}^{(t)}} (x_{S1}^{(t)} + p_1^{(t)} + r_{11}^{(t)} + r_{21}^{(t)} - D_1) f(D_1^{(t)}, \mu_1, \sigma_1) dD_1^{(t)} \\
&\quad +(v_1 - u_1^{(t)})\int_{x_{S1}^{(t)}+p_1^{(t)}+r_{11}^{(t)}+r_{21}^{(t)}}^{\infty} (D_1^{(t)} - x_{S1}^{(t)} - p_1^{(t)} - r_{11}^{(t)} - r_{21}^{(t)}) f(D_1^{(t)}, \mu_1, \sigma_1) dD_1^{(t)} \\
&\quad +(h_2 + u_2^{(t)})\int_0^{x_{S2}^{(t)}+p_2^{(t)}+r_{12}^{(t)}+r_{22}^{(t)}} (x_{S2}^{(t)} + p_2^{(t)} + r_{12}^{(t)} + r_{22}^{(t)} - D_2^{(t)}) f(D_2^{(t)}, \mu_2, \sigma_2) dD_2^{(t)} \\
&\quad +(v_2 - u_2^{(t)})\int_{x_{S2}^{(t)}+p_2^{(t)}+r_{12}^{(t)}+r_{22}^{(t)}}^{\infty} (D_2^{(t)} - x_{S2}^{(t)} - p_2^{(t)} - r_{12}^{(t)} - r_{22}^{(t)}) f(D_2^{(t)}, \mu_2, \sigma_2) dD_2^{(t)}.
\end{aligned}
\tag{6.9}
$$

The functions EC_{M} and $\tilde{ETC}_t$, expressed in Formulae (6.5) and (6.9) respectively, are found to be similar to each other except for some coefficient differences. Therefore, we can prove the convex property of the function $\tilde{ETC}_t$ similar to the function EC_{M}. Thus, we can find the optimal solution to minimize the function $\tilde{ETC}_t$ by solving KKT conditions. The optimal solution has the same structures as that for the single-period problem in Appendix B. Thus, the policy of the multi-period problem by solving the ADP model is similar to the optimal policy of the single-period problem. However, due to coefficient differences, the threshold levels of the policy for the multi-period problem need to be re-computed. For example, the threshold level $AL_0^{(t)}$ can be determined as follows:

$$
\begin{aligned}
&\left.\frac{\partial \tilde{ETC}_t}{\partial p_1^{(t)}}\right|_{x_1^{(t)}=AL_0^{(t)}} = 0 \\
&\Rightarrow u_1^{(t)} - v_1 + c_{P1} + (v_1 + h_1)F(AL_0^{(t)}, \mu_1, \sigma_1) = 0 \\
&\Rightarrow AL_0^{(t)} = F^{-1}\left(\frac{v_1 - u_1^{(t)} - c_{P1}}{v_1 + h_1}, \mu_1, \sigma_1\right).
\end{aligned}
\tag{6.10}
$$

Given the gradients $u_1^{(t)}$ and $u_2^{(t)}$, we can determine all the threshold levels in period t for the multi-period problem. In Table 6.1, we have listed the formulae of determining the threshold levels for the single-period problem and the multi-period problem respectively.

Table 6.1 The formulae of determining the threshold levels for the single-period problem and the multi-period problem

	The single-period problem	**The multi-period problem**
$AL_0^{(t)}$	$F^{-1}(\frac{v_1}{v_1+h_1+c_{P1}},\mu_1,\sigma_1)$	$F^{-1}(\frac{v_1-u_1^{(t)}-c_{P1}}{v_1+h_1},\mu_1,\sigma_1)$
$AL_1^{(t)}$	$F^{-1}(\frac{v_1+c_{P1}-c_{R11}}{v_1+h_1+c_{P1}},\mu_1,\sigma_1)$	$F^{-1}(\frac{v_1-u_1^{(t)}-c_{R11}}{v_1+h_1},\mu_1,\sigma_1)$
$AL_2^{(t)}$	$F^{-1}(\frac{v_1+c_{P1}-c_{R21}}{v_1+h_1+c_{P1}},\mu_1,\sigma_1)$	$F^{-1}(\frac{v_1-u_1^{(t)}-c_{R21}}{v_1+h_1},\mu_1,\sigma_1)$
$SW_1^{(t)}$	$F^{-1}(\frac{v_1+c_{P1}+c_{R12}-c_{R11}-c_{P2}}{v_1+h_1+c_{P1}},\mu_1,\sigma_1)$	$F^{-1}(\frac{v_1-u_1^{(t)}+c_{R12}-c_{R11}-c_{P2}}{v_1+h_1},\mu_1,\sigma_1)$
$SW_2^{(t)}$	$F^{-1}(\frac{v_1+c_{P1}+c_{R22}-c_{R21}-c_{P2}}{v_1+h_1+c_{P1}},\mu_1,\sigma_1)$	$F^{-1}(\frac{v_1-u_1^{(t)}+c_{R22}-c_{R21}-c_{P2}}{v_1+h_1},\mu_1,\sigma_1)$
$RP^{(t)}$	$F^{-1}(\frac{v_1+c_{P1}+c_{R22}-c_{R21}-c_{R12}}{v_1+h_1+c_{P1}},\mu_1,\sigma_1)$	$F^{-1}(\frac{v_1-u_1^{(t)}+c_{R22}-c_{R21}-c_{R12}}{v_1+h_1},\mu_1,\sigma_1)$
$BL_0^{(t)}$	$F^{-1}(\frac{v_2}{v_2+h_2+c_{P2}},\mu_2,\sigma_2)$	$F^{-1}(\frac{v_2-u_2^{(t)}-c_{P2}}{v_2+h_2},\mu_2,\sigma_2)$
$BL_1^{(t)}$	$F^{-1}(\frac{v_2+c_{P2}-c_{R12}}{v_2+h_2+c_{P2}},\mu_2,\sigma_2)$	$F^{-1}(\frac{v_2-u_2^{(t)}-c_{R12}}{v_2+h_2},\mu_2,\sigma_2)$
$BL_2^{(t)}$	$F^{-1}(\frac{v_2+c_{P2}-c_{R22}}{v_2+h_2+c_{P2}},\mu_2,\sigma_2)$	$F^{-1}(\frac{v_2-u_2^{(t)}-c_{R22}}{v_2+h_2},\mu_2,\sigma_2)$

Similar to Chapter 5, we determine each threshold level for the multi-period problem in Table 6.1 by an iterative learning algorithm, which uses the gradients of the cost-to-go function at the points of interest. The details can be referred to in Chapter 5. In the following, we will introduce how to determine the two gradients at the point of interest.

6.3 The determination of the gradient at the points of interest in the multi-period context

Without closed-form formula of the gradient $u_j^{(t)}$ at the point of interest ($x_1^{(t)}$, $x_2^{(t)}$), we need to run Monte Carlo simulation, then estimate the gradient based on the simulation results. Before that, we need to approximate the cost-to-go function by Monte Carlo formulation. In Monte Carlo sampling, sample k is about the realization of stochastic returns and demands in each period from period t till period M-1, which is expressed as $\begin{pmatrix} R_{1,k}^{(t+1)} & R_{2,k}^{(t+1)} & D_{1,k}^{(t)} & D_{2,k}^{(t)} \\ .... & & & \\ R_{1,k}^{(M)} & R_{2,k}^{(M)} & D_{1,k}^{(M-1)} & D_{2,k}^{(M-1)} \end{pmatrix}$. The cost-to-go function is approximated as follows:

$$\begin{aligned} & \underset{D_1^{(t)},D_2^{(t)}}{E}[f_{t+1}(x_{S1}^{(t+1)},x_{S2}^{(t+1)})] \\ & \approx \frac{1}{N}\sum_{k=1}^{N}(\sum_{\tau=t+1}^{M-1} ATC_k^{(\tau)*} + MEC_k). \end{aligned} \tag{6.11}$$

In the above formula, the sample value for sample k, is obtained by summing up the cost for the realization from period t+1 till period M-1 and the minimum expected cost at period M. Without a closed-form formula for the function MEC_k, we compute it by minimizing the expected cost $EC_{\mathrm{M,\,k}}$. In addition, the function $ATC_k^{(\tau)*}$ is used to calculate the cost of period τ ($t < \tau < M$) for sample k. The function $ATC_k^{(\tau)*}$ is expressed as follows:

$$
\begin{aligned}
& ATC_k^{(\tau)*}(x_{S1}^{(\tau)}, x_{S2}^{(\tau)}, p_1^{(\tau)*}, p_2^{(\tau)*}, r_{11}^{(\tau)*}, r_{12}^{(\tau)*}, r_{21}^{(\tau)*}, r_{22}^{(\tau)*}, D_{1,k}^{(\tau)*}, D_{2,k}^{(\tau)*}) \\
& = c_{P1} p_1^{(\tau)*} + c_{P2} p_2^{(\tau)*} + c_{R11} r_{11}^{(\tau)*} + c_{R12} r_{12}^{(\tau)*} + c_{R21} r_{21}^{(\tau)*} + c_{R22} r_{22}^{(\tau)*} \\
& + h_{s1}[x_{S1}^{(\tau)} + p_1^{(\tau)*} + r_{11}^{(\tau)*} + r_{21}^{(\tau)*} - D_{1,k}^{(\tau)*}]^+ + v_1[D_{1,k}^{(\tau)*} - x_{S1}^{(\tau)} - p_1^{(\tau)*} - r_{11}^{(\tau)*} - r_{21}^{(\tau)*}]^+ \\
& + h_{s2}[x_{S2}^{(\tau)} + p_2^{(\tau)*} + r_{12}^{(\tau)*} + r_{22}^{(\tau)*} - D_{2,k}^{(\tau)*}]^+ + v_2[D_{2,k}^{(\tau)*} - x_{S2}^{(\tau)} - p_2^{(\tau)*} - r_{12}^{(\tau)*} - r_{22}^{(\tau)*}]^+ .
\end{aligned}
\tag{6.12}
$$

At period M, the minimum expected cost MEC_{k} is calculated by minimizing the expected cost $EC_{\mathrm{M,k}}$ in Formula (6.3).

As the functions $EC_{\mathrm{M,\,k}}$ and $ATC_k^{(\tau)*}$ are both continuous functions, it is suitable to approximate the cost-to-go function by Monte Carlo sampling method. Furthermore, the gradient $u_j^{(t)}$ can be approximated by sample average. Corresponding to the gradient $u_j^{(t)}$, $grad_{j,k}^{(t)}$ denotes the sample gradient for sample k at the point ($x_1^{(t)}, x_2^{(t)}$). Thus, the approximation can be expressed as follows:

$$
u_j^{(t)}(x_1^{(t)}, x_2^{(t)}) \approx \frac{1}{N} \sum_{k=1}^{N} grad_{j,k}^{(t)}(x_1^{(t)}, x_2^{(t)}). \tag{6.13}
$$

Starting with the two-period problem, we would introduce the determination of the above-mentioned sample gradient. Then, we would extend from the two-period problem to the three-period problem. Finally, we can determine the sample gradient for any multi-period problem by induction. In the determination of the sample gradient, we have taken advantage of an Infinite Perturbation Analysis (IPA) based approach.

6.3.1 The determination of sample gradient in the two-period problem

For the two-period problem, the threshold levels of the last period can be obtained by referring to the single-period problem. The threshold levels of the first period are determined by using the gradients of the cost-to-go function estimated by Monte Carlo simulation. Before that, the sample gradient of the cost-to-go function needs to be determined. The sample k for Monte Carlo sampling is expressed as $(R_{1,k}^{(2)}, R_{2,k}^{(2)}, D_{1,k}^{(1)}, D_{2,k}^{(1)})$. The sample gradient can be calculated as follows (j =1, 2):

$$grad_{j,k}^{(1)}(x_1^{(1)}, x_2^{(1)}) = \frac{\partial MEC_k}{\partial x_j^{(1)}} = \frac{\partial MEC_k}{\partial x_{Sj}^{(2)}}. \tag{6.14}$$

In the above formula, the term $\frac{\partial MEC_k}{\partial x_{Sj}^{(2)}}$ can be determined as follows (j = 1, 2):

$$\begin{aligned}
&\frac{\partial MEC_k}{\partial x_{Sj}^{(2)}} \\
&= \frac{\partial EC_{2,k}{}^*}{\partial x_{Sj}^{(2)}} + \left(\frac{\partial EC_{2,k}{}^*}{\partial p_1^{(2)*}} \cdot \frac{\partial p_1^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial EC_{2,k}{}^*}{\partial r_{11}^{(2)*}} \cdot \frac{\partial r_{11}^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial EC_{2,k}{}^*}{\partial r_{21}^{(2)*}} \cdot \frac{\partial r_{21}^{(2)*}}{\partial x_{Sj}^{(2)}}\right) \\
&\qquad + \left(\frac{\partial EC_{2,k}{}^*}{\partial p_2^{(2)*}} \cdot \frac{\partial p_2^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial EC_{2,k}{}^*}{\partial r_{12}^{(2)*}} \cdot \frac{\partial r_{12}^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial EC_{2,k}{}^*}{\partial r_{22}^{(2)*}} \cdot \frac{\partial r_{22}^{(2)*}}{\partial x_{Sj}^{(2)}}\right).
\end{aligned} \tag{6.15}$$

According to Formula (6.5), the partial derivatives of the function $EC_{2,k}{}^{*}$ can be determined as follows ($j = 1, 2$):

$$\begin{aligned}
&\frac{\partial EC_{2,k}{}^{*}}{\partial x_{Sj}^{(2)}} = (v_j + h_j + c_{Pj})F(x_j^{(2)*}, \mu_j, \sigma_j) - c_{Pj} - v_j;\\
&\frac{\partial EC_{2,k}{}^{*}}{\partial p_j^{(2)*}} = \frac{\partial EC_{2,k}{}^{*}}{\partial x_{Sj}^{(2)}} + c_{Pj};\\
&\frac{\partial EC_{2,k}{}^{*}}{\partial r_{1j}^{(2)*}} = \frac{\partial EC_{2,k}{}^{*}}{\partial x_{Sj}^{(2)}} + c_{R1j};\\
&\frac{\partial EC_{2,k}{}^{*}}{\partial r_{2j}^{(2)*}} = \frac{\partial EC_{2,k}{}^{*}}{\partial x_{Sj}^{(2)}} + c_{R2j}.
\end{aligned} \tag{6.16}$$

Based on Formulae (6.15) and (6.16), Formula (6.14) is further expressed as follows ($j = 1, 2$):

$$\begin{aligned}
grad_{j,k}^{(1)}(x_1^{(1)}, x_2^{(1)}) = {}& \frac{\partial EC_{2,k}{}^{*}}{\partial x_{S1}^{(2)}}\frac{\partial x_1^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial EC_{2,k}{}^{*}}{\partial x_{S2}^{(2)}}\frac{\partial x_2^{(2)*}}{\partial x_{Sj}^{(2)}} + c_{P1}\frac{\partial p_1^{(2)*}}{\partial x_{Sj}^{(2)}} + c_{R11}\frac{\partial r_{11}^{(2)*}}{\partial x_{Sj}^{(2)}}\\
&+ c_{R21}\frac{\partial r_{21}^{(2)*}}{\partial x_{Sj}^{(2)}} + c_{P2}\frac{\partial p_2^{(2)*}}{\partial x_{Sj}^{(2)}} + c_{R12}\frac{\partial r_{12}^{(2)*}}{\partial x_{Sj}^{(2)}} + c_{R22}\frac{\partial r_{22}^{(2)*}}{\partial x_{Sj}^{(2)}}.
\end{aligned} \tag{6.17}$$

The above formula involves the partial derivatives of the optimal replenishment decisions with respect to initial inventory. These partial derivatives can be obtained according to the corresponding structure in Appendix B. Suppose that the optimal replenishment decisions for sample k match the structure S7 in the Appendix as mentioned in Chapter 5. Therefore, the sample gradient in Formula (6.17) is calculated as follows:

$$grad_{1,k}^{(1)}(x_1^{(1)},x_2^{(1)})=c_{R12}-c_{R11}-c_{P2};$$

$$grad_{2,k}^{(1)}(x_1^{(1)},x_2^{(1)})=-c_{P2}.$$

6.3.2 The determination of sample gradient in the three-period problem

For the three-period problem, the threshold levels of the second and the last period can be obtained by referring to the two-period problem and the single-period problem respectively. The threshold levels of the first period are determined by using the two gradients of the cost-to-go function with respect to inventory levels of the two products after replenishment. As the two gradients are estimated by Monte Carlo simulation, the sample gradient of the cost-to-go function needs to be determined at first. The sample k for Monte Carlo sampling is expressed as $\begin{pmatrix} R_{1,k}^{(2)} & R_{2,k}^{(2)} & D_{1,k}^{(1)} & D_{2,k}^{(1)} \\ R_{1,k}^{(3)} & R_{2,k}^{(3)} & D_{1,k}^{(2)} & D_{2,k}^{(2)} \end{pmatrix}$.

The sample gradient can be calculated as follows (j = 1, 2):

$$\begin{aligned}
& grad_{j,k}^{(1)}(x_1^{(1)},x_2^{(1)}) \\
&= \frac{\partial}{\partial x_j^{(1)}}\{ATC_k^{(2)*}+MEC_k\} \\
&= \frac{\partial}{\partial x_{Sj}^{(2)}}\{ATC_k^{(2)*}+MEC_k\} \\
&= \frac{\partial ATC_k^{(2)*}}{\partial x_{Sj}^{(2)}}+\frac{\partial MEC_k}{\partial x_{Sj}^{(2)}} \\
&= \frac{\partial ATC_k^{(2)*}}{\partial x_{Sj}^{(2)}}+\frac{\partial MEC_k}{\partial x_1^{(2)*}}\frac{\partial x_1^{(2)*}}{\partial x_{Sj}^{(2)}}+\frac{\partial MEC_k}{\partial x_2^{(2)*}}\frac{\partial x_2^{(2)*}}{\partial x_{Sj}^{(2)}} \\
&= \frac{\partial ATC_k^{(2)*}}{\partial x_{Sj}^{(2)}}+\frac{\partial x_1^{(2)*}}{\partial x_{Sj}^{(2)}}grad_{1,k}^{(2)}(x_1^{(2)*},x_2^{(2)*})+\frac{\partial x_2^{(2)*}}{\partial x_{Sj}^{(2)}}grad_{2,k}^{(2)}(x_1^{(2)*},x_2^{(2)*}).
\end{aligned} \tag{6.18}$$

In the above formula, the term $\frac{\partial x_i^{(2)*}}{\partial x_{Sj}^{(2)}}$ is calculated as follows (i, j = 1, 2):

$$\frac{\partial x_i^{(2)*}}{\partial x_{Sj}^{(2)}} = \frac{\partial x_{Si}^{(2)}}{\partial x_{Sj}^{(2)}} + \frac{\partial p_i^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial r_{1i}^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial r_{2i}^{(2)*}}{\partial x_{Sj}^{(2)}}. \tag{6.19}$$

The above formula involves the partial derivatives of the optimal replenishment decisions with respect to initial inventory of the two products at period *2*. Similar to the two-period problem, these derivatives can be obtained by referring to the corresponding solution structure in Appendix B. Before that, the threshold levels of the optimal policy at period *2* are determined on the basis of the objective function $\tilde{ETC}_t$ considering both period *2* and period *3*. The two gradients of $\tilde{ETC}_t$ at the point of interest need to be estimated in the two-period context. Suppose that the optimal replenishment decisions at period *2* for sample *k* match the above-mentioned structure S7 with the threshold levels at period *2*. Therefore, the sample gradient for sample *k* is calculated as $grad_{j,k}^{(1)}(x_1^{(1)}, x_2^{(1)}) = \frac{\partial ATC_k^{(2)*}}{\partial x_{Sj}^{(2)}}$ by Formula (6.18). The partial derivatives of the function $ATC_k^{(2)*}$ can be determined as follows (j = 1, 2):

$$\begin{aligned}
&\frac{\partial ATC_k^{(2)*}}{\partial x_{Sj}^{(2)}} \\
&= \frac{\partial ATC_k^{(2)*}}{\partial x_{Sj}^{(2)}} + \left(\frac{\partial ATC_k^{(2)*}}{\partial p_1^{(2)*}}\frac{\partial p_1^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial ATC_k^{(2)*}}{\partial r_{11}^{(2)*}}\frac{\partial r_{11}^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial ATC_k^{(2)*}}{\partial r_{21}^{(2)*}}\frac{\partial r_{21}^{(2)*}}{\partial x_{Sj}^{(2)}}\right) \\
&\qquad + \left(\frac{\partial ATC_k^{(2)*}}{\partial p_2^{(2)*}}\frac{\partial p_2^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial ATC_k^{(2)*}}{\partial r_{12}^{(2)*}}\frac{\partial r_{12}^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial ATC_k^{(2)*}}{\partial r_{22}^{(2)*}}\frac{\partial r_{22}^{(2)*}}{\partial x_{Sj}^{(2)}}\right).
\end{aligned} \tag{6.20}$$

According to Formula (6.12), the related partial derivatives of the function $ATC_k^{(2)*}$ are determined and listed in Table 6.2 as follows. While calculating these partial derivatives, we have considered all the combinations of demand satisfaction. In order to summarize all the possible expressions, the related index and indicator are excluded from the notations in Table 6.2.

Table 6.2 The partial derivatives of the function $ATC_k^{(\tau)*}$ with respect to initial inventory and replenishment decisions

	$\frac{\partial ATC}{\partial x_{S1}}$	$\frac{\partial ATC}{\partial x_{S2}}$	$\frac{\partial ATC}{\partial p_1}$	$\frac{\partial ATC}{\partial r_{11}}$	$\frac{\partial ATC}{\partial r_{21}}$	$\frac{\partial ATC}{\partial p_2}$	$\frac{\partial ATC}{\partial r_{12}}$	$\frac{\partial ATC}{\partial r_{22}}$
Product 1 underage product 2 underage	$-v_1$	$-v_2$	$c_{P1}-v_1$	$c_{R11}-v_1$	$c_{R21}-v_1$	$c_{P2}-v_2$	$c_{R12}-v_2$	$c_{R22}-v_2$
Product 1 underage product 2 overage	$-v_1$	h_2	$c_{P1}-v_1$	$c_{R11}-v_1$	$c_{R21}-v_1$	h_2+c_{P2}	h_2+c_{R12}	h_2+c_{R22}
Product 1 overage product 2 underage	h_1	$-v_2$	h_1+c_{P1}	h_1+c_{R11}	h_1+c_{R21}	$c_{P2}-v_2$	$c_{R12}-v_2$	$c_{R22}-v_2$
Product 1 overage product 2 overage	h_1	h_2	h_1+c_{P1}	h_1+c_{R11}	h_1+c_{R21}	h_2+c_{P2}	h_2+c_{R12}	h_2+c_{R22}

With reference to Formula (6.20), the two gradients for sample k are calculated as follows. According to the situation of demand satisfaction in period 2, we can refer to Table 6.2 to obtain the values of the related partial derivatives in the following formulae.

$$grad_{1,k}^{(1)}(x_1^{(1)}, x_2^{(1)}) = \frac{\partial ATC_k^{(2)*}}{\partial x_{S1}^{(2)}} + \frac{\partial ATC_k^{(2)*}}{\partial r_{12}^{(2)*}} - \frac{\partial ATC_k^{(2)*}}{\partial r_{11}^{(2)*}} - \frac{\partial ATC_k^{(2)*}}{\partial p_2^{(2)*}} = c_{R12} - c_{R11} - c_{P2};$$

$$grad_{2,k}^{(1)}(x_1^{(1)}, x_2^{(1)}) = \frac{\partial ATC_k^{(2)*}}{\partial x_{S2}^{(2)}} - \frac{\partial ATC_k^{(2)*}}{\partial p_2^{(2)*}} = -c_{P2}.$$

6.3.3 The determination of sample gradient in the *N*-period problem

By backward induction as mentioned in Chapter 5, the sample gradient of the first period for the *N*-period problem can be calculated as follows (j = 1, 2):

$$
\begin{aligned}
& grad_{j,k}^{(1)}(x_1^{(1)}, x_2^{(1)}) \\
&= \frac{\partial}{\partial x_j^{(1)}}\{\sum_{\tau=2}^{N-1} ATC_k^{(\tau)*} + MEC_k\} \\
&= \frac{\partial}{\partial x_{Sj}^{(2)}}\{\sum_{\tau=2}^{N-1} ATC_k^{(\tau)*} + MEC_k\} \\
&= \frac{\partial ATC_k^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial}{\partial x_{Sj}^{(2)}}\{\sum_{\tau=3}^{N-1} ATC_k^{(\tau)*} + MEC_k\} \\
&= \frac{\partial ATC_k^{(2)*}}{\partial x_{Sj}^{(2)}} + \frac{\partial x_1^{(2)*}}{\partial x_{Sj}^{(2)}} grad_{1,k}^{(2)}(x_1^{(2)*}, x_2^{(2)*}) + \frac{\partial x_2^{(2)*}}{\partial x_{Sj}^{(2)}} grad_{2,k}^{(2)}(x_1^{(2)*}, x_2^{(2)*}).
\end{aligned}
\tag{6.21}
$$

For the sample gradient at period t ($1 \le t \le M$) in the M-period horizon of the two-product recovery system, we can take advantage of backward way to determine it. The sample gradient at period M *-1* can be determined by solving the two-period problem considering the final two periods. Then, the sample gradient at period M*-2* can be determined by solving the three-period problem considering the final three periods. In this way of backward induction, the sample gradient at period t can be finally determined. The process of determining the sample gradient at period t can be referred to in Appendix C.

6.4 Computational results

6.4.1 The impact of stochastic returns and demands on the threshold levels

Based on the same three sets of system parameters as Chapter 5, we will investigate the impact of stochastic returns and demands on the threshold levels, which have converged in the multi-period context. Firstly, we will investigate how the expected value of returned items affects the threshold levels. Secondly, we will investigate the impact of demand variability on the threshold levels.

6.4.1.1 The impact of the expected value of returned items on the threshold levels

We will investigate the impact of the expected value of returned items in two groups on the threshold levels based on the following stochastic demands:

$$E[D_1] = 200, StDev[D_1] = 60; \quad E[D_2] = 100, StDev[D_2] = 30.$$

The impact of the expected value of returned items in group 1

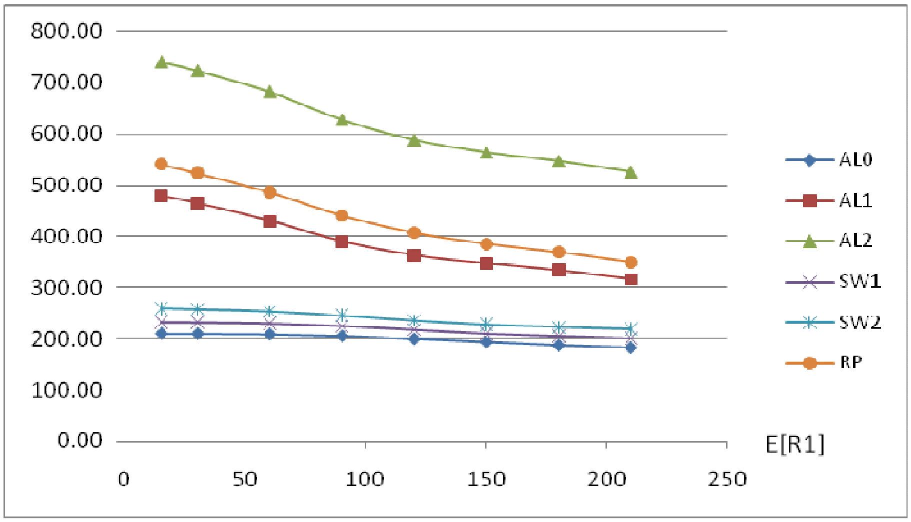
Firstly, the impact of the expected value of returned items in group 1 will be investigated on the same scenarios as Table 5.6.

In the following, the threshold levels of the two products from solving the approximate dynamic programming model are shown in Table 6.3. Furthermore, the trend of the threshold levels is shown in Figure 6.1. The results have shown that all the threshold levels decrease with the expected value of returned items in group 1. For product 1, the threshold levels AL_1, AL_2 and RP are decreasing faster than its other threshold levels. On the other hand, for product 2, the threshold levels BL_1 and BL_2 are decreasing faster than BL_0. As more returned items are available for the recovery in each period, the threshold levels would be decreased.

As the interactive allocation of the returned items in two groups, the expected value of returned items in group 1 would impact the threshold levels of the recovery processes using the returned items in each group. In addition, the expected value of returned items in group 1 has less impact on the threshold levels related to production and switching. As production never uses the returned items, it would not be impacted by the expected value of returned items. In addition, the two switching levels related to product 1, i.e. SW_1 and SW_2, are from the comparison of marginal profits of the recovery using returned items in group 1 and group 2 while the inventory level of product 2 is at the threshold level BL_0. Therefore, the expected value of returned items in group 1 has less impact on the two switching levels. However, there is remarkable impact on the threshold level RP, which is from the comparison of marginal profits of the recovery using returned items in group 2 while the inventory level of product 2 is at the threshold level BL_1.

Table 6.3 The threshold levels in different scenarios of returned items in group 1 with parameter set 1

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
$E[R_1]=15$	210.1	478.9	740.8	232.4	259.6	539.4	112.3	216.5	319.3
$E[R_1]=30$	209.7	464.7	724.0	231.7	257.9	522.8	112.1	209.2	313.6
$E[R_1]=60$	208.8	430.7	682.8	229.8	253.8	485.5	111.3	198.4	304.3
$E[R_1]=90$	205.4	390.3	628.1	224.7	245.6	440.2	108.4	190.2	297.5
$E[R_1]=120$	199.5	363.4	588.8	217.5	236.0	406.4	107.6	184.2	286.8
$E[R_1]=150$	193.1	347.6	564.3	210.0	228.2	384.9	107.5	179.2	274.7
$E[R_1]=180$	187.1	333.9	547.2	205.1	223.4	368.8	107.2	174.1	263.1
$E[R_1]=210$	182.6	316.9	526.4	200.7	219.3	349.7	105.6	165.9	248.3



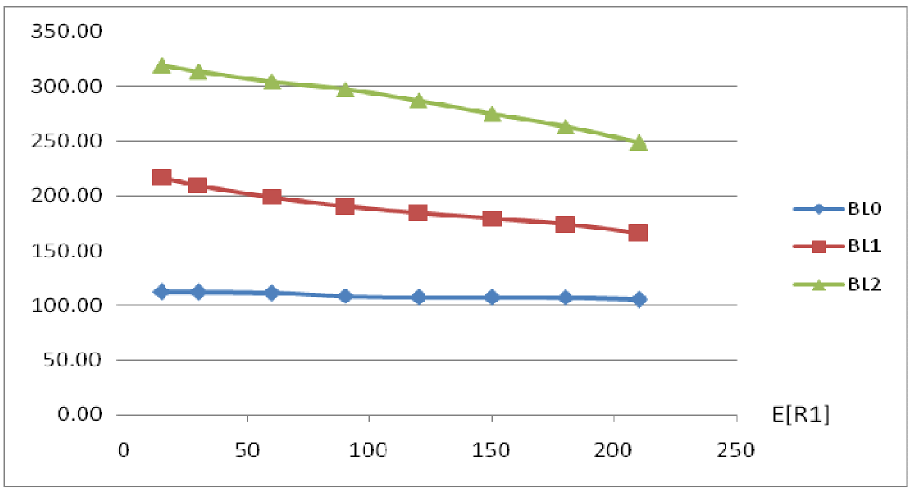
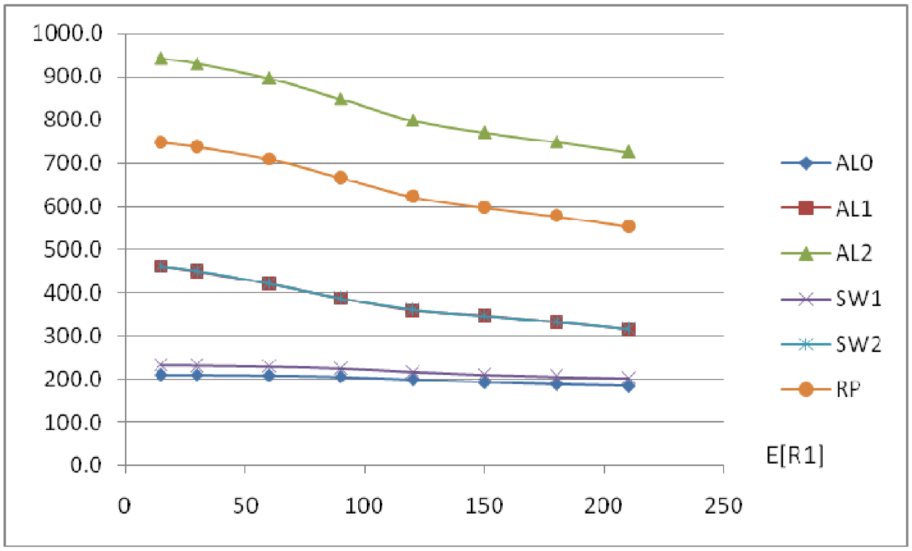


Figure 6.1 The trend of the threshold levels in different scenarios of returned items in group 1 with parameter set 1

Table 6.4 The threshold levels in different scenarios of returned items in group 1 with parameter set 2

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
$E[R_1]=15$	210.0	461.4	943.1	232.5	460.8	746.5	112.3	219.2	318.9
$E[R_1]=30$	209.7	449.6	930.9	231.7	450.4	736.9	111.5	211.3	312.3
$E[R_1]=60$	208.2	421.7	897.0	229.2	421.8	708.6	108.7	197.4	297.8
$E[R_1]=90$	205.4	387.7	848.8	224.8	387.7	665.2	106.3	188.7	287.6
$E[R_1]=120$	199.0	360.3	798.9	216.9	361.7	621.7	105.1	181.8	274.7
$E[R_1]=150$	192.9	347.5	771.0	210.3	347.0	596.7	104.3	177.2	264.7
$E[R_1]=180$	188.3	332.8	749.6	205.9	333.7	577.2	103.9	172.0	255.6
$E[R_1]=210$	184.7	316.0	726.1	202.3	316.8	553.3	102.3	164.3	242.8



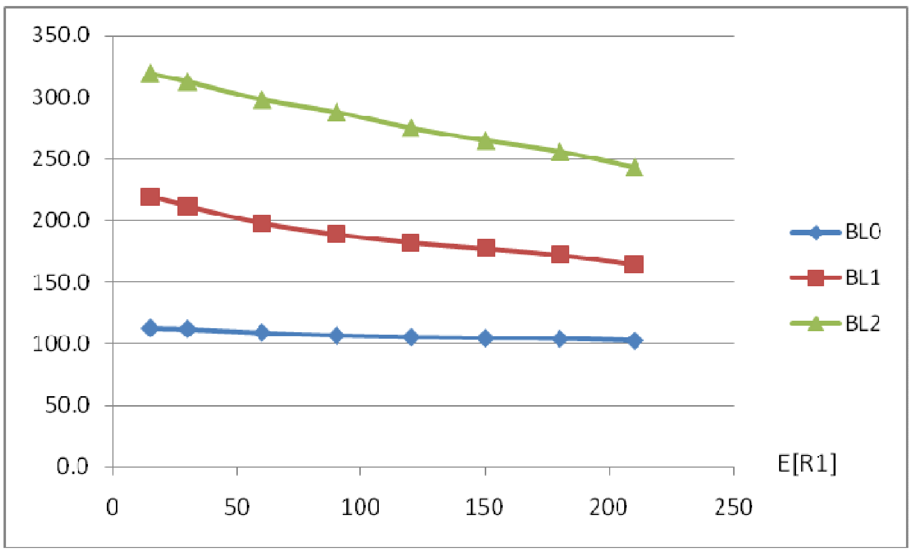
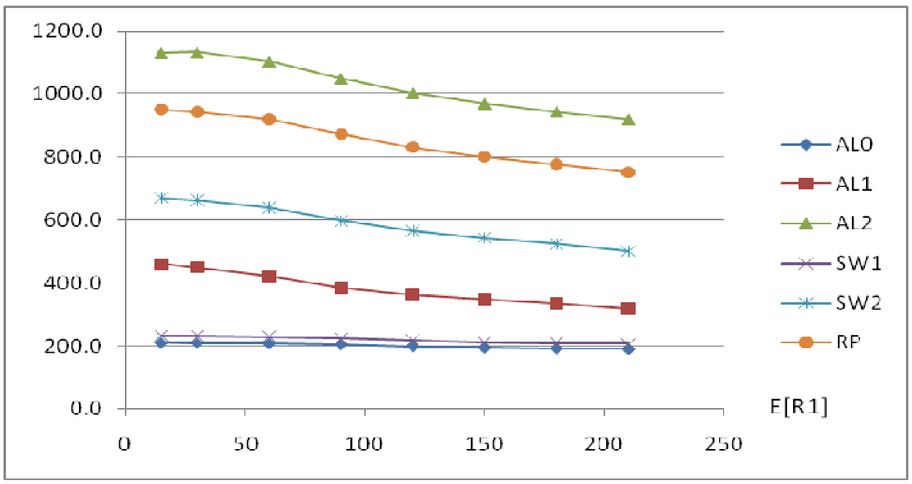


Figure 6.2 The trend of the threshold levels in different scenarios of returned items in group 1 with parameter set 2

Table 6.5 The threshold levels in different scenarios of returned items in group 1 with parameter set 3

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
$E[R_1]=15$	210.1	461.0	1131.5	232.4	670.8	951.3	112.4	219.4	321.4
$E[R_1]=30$	209.7	449.5	1132.9	231.6	662.4	943.6	111.5	210.8	313.2
$E[R_1]=60$	208.3	422.5	1103.4	229.2	639.3	919.9	108.5	197.0	296.5
$E[R_1]=90$	204.9	385.6	1050.3	224.1	598.2	872.7	105.8	187.7	280.6
$E[R_1]=120$	199.6	361.5	1003.3	217.6	565.7	830.8	104.3	181.1	270.0
$E[R_1]=150$	194.8	347.9	970.0	212.0	543.9	801.2	103.4	176.6	261.9
$E[R_1]=180$	192.0	335.0	944.9	209.6	524.7	776.2	103.1	172.7	254.2
$E[R_1]=210$	190.0	318.5	919.2	207.2	501.3	751.7	102.0	166.4	244.7



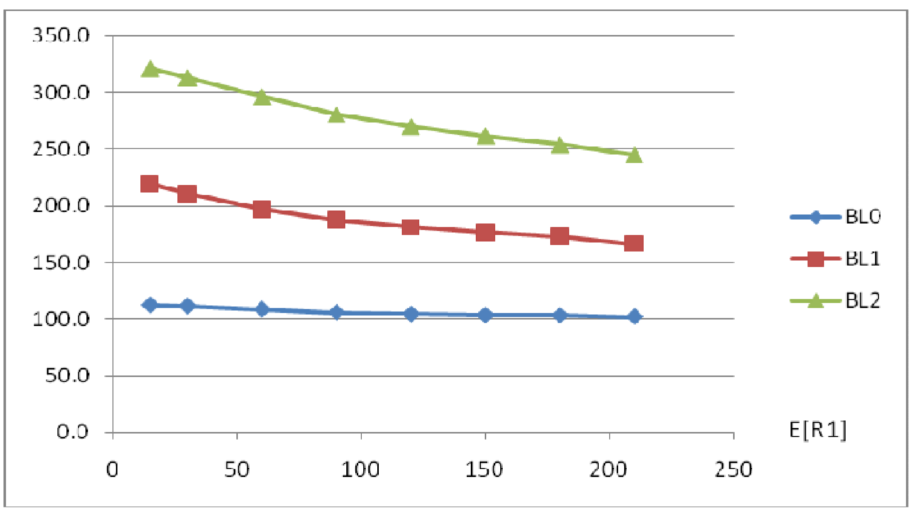


Figure 6.3 The trend of the threshold levels in different scenarios of returned items in group 1 with parameter set 3

The impact of the expected value of returned items in group 2

Secondly, the impact of the expected value of returned items in group 2 will be investigated on the same scenarios as Table 5.10.

In the following, the threshold levels of the two products from solving the approximate dynamic programming model are shown in Table 6.6. Furthermore, the trend of the threshold levels is shown in Figure 6.4. The results have shown that all the threshold levels decrease with the expected value of returned items in group 2. For product 1, the threshold levels AL_1, AL_2 and RP are decreasing faster than its other threshold levels. On the other hand, for product 2, the threshold levels BL_1 and BL_2 are decreasing faster than BL_0. The results can be explained with reference to the above-mentioned impact of the expected value of the returned items in group 1.

Table 6.6 The threshold levels in different scenarios of returned items in group 2 with parameter set 1

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
$E[R_2]=30$	207.1	410.1	660.1	227.4	249.8	462.1	110.0	193.2	302.4
$E[R_2]=45$	205.4	390.3	628.1	224.7	245.6	440.2	108.4	190.2	297.5
$E[R_2]=60$	202.9	373.8	596.4	221.1	240.9	418.9	106.8	187.1	290.8
$E[R_2]=75$	199.8	356.0	567.7	217.6	236.1	398.3	105.5	184.6	284.3
$E[R_2]=90$	196.1	342.3	538.3	213.4	231.5	380.2	104.4	181.8	276.2
$E[R_2]=105$	192.3	327.1	509.3	209.2	226.3	361.6	103.6	178.7	267.4
$E[R_2]=120$	187.6	309.2	475.7	203.7	220.4	341.2	102.9	173.4	255.9

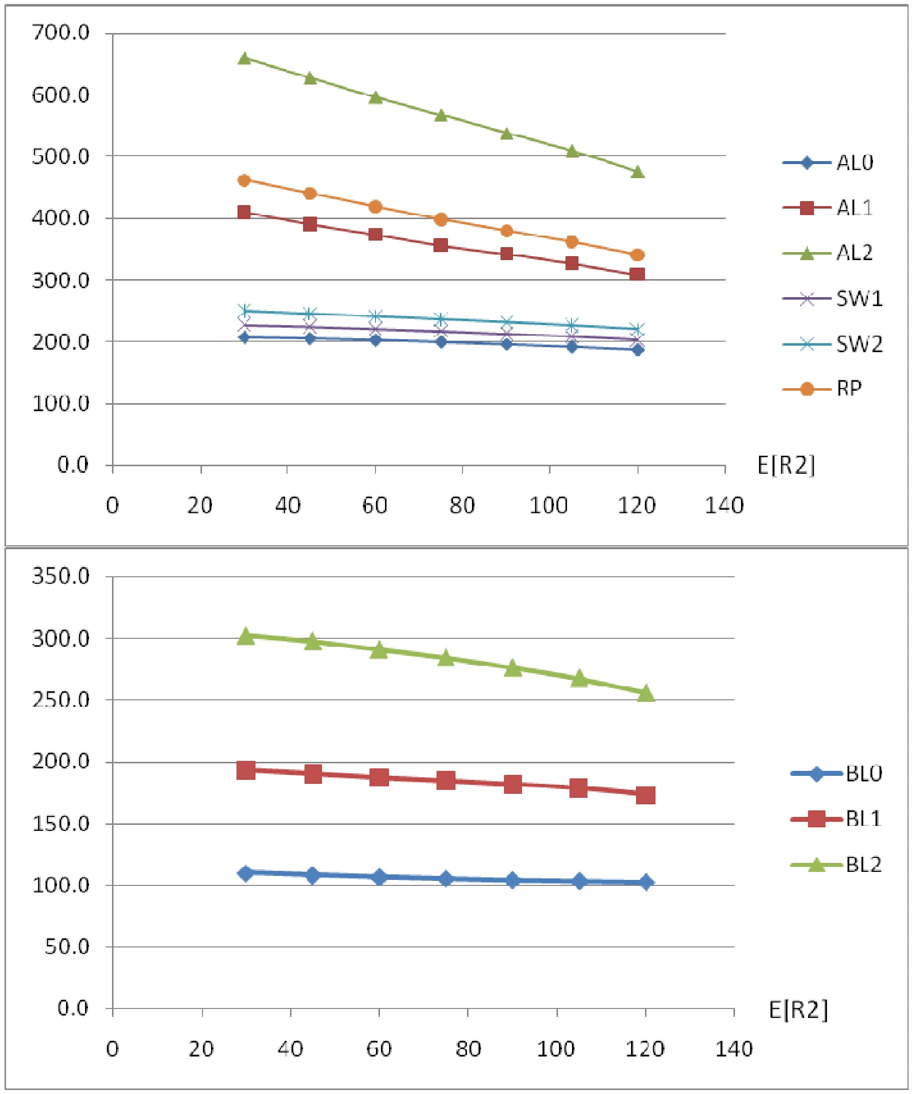
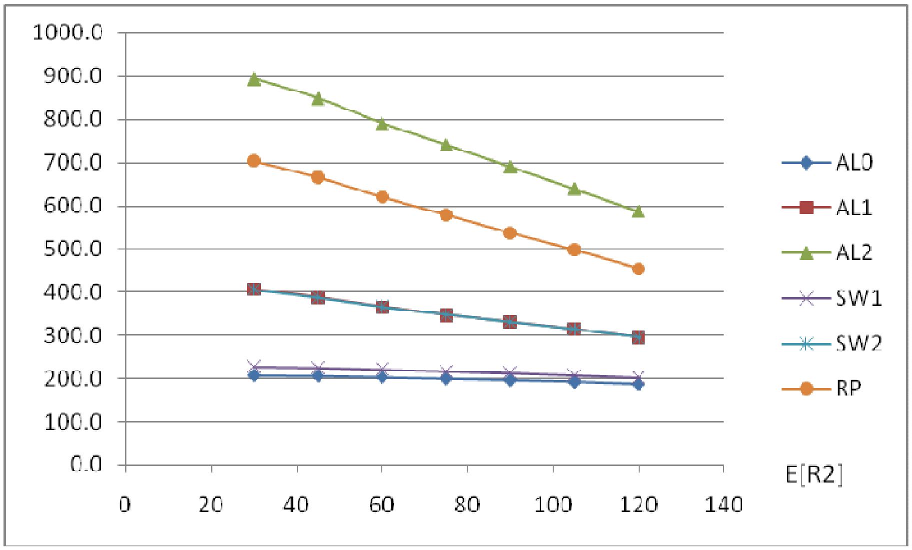


Figure 6.4 The trend of the threshold levels in different scenarios of returned items in group 2 with parameter set 1

Table 6.7 The threshold levels in different scenarios of returned items in group 2 with parameter set 2

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
$E[R_2]=30$	206.7	406.3	894.0	226.9	405.6	702.1	107.7	58.5	288.1
$E[R_2]=45$	205.4	387.7	848.8	224.8	387.7	665.2	106.3	188.7	287.6
$E[R_2]=60$	202.3	365.6	790.5	220.7	365.8	619.4	105.4	186.9	282.1
$E[R_2]=75$	198.9	346.9	740.9	216.9	348.3	578.6	104.5	185.3	276.6
$E[R_2]=90$	195.4	330.3	689.8	212.7	330.6	536.8	103.9	182.5	270.2
$E[R_2]=105$	191.5	313.9	638.8	208.0	314.0	497.7	103.3	178.7	262.4
$E[R_2]=120$	186.1	295.1	586.6	202.4	295.4	454.8	102.8	174.5	254.4



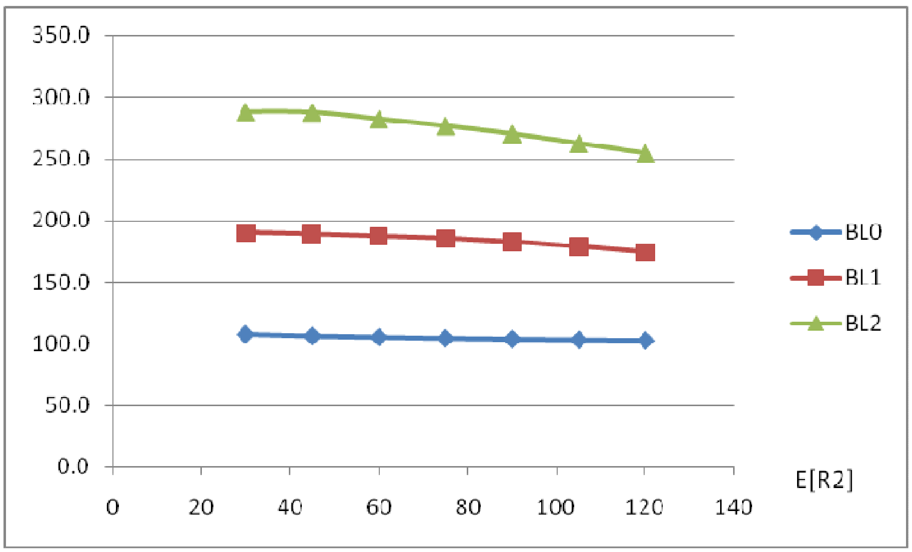
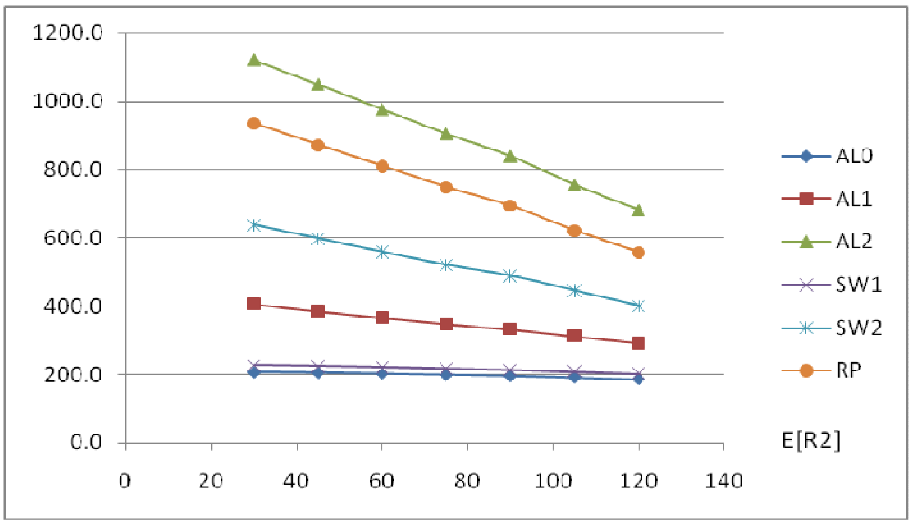


Figure 6.5 The trend of the threshold levels in different scenarios of returned items in group 2 with parameter set 2

Table 6.8 The threshold levels in different scenarios of returned items in group 2 with parameter set 3

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
$E[R_2]=30$	206.7	406.7	1121.3	226.7	637.6	935.1	107.1	189.1	282.1
$E[R_2]=45$	204.9	385.6	1050.3	224.1	598.2	872.7	105.8	187.7	280.6
$E[R_2]=60$	202.3	366.8	977.4	220.9	559.3	810.5	105.1	186.4	278.6
$E[R_2]=75$	199.1	348.2	904.8	217.0	521.5	748.1	104.1	184.4	273.7
$E[R_2]=90$	196.3	332.6	839.4	213.4	490.0	693.3	103.8	182.6	269.3
$E[R_2]=105$	191.6	312.1	754.7	208.0	447.0	621.3	103.4	179.1	261.2
$E[R_2]=120$	186.5	292.7	681.0	202.4	402.6	557.5	103.0	175.1	252.8



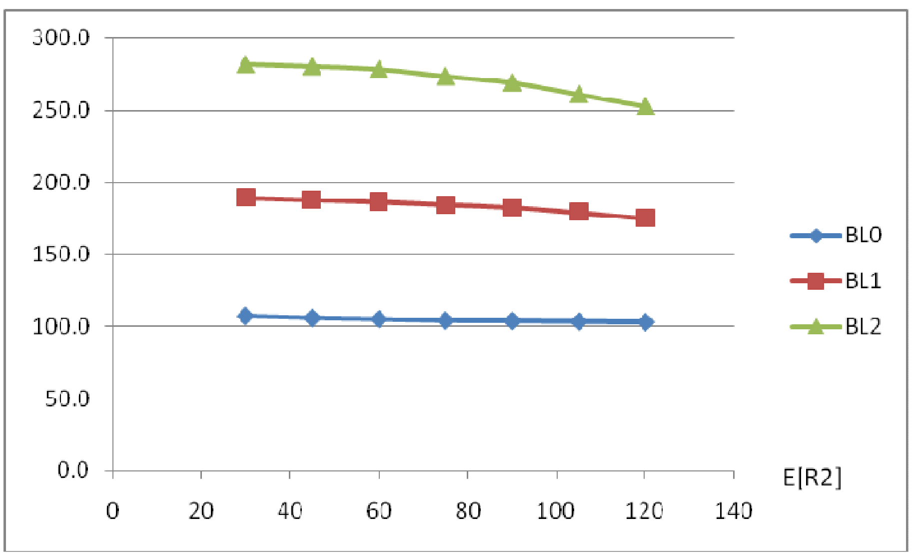


Figure 6.6 The trend of the threshold levels in different scenarios of returned items in group 2 with parameter set 3

6.4.1.2 The impact of demand variability of two products on the threshold levels

We will investigate the impact of demand variability of two products on the threshold levels with the following set of parameters on returned items:

$$E[R_1] = 90, StDev[R_1] = 30; \quad E[R_2] = 45, StDev[R_2] = 15.$$

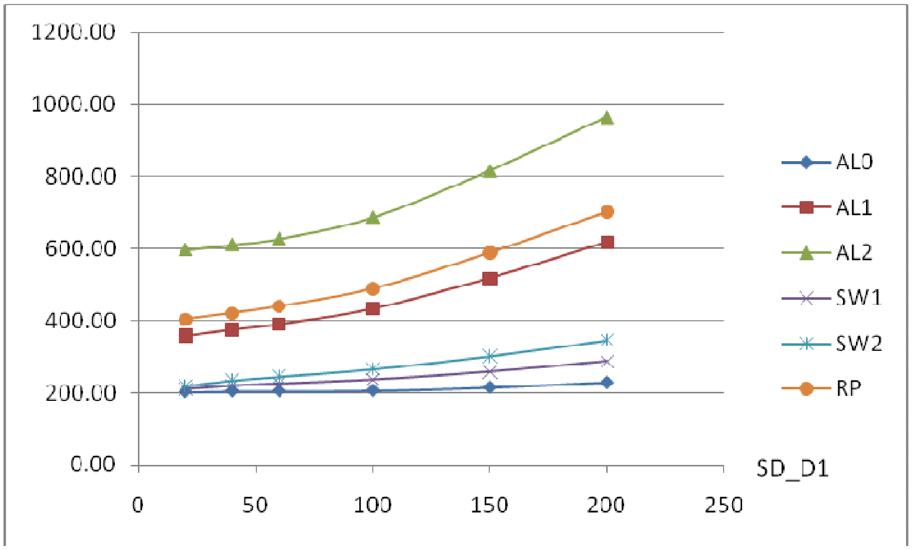
The impact of demand variability of product 1

Firstly, the impact of demand variability of product 1 will be investigated on the same scenarios as Table 5.14.

In the following, the threshold levels of the two products from solving the approximate dynamic programming model are shown in Table 6.9. Furthermore, the trend of the threshold levels is shown in Figure 6.7. The results have shown that all the threshold levels related to product 1 increase with the demand variability of product 1 whereas the threshold levels related to product 2 seem unaffected. As the demands for the two products are independent of each other, the impact of demand variability of product 1 would only affect the threshold levels related to product 1. Furthermore, the higher demand variability results in the higher threshold levels to avoid possible stock shortage.

Table 6.9 The threshold levels in different scenarios of demand for product 1 with parameter set 1

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
COV_1=0.1	203.1	358.0	597.2	210.5	219.6	404.0	109.6	190.9	301.7
COV_1=0.2	205.2	375.2	611.6	219.1	234.8	421.3	109.2	189.7	299.1
COV_1=0.3	205.4	390.3	628.1	224.7	245.6	440.2	108.4	190.2	297.5
COV_1=0.5	206.6	433.1	687.2	236.3	266.5	488.9	108.4	190.7	298.4
COV_1=0.75	214.9	518.1	816.6	258.5	302.3	589.2	109.4	193.1	300.7
COV_1=1.0	227.8	617.7	965.0	287.7	346.1	702.9	110.1	194.1	304.3



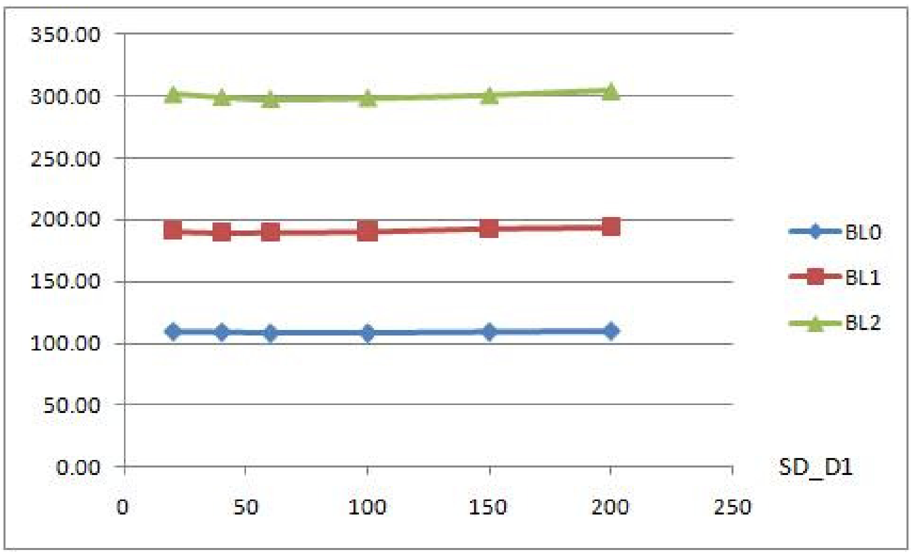
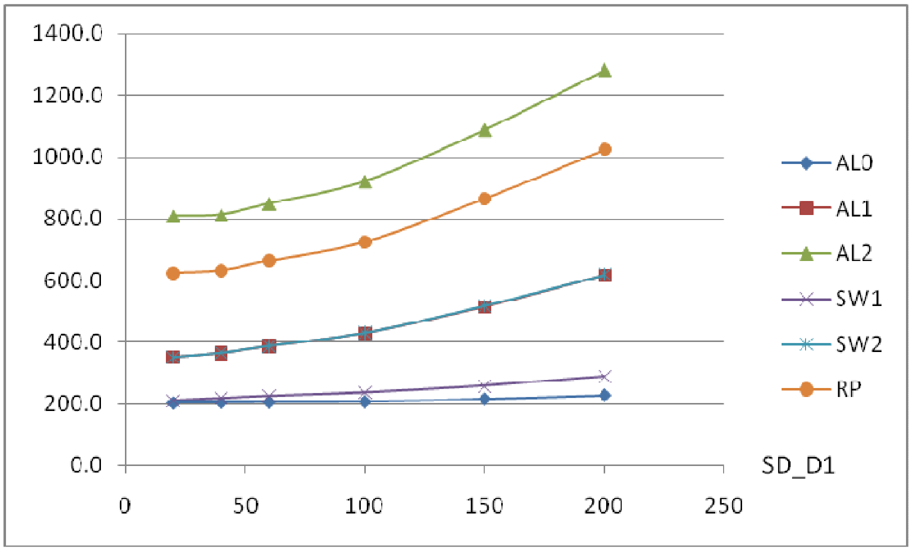


Figure 6.7 The trend of the threshold levels in different scenarios of demand for product 1 with parameter set 1

Table 6.10 The threshold levels in different scenarios of demand for product 1 with parameter set 2

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
COV_1=0.1	202.6	352.0	808.5	209.9	351.5	625.5	106.8	189.2	292.4
COV_1=0.2	203.7	365.5	813.1	217.4	365.3	633.2	106.5	188.5	287.9
COV_1=0.3	205.4	387.7	848.8	224.8	387.7	665.2	106.3	188.7	287.6
COV_1=0.5	206.7	429.8	920.6	236.3	429.8	725.9	106.5	189.1	284.7
COV_1=0.75	215.1	515.1	1087.6	258.7	517.3	865.6	106.5	190.6	286.8
COV_1=1.0	227.3	616.1	1279.9	288.8	617.0	1022.7	106.9	191.4	288.1



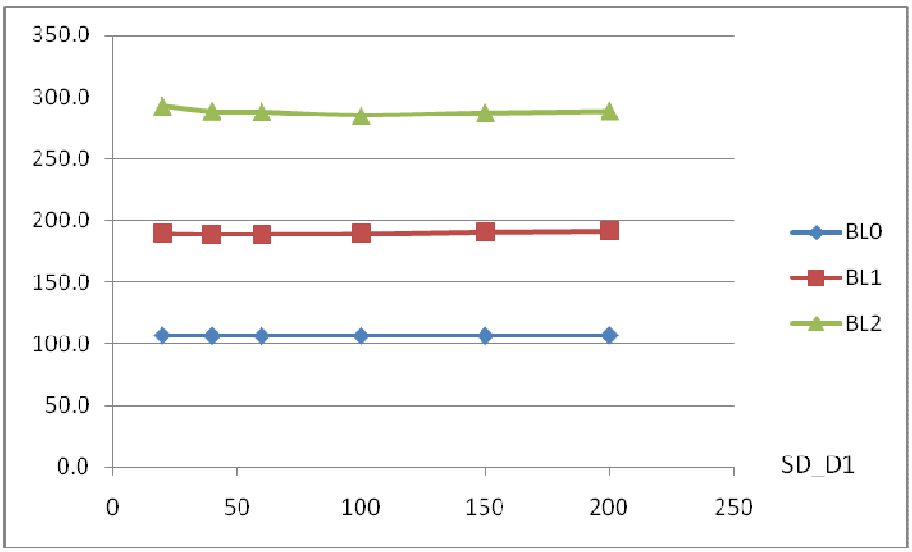
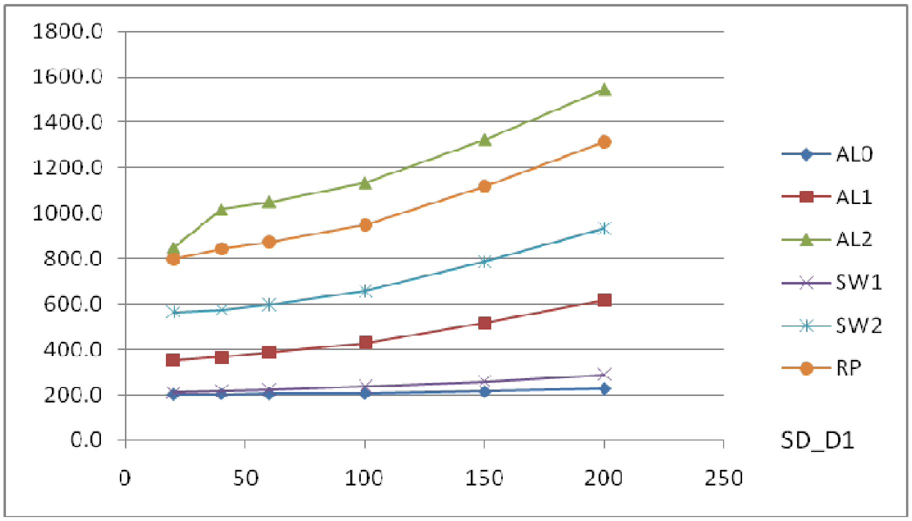


Figure 6.8 The trend of the threshold levels in different scenarios of demand for product 1 with parameter set 2

Table 6.11 The threshold levels in different scenarios of demand for product 1 with parameter set 3

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
COV_1=0.1	202.7	352.2	846.2	210.0	564.1	799.3	106.4	187.9	289.8
COV_1=0.2	203.9	366.2	1016.8	217.3	575.3	842.8	106.2	187.6	282.4
COV_1=0.3	204.9	385.6	1050.3	224.1	598.2	872.7	105.8	187.7	280.6
COV_1=0.5	206.6	429.5	1133.9	236.4	658.5	949.2	105.9	188.4	279.7
COV_1=0.75	214.8	516.0	1324.2	258.9	787.1	1119.1	106.1	190.3	281.1
COV_1=1.0	228.0	617.8	1546.3	288.6	934.2	1315.4	106.6	191.3	283.1



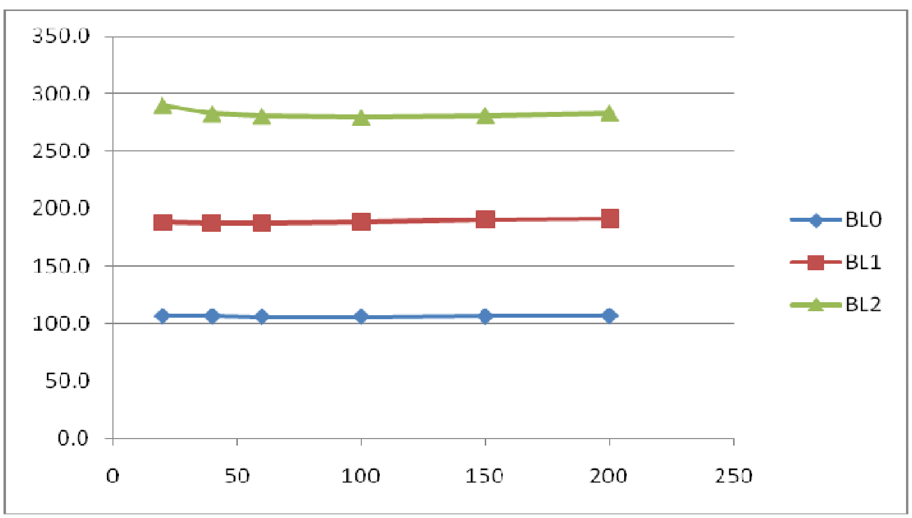


Figure 6.9 The trend of the threshold levels in different scenarios of demand for product 1 with parameter set 3

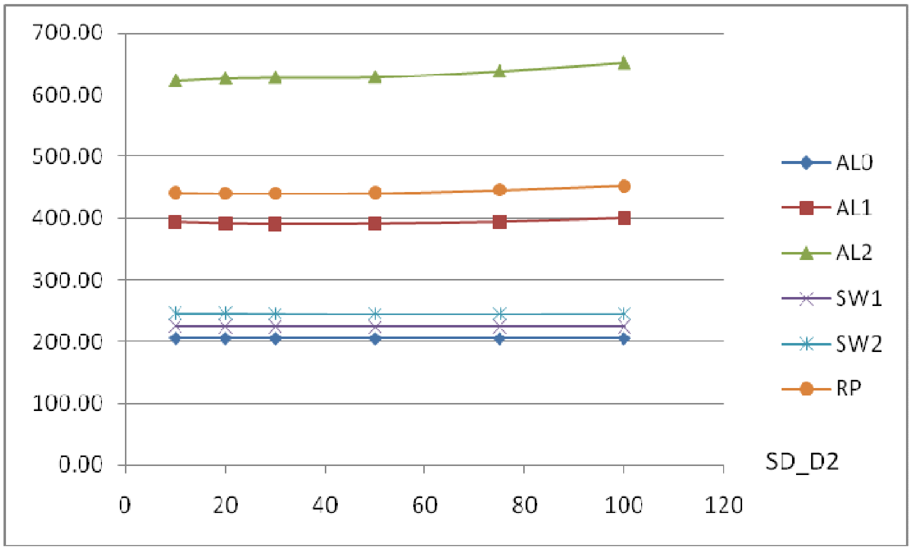
The impact of demand variability of product 2

Secondly, the impact of demand variability of product 2 will be investigated on the same scenarios as Table 5.18.

In the following, the threshold levels of the two products from solving the approximate dynamic programming model are shown in Table 6.12. Furthermore, the trend of the threshold levels is shown in Figure 6.10. The results have shown that all the threshold levels related to product 2 increase with the demand variability of product 2 whereas the threshold levels related to product 1 seem unaffected. The explanation to the results is similar to that on the impact of demand variability of product 1.

Table 6.12 The threshold levels in different scenarios of demand for product 2 with parameter set 1

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
COV_2=0.1	205.7	393.7	623.1	225.3	246.0	441.4	103.9	165.0	276.3
COV_2=0.2	205.6	391.5	627.0	224.7	246.0	440.0	106.5	177.0	286.7
COV_2=0.3	205.4	390.3	628.1	224.7	245.6	440.2	108.4	190.2	297.5
COV_2=0.5	205.2	391.0	628.9	224.3	245.1	440.7	114.3	214.1	321.6
COV_2=0.75	205.2	393.9	638.4	224.5	245.2	445.5	124.6	253.3	375.8
COV_2=1.0	205.3	399.9	651.5	224.5	245.4	452.3	137.3	307.2	456.1



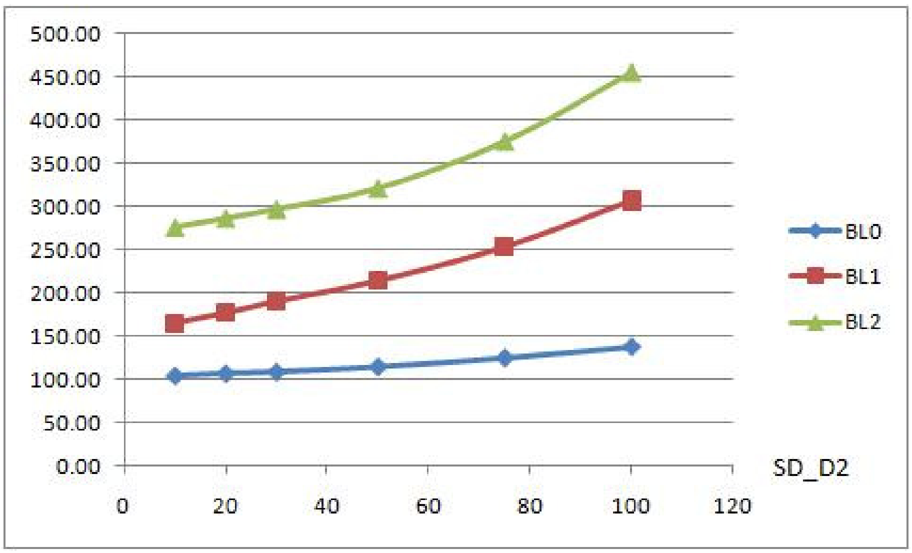
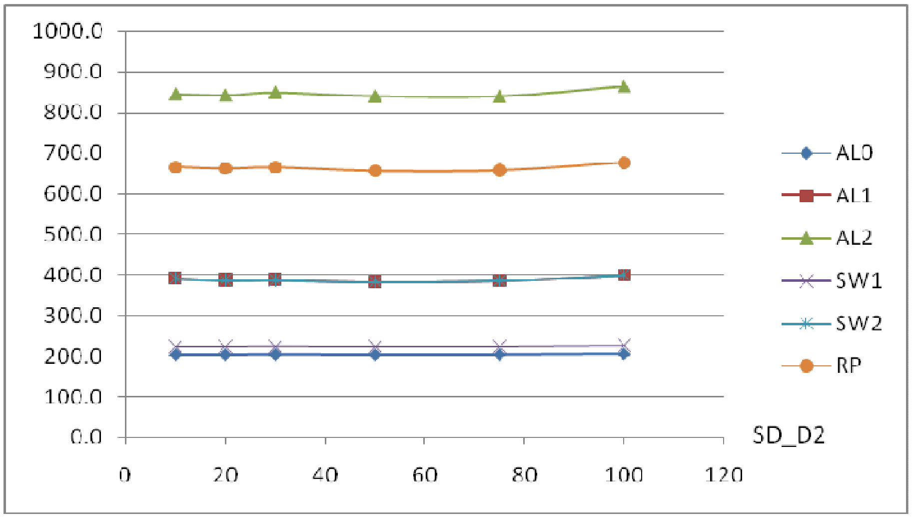


Figure 6.10 The trend of the threshold levels in different scenarios of demand for product 2 with parameter set 1

Table 6.13 The threshold levels in different scenarios of demand for product 2 with parameter set 2

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
COV_2=0.1	205.1	391.3	845.0	224.5	391.4	666.1	102.6	164.6	262.9
COV_2=0.2	204.8	386.9	842.2	224.4	387.3	662.6	104.7	176.6	274.1
COV_2=0.3	205.4	387.7	848.8	224.8	387.7	665.2	106.3	188.7	287.6
COV_2=0.5	204.7	383.2	840.1	224.0	383.7	657.5	110.8	207.4	305.2
COV_2=0.75	205.0	385.9	840.0	224.3	386.9	658.4	114.5	224.8	323.1
COV_2=1.0	206.3	399.1	864.5	226.2	399.6	676.7	114.0	236.3	332.8



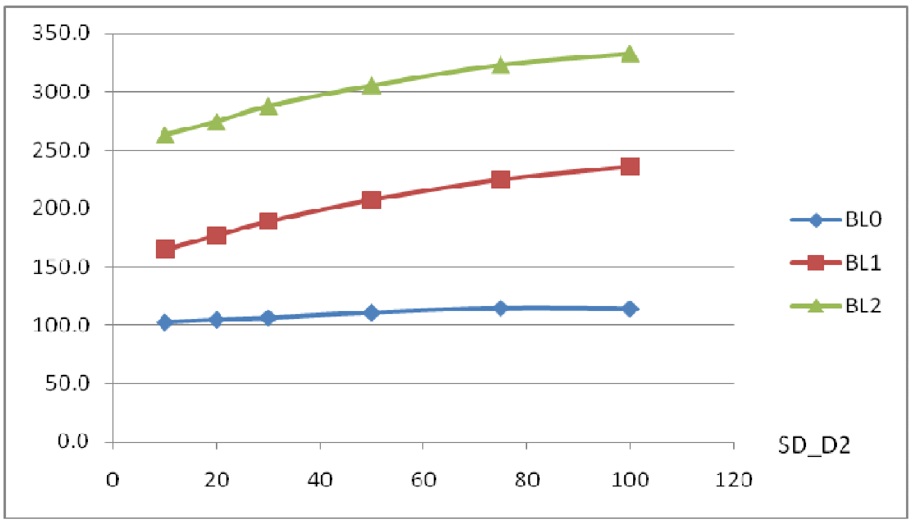


Figure 6.11 The trend of the threshold levels in different scenarios of demand for product 2 with parameter set 2

Table 6.14 The threshold levels in different scenarios of demand for product 2 with parameter set 3

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
COV_2=0.1	205.2	391.6	1055.2	224.4	605.2	877.3	102.2	164.7	254.4
COV_2=0.2	205.0	387.6	1051.3	224.3	600.6	873.8	104.3	175.4	266.1
COV_2=0.3	204.9	385.6	1050.3	224.1	598.2	872.7	105.8	187.7	280.6
COV_2=0.5	204.7	383.6	1047.9	224.0	592.7	869.0	110.1	206.9	302.8
COV_2=0.75	205.1	387.0	1047.9	224.5	592.5	868.0	113.9	224.0	321.5
COV_2=1.0	205.7	397.3	1066.9	225.5	602.9	881.3	113.6	235.3	330.1

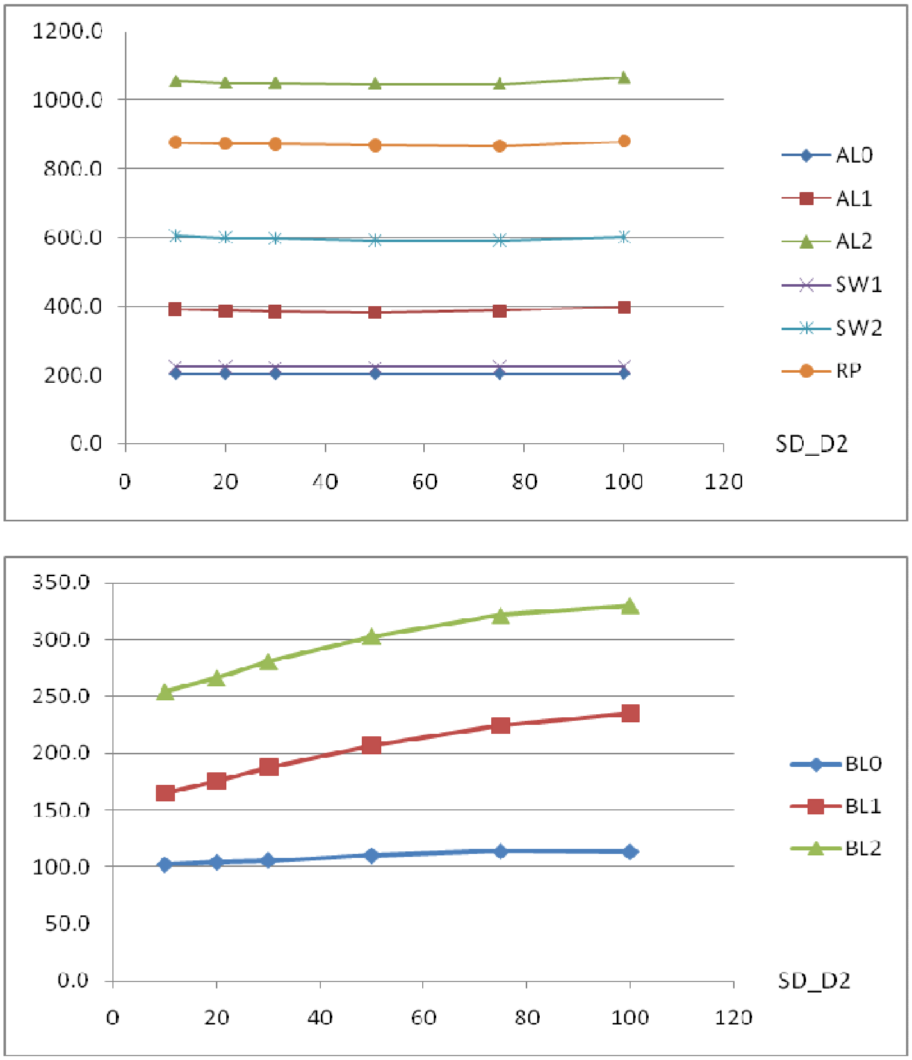


Figure 6.12 The trend of the threshold levels in different scenarios of demand for product 2 with parameter set 3

6.4.2 The comparison of three heuristic policies with respect to the expected average cost

While using the threshold levels to help to make production and recovery decisions in a relatively long horizon, the resulting expected average cost is compared with those values obtained by using two heuristic policies from the single-period problem. The following symbols will be used in the presentation of numerical results:

H1 – Heuristic policy from the single-period problem disregarding scrap values of the remaining finished products;

H2 – Heuristic policy from the single-period problem assuming scrap value of the remaining product 1 and product 2 to be equal to c_{R21} and c_{R22} respectively;

H3 – Heuristic policy from solving the ADP model;

EAC_H1 – Expected average cost calculated while the heuristic policy ***H1*** is used in a relatively long horizon;

EAC_H2 – Expected average cost calculated while the heuristic policy ***H2*** is used in a relatively long horizon;

EAC_H3 – Expected average cost calculated while the heuristic policy ***H3*** is used in a relatively long horizon.

The optimal policy of the single-period problem is used as heuristic policy for the multi-period problem. Two heuristic policies, denoted as *H1* and *H2* respectively, are derived from solving the single-period problem. The policy *H1* disregards the scrap values of the remaining finished products whereas the policy *H2* assumes the scrap value of product 1 and product 2 to be equal to c_{R21} and c_{R22} respectively. In order to compare the threshold levels of the policies *H1* and *H2* with those of the policy *H3* from solving the approximate dynamic programming model, we have selected the set of threshold levels when $E[R_1] = 210$ in Table 6.3. The threshold levels of the three heuristic policies have been shown in Table 6.15 and further compared in Figure 6.13. The results have shown that the corresponding threshold levels of the policy *H3* are highest whereas the threshold levels of the policy *H1* are lowest. The difference of each corresponding threshold level between the policies *H1* and *H2* is small whereas the difference between the policies *H1* and *H3* is obviously large.

Table 6.15 The threshold levels in three heuristic policies

	AL_0	AL_1	AL_2	SW_1	SW_2	RP	BL_0	BL_1	BL_2
H1	151.7	204.0	238.0	162.0	171.2	212.0	79.8	96.9	106.3
H2	156.7	213.4	255.7	167.5	177.4	222.6	88.7	111.3	127.9
H3	182.6	316.9	526.4	200.7	219.3	349.7	105.6	165.9	248.3

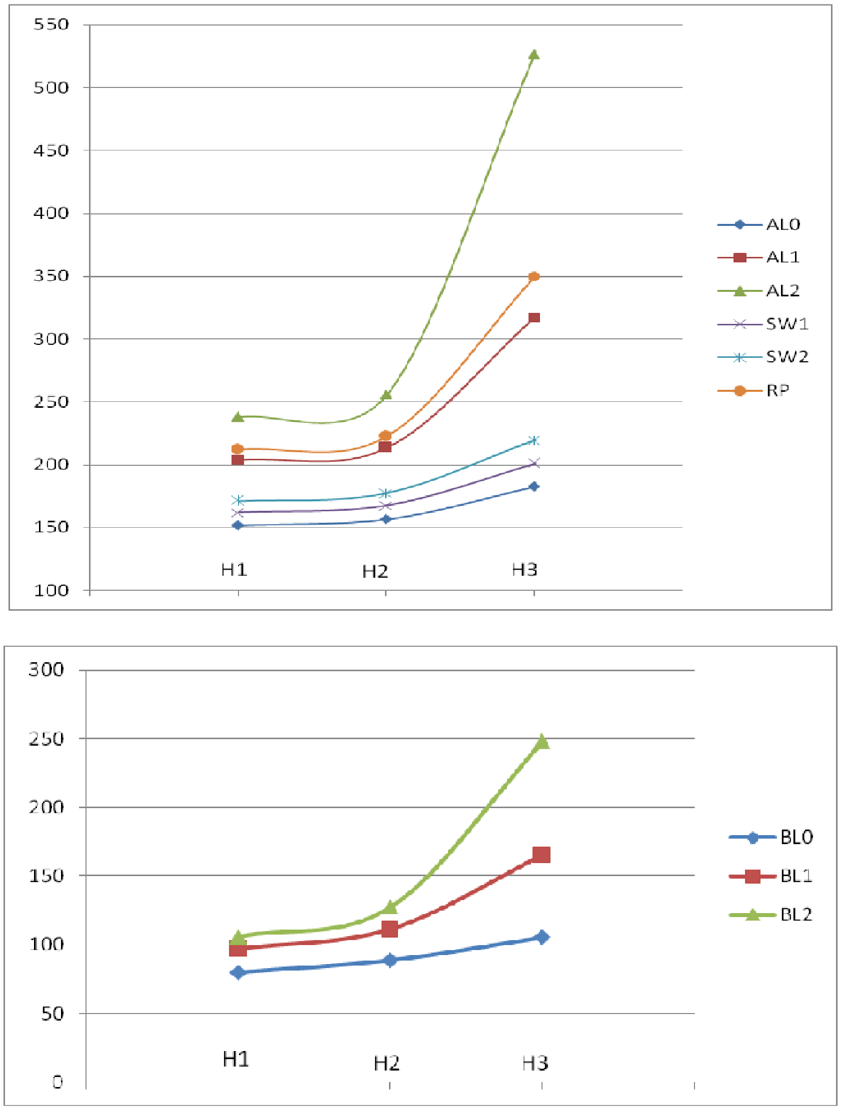


Figure 6.13 The comparison of the threshold levels in different heuristic policies

While using the three heuristic policies to make production and recovery decisions of the two-product recovery system in a relatively long horizon, the resulting expected average costs are shown in Table 6.16. The results have shown that the policy *H3* performs best, secondly the policy *H2* and finally the policy *H1*. By comparing the expected average costs between the policies *H1* and *H3*, 3.3% decrement can be achieved while using the policy *H3* to replace the policy *H1*.

Table 6.16 The expected average costs using different heuristic policies

EAC_H1	2738.7
EAC_H2	2698.1
EAC_H3	2649.0

6.5 Summary

In this Chapter, we have developed the ADP model of the two-product recovery system in the situation of backorder over a finite horizon. The model aims to determine the threshold levels as the multi-level threshold policy from the single-period problem is assumed to be used for the multi-period problem. In the multi-period situation, the threshold levels are found to be only dependent on the gradient of the cost-to-go function at the points of interest.

The impact of system parameters on the threshold levels has been investigated. The numerical results have shown that the more returned items from either group in each period would make the threshold levels lower. Among them, the threshold levels AL_1, AL_2, BL_1, BL_2 and RP, related to recovery processes, would obviously decrease with returned items increasing. However, there are small decreases on the threshold levels AL_0 and BL_0, related to production processes, and the threshold levels SW_1 and SW_2, related to switching the allocation of returned items to the recovery processes between the two products. On the other hand, with the increasing demand variability of a certain product, the threshold levels related to this product would increase at the same time whereas the threshold levels related to the other product seem unaffected.

After determining the threshold levels, we can use the threshold policy to control the two-product recovery system in the multi-period context. The performance of this policy is compared with the two heuristic policies derived from the optimal policy of the single-period problem. Through the comparison of the resulting expected average cost, the policy from solving the approximate dynamic programming model outperforms the other two heuristic policies.

Chapter 7 The study on two-product recovery system in a finite horizon with backorder and nonzero constant lead time

Chapter 7 focuses on the two-product recovery system in a finite horizon, in which backorder is allowed. In the recovery system, all the lead times of production and recovery processes are assumed to be the same nonzero constant. Section 7.1 introduces the recovery system. In Section 7.2, an ADP model of the recovery system is developed in order to minimize the expected total cost in a finite horizon. In the model, the lead time effect is considered. The model aims to derive the threshold levels, which are only dependent on the gradient of the cost-to-go function at the points of interest. Section 7.3 provides the details about how to determine the gradient at the points of interest. Section 7.4 gives the computational results about the performance of the policy from solving the ADP model. Finally, Section 7.5 summarizes the main findings.

7.1 Introduction

Chapter 6 has studied the two-product recovery system in a finite horizon, in which backorder is allowed. In addition, production and recovery processes are assumed to have zero lead time. However, these processes often have nonzero lead time in practice. Therefore, this Chapter investigates the lead time effect of production and recovery processes. Hereafter, all the lead times of production and recovery processes are assumed to be the same nonzero constant. Due to the existence of lead times, the initial inventory position at the beginning of each period in the planning horizon needs to include pipeline inventory.

The objective of modeling the recovery system is to minimize the expected total cost in a finite horizon. In order to fulfill the aim, we need to obtain the optimal policy, which helps to make the optimal production and recovery decisions in each period of the planning horizon.

7.2 Approximate dynamic programming model of the two-product recovery system in the multi-period context

Since the related assumptions and notations for the two-product recovery system can be referred to in Chapter 3 and Chapter 6, we will not repeat here. The only different notation is as follows:

L lead time of production and recovery processes for each product.

The inventory position at the beginning of period t is net stock plus pipeline inventory. The inventory state transition equations between two consecutive periods can be written as follows (j = 1, 2):

$$x_{Sj}^{(t+1)} = x_{Sj}^{(t)} + p_j^{(t)} + r_{1j}^{(t)} + r_{2j}^{(t)} - D_j^{(t)}. \tag{7.1}$$

Due to lead times existing in the system, production and recovery decisions made in period t will affect the joint inventory holding cost and penalty cost of shortage in period $t + L$. Thus, we take into account production and recovery costs of period t, together with the joint inventory holding cost and penalty cost of shortage in period $t + L$. The expected cost in period t is calculated as follows:

$$\begin{aligned} & EC_t(x_{S1}^{(t)}, x_{S2}^{(t)}, p_1^{(t)}, p_2^{(t)}, r_{11}^{(t)}, r_{12}^{(t)}, r_{21}^{(t)}, r_{22}^{(t)}) \\ & = c_{P1}p_1^{(t)} + c_{P2}p_2^{(t)} + c_{R11}r_{11}^{(t)} + c_{R12}r_{12}^{(t)} + c_{R21}r_{21}^{(t)} + c_{R22}r_{22}^{(t)} \\ & + h_1\int_0^{x_{S1}^{(t)}+p_1^{(t)}+r_{11}^{(t)}+r_{21}^{(t)}} (x_{S1}^{(t)} + p_1^{(t)} + r_{11}^{(t)} + r_{21}^{(t)} - D_1^{'})f(D_1^{'}, \mu_1^{'}, \sigma_1^{'})dD_1^{'} \\ & + v_1\int_{x_{S1}^{(t)}+p_1^{(t)}+r_{11}^{(t)}+r_{21}^{(t)}}^{\infty} (D_1^{'} - x_{S1}^{(t)} - p_1^{(t)} - r_{11}^{(t)} - r_{21}^{(t)})f(D_1^{'}, \mu_1^{'}, \sigma_1^{'})dD_1^{'} \\ & + h_2\int_0^{x_{S2}^{(t)}+p_2^{(t)}+r_{12}^{(t)}+r_{22}^{(t)}} (x_{S2}^{(t)} + p_2^{(t)} + r_{12}^{(t)} + r_{22}^{(t)} - D_2^{'})f(D_2^{'}, \mu_2^{'}, \sigma_2^{'})dD_2^{'} \\ & + v_2\int_{x_{S2}^{(t)}+p_2^{(t)}+r_{12}^{(t)}+r_{22}^{(t)}}^{\infty} (D_2^{'} - x_{S2}^{(t)} - p_2^{(t)} - r_{12}^{(t)} - r_{22}^{(t)})f(D_2^{'}, \mu_2^{'}, \sigma_2^{'})dD_2^{'}. \end{aligned} \tag{7.2}$$

In Formula (7.2), the transformed demand $D_j^{'}$ is the aggregation of demands from period t till period $t + L$. The related characteristic parameters of the transformed demand are calculated as follows (j = 1, 2):

$$\mu_j^{'} = (L+1)\mu_j; \quad \sigma_j^{'} = \sqrt{L+1}\sigma_j. \tag{7.3}$$

Let $f_t(x_{S1}^{(t)}, x_{S2}^{(t)})$ denote the expected total cost from period t till period M - L. Assume $f_{M-L+1}(x_{S1}^{(M-L+1)}, x_{S2}^{(M-L+1)}) = \sum_{j=1}^{2} c_{Pj}[-x_{Sj}^{(M-L+1)}]^{+}$. The Bellman's equation of dynamic programming can be written as follows (t = 1, 2,..., M - L):

$$\begin{aligned} f_t(x_{S1}^{(t)}, x_{S2}^{(t)}) = \underset{R_1^{(t-1)}, R_2^{(t-1)}}{E}\{ & \min_{p_1^{(t)}, p_2^{(t)}, r_{11}^{(t)}, r_{12}^{(t)}, r_{21}^{(t)}, r_{22}^{(t)}} \{EC_t(x_{S1}^{(t)}, x_{S2}^{(t)}, p_1^{(t)}, p_2^{(t)}, r_{11}^{(t)}, r_{12}^{(t)}, r_{21}^{(t)}, r_{22}^{(t)}) \\ & + \underset{D_1^{(t)}, D_2^{(t)}}{E}[f_{t+1}(x_{S1}^{(t+1)}, x_{S2}^{(t+1)})]\}\}. \end{aligned} \tag{7.4}$$

Similar to Chapter 6, the approximation is made to the cost-to-go function in the above formula. The approximate dynamic programming model considering the lead time effect is the same as that in Chapter 6 except that the expected cost of period t is calculated on the basis of the above-mentioned transformed demands. Therefore, similar threshold level policy can be obtained by solving the approximate dynamic programming model. Thus, the threshold levels of the policy are determined on the basis of the transformed demands. In addition, similar to Chapter 6, two gradients used for the approximation are estimated at the point of interest by Monte Carlo simulation.

7.3 The determination of the gradient at the points of interest in the multi-period context

Without a closed-form formula of the gradient $u_j^{(t)}$ at the point of interest ($x_1^{(t)}$, $x_2^{(t)}$), we need to run Monte Carlo simulation, and estimate the gradient based on the simulation results. Before that, we need to approximate the cost-to-go function of dynamic programming by Monte Carlo formulation. In Monte Carlo sampling, sample k is about the realization of stochastic returns in each period from period t + 1 till period M - L, and the realization of stochastic demands in each period from period t till period M - 1. The sample k is expressed as:

$$\begin{pmatrix} R_{1,k}^{(t+1)} & R_{2,k}^{(t+1)} & D_{1,k}^{(t)} & D_{2,k}^{(t)} \\ \vdots & \vdots & \vdots & \vdots \\ R_{1,k}^{(M-L)} & R_{2,k}^{(M-L)} & D_{1,k}^{(M-L-1)} & D_{2,k}^{(M-L-1)} \\ 0 & 0 & D_{1,k}^{(M-L)} & D_{2,k}^{(M-L)} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & D_{1,k}^{(M-1)} & D_{2,k}^{(M-1)} \end{pmatrix}.$$

The cost-to-go function of dynamic programming is approximated as follows:

$$\underset{D_1^{(t)},D_2^{(t)}}{E}[f_{t+1}(x_{S1}^{(t+1)},x_{S2}^{(t+1)})] \approx \frac{1}{N}\sum_{k=1}^{N}(\sum_{\tau=t+1}^{M-L-1}ATC_k^{(\tau)*}+MEC_k). \quad (7.5)$$

In the above formula, the sample value for sample k, is obtained by summing up the cost in each period from period t + 1 till period M – L - 1 and the minimum expected cost in period M - L. Without a closed-form formula of the function MEC_k, we would compute it by minimizing the expected cost $EC_{M-L,\ k}$. In addition, the function $ATC_k^{(\tau)*}$ is to calculate the cost of period τ at the optimum. The function $ATC_k^{(\tau)}$ is expressed as follows ($[X]^+ := \max\{X, 0\}$; $t < \tau < M - L$):

$$\begin{aligned}&ATC_k^{(\tau)}(x_{S1}^{(\tau)},x_{S2}^{(\tau)},p_1^{(\tau)},p_2^{(\tau)},r_{11}^{(\tau)},r_{12}^{(\tau)},r_{21}^{(\tau)},r_{22}^{(\tau)},D_{1,k}^{(i),i=\tau,\ldots,\tau+L},D_{2,k}^{(i),i=\tau,\ldots,\tau+L})\\&=c_{P1}p_1^{(\tau)}+c_{P2}p_2^{(\tau)}+c_{R11}r_{11}+c_{R12}r_{12}+c_{R21}r_{21}+c_{R22}r_{22}\\&+h_1[x_{S1}^{(\tau)}+p_1^{(\tau)}+r_{11}^{(\tau)}+r_{21}^{(\tau)}-\sum_{i=\tau}^{\tau+L}D_{1,k}^{(i)}]^+ +v_1[\sum_{i=\tau}^{\tau+L}D_{1,k}^{(i)}-x_{S1}^{(\tau)}-p_1^{(\tau)}-r_{11}^{(\tau)}-r_{21}^{(\tau)}]^+\\&+h_2[x_{S2}^{(\tau)}+p_2^{(\tau)}+r_{12}^{(\tau)}+r_{22}^{(\tau)}-\sum_{i=\tau}^{\tau+L}D_{2,k}^{(i)}]^+ +v_2[\sum_{i=\tau}^{\tau+L}D_{2,k}^{(i)}-x_{S2}^{(\tau)}-p_2^{(\tau)}-r_{12}^{(\tau)}-r_{22}^{(\tau)}]^+.\end{aligned} \quad (7.6)$$

At period M – L, we can calculate the minimum expected cost MEC_k by minimizing the expected cost $EC_{\text{M-L, k}}$, which is calculated by Formula (7.2).

As the function $EC_{\text{M-L, k}}$ and the function $ATC_k^{(\tau)}$ are both continuous functions, it is suitable to approximate the cost-to-go function by Monte Carlo sampling method. Furthermore, the two gradients used for the approximate dynamic programming model can be approximated by sample average.

In the following, we present the computational results about applying the policy from solving the approximate dynamic programming model to the two-product recovery system in a finite horizon. At the same time, this policy is compared with the other two heuristic policies derived from the optimal policy of the single-period problem.

7.4 Computational results

The following symbols will be used in the presentation of numerical results:

H1 – Heuristic policy from the single-period problem disregarding scrap values of the remaining finished products;

H2 – Heuristic policy from the single-period problem assuming scrap value of the remaining product 1 and product 2 to be equal to c_{R21} and c_{R22} respectively;

H3 – Heuristic policy from solving the ADP model;

EAC_H1 – Expected average cost calculated while the heuristic policy ***H1*** is used in a relatively long horizon;

EAC_H2 – Expected average cost calculated while the heuristic policy ***H2*** is used in a relatively long horizon;

EAC_H3 – Expected average cost calculated while the heuristic policy ***H3*** is used in a relatively long horizon.

A set of system parameters is given as follows:

Cost: $h_1 = 3, h_2 = 3, v_1 = 4, v_2 = 6, s_1 = 15, s_2 = 20,$
$c_{P1} = 12, c_{P2} = 15, c_{R11} = 6, c_{R12} = 10, c_{R21} = 2, c_{R22} = 7;$

Demand: $E[D_1] = 200, StDev[D_1] = 60; E[D_2] = 100, StDev[D_2] = 30;$

Return : $E[R_1] = 90, StDev[R_1] = 30; E[R_2] = 45, StDev[R_2] = 15.$

In Table 7.1, we have shown the threshold levels of the policy from solving the approximate dynamic programming model and the two heuristic policies from the single-period problem under different values of fixed lead time L (L = 0, 1, 2).

Table 7.1 The threshold levels in different heuristic policies (L=0, 1, 2)

		AL_1	AL_2	AL_3	SW_1	SW_2	RP	BL_1	BL_2	BL_3
L=0	**H1**	151.7	204.0	238.0	162.0	171.2	212.0	79.8	96.9	106.3
	H2	156.7	213.4	255.7	167.5	177.4	222.6	88.7	111.3	127.9
	H3	205.4	390.3	628.1	224.7	245.6	440.2	108.4	190.2	297.5
L=1	**H1**	331.7	405.6	453.8	346.2	359.3	416.9	171.4	195.6	208.9
	H2	338.8	418.9	478.8	354.1	368.0	432.0	184.0	216.0	239.4
	H3	409.0	627.1	869.5	435.9	463.9	677.6	207.6	299.8	407.5
L=2	**H1**	516.4	606.9	665.9	534.2	550.2	620.7	265.0	294.6	310.9
	H2	525.0	623.2	696.5	543.7	560.8	639.2	280.4	319.6	348.3
	H3	612.2	847.2	1089.1	644.1	676.6	897.7	310.3	410.9	519.4

Furthermore, the expected average costs are calculated and shown in Table 7.2 while using the above three policies to control the two-product recovery system in a relatively long horizon. The percentage of increment is calculated on the basis of the expected average cost by using the policy *H1*. Figure 7.1 has shown the trend of the expected average cost with the lead time. With the larger value of the lead time, the expected average cost is higher. The average inventory level per period increases with the lead time so as to reduce possible stock shortage. Therefore, the expected average cost will increase with the lead time as more inventory holding cost is incurred.

Table 7.2 The expected average cost using different heuristic policies (L=0, 1, 2)

	L =0	Increment(%)	L=1	Increment(%)	L=2	Increment(%)
EAC_H1	3336.9		3409.5		3668.9	
EAC_H2	3297.9	*-1.2*	3357	*-1.5*	3598.8	*-1.9*
EAC_H3	3208.7	*-3.8*	3247.5	*-4.8*	3445.2	*-6.1*

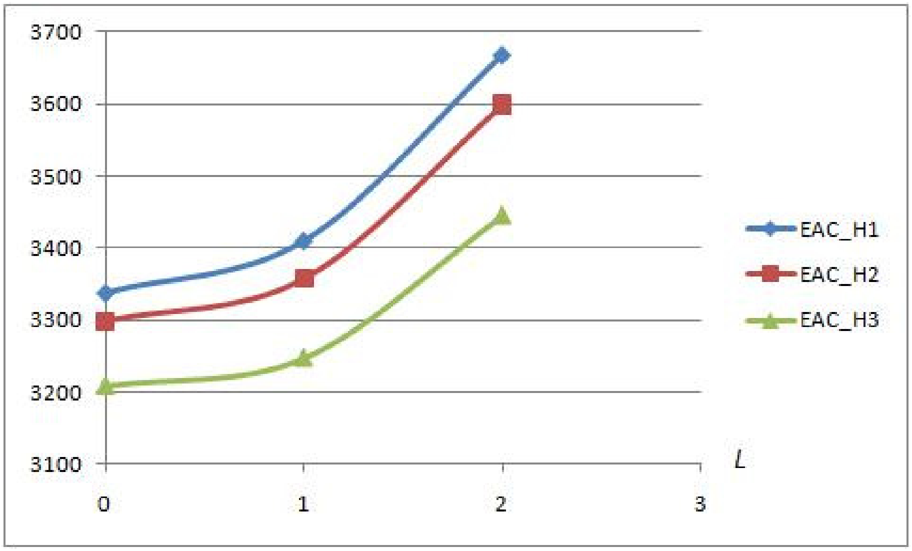


Figure 7.1 The expected average cost using different heuristic policies (L=0, 1, 2)

7.5 Summary

In this Chapter, we have studied the two-product recovery system, in which backorder is allowed. For the system, the lead time effect has been investigated by assuming all the lead times of production and recovery processes to be the same nonzero constant. We have developed the ADP model of the system in order to minimize the expected total cost in the finite horizon. The model is used to derive the threshold levels as the multi-level threshold policy from the single-period problem is assumed to be used for the multi-period problem. In the multi-period situation, the threshold levels are found to be only dependent on the gradient of the cost-to-go function at the points of interest.

The computational results have shown that the policy from solving the approximate dynamic programming model outperforms the other two heuristic policies from the single-period problem. Between the two heuristic policies, the heuristic policy, which considers the scrap values of the remaining finished products, performs better. In addition, the expected average cost increases with the lead time as the average inventory of the system increases with the lead time.

Chapter 8 Conclusion

The main purpose of this thesis is to develop mathematical models on the two-product recovery system in a finite horizon in order to obtain the optimal or near-optimal policy for production planning and inventory control. This chapter concludes the study by presenting a summary of research findings and discussing the implications and limitations of this research, as well as suggesting several directions for future research.

8.1 Main findings

In Chapter 3, we have developed a dynamic programming model for the two-product recovery system in a finite horizon. The aim is to maximize the expected total profit in a finite horizon. However, the dynamic programming model is found to be difficult to be solved efficiently due to no nice property. Therefore, we have studied the single-period problem as the special case of the multi-period problem in Chapter 4. After modeling and solving the single-period problem, an optimal multi-level threshold policy is obtained. The related threshold levels are discovered and their insights are further explained.

Even though this multi-level threshold policy might not be optimal for the multi-period problem, it is intuitive, easy to use and provides good managerial perspectives. Hence, we apply this policy to the multi-period problem

In Chapter 5, we have proposed an ADP model to derive the threshold levels. We have found that different from the single-period problem, the threshold will not only depend on the current-period cost parameters, but also on the future cost-to-go function. The threshold levels are further found to be only dependent on the gradient of the cost-to-go function at the points of interest. Unlike the usual approach which uses a single function (or piecewise function) to represent the cost-to-go function, we just need to estimate the gradient of the cost-to-go function at the points of interest. These gradients will be used to compute the threshold level. As the threshold level and the gradient are dependent on each other, we have determined the threshold levels via an iterative algorithm. When estimating the gradient by a Monte Carlo simulation-based technique, i.e. Sample Average Approximation (SAA), we develop an Infinitesimal Perturbation Analysis (IPA) based approach to determine the sample gradient. This approach not only uses the least computing resources but also its estimation quality is better.

The threshold policy from solving the ADP model is compared with the two heuristic policies, which are derived from the optimal policy of the single-period problem. One heuristic policy assumes the scrap values of the two products to be nonzero fixed values whereas the other heuristic policy assumes the scrap values of the two products to be zero. By the comparison of the resulting expected average profits, we find that the policy from solving the ADP model performs best, followed by the heuristic policy considering the scrap value of finished products, and finally the heuristic policy disregarding the scrap value. Furthermore, with the best policy, the impact of system parameters has been investigated. The computational results have shown that the larger expected value of returned items in either group brings more expected average profit. In addition, the higher the demand variability, the less the expected average profit.

In addition, Chapter 6 and Chapter 7 focus on the two-product recovery system in the situation of backorder over a finite horizon. The model aims to minimize the expected total cost over the finite horizon. Chapter 6 has done similar work as Chapter 5 to investigate the performance of the threshold policy and the impact of system parameters under different scenarios. Chapter 7 investigates the lead time effect of production and recovery processes. By assuming all the lead times of production and recovery processes to be the same nonzero constant, the expected average cost of the system is found to be increasing with the value of the constant lead time. This results from the aggregation of demand variability.

8.2 Discussion about the relaxation of certain assumptions

We would discuss about the relaxation of certain assumptions mentioned in Chapter 3.

In Chapter 3, lead time is assumed to be equal to zero. In Chapter 7, the lead time effect has been considered in the situation of backorder based on a simple case that both production lead time and recovery lead time of each product are equal to the same nonzero constant. The threshold level is computed based on the gradient of the cost-to-go function which considers the constant lead time. If the simple case would be extended to a more complex case that production lead time and recovery lead time are different, the state space of dynamic programming will be increased due to the lead time difference. If we still assume the multi-level threshold policy to be used for this case, the threshold levels need to be computed based on the cost-to-go function which considers not only the lead time effect but also the lead time difference between production and recovery. This could be further studied as one of the future directions.

In Chapter 3, disposal cost is assumed to be negligible. Otherwise, disposal cost needs to be included in the total cost. For the single-period problem, the threshold policy remains unchanged expect that some threshold levels need to be recomputed. For example, the order-up-to level of product 1 by recovering the returns in group 1, is calculated as $AL_1 = F^{-1}(\frac{s_1 + v_1 - c_{R11}}{s_1 + v_1 + h_1}, \mu_1, \sigma_1)$. If disposal cost is considered and its cost rate is assumed as c_{D1}, the order-up-to level is recomputed as $AL_1 = F^{-1}(\frac{s_1 + v_1 - (c_{R11} - c_{D1})}{s_1 + v_1 + h_1}, \mu_1, \sigma_1)$. For the multi-period problem studied over a long horizon, as the threshold policy is evaluated by measuring the expected average cost, the disposal cost would be regarded as negligible and need not be considered.

In Chapter 3, it is assumed that there is no stocking of the returned products. This assumption is reasonable in some practical situations. For example, it might be cost-saving without establishing extra storage capacity for returned items. The same situation occurs if these returned items cannot be stored over a longer period because of environmental or similar reasons. Otherwise, if this assumption is relaxed to allow the stocking of returned items for future periods, disposal of unused returned items is optional and depends on the inventory states of both returned items and finished items. Thus, the stock holding cost of two groups of returned items will be considered in the modeling and the threshold level for the disposal of returned items might be necessary to characterize the multi-level threshold policy. Furthermore, we need to investigate how to compute the threshold levels for both single-period case and multi-period case in the situations of having stocks of returned items. It could be further studied as one of the future directions.

8.3 Suggestions for Future Work

Demand substitution

One-way (downward) substitution often exists in practice, especially in high-tech industry. Inderfurth (2004) and Bayindir et al. (2007) considered one-way substitution of the finished product from production for that from recovery in the single-product recovery system. The one-way demand substitution will reduce shortages, and also incur additional substitution cost. If this one-way demand substitution is allowed in the two-product recovery system of this study, the optimal production and recovery decisions need to be re-considered.

Capacitated production

In the two-product recovery system of this study, production capacity is assumed to be unlimited. Production process will be used for the replenishment of finished item inventory if the recovery of returned items is not enough to achieve replenishment requirement. However, if production process is capacitated, how to determine the related threshold levels needs to be re-considered. In most existing models of this field, capacitated production is considered together with demand substitution. For example, Li et al. (2007) considered the capacitated production planning problem in the single-product recovery system.

Different lead times between production and recovery

In the two-product recovery system of this study, if production lead time and recovery lead time are different, the state space of dynamic programming will be increased due to the lead time difference. If we still assume the multi-level threshold policy to be used for this case, how to compute the threshold levels based on the cost-to-go function in the more complex case is one of the future directions.

Stocking of returned items

In the two-product recovery system of this study, if there is the stocking of returned items, disposal of unused returned items will depend on the inventory states of both returned items and finished items. The threshold level for the disposal of returned items needs to be determined for the multi-level threshold policy for both single-period case and multi-period case. It could be further studied as one of the future directions.

Approximate dynamic programming model with neural network

In the approximate dynamic programming model of this study, we have taken advantage of simple linear models to fulfill the approximation. On the other hand, neural networks represent a powerful and general class of approximation strategies used in approximate dynamic programming. By means of neural networks, a much richer class of nonlinear functions can be trained in an iterative way, which is matching the needs of approximate dynamic programming. If neural network approximation would be used for the approximate dynamic programming model of this research, we need to take advantage of the problem structure. The advantage of this approximation method should be shown by the comparison with linear approximation.

References

Ahiska, S.S., King, R.E., 2010. Inventory optimization in a one product recoverable manufacturing system". International Journal of Production Economics 124, 11-19.

Bayindir, Z.P., Erkip, N., Güllü, R., 2007. Assessing the benefits of remanufacturing option under one-way substitution and capacity constraint. Computers & Operations Research 34, 487-514.

Behret, H., Korugan, A., 2009. Performance analysis of a hybrid system under quality impact of returns. Computers and Industrial Engineering 56, 507-520.

Beltran, J.L., Krass, D., 2002. Dynamic lot sizing with returning items and disposals. IIE Transactions 34, 437-448.

Brayman, R.B., 1992. How to implement MRP II successfully the second time: getting people involved in a remanufacturing environment. APICS Remanufacturing Seminar Proceedings, September, 82-88.

Buchanan, D.J., Abad, P.L., 1998. Optimal policy for a periodic review returnable inventory system. IIE Transactions 30, 1049-1055.

Cohen, M.A., Nahmias, S., Pierskalla, W.P., 1980. A dynamic inventory system with recycling. Naval Research Logistics Quarterly 27, 289-296.

De Brito, M., 2003. Managing reverse logistics or reversing logistics management. Ph.D. Thesis, Erasmus University Rotterdam, The Netherlands.

De Brito, M., Dekker, R., 2003. Modeling products return in inventory control: Exploring the validity of general assumptions. International Journal of Production Economics 81-82, 225-241.

De Brito, M., Dekker, R., 2004. A framework for reverse-logistics. In Reverse Logistics: Quantitative Models in Closed-Loop Supply Chains, ed by Dekker, R., Inderfurth, K., Van Wassenhove, L.N., Fleischmann, M., 3-27. Berlin: Springer-Verlag.

DeCroix, G., Jing-Sheng, S., Zipkin, P., 2005. A series system with returns: stationary analysis. Operations Research 53, 350–362.

DeCroix, G.A., 2006. Optimal policy for a multiechelon inventory system with remanufacturing. Operations Research 54, 532–543.

Dekker, R., Inderfurth, K., Van Wassenhove, L.N., Fleischmann, M. (ed), 2004. Reverse Logistics: Quantitative Models in Closed-Loop Supply Chains. Berlin: Springer-Verlag.

Dowlatshahi, S., 2005. A strategic framework for the design and implementation of remanufacturing operations in reverse-logistics. International Journal of Production Research 43, 3455-3480.

Drake, M.J., Ferguson M., 2008. Closed-loop supply chain management for global sustainability. In Global Sustainability Initiatives: New Models and New Approaches, ed by Stoner, J., Wankel, C., 173-191. Information Age Publishing.

Ferguson, M., 2010. Strategic and tactical aspects of closed-loop supply chains. Foundations and Trends in Technology, Information and Operations Management 3, 101-199.

Ferguson, M., Souza, G., 2010. Closed-Loop Supply Chains: New Developments to Improve the Sustainability of Business Practices. London: Taylor and Francis Publishing.

Fleischmann, M., Bloemhof-Ruwaard, J.M., Dekker, R., Van der Laan, E., Van Nunen, J.A.E.E., Van Wassenhove, L.N., 1997. Quantitative models for reverse logistics: A review. European Journal of Operational Research 103, 1-17.

Fleischmann, M., 2001. Quantitative models for reverse logistics. Springer.

Fleischmann, M., Kuik, R., Dekker, R., 2002. Controlling inventories with stochastic item returns: a basic model. European Journal of Operational Research 138, 63-75.

Fleischmann, M., Minner, S., 2004. Inventory management in closed loop supply chains. In Supply Chain Management and Reverse-Logistics, ed by Dyckhoff, H., Lackes, R., Reese, J., 115-138. Berlin: Springer.

Gupal, A.M., Bazhenov, L.G., 1972. A stochastic method of linearization. Cybernetics 8, 482-484.

Heyman, D.P., 1977. Optimal disposal policies for a single-item inventory system with returns. Naval Research Logistics Quarterly 24, 385-405.

Hillier, F.S., Lieberman, G.J., 1990. Introduction to Operations Research, fifth ed. New York: McGraw-Hill.

Ilgin, M.A., Gupta, S.M., 2010. Environmentally conscious manufacturing and product recovery (ECMPRO): A Review of the state of the art. Journal of Environmental Management 91, 563-591.

Inderfurth, K., 1997. Simple optimal replenishment and disposal policies for a product recovery system with leadtimes. OR Spektrum 19, 111-122.

Inderfurth, K., De Kok, A.G., Flapper, S.D.P., 2001. Product recovery in stochastic remanufacturing systems with multiple reuse options. European Journal of Operational Research 133, 130-152.

Inderfurth, K., Van der Laan, E.A., 2001. Leadtime effects and policy improvement for stochastic inventory control with remanufacturing. International Journal of Production Economics 71, 381-390.

Inderfurth, K., 2004. Optimal policies in hybrid manufacturing/remanufacturing systems with product substitution. International Journal of Production Economics 90, 325-343.

Kandebo, S.W., 1990. Grumman, U.S. Navy under way in F-14 remanufacturing program. Aviation Week & Space Technology, December, 44-45.

Kelle, P., Silver, E.A., 1989. Purchasing policy of new containers considering the random returns of previously issued containers. IIE Transactions 21, 349-354.

Kiesmüller, G.P., Minner, S., Kleber, R., 2000. Optimal control of a one product recovery system with backlogging. IMA Journal of Mathematics Applied in Business and Industry 11, 189-207.

Kiesmüller, G.P., Van der Laan, E., 2001. An inventory model with dependent product demands and returns. International Journal of Production Economics 72, 73-87.

Kiesmüller, G.P., 2003a. A new approach for controlling a hybrid stochastic manufacturing/remanufacturing system with inventories and different leadtimes. European Journal of Operational Research 147, 62-71.

Kiesmüller, G.P., 2003b. Optimal control of a one product recovery system with leadtimes. International Journal of Production Economics 81-82, 333-340.

Kiesmüller, G.P., Scherer, C.W., 2003. Computational issues in a stochastic finite horizon one product recovery inventory model. European Journal of Operational Research 146, 553-579.

Kleber, R., Minner, S., Kiesmüller, G.P., 2002. A continuous time inventory model for a product recovery system with multiple options. International Journal of Production Economics 79, 121-141.

Koh, S.G., Hwang, H., Sohn, K.I., Ko, C.S., 2002. An optimal ordering and recovery policy for reusable items. Computers and Industrial Engineering 43, 59-73.

Konstantaras, I., Papachristos, S., 2007. Optimal policy and holding cost stability regions in a periodic review inventory system with manufacturing and remanufacturing options. European Journal of Operational Research 178, 433-448.

Konstantaras, I., Papachristos, S., 2008. Note on: An optimal ordering and recovery policy for reusable items. Computers and Industrial Engineering 55, 729-734.

Li, Y., Chen, J., Cai, X., 2007. Heuristic genetic algorithm for capacitated production planning problems with batch processing and remanufacturing. International Journal of Production Economics 105, 301-317.

Mahadevan, B., Pyke, D.F., Fleischmann, M., 2003. Periodic review, push inventory policies for remanufacturing. European Journal of Operational Research 151, 536-551.

Minner, S., Kleber, R., 2001. Optimal control of production and remanufacturing in a simple recovery model with linear cost functions. OR Spektrum 23, 3-24.

Minner, S., 2003. Multiple-supplier inventory models in supply chain: A review. International Journal of Production Economics 81-82, 265-279.

Mostard, J., Teunter, R.H., 2006. The newsboy problem with resalable returns: A single period model and case study. European Journal of Operational Research 169, 81-96.

Muckstadt, J.A., Isaac, M.H., 1981. An analysis of single item inventory systems with returns. Naval Research Logistics Quarterly 28, 237-254.

Ouyang, H., Zhu, X., 2008. An inventory control model with disposal for manufacturing/remanufacturing hybrid system. Computers and Industrial Engineering 147, 62-71.

Paksoy, T., Bektaş, T., Özceylan, E., 2011. Operational and environmental performance measures in a multi-product closed-loop supply chain. Transportation Research Part E: Logistics and Transportation Review 47, 532-546.

Prahinski, C., Kocabasoglu, C., 2006. Empirical research opportunities in reverse supply chains. Omega 34, 519-532.

Ravi, V., Shankar, R., Tiwari, M.K., 2005. Analyzing alternatives in reverse-logistics for end-of-life computers: ANP and balanced scorecard approach. Computers and Industrial Engineering 48, 327-356.

Richter, K., 1994. An EOQ repair and waste disposal model. Proceedings of the 8th International Working Seminar on Production Economics, 83-91, Igls/Innsbruck, Austria.

Richter, K., 1996. The extended EOQ repair and waste model. International Journal of Production Economics 1-3, 443-448.

Richter, K., 1997. Pure and mixed strategies for the EOQ repair and waste disposal problem. OR Spektrum 19, 123-129.

Richter, K., Weber, J., 2001. The reverse Wagner/Whitin model with variable manufacturing and remanufacturing cost. International Journal of Production Economics 71, 447-456.

Rubio, S., Chamorro, A., Miranda, F.J., 2008. Characteristics of the research on reverse-logistics. International Journal of Production Research 46, 1099-1120.

Schrady, D.A., 1967. A deterministic inventory model for repairable items. Naval Research Logistics Quarterly 14, 391-398.

Simpson, V.P., 1970. An ordering model for recoverable stock items. AIIE Transactions 2, 315-320.

Simpson, V.P., 1978. Optimum solution structure for a repairable inventory problem. Operations Research 26, 270-281.

Sivinski, J.A., Meegan, S., 1993. Case study: Abbott labs formalized approach to remanufacturing. APICS Remanufacturing Seminar Proceedings, May, 27-30.

Sprow, E., 1992. The mechanics of remanufacturing. Manufacturing Engineering, March 38-52.

Stock, J.R., 1992. Reverse Logistics. Council of Logistics Management, Oak Brook, IL.

Teunter, R.H., 2001. Economic ordering quantities for recoverable item inventory systems. Naval Research Logistics 48, 484-495.

Teunter, R.H., 2002. Economic Order quantities for stochastic discounted cost inventory systems with remanufacturing. International Journal of Logistics Research and Applications 5, 161-175.

Teunter, R.H., Vlachos, D., 2002. On the necessity of a disposal option for returned items that can be remanufactured. International Journal of Production Economics 75, 257-266.

Teunter, R.H., Bayindir, Z.P., Heuvel, W.V.D., 2006. Dynamic lot sizing with product returns and remanufacturing. International Journal of Production Research 44, 4377–4400.

Thierry, M.C., Salomon, M., Van Nunen, J.A.E.E., Van Wassenhove, L.N., 1995. Strategic issues in product recovery management. California Management Review 37, 114-135.

Van der Laan, E.A., Dekker, R., Salomon, M., 1996a. An (*s, Q*) inventory model with remanufacturing and disposal. International Journal of Production Economics 46-47, 339-350.

Van der Laan, E.A., Dekker, R., Salomon, M., 1996b. Product remanufacturing and disposal: A numerical comparison of alternative control strategies. International Journal of Production Economics 45, 489-498.

Van der Laan, E.A., 1997. The effects of remanufacturing on inventory control. Ph.D. Thesis, Erasmus University Rotterdam, The Netherlands.

Van der Laan, E.A., Salomon, M., 1997. Production planning and inventory control with remanufacturing and disposal. European Journal of Operational Research 102, 264-278.

Van der Laan, E.A., Salomon, M., Dekker, R., 1999a. Leadtime effects in push and pull controlled manufacturing/remanufacturing systems. European Journal of Operational Research 115, 195-214.

Van der Laan, E.A., Salomon, M., Dekker, R., Van Wassenhove, L.N., 1999b. Inventory control in hybrid systems with remanufacturing. Management Science 45, 733-747.

Van der Laan, E.A., Teunter R.H., 2006. Simple heuristics for push and pull remanufacturing policies. European Journal of Operational Research 175, 1084-1102.

Vandermerwe, S., Oliff, M.D., 1991. Corporate challenges for an age of reconsumption. Columbia Journal of World Business, Fall vol., 7-25.

Vlachos, D., Dekker, R., 2003. Return handling options and order quantities for single period products. European Journal of Operational Research 151, 38-52.

Wagner, H.M., Whitin, T.M., 1958. Dynamic version of the economic lot size model. Management Science 5, 212-219.

Whisler, W.D., 1967. A stochastic inventory model for returned equipment. Management Science 13, 640-647.

Yuan, X., Cheung, K.L., 1998. Modeling returns of merchandise in an inventory system. OR Spektrum 20, 147-154.

Appendix A The threshold levels for the optimal inventory control of the two-product recovery system in a single period

The related threshold levels for the optimal inventory control of the two-product recovery system in a single period are listed in Table A.1 and Table A.2. Table A.1 includes the order-up-to levels for the inventory replenishment of the two products respectively. In addition, Table A.2 includes the threshold levels for the interactive inventory control of the two products.

Table A.1 The order-up-to levels for the optimal inventory control of the two-product recovery system in a single period

	Formula	Insight
***AL*₀**	$F^{-1}(\frac{s_1+v_1-c_{P1}}{s_1+v_1+h_1},\mu_1,\sigma_1)$	Is to balance between the opportunity loss due to one unit of product 1 short arising from production process ($s_1+v_1-c_{P1}$) and the opportunity loss due to having one unit excess of product 1 arising from production process ($c_{P1}+h_1$).
***AL*₁**	$F^{-1}(\frac{s_1+v_1-c_{R11}}{s_1+v_1+h_1},\mu_1,\sigma_1)$	Is to balance between the opportunity loss due to one unit of product 1 short arising from recovery process by returned items in group 1 ($s_1+v_1-c_{R11}$) and the opportunity loss due to having one unit excess of product 1 arising from the recovery process ($c_{R11}+h_1$).
***AL*₂**	$F^{-1}(\frac{s_1+v_1-c_{R21}}{s_1+v_1+h_1},\mu_1,\sigma_1)$	Is to balance between the opportunity loss due to one unit of product 1 short arising from recovery process by returned items in group 2 ($s_1+v_1-c_{R21}$) and the opportunity loss due to having one unit excess of product 1 arising from the recovery process ($c_{R21}+h_1$).
***BL*₀**	$F^{-1}(\frac{s_2+v_2-c_{P2}}{s_2+v_2+h_2},\mu_2,\sigma_2)$	Is to balance between the opportunity loss due to one unit of product 2 short arising from production process ($s_2+v_2-c_{P2}$) and the opportunity loss due to having one unit excess of product 2 arising from production process ($c_{P2}+h_2$).
***BL*₁**	$F^{-1}(\frac{s_2+v_2-c_{R12}}{s_2+v_2+h_2},\mu_2,\sigma_2)$	Is to balance between the opportunity loss due to one unit of product 2 short arising from recovery process by returned items in group 1 ($s_2+v_2-c_{R12}$) and the opportunity loss due to having one unit excess of product 2 arising from the recovery process ($c_{R12}+h_2$).
***BL*₂**	$F^{-1}(\frac{s_2+v_2-c_{R22}}{s_2+v_2+h_2},\mu_2,\sigma_2)$	Is to balance between the opportunity loss due to one unit of product 2 short arising from recovery process by returned items in group 2 ($s_2+v_2-c_{R22}$) and the opportunity loss due to having one unit excess of product 2 arising from the recovery process ($c_{R22}+h_2$).

Table A.2 The threshold levels for the interactive inventory control of the two-product recovery system in a single period

	Formula	Insight
***SW*₁**	$F^{-1}(\frac{s_1+v_1+c_{R12}-c_{R11}-c_{P2}}{s_1+v_1+h_1},\mu_1,\sigma_1)$	Is to balance between the opportunity loss due to one unit of product 1 short arising from switching of returned items in group 1 from the recovery process for product 1 to that for product 2 in place of production process (s_1+v_1+c_{R12}–c_{R11}–c_{P2}) and the opportunity loss due to having one unit excess of product 1 arising from no switching (c_{R11}+c_{P2}–c_{R12}+h_1).
***SW*₂**	$F^{-1}(\frac{s_1+v_1+c_{R22}-c_{R21}-c_{P2}}{s_1+v_1+h_1},\mu_1,\sigma_1)$	Is to balance between the opportunity loss due to one unit of product 1 short arising from switching of returned items in group 2 from the recovery process for product 1 to that for product 2 in place of production process (s_1+v_1+c_{R22}–c_{R21}–c_{P2}) and the opportunity loss due to having one unit excess of product 1 arising from no switching (c_{R21}+c_{P2}–c_{R22}+h_1).
RP	$F^{-1}(\frac{s_1+v_1+c_{R22}-c_{R21}-c_{R12}}{s_1+v_1+h_1},\mu_1,\sigma_1)$	Is to balance between the opportunity loss due to one unit of product 1 short arising from no replacement and reallocation of returned items in group 2 by returned items in group 1 (s_1+v_1+c_{R22}-c_{R21}-c_{R12}) and the opportunity loss due to having one unit excess of product 1 arising from the replacement and reallocation process (c_{R12}+c_{R21}-c_{R22}+h_1).

Appendix B The structures of the optimal solution to the single-period problem on the two-product recovery system

To maximize the expected profit in a single period, the optimal solution to the single-period problem on the two-product recovery system has different structural forms due to different combinations of the initial inventory of the two products and the availability of returned items. The solution structures involve the threshold levels, which have been explained in Chapter 4. In addition, there are notations: R_1 and R_2 denote the availability of returned items in group 1 and group 2 respectively; x_{S1} and x_{S2} denote the initial inventory of product 1 and product 2 respectively; RL_1 and RL_2 denote the replenishment level of product 1 and product 2 respectively. As some structures involve the comparison of marginal profits from allocating returned items to the recovery for the two products, we list the formulae of the related marginal profits as follows ($j = 1, 2$):

$$\begin{aligned} \frac{\partial EP}{\partial r_{1j}} &= s_j + v_j - c_{R1j} - (s_j + v_j + h_j)F(x_j, \mu_j, \sigma_j); \\ \frac{\partial EP}{\partial r_{2j}} &= s_j + v_j - c_{R2j} - (s_j + v_j + h_j)F(x_j, \mu_j, \sigma_j). \end{aligned} \tag{B.1}$$

In order to obtain the perturbation effect of the initial inventory of the two products on the optimal replenishment decisions, we have listed nonzero values of the first-order derivatives of the optimal replenishment decisions with respect to the initial inventory of the two products in Table B.1. In more details, the solution structures are listed as follows:

S1. $R_1 + R_2 + x_{S1} < AL_0,\ \ x_{S2} < BL_0$:

$$p_1^* = AL_0 - R_1 - R_2 - x_{S1},\quad r_{11}^* = R_1,\quad r_{21}^* = R_2;\quad (RL_1 = AL_0)$$
$$p_2^* = BL_0 - x_{S2},\qquad r_{12}^* = 0,\quad r_{22}^* = 0.\quad (RL_2 = BL_0)$$

S2. $AL_0 \le R_1 + R_2 + x_{S1} \le SW_1,\ \ x_{S2} < BL_0$:

$$p_1^* = 0,\qquad r_{11}^* = R_1,\quad r_{21}^* = R_2;\quad (AL_0 \le RL_1 \le SW_1)$$
$$p_2^* = BL_0 - x_{S2},\quad r_{12}^* = 0,\quad r_{22}^* = 0.\quad (RL_2 = BL_0)$$

S3. $R_1 + R_2 + x_{S1} < AL_0,\ \ x_{S2} \ge BL_0$:

$$p_1^* = AL_0 - R_1 - R_2 - x_{S1},\quad r_{11}^* = R_1,\quad r_{21}^* = R_2;\quad (RL_1 = AL_0)$$
$$p_2^* = 0,\qquad r_{12}^* = 0,\quad r_{22}^* = 0.\quad (BL_0 \le RL_2 \le BL_2)$$

S4. $AL_0 \le R_1 + R_2 + x_{S1} \le AL_1,\ \ x_{S2} \ge BL_0,\ \ \left.\dfrac{\partial EP}{\partial r_{11}}\right|_{x_1 = R_1 + R_2 + x_{S1}} \ge \left.\dfrac{\partial EP}{\partial r_{12}}\right|_{x_2 = x_{S2}}$:

$$p_1^* = 0,\quad r_{11}^* = R_1,\quad r_{21}^* = R_2;\quad (AL_0 \le RL_1 \le AL_1)$$
$$p_2^* = 0,\quad r_{12}^* = 0,\quad r_{22}^* = 0.\quad (BL_0 \le RL_2 \le BL_2)$$

S5. $R_1 + R_2 + x_{S1} > AL_1,\ \ R_2 + x_{S1} \le AL_1,\ \ x_{S2} \ge BL_1$:

$$p_1^* = 0,\quad r_{11}^* = AL_1 - R_2 - x_{S1},\quad r_{21}^* = R_2;\quad (RL_1 = AL_1)$$
$$p_2^* = 0,\quad r_{12}^* = 0,\qquad r_{22}^* = 0.\quad (BL_1 \le RL_2 \le BL_2)$$

S6. $R_2 + x_{S1} > AL_1,\ \ x_{S2} \ge BL_1,\ \ \left.\dfrac{\partial EP}{\partial r_{21}}\right|_{x_1 = R_2 + x_{S1}} \ge \left.\dfrac{\partial EP}{\partial r_{22}}\right|_{x_2 = x_{S2}}$:

$$p_1^* = 0,\quad r_{11}^* = 0,\quad r_{21}^* = R_2;\quad (AL_1 < RL_1 \le AL_2)$$
$$p_2^* = 0,\quad r_{12}^* = 0,\quad r_{22}^* = 0.\quad (BL_1 \le RL_2 \le BL_2)$$

S7. $R_2 + x_{S1} < SW_1,\ \ R_1 + R_2 + x_{S1} > SW_1,\ \ R_1 + R_2 + x_{S1} + x_{S2} \le SW_1 + BL_0$:

$$p_1^* = 0,\qquad r_{11}^* = SW_1 - R_2 - x_{S1},\qquad r_{21}^* = R_2;\quad (RL_1 = SW_1)$$
$$p_2^* = SW_1 + BL_0 - R_1 - R_2 - x_{S1} - x_{S2},\quad r_{12}^* = R_1 + R_2 + x_{S1} - SW_1,\quad r_{22}^* = 0.\quad (RL_2 = BL_0)$$

S8. $SW_1 \le R_2 + x_{S1} \le SW_2,\ \ R_1 + x_{S2} < BL_0$:

$$p_1^* = 0, \qquad r_{11}^* = 0, \quad r_{21}^* = R_2; \ \ (SW_1 \le RL_1 \le SW_2)$$
$$p_2^* = BL_0 - R_1 - x_{S2}, \ \ r_{12}^* = R_1, \ \ r_{22}^* = 0. \quad (RL_2 = BL_0)$$

S9. $BL_0 \le R_1 + x_{S2} \le BL_1,\ \ \left.\frac{\partial EP}{\partial r_{21}}\right|_{x_1 = R_2 + x_{S1}} \ge \left.\frac{\partial EP}{\partial r_{22}}\right|_{x_2 = R_1 + x_{S2}},\ \ \left.\frac{\partial EP}{\partial r_{11}}\right|_{x_1 = R_2 + x_{S1}} \le \left.\frac{\partial EP}{\partial r_{12}}\right|_{x_2 = R_1 + x_{S2}}$:

$$p_1^* = 0, \ \ r_{11}^* = 0, \quad r_{21}^* = R_2; \ \ (SW_1 \le RL_1 \le RP)$$
$$p_2^* = 0, \ \ r_{12}^* = R_1, \ \ r_{22}^* = 0. \quad (BL_0 \le RL_2 \le BL_1)$$

S10. $AL_1 \le R_2 + x_{S1} \le RP,\ \ x_{S2} < BL_1,\ \ R_1 + x_{S2} > BL_1$:

$$p_1^* = 0, \ \ r_{11}^* = 0, \qquad r_{21}^* = R_2; \ \ (AL_1 \le RL_1 \le RP)$$
$$p_2^* = 0, \ \ r_{12}^* = BL_1 - x_{S2}, \ \ r_{22}^* = 0. \quad (RL_2 = BL_1)$$

S11. $SW_1 + BL_0 < R_1 + R_2 + x_{S1} + x_{S2} \le AL_1 + BL_1,$

$$\left.\frac{\partial EP}{\partial r_{11}}\right|_{x_1 = R_1 + R_2 + x_{S1}} < \left.\frac{\partial EP}{\partial r_{12}}\right|_{x_2 = x_{S2}},\ \ \left.\frac{\partial EP}{\partial r_{11}}\right|_{x_1 = R_2 + x_{S1}} > \left.\frac{\partial EP}{\partial r_{12}}\right|_{x_2 = R_1 + x_{S2}} :$$

$$p_1^* = 0, \ \ r_{21}^* = R_2; \ \ p_2^* = 0, \ \ r_{22}^* = 0;$$

solve $\left.\frac{\partial EP}{\partial r_{11}}\right|_{x_1 = r_{11}^* + R_2 + x_{S1}} = \left.\frac{\partial EP}{\partial r_{12}}\right|_{x_2 = r_{12}^* + x_{S2}}$ and $r_{11}^* + r_{12}^* = R_1$ to obtain r_{11}^*, r_{12}^*.

$(SW_1 < RL_1 \le AL_1,\ \ BL_0 < RL_2 \le BL_1)$

S12. $R_2 + x_{S1} < AL_1,\ \ x_{S2} < BL_1,\ \ R_1 + R_2 + x_{S1} + x_{S2} > AL_1 + BL_1$:

$$p_1^* = 0, \ \ r_{11}^* = AL_1 - R_2 - x_{S1}, \ \ r_{21}^* = R_2; \ \ (RL_1 = AL_1)$$
$$p_2^* = 0, \ \ r_{12}^* = BL_1 - x_{S2}, \qquad r_{22}^* = 0. \quad (RL_2 = BL_1)$$

S13. $x_{S1} < SW_2,\ \ R_2 + x_{S1} > SW_2,\ \ R_1 + R_2 + x_{S1} + x_{S2} \le SW_2 + BL_0$:

$$p_1^* = 0, \qquad r_{11}^* = 0, \quad r_{21}^* = SW_2 - x_{S1}; \qquad (RL_1 = SW_2)$$
$$p_2^* = SW_2 + BL_0 - R_1 - R_2 - x_{S1} - x_{S2},\ \ r_{12}^* = R_1,\ \ r_{22}^* = R_2 + x_{S1} - SW_2. \quad (RL_2 = BL_0)$$

S14. $x_{S1} \ge SW_2,\ \ R_1 + R_2 + x_{S2} < BL_0$:

$$p_1^* = 0, \qquad r_{11}^* = 0, \quad r_{21}^* = 0; \quad (SW_2 \le RL_1 \le AL_2)$$
$$p_2^* = BL_0 - R_1 - R_2 - x_{S2},\ \ r_{12}^* = R_1,\ \ r_{22}^* = R_2. \quad (RL_2 = BL_0)$$

S15. $BL_0 \le R_1 + R_2 + x_{S2} \le BL_1,\ \ \left.\dfrac{\partial EP}{\partial r_{21}}\right|_{x_1 = x_{S1}} \le \left.\dfrac{\partial EP}{\partial r_{22}}\right|_{x_2 = R_1 + R_2 + x_{S2}}$:

$$p_1^* = 0,\ \ r_{11}^* = 0, \quad r_{21}^* = 0; \quad (SW_2 \le RL_1 \le AL_2)$$
$$p_2^* = 0,\ \ r_{12}^* = R_1,\ \ r_{22}^* = R_2. \quad (BL_0 \le RL_2 \le BL_1)$$

S16. $x_{S1} \ge RP,\ \ R_2 + x_{S2} \le BL_1,\ \ R_1 + R_2 + x_{S2} > BL_1$:

$$p_1^* = 0,\ \ r_{11}^* = 0, \qquad r_{21}^* = 0; \quad (RP \le RL_1 \le AL_2)$$
$$p_2^* = 0,\ \ r_{12}^* = BL_1 - R_2 - x_{S2},\ \ r_{22}^* = R_2. \quad (RL_2 = BL_1)$$

S17. $SW_2 + BL_0 < R_1 + R_2 + x_{S1} + x_{S2} \le RP + BL_1$,

$$\left.\frac{\partial EP}{\partial r_{21}}\right|_{x_1 = x_{S1}} > \left.\frac{\partial EP}{\partial r_{22}}\right|_{x_2 = R_2 + R_1 + x_{S2}},\quad \left.\frac{\partial EP}{\partial r_{21}}\right|_{x_1 = R_2 + x_{S1}} < \left.\frac{\partial EP}{\partial r_{22}}\right|_{x_2 = R_1 + x_{S2}} :$$

$$p_1^* = 0,\ \ r_{11}^* = 0;\ \ p_2^* = 0,\ \ r_{12}^* = R_1;$$

solve $\left.\dfrac{\partial EP}{\partial r_{21}}\right|_{x_1 = r_{21}^* + x_{S1}} = \left.\dfrac{\partial EP}{\partial r_{22}}\right|_{x_2 = r_{22}^* + R_1 + x_{S2}}$ and $r_{21}^* + r_{22}^* = R_2$ to obtain r_{21}^*, r_{22}^*.

$(SW_2 < RL_1 \le RP,\ \ BL_0 < RL_2 \le BL_1)$

S18. $x_{S1} < RP,\ \ R_2 + x_{S1} > RP,\ \ R_2 + x_{S1} + x_{S2} < RP + BL_1,\ \ R_1 + R_2 + x_{S1} + x_{S2} > RP + BL_1:$

$$p_1^* = 0,\ \ r_{11}^* = 0,\qquad r_{21}^* = RP - x_{S1};\qquad (RL_1 = RP)$$
$$p_2^* = 0,\ \ r_{12}^* = RP + BL_1 - R_2 - x_{S1} - x_{S2},\quad r_{22}^* = R_2 + x_{S1} - RP.\quad (RL_2 = BL_1)$$

S19. $RP + BL_1 \le R_2 + x_{S1} + x_{S2} \le AL_2 + BL_2,$

$$\left.\frac{\partial EP}{\partial r_{21}}\right|_{x_1 = R_2 + x_{S1}} < \left.\frac{\partial EP}{\partial r_{22}}\right|_{x_2 = x_{S2}},\quad \left.\frac{\partial EP}{\partial r_{21}}\right|_{x_1 = x_{S1}} > \left.\frac{\partial EP}{\partial r_{22}}\right|_{x_2 = R_2 + x_{S2}}:$$

$p_1^* = 0,\ \ r_{11}^* = 0;\ \ p_2^* = 0,\ \ r_{12}^* = 0;$

solve $\left.\frac{\partial EP}{\partial r_{21}}\right|_{x_1 = r_{21}^* + x_{S1}} = \left.\frac{\partial EP}{\partial r_{22}}\right|_{x_2 = r_{22}^* + x_{S2}}$ and $r_{21}^* + r_{22}^* = R_2$ to obtain r_{21}^*, r_{22}^*.

$(RP \le RL_1 \le AL_2,\ \ BL_1 \le RL_2 \le BL_2)$

S20. $R_2 + x_{S2} > BL_1,\ \ \left.\frac{\partial EP}{\partial r_{21}}\right|_{x_1 = x_{S1}} \le \left.\frac{\partial EP}{\partial r_{22}}\right|_{x_2 = R_2 + x_{S2}}:$

$$p_1^* = 0,\ \ r_{11}^* = 0,\ \ r_{21}^* = 0;\quad (RP < RL_1 \le AL_2)$$
$$p_2^* = 0,\ \ r_{12}^* = 0,\ \ r_{22}^* = R_2.\ \ (BL_1 < RL_2 \le BL_2)$$

S21. $R_2 + x_{S1} + x_{S2} > AL_2 + BL_2:$

$$p_1^* = 0,\ \ r_{11}^* = 0,\ \ r_{21}^* = AL_2 - x_{S1};\ \ (RL_1 = AL_2)$$
$$p_2^* = 0,\ \ r_{12}^* = 0,\ \ r_{22}^* = BL_2 - x_{S2}.\ \ (RL_2 = BL_2)$$

According to the above solution structures, the nonzero values of the first-order derivatives of the optimal replenishment decisions with respect to the initial inventory of the two products are listed in Table B.1.

Table B.1 The nonzero values of the first-order derivatives of the optimal replenishment decisions with respect to the initial inventory of the two products

	$\frac{\partial p_1}{\partial x_{S1}}$	$\frac{\partial p_1}{\partial x_{S2}}$	$\frac{\partial r_{11}}{\partial x_{S1}}$	$\frac{\partial r_{11}}{\partial x_{S2}}$	$\frac{\partial r_{21}}{\partial x_{S1}}$	$\frac{\partial r_{21}}{\partial x_{S2}}$	$\frac{\partial p_2}{\partial x_{S1}}$	$\frac{\partial p_2}{\partial x_{S2}}$	$\frac{\partial r_{12}}{\partial x_{S1}}$	$\frac{\partial r_{12}}{\partial x_{S2}}$	$\frac{\partial r_{22}}{\partial x_{S1}}$	$\frac{\partial r_{22}}{\partial x_{S2}}$
S1	-1							-1				
S2								-1				
S3	-1											
S5			-1									
S7			-1				-1	-1	1			
S8								-1				
S10										-1		
S11			$-C_1$	$1-C_1$					C_1	C_1-1		
S12			-1							-1		
S13					-1		-1	-1			1	
S14								-1				
S16										-1		
S17					$-C_2$	$1-C_2$					C_2	C_2-1
S18					-1				-1	-1	1	
S19					$-C_3$	$1-C_3$					C_3	C_3-1
S21					-1							-1

In Table B.1, the variables C_1, C_2 and C_3 can be calculated as follows:

$$C_1 = \frac{(s_1+v_1+h_1)f(R_2+x_{S1}+r_{11}^*,\mu_1,\sigma_1)}{(s_1+v_1+h_1)f(R_2+x_{S1}+r_{11}^*,\mu_1,\sigma_1)+(s_2+v_2+h_2)f(x_{S2}+r_{12}^*,\mu_2,\sigma_2)};$$

$$C_2 = \frac{(s_1+v_1+h_1)f(x_{S1}+r_{21}^*,\mu_1,\sigma_1)}{(s_1+v_1+h_1)f(x_{S1}+r_{21}^*,\mu_1,\sigma_1)+(s_2+v_2+h_2)f(R_1+x_{S2}+r_{22}^*,\mu_2,\sigma_2)};$$

$$C_3 = \frac{(s_1+v_1+h_1)f(x_{S1}+r_{21}^*,\mu_1,\sigma_1)}{(s_1+v_1+h_1)f(x_{S1}+r_{21}^*,\mu_1,\sigma_1)+(s_2+v_2+h_2)f(x_{S2}+r_{22}^*,\mu_2,\sigma_2)}.$$

(B.2)

Appendix C The process of determining the sample gradient for the approximate dynamic programming models

In Chapter 5 and Chapter 6, we have developed the approximate dynamic programming models for the two-product recovery system considering lost sale and backorder respectively. In addition, production and recovery processes are assumed to have zero lead time. Due to a liner approximation involved in modeling, the two gradients of the cost-to-go function of dynamic programming with respect to the inventory level of the two products after replenishment need to be estimated by sample average through Monte Carlo simulation. Furthermore, the sample gradient $grad_{j,k}$ (j = 1, 2) is to be determined for the M-period problem given the sample k about the realization of stochastic returns and demands as:

$$\begin{pmatrix} R_{1,k}^{(2)} & R_{2,k}^{(2)} & D_{1,k}^{(1)} & D_{2,k}^{(1)} \\ R_{1,k}^{(3)} & R_{2,k}^{(3)} & D_{1,k}^{(2)} & D_{2,k}^{(2)} \\ \ldots\ldots & & & \\ R_{1,k}^{(M)} & R_{2,k}^{(M)} & D_{1,k}^{(M-1)} & D_{2,k}^{(M-1)} \end{pmatrix}.$$

As introduced in the two chapters, it is similar for the two products to determine the sample gradient of the cost-to-go function with regard to their respective inventory level after replenishment. We would take product 1 as an example to introduce the process of determining the sample gradient in Figure C.1 and Figure C.2 for the two-product recovery system considering lost sale and backorder respectively.

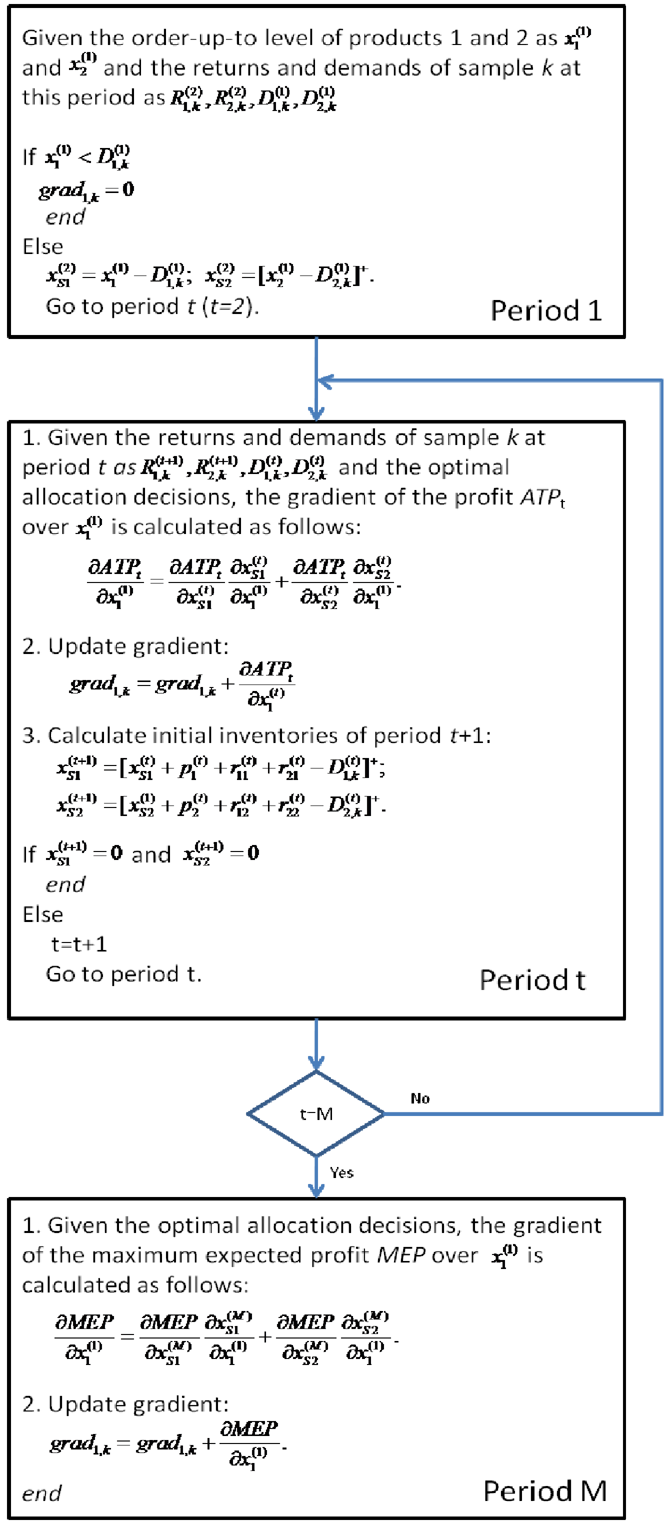


Figure C.1 The determination of the sample gradient for the two-product recovery system assuming lost sale and zero lead time

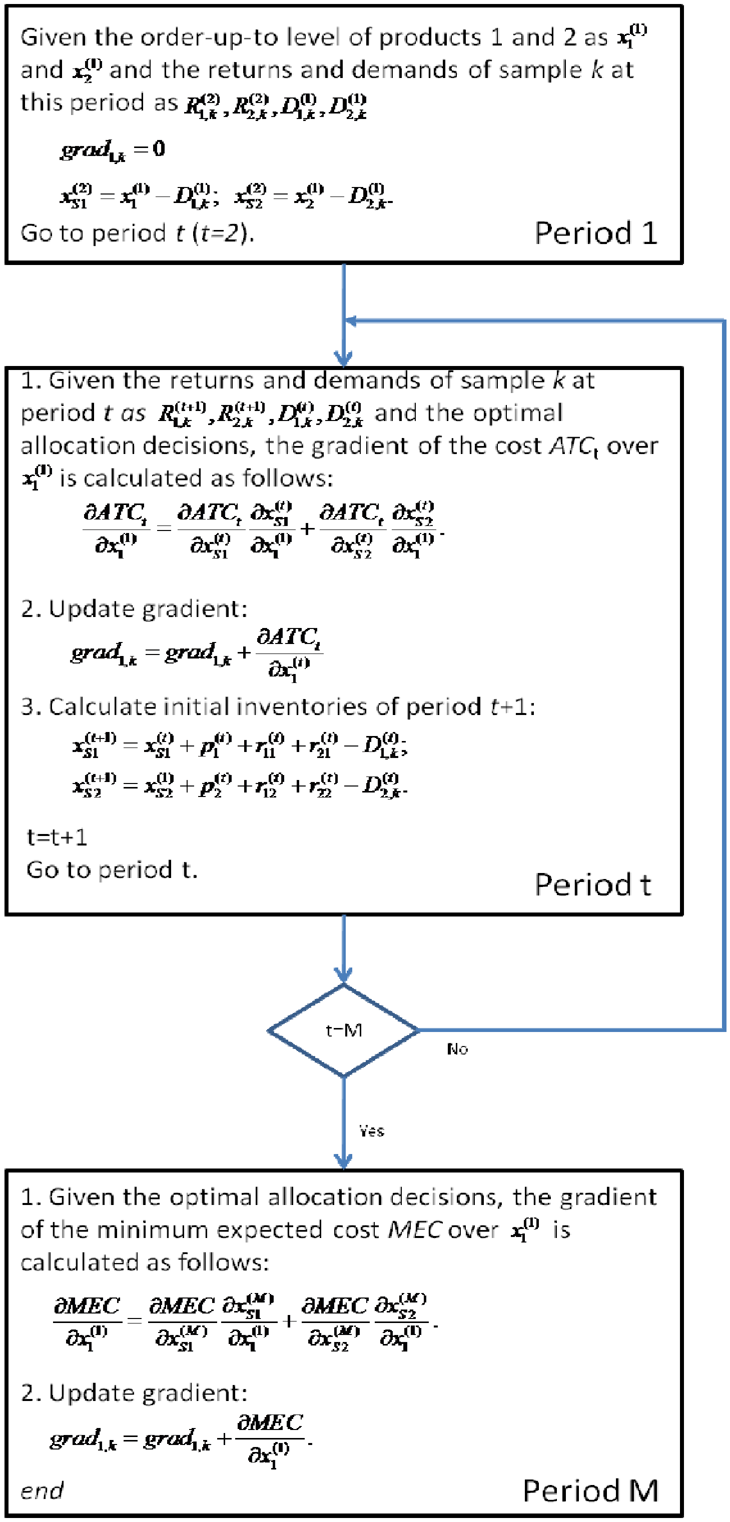


Figure C.2 The determination of the sample gradient for the two-product recovery system assuming backorder and zero lead time