The significance of product recovery through remanufacturing has been widely recognized and has compelled manufacturers to incorporate product recovery activities into normal manufacturing processes. Consequently, increasing attention has been paid to production and inventory management of the product recovery system where demand is satisfied through either manufacturing brand-new products or remanufacturing returned products into new ones. In this work, we investigate a recovery system with two product types and two return flows. A periodic-review inventory problem is addressed in the two-product recovery system and an approximate dynamic programming approach is proposed to obtain production and recovery decisions. A single-period problem is first solved and the optimal solution is characterized by a multilevel threshold policy. For the multi-period problem, we show that the threshold levels of each period are solely dependent on the gradients of the cost-to-go function at points of interest after approximation. The gradients are estimated by an infinitesimal perturbation analysis–based method and a backward induction approach is then applied to derive the threshold levels of each period. Numerical experiments are conducted under different scenarios and the threshold policy is shown to outperform two other heuristic policies.

**Keywords:** Two-product recovery system, product returns, remanufacturing, multilevel threshold policy, approximate dynamic programming

### 1. Introduction

In recent years, product recovery has emerged as an important field in reverse logistics and supply chain management, with increasing attention on the management of return flows in parallel with the conventional manufacturing process. Motivations behind the product recovery activities mainly include economic incentives and growing environmental concerns from society. Originally, engagement in recovery activities was often driven by government legislation that forced manufacturers to take responsibility for the whole product life cycle and waste reduction. However, direct and indirect economic benefits have now become the main reason, as raw material savings and reduction in disposal cost have gradually been recognized as significant. Furthermore, an environmentally friendly image also provides manufacturers a business advantage in the market.

Returns of used products complicate product recovery systems compared with traditional inventory systems as there are two options to fulfill customer demand: manufacturing brand-new products from virgin components by normal production and remanufacturing returned products into new ones by recovery. In other words, both virgin components and returns are available resources that can be used to make products. Thus, production planning and inventory management become a challenging issue for manufacturers as they need to coordinate between these two options in order to optimize the system’s performance. In current practice, virgin components are usually purchased from upstream suppliers and returned products can be acquired through a variety of reverse channels (Pokharel and Mutha, 2009; Ilgin and Gupta, 2010). In some cases, manufacturers collect the returned products directly from customers, whereas in some other cases manufacturers rely on retailers to return their used products (Wojanowski et al., 2007). In other cases, it is also very common for manufacturers to outsource used-product collection activities to independent third parties who are solely engaged in collecting and reprocessing used products from the market, allowing the manufacturers to focus on their core business.
Our work is motivated by a real case in which a manufacturer uses third parties as suppliers of returns for remanufacturing together with normal production. The manufacturer does not own any returned products and needs to acquire them from various third-party suppliers. Usually, the condition of returned products is highly variable, so different returns may require different responses and thus incur different remanufacturing costs (Blackburn et al., 2004; Zhou et al., 2011). The problem the manufacturer faces is how to jointly make remanufacturing decisions on different sources of returns and normal production decisions to minimize its total cost.

Considerable attention from academia has been paid to the study of product recovery; published papers are generally listed under the categories of reverse logistics and closed-loop supply chain. We refer readers to Prahinski and Kocabasoglu (2006), Pokharel and Mutha (2009), Ferguson and Souza (2010), and Ilgin and Gupta (2010) for complete reviews. Our work draws on and contributes to the primary stream of literature concerning the management of production and inventory of recovery systems that have been attracting growing research efforts. Many articles have been published that explore the structure of the optimal policy or propose a well-designed heuristic approach to draw near-optimal solutions (Thierry et al., 1995; Fleischmann et al., 1997; Ilgin and Gupta, 2010).

Product recovery systems can be classified by the number of return flows and demand flows. In practice, returns may be different in terms of type and quality condition, and thus fall into different groups. On the other hand, demand can also be divided based on different customer segments, service levels, etc. Specific recovery options are then applied to each different demand flow. For example, replacing a PC machine’s hard disk or processor or adding some extra space memory or a modem may be reasonable ways to exploit most of the value a used machine still contains (Inderfurth et al., 2001). Single-return-flow and single-demand-flow recovery systems have been widely studied, whereas there have only been a few studies on single-return-flow and multi-demand-flow recovery systems.

For recovery systems with one return and one demand flow, both deterministic and stochastic models have been developed. In deterministic models, the demand and return are known a priori for the entire planning horizon. Schrady (1967) extends the classic Economic Order Quantity (EOQ) model by considering return flow and develops a model on a system with constant demand and return rate. Similar models are proposed by Richter (1994, 1996) and Teunter (2001). Koh et al. (2002) consider a joint EOQ and Economic Production Quantity (EPQ) model and assume a proportion of the used products are returned. Their approach is improved by Konstantaras and Papachristos (2008).

Continuous- and periodic-review policies are studied in stochastic models and can be further divided into two groups, with one considering a single aggregated stock point and the other distinguishing recoverable and serviceable inventories. Within the former class, Yuan and Cheung (1998) model dependent demand and return by assuming an exponentially distributed market sojourn time. Van der Laan and Teunter (2006) consider a recovery system controlled by certain extensions of the \((s, Q)\) policy. The traditional \((s, Q)\) model is extended into an order-disposal strategy to control a hybrid system by Ouyang and Zhu (2008). Some other researchers distinguish serviceable and recoverable inventory in their works. Different continuous-review policies for controlling serviceable and recoverable inventory are analyzed in Van der Laan, Salomon, and Dekker (1999) and van der Laan, Salomon, Dekker, and van Wassenhove (1999). Inderfurth and van der Laan (2001) further improve this model by including a modified inventory position. As for periodic-review policies, Simpson (1978) proposes an inventory model based on a fixed periodic-review policy and obtains the optimal solution structure for the multi-period problem. Inderfurth (1997) extends this model by considering the impact of non-zero lead times for production and recovery. Kiesmüller (2003) presents a method for the exact computation of the parameters that determine the optimal policy in Simpson (1978). DeCroix (2006) extends the studies of Simpson (1978) and Inderfurth (1997) by identifying the structure of the optimal remanufacturing/ordering/disposal policy. DeCroix et al. (2005) propose a stochastic model of a multistage system with stationary costs and stochastic demand over an infinite horizon. Ahiska and King (2010) discuss inventory optimization of a single-product stochastic system with two stocking points and find optimal inventory policies by using Markov decision processes.

A key assumption in the above cited publications on recovery systems is that only one remanufacturing option is considered, which indicates the number of demand flows is one. However, there exist different options of reusing returns in practice; e.g., the manufacturer may be required to make different products by recovery that are of different types or have different qualities. Some works have addressed this problem. Inderfurth et al. (2001) present a periodic-review model for product recovery in a system with returns of a single product and multiple alternating reuse options. A fairly simple near-optimal policy, characterized by a dispose-down-to level and a recover-up-to level for each option is obtained by taking advantage of a linear allocation rule on returns. Kleber et al. (2002) propose a continuous model of a recovery system with multiple recovery options, with each corresponding to a different demand class.

To the best of our knowledge, no work has been published to investigate a product recovery system that involves multiple demand flows or multiple return flows. This is due to this type of problem being significantly more complex than its single-demand-flow and single-return-flow
counterparts. However, studies on systems with multiple streams of demands and returns is of practical importance. For example, returns of even the same product type may vary in quality condition. Furthermore, different providers of returned products may charge differently. Aras et al. (2004) assess the impact of quality-based categorization of returned products and show that incorporating returned product quality into the remanufacturing and disposal decisions can lead to significant cost savings. Behret and Korgan (2009) construct a simulation model to analyze the effect of quality classification of returns and indicate that quality-based classification of returns can result in significant cost savings, especially when return rates are high. Zhou et al. (2011) study an inventory system with multiple types of remanufacturable products and characterize a manufacturing–remanufacturing–disposal policy.

In this article we investigate a two-product recovery system, in which two types of products are demanded and two return flows are involved. Returns can be remanufactured into new products to satisfy two respective demand flows in parallel with normal production over a finite horizon. Our proposed periodic-review inventory model is used to determine the production quantities of each manufacturing/remanufacturing channel and characterize the inventory policy for the system with uncertain returns and demands. A dynamic programming model is developed to maximize the expected total profit and determine the optimal production and inventory decisions. Since the last period of multi-period problem actually forms a single-period problem, it is natural to investigate the single-period problem first. We prove the concavity of the single-period model and obtain global optimum by solving Karush–Kuhn–Tucker (KKT) conditions. The optimal solution contains 21 cases in total that are further analyzed and represented by a multilevel threshold policy. Although this threshold policy is not necessarily optimal for the multi-period problem, it is instructive and easy to extend. An approximate dynamic programming approach is then proposed for the multi-period problem and we show through approximation that the threshold levels of each period depend only on the gradients of the cost-to-go function at the points of interest to free us from assuming any explicit form for the cost-to-go function. The gradients are estimated by an Infinitesimal Perturbation Analysis (IPA)-based method and a backward induction approach is then conducted from the second-to-last period to the first period to sequentially derive the threshold policy of each period. The threshold policy is further experimentally proven to outperform two other heuristic policies under a wide range of settings and numerical examples are tested to assess the impact of uncertainty.

The rest of the article is organized as follows. Section 2 describes the two-product recovery system over a finite horizon. A dynamic programming model of the multi-period problem is developed. Section 3 studies the system in a single period. An optimal multilevel threshold policy is obtained and managerial insights into the policy are provided. In Section 4, we focus on the study of multi-period problem. The multilevel threshold policy of single-period system is extended and a backward induction approach is proposed to derive the threshold levels of each period. Numerical experiments are conducted in Section 5 to show the performance of our policy and the impact of uncertainties is investigated. We conclude our work in Section 6.

2. A dynamic programming model

2.1. Problem description

A periodic-review two-product recovery system with a planning horizon of $T$ periods indexed by $t \in \{1, \ldots, T\}$ is considered. Two product flows are provided to fulfill stochastic demand following independent stationary general distributions. Products can be either manufactured from virgin components or remanufactured from returned products and are assumed to be indistinguishable between manufacturing and remanufacturing. Two groups of returns, which can be acquired from the suppliers, are available to the manufacturer for remanufacturing. Each type of returns can be recovered to either of the two products and one unit of returns is remanufactured to one unit of new product. Returns of the same group are recovered at the same remanufacturing cost. Moreover, normal

| Table 1. Parameters and decision variables |
|---|
| **Parameters:** |
| $I_{jt}$: initial inventory of product $j$ during period $t$ |
| $s_j$: selling price of product $j$ |
| $rc_{ij}$: remanufacturing cost of one unit of returned products from group $i$ to produce $j$ |
| $pc_j$: manufacturing cost for one unit of product $j$ |
| $h_j$: holding cost for one unit of product $j$ |
| $v_j$: penalty cost of one unit of shortage of product $j$ |
| $R_i$: available returns in group $i$ during period $t$ |
| $\{R_i\}$: set of returns $R_i$ |
| $D_{jt}$: demand for product $j$ during period $t$ |
| $\mu_{jt}$: the mean of demand for product $j$ during period $t$ |
| $f_{jt}(\cdot), F_{jt}(\cdot)$: pdf and cdf of the distribution of $D_{jt}$ for product $j$ during period $t$ |
| $F_{jt}^{-1}(\cdot)$: inverse function of $F_{jt}(\cdot)$ |
| **Decision variables:** |
| $p_{jt}$: amount of product $j$ produced during period $t$ |
| $r_{ijt}$: amount of returned products of $\{R_i\}$ recovered to produce product $j$ during period $t$ |
Optimal control for a two-product recovery system

Fig. 1. Two-product recovery system with two return flows and two demand flows.

production is more costly than recovery, so normal production is used only when returns for recovery are insufficient. The capacity for normal production is unlimited.

A simplified operation flow of the manufacturer at each period is described as follows. After observing the on-hand inventory, the manufacturer contacts the suppliers to check on the availability of the two groups of returned products. Based on the information about inventory and availability of returns, the manufacturer makes recovery and normal production decisions and then orders a certain amount of returned products of each group from the suppliers. Inventory is replenished instantly and is used to satisfy demand later in the same period. If demand cannot be fully satisfied, unsatisfied demand is lost forever and penalty cost on the shortage is incurred. Otherwise, remaining inventory is carried to subsequent periods and holding cost is counted. Revenue of the system is generated by selling new products and the total cost consists of manufacturing cost, remanufacturing cost, holding cost, and penalty cost. Since the manufacturer does not actually own the returns, it does not need to keep inventory for the returns. Moreover, we assume that the remanufacturing lead times for different groups of returns are identical and identical to the normal production lead time (see Inderfurth (1997), DeCroix (2006), and Zhou et al. (2011) for similar assumptions). Since the model under these assumptions is equivalent to the one with zero lead times, in this work we focus on the problem with zero lead times for ease of exposition.

The two products are denoted as product 1 and product 2 and the two groups of returns are denoted as group 1 and group 2. We use indexes $i \in \{1, 2\}$ and $j \in \{1, 2\}$ to distinguish between two groups of returns and two types of products, respectively. Parameters and decision variables are specified in Table 1. The structure of the two-product recovery system is depicted in Fig. 1 where index $t$ is omitted for simplicity.

In addition, some restrictions on cost parameters are imposed to ensure the economic validity of this study. First, for each product, the selling price is higher than the manufacturing cost, and the penalty cost is higher than the profit from normal production:

$$s_1 > pc_1, \ s_2 > pc_2, \ v_1 > s_1 - pc_1, \ v_2 > s_2 - pc_2.$$  \hfill (1)

The manufacturing cost is higher than remanufacturing cost or else recovery is unnecessary:

$$pc_1 > rc_{11}, \ pc_1 > rc_{21}, \ pc_2 > rc_{12}, \ pc_2 > rc_{22}. \hfill (2)$$

Finally, without loss of generality, the recovery using returns of $\{R_2\}$ is assumed to be cheaper than using returns of $\{R_1\}$:

$$rc_{21} < rc_{11}, \ rc_{22} < rc_{12}. \hfill (3)$$

Therefore, to fulfill the demand, group 2 always has the highest priority to be selected to make products. Group 1 takes the second place and normal production is used only when returns in groups 1 and 2 are used up.

2.2. Model formulation

A dynamic programming model of the two-product recovery system is developed with the aim of maximizing the expected profit over a finite time horizon. In practice, once production and recovery decisions have been made, inventory gets replenished instantly and is used
to satisfy product demands later in the same period. We define $I_{t1}$ and $I_{t2}$ as the inventory level after replenishment at period $t$ and thus we have $I_{t1} = I_{t1} + p_{t1} + r_{11t} + r_{21t}$, $I_{t2} = I_{t2} + p_{t2} + r_{12t} + r_{22t}$. Therefore, the expected revenue of period $t$ (denoted as $ER_t$) is formulated as follows:

$$ER_t = s_1 \mu_{t1} + s_2 \mu_{t2} - s_1 \int_{I_{t1}}^{+\infty} \left( D_{t1} - I_{t1}' \right) f_{t1}(\cdot) dD_{t1}$$

$$- s_2 \int_{I_{t2}}^{+\infty} \left( D_{t2} - I_{t2}' \right) f_{t2}(\cdot) dD_{t2}.$$  \hspace{1cm} (4)

The total cost includes manufacturing cost, remanufacturing cost, holding cost, and penalty cost of shortage. Hence, the expected total cost of period $t$ is formulated as

$$EC_t = pc_1 p_{1t} + pc_2 p_{2t} + rc_{t1} r_{11t} + rc_{t2} r_{12t} + rc_{t1} r_{11t} + rc_{t2} r_{12t}$$

$$+ h_1 \int_{0}^{I_{t1}} \left( I_{t1}' - D_{t1} \right) f_{t1}(\cdot) dD_{t1} + v_1 \int_{I_{t1}}^{+\infty} \left( D_{t1} - I_{t1}' \right) f_{t1}(\cdot) dD_{t1}$$

$$+ h_2 \int_{0}^{I_{t2}} \left( I_{t2}' - D_{t2} \right) f_{t2}(\cdot) dD_{t2} + v_2 \int_{I_{t2}}^{+\infty} \left( D_{t2} - I_{t2}' \right) f_{t2}(\cdot) dD_{t2}. \hspace{1cm} (5)

The optimal cost-to-go function from period $t$ until the final period is denoted as $Q_t(I_{t1}, I_{t2})$. For the multi-period problem, Bellman’s equation of a dynamic programming model can be written as follows ($t = 1, 2, \ldots, T$):

$$Q_t(I_{t1}, I_{t2}) = \max EP_t + E[Q_{t+1}(I_{t1+1}, I_{t2+1})]. \hspace{1cm} (7)$$

Consumption of the returns of the two groups is no more than availability. Thus,

$$r_{11t} + r_{12t} \leq R_{t1}, \ r_{21t} + r_{22t} \leq R_{t2}, \ t = 1, 2, \ldots, T. \hspace{1cm} (8)$$

Finally, all decision variables should be non-negative; i.e., $p_{t1} \geq 0$, $p_{t2} \geq 0$, $r_{11t} \geq 0$, $r_{12t} \geq 0$, $r_{21t} \geq 0$, $r_{22t} \geq 0$, $t = 1, 2, \ldots, T$. Due to the curse of dimensionality, a dynamic programming problem involving more than two states is intractable, although for some problems where there is enough structure, the multidimensional dynamic programming problem can be solved. Unfortunately, our problem does not exhibit enough structure to avoid the curse of dimensionality because the state space of each state variable being too large and the probability distributions of the random variables are unknown. Thus, in our work, we seek to use an approximate dynamic programming technique to solve the problem. We will first investigate the single-period problem as a starting point and then apply the properties found to solve the multi-period dynamic programming model.

### 3. Recovery system in a single period

#### 3.1. Solving the single period problem

At the beginning of a single period, information about returns and the initial inventory of the two products are known. Based on this information, production and recovery decisions are made to maximize the expected profit. As recovery decisions are subject to the availability of the returns, the model for the single-period problem is formulated as

$$\min EP,$$  \hspace{1cm} (9)

subject to

$$r_{11} + r_{12} \leq R_1, \ r_{21} + r_{22} \leq R_2,$$  \hspace{1cm} (10)

$$p_{1} \geq 0, \ p_{2} \geq 0, \ r_{11} \geq 0, \ r_{12} \geq 0, \ r_{21} \geq 0, \ r_{22} \geq 0,$$  \hspace{1cm} (11)

where

$$EP = s_1 \mu_{t1} + s_2 \mu_{t2} - pc_1 p_{1t} - pc_2 p_{2t}$$

$$- rc_{t1} r_{11t} - rc_{t2} r_{12t} - rc_{t1} r_{11t} - rc_{t2} r_{12t} \hspace{1cm} (12)$$

and $I_j'$ denotes the inventory level of each product type after replenishment: $I_j' = I_j + p_j + r_{j1} + r_{j2}, \ j = 1, 2$. To simplify the expression, let $L_1(p_{1t}, r_{11t}, r_{21t})$ and $L_2(p_{2t}, r_{12t}, r_{22t})$ denote the terms related to stochastic demands for products 1 and 2, respectively; that is,

$$L_1(p_{1t}, r_{11t}, r_{21t}) = -h_1 \int_{0}^{I_{t1}} \left( I_{t1}' - D_{t1} \right) f_{t1}(\cdot) dD_{t1}$$

$$- (S_t + v_1) \int_{I_{t1}}^{+\infty} \left( D_{t1} - I_{t1}' \right) f_{t1}(\cdot) dD_{t1}, \hspace{1cm} (13)$$

$$L_2(p_{2t}, r_{12t}, r_{22t}) = -h_2 \int_{0}^{I_{t2}} \left( I_{t2}' - D_{t2} \right) f_{t2}(\cdot) dD_{t2}$$

$$- (S_t + v_2) \int_{I_{t2}}^{+\infty} \left( D_{t2} - I_{t2}' \right) f_{t2}(\cdot) dD_{t2}. \hspace{1cm} (14)$$
By substituting Equation (13) into Equation (12), the objective function can then be rewritten as

$$EP = s_1\mu_1 + s_2\mu_2 - pc_1 p_1 - pc_2 p_2 - r c_11 r_{11} - r c_12 r_{12} - r c_22 r_{22} + L_1(p_1, r_{11}, r_{21}) + L_2(p_2, r_{12}, r_{22}).$$

(15)

As decision variables $p_1$, $p_2$, $r_{11}$, $r_{12}$, $r_{21}$, $r_{22}$ act similarly in the functions, the first-order partial derivatives of $L_1$ and $L_2$ with respect to them are equal to each other:

$$\frac{\partial L_j}{\partial p_j} = \frac{\partial L_j}{\partial r_{1j}} = \frac{\partial L_j}{\partial r_{2j}} = s_j + v_j - (s_j + v_j + h_j)F_j(I_j'),$$

$$j = 1, 2.$$

(16)

Therefore, we have

$$\frac{\partial EP}{\partial p_j} = -pc_j + \frac{\partial L_j}{\partial p_j} = -pc_j + s_j + v_j$$

$$- (s_j + v_j + h_j)F_j(I_j'),$$

$$\frac{\partial EP}{\partial r_{1j}} = -rc_{1j} + \frac{\partial L_j}{\partial r_{1j}} = -rc_{1j} + s_j + v_j$$

$$- (s_j + v_j + h_j)F_j(I_j'),$$

$$\frac{\partial EP}{\partial r_{2j}} = -rc_{2j} + \frac{\partial L_j}{\partial r_{2j}} = -rc_{2j} + s_j + v_j$$

$$- (s_j + v_j + h_j)F_j(I_j').$$

(17)

**Lemma 1.** The objective function of the dynamic programming model (9)–(11) is jointly concave on all the decision variables for the single-period two-product recovery system if the salvage value of the unused returned products is ignored.

**Proof.** Lemma 1 holds if and only if the nonlinear part of the objective function—i.e., functions $L_1$ and $L_2$—is concave on all of the decision variables.

Take $L_1$ for example. The second-order partial derivative of function $L_1$ with respect to production decision $p_1$ is calculated as follows (let $\theta$ denote its value):

$$\theta = \frac{\partial^2 L_1}{\partial p_1^2} = - (s_1 + v_1 + h_1)f_1(I_1') \leq 0.$$  

(18)

The inequality above also holds for variables $r_{11}$ and $r_{12}$. Thus, the concavity of $L_1$ on decision variables is then proven. Similarly, concavity of $L_2$ can also be proven in the same way. Therefore, Lemma 1 is proven.

The method of Lagrange multipliers is then applied to find the maximum of the model (9)–(11). The Lagrangian function (denoted as $L$) is expressed as follows:

$$L(p_1, p_2, r_{11}, r_{12}, r_{21}, r_{22}, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \lambda_8)$$

$$= EP + \lambda_1(r_{11} + r_{12} - R_1) + \lambda_2(r_{21} + r_{22} - R_2)$$

$$- \lambda_3 p_{11} - \lambda_4 p_{22} - \lambda_5 r_{11} - \lambda_6 r_{21} - \lambda_7 r_{12} - \lambda_8 r_{22}.$$  

(19)

In order to obtain the optimal solution, we need to consider the KKT conditions for the maximum of $L$. Lemma 1 has shown the concavity of the objective function (9). In addition, constraint (10) is linear on decision variables.

Therefore, the solution to the KKT conditions is also the global maximum of the model. By solving these KKT conditions, we can then obtain the optimal production and recovery decisions for the two products.

The values of the decision variables are dependent on the availability of each resource, initial inventory position, and cost structure. Under a certain cost structure, the optimal solution of the model usually contains a number of cases given different inventories and returns. After analyzing the optimal solution cases, we realize that they can be represented by an optimal multilevel threshold policy that is characterized by a set of threshold levels. Therefore, we can employ these threshold levels to help make optimal decisions once they are determined. In the following, we will explain how to determine the multilevel threshold policy and make optimal decisions with its help.

### 3.2. Optimal multilevel threshold policy

There exist three channels to make a product: i.e., normal production and recovery from returns of $\{R_1\}$ or $\{R_2\}$. To fulfill the demand and maximize the expected profit, we need to decide when to choose each channel and how many units of each product to make from each channel. In other words, recovery from returns provides alternative options to make products in the product recovery system. Resources of each channel need to be sequentially allocated between two products based on the cost structure.

As different cost structures usually result in different forms of production and recovery decisions, we choose one structure as an example to illustrate how the optimal solution is obtained. For other cost structures, the problem can be solved by adapting the modeling and solution process presented here. Note that we have already made assumptions (2) and (3) in Equations (2) and (3), which indicate that among the three channels, the sequence of being selected for replenishment is recovery from $\{R_3\}$, recovery from $\{R_1\}$, and normal production. Furthermore, to simplify the illustration we impose the following restrictions on the cost parameters:

$$pc_1 - r c_{11} > pc_2 - r c_{12}, \quad pc_1 - r c_{21} > pc_2 - r c_{22}.$$  

(20)

For instance, the first inequality above indicates that replenishment for one unit of product 1 from $\{R_1\}$ instead of normal production brings more cost saving than replenishment for one unit of product 2 from $\{R_1\}$ instead of normal production. Therefore, the three inequalities in Equation (20) indicate that if returns from $\{R_2\}$ or $\{R_1\}$ are not enough, the manufacturer will allocate them to product 1 rather than product 2.

At the beginning of the period, information about initial inventories and the availability of resources in each channel is known. Each resource is then sequentially allocated between two products until the corresponding resource is
used up or the amount of the product reaches a certain threshold level beyond which any further replenishment is not economically worthwhile any longer (i.e., order-up-to level) or replenishment for the other product becomes more profitable (i.e., switching level) so that the replenishment should switch. These threshold levels are used to control the process of resource allocation. Under the given cost structure, the optimal solution has 21 cases in total based on the different combinations of initial inventories and amount of returns. Details of the 21 cases are given in Appendix A1.

Among these threshold levels, there are three order-up-to levels for each product corresponding to three different resources. The order-up-to level is the highest inventory level when the product is replenished by a certain resource. It means any further replenishment after the order-up-to level is not profitable and the replenishment should stop. These order-up-to levels can be obtained by solving the related KKT conditions as follows.

1. Order-up-to level by production.

For each product, the order-up-to level by production is defined as the maximum inventory level by normal production. At this order-up-to level, the marginal profit of further replenishment is equal to zero. Let \( AL_0 \) and \( BL_0 \) denote the order-up-to levels by production for products 1 and 2, respectively. We have

\[
\frac{\partial EP}{\partial p_1^*} \bigg| _{I_1^*=AL_0} = \lambda_3 = 0, \quad \lambda_3 p_1^* = 0,
\]

\[p_1^* > 0 \implies AL_0 = F_1^{-1} \left( \frac{s_1 + v_1 - p c_1}{s_1 + v_1 + h_1} \right).	ag{21}\]

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

\[BL_0 = F_2^{-1} \left( \frac{s_2 + v_2 - p c_2}{s_2 + v_2 + h_2} \right).	ag{22}\]

2. Order-up-to level by group 1.

For each product, the order-up-to level by group 1 is defined as the maximum inventory level by replenishment using returns of \( \{R_1\} \). If they are enough for the recovery of two products, their inventories will be replenished until the order-up-to levels (denoted as \( AL_1 \) and \( BL_1 \)), at which the marginal profits of further replenishment are equal to zero. Thus, we have

\[
\frac{\partial EP}{\partial r_{11}^*} + \lambda_1 = 0, \quad \frac{\partial EP}{\partial r_{12}^*} + \lambda_1 = 0,
\]

\[\implies AL_1 = F_1^{-1} \left( \frac{s_1 + v_1 - r c_{11}}{s_1 + v_1 + h_1} \right),
\]

\[BL_1 = F_2^{-1} \left( \frac{s_2 + v_2 - r c_{12}}{s_2 + v_2 + h_2} \right).	ag{23}\]

3. Order-up-to level by group 2.

Similarly, for each product, the order-up-to levels by group 2 (denoted as \( AL_2 \) and \( BL_2 \)) are defined as the maximum inventory level by replenishment using \( \{R_2\} \). Thus, we have

\[AL_2 = F_1^{-1} \left( \frac{s_1 + v_1 - r c_{21}}{s_1 + v_1 + h_1} \right), \quad BL_2 = F_2^{-1} \left( \frac{s_2 + v_2 - r c_{22}}{s_2 + v_2 + h_2} \right).\tag{24}\]

If there is only one resource available, the above order-up-to level is the highest level to which the resource can replenish the inventory of a product to. Since \( \{R_2\} \) is the cheapest resource, \( \{R_1\} \) is the second and normal production is more expensive than recovery, according to Equations (21) to (24), we have \( AL_2 > AL_1 > AL_0, BL_2 > BL_1 > BL_0 \). The replenishment procedure first chooses \( \{R_2\} \) to make products. If the returns of \( \{R_2\} \) are enough to replenish products 1 and 2 to more than \( AL_1 \) and \( BL_1 \), recovery from \( \{R_1\} \) and normal production will be unnecessary because replenishment by these two channels is not profitable. If the inventory levels of the products by recovery from \( \{R_2\} \) are between \( AL_0(BL_0) \) and \( AL_1(BL_1) \), \{R_2\} is used up, \{R_1\} will be used, since the inventory levels have not reached their order-up-to levels and normal production is avoided. If the inventory levels are still lower than \( AL_0 \) and \( BL_0 \) after \( \{R_1\} \) and \( \{R_2\} \) is used up, normal production will be needed to replenish the inventories.

Since two products are considered, if a certain channel is not able to replenish the inventories of both products to their corresponding order-up-to levels, they will compete for the resource. Therefore, in addition to the order-up-to levels, three switching levels, denoted as \( SW_1, SW_2, \) and \( RP \), are used to control the allocation of possibly limited resources between the two products. Allocation of a limited resource involves the comparison of marginal profits with respect to each channel. Returns of \( \{R_2\} \) or \( \{R_1\} \) will be assigned to the product with the highest marginal profit which means replenishment for this product is more profitable than the other one under the current situation. Thus, the inventory of the product with the highest marginal profit increases and its marginal profit is decreasing at the same time. On the other hand, the inventory of the other product remains unchanged. The recovery process continues until the two products have equal marginal profits. This is when the recovery process switches from one product to the other since the marginal profit of the product is no longer higher than the other. After this, the recovery process simultaneously allocates the resource between two products and maintains the equality of the two marginal profits until the resource is used up or the inventories reach their corresponding order-up-to levels. This type of allocation policy is defined as the fair allocation rule, which aims to balance the marginal profits of the two products during replenishment by a recovery channel. The introduction of the switching levels helps determine at which point recovery switches from one product to the other and the process of fair allocation begins.
In the following, we will illustrate how to determine the switching levels in detail by using the “replacement” concept. Suppose that the initial inventories of products 1 and 2 are below \( AL_0 \) and \( BL_0 \) and normal production is the only channel. Due to an unlimited production capacity, products 1 and 2 can be replenished to \( AL_0 \) and \( BL_0 \). When the two recovery resources become available, they will be sequentially used to replace the products, which are originally made by normal production. Since the remanufacturing cost of \( \{ R_2 \} \) is lower than \( \{ R_1 \} \), \( \{ R_2 \} \) will be used first. Details of these switching levels are further explained as in the following:

4. Switching level \( SW_2 \).

The switching level \( SW_2 \) is the inventory level of product 1 and corresponds to the inventory level \( BL_0 \) of product 2. It denotes the inventory level of product 1 at which the recovery process from group 2 switches from product 1 to product 2.

When recovery resources are available, \( \{ R_2 \} \) will first be used to replace the new items of product 1 by normal production. During this replacement process, the recovery amount of product 1 from \( \{ R_2 \} \); i.e., \( r_{21} \) is increasing and the amount by normal production—i.e., \( p_1 \)—decreases. In other words, a portion of the new items originally completed by normal production is now made by \( \{ R_2 \} \) instead. Before the new items produced by normal production are all replaced, its inventory—i.e., \( I' \)—remains unchanged. Therefore, according to Equations (17) and (21), the marginal profit with respect to \( r_{21} \) (denoted as \( \frac{\partial EP}{\partial r_{21}} \big|_{I'} \)) does not change and we have \( \frac{\partial EP}{\partial r_{21}} \big|_{I' = AL_0} = p_c - r_{c21} \).

On the other hand, the inventory of product 2—i.e., \( I'_2 \)—also stays unchanged and the marginal profit with respect to \( r_{22} \) at \( BL_0 \) (denoted as \( \frac{\partial EP}{\partial r_{22}} \big|_{I'_2 = BL_0} \)) is \( p_c - r_{c22} \). According to Equation (20), we have

\[
\frac{\partial EP}{\partial r_{21}} \big|_{I' = AL_0} > \frac{\partial EP}{\partial r_{22}} \big|_{I'_2 = BL_0}.
\]

Therefore, if \( \{ R_2 \} \) is sufficient to replace all new items of product 1, the replenishment process will continue to make product 1. \( r_{21} \) and \( I'_1 \) will increase at the same time. \( \frac{\partial EP}{\partial r_{21}} \big|_{I'_1} \) decreases and the process continues until \( I'_1 \) reaches a certain level—i.e., \( SW_2 \)—at which \( \frac{\partial EP}{\partial r_{21}} \big|_{I'_1 = SW_2} \) is equal to \( \frac{\partial EP}{\partial r_{22}} \big|_{I'_2 = BL_0} \). After that, the process will switch to replenish product 2 in order to replace the new items originally made by normal production. During this replacement process, \( r_{22} \) is increasing and \( p_2 \) decreases. \( I'_2 \) remains unchanged and \( \frac{\partial EP}{\partial r_{22}} \big|_{I'_2} \) does not change. If there are still items of \( \{ R_2 \} \) left after all new items of product 2 by normal production have been replaced, the remaining \( \{ R_2 \} \) will be simultaneously allocated between two products following the fair allocation rule.

Therefore, when the inventory levels of products 1 and 2 are at \( SW_2 \) and \( BL_0 \), respectively, the two products have equal marginal profits with respect to the recovery amount from \( \{ R_2 \} \). Thus, we have

\[
\frac{\partial EP}{\partial r_{21}} \big|_{I'_1 = SW_2} = \frac{\partial EP}{\partial r_{22}} \big|_{I'_2 = BL_0} = p_c - r_{c22} \implies SW_2\]
\[
= F^{-1}\left(\frac{s_1 + v_1 + r_{c22} - r_{c21} - pc}{s_1 + v_1 + h_1}\right).
\] (25)

5. Switching level \( SW_1 \).

The switching level \( SW_1 \) is also the inventory level of product 1, which corresponds to the inventory level \( BL_0 \) of product 2. It denotes the inventory level of product 1, at which recovery process from group 1 switches from product 1 to product 2. When the inventory levels of products 1 and 2 are at \( SW_1 \) and \( BL_0 \), respectively, the two products have equal marginal profits with respect to the recovery amount from \( \{ R_1 \} \). Thus, we have

\[
\frac{\partial EP}{\partial r_{11}} \big|_{I'_1 = SW_1} = \frac{\partial EP}{\partial r_{12}} \big|_{I'_2 = BL_0} = p_c - r_{c12} \implies SW_1\]
\[
= F^{-1}\left(\frac{s_1 + v_1 + r_{c12} - r_{c11} - pc}{s_1 + v_1 + h_1}\right).
\] (26)

6. Switching level \( RP \).

The switching level \( RP \) is the inventory level of product 1 that corresponds to the inventory level \( BL_1 \) of product 2. It denotes the inventory level of product 1 at which recovery process from group 1 switches from product 1 to product 2. When the inventory levels of products 1 and 2 are at \( RP \) and \( BL_1 \), the two products will have equal marginal profits from recovering the returns of \( \{ R_1 \} \). Thus, we have

\[
\frac{\partial EP}{\partial r_{11}} \big|_{I'_1 = RP} = \frac{\partial EP}{\partial r_{12}} \big|_{I'_2 = BL_1} = r_{c12} - r_{c21} \implies RP\]
\[
= F^{-1}\left(\frac{s_1 + v_1 + r_{c12} - r_{c21} - pc}{s_1 + v_1 + h_1}\right).
\] (27)

All threshold levels of products 1 and 2 are illustrated in Fig. 2. Order-up-to levels \( AL_1 \), \( AL_2 \), \( BL_1 \), and \( BL_2 \) are the highest replenishment levels for products 1 and 2 when recovery resources are used. \( AL_0 \) and \( BL_0 \) are the highest replenishment levels when normal production is used and they are also the lowest replenishment levels, as production capacity is unlimited. Among the order-up-to levels of product 1 (product 2), \( AL_2 \) (\( BL_2 \)) is the highest, whereas \( AL_0 \) (\( BL_0 \)) is the lowest. \( AL_1 \) (\( BL_1 \)) locates between them. \( SW_1 \), \( SW_2 \), and \( RP \) are involved to control the allocation of recovery resources. \( SW_1 \) is between \( AL_0 \) and \( AL_1 \), whereas \( RP \) is between \( AL_1 \) and \( AL_2 \). \( SW_2 \) is between \( SW_1 \) and \( RP \). In Fig. 2, \( SW_2 \) is shown to be lower than \( AL_1 \). However, \( SW_2 \) may be higher than \( AL_1 \) under other cost structures.
4. Recovery system over a finite horizon

In the course of solving the single-period problem, an optimal multilevel threshold policy, which is characterized by six order-up-to levels and three switching levels, has been obtained to help make optimal production and inventory decisions. In addition, it can guide us to address the dynamic programming model of the multi-period problem. We assume that this threshold policy is also applicable to the multi-period problem. As the last period of the multi-period problem is actually a single-period problem, we can start from the last period and take advantage of the threshold levels previously obtained. In the following, an approximate dynamic programming approach is applied and a backward induction method, which is conducted from the last period to the first period, is used to determine the threshold levels.

4.1. An approximate dynamic programming approach

Recall that Bellman’s equation of dynamic programming is

\[ Q_t(I_{1t}, I_{2t}) = \max \{ EP_t + E[Q_{t+1}(I_{1,t+1}, I_{2,t+1})] \} \quad (28) \]

We cannot expect to solve the above cost-to-go function to optimality. However, we assume that the multilevel threshold policy can also be obtained here and nine threshold levels can be determined to help make decisions for the multi-period problem.

In order to solve Equation (28), we take a first-order partial derivative of it with respect to each decision variable. For example, to decide \( p_{1t} \), its first-order partial derivatives are expressed as

\[
\frac{\partial (EP_t + E[Q_{t+1}(I_{1,t+1}, I_{2,t+1})])}{\partial p_{1t}} = \frac{\partial EP_t}{\partial p_{1t}} + \frac{\partial E[Q_{t+1}(I_{1,t+1}, I_{2,t+1})]}{\partial p_{1t}}.
\quad (29)
\]

According to Equation (17), the first term of Equation (29) can be calculated as

\[
\frac{\partial EP_t}{\partial p_{1t}} = -pc_1 + s_1 + v_1 - (s_1 + v_1 + h_1)F_{1t}(I_{1t}^{'}). \quad (30)
\]

Since there is no closed-form solution for \( E[Q_{t+1}(I_{1,t+1}, I_{2,t+1})] \), we approximate its partial derivative with respect to \( p_{1t} \) as the gradient of the inventory level at period \( t \) (denoted as \( u_{1t}(I_{1t}^{'}, I_{2t}^{'}) \)):

\[
\frac{\partial (E[Q_{t+1}(I_{1,t+1}, I_{2,t+1})])}{\partial p_{1t}} \approx u_{1t}(I_{1t}^{'}, I_{2t}^{'}). \quad (31)
\]

Let the approximation of partial derivative (29) be equal to zero. Thus, we have

\[
\frac{\partial EP_t}{\partial p_{1t}} + u_{1t}(I_{1t}^{'}, I_{2t}^{'}) = 0 \quad \Rightarrow \quad s_1 + v_1 - pc_1 - (s_1 + v_1 + h_1)F_{1t}(I_{1t}^{'}) + u_{1t}(I_{1t}^{'}, I_{2t}^{'}) = 0. \quad (32)
\]

By solving Equation (32), we can obtain the inventory level of product 1 when normal production is involved. This inventory level is called the order-up-to level, which was defined in Section 3. The order-up-to level of product 1 by normal production at period \( t \) is calculated as

\[
AL_{0t} = F_{1t}^{-1}\left( u_{1t}(I_{1t}^{'}, I_{2t}^{'}) | I_{1t} = AL_{0t}, s_1 + v_1 - pc_1 \right) \div \left( s_1 + v_1 + h_1 \right). \quad (33)
\]

Similarly, the other threshold levels of period \( t \) (i.e., \( AL_{1t}, AL_{2t}, BL_{0t}, BL_{1t}, BL_{2t}, SW_{1t}, SW_{2t}, RP_t \)) can be determined in the same way. The results are presented in Appendix A2.

It can be seen that the threshold levels of the multi-period problem are only dependent on the gradients of the cost-to-go function at the corresponding threshold levels after approximation. Therefore, unlike usual approaches that represent the cost-to-go function by a single function (or piecewise function) across the whole state space, we only need to estimate the gradients at the points of interest. Hence, the performance of the solution is not affected by the function we assume, which can pose a challenge for most approximate dynamic programming approaches.

If the cost-to-go function is separable with the inventory levels of the two products, the gradient \( u_{1t} \) is independent of \( I_{2t} \). However, this is not always possible. We have performed some numerical experiments on a number of samples and the results show that the position of \( I_{2t} \) does not have a major impact on \( u_{1t} \). Therefore, we use the average of the gradients at the three order-up-to levels of product 2 as the approximation of \( u_{1t} \) at \( AL_{0t} \). Thus, in our approach, \( u_{1t} \) is estimated as

\[
u_{1t}(I_{1t}^{'}, I_{2t}^{'}) | I_{1t} = AL_{0t} \approx \frac{1}{3} \sum_{l=0}^{2} u_{1t}(I_{1t}^{'}, I_{2t}^{'}) | I_{1t} = AL_{0t}, E_{2t} = BL_{0t}. \quad (34)
\]

In addition, to determine a threshold level, say \( AL_{0t} \), the gradient at \( AL_{0t} \) needs to be obtained first. However, the gradient at \( AL_{0t} \) can be computed only after \( AL_{0t} \) is
decided. Therefore, an iterative algorithm is employed to search for the threshold levels starting with the predetermined threshold levels of period $t+1$. The threshold levels of $t+1$ are obtained in advance through a backward induction approach that is elaborated in next subsection.

Initially, $u_{11}$ is obtained by taking predetermined threshold levels of $t+1$ as initial values; i.e., $I'_{11} = AL_{0,t+1}$, $I''_{11} = BL_{t,t+1} (l = 0, 1, 2)$. $u_{11}$ is then smoothed with the previous value to get a weighted average gradient, which is used to calculate the corresponding threshold level. After that, the newest threshold level is also smoothed by certain rule to obtain a weighted average. The procedure is repeated until it converges. Due to the time-consuming computation, the algorithm generally stops if the total absolute change of the threshold level over a certain number of iterations is small.

Due to the time-consuming computation, the algorithm generally stops if the total absolute change of the threshold level over a certain number of iterations is small. For example, if $\sum_{i=m-L+1,m+L} || AL_{0t} - AL_{1t} || < \delta (L$: the number of consecutive results used for calculating the absolute change; $\delta$: a small number), the algorithm steps. In addition, as 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. Thus, we need the averaging steps—i.e., Steps 4 and 6—to help stabilize the solution. According to Gupal and Bazhenov (1972), the solution converges to an optimal solution under certain conditions on the step sizes $\alpha_m$ and $\beta_m$, such as

$$\alpha_m \geq 0, \beta_m \geq 0, \alpha_m/\beta_m \rightarrow 0, \Sigma_{m=1}^{\infty} \alpha_m = \infty,$$

$$\Sigma_{m=1}^{\infty} \beta_m^2 < \infty. \quad (35)$$

The procedure for determining $AL_{0t}$ is summarized below. All the other order-up-to levels can be obtained in a similar way:

**Algorithm 1** Procedure of determining order-up-to levels.

1. **Step 1.** Let $m$ be the index for an iteration, $u_{m}^{\prime\prime}$ be a stochastic gradient obtained at iteration $m$, $\bar{u}_{m}^{\prime\prime}$ be a weighted average of gradient up to iteration $m$, $AL_{0m}^{\prime\prime}$ be the threshold level corresponding to a weighted average of gradient at iteration $m$, and $AL_{1m}^{\prime\prime}$ be a weighted average of the threshold level at iteration $m$.

2. **Step 2.** Set $m=1; \bar{u}_{11}^{\prime\prime} = u_{11,t+1}$, $AL_{01}^{\prime\prime} = AL_{0,t+1}$.

3. **Step 3.** Obtain $u_{11}^{\prime\prime}$ by Equation (34).

4. **Step 4.** Update $\bar{u}_{11}^{\prime\prime}$ by

$$\bar{u}_{11}^{\prime\prime} = \bar{u}_{11}^{\prime\prime-1} + \beta_m (u_{11}^{\prime\prime} - \bar{u}_{11}^{\prime\prime-1}). \quad (36)$$

5. **Step 5.** Obtain $AL_{01}^{\prime\prime}$ by

$$AL_{01}^{\prime\prime} = F_{11}^{\prime\prime} \left( \frac{\bar{u}_{11}^{\prime\prime} + s_1 + v_1 + pc_1}{s_1 + v_1 + h_1} \right). \quad (37)$$

6. **Step 6.** Update $AL_{11}^{\prime\prime}$ by

$$AL_{11}^{\prime\prime} = AL_{01}^{\prime\prime-1} + \alpha_m (AL_{11}^{\prime\prime} - AL_{01}^{\prime\prime-1}). \quad (38)$$

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

### 4.2. Determination of gradients in the multi-period context

To solve the multi-period problem, we start from period $T$ and then take advantage of a backward induction approach to determine the threshold levels of each period from the second-to-last period; i.e., period $T - 1$ until period 1. Since the threshold levels of each period, except period $T$, are actually dependent on their corresponding gradients, they can be obtained by determining the gradients at points of interest through the backward induction approach from period $T - 1$ to period 1.

Threshold levels of period $T$ are obtained by solving a single-period problem. Then, for period $T - 1$, the threshold levels can be determined after the gradients of the optimal cost-to-go function—i.e., the gradients of the maximum expected profit of period $T$—have been estimated. Suppose we are now at period $t$. Up to now, the threshold levels from period $t+1$ until $T$ have been decided and replenishment decisions have been made. In order to determine its threshold levels, we need to estimate the gradients of the optimal cost-to-go function; i.e., the gradients of maximum expected total profit earned from period $t+1$ until $T$. After that, replenishment decisions are made under the optimal policy and the algorithm moves to $t - 1$.

The gradients $u_{jt}$ are estimated by Monte Carlo simulation. $N$ sets of random realizations of stochastic returns and demands in each period from $t$ until $T$ are generated. By looking into the realization of the return at each period, we can tell which one out of the 21 cases presented in Appendix A1 it matches and then calculate the derivatives of the optimal replenishment decisions as well as optimal cost-to-go function of period $t$ together with the realization of the demand. The value of the optimal cost-to-go function of sample $k$ is obtained by summing the profit of the realization from period $t+1$ until period $T - 1$ after applying the optimal policy for these periods and the expected profit of period $T$. Let $P^*_{k\tau}$ denote the maximum profit of period $\tau$ of sample $k$, which is $(\eta X^+ := \max \{X, 0\})$:

$$P_{k\tau}^* = s_1 (I_{1\tau} + p_{1\tau}^* + r_{1\tau}^* + r_{2\tau}^*) + s_2 (I_{2\tau} + p_{2\tau}^* + r_{1\tau}^* + r_{2\tau}^* + r_{2\tau}^* + r_{2\tau}^*)$$

$$- (p_{c1} p_{1\tau}^* + p_{c2} p_{2\tau}^* + r_{c1} r_{1\tau}^* + r_{c2} r_{1\tau}^* + r_{c2} r_{2\tau}^* + r_{c2} r_{2\tau}^* + r_{c2} r_{2\tau}^*)$$

$$- (h_1 + s_1) [I_{1\tau} + p_{1\tau}^* + r_{1\tau}^* + r_{2\tau}^* - D_{1\tau}^k]^+$$

$$- v_1 [D_{1\tau}^k - (I_{1\tau} + p_{1\tau}^* + r_{1\tau}^* + r_{2\tau}^*)]^+$$

$$- (h_2 + s_2) [I_{2\tau} + p_{2\tau}^* + r_{1\tau}^* + r_{2\tau}^* + r_{2\tau}^* - D_{2\tau}^k]^+$$

$$- v_2 [D_{2\tau}^k - (I_{2\tau} + p_{2\tau}^* + r_{1\tau}^* + r_{2\tau}^* + r_{2\tau}^*)]^+. \quad (39)$$

Let $EP_k^*\tau$ denote the maximum expected profit of period $T$ of sample $k$. Thus, the gradient with respect to $p_{1\tau}$ of...
sample \( k \) (denoted as \( \text{grad}^k_i'(I_{t+1}, I_t) \)) is expressed as

\[
\text{grad}^k_i'(I_{t+1}, I_t) = \sum_{j=1}^{2} \frac{\partial (\sum_{t=1}^{T-1} p_{k,j} + EP_{k,j}^T)}{\partial I_{j,t}} \partial I_{j,t}.
\]

\[
= \sum_{j=1}^{2} \left( \frac{\partial p_{k,j+1}}{\partial I_{j,t}} + \frac{\partial (\sum_{t=2}^{T-1} p_{k,j} + EP_{k,j+1}^T)}{\partial I_{j,t}} \right) \partial I_{j,t}.
\]

\[
= \sum_{j=1}^{2} \left( \frac{\partial p_{k,j+1}}{\partial I_{j,t}} + \frac{\partial r_{t+1}}{\partial I_{j,t}} \cdot \text{grad}^k_i(I_{t+1}, I_t) \right) \partial I_{j,t} + \frac{\partial I_{t+1}^*}{\partial I_{j,t}} \text{grad}^k_i(I_{t+1}, I_t),
\]

\[
\partial I_{j,t+1} = \frac{\partial I_{t+1}^*}{\partial I_{j,t}} \partial I_{t+1} + \frac{\partial p_{t+1}}{\partial I_{j,t}} \partial I_{j,t} + \frac{\partial r_{t+1}}{\partial I_{j,t}} \partial I_{j,t} + \frac{\partial I_{t+1}^*}{\partial I_{j,t}} \text{grad}^k_i(I_{t+1}, I_t) \partial I_{t+1} + \frac{\partial I_{t+1}^*}{\partial I_{j,t}} \text{grad}^k_i(I_{t+1}, I_t), \quad (41)
\]

Partial derivatives of the optimal replenishment decisions with respect to the initial inventory are involved in the calculation of Equation (41). Therefore, IPA is applied to compute the derivatives. The perturbation of the initial inventory of product \( j \) of period \( t + 1 \) is propagated to the final inventory of product \( i \) in the same period \((i, j = 1, 2; i \neq j)\):

\[
\frac{\partial I_{t+1}^*}{\partial I_{j,t+1}} = \frac{\partial I_{t+1}}{\partial I_{j,t+1}} + \frac{\partial p_{t+1}}{\partial I_{j,t+1}} + \frac{\partial r_{t+1}}{\partial I_{j,t+1}} + \frac{\partial I_{t+1}^*}{\partial I_{j,t+1}} \cdot \text{grad}^k_i(I_{t+1}, I_t). \quad (42)
\]

Suppose the current condition at period \( t + 1 \) matches case S7 in Appendix A1, which is

\[
R_{t+1}^1 + I_{t+1}^1 < SW_{t+1}^1, \quad R_{t+1}^2 + I_{t+1}^2 + I_{t+1}^1 > SW_{t+1}^1.
\]

\[
R_{t+1}^1 + R_{t+1}^2 + I_{t+1}^2 + I_{t+1}^1 \leq SW_{t+1}^1 + BL_{0,t+1}. \quad (43)
\]

Then, the optimal replenishment decisions for sample \( k \) are as follows:

\[
p_{t+1}^{1} = 0, \quad r_{t+1}^{1} = SW_{t+1} - R_{t+1}^1 - I_{t+1}^1, \quad r_{t+1}^{21} = R_{t+1}^2 - I_{t+1}^1 - I_{t+1}^2,
\]

\[
p_{t+1}^{2} = SW_{t+1} + BL_{0,t+1} - R_{t+1}^2 - I_{t+1}^1 - I_{t+1}^2, \quad r_{t+1}^{22} = R_{t+1}^2 - I_{t+1}^1 + I_{t+1}^2 - SW_{t+1} + BL_{0,t+1} = 0,
\]

and

\[
I_{t+1}^{1*} = SW_{t+1} + I_{t+1}^1 = BL_{0,t+1}. \quad (44)
\]

We can tell from Equation (44) that the inventory levels of the two products after replenishment have reached the threshold levels \( SW_{t+1} \) and \( BL_{0,t+1} \), respectively. Therefore, the perturbation on the initial inventory of the two products is not propagated to the order-up-to levels of the two products; that is,

\[
\frac{\partial I_{t+1}^{1*}}{\partial I_{t+1}} = \frac{\partial I_{t+1}^{2*}}{\partial I_{t+1}} = \frac{\partial I_{t+1}^{2*}}{\partial I_{t+1}} = \frac{\partial I_{t+1}^{2*}}{\partial I_{t+1}} = 0.
\]

However, the perturbation affects the related replenishment decisions and the impact can be determined as

\[
\frac{\partial p_{t+1}^{2*}}{\partial I_{t+1}} = -1, \quad \frac{\partial r_{t+1}^{2*}}{\partial I_{t+1}} = 1.
\]

Thus, we can conclude that if the initial inventory of product 1 is increased by a small amount \( \Delta \), the same amount of normal production for product 2 will be saved and the recovery amount from \( R_i \) for product 2 will be increased by the same amount \( \Delta \). Despite the impact on related replenishment decisions, the order-up-to levels of each product remain unaffected.

### 5. Numerical experiments

Extensive numerical experiments are conducted to validate the effectiveness of the multilevel threshold policy. Our approach to solve the two-product recovery problem over a finite horizon is run on a PC (Intel Core i2 Duo, 2.66 GHz; Memory, 4G). Replications of stochastic demands and returns are generated by assuming both of demands and returns follow normal distributions in our experiments. Let \( E(X) \) and \( StDev(X) \) denote the expectation and standard deviation of \( X \), respectively. Note that while deriving the threshold levels of each period we do not need to assume the stochastic demands and returns follow any distribution.

| Parameters | \( h_1 \) | \( h_2 \) | \( u_1 \) | \( u_2 \) | \( s_1 \) | \( s_2 \) | \( p_{c1} \) | \( p_{c2} \) | \( r_{c11} \) | \( r_{c12} \) | \( r_{c21} \) | \( r_{c22} \) |
|------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| Set 1      | 3        | 3        | 4        | 6        | 15       | 20       | 12       | 15       | 6         | 10        | 2         | 7         |
| Set 2      | 3        | 3        | 4        | 6        | 18       | 20       | 16       | 15       | 10        | 10        | 2         | 7         |
| Set 3      | 3        | 3        | 4        | 6        | 21       | 20       | 15       | 14       | 10        | 2         | 7         |
Table 3. Threshold levels with different scenarios of $R_1$ with parameter set 1

| Scenario | $E(R_1)$ | StDev$(R_1)$ | $AL_0$ | $AL_1$ | $AL_2$ | $SW_0$ | $SW_1$ | $SW_2$ | $BL_0$ | $BL_1$ | $BL_2$ |
|----------|----------|--------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 15       | 5        | 229.1        | 500.3  | 761.4  | 245.9  | 267.8  | 560.7  | 122.1  | 228.3  | 332.1  |
| 30       | 10       | 228.6        | 485.6  | 747.6  | 245.3  | 266.9  | 544.4  | 121.5  | 224.3  | 328.4  |
| 60       | 20       | 227.8        | 467.6  | 713.2  | 243.8  | 263.4  | 511.9  | 120.5  | 213.1  | 319.6  |
| 90       | 30       | 225.3        | 418.2  | 664.4  | 240.6  | 257.7  | 469.0  | 119.5  | 203.9  | 310.5  |
| 120      | 40       | 221.1        | 376.6  | 613.1  | 234.8  | 251.1  | 436.4  | 118.1  | 195.9  | 299.2  |
| 150      | 50       | 216.1        | 340.8  | 566.1  | 228.8  | 242.1  | 417.7  | 116.2  | 183.6  | 280.6  |
| 180      | 60       | 210.6        | 308.7  | 529.8  | 222.1  | 234.2  | 375.2  | 114.1  | 169.4  | 257.2  |
| 210      | 70       | 205.4        | 287.7  | 503.9  | 216.0  | 227.2  | 317.1  | 112.3  | 154.9  | 235.8  |

We perform some pre-experiments to verify the convergence of the threshold levels obtained from our solution approach. Once the differences of all of the threshold levels between two consecutive periods are no more than 1%, convergence is assumed to have been achieved. Our pre-experiments show that under a certain cost structure, the threshold levels converge faster when the holding cost is higher. Therefore, in practice, the converged threshold levels can be applied while making production and recovery decisions of each period over a relatively long finite horizon.

5.1. Impact of stochastic returns and demands

We investigate the impact of stochastic returns and demands on the threshold levels, which have converged in the multiperiod context. We first assess how the expected value of returned items affects the threshold levels. Second, we evaluate the impact of demand variability on the threshold levels. Experiments are conducted on three sets of system parameters, which are shown in Table 2.

5.1.1. Impact of stochastic returns on threshold levels

Eight instances are generated and tested for each set of system parameters. Among the eight instances, demand for products 1 and 2 and returns of product 2 are the same. They are generated according to the following parameters: $E(D_1) = 200$, $StDev(D_1) = 60$; $E(D_2) = 100$, $StDev(D_2) = 30$; $E(R_2) = 45$, $StDev(R_2) = 15$. Furthermore, eight sets of returns of group 1, one for each of the eight instances, are generated with $E(R_1)$ varying from 15 to 210 and $StDev(R_1)$ from five to 70, respectively.

Threshold levels for the two products obtained on three test sets are shown in Tables 3, 4, and 5, where the first and second columns present eight scenarios of the returns of group 1 and the next nine columns report the converged threshold levels of each scenario. Furthermore, the trends of the threshold levels with each parameter set are illustrated in Figs. 3, 4, and 5. The results show that for the three parameter sets, all threshold levels decrease with the increasing expected value of $R_1$. For product 1, the threshold levels $AL_1$, $AL_2$, and $RP$ decrease faster than other threshold levels. And for product 2, the threshold levels $BL_1$ and $BL_2$ decrease faster than $BL_0$. Threshold levels of each period decrease when there are more returns available. As $R_1$ is involved in recovery process, it affects the threshold levels related to recovery; i.e., $AL_1$, $AL_2$, $BL_1$, and $BL_2$. On the other hand, $AL_0$ and $BL_0$ are not affected by $R_1$. Consequently, $R_1$ has a significant impact on $RP$, which is obtained by the comparison of marginal profits of the recovery using $R_2$ while the inventory of product 2 is at the threshold level $BL_1$. However, $SW_1$ and $SW_2$ are obtained by the comparison of marginal profits of the recovery using $R_1$ and $R_2$ while the...
inventory of product 2 at the threshold level $BL_0$ is only slightly impacted by $R_1$.

5.1.2. Impact of demand variability on threshold levels

Six instances are generated and tested for each set of parameters. Among the six instances, demand of product 2 and returns of products 1 and 2 are the same. They are generated according to the following parameters: $E(D_2) = 100$, $StDev(D_2) = 30$; $E(R_1) = 90$, $StDev(R_1) = 30$; $E(R_2) = 45$, $StDev(R_2) = 15$. Furthermore, six sets of the demand for product 1, one for each of the six instances, are generated with $StDev(D_1)$ varying from 20 to 200 and the coefficient of variation of $D_1$ (denoted as $CV_1$) from 0.1 to 1.0, respectively, where $E(D_1)$ is fixed at 200.

Threshold levels of the two products obtained on three test sets are shown in Tables 6, 7, and 8, where the first and second columns present six scenarios of the demand for product 1 and the next nine columns report the converged threshold levels of each scenario. Furthermore, the trends of the threshold levels with each test set are illustrated in Figs. 6, 7, and 8. The results show that for the three test sets, all threshold levels related to product 1 increase with the increasing demand variability of product 1, whereas the threshold levels related to product 2 stay unaffected. As the demands for the two products are independent of each other, the impact of demand variability of product 1 will only affect the threshold levels related to product 1. Furthermore, higher demand variability results in higher threshold levels, so that possible stock shortage can be avoided. Note that in some cases the threshold level does not follow exact increasing trend; e.g., $AL_0$ in Fig. 8. This is because when the demand variability reaches a very high level, there is high chance that values of some demand realizations may be negative and thus need to be adjusted to zero to ensure the meaningfulness of the experiments. As a result, the expected demand increases but the demand variability drops and may be less than its previous scenario. Then, the corresponding threshold level may be lower than the previous one.

5.2. Comparison with other heuristic policies

To assess the benefit of our proposed approach, we compare it with two heuristic methods, both of which consider each period of the multi-period problem as a series of single-period problems and solve them to optimality period by period. The first threshold policy, denoted as $H_1$, is derived from the first method and disregards scrap value
Optimal control for a two-product recovery system

Fig. 4. Trends of the threshold levels with different scenarios of $R_1$ with parameter set 2.

Fig. 5. Trends of the threshold levels with different scenarios of $R_1$ with parameter set 3.

Table 6. Threshold levels with different scenarios of $D_1$ with parameter set 1

| Scenario | Std Dev($D_1$) | $CV_1$ | $AL_0$ | $AL_1$ | $AL_2$ | $SW_0$ | $SW_1$ | $SW_2$ | $BL_0$ | $BL_1$ | $BL_2$ |
|----------|----------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
|          | 20             | 0.1    | 210.0  | 378.7  | 625.5  | 216.1  | 223.9  | 429.1  | 120.2  | 203.0  | 312.9  |
|          | 40             | 0.2    | 218.6  | 399.6  | 644.8  | 229.6  | 242.8  | 447.7  | 120.0  | 205.6  | 312.2  |
|          | 60             | 0.3    | 225.3  | 418.2  | 664.4  | 240.6  | 257.7  | 469.0  | 119.5  | 203.9  | 310.5  |
|          | 100            | 0.5    | 236.7  | 470.4  | 734.9  | 260.1  | 287.2  | 532.6  | 118.9  | 202.4  | 310.2  |
|          | 150            | 0.75   | 243.6  | 577.3  | 892.2  | 281.9  | 324.0  | 658.0  | 118.8  | 206.8  | 314.6  |
|          | 200            | 1.0    | 246.2  | 708.4  | 1073.8 | 305.2  | 370.5  | 803.7  | 118.8  | 207.0  | 317.1  |

Table 7. Threshold levels with different scenarios of $D_1$ with parameter set 2

| Scenario | Std Dev($D_1$) | $CV_1$ | $AL_0$ | $AL_1$ | $AL_2$ | $SW_0$ | $SW_1$ | $SW_2$ | $BL_0$ | $BL_1$ | $BL_2$ |
|----------|----------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
|          | 20             | 0.1    | 208.1  | 408.0  | 851.5  | 214.6  | 408.3  | 666.7  | 119.4  | 207.3  | 301.9  |
|          | 40             | 0.2    | 214.8  | 409.0  | 869.7  | 226.2  | 408.9  | 686.0  | 119.4  | 205.2  | 307.2  |
|          | 60             | 0.3    | 219.3  | 417.1  | 883.5  | 235.1  | 417.1  | 700.0  | 118.9  | 203.8  | 305.1  |
|          | 100            | 0.5    | 225.9  | 430.7  | 907.9  | 249.2  | 430.3  | 723.0  | 118.7  | 202.1  | 300.6  |
|          | 150            | 0.75   | 227.1  | 444.4  | 926.3  | 256.1  | 445.98 | 746.4  | 119.3  | 201.0  | 302.3  |
|          | 200            | 1.0    | 221.6  | 457.0  | 935.6  | 257.0  | 455.9  | 749.6  | 120.6  | 206.4  | 305.8  |
Table 8. Threshold levels with different scenarios of $D_1$ with parameter set 3

| Scenario | Std Dev($D_1$) | $CV_1$ | $AL_0$ | $AL_1$ | $AL_2$ | $SW_0$ | $SW_1$ | $SW_2$ | $BL_0$ | $BL_1$ | $BL_2$ |
|----------|----------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 20       | 0.1            | 205.7  | 409.5  | 1060.1 |        | 212.3  | 618.4  | 881.8  | 116.8  | 207.3  | 293.3  |
| 40       | 0.2            | 210.4  | 409.3  | 1079.9 |        | 221.9  | 630.6  | 902.3  | 117.8  | 204.5  | 297.2  |
| 60       | 0.3            | 212.5  | 418.9  | 1091.6 |        | 228.7  | 642.2  | 915.8  | 117.8  | 204.4  | 297.3  |
| 100      | 0.5            | 215.1  | 418.4  | 1089.2 |        | 237.9  | 649.5  | 918.7  | 118.3  | 201.7  | 296.5  |
| 150      | 0.75           | 209.1  | 420.0  | 1086.0 |        | 238.5  | 654.7  | 921.6  | 119.2  | 200.7  | 301.8  |
| 200      | 1.0            | 198.2  | 424.4  | 1081.9 |        | 231.7  | 631.5  | 918.2  | 120.8  | 203.7  | 304.4  |

of remaining products, and the second threshold policy, denoted as $H_2$, is derived from the second method and assumes the scrap value of remaining products 1 and 2 to be $r_{c21}$ and $r_{c22}$, respectively. We denote our threshold policy as $H_3$ and average expected profit over all periods of each method as $AE_{PH_1}$, $AE_{PH_2}$, and $AE_{PH_3}$, respectively.

The performance of the three policies is experimentally compared in the last scenario of Table 3, where $E(R_1) = 210$ and $StDev(R_1) = 70$. The threshold levels of the three policies are shown in Table 9. It can be seen that the threshold levels of $H_3$ are the highest, whereas the threshold levels of the $H_1$ are the lowest. The difference in corresponding threshold levels between $H_1$ and $H_2$ is small, whereas the difference between $H_1$ and $H_3$ is relatively large.

The three methods help make production and recovery decisions of the two-product recovery system in a relatively long horizon, and average expected profit over all the periods is also calculated and reported in Table 10.
Fig. 8. Trends of the threshold levels with different scenarios of $D_t$ with parameter set 3.

Table 9. Threshold levels of three policies

|          | $AL_0$ | $AL_1$ | $AL_2$ | $SW_0$ | $SW_1$ | $SW_2$ | $BL_0$ | $BL_1$ | $BL_2$ |
|----------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| $H_1$    | 171.6  | 213.8  | 244.9  | 179.1  | 186.2  | 220.9  | 90.8   | 103.9  | 112.0  |
| $H_2$    | 176.9  | 223.1  | 262.2  | 184.8  | 192.5  | 231.5  | 100.0  | 118.1  | 132.9  |
| $H_3$    | 205.4  | 287.7  | 503.9  | 216.0  | 227.2  | 317.1  | 112.3  | 154.9  | 235.8  |

Our method outperforms $H_1$ by 7.19% and $H_2$ by 3.35%. The threshold levels obtained by our approach perform the best among three policies. Since the optimal solution to the multi-period two-product recovery problem is difficult to obtain, our approach provides a promising way to obtain a good solution.

Table 10. Average expected profit of different policies

|          | $AEP_{H_1}$ | $AEP_{H_2}$ | $AEP_{H_3}$ |
|----------|-------------|-------------|-------------|
|          | 2115.3      | 2186.2      | 2267.4      |

6. Conclusions

In this study, we consider a two-product recovery system over a finite horizon. Based on the multilevel threshold policy derived from solving the single-period problem, an approximate dynamic programming approach is proposed to solve the multi-period problem, which only needs to estimate the gradients of the cost-to-go function at points of interest and frees us from assuming a form for the cost-to-go function. Numerical experiments are conducted to validate the proposed approach. Compared with existing researches in related areas, this work contributes the following:

1. Our article considers production planning and inventory control of a recovery system with two products and two respective return flows, which, as far as we know, is the first study on recovery systems involving multiple products and multiple returns.
2. A single-period two-product recovery problem is solved to optimality and a multilevel threshold policy is derived to help make production and inventory decisions. This threshold policy serves as a good starting point for solving the multi-period problem.
3. An approximate dynamic programming approach is used to address the multi-period two-product recovery problem. We show that the threshold levels of each period only depend on the gradients of cost-to-go function at the points of interest through a simple approximation. Therefore, we do not need to assume any explicit form of the cost-to-go function, which can be a challenge in most approximate dynamic programming approaches.
4. To solve the multi-period problem, it is shown that the gradients of each period can be sequentially calculated period by period through a backward induction approach starting from the last period and then moving backwards to the first period.

We note that the multilevel threshold policy and the proposed solution approach are also applicable to general multi-product recovery systems. Detailed discussion is provided in Appendix A3. However, there are limitations of the current study. First, we assume that there is no setup cost and backorders are not allowed, both of which are commonly seen in industry. Also, due to the high complexity of the studied multi-period two-product recovery system, the sub-optimality of the proposed approach needs to be explored in the future.
References

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

Aras, N., Boyaci, T., and Verter, V. (2004) The effect of categorizing returned products in remanufacturing. IIE Transactions, 35(4), 319–331.

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

Blackburn, J.D., Guide, V.D., Souza, G.C. and Van Wassenhove, L.N. (2004) Reverse supply chains for commercial returns. California Management Review, 46(2), 5–22.

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

DeCroix, G., Song, J. and Zipkin, P. (2005) A series system with returns: stationary analysis. Operations Research, 53, 350–362.

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

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

Gupta, A.M. and Bazhenov, L.G. (1972) A stochastic method of linearization. Cybernetics, 8, 482–484.

Ilgin, M.A. and 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 Spectrum, 19(2), 111–122.

Inderfurth, K., De Kok, A.G. and 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. and 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.

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

Kleber, R., Minner, S. and 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. and Ko, C.S. (2002) An optimal ordering and recovery policy for reusable items. Computers and Industrial Engineering, 43, 59–73.

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

Krumwiede, D.W. and Sheu, C. (2002) A model for reverse logistics entry by third-party providers. Omega, 30, 325–333.

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

Pokharel, S. and Mutha, A. (2009) Perspectives in reverse logistics: a review. Resources, Conservation and Recycling, 53, 175–182.

Prahinski, C. and Kocabasoglu, C. (2006) Empirical research opportunities in reverse supply chains. Omega, 34, 519–532.

Richter, K. (1994) An EOQ repair and waste disposal model, in Proceedings of the Eighth International Working Seminar on Production Economics, pp. 83–91, Innsbruck, Austria, 21–25 February 1994. Elsevier, Amsterdam, The Netherlands.

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

Savaskan, R.C., Bhattacharya, S. and van Wassenhove, L.N. (2004) Closed-loop supply chain models with product remanufacturing. Management Science, 50(2), 239–252.

Schadry, D.A. (1967) A deterministic inventory model for repairable items. Naval Research Logistics Quarterly, 14, 391–398.

Serrato, M.A., Ryan, S.M. and Gaytán, J. (2007) A Markov decision model to evaluate outsourcing in reverse logistics. International Journal of Production Research, 45(18–19), 4289–4315.

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

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

Thierry, M.C., Salomon, M., van Nunen, J.A.E.E. and van Wassenhove, L.N. (1995) Strategic issues in product recovery management. California Management Review, 37, 114–135.

Van der Laan, E.A., Salomon, M. and Dekker, R. (1999) 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., and Dekker, R. and van Wassenhove, L.N. (1999) Inventory control in hybrid systems with remanufacturing. Management Science, 45, 733–747.

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

Wojanowski, R., Verter, V. and Boyaci, T. (2007) Retail-collection network design under deposit-refund. Computers & Operations Research, 34(2), 324–345.

Yuan, X. and Cheung, K.L. (1998) Modeling returns of merchandise in an inventory system. OR Spectrum, 20, 147–154.

Zhou, S.X., Tao, Z. and Chao, X. (2001) Optimal control of inventory systems with multiple types of remanufacturable products. Manufacturing & Service Operations Management, 13(1), 20–34.

Appendix

A1: Cases of the single-period problem

Details of the 21 cases derived from the combinations of different initial inventory and return for single-period problem are listed in Table A1. The first row of each case shows the specific condition and the second row presents the corresponding production and recovery decisions. The last column of each case points out the locations of the inventory levels of product 1 and product 2 after replenishment.
Table A1. The 21 cases derived from the combinations of different initial inventory and returns for the single-period problem

| Index | Conditions |
|-------|-------------|
| S1    | $R_1 + R_2 + I_1 < AL_0$, $I_2 < BL_0$  
$p_1^* = AL_0 - R_1 - R_2 - I_1$, $r_{11}^* = R_1$, $r_{21}^* = R_2$  
$p_2^* = BL_0 - I_2$, $r_{12}^* = 0$, $r_{22}^* = 0$  
$I_1^* = AL_0$  
$I_2^* = BL_0$ |
| S2    | $AL_0 \leq R_1 + R_2 + I_1 \leq SW_1$, $I_2 < BL_0$  
$p_1^* = 0$, $r_{11}^* = R_1$, $r_{21}^* = R_2$  
$p_2^* = BL_0 - I_2$, $r_{12}^* = 0$, $r_{22}^* = 0$  
$AL_0 \leq I_1^* \leq SW_1$  
$I_2^* = BL_0$ |
| S3    | $R_1 + R_2 + I_1 < AL_0$, $I_2 \geq BL_0$  
$p_1^* = AL_0 - R_1 - R_2 - I_1$, $r_{11}^* = R_1$, $r_{21}^* = R_2$  
$p_2^* = 0$, $r_{12}^* = 0$, $r_{22}^* = 0$  
$I_1^* = AL_0$ |
| S4    | $AL_0 \leq R_1 + R_2 + I_1 < AL_1$, $I_2 \geq BL_0$, $\frac{\partial EP}{\partial r_{11}} |_{I_1=R_1+I_1} = \frac{\partial EP}{\partial r_{12}} |_{I_2=I_2}$  
$p_1^* = 0$, $r_{11}^* = R_1$, $r_{21}^* = R_2$  
$p_2^* = 0$, $r_{12}^* = 0$, $r_{22}^* = 0$  
$AL_0 \leq I_1^* \leq AL_1$  
$BL_0 \leq I_2^* \leq BL_2$ |
| S5    | $R_1 + R_2 + I_1 > AL_1$, $R_2 + I_1 \leq AL_1$, $I_2 \geq BL_1$  
$p_1^* = 0$, $r_{11}^* = AL_1 - R_2 - I_1$, $r_{21}^* = R_2$  
$p_2^* = 0$, $r_{12}^* = 0$, $r_{22}^* = 0$  
$BL_1 \leq I_1^* \leq BL_2$ |
| S6    | $R_2 + I_1 > AL_1$, $I_2 \geq BL_1$, $\frac{\partial EP}{\partial r_{11}} |_{I_1=R_1+I_1} = \frac{\partial EP}{\partial r_{12}} |_{I_2=I_2}$  
$p_1^* = 0$, $r_{11}^* = 0$, $r_{21}^* = R_2$  
$p_2^* = 0$, $r_{12}^* = 0$, $r_{22}^* = 0$  
$AL_1 < I_1^* \leq AL_2$  
$BL_1 \leq I_2^* \leq BL_2$ |
| S7    | $R_2 < SW_1$, $R_1 + R_2 + I_1 > SW_1$, $R_2 + R_1 + I_2 \leq SW_1 + BL_0$  
$p_1^* = 0$, $r_{11}^* = SW_1 - R_2 - I_1$, $r_{21}^* = R_2$  
$p_2^* = SW_1 + BL_0 - R_2 - R_1 - I_2$, $r_{12}^* = R_2 + R_1 + I_1 - SW_1$, $r_{22}^* = 0$  
$I_1^* = SW_1$  
$I_2^* = BL_0$ |
| S8    | $SW_1 \leq R_2 + I_1 \leq SW_2$, $R_1 + I_2 < BL_0$  
$p_1^* = 0$, $r_{11}^* = 0$, $r_{21}^* = R_2$  
$p_2^* = BL_0 - R_1 - I_2$, $r_{12}^* = R_2$, $r_{22}^* = 0$  
$SW_1 \leq I_1^* \leq SW_2$  
$I_2^* = BL_0$ |
| S9    | $BL_0 \leq R_1 + I_2 \leq BL_1$, $\frac{\partial EP}{\partial r_{11}} |_{I_1=R_1+I_1} > \frac{\partial EP}{\partial r_{12}} |_{I_2=I_2}$  
$p_1^* = 0$, $r_{11}^* = 0$, $r_{21}^* = R_2$  
$p_2^* = 0$, $r_{12}^* = R_2$, $r_{22}^* = 0$  
$SW_1 \leq I_1^* \leq RP$  
$BL_0 \leq I_2^* \leq BL_1$ |
| S10   | $AL_1 \leq R_1 + I_1 \leq RP$, $I_2 < BL_1$, $R_1 + I_2 > BL_1$  
$p_1^* = 0$, $r_{11}^* = 0$, $r_{12}^* = R_2$  
$p_2^* = 0$, $r_{12}^* = BL_1 - I_2$, $r_{22}^* = 0$  
$AL_1 \leq I_1^* \leq RP$  
$I_2^* = BL_1$ |
| S11   | $SW_1 + BL_0 < R_1 + R_2 + I_1 + I_2 \leq AL_1 + BL_1$  
$\frac{\partial EP}{\partial r_{11}} |_{I_1=R_1+I_1} < \frac{\partial EP}{\partial r_{12}} |_{I_2=I_2}$  
$\frac{\partial EP}{\partial r_{11}} |_{I_1=R_1+I_1} > \frac{\partial EP}{\partial r_{12}} |_{I_2=I_2}$  
$p_1^* = 0$, $r_{21}^* = R_2$, $p_2^* = 0$, $r_{12}^* = 0$, solve $r_{11}^* + r_{12}^* = R_1$  
and $\frac{\partial EP}{\partial r_{11}} |_{I_1=R_1+I_1} < \frac{\partial EP}{\partial r_{12}} |_{I_2=I_2}$  
$SW_1 < I_1^* \leq AL_1$  
$BL_0 < I_2^* \leq BL_1$ |
| S12   | $R_2 + I_1 < AL_1$, $I_2 < BL_1$, $R_1 + R_2 + I_1 + I_2 > AL_1 + BL_1$  
$p_1^* = 0$, $r_{11}^* = AL_1 - R_2 - I_1$, $r_{21}^* = R_2$  
$p_2^* = 0$, $r_{12}^* = BL_1 - I_2$, $r_{22}^* = 0$  
$I_1^* = AL_1$  
$I_2^* = BL_1$ |

(Continued on next page)
| Index | Conditions |
|-------|------------|
| S13   | $I_1 < SW_2$, $R_2 + I_1 > SW_2$, $R_1 + R_2 + I_1 + I_2 \leq SW_2 + BL_0$ |
|       | $p_1^* = 0$, $r_{11}^* = 0$, $r_{21}^* = SW_2 - I_1$ |
|       | $p_2^* = SW_2 + BL_0 - R_1 - R_2 - I_1 - I_2$, $r_{12}^* = R_3$, $r_{22}^* = R_2 - SW_2$ |
|       | $I_1 \geq SW_2$, $R_1 + R_2 + I_2 > BL_0$ |
|       | $p_1^* = 0$, $r_{11}^* = 0$, $r_{21}^* = 0$, $SW_2 \leq I_1 \leq AL_2$ |
|       | $p_2^* = BL_0 - R_1 - R_2 - I_2$, $r_{12}^* = R_3$, $r_{22}^* = R_2$ |
| S15   | $BL_0 \leq R_1 + R_2 + I_1 \leq BL_1$, $\frac{dEP}{d\zeta_1}|_{\zeta_1 = I_1} \leq \frac{dEP}{d\zeta_2}|_{\zeta_2 = R_1 + R_2 + I_1}$ |
|       | $p_1^* = 0$, $r_{11}^* = 0$, $r_{21}^* = 0$, $RP \leq I_1 \leq AL_2$ |
|       | $p_2^* = 0$, $r_{12}^* = BL_1 - R_1 - I_2$, $r_{22}^* = R_2$, $BL_0 \leq I_2 \leq BL_1$ |
| S16   | $I_1 \geq RP$, $R_2 + I_2 \leq BL_1$, $R_1 + R_2 + I_1 > BL_1$ |
|       | $p_1^* = 0$, $r_{11}^* = 0$, $r_{21}^* = 0$, $RP \leq I_1 \leq AL_2$ |
|       | $p_2^* = 0$, $r_{12}^* = BL_1 - R_1 - I_2$, $r_{22}^* = R_2$, $I_2 = BL_1$ |
| S17   | $SW_2 + BL_0 < R_1 + R_2 + I_1 + I_2 \leq RP + BL_1$ |
|       | $\frac{dEP}{d\zeta_1}|_{\zeta_1 = I_1} > \frac{dEP}{d\zeta_2}|_{\zeta_2 = R_1 + R_2 + I_2}$, $\frac{dEP}{d\zeta_1}|_{\zeta_1 = R_1 + I_1} < \frac{dEP}{d\zeta_2}|_{\zeta_2 = R_1 + I_2}$ |
|       | $p_1^* = 0$, $r_{11}^* = 0$, $p_2^* = 0$, $r_{12}^* = R_3$, $r_{22}^* = R_2$ |
|       | $SW_2 < I_1 \leq RP$, $BL_0 \leq I_2 \leq BL_1$ |
| S18   | $I_1 < RP$, $R_2 + I_1 > RP$, $R_2 + I_1 + I_2 < RP + BL_1$, $R_1 + R_2 + I_1 + I_2 > RP + BL_1$ |
|       | $p_1^* = 0$, $r_{11}^* = 0$, $r_{21}^* = RP - I_1$, $I_1 = RP$ |
|       | $p_2^* = 0$, $r_{12}^* = RP + BL_1 - R_2 - I_1 - I_2$, $r_{22}^* = R_2 - I_1 - RP$, $I_2 = BL_1$ |
| S19   | $RP + BL_1 < R_2 + I_1 + I_2 \leq AL_2 + BL_2$ |
|       | $\frac{dEP}{d\zeta_1}|_{\zeta_1 = R_1 + I_1} < \frac{dEP}{d\zeta_2}|_{\zeta_2 = I_1}$, $\frac{dEP}{d\zeta_1}|_{\zeta_1 = I_2} > \frac{dEP}{d\zeta_2}|_{\zeta_2 = R_1 + I_2}$ |
|       | $p_1^* = 0$, $r_{11}^* = 0$, $p_2^* = 0$, $r_{12}^* = 0$, solve $r_{21}^* + r_{22}^* = R_2$ |
|       | $RP \leq I_1 \leq AL_2$, $BL_1 \leq I_2 \leq BL_2$ |
| S20   | $R_2 + I_2 > BL_1$, $\frac{dEP}{d\zeta_1}|_{\zeta_1 = I_1} \leq \frac{dEP}{d\zeta_2}|_{\zeta_2 = R_1 + I_2}$ |
|       | $p_1^* = 0$, $r_{11}^* = 0$, $r_{21}^* = 0$, $RP < I_1 \leq AL_2$ |
|       | $p_2^* = 0$, $r_{12}^* = 0$, $r_{22}^* = 0$, $BL_1 < I_2 \leq BL_2$ |
| S21   | $R_2 + I_1 + I_2 > AL_2 + BL_2$ |
|       | $p_1^* = 0$, $r_{11}^* = 0$, $r_{21}^* = AL_2 - I_1$, $I_1 = AL_2$ |
|       | $p_2^* = 0$, $r_{12}^* = 0$, $r_{22}^* = BL_2 - I_2$, $I_2 = BL_2$ |
Optimal control for a two-product recovery system

A2: Threshold levels of the multi-period problem

All the threshold levels are listed in Table A2.

Table A2. Threshold levels of the multi-period problem

| Threshold | Formula |
|-----------|---------|
| $AL_{0i}$ | $F_{1i}^{-1}\left(\frac{u_i(x_i, y_i) - st + v_t - pc}{x_t + v_t + h_t}\right)$ |
| $BL_{0i}$ | $F_{2i}^{-1}\left(\frac{u_i(x_i, y_i) - st + v_t - pc}{x_t + v_t + h_t}\right)$ |
| $AL_{1i}$ | $F_{1i}^{-1}\left(\frac{u_i(x_i, y_i) - st + v_t - rc}{x_t + v_t + h_t}\right)$ |
| $BL_{1i}$ | $F_{2i}^{-1}\left(\frac{u_i(x_i, y_i) - st + v_t - rc}{x_t + v_t + h_t}\right)$ |
| $AL_{2i}$ | $F_{1i}^{-1}\left(\frac{u_i(x_i, y_i) - st + v_t - rc}{x_t + v_t + h_t}\right)$ |
| $BL_{2i}$ | $F_{2i}^{-1}\left(\frac{u_i(x_i, y_i) - st + v_t - rc}{x_t + v_t + h_t}\right)$ |
| $SW_{1i}$ | $F_{1i}^{-1}\left(\frac{u_i(x_i, y_i) - st + v_t + rc - rc - pc}{x_t + v_t + h_t}\right)$ |
| $SW_{2i}$ | $F_{1i}^{-1}\left(\frac{u_i(x_i, y_i) - st + v_t + rc - rc - pc}{x_t + v_t + h_t}\right)$ |
| $RP_{i}$ | $F_{1i}^{-1}\left(\frac{u_i(x_i, y_i) - st + v_t + rc - rc - pc}{x_t + v_t + h_t}\right)$ |

A3: The extension to a general multi-product recovery system

After studying the two-product recovery system, we extend our work to a general multi-product recovery system. The structure of a multi-period system is depicted in Fig. A1. Similar to the two-product problem, some restrictions on cost parameters are also imposed to ensure the economic validity of the study on the N-product recovery system.

First, for each product, the selling price is higher than the production cost, and the penalty cost of shortage is higher than the profit from production. Therefore, $s_j > pc_j$ and $v_j > s_j - pc_j$ ($j = 1, 2, \ldots, N$). Second, for each product, the production cost is higher than recovery cost. Otherwise, recovery is unnecessary. Therefore, $pc_j > rc_{1j}$ and $pc_j > rc_{2j}$ ($j = 1, 2, \ldots, N$). Lastly, the recovery using returns in group 2 is cheaper than that using returns in group 1. Therefore, $rc_{2j} < rc_{1j}$ ($j = 1, 2, \ldots, N$).

The N-product recovery system is first studied in a single period by referring to the two-product recovery system. Similarly, the optimal solution can be found by solving the KKT conditions of a Lagrangian function. The optimal solution to the single-period problem on the two-product recovery system has been shown in Appendix A1 based on a choice of cost structure: $pc_1 - rc_{11} > pc_2 - rc_{12}$, $pc_1 - rc_{21} > pc_2 - rc_{22}$, and $rc_{11} - rc_{21} > rc_{12} - rc_{22}$. For the N-product recovery system, the optimal solution can be similarly shown based on the cost structure ($j = 1, 2, \ldots, N - 1$): $pc_j - rc_{1j} > pc_{j+1} - rc_{1,j+1}$, $pc_j - rc_{2j} > pc_{j+1} - rc_{2,j+1}$, and $rc_{1j} - rc_{2j} > rc_{1,j+1} - rc_{2,j+1}$. For the N-product recovery system, there are three order-up-to levels for each product with respect to the three replenishment sources. Based on the selected cost structure, there are three additional switching levels to control the interactive allocation of possibly limited recovery sources between product $j$ and product $j + 1$. Therefore, the total number of threshold levels for the N-product recovery system is calculated as $3 \times N + 3 \times (N - 1) = 6 \times N - 3$. Although the complexity of the optimal solution will increase with the number of products in the recovery system, the N-product case can still be formulated as a nonlinear programming problem and solved using the KKT conditions. When the recovery system is extended from the single-period context to the multi-period problem, we can still consider the use of the multi-level threshold policy, as it is easy to implement. For example, for an M-period problem of an 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 system.

Fig. A1. A general multi-product system.

Biographies

Jie Pan received his B.Eng. and M.Eng. degrees from Beihang University, China, in 1999 and 2002, respectively. He received his Ph.D. in Industrial and Systems Engineering at the National University of Singapore in 2011. He joined for Integrated Decision Systems Consultancy Pte Ltd (Singapore) as Senior Optimization Consultant in 2013. Prior to this, he worked as a Business Optimization Analyst for Alexandra Health Pte Ltd (Singapore) between 2010 and 2013. He also worked as a research engineer in the School of Information Systems at Singapore Management University in 2009. His research interests include operations research, logistics and supply chain management, forecasting, sales analytics, and inventory optimization.
Yi Tao is currently an Assistant Professor in the School of Management at Guang Dong University of Technology, China. He received his B.S. (Computer Science) degree from the Beihang University, China, in 2005 and his Ph.D. from Sun Yat-Sen University, China, in 2010. His research interests include logistics and supply chain management, port operations and maritime transportation, and optimization.

Loo Hay Lee is an Associate Professor in the Department of Industrial and Systems Engineering at the National University of Singapore. He received his B.S. (Electrical Engineering) degree from the National Taiwan University in 1992 and his S.M. and Ph.D. degrees in 1994 and 1997 from Harvard University. His research focuses on the simulation-based optimization, maritime logistics, which includes port operations, and modeling and analysis of logistics and supply chain systems.

Ek Peng Chew is currently an Associate Professor and Deputy Head (Undergraduate Studies) in the Department of Industrial and Systems Engineering at the National University of Singapore. He received his Ph.D. in Industrial Engineering from the Georgia Institute of Technology. His current research interests include port logistics and maritime transportation, simulation optimization, and inventory management.