Joint Production and Disposal Decisions for Sustainable Operations of the Hybrid Production System

Eungab Kim[†]

College of Business Administration, Ewha Womans University

혼성 생산 시스템의 지속 가능 운영을 위한 신제품 생산과 회수제품 수용 통제의 통합 구현

김은갑

이화여자대학교 경영대학 경영전문대학원

We consider a reverse supply chain with a production facility and a recovery facility, and address the joint control of production and disposal decisions for sustainable operations. Demands are satisfied from on-hand inventory of serviceable products, replenished via manufacturing or remanufacturing. Sold products may be returned after usage and each returned product is disposed of or accepted for recovery. Accepted returned products are converted into serviceable products after remanufacturing process. Formulating the model as a Markov decision process, we characterized the structure of the optimal production and disposal policy as two monotone switching curves under a special condition. Three types of heuristic policies are presented and their performance is numerically compared.

Keywords: Reverse Supply Chain, Product Return, Sustainable Operations, Recovery, Disposal

1. Introduction

Remanufacturing is not only an important step toward sustainability, but is also economically beneficial for firms. Therefore, many firms implement collection channels for used product return and integrate corresponding remanufacturing strategies with their regular production system (Li *et al*., 2013). Considering the economic benefits from a cheaper remanufacturing, it might be also cost-effective in some industries to consider disposing of returned items, since an increase in the inventory of returned products may be too costly (Teunter and Vlachos, 2002; van der Lann *et al*., 1996a). This research is a step toward rationalizing such decisions. In this paper, we generate insights into how production of new products and

disposal of returned items can be effectively coordinated in a hybrid production system with manufacturing and remanufacturing. Such coordination will provide financial incentives for firms to move toward sustainable operations.

This paper considers a single type of product stocked in order to meet random demand from customers who may return products after usage. Demands are satisfied from on-hand inventory of products, replenished via manufacturing or remanufacturing. Unmet demand is lost. Product returns occur randomly and each returned item is disposed of or accepted for recovery. Accepted ones are converted into products and are as good as newly produced ones. The assumption of the equal quality is valid in some real cases. For example, Zhou *et al*. (2011) considered a large energy company which provides service on meters and transformers for private houses

and commercial buildings and recovers all failed meters/transformers. In this case, customers are indifferent between a new product and a remanufactured one.

Although there has been substantial research into the production planning and inventory management in a hybrid system, relatively few papers have addressed the product recovery with a disposal option. Heyman (1977) presented a continuous inventory review model with zero lead times for remanufacturing and procurement processes. Under Poisson demand and return processes, the optimality of a single parameter disposal policy is shown. Simpson (1978) studied a periodic inventory review model where unmet demand is backlogged and zero lead times are assumed for purchase and recovery processes, and show that a three-parameter policy is optimal to control manufacturing, remanufacturing, and disposal. Inderfurth (1997) studied the effect of deterministic replenishment and remanufacturing lead times on the optimal policy, and show that the policy identified in Simpson (1978) is also optimal for identical deterministic lead times. van der Lann *et al.* (1996a) studied a continuous inventory review model with Poisson demand and returns in which the number of remanufacturable products is limited to a level. van der Lann *et al.* (1996b) show, based on a numerical study, that there is an advantage of making disposal decisions based on both remanufacturable and serviceable inventory levels rather than only based on the remanufacturable inventory level. Using simulation, Teunter and Vlachos (2002) showed that

disposal option can result in a significant cost reduction only when an item is very slow moving, the recovery rate is high, and remanufacturing is almost as expensive as manufacturing.

Detailed literature reviews for inventory management of a hybrid production system are summarized in <Table 1>. We classify the references using the following characteristics : 1. Inventory review (periodic or continuous), 2. Production and remanufacturing lead time (stochastic, constant, or zero), 3. Type of inventory holding (serviceable, remanufacturable), 4. Type of unit costs (production, remanufacturing, disposal), 5. Type of control (production, disposal), 6. Identification of optimal policy (yes or no), 7. Stockout treatment (backlog or lost sales).

Our model contributes to this stream of research in several ways. First, we coordinate decisions of when to produce a new product and dispose of a returned item based on the levels of both serviceable and remanufacturable inventories. Second, we identify the optimal policy under a continuous inventory review. Previous studies with continuous inventory review have mainly tried to find optimal parameters for given policies rather than characterizing the optimal policy. This is partially due to the complexity of the underlying control problem when lead times are stochastic. Heyman (1977) and Inderfurth (2001) identified the optimal policy under a continuous inventory review. But Heyman (1977) did not consider lead time and Inderfurth (2001) assumed constant lead time. Third, we develop effective and implementable heuristic policies

Table 1. Summary of literature survey (PR = periodic review, CR = continuous review, M = manufacturing, R = remanufacturing, D = disposal, H_1 = serviceable inventory, H_2 = remanufacturable inventory, B = backlog, L = lost sales, Exp = Exponential random variable, Dis = Discrete random variable)

Articles (indexed by the first author)	Inventory review		Lead Time						Holding		Unit cost			Control		Opt. policy	Stock out	
			Stochastic		Constant		Zero											
	PR	CR	M	R	M	R	M	R	H_1	H_2	M	R	D	M	D		B	L
DeCroix, 2006	O				O	O			O	O	O	O	O	O	O	O	O	
Fleischmann, 2003		O			O				O					O			O	
Heyman, 1977		O					O	O	O		O		O	O	O	O	O	
Inderfurth, 1997	O				O	O			O	O	O	O	O	O	O	O	O	
Inderfurth, 2001		O			O	O			O	O	O	O	O	O	O	O		
Mahadevan, 2003		O			O	O			O	O				O			O	
Muckstadt, 1981		O			O				O					O			O	
Simpson, 1978	O						O	O	O	O	O	O		O	O	O	O	
Teunter, 2002	O				O	O			O	O	O	O	O	O	O		O	
Teunter, 2004		O			O	O			O	O				O			O	
van der Laan, 1996a		O		Exp	O				O	O				O	O		O	
van der Laan, 1996b		O		Exp	O				O	O	O	O	O	O	O		O	
van der Laan, 1997		O			O	O			O					O	O		O	
van der Laan, 1999a		O	Dis	Dis					O	O	O	O	O	O			O	
van der Laan, 1999b		O			O	O			O	O	O	O		O			O	
van der Laan, 2006		O			O	O			O	O	O			O			O	
Wei, 2009	O				O	O			O	O	O	O	O	O	O		O	
Our model		O	Exp	Exp					O	O	O	O	O	O	O	O		O

which provide firms with a simple tool to achieve higher profits while encouraging the use of sustainable operations.

The remainder of the paper is organized as follows. In Section 2, we present our model. Analysis of the optimal policy is given in Section 3. In Section 4, we implement a sensitivity analysis on the optimal policy. In Section 5, we present a computational study of the optimal and heuristic policies. Finally, we conclude in Section 6.

2. Model Assumptions and Problem Formulation

Demands for a single product arrive according to a Poisson process with rate λ_1. If demand is satisfied from on-hand inventory, a revenue of R is generated. Hereafter, we refer to the on-hand inventory as serviceable inventory. Each unmet demand due to stockout is lost. The production time of a new item is exponentially distributed with mean μ_1^{-1}. Arrival of returned products follows a Poisson process with rate λ_2. Similar to the production process, recovery is performed item by item (i.e., no batching). The remanufacturing time is exponentially distributed with mean μ_2^{-1}. Holding costs are assessed at rate h_1 and h_2 for each unit in serviceable and remanufacturable inventory, respectively. We assume that $h_1 \geq h_2$. Unit production and remanufacturing costs are c_M and c_R, respectively. Disposing of a returned product incurs c_D where $c_D < 0$ implies a salvage value.

Although the return process may be affected by the previous sales, we assume that demand process does not influence the rate of return. If there are large quantities of sold products in the market, the independence assumption will be reasonable. It can also serve as a reasonable approximation in solving many practical problems by yielding easily implementable control policies. See Fleischmann (2000) for detailed justification of the independence assumption.

At each epoch of a demand arrival, a product return, a production completion, or a remanufacturing completion, a control policy specifies whether or not to produce a new item. Upon a product return, a decision must be made regarding whether to dispose of or accept it. A state is described by the vector (x_1, x_2) where x_1 and x_2 represent the serviceable and remanufacturable inventory level, respectively. We denote the state space by Γ.

The goal of this paper is to identify a joint production and disposal control policy that maximizes the long-run average profit. To this end, we formulate our model as a discrete-time Markov Decision Process (MDP) problem with a transition rate $\gamma \equiv \lambda_1 + \lambda_2 + \mu_1 + \mu_2$. The properties and insights provided in this paper can be useful for understanding and addressing more realistic models with non-exponential distributions where it is not tractable to identify the optimal policy.

Let $v(x_1, x_2)$ be the optimal profit function in state (x_1, x_2). We define the value iteration operator T on v as :

$$
\begin{aligned}
Tv(x_1, x_2) = \frac{1}{\gamma}[&-(x_1 h_1 + x_2 h_2) \qquad\qquad (1)\\
&+ \lambda_1\{(R + v(x_1 - 1, x_2))1(x_1 > 0) + v(x_1, x_2)1(x_1 = 0)\}\\
&+ \mu_1 \max\{-c_M + v(x_1 + 1, x_2), v(x_1, x_2)\}\\
&+ \mu_2\{(v(x_1 + 1, x_2 - 1) - c_R)1(x_2 > 0) + v(x_1, x_2)1(x_2 = 0)\}\\
&+ \lambda_2 \max\{v(x_1, x_2 + 1), v(x_1, x_2) - c_D\}]
\end{aligned}
$$

where $1(a) = 1$ if a is true, otherwise, 0. In (1), $-(x_1 h_1 + x_2 h_2)/\gamma$ is the average holding costs of x_1 units and x_2 units until the next state transition. The terms multiplied by λ_1 and λ_2 represent the sales revenue and transition generated with a demand arrival and the cost and transition associated with a product return. The terms multiplied by μ_1 and μ_2 implies the costs and transitions associated with a production completion and a recovery completion. Then, the optimality equation of our MDP model can be written as

$$
g + v(x_1, x_2) = Tv(x_1, x_2)
$$

where g is the optimal average profit during the expected state transition time γ^{-1}.

3. Structure of the Optimal Production and Disposal Policy Under the Special Case

In this section, we characterize the structure of the optimal production and disposal policy under the assumption of $h_1 = h_2$. When $h_1 > h_2$, we could not complete the proof of the existence of the optimal policy, although it is a more practical assumption. The condition $h_1 = h_2$ guarantees that it is always advantageous to convert the returned products into the serviceable ones as long as $x_2 > 0$, that is, $f(x_1 + 1, x_2) - c_R \geq f(x_1, x_2 + 1)$. When $h_1 > h_2$, however, the above equation cannot be confirmed. One might think that it is a better policy to postpone the remanufacturing until it is required when $h_1 > h_2$. For example, when x_1 is large enough to cope with demand, the replenishment of serviceable product via remanufacturing could not be cost-effective. This indicates that the control of remanufacturing as well as production and disposal will become important for more efficient operation of a hybrid production system. We will cover this issue in the future research.

To prove the structure of the optimal production and disposal policy, we identify a set of structured profit functions is preserved under the value iteration operator T (Porteus, 1982). Define $D_i f$ $i = 1, 2$, on any real valued function f :

$$D_1 f(x_1, x_2) \equiv f(x_1 + 1, x_2) - f(x_1, x_2),$$
$$D_2 f(x_1, x_2) \equiv f(x_1, x_2 + 1) - f(x_1, x_2).$$

$D_1 f$ ($D_2 f$) is the marginal value of having one more unit in serviceable (remanufacturable) inventory. Let F be the set of all functions defined on state space Γ such that if $f \in F$, then

$$D_i f(x_i, x_j) \geq D_i f(x_i, x_j + 1),$$
$$i \neq j, i = 1, 2, j = 1, 2, \tag{2}$$

$$D_i f(x_i, x_j + 1) \geq D_i f(x_i + 1, x_j),$$
$$i \neq j, i = 1, 2, j = 1, 2, \tag{3}$$

$$D_i f(x_i, x_j) \geq D_i f(x_i + 1, x_j),$$
$$i \neq j, i = 1, 2, j = 1, 2, \tag{4}$$

$$D_1 f(x_1, x_2) \leq R, \tag{5}$$

$$f(x_1 + 1, x_2) - c_R \geq f(x_1, x_2 + 1). \tag{6}$$

Equation (2) states that f is submodular, which implies that the marginal benefit of holding one more unit of one inventory decreases as the other inventory increases. Equation (3) implies that the marginal value of having one more unit of one inventory decreases faster in its inventory than in the other inventory. Equation (4) says that f is concave in its coordinates, which means that the incremental benefit with one more unit of one inventory is decreasing in its inventory. Note that (2) and (3) together imply concavity. Equation (6) says that it is always advantageous to convert the returned product into the serviceable one.

The following lemma guarantees that equations (2)-(6) are preserved under T. The proof of this and all subsequent results are included in the Appendix.

Lemma 1 : If $f \in F$, $Tf \in F$.

The following theorem identifies the structural properties of the optimal policy :

Theorem 1.
(i) There exists an optimal average profit, g, and an optimal profit function, v, which satisfy (1), and $v \in F$.

(ii) Let
$$P(x_2) \equiv \max\{x_1 : v(x_1, x_2) \leq v(x_1 + 1, x_2) - c_M\},$$
$$D(x_1) \equiv \min\{x_2 : v(x_1, x_2 + 1) \leq v(x_1, x_2) - c_D\}$$

In state (x_1, x_2), it is optimal to produce an item if $x_1 \leq P(x_2)$. When a product return occurs, it is optimal to disposes of this return if $x_2 \geq D(x_1)$, otherwise, accepts it for recovery.

(iii) $P(x_2)$ decreases in x_2.
(iv) $D(x_1)$ decreases in x_1.

The structure of the optimal policy is characterized by two monotone stitching curves $P(x_2)$ and $D(x_1)$. Part (ii) states that if $D_1 v \geq c_M$, it is beneficial to increase x_1 by one unit. It also says that if $D_2 v \leq -c_D$, it should dispose of a returned item. Parts (iii) and (iv) state that it is optimal to keep less of one inventory as the other inventory increases, which follows from the submodularity of v. Since a returned item, once recovered, is perfectly converted into a serviceable product, if more units are stored in one inventory, it is cost-effective to store less units of the other inventory.

4. Marginal Effects of Problem Parameters on the Optimal Policy

Suppose that product returns are controlled by a threshold, M. In other words, a product return is accepted for recovery only when $x_2 < M$ and, otherwise, disposed of. We then investigate how the system parameters affect $P(x_2)$. Consider the following optimization problem :

$$v(x_1, x_2) + g = \frac{1}{\gamma}[-x_1 h_1 - x_2 h_2 \tag{7}$$
$$+ \lambda_1\{(R + v(x_1 - 1, x_2))1(x_1 > 0) + v(x_1, x_2)1(x_1 = 0)\}$$
$$+ \mu_1 \max\{-c_M + v(x_1 + 1, x_2), v(x_1, x_2)\}$$
$$+ \mu_2\{(v(x_1 + 1, x_2 - 1) - c_R)1(x_2 > 0) + v(x_1, x_2)1(x_2 = 0)\}$$
$$+ \lambda_2\{f(x_1, x_2 + 1)1(x_2 < M) + (f(x_1, x_2) - c_D)1(x_2 = M)\}]$$

Then, using the same arguments as in Section 3, one can show the existence of $P(x_2)$ as it is defined in Theorem 1. We first present the effects of the cost parameters on $P(x_2)$:

Theorem 2.
(a) If unit sales price R increases, then $P(x_2)$ increases.
(b) If unit manufacturing cost c_M increases, then $P(x_2)$ decreases.
(c) If unit holding cost rate h_1 increases, then $P(x_2)$ decreases.

The monotonicity of $P(x_2)$ with respect to the time rate parameters is presented in the following theorem. Since the results can be proven using the arguments similar to the ones in Theorem 2, we omit the proof.

Theorem 3.
(a) If demand rate λ_1 increases, then $P(x_2)$ increases.
(b) If manufacturing rate μ_1 increases, then $P(x_2)$ decreases.
(c) If recovery rate μ_2 increase, then $P(x_2)$ decreases.
(d) If product return rate λ_2 increases, then $P(x_2)$ decreases.

5. Numerical Study

$P(x_2)$ and $D(x_1)$, determined by Theorem 1, are fairly complex because they are characterized by state dependent threshold levels. Since the number of system states grows exponentially as x_1 and x_2 get larger, computing the optimal policy becomes intractable. Hence, motivated by the prohibitive inefficiencies associated with problems of large sizes, we introduce the heuristic solution approaches. The heuristics differ in the rule that uses x_1 and x_2 in determining production and disposal controls. Also, unlike $P(x_2)$ and $D(x_1)$ applied only to the MDP problem, the following heuritics are applicable to the non-MDP problem with general probability distributions.

The first heuristic is a base stock policy for the two-stage production system (Veach and Wein, 1994) :

- In (x_1, x_2), a new item is produced when $x_1 < H_S^1$.
- If a product return occurs in (x_1, x_2), then it is accepted when $x_1 + x_2 < H_S^1 + H_R^1$. When $x_1 + x_2 = H_S^1 + H_R^1$, each returned item is disposed of.

The optimality equation for the base stock policy is

$$
\begin{aligned}
v(x_1,x_2) + g_{H_1} = \frac{1}{\gamma}[&-x_1 h_1 - x_2 h_2 \qquad\qquad (8)\\
&+ \lambda_1\{(R + v(x_1-1,x_2))1(x_1 > 0) + v(x_1,x_2)1(x_1 = 0)\}\\
&+ \mu_1\{(v(x_1+1,x_2)-c_M)1(x_1 < H_S^1), v(x_1,x_2)1(x_1 = H_S^1)\}\\
&+ \mu_2\{(v(x_1+1,x_2-1)-c_R)1(x_2 > 0) + v(x_1,x_2)1(x_2 = 0)\}\\
&+ \lambda_2\{f(x_1,x_2+1)1(x_1+x_2 < H_S^1+H_R^1) +\\
&\quad (f(x_1,x_2)-c_D)1(x_1+x_2 = H_S^1+H_R^1)\}]
\end{aligned}
$$

The second heuristic is a fixed buffer policy for the two-stage production system (Veach and Wein, 1994) :

- In (x_1, x_2), a new item is produced when $x_1 < H_S^2$.
- If a product return occurs in (x_1, x_2), then it is accepted when $x_2 < H_R^2$. When $x_2 = H_R^2$, each returned item is disposed of.

The optimality equation for the fixed buffer policy is

$$
\begin{aligned}
v(x_1,x_2) + g_{H_2} = \frac{1}{\gamma}[&-x_1 h_1 - x_2 h_2 \qquad\qquad (9)\\
&+ \lambda_1\{(R + v(x_1-1,x_2))1(x_1 > 0) + v(x_1,x_2)1(x_1 = 0)\}\\
&+ \mu_1\{(v(x_1+1,x_2)-c_M)1(x_1 < H_S^2), v(x_1,x_2)1(x_1 = H_S^2)\}\\
&+ \mu_2\{(v(x_1+1,x_2-1)-c_R)1(x_2 > 0) + v(x_1,x_2)1(x_2 = 0)\}\\
&+ \lambda_2\{f(x_1,x_2+1)1(x_2 < H_R^2) + (f(x_1,x_2)-c_D)1(x_2 = H_R^2)\}]
\end{aligned}
$$

The third heuristic uses two linear switching functions to make production and disposal decisions as follows :

- In (x_1, x_2), a new item is produced when $x_1 + x_2 < H_S^3$.
- If a product return occurs in (x_1, x_2), then it is accepted when $x_1 + x_2 < H_R^3$. When $x_1 + x_2 = H_R^3$, each returned item is disposed of.

The optimality equation for the third control policy is

$$
\begin{aligned}
v(x_1,x_2) + g_{H_3} = \frac{1}{\gamma}[&-x_1 h_1 - x_2 h_2 \qquad\qquad (10)\\
&+ \lambda_1\{(R + v(x_1-1,x_2))1(x_1 > 0) + v(x_1,x_2)1(x_1 = 0)\}\\
&+ \mu_1\Big\{(v(x_1+1,x_2)-c_M)1(x_1+x_2 < H_S^3),\\
&\qquad v(x_1,x_2)1(x_1+x_2 = H_S^3)\Big\}\\
&+ \mu_2\Big\{(v(x_1+1,x_2-1)-c_R)1(x_2 > 0) +\\
&\qquad v(x_1,x_2)1(x_2 = 0)\Big\}\\
&+ \lambda_2\Big\{f(x_1,x_2+1)1(x_1+x_2 < H_R^3) +\\
&\qquad (f(x_1,x_2)-c_D)1(x_1+x_2 = H_R^3)\Big\}]
\end{aligned}
$$

Best values of $(H_S^{1^*}, H_R^{1^*})$, $(H_S^{2^*}, H_R^{2^*})$, and $(H_S^{3^*}, H_R^{3^*})$ to maximize g_{H_1}, g_{H_2}, and g_{H_3}, respectively, are found using a two-dimensional search on possible integer values of each pair via the value iteration algorithm presented in the appendix. Performance comparison between the optimal and heuristic policies is reported in <Table 2> and <Table 3>. In these tables, % are obtained from $\frac{g - g_{H_i}}{g} \times 100$, $i = 1, 2, 3$. We set $\lambda_1 = 0.5$, $\mu_1 = 0.6$, $\mu_2 = 0.9$, and $\lambda_2 = 0.25$ in <Table 2>, and $R = 100$, $h_1 = 2$, $h_2 = 1$, $c_R = 5$, $c_M = 10$, and $c_D = 3$ in <Table 3>.

The average % differences of three heuristic policies over the optimal policy show that the base stock policy and the linear switching rule performs very well. This indicates that the information of both x_1 and x_2 is very useful in determining production and disposal controls.

It is worth noting that $(H_S^{1^*}, H_R^{1^*}), (H_S^{2^*}, H_R^{2^*})$, and $(H_S^{3^*}, H_R^{3^*})$ are sensitive to the change of h_1, c_M, λ_1, and λ_2. For example, from <Table 2>, $(H_S^{1^*}, H_R^{1^*})$, $(H_S^{2^*}, H_R^{2^*})$, and $(H_S^{3^*}, H_R^{3^*})$ tend to decrease as h_1 increases, which leads to less production of new items and more disposal of returned items. It is also observed that as c_M increases, H_R^1, H_R^2, and H_R^3 are increasing, which means each heuristic accepts more returned items. This makes sense because the product recovery becomes more valuable in replenishing serviceable inventory if c_M increases. From <Table 3>, as λ_1 increases, $(H_S^{1^*}, H_R^{1^*})$, $(H_S^{2^*}, H_R^{2^*})$, and $(H_S^{3^*}, H_R^{3^*})$ tend to increase. With increased λ_1, the chance of enhancing sales revenue also increases and thus serviceable inventory becomes more valuable. Hence, the firm has an incentive to manufacture more new products and accept more returned products. In contrast, an increase in λ_2 tends to decrease $(H_S^{1^*}, H_R^{1^*})$, $(H_S^{2^*}, H_R^{2^*})$, and $(H_S^{3^*}, H_R^{3^*})$. When λ_2 is increased, more serviceable products can be stocked because of increased returned products unless the control policy would be changed. Therefore, the production and disposal thresholds must be lowered to avoid excessive units in both serviceable and remanufacturable inventory.

Table 2. Comparison of the optimal and heuristic policies : Cost parameters

	R	h_1	h_2	c_R	c_M	c_D	g	Base stock				Fixed buffer				Linear switching			
								g_{H_1}	$H_S^{1^*}$	$H_R^{1^*}$	%	g_{H_2}	$H_S^{2^*}$	$H_R^{2^*}$	%	g_{H_3}	$H_S^{3^*}$	$H_R^{3^*}$	%
1	100	2	1	5	10	3	37.05	37.02	3	2	0.08	36.86	3	2	0.51	37.01	4	5	0.11
2	125						48.91	48.75	4	2	0.33	48.51	4	2	0.82	48.88	4	6	0.06
3	150						60.83	60.73	4	2	0.16	60.52	4	2	0.51	60.78	4	6	0.08
4	175						72.75	72.71	4	2	0.05	72.53	4	2	0.30	72.68	4	6	0.10
5	100	1.5	1	5	10	3	38.62	38.5	4	2	0.31	38.36	4	2	0.67	38.6	4	6	0.05
6		2					37.05	37.02	3	2	0.08	36.86	3	2	0.51	37.01	4	5	0.11
7		2.5					35.79	35.71	3	2	0.22	35.43	3	2	1.01	35.69	3	5	0.28
8		3					34.6	34.45	3	1	0.43	34.13	3	1	1.36	34.53	3	4	0.20
9	100	2	0.5	5	10	3	37.21	37.17	3	2	0.11	37.02	3	2	0.51	37.15	4	5	0.16
10			1				37.05	37.02	3	2	0.08	36.86	3	2	0.51	37.01	4	5	0.11
11			1.5				36.89	36.87	3	2	0.05	36.7	3	2	0.52	36.87	4	5	0.05
12			2				36.75	36.72	3	2	0.08	36.54	3	2	0.57	36.73	4	5	0.05
13	100	2	1	2	10	3	37.82	37.77	3	3	0.13	37.64	3	2	0.48	37.74	4	6	0.21
14				4			37.3	37.26	3	2	0.11	37.12	3	2	0.48	37.24	4	5	0.16
15				6			36.8	36.78	3	2	0.05	36.6	3	2	0.54	36.78	4	5	0.05
16				8			36.35	36.29	3	2	0.17	36.07	3	2	0.77	36.33	4	5	0.06
17	100	2	1	5	8	3	37.57	37.51	3	2	0.16	37.32	3	2	0.67	37.55	4	5	0.05
18					11		36.81	36.77	3	2	0.11	36.63	3	2	0.49	36.74	4	5	0.19
19					14		36.12	36.07	3	3	0.14	35.94	3	2	0.50	36.01	3	6	0.30
20					17		35.44	35.37	3	3	0.20	35.25	3	3	0.54	35.35	3	6	0.25
21	100	2	1	5	10	1	37.12	37.08	3	2	0.11	36.89	3	2	0.62	37.1	4	5	0.05
22						2	37.08	37.05	3	2	0.08	36.87	3	2	0.57	37.06	4	5	0.05
23						3	37.05	37.02	3	2	0.08	36.86	3	2	0.51	37.01	4	5	0.11
24						4	37.02	36.99	3	2	0.08	36.84	3	2	0.49	36.97	4	5	0.14
Average % of optimality gap											0.14				0.6				0.13

Table 3. Comparison of the optimal and heuristic policies : Time parameters

	λ_1	μ_1	μ_2	λ_2	g	Base stock				Fixed buffer				Linear switching			
						g_{H_1}	$H_S^{1^*}$	$H_R^{1^*}$	%	g_{H_2}	$H_S^{2^*}$	$H_R^{2^*}$	%	g_{H_3}	$H_S^{3^*}$	$H_R^{3^*}$	%
1	0.35	0.6	0.9	0.25	25.39	25.37	2	2	0.08	24.88	2	1	2.01	25.26	3	4	0.51
2	0.45				33.31	33.21	3	2	0.30	32.96	3	1	1.05	33.24	3	5	0.21
3	0.55				40.88	40.74	4	2	0.34	40.6	4	2	0.68	40.85	4	6	0.07
4	0.65				48.09	47.95	5	2	0.29	47.9	5	3	0.40	48.07	5	7	0.04
5	0.5	0.55	0.9	0.25	36.78	36.67	3	2	0.30	36.53	3	2	0.68	36.76	4	5	0.05
6		0.65			37.34	37.3	3	2	0.11	37.12	3	2	0.59	37.22	3	5	0.32
7		0.75			37.82	37.73	3	2	0.24	37.53	3	1	0.77	37.72	3	5	0.26
8		0.85			38.17	38.03	3	2	0.37	37.87	3	1	0.79	38.07	3	5	0.26
9	0.5	0.6	0.6	0.25	36.26	36.15	3	2	0.30	36.02	3	2	0.66	36.13	4	5	0.36
10			0.8		36.84	36.8	3	2	0.11	36.65	3	2	0.52	36.81	4	5	0.08
11			1		37.22	37.19	3	2	0.08	37.02	3	2	0.54	37.16	4	5	0.16
12			1.2		37.47	37.44	3	2	0.08	37.27	3	2	0.53	37.44	3	5	0.08
13	0.5	0.6	0.9	0.1	35.88	35.84	4	2	0.11	35.84	4	3	0.11	35.87	4	6	0.03
14				0.2	36.77	36.62	4	2	0.41	36.57	3	2	0.54	36.75	4	6	0.05
15				0.3	37.33	37.29	3	2	0.11	36.9	3	1	1.15	37.18	4	5	0.40
16				0.4	37.66	37.53	3	2	0.35	36.99	3	1	1.78	37.43	3	5	0.61
Average % of optimality gap									0.22				0.8				0.22

6. Conclusions

This paper studied the joint control of production and disposal decisions for a hybrid system with a disposal option. When both holding cost rates of serviceable and remanufacturable inventories are equal, we showed that the optimal joint decisions are controlled by two monotone switching curves. In addition, when a disposal control is a threshold rule, we showed that the optimal production switching curve is monotonically affected by the changes in system parameters. From the inventory policies for forward supply chain models, we developed three heuristic policies. Numerical experiment demonstrated that the heuristics defined by both serviceable and remanufacturable inventories are effective.

One can extend the model into a situation where the quality of remanufactured product is not the same as that of a new product. In such a situation, a portfolio of new and remanufactured products will become important because the remanufactured product reduces the sales of the new product when sold on the same market. Another important future direction is to consider situations where the product return intensity is varied by the compensation given to customers and there is a random yield in the remanufacturing process due to the uncertain quality of the returned items.

References

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

Fleischmann, M. (2000), Quantitative models for reverse logistics, Ph.D. thesis, Erasmus University, Rotterdam, The Netherlands.

Fleischmann, M. and Kuik, R. (2003), On optimal inventory control with independent stochastic item returns, *European Journal of Operational Research*, **151**, 25-37.

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

Inderfurth, K. (1997), Simple optimal replenishment and disposal policies for a product recovery system with lead times, *OR Spektrum*, **19**, 111-122.

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

Li, X., Li, Y., and Saghafian, S. (2013), A hybrid manufacturing/remanu-
facturing system with random remanufacturing yield and market-driven product acquisition, To apear in *IEEE Transactions on Engineering Management*.

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

Muckstadt, J. and Isaac, M. (1981), An analysis of single item inventory systems with returns, *Naval Research Logistics*, **28**, 237-254.

Porteus, E. (1982), Conditions for characterizing the structure of optimal strategies in infinite-horizon dynamic programs, *Journal of Optimization Theory and Applications*, **36**, 419-432.

Puterman, M. (2005), *Markov Decision Processes*, John Wiley and Sons.

Simpson, V. (1978), Optimal solution structure for a repairable inventory problem, *Operations Research*, **26**, 270-281.

Teunter, R. and 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., van der Laan, E., and Vlachos, D. (2004), Inventory strategies for systems with fast remanufacturing, *The Journal of Operational Research Society*, **55**, 475-484.

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

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

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

van der Laan, E., Salomon, M., and Dekker, R. (1999a), An investigation of lead-time effects in manufacturing/remanufacturing systems under simple PUSH and PULL control strategies, *European Journal of Operational Research*, **115**, 195-214.

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

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

Veatch, M. and Wein, L. (1994), Optimal control of a two-station tandem production/inventory system, *Operations Research*, **42**, 337-350.

Wei, C., Li, Y., and Cai, X. (2009), Robust optimal policies of production and inventory with uncertain returns and demand, *International Journal of Production Economics*, **134**, 357-367.

Zhou, S., Tao, Z. and Chao, X. (2011), Optimal control of inventory systems with multiple types of remanufacturable products, *Manufacturing and Service Operations Management*, **13**, 20-34.

<Appendix A>

We denote by P, N, D, and A the optimal actions in state (x_1, x_2) where P, N, D, and A respectively represent *Produce, Do not produce, Dispose of, and Accept*. Define

$$T_1 f(x_1, x_2) = (R + f(x_1 - 1, x_2))1(x_1 > 0) + f(x_1, x_2)1(x_1 = 0),$$
$$T_2 f(x_1, x_2) = \max\{-c_M + f(x_1 + 1, x_2), f(x_1, x_2)\},$$
$$T_3 f(x_1, x_2) = (f(x_1 + 1, x_2 - 1) - c_R)1(x_2 > 0) + f(x_1, x_2)1(x_2 = 0),$$
$$T_4 f(x_1, x_2) = \max\{f(x_1, x_2 + 1), f(x_1, x_2) - c_D\}.$$

Then, the value iteration operator T on f can be written as

$$Tf(x_1, x_2) = \frac{1}{\gamma}\left[-x_1 h_1 - x_2 h_2 + \lambda_1 T_1 f(x_1, x_2) + \mu_1 T_2 f(x_1, x_2) + \mu_2 T_3 f(x_1, x_2) + \lambda_2 T_4 f(x_1, x_2)\right].$$

Proof of Lemma 1 :

(i) Since $D_1 f(x_1, x_2) - D_1 f(x_1, x_2 + 1) = D_2 f(x_1, x_2) - D_2 f(x_1 + 1, x_2)$, we show the case with $i = 1$ and $j = 2$.

Let $\Delta^k = D_1 T_k f(x_1, x_2) - D_1 T_k f(x_1, x_2 + 1)$.

First, we show $\Delta^k \geq 0$, $k = 1, 2, 3, 4$.

$\Delta^1 \geq 0$ by (2) if $x_1 > 0$ and $\Delta^1 = 0$ if $x_1 = 0$.

$\Delta^3 \geq 0$ by (2) if $x_2 > 0$ and $\Delta^3 = D_1 f(x_1, 0) - D_1 f(x_1 + 1, 0) \geq 0$ by (4) if $x_2 = 0$.

To show $\Delta^2 \geq 0$ and $\Delta^3 \geq 0$, consider production and disposal controls in $(x_1 + 1, x_2)$, (x_1, x_2), $(x_1 + 1, x_2 + 1)$, and $(x_1, x_2 + 1)$:

- Cases $(P, N, \cdot, \cdot)$ and $(\cdot, \cdot, P, N)$ are excluded by (4), $(P, \cdot, \cdot, N)$ by (3), and $(\cdot, N, \cdot, P)$ and $(N, \cdot, P, \cdot)$ by (2).
  Cases (N, N, N, N) and (P, P, P, P) follow by (2).
  For (N, P, N, P), (N, P, N, N), and (P, P, N, P),
  $\Delta^2 \geq f(x_1 + 1, x_2) - (f(x_1 + 1, x_2) - c_M) - [f(x_1 + 1, x_2 + 1) - (f(x_1 + 1, x_2 + 1) - c_M)] = 0.$

- Cases $(A, D, \cdot, \cdot)$ and $(\cdot, \cdot, A, D)$ are excluded by (2), $(D, \cdot, A, \cdot)$ and $(\cdot, D, \cdot, A)$ by (4), and $(D, \cdot, \cdot, A)$ by (3).
  Cases (A, A, A, A) and (D, D, D, D) follow by (2).
  For (A, A, D, D), (D, A, D, D), and (A, A, D, A),
  $\Delta^4 \geq f(x_1 + 1, x_2 + 1) - f(x_1, x_2 + 1) - [(f(x_1 + 1, x_2 + 1) - c_D) - (f(x_1, x_2 + 1) - c_D)] = 0.$

Therefore, $TD_1 f(x_1, x_2) - TD_1 f(x_1, x_2 + 1) = 1/\gamma[\lambda_1 \Delta^1 + \mu_1 \Delta^2 + \mu_2 \Delta^3 + \lambda_2 \Delta^4] \geq 0.$

(ii) Suppose that $i = 1$ and $j = 2$.

Let $\Delta^k = D_1 T_k f(x_1, x_2 + 1) - D_1 T_k f(x_1 + 1, x_2)$.

$\Delta^1 \geq 0$ by (3) if $x_1 > 0$ and $\Delta^1 = R + f(0, x_2 + 1) - f(0, x_2 + 1) - D_1 f(0, x_2) \geq 0$ by (5) if $x_1 = 0$.

$\Delta^3 \geq 0$ by (3) if $x_2 > 0$ and $\Delta^3 = D_1 f(x_1 + 1, 0) - D_1 f(x_1 + 1, 0) = 0$ if $x_2 = 0$.

To show $\Delta^2 \geq 0$ and $\Delta^4 \geq 0$, consider production and disposal controls in $(x_1 + 1, x_2 + 1)$, $(x_1, x_2 + 1)$, $(x_1 + 2, x_2)$, and $(x_1 + 1, x_2)$:

- Cases $(P, N, \cdot, \cdot)$ and $(\cdot, \cdot, P, N)$ are excluded by (4), $(P, \cdot, \cdot, N)$ by (2), and $(\cdot, N, \cdot, P)$ and $(N, \cdot, P, \cdot)$ by (3).
  Cases (P, P, P, P) and (N, N, N, N) follow by (3).
  For (N, P, N, P), (P, P, N, P), and (N, P, N, N), $\Delta^2 \geq c_M - c_M = 0.$

- Cases $(A, D, \cdot, \cdot)$ and $(\cdot, \cdot, A, D)$ are excluded by (2), $(A, \cdot, D, \cdot)$ and $(\cdot, A, \cdot, D)$ by (3), and $(A, \cdot, \cdot, D)$ by (4).
  Cases (A, A, A, A), (D, D, D, D), (D, D, D, A), and (D, A, A, A) follow by (3).
  For (D, A, D, A), $\Delta^4 = -c_D + f(x_1 + 1, x_2 + 1) - f(x_1, x_2 + 2) - [-c_D + f(x_1 + 2, x_2) - f(x_1 + 1, x_2 + 1)] \geq$
  $f(x_1, x_2 + 1) - f(x_1 + 1, x_2) - [f(x_1, x_2 + 2) - f(x_1 + 1, x_2 + 1)] \geq 0.$ The inequalities follows by (3).

Suppose that $i = 2$ and $j = 1$. Let $\Delta^k = D_2 T_k f(x_1 + 1, x_2) - D_2 T_k f(x_1, x_2 + 1)$.

$\Delta^1 \geq 0$ by (3) if $x_1 > 0$ and $\Delta^1 = D_2 f(0, x_2) - D_2 f(0, x_2 + 1) \geq 0$ by (4) if $x_1 = 0$.

$\Delta^3 \geq 0$ by (3) if $x_2 > 0$ and $\Delta^3 = f(x_1 + 2, 0) - c_R - f(x_1 + 1, 1) \geq 0$ by (6) if $x_2 = 0$.

To show $\Delta^2 \geq 0$ and $\Delta^4 \geq 0$, consider production and disposal controls in $(x_1 + 1, x_2 + 1)$, $(x_1 + 1, x_2)$, $(x_1, x_2 + 2)$, and $(x_1, x_2 + 1)$:

- Cases $(P, N, \cdot, \cdot)$ and $(\cdot, \cdot, P, N)$ are excluded by (2), $(\cdot, P, \cdot, N)$ and $(P, \cdot, N, \cdot)$ by (3). and $(P, \cdot, \cdot, N)$ by (4).
  Cases (P, P, P, P), (N, N, N, N), (N, P, P, P), and (N, N, N, P) follow by (3).
  For (N, N, P, P), $\Delta^2 = D_2 f(x_1 + 1, x_2) - D_2 f(x_1 + 1, x_2 + 1) \geq 0$ by (4).
  For (N, P, N, P), $\Delta^2 = f(x_1 + 1, x_2 + 1) - (f(x_1 + 2, x_2) - c_M) - [f(x_1, x_2 + 2) - (f(x_1 + 1, x_2 + 1) - c_M)] \geq f(x_1 + 1, x_2) - f(x_1 + 2, x_2) - (f(x_1, x_2 + 1) - f(x_1 + 1, x_2 + 1)) \geq 0$.
  The inequalities follow by (3).

- Cases $(A, D, \cdot, \cdot)$ and $(\cdot, \cdot, A, D)$ are excluded by (4), $(D, \cdot, A, \cdot)$ and $(\cdot, D, \cdot, A)$ by (3), and $(A, \cdot, \cdot, D)$ by (2).
  Cases (A, A, A, A) and (D, D, D, D) follow by (3).
  For (D, A, D, A), (D, A, D, D), and (A, A, D, A),
  $\Delta^4 \geq f(x_1 + 1, x_2 + 1) - c_D - f(x_1 + 1, x_2 + 1) - [f(x_1, x_2 + 2) - c_D - f(x_1, x_2 + 2)] = 0$.
  Since $\Delta^k \geq 0$,
  $k = 1, 2, 3, 4$, $\Delta_i T f(x_i, x_j + 1) - \Delta_i T f(x_i + 1, x_j) = 1/\gamma [\lambda_1 \Delta^1 + \mu_1 \Delta^2 + \mu_2 \Delta^3 + \lambda_2 \Delta^4] \geq 0$.

(iii) $D_i T f(x_i, x_j) \geq D_i T f(x_i, x_j + 1)$ (by (2)) $\geq D_i T f(x_i + 1, x_j)$ by (3).

(iv) $D_1 T_1 f(x_1, x_2) = R1(x_1 = 0) + D_1 f(x_1 - 1, x_2) 1(x_1 > 0) \leq R$ by (5).
For $k = 2, 3, 4$, $D_1 T_k f(x_1, x_2) \leq R$ by (5).
Therefore, $D_1 T f(x_1, x_2) \leq 1/\gamma [-h_1 + (\lambda_1 + \mu_1 + \mu_2 + \lambda_2) R] \leq R$.

(v) Let $\Delta^k = T_k f(x_1 + 1, x_2) - T_k f(x_1, x_2 + 1)$. If $x_1 > 0$, $\Delta^1 \geq c_R$ by (6).
If $x_1 = 0$, $\Delta^1 = R + f(0, 0) - f(0, 1) \geq f(1, 0) - f(0, 1)$ (by (5)) $\geq c_R$ by (6).
It can be easily shown that $\Delta^2 \geq c_R$ and $\Delta^4 \geq c_R$ by (6).
$\Delta^3 \geq c_R$ by (6) when $x_2 > 0$ and $\Delta^3 = f(x_1 + 1, 0) - (f(x_1 + 1, 0) - c_R) = c_R$ when $x_2 = 0$.
Therefore, $T f(x_1 + 1, x_2) - T f(x_1, x_2 + 1) = 1/\gamma [-h_1 + h_2 + \lambda_1 \Delta^1 + \mu_1 \Delta^2 + \mu_2 \Delta^3 + \lambda_2 \Delta^4]$
$\geq 1/\gamma [-h_1 + h_2 + \gamma c_R] = c_R$, since $h_1 = h_2$.

Proof of Theorem 1 :

(i) Since the control with allowing the serviceable and remanufacturable inventories to go to the infinity cannot be optimal, the original problem with infinite state space can be converted into one with finite state space. Since the model has a finite action space and is unichain and aperiodic, the result follows from Theorem 8.4.5 of Puterman (2005).

(ii) From the definition of $P(x_2)$, $-c_M + v(P(x_2) + 1, x_2) \geq v(P(x_2), x_2)$.
Then, by (4), $-c_M + v(P(x_2), x_2) \geq v(P(x_2) - 1, x_2)$, which means it is optimal to produce in $(P(x_2) - 1, x_2)$.
From the definition of $D(x_1)$, $v(x_1, D(x_1) + 1) \leq v(x_1, D(x_1)) - c_D$.
Then, by (4), $v(x_1, D(x_1) + 2) \leq v(x_1, D(x_1) + 1) - c_D$, which means it is optimal to dispose of a returned item in $(x_1, D(x_1) + 1)$.

(iii) From the definition of $P(x_2 + 1)$, $-c_M + v(P(x_2 + 1) + 1, x_2 + 1) \geq v(P(x_2 + 1), x_2 + 1)$.
Then, from (2), $-c_M + v(P(x_2 + 1) + 1, x_2) \geq v(P(x_2 + 1), x_2)$, which means $P(x_2) \geq P(x_2 + 1)$.

(iv) From the definition of $D(x_1)$, $v(x_1, D(x_1) + 1) \leq v(x_1, D(x_1)) - c_D$.
Then, from (2), $v(x_1 + 1, D(x_1) + 1) \leq v(x_1 + 1, D(x_1)) - c_D$, which means $D(x_1) \geq D(x_1 + 1)$.

Proof of Theorem 2 :

We prove (b). Results (a) and (c) can be shown in a similar way. Consider two hybrid systems with c_M^A and c_M^B where $c_M^A < c_M^B$. Consider the following equation established by f^A and f^B :

$$D_1 f^A(x_1, x_2) - c_M^A \geq D_1 f^B(x_1, x_2) - c_M^B. \tag{12}$$

(12) implies that the marginal benefit of producing a new product is smaller with a higher c_M.

T_1, T_2, and T_3 are the same as before and let $T_4 f(x_1, x_2) \equiv f(x_1, x_2 + 1)1(x_2 < M) + (f(x_1, x_2) - c_D)1(x_2 = M)$.

Let $\Delta^k = D_1 T_k f^A(x_1, x_2) - D_1 T_k f^B(x_1, x_2)$.

Then, Δ^1 : If $x_1 > 0$, $\Delta^1 \geq c_M^A - c_M^B$ by (12).

        If $x_1 = 0$, $\Delta^1 = 0 \geq c_M^A - c_M^B$.

   Δ^2 : Consider production controls in states $(x_1 + 1, x_2)^A$, $(x_1, x_2)^A$, $(x_1 + 1, x_2)^B$, and $(x_1, x_2)^B$.

      Cases $(P, N, \cdot, \cdot)$ and $(\cdot, \cdot, P, N)$ are excluded by (4), and $(N, \cdot, P, \cdot)$ and $(\cdot, N, \cdot, P)$ by (12).

      Cases (P, P, P, P) and (N, N, N, N) follow by (12).

      For (N, P, N, P) and (N, P, N, N),

      $\Delta^2 \geq f^A(x_1 + 1, x_2) - (f^A(x_1 + 1, x_2) - c_M^A) - [f^B(x_1 + 1, x_2) - (f^B(x_1 + 1, x_2) - c_M^B)] = c_M^A - c_M^B$.

      Cases (P, P, N, P) and (P, P, N, N) can be shown in a similar way.

   Δ^3 : $\Delta^3 \geq c_M^A - c_M^B$ follows by (12).

   Δ^4 : $\Delta^4 = D_1 f^A(x_1, x_2 + 1) - D_1 f^B(x_1, x_2 + 1)$ when $x_2 < M$. When $x_2 = M$,

      $\Delta^4 = D_1 f^A(x_1, M) - D_1 f^B(x_1, M)$. By (12), both cases $\geq c_M^A - c_M^B$.

Since (12) is preserved under T_1, T_2, T_3, and T_4, (12) is also preserved under T. Finally, by the definition of $P^B(x_2)$, $f^B(P^B(x_2) + 1, x_2) - c_M^B \geq f^B(P^B(x_2), x_2)$. Then, from (12), $f^A(P^B(x_2) + 1, x_2) - c_M^A \geq f^A(P^B(x_2), x_2)$, which means $P^A(x_2) \geq P^B(x_2)$.

\<Appendix B\>

We present the value iteration algorithm to find g. To find g_{H_1}, g_{H_2}, and g_{H_3}, (12) should be replaced with the value iteration equation corresponding to (8), (9), and (10), respectively. Denote by ϵ the termination criterion of the algorithm.

1. **Initialization** : Set $k = 0$. For each state (x_1, x_2), set $v^0(x_1, x_2) = 0$.

2. **Value iteration** : Implement a VI on the current value function estimate v^k:

$$\begin{aligned}
Tv^k(x_1, x_2) = \frac{1}{\gamma}[&- (x_1 h_1 + x_2 h_2) \\
&+ \lambda_1 \{(R + v^k(x_1 - 1, x_2))1(x_1 > 0) + v^k(x_1, x_2)1(x_1 = 0)\} \\
&+ \mu_1 \max\{- c_M + v^k(x_1 + 1, x_2), v^k(x_1, x_2)\} \\
&+ \mu_2 \{(v^k(x_1 + 1, x_2 - 1) - c_R)1(x_2 > 0) + v^k(x_1, x_2)1(x_2 = 0)\} \\
&+ \lambda_2 \max\{v^k(x_1, x_2 + 1), v^k(x_1, x_2) - c_D\}]
\end{aligned} \tag{12}$$

3. **Termination test** : Let $\underline{b_k} = \min_{(x_1, x_2)}\{Tv^k(x_1, x_2) - v^k(x_1, x_2)\}$ and $\overline{b_k} = \max_{(x_1, x_2)}\{Tv^k(x_1, x_2) - v^k(x_1, x_2)\}$.

        If $(\overline{b_k} - \underline{b_k}) \geq \epsilon$ for every state (x_1, x_2), set $v^{k+1}(x_1, x_2) \equiv Tv^k(x_1, x_2) - Tv^k(0, 0)$,

        increase k by one, and go to *Value iteration*. Otherwise, go to *Evaluation*.

4. **Evaluation :** The optimal average profit is approximated by $g \approx (\overline{b_n} + \underline{b_n})/2$. Stop the procedure.