Optimal control of a production-inventory system with product returns

Samuel Vercraene, Jean-Philippe Gayon

Abstract

We consider a production-inventory system that consists of n stages. Each stage has a finite production capacity modelled by an exponential server. The downstream stage faces a Poisson demand. Each stage receives returns of products according to independent Poisson processes that can be used to serve demand. The problem is to control production to minimize discounted (or average) holding and backordering costs. For the single-stage problem ($n = 1$), we fully characterize the optimal policy. We show that the optimal policy is base-stock and we derive an explicit formula for the optimal base-stock level. For the general n-stage problem, we show that the optimal policy is characterized by state-dependent base-stock levels. In a numerical study, we investigate three heuristic policies: the base-stock policy, the Kanban policy and the fixed buffer policy. The fixed-buffer policy obtains poor results while the relative performances of base-stock and Kanban policies depend on bottlenecks. We also show that returns have a non-monotonic effect on average costs and strongly affect the performances of heuristics. Finally, we observe that having returns at the upstream stage is preferable in some situations.

Keywords: Inventory control, Product returns, Multi-echelon systems, Queueing, Markov decision process

1. Introduction

The importance of product returns is growing in supply chains. Customers often can return products a short time after purchase, due to take-back commitments of the supplier. For instance, the proportion of returns is particularly important in electronic business where customers can not touch a product before purchasing it. Customers might also return used products a long time after purchase. This type of return has increased in recent years due to new regulations on waste reduction, especially in Europe. Some industries also encourage returns for economical and marketing reasons. Though different in nature, these two types of returns are similar from an inventory control point of view since they constitute a reverse flow which complicates decision making.

The inventory control literature on product returns is quite abundant (see e.g. Fleischmann et al. (1997); Ilgin and Gupta (2010); Zhou and Yu (2011)). However, most of the literature focusses on single-echelon systems with infinite production capacity. In this paper, we fill this gap by considering a n-stage production/inventory system with finite production capacity and product returns at each stage (see Figure 1). The flow of returns at the finished good (FG) inventory may result from remanufacturing, recycling, repairing or simply returning new products. The flows of returns at the work-in-process (WIP) inventories can also result from disassembly operations. For instance, the Kodak company reuses only some parts of cameras like circuit board, plastic body and lens aperture (Toktay et al., 2000).

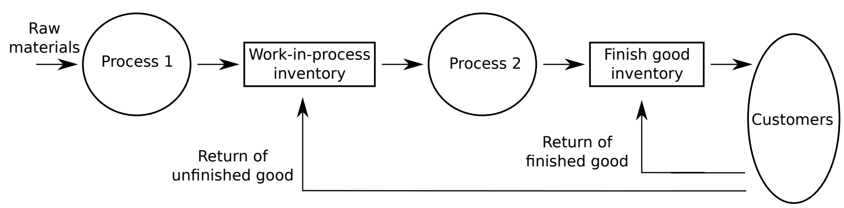

Figure 1: A two-stage production/inventory system with returns.

More precisely, we adopt a queueing framework to model production capacity. Items are produced by servers one by one and each unit requires a random lead-time to be produced. We assume that each stage consists of a single exponential server and an output inventory. The downstream stage faces a Poisson demand. Each stage receives returns of products, according to independent Poisson processes, that can be used to serve demand. The problem is then to control production at each stage, in order to minimize discounted/average holding and backordering costs. We also study the single-stage problem which has not been studied in the literature. In what follows, we review the literature on single-echelon and multi-echelon systems with returns, before presenting in detail our contributions.

The literature on single-echelon systems is quite mature. Heyman (1977) considers an inventory system with independent Poisson demand and Poisson returns. Unsatisfied demands are backordered. Heyman assumes zero lead-times and linear costs for both manufacturing and remanufacturing. These strong assumptions imply that the optimal production policy is a make-to-order policy and that the optimal disposal policy is a simple threshold policy: when the inventory level exceeds a certain disposal threshold R, every returned item is disposed upon arrival. An explicit expression for the optimal disposal threshold is also derived. For a lost sale problem with exponential service times, Poisson demand and returns, Zerhouni et al. (2010) investigate the impact of ignoring dependency between demands and returns.

Fleischmann et al. (2002) consider a similar setting with deterministic manufacturing lead-time and fixed order cost. Again, remanufacturing lead-time and remanufacturing costs are neglected. They extend results standing for a system without returns by showing that the optimal policy is (s, Q) for the average-cost problem. For the periodic review version with a stochastic demand either positive or negative in each period, Fleischmann and Kuik (2003) show the average-cost optimality of an (s, S) policy. Simpson (1978) and Inderfurth (1997) consider a periodic-review problem where returns are held in a separate buffer until they are remanufactured or disposed of. When the remanufacturing lead-time is equal to the production leadtime and the costs are linear, they show that a three-parameter policy is optimal.

Apart from these optimal control papers, several heuristic policies have been investigated in the literature. Van der Laan et al. (1996b) model the remanufacturing shop as an $M/M/c/(c+N)$ queue with c parallel servers and introduce the (s_p, Q_p, N) policy where s_p is the reorder point, Q_p the order quantity and any return is disposed whenever the number of products waiting for repair equals N. Van der Laan et al. (1996a) extend this policy with the (s_p, Q_p, s_d, N) policy where returns are disposed when the stock level is above s_d. Van der Laan and Salomon (1997) consider a model with correlated demand process and return process. They compare an (s_p, Q_p, Q_r, s_d) push-disposal policy with an $(s_p, Q_p, s_r, S_r, s_d)$ pull-disposal policy to coordinate manufacturing and remanufacturing decisions. For the push-disposal policy, returned products are remanufactured with batch size Q_r. For the pull-disposal policy, remanufacturing is initiated only when the finished good inventory is below s_r and the remanufacturable inventory is above S_r. Teunter and Vlachos (2002) complement the numerical study of the above model.

The literature on multi-echelon systems with returns is much more limited. In their seminal work (without returns), Clark and Scarf (1960) studies a series inventory system with n stages, periodic review, linear holding and backorder cost, no setup cost and stochastic demand at the downstream stage. They prove that a base-stock policy is optimal. DeCroix et al. (2005) extend the results of Clark and Scarf (1960) to the case where demand can be negative. They also propose a method to compute a near optimal policy, explain how to extend their model when returns occur at different stages and compare the base-stock policies to fixed-buffer policies. DeCroix (2006) combines the multi-echelon structure of DeCroix et al. (2005) and the remanufacturing structure of Inderfurth (1997). DeCroix and Zipkin (2005) and Decroix et al. (2009) consider assemble-to-order systems with returns of components or finished product.

In production-inventory systems, replenishment is modelled in a different way than in pure inventory systems. Items are produced by servers one by one, or possibly by batches. Each unit, or batch, requires a random lead-time to be produced. Hence replenishments are capacitated in production-inventory systems. In line with this approach, Veatch and Wein (1992) consider a n-stage system with exponential server at each stage. Otherwise, their assumptions are similar to Clark and Scarf (1960). They prove that the optimal policy is never a state-dependent base-stock policy. In another paper, Veatch and Wein (1994) studies the case $n = 2$. They investigate several classes of policies and compare them to the optimal policy. They conclude that the base-stock policy is generally the best heuristic. However, when the downstream station is the bottleneck, the Kanban policy is better. Dallery and Liberopoulos (2003) investigates a generalized Kanban policy being a mix between Kanban and base-stock policy. In a deterministic environment, several papers have investigated capacitated production and/or remanufacturing (see e.g. Nahmias and Rivera (1979); Teunter (2001, 2004); Li et al. (2007)).

In this paper, we extend the model of Veatch and Wein (1992) by including Poisson returns at each stage. We show that the optimal policy is still a complex state-dependent base-stock policy and we derive several monotonicity results for the base-stock levels. Interestingly, the single-echelon problem has not been treated in the literature, when including Poisson returns. In this case, the optimal policy reduces to a simple base-stock policy and we are able to derive an explicit formula for the optimal base-stock level for both average-cost and discounted-cost problems. Such explicit formulas are very rare in inventory control theory, especially when returns are included. When service times, inter-arrival times and inter-return times are not exponential but have general i.i.d. distributions, we explain how to compute the optimal base-stock level by using results from the newsvendor problem.

The optimal policy of the n-stage problem has a complex form and might be difficult to implement in practice. To counter this, we evaluate the performances of three classes of heuristic policies (fixed buffer, base-stock and Kanban) which are reasonable with respect to the optimal policy structure. The fixed-buffer policy obtains poor results while the relative performances of base-stock and Kanban policies depend on bottlenecks, consistently with Veatch and Wein (1996). Moreover, we observe that return rates strongly affect the relative performances of heuristics.

Section 2 describes in detail the n-stage problem. Section 3 provides a full characterization of the optimal policy for the single-stage system. Section 4 shows that the optimal policy for the n-stage system is a state-dependent base-stock policy. Section 5 investigates the performances of three heuristic policies. Finally, we conclude and discuss avenues for research in Section 6.

2. Assumptions and notations

We consider a n-stage production/inventory system in series which satisfies end-customer demand (see Figure 2). Station M_i, $i \in \{1, \ldots, n\}$, produces items one by one. The production lead-time of station M_i is exponentially distributed with rate μ_i. Preemption is allowed and works as follows. The processing of a job at station M_i can be interrupted at any point in time and continued latter. Because of the memoryless property of the exponential distribution, continuing a job is equivalent to restarting it from the beginning. Produced items are stocked in a buffer B_i just after M_i. The end buffer B_n sees customer demands arriving according to a Poisson process with rate λ. We assume that backorders are allowed. At time t, the on-hand inventory at B_i ($1 \leq i < n$) is denoted by $X_i(t)$ and the net on-hand inventory at B_n is denoted by $X_n(t)$. When buffer B_i ($1 \leq i < n$) is empty, M_{i+1} can not start to produce.

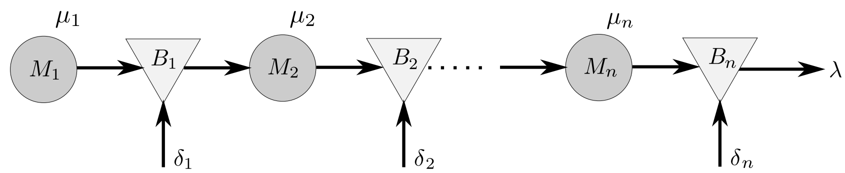

Figure 2: The n-stage $M/M/1$ make-to-stock queue with product returns.

Returns of products occur at buffer B_i according to an independent Poisson process with rate δ_i. When a return is accepted in buffer B_i, it can be used immediately as a new product (we neglect the remanufacturing lead-time). Another way to see these n return flows is to consider a situation where there is a single flow of returns for the whole system (rate $\sum_{i=1}^n \delta_i$) and, after an inspection, returned products are routed to the inventory B_i with probability $p_i = \delta_i/(\sum_{i=1}^n \delta_i)$.

The system is stable if we have the following conditions on the parameters.

$$\sum_{i=1}^{n} \delta_i < \lambda, \tag{1a}$$

$$\lambda < \mu_j + \sum_{i=j}^{n} \delta_i, \quad \forall j \in \{1, \ldots, n\}. \tag{1b}$$

Condition (1a) requires that the demand rate must be larger than the total return rate. Condition (1b) requires that all echelon have to be able to serve the demand. These two conditions can be aggregated into the following inequalities

$$\sum_{i=1}^{n} \delta_i < \lambda < \min_{j \in \{1,\ldots,n\}} \left\{ \mu_j + \sum_{i=j}^{n} \delta_i \right\}. \tag{2}$$

The system incurs in state $\mathbf{X}(t) = (X_1(t), \ldots, X_n(t))$ a cost rate

$$c(\mathbf{X}) = \sum_{i=1}^{n-1} h_i X_i + h_n X_n^+ + b X_n^-$$

where h_i is the inventory holding cost per unit in stock per unit of time at buffer B_i, b is the backorder cost per unit of waiting demand per unit of time, $x^+ = \max[0, x]$ and $x^- = \max[0, -x]$. The unit return cost is c_i^r at stage i.

As the optimal policy is independent of c_i^r, we set without loss of generality $c_i^r = 0$ for $i \in \{1, \ldots, n\}$.

A production policy π specifies when to produce for each stage. The discounted expected cost over an infinite horizon of a policy π, with initial state $\mathbf{x} = (x_1, \ldots, x_n)$ and discount rate $\alpha > 0$, is

$$v_\alpha^\pi(\mathbf{x}) \quad = \quad E\left[\int_0^{+\infty} e^{-\alpha t} c(\mathbf{X}(t))\mathrm{d}t | \mathbf{X}(0) = \mathbf{x}\right].$$

Our objective is to find the optimal policy, denoted by $\pi^\star$, that minimizes the expected discounted cost $v_\alpha^\pi(\mathbf{x})$ over an infinite horizon. We denote by $v_\alpha^\star(\mathbf{x})$ the optimal value function:

$$v_\alpha^\star(\mathbf{x}) = \min_\pi v_\alpha^\pi(\mathbf{x}).$$

We are also interested in the average-cost problem

$$g^\star = \min_\pi \lim_{T\to\infty} \frac{E_\mathbf{x}^\pi\left[\int_0^T c(\mathbf{X}(t))\mathrm{d}t\right]}{T}.$$

There is a strong link between the discounted-cost problem and the average-cost problem. The average-cost optimal policy can be obtained as the limit of the discounted-cost optimal policy when α goes to zero. Moreover, the optimal average cost $g^\star$, is the limit of $\alpha v_\alpha^\star(\mathbf{x})$ when α goes to 0, for each $\mathbf{x}$. To justify these two properties, we use the results of Weber and Stidham (1987) which apply to problems with infinite state space and unbounded costs.

3. A full characterization of the optimal policy for the single-stage problem ($n = 1$)

Before considering the n-stage problem, we analyze the single-stage problem (see Figure 3), for which we are able to fully characterize the optimal policy. Veatch and Wein (1996) and Dusonchet and Hongler (2003) have investigated the single-stage problem. In this section, we extend their results to a single-stage problem including product returns.

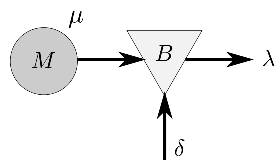

Figure 3: The single-stage $M/M/1$ make-to-stock queue with product returns ($n = 1$).

For the single-stage problem, we denote the system parameters by λ, μ, δ, h, b, α (demand rate, production rate, return rate, holding cost, backorder cost, discount rated). The net on-hand inventory is denoted by x and the cost rate is $c(x) = hx^+ + bx^-$. The problem is again to control production in order to minimize discounted or average costs.

3.1. Structure of the optimal policy

The problem of finding the optimal control policy can be formulated as a continuous-time Markov Decision Process (MDP). After uniformizing the MDP with rate $\tau = \lambda + \mu + \delta$, we can transform the continuous-time MDP into a discrete time MDP (Puterman, 1994). The optimal value function $v_\alpha^\star$ satisfies the following optimality equations:

$$v_\alpha^\star(x) = Tv_\alpha^\star(x), \ \forall x$$

with

$$Tv(x) = \frac{1}{\tau + \alpha}[c(x) + \mu \min[v(x), v(x+1)] + \lambda v(x-1) + \delta v(x+1)]. \quad (3)$$

Theorem 1. *The optimal value function $v_\alpha^\star(x)$ is convex in x. The optimal policy for the discounted-cost problem (respectively the average-cost problem) is base-stock: there exists a base-stock level $z_\alpha^\star$ (respectively $z^\star$) such that it is optimal to produce if the stock level is smaller than $z_\alpha^\star$ (respectively $z^\star$) and to idle production otherwise.*

Proof. We first prove that operator T preserves convexity. Consider a convex value function v, such that $\Delta v(x) = v(x+1) - v(x)$ is non-decreasing in x. By assumption, the cost rate $c(x)$ is also convex. As mentioned by Koole (1998), the function $\min[v(x), v(x+1)]$ is also convex. The functions $v(x-1)$ and $v(x+1)$ are also convex. Finally Tv, as a non-negative linear combination of convex functions, is also convex.

As operator T is a contraction mapping, the fixed point theorem in a Banach space (Puterman, 1994) ensures that any sequence of value functions (v_n) defined as $v_{n+1} = Tv_n$ will converge to the optimal value function $v^\star_\alpha$, the unique solution of the optimality equations $v^\star_\alpha = Tv^\star_\alpha$.

If we take a null value function v_0, it is clear that v_0 is convex. By induction, we conclude that $v^\star_\alpha$ is convex. This property allows to define the threshold $z^\star_\alpha = \min[x : \Delta v^\star_\alpha(x) > 0]$ such that $\Delta v^\star_\alpha(x) \leq 0$ (produce) when $x < z^\star_\alpha$ and $\Delta v^\star_\alpha(x) > 0$ (idle) when $x \geq z^\star_\alpha$. For the average-cost problem, it suffices to use the property that the discounted-cost policy converges to the average-cost policy when α goes to 0 (Weber and Stidham, 1987). $\square$

3.2. Steady-state probabilities

In this subsection, we derive the steady state probabilities when the control policy is base-stock with a base-stock level z. In this case, the net on-hand inventory $X(t)$ evolves according to a continuous-time Markov chain (Figure 4).

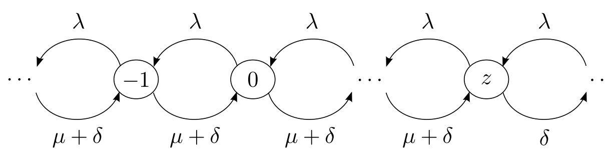

Figure 4: Markov chain in the backorder case with returns case.

Define the ratios $\rho_1 = \frac{\lambda}{\mu+\delta}$ and $\rho_2 = \frac{\delta}{\lambda}$ where ρ_2 will be referred to as the return ratio. To ensure the stability of the number of backorders and the inventory level, we assume that $\rho_1 < 1$ and $\rho_2 < 1$. Let $p(i)$ be the steady-state probability to be in state i. We have

$$p(i) = \begin{cases} \rho_1^{z-i} p(z) & \text{if } i \leq z, \\ \rho_2^{i-z} p(z) & \text{if } i \geq z. \end{cases} \tag{4}$$

Using the normalization condition, $\sum_{i=-\infty}^{\infty} p(i) = 1$, we obtain

$$p(z) = \frac{(1-\rho_1)(1-\rho_2)}{1-\rho_1\rho_2}.$$

3.3. Average-cost problem

When computing the average cost, me must distinguish two cases: $z \geq 0$ and $z \leq 0$. These two cases are symmetrical (interchange h and b, μ_1 and μ_2 and replace z by $-z$). The average on-hand inventory $\bar{X}^+$ and the average number of backlogs $\bar{X}^-$ are given by

$$
\begin{aligned}
\bar{X}^+ &= \sum_{i=0}^{+\infty} ip(i) \\
&= \begin{cases} \sum_{i=0}^{z} i\rho_1^{z-i} p(z) + \sum_{i=z+1}^{+\infty} i\rho_2^{i-z} p(z) & \text{if } z \geq 0, \\ \sum_{i=1}^{+\infty} i\rho_2^{i-z} p(z) & \text{if } z \leq 0, \end{cases} \\
&= \begin{cases} z + p(z)\left[\frac{\rho_1}{(1-\rho_1)^2}\left(\rho_1^z - 1\right) + \frac{\rho_2}{(1-\rho_2)^2}\right] & \text{if } z \geq 0, \\ p(z)\frac{\rho_2^{-z+1}}{(1-\rho_2)^2} & \text{if } z \leq 0, \end{cases} \\
\bar{X}^- &= -\sum_{i=-\infty}^{0} ip(i) \\
&= \begin{cases} -p(z)\sum_{i=-\infty}^{0} i\rho_1^{z-i} & \text{if } z \geq 0, \\ -p(z)\sum_{i=-\infty}^{z-1} i\rho_1^{z-i} - p(z)\sum_{i=z}^{0} i\rho_2^{i-z} & \text{if } z \leq 0, \end{cases} \\
&= \begin{cases} p(z)\frac{\rho_1^{z+1}}{(1-\rho_1)^2} & \text{if } z \geq 0, \\ -z + p(z)\left[\frac{\rho_2}{(1-\rho_2)^2}\left(\rho_2^{-z} - 1\right) + \frac{\rho_1}{(1-\rho_1)^2}\right] & \text{if } z \leq 0. \end{cases}
\end{aligned}
$$

After some algebraic operations, the average cost $g(z) = h\bar{X}^+ + b\bar{X}^-$ can be expressed as

$$
g(z) = \begin{cases} h\left\{z + p(z)\left[\frac{\rho_1}{(1-\rho_1)^2}\left(-1 + \frac{h+b}{h}\rho_1^z\right) + \frac{\rho_2}{(1-\rho_2)^2}\right]\right\} & \text{if } z \geq 0, \\ b\left\{-z + p(z)\left[\frac{\rho_2}{(1-\rho_2)^2}\left(-1 + \frac{b+h}{b}\rho_2^{-z}\right) + \frac{\rho_1}{(1-\rho_1)^2}\right]\right\} & \text{if } z \leq 0, \end{cases} \tag{5}
$$

and

$$
\begin{cases} g(z+1) - g(z) = \frac{(1-\rho_2)[h-(h+b)\rho_1^{z+1}]+h(1-\rho_1)\rho_2}{1-\rho_1\rho_2} & \text{if } z \geq 0, \\ g(z-1) - g(z) = -\frac{(1-\rho_1)[b-(b+h)\rho_2^{-z+1}]+b(1-\rho_2)\rho_1}{1-\rho_2\rho_1} & \text{if } z \leq 0. \end{cases}
$$

The quantity $\Delta g(z) = g(z+1) - g(z)$ is increasing in z, which implies that $g(\cdot)$ is convex. Hence the average cost is minimized for $z^\star = \min[z : \Delta g(z) > 0]$. This property implies the following theorem.

Theorem 2. *The average-cost optimal base-stock level $z^\star$ is*

$$z^* = \begin{cases} \left\lfloor \dfrac{\ln\left(\frac{1-\rho_1\rho_2}{1-\rho_2}\frac{h}{h+b}\right)}{\ln \rho_1} \right\rfloor \geq 0 & \text{if} \quad \frac{1-\rho_1\rho_2}{1-\rho_2}\frac{h}{h+b} \leq 1, \\ \left\lceil -\dfrac{\ln\left(\frac{1-\rho_2\rho_1}{1-\rho_1}\frac{b}{b+h}\right)}{\ln \rho_2} \right\rceil \leq 0 & \text{else.} \end{cases}$$

Based on Theorem 2, we can easily establish several properties of the optimal base-stock level. First, $z^\star$ is a decreasing function of the return rate δ. When the return rate is increasing, it is better off diminishing the base-stock level in order to limit excess inventory. When $\delta = 0$, we re-obtain the result obtained by Veatch and Wein (1996) in a system without returns:

$$z^\star = \left\lfloor \frac{\ln \frac{h}{h+b}}{\ln \frac{\lambda}{\mu}} \right\rfloor \text{ if } \delta = 0.$$

When δ goes to λ, ρ_2 goes to 1 and $z^\star$ goes to infinity. In presence of returns, the base-stock level can take any negative integer value. Without returns, the optimal base-stock level is always non-negative.

3.4. Discounted-cost problem

It is more complex to compute analytically the optimal base-stock level in the discounted cost case. Denote by $v_\alpha^z(x)$ the expected discounted cost when the base-stock level is z, the initial inventory level is x and the discount rate is α. The following lemma establishes an explicit formula for the discounted cost.

Lemma 1.

$$v_\alpha^z(z) = \begin{cases} \frac{h}{\alpha}\left(z + \alpha B\left[\frac{\beta_1}{(1-\beta_1)^2}\left(-1+\frac{h+b}{h}\beta_1^z\right) + \frac{\beta_2}{(1-\beta_2)^2}\right]\right) & \text{if } z \geq 0, \\ \frac{b}{\alpha}\left(-z + \alpha B\left[\frac{\beta_2}{(1-\beta_2)^2}\left(-1+\frac{b+h}{b}\beta_2^{-z}\right) + \frac{\beta_1}{(1-\beta_1)^2}\right]\right) & \text{if } z \leq 0. \end{cases} \tag{6}$$

where

$$B = \frac{1}{\alpha}\frac{(1-\beta_1)(1-\beta_2)}{1-\beta_1\beta_2},$$
$$\beta_1 = \frac{\alpha+\lambda+\delta+\mu-\sqrt{(\alpha+\lambda+\mu+\delta)^2-4\lambda(\mu+\delta)}}{2(\mu+\delta)},$$
$$\beta_2 = \frac{\alpha+\lambda+\delta-\sqrt{(\alpha+\lambda+\delta)^2-4\lambda\delta}}{2\lambda}.$$

The proof of this lemma is provided in appendix.

When α goes to 0, β_1 goes to ρ_1, β_2 goes to ρ_2 and αB goes to $p(z)$. Therefore $\alpha v_\alpha^z(z)$ goes to the average cost $g(z)$, given in Equation (5), consistently with Weber and Stidham (1987).

We have $v_\alpha^\star(z_\alpha^\star) = \min_z v_\alpha^z(z)$. Similarly to the average-cost problem, we have

$$\begin{cases} v_\alpha^z(z+1) - v_\alpha^z(z) = \frac{1}{\alpha}\frac{(1-\beta_2)[h-(h+b)\beta_1^{z+1}]+h(1-\beta_1)\beta_2}{1-\beta_1\beta_2} & \text{if } z \geq 0, \\ v_\alpha^z(z-1) - v_\alpha^z(z) = -\frac{1}{\alpha}\frac{(1-\beta_1)[b-(b+h)\beta_2^{-z+1}]+b(1-\beta_2)\beta_1}{1-\beta_2\beta_1} & \text{if } z \leq 0. \end{cases}$$

The quantity $\Delta v_\alpha^z(z) = v_\alpha^z(z+1) - v_\alpha^z(z)$ is increasing in z and again $z_\alpha^\star = \min[z : \Delta v_\alpha^z(z) > 0]$.

Theorem 3. *The optimal base-stock level $z_\alpha^\star$ of the discounted problem is*

$$z_\alpha^\star = \begin{cases} \left\lfloor \dfrac{\ln\left(\frac{1-\beta_1\beta_2}{1-\beta_2}\frac{h}{h+b}\right)}{\ln \beta_1} \right\rfloor \geq 0 & \textit{if} \quad \frac{1-\beta_1\beta_2}{1-\beta_2}\frac{h}{h+b} \leq 1, \\ \left\lceil -\dfrac{\ln\left(\frac{1-\beta_2\beta_1}{1-\beta_1}\frac{b}{b+h}\right)}{\ln \beta_2} \right\rceil \leq 0 & \textit{else.} \end{cases}$$

Theorem 3 is consistent with Theorem 2: When α goes to 0, $z_\alpha^\star$ goes to $z^\star$. Theorem 3 is also consistent with the results of Dusonchet and Hongler (2003) who consider the case without returns ($\delta = 0$).

3.5. General distributions

In this subsection only, we relax the assumption of exponential distributions and simply assume that service times, inter-arrival times and inter-return times are identically and independently distributed. In this case, the optimal policy can be very complicated and we focus on the class of base-stock policies. Consider a base-stock policy with base-stock level z. Let $X(t)$ be the net on-hand inventory level at time t and define $N(t) = z - X(t)$. The probability distribution of $N(t)$ is independent of z. Denote by $p(i)$ the steady-state probability of $N(t)$ and by $F(i) = \sum_{x=-\infty}^{i} p(x)$ the cumulative distribution function. The average cost is then

$$\begin{aligned} g(z) &= hE(X^+) + bE(X^-) = hE(z-N)^+ + bE(z-N)^- \\ &= h\sum_{i=0}^{z}(z-i)p(i) + b\sum_{i=z}^{\infty}(i-z)p(i). \end{aligned}$$

We recognize the objective function of a newsboy problem where the order quantity is z, the stochastic demand is N, the shortage cost is b and the holding cost is h. The optimal order quantity for the newsboy model is

$$z^\star = \min\left[z : F(z) > \frac{b}{h+b}\right]. \tag{7}$$

In the special case of an $M/M/1$ make-to-stock queue with Poisson returns, (7) yields to Theorem 2. For other distributions, numerical methods or simulation can be used to compute $F(\cdot)$.

When considering a lost-sale version of our problem, it can be shown that the optimal policy is base-stock. However, it is not possible to derive closed-form expressions for the optimal base-stock level.

4. A partial characterization of the optimal policy for the n-stage problem

The n-stage problem is more complex to analyze and it seems intractable to fully characterize the optimal policy. In this section, we provide some characteristics of the optimal policy, before investigating the performances of several heuristics, in the next section.

Again, the n-stage problem can be formulated as a continuous-time Markov Decision Process (MDP). After uniformizing the MDP with rate

$$\tau = \lambda + \sum_{i=1}^{n}(\mu_i + \delta_i),$$

we can transform the continuous-time MDP into a discrete time MDP (Puterman, 1994). The optimal value function $v_\alpha^\star$ satisfies the following optimality equations:

$$v_\alpha^\star(\mathbf{x}) = Tv_\alpha^\star(\mathbf{x}),\ \forall \mathbf{x}$$

with

$$\begin{array}{l} Tv(\mathbf{x}) = \frac{1}{\tau+\alpha}\left[c(\mathbf{x}) + \lambda v(\mathbf{x}-\mathbf{e}_n) + \sum_{i=1}^n (\delta_i v(\mathbf{x}+\mathbf{e}_i) + \mu_i T_i v(\mathbf{x}))\right], \\ T_1 v(\mathbf{x}) = \min[v(\mathbf{x}), v(\mathbf{x}+\mathbf{e}_1)], \\ T_i v(\mathbf{x}) = \begin{cases} \min[v(\mathbf{x}), v(\mathbf{x}-\mathbf{e}_{i-1}+\mathbf{e}_i)] & \text{if } x_{i-1} > 0, \\ v(\mathbf{x}) & \text{else,} \end{cases} \qquad \forall i \in \{2,\dots,n\}. \end{array}$$

In the optimality equations, let $\mathbf{e}_i = (0,\dots,0,1,0,\dots,0)$ be for vectors of dimension n with the "1" in the i^{th} position. In order to derive structural properties of the optimal policy, we will show that the optimal value function belongs to the following set of value functions V.

Definition 1. *A value function v belongs to V if for all $\mathbf{x}+\mathbf{d}_i, \mathbf{x}+\mathbf{d}_j \in \mathbb{N}^{n-1}\times\mathbb{Z}$, and for all $0 \le i < j \le n$*

$$v(\mathbf{x}) + v(\mathbf{x}+\mathbf{d}_i+\mathbf{d}_j) \le v(\mathbf{x}+\mathbf{d}_i) + v(\mathbf{x}+\mathbf{d}_j), \tag{8}$$

with $\mathbf{d}_0 = \mathbf{e}_1$, $\mathbf{d}_k = \mathbf{e}_{k+1} - \mathbf{e}_k$ for all $k \in \{1,\dots,n-1\}$, and $\mathbf{d}_n = -\mathbf{e}_n$.

In Koole (2006), the property presented in equation (8) is called multimodularity in direction i and j and denoted by $MM(i,j)$.

The following theorem shows that the optimal value function satisfies this property and consequently provides a characterization of the optimal policy.

Theorem 4. *The optimal value function $v_\alpha^\star$ belongs to V and the discounted-cost optimal policy is a state-dependent base-stock policy. There exists n switching surfaces $z_i^\star(\mathbf{x})$ such that*

- *Produce at M_1 if and only if $x_1 < z_1^\star(\mathbf{x})$. Moreover $z_1^\star(\mathbf{x}) - 1 \le z_1^\star(\mathbf{x}+\mathbf{e}_2) \le z_1^\star(\mathbf{x})$.*

- *For all $i \in \{2,\dots,n\}$, produce at M_i if and only if $x_i < z_i^\star(\mathbf{x})$. Moreover $z_i^\star(\mathbf{x}) \le z_i^\star(\mathbf{x}+\mathbf{e}_{i-1})$.*

Proof. The proof is again by induction (see proof of Theorem 1). The operator T is a linear combination of operators that propagate $MM(i,j)$ for all $0 \leq i < j < n$ (Koole, 2006). As a result, if a value function v is in V, then the value function Tv is also in V. By induction, we conclude that $v_\alpha^\star \in V$.

Koole (2006) shows that multimodularity $MM(i,j)$ for all $0 \leq i < j \leq n$ implies the following properties for all $i, j \in \{1, \ldots n\}$:

(a) $v(\mathbf{x} + \mathbf{e}_i) - v(\mathbf{x}) \leq v(\mathbf{x} + \mathbf{e}_i + \mathbf{e}_j) - v(\mathbf{x} + \mathbf{e}_j)$,

(b) $v(\mathbf{x} + \mathbf{e}_j) - v(\mathbf{x} + \mathbf{e}_i) \geq v(\mathbf{x} + \mathbf{e}_j + \mathbf{e}_i) - v(\mathbf{x} + 2\mathbf{e}_i)$,

Property (a) is called supermodularity and denoted by $Super(i,j)$. Property (b) is called superconvexity and denoted by $SuperC(i,j)$.

As $v_\alpha^\star \in V$, we can define the threshold $z_i^\star(\mathbf{x})$ for all $i \in \{1, \ldots, n\}$. The threshold $z_1^\star(\mathbf{x}) = \min[x_1 | v(\mathbf{x} + \mathbf{e}_1) - v(\mathbf{x}) > 0]$ is well defined since v is $Super(1,1)$. In the same way, the threshold $z_i^\star(\mathbf{x}) = \min[x_i | v(\mathbf{x} + \mathbf{e}_i - \mathbf{e}_{i-1}) - v(\mathbf{x}) > 0]$ for all $i \in \{2, \ldots, n\}$ is also well defined since v is $SuperC(i-1,i)$.

The monotonicity results on the switching curves are also implied by the fact that $v_\alpha^\star \in V$. For instance, $Super(1,2)$ ensures that $z_1^\star(\mathbf{x} + \mathbf{e}_2) \leq z_1^\star(\mathbf{x})$ and $SuperC(1,2)$ ensures that $z_1^\star(\mathbf{x}) - 1 \geq z_1^\star(\mathbf{x} + \mathbf{e}_2)$. The other monotonicity results are obtained in a similar way. $\square$

Note that in Theorem 4, the switching surface $z_i^\star$ does not depend on x_i. We write it as a function of $\mathbf{x}$ to simplify the notations.

The structure of the optimal policy pertains to the average-cost problem as explained at the end of Section 2. With lost sales instead of backorders, it can be shown similarly that the optimal policy has the same structure. Figure 5 illustrates Theorem 4 on a numerical example with two stages. The computational procedure to obtain this curve is explained in appendix.

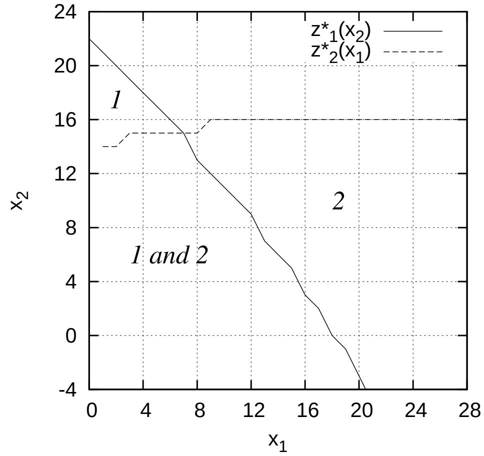

Figure 5: An illustration of the average cost optimal policy with $n = 2$ ($\mu_1 = 0.5$, $\mu_2 = 0.8$, $\delta_1 = 0.3$, $\delta_2 = 0.3$, $\lambda = 1$, $h_1 = 1$, $h_2 = 2$, $b = 4$). Number i means that station M_i produces in this region.

5. Heuristic policies

We can observe on Figure 5 that, the optimal policy of the multi-echelon problem has a complex form and might be difficult to implement in practice. To counter this, we evaluate the performances of three simple and classical heuristic policies: the fixed buffer policy, the base-stock policy and the Kanban policy. To be numerically tractable, we limit our study to the case with two stages ($n = 2$). In this case, each heuristic can be described by two parameters z_1 and z_2. The production control of each class of policies is detailed in Table 1 and illustrated in Figure 6.

Policy	Produce at station M_1 when	Produce at station M_2 when $x_1 > 0$ and
Fixed-buffer	$x_1 < z_1$	$x_2 < z_2$
Base-stock	$x_1 + x_2 < z_1$	$x_2 < z_2$
Kanban	$x_1 + x_2^+ < z_1$	$x_2 < z_2$
Optimal	$x_1 < z_1^\star(x_2)$	$x_2 < z_2^\star(x_1)$

Table 1: Production control policies.

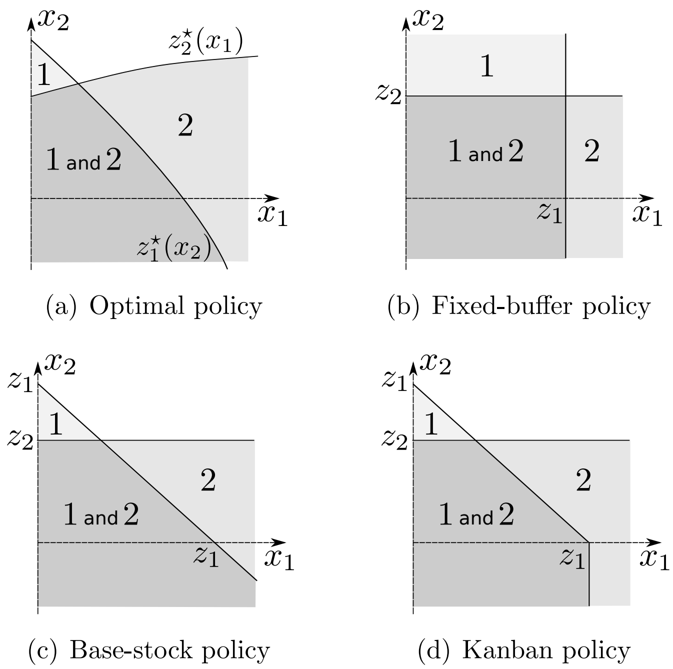

Figure 6: Illustration of policies. Number i means that station M_i produces in this region.

In each class of policies, we compute the optimal average-cost policy parameter values (details on the computational procedure are given at the end of the appendix) for all combinations of the following values:

$$\lambda = \{1\}, \mu_1 = \{1, 1.5, 2\}, \mu_2 = \{1, 1.5, 2\},$$

$$\delta_1 = \{0, 0.3, 0.6, 0.8\}, \delta_2 = \{0, 0.3, 0.6, 0.8\},$$

$$h_1 = \{1\}, h_2 = \{0.5, 1, 10\}, b = \{0.5, 1, 10, 100\}.$$

If we restrict to systems satisfying the stability condition (2), we obtain results for 912 instances summarized in Table 2. In this table, Δg is the % cost increase for using a heuristic policy (with parameters set optimally) instead of the optimal policy.

	Fixed-buffer	Base-stock	Kanban
% of instances where the heuristic is the best	0.6	68.0	31.4
Average Δg (%)	29.0	3.8	9.8
Minimum Δg (%)	0.60	0.0	0.0
Maximum Δg (%)	590	51.4	150
% of instances with $\Delta g \in [0\%; 1\%[$	0.4	45.6	25.4
% of instances with $\Delta g \in [1\%; 5\%[$	14.2	31.4	26.6
% of instances with $\Delta g \in [5\%; 10\%[$	5.6	11.2	13.7
% of instances with $\Delta g > 5\%$	79.8	11.8	34.3

Table 2: Quantitative comparison of heuristics.

We observe that the base-stock policy is generally the best heuristic and outperforms other policies in 68% of cases, with a Δg less than 10% in 88% of cases. The fixed-buffer policy is the worst by far and is outperformed by other policies in 99.4% of cases. The Kanban policy is the best heuristic in 31.4%.

When station M_1 is the bottleneck ($\mu_2/(\mu_1 + \delta_1) \ll 1$), the base-stock policy generally performs better than the Kanban policy. It is the reverse when station M_2 is the bottleneck. These results are consistent with Veatch and Wein (1994) who treat the problem without returns. In Figure 7, we consider the influence of δ_1 on the performance of base-stock and Kanban policies. We have chosen an instance such that Station 1 is the bottleneck when δ_1 is small and Station 2 is the bottleneck when δ_1 is large. Expectedly, we observe that the base-stock policy performs better when Station 1 is the bottleneck and the Kanban policy performs better when Station 2 is the bottleneck.

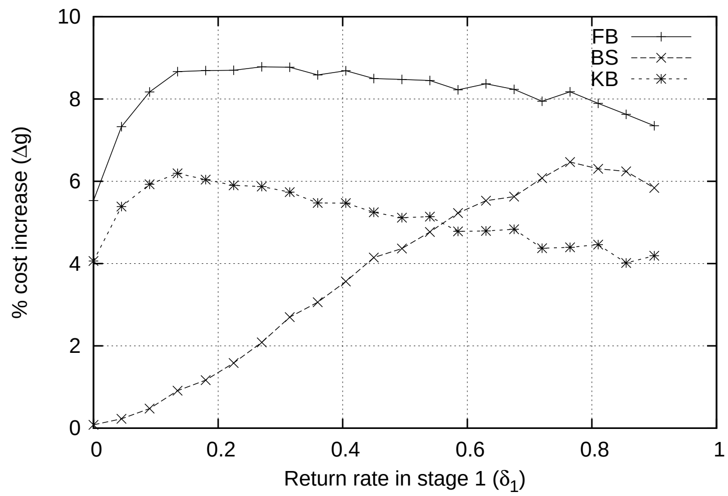

Figure 7: Influence of returns on the performances of heuristics ($\mu_1 = 1.1, \mu_2 = 1.2, \lambda = 1, \delta_2 = 0, h_1 = 1, h_2 = 5, b = 4$).

In Figure 8, we observe that return rates have a non-monotonic effect on average costs. To explain this behavior, let's rewrite the stability condition (2) as

$$\lambda - \mu_1 - \delta_2 < \delta_1 < \lambda - \delta_2,$$
$$\lambda - \mu_2 < \delta_2 < \lambda - \delta_1.$$

When δ_1 (resp. δ_2) goes to $\lambda - \delta_2$ (resp. to $\lambda - \delta_1$), the total average on-hand inventory goes to infinity, so does the average cost. On the other hand, when δ_1 (respectively δ_2) decreases to $\lambda - \mu_1 - \delta_2$ (respectively to $\lambda - \mu_2$), the average number of backorders goes to infinity, so does the average cost. These phenomenons do not appear when production is uncapacitated (Fleischmann et al., 2002).

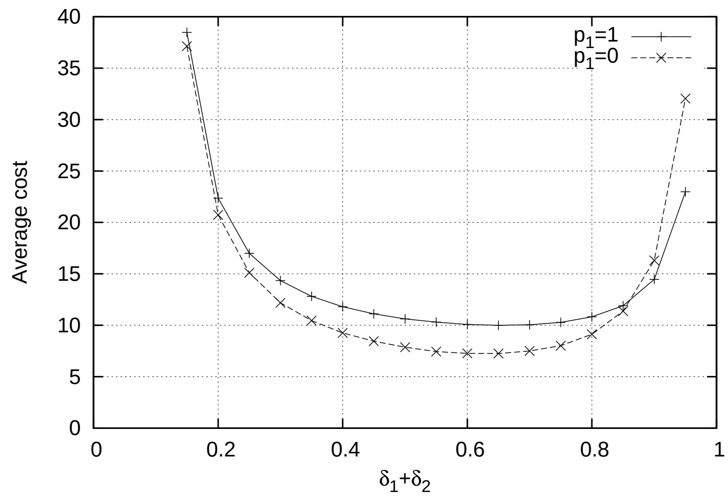

Figure 8: Optimal average cost in function of type and quantity of returns ($\mu_1 = 0.9$, $\mu_2 = 1.5$, $\lambda = 1$, $h_1 = 1$, $h_2 = 2$, $b = 4$), with $p_1 = \delta_1/(\delta_1 + \delta_2)$.

On the same figure, we observe that when the return rate is high ($\delta_1+\delta_2 \geq 0.9$), it is preferable to return products in the first stage. In this case, returned products have to stay a long time in the system before being consumed by demand. So the system prefers to keep returns in the queue with the lowest holding cost ($h_1 < h_2$). When the return rate is smaller ($\delta_1 + \delta_2 \leq 0.9$), it is preferable to have returns at stage 2 in order to satisfy the demand quickly.

6. Conclusions and future research

In this paper, we consider a n-stage production-inventory system with returns. Unlike most of the literature on inventory control with returns, we assume that production is capacitated. To model production capacity, we adopt a queuing framework.

Interestingly, the single-echelon make-to-stock queue problem has not been treated in the literature, when including Poisson returns. In this case, the optimal policy reduces to a simple base-stock policy and we are able to derive an explicit formula for the optimal base-stock level for both average-cost and discounted-cost problems.

For the n-stage problem, we show that the optimal policy is characterized by n switching surfaces with several monotonicity properties. Based on this characterization, we investigate the performances of three heuristics for the case $n = 2$. The fixed-buffer policy obtains poor results while the relative performances of base-stock and Kanban policies depend on bottlenecks. We also show that returns have a non-monotonic effect on average costs and strongly affect the performances of heuristics. Finally, we observe that having returns at the upstream stage is preferable in some situations.

In this paper, we have assumed that returns where always accepted in the system. A first avenue for research is to control arrivals of returns. A return can be either accepted with an acceptance cost or rejected with a rejection cost. For the single-stage problem, the optimal policy is likely to be an (R, S) policy stating to accept returns when the inventory level is below R and to produce when the inventory level is below S. For the n-stage problem, the optimal policy should be characterized by n production/idle switching surfaces and n accept/reject switching surfaces. Another avenue for research is to model explicitly the remanufacturing process. In this case, returned products are first kept in a remanufacturable inventory, before being remanufactured.

A. Clark and H. Scarf. Optimal policies for a multi-echelon inventory problem. *Management Science*, 6(4):475–490, 1960.

Y. Dallery and G. Liberopoulos. Comparative modelling of multi-stage production-inventory control policies with lot sizing. *International Journal of Production Research*, 41(6):1273–1298, 2003.

G. DeCroix. Optimal policy for a multiechelon inventory system with remanufacturing. *Operations Research*, 54(3):532–543, May 2006.

G. DeCroix and P. Zipkin. Inventory management for an assembly system with product or component returns. *Management Science*, 51(8):1250–1265, 2005.

G. DeCroix, J. Song, and P. Zipkin. A series system with returns: Stationary analysis. *Operations Research*, 53(2):350–362, 2005.

G. Decroix, J. Song, and P. Zipkin. Managing an assemble-to-order system with returns. *Manufacturing & service operations management*, 11(1):144–159, 2009.

F. Dusonchet and M. Hongler. Continuous time restless bandit and dynamic scheduling for make to stock production. *IEEE Transactions on Robotics and Automation*, 19(6):977–980, 2003.

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

M. Fleischmann, J. Bloemhof-ruwaard, R. Dekker, J. Van Nunen, and L. Van Wassenhove. Quantitative model for reverse logistics : A review. *European Journal of Operational Research*, 103(1):1–17, 1997.

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

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

M. Ilgin and S. Gupta. Environmentally conscious manufacturing and product recovery (ecmpro): A review of the state of the art. *Journal of Environmental Management*, 91(3):563 – 591, 2010.

K. Inderfurth. Simple optimal replenishment and disposal policies for a product recovery system with leadtimes. *OR Spectrum*, 19:111–122, 1997.

Koole. Monotonicity in markov reward and decision chains: Theory and applications. *Foundations and Trends in Stochastic Systems*, 1:1–76, January 2006.

G. Koole. Structural results for the control of queueing systems using event-based dynamic programming. *Queueing Systems*, 30(3):323–339, December 1998.

Y. Li, J. Chen, and X. Cai. Heuristic genetic algorithm for capacitated production planning problems with batch processing and remanufacturing. *International Journal of Production Economics*, 105(2):301 – 317, 2007.

S. Nahmias and H. Rivera. A deterministic model for a repairable item inventory system with a finite repair rate. *International Journal of Production Research*, 17:215–221, 1979.

M. Puterman. *Marckov Decision Processes, Discrete stochastic, Dynamic programming.* ed. John Wiley & Sons, Inc., 1994.

V. Simpson. Optimum solution structure for a repairable inventory problem. *Operations Research*, 26(2):270–281, 1978.

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

R. Teunter. Lot-sizing for inventory systems with product recovery. *Computers & Industrial Engineering*, 46:431–441, 2004.

R. Teunter and D. Vlachos. On the necessity of a disposal option for returned items that can be remanufactured. *International journal of production economics*, 75:257–266, 2002.

L. Toktay, L. Wein, and S. Zenios. Inventory management of remanufacturable products. *Management Science*, 46(11):1412–1426, 2000.

E. Van der Laan and M. Salomon. Production planning and inventory control with remanufacturing and disposal. *European Journal of Operational Research*, 102:264–678, 1997.

E. Van der Laan, R. Dekker, and M. Salomon. Product remanufacturing and disposal: A numerical comparison of alternative control strategies. *international journal of production economics*, 45:489–498, 1996a.

E. Van der Laan, R. Dekker, M. Salomon, and A. Ridder. An (s,q) inventory model with remanufacturing and disposal. *International journal of production economics*, 46-47:339–350, 1996b.

M. Veatch and L. Wein. Monotone control of queueing networks. *Queueing Systems*, 12(3):391–408, 1992.

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

M. Veatch and L. Wein. Scheduling a make-to-stock queue: index policies and hedging points. *Operations Research*, 44(4):634–647, 1996.

R. Weber and S. Stidham. Optimal control of service rates in networks of queues. *Advanced applied probabilities*, 19, 1987.

H. Zerhouni, J.-P. Gayon, and Y. Frein. Influence of dependency between demands and returns in a reverse logistics system. *International Journal of Production Economics*, In Press, Corrected Proof, 2010.

S. Zhou and Y. Yu. Optimal product acquisition, pricing, and inventory management for systems with remanufacturing. *Operations Research*, 59 (2):514–521, 2011.

Appendix A. Proof of Lemma 1

We have

$$
\begin{aligned}
v_\alpha^z(z) &= E\left[\int_0^\infty e^{-\alpha t} c(X(t))dt | X(0) = z\right] \\
&= \int_0^\infty e^{-\alpha t} E\left[c(X(t))|X(0) = z\right] dt \\
&= \int_0^\infty e^{-\alpha t} \left\{\sum_y c(y) P[X(t) = y|X(0) = z]\right\} dt \\
&= \sum_y \left\{c(y) \int_0^\infty e^{-\alpha t} P[X(t) = y|X(0) = z] dt\right\}.
\end{aligned}
$$

Let $p_y(t) = P[X(t) = y|X(0) = z]$ be the transient probability to be in state y at time t, when the initial state is z. Let $\tilde{p}_y(\alpha)$ be the Laplace transform of $p_y(t)$:

$$\tilde{p}_y(\alpha) = \int_0^\infty e^{-\alpha t} p_y(t) dt.$$

Then

$$
\begin{aligned}
v_\alpha^z(z) &= \sum_y c(y) \int_0^\infty e^{-\alpha t} p_y(t) \\
&= \sum_y c(y) \tilde{p}_y(\alpha).
\end{aligned}
$$

In order to compute $\tilde{p}_y(\alpha)$, we write the differential equations on transient probabilities:

$$
\begin{aligned}
p_y' &= -(\lambda + \delta) p_y + \lambda p_{y+1} + \delta p_{y-1} && \text{if } y > z, \\
p_y' &= -(\lambda + \delta) p_y + \lambda p_{y+1} + (\mu + \delta) p_{y-1} && \text{if } y = z, \\
p_y' &= -(\lambda + \mu + \delta) p_y + \lambda p_{y+1} + (\mu + \delta) p_{y-1} && \text{if } y < z,
\end{aligned}
$$

where $p_y'(t)$ denotes the first derivative of $p_y(t)$. By taking the Laplace transform of the previous set of differential equations, we obtain

$$(\alpha + \lambda + \delta)\tilde{p}_y = \lambda \tilde{p}_{y+1} + \delta \tilde{p}_{y-1} \qquad \text{if } y > z, \tag{A.1}$$

$$(\alpha + \lambda + \delta)\tilde{p}_y = 1 + \lambda \tilde{p}_{y+1} + (\mu + \delta)\tilde{p}_{y-1} \qquad \text{if } y = z, \tag{A.2}$$

$$(\alpha + \lambda + \mu + \delta)\tilde{p}_y = \lambda \tilde{p}_{y+1} + (\mu + \delta)\tilde{p}_{y-1} \qquad \text{if } y < z. \tag{A.3}$$

(A.1) and (A.3) are second-order linear recurrence and have the following solutions

$$\begin{cases} \tilde{p}_y(\alpha) = A_1\alpha_1^{z-y} + B_1\beta_1^{z-y} & \text{if } y \le z, \\ \tilde{p}_y(\alpha) = A_2\alpha_2^{y-z} + B_2\beta_2^{y-z} & \text{if } y \ge z, \end{cases} \tag{A.4}$$

where α_1, β_1 are the roots of the characteristic equation

$$(\mu+\delta)x^2 - (\alpha+\lambda+\mu+\delta)x + \lambda = 0, \tag{A.5}$$

and α_2, β_2 are the roots of another characteristic equation

$$\lambda x^2 - (\alpha+\delta+\lambda)x + \delta = 0. \tag{A.6}$$

Solving quadratic equations (A.5) and (A.6) gives

$$\begin{matrix}\alpha_1 \\ \beta_1\end{matrix} = \frac{\alpha+\lambda+\delta+\mu \pm \sqrt{(\alpha+\lambda+\mu+\delta)^2 - 4\lambda(\mu+\delta)}}{2(\mu+\delta)},$$

and

$$\begin{matrix}\alpha_2 \\ \beta_2\end{matrix} = \frac{\alpha+\lambda+\delta \pm \sqrt{(\alpha+\lambda+\delta)^2 - 4\lambda\delta}}{2\lambda}.$$

We observe that $\alpha_i > 1$ and $0 < \beta_i < 1$ for $i = 1, 2$.

On one hand, we have $\sum_y \tilde{p}_y(\alpha) = 1/\alpha$ since $\sum_y p_y(t) = 1$. On the other hand, we have

$$\sum_y \tilde{p}_y(\alpha) = \sum_{y=-\infty}^{z} (A_1\alpha_1^{z-y} + B_1\beta_1^{z-y}) + \sum_{y=z+1}^{+\infty} (A_2\alpha_2^{y-z} + B_2\beta_2^{y-z}).$$

The convergence of $\sum_y \tilde{p}_y(\alpha)$ implies that $A_1 = A_2 = 0$.

Using (A.4) when $y = z$ gives $\tilde{p}_z(\alpha) = B_1 = B_2 = B$ and we get

$$\begin{cases} \tilde{p}_y(\alpha) = B\beta_1^{z-y} & \text{if } y \le z, \\ \tilde{p}_y(\alpha) = B\beta_2^{y-z} & \text{if } y \ge z. \end{cases}$$

Then (A.2) gives

$$B = \frac{1}{\alpha+\lambda+\delta-(\mu+\delta)\beta_1 - \lambda\beta_2}.$$

As β_1, β_2 respectively satisfy the quadratic equations (A.5) and (A.6), we have

$$\lambda - (\mu + \delta)\beta_1 = \alpha \frac{\beta_1}{1 - \beta_1}, \quad \delta - \lambda\beta_2 = \alpha \frac{\beta_2}{1 - \beta_2}.$$

Then

$$B = \frac{1}{\alpha + \alpha \frac{\beta_1}{1-\beta_1} + \alpha \frac{\beta_2}{1-\beta_2}} = \frac{1}{\alpha} \frac{(1 - \beta_1)(1 - \beta_2)}{1 - \beta_1\beta_2}.$$

Finally, for $z \geq 0$ we obtain

$$\begin{aligned} v_\alpha^z(z) &= -b \sum_{i=-\infty}^{0} i\tilde{p}_i(\alpha) + h \sum_{x=0}^{+\infty} i\tilde{p}_i(\alpha) \\ &= -b \sum_{i=-\infty}^{0} i\beta_1^{z-i} B + h \sum_{i=0}^{z} i\beta_1^{z-i} B + h \sum_{i=z+1}^{+\infty} i\beta_2^{i-z} B \\ &= \frac{h}{\alpha} \left\{ z + \alpha B \left[\frac{\beta_1}{(1 - \beta_1)^2} \left(-1 + \frac{h+b}{h} \beta_1^z \right) + \frac{\beta_2}{(1 - \beta_2)^2} \right] \right\}, \end{aligned}$$

and for $z \leq 0$ we obtain

$$v_\alpha^z(z) = \frac{b}{\alpha} \left\{ -z + \alpha B \left[\frac{\beta_2}{(1 - \beta_2)^2} \left(-1 + \frac{b+h}{b} \beta_2^{-z} \right) + \frac{\beta_1}{(1 - \beta_1)^2} \right] \right\}.$$

Appendix B. Computational procedure

To compute the optimal policy, we truncate the state space in three directions. Let Γ_1 and Γ_2^+ two positive integers and Γ_2^- a negative integer :

$$0 \leq x_1 \leq \Gamma_1 \text{ and } \Gamma_2^- \leq x_2 \leq \Gamma_2^+.$$

We can then apply a value iteration algorithm (Puterman, 1994) to this truncated state. We increase the state space until the average cost is no more sensitive to the truncation level.

In order to evaluate a heuristic policy with parameters (z_1, z_2), we apply the same procedure except that we must change the production operators. For all heuristics, the control is similar at stage 2 and operator T_2 is be replaced by

$$\tilde{T}_2 v(\mathbf{x}) = \begin{cases} v(\mathbf{x} - \mathbf{e}_1 + \mathbf{e}_2) & \text{if } x_1 > 0, \text{ and } x_2 < z_2, \\ v(\mathbf{x}) & \text{else.} \end{cases}$$

At stage 1, the control policy depends on the policy. For base-stock policy, Kanban policy and fixed-buffer policy, operator T_1 is respectively replaced by

$$\tilde{T}_1^{BS} v(\mathbf{x}) = \begin{cases} v(\mathbf{x}+\mathbf{e}_1) & \text{if } x_1 + x_2 < z_1, \\ v(\mathbf{x}) & \text{otherwise,} \end{cases}$$

$$\tilde{T}_1^{KB} v(\mathbf{x}) = \begin{cases} v(\mathbf{x}+\mathbf{e}_1) & \text{if } x_1 + x_2^+ < z_1, \\ v(\mathbf{x}) & \text{otherwise,} \end{cases}$$

$$\tilde{T}_1^{FB} v(\mathbf{x}) = \begin{cases} v(\mathbf{x}+\mathbf{e}_1) & \text{if } x_1 < z_1, \\ v(\mathbf{x}) & \text{otherwise.} \end{cases}$$

Denote by $C^\pi(z_1, z_2)$ the average cost of policy π, with $\pi =$ Kanban, fixed-buffer, base-stock. For each class of policies, we want to find the parameters $z_1^\star, z_2^\star$ minimizing $C(z_1, z_2)$. This optimization problem is a non linear problem with integer variables that might be long to solve since evaluating a given policy might already takes time. Therefore, we make the plausible assumption that the function $C(z_1, z_2)$ is unimodal. This assumption has been checked on several instances. Based on the unimodularity assumption, we can solve efficiently the problem with the maximal gradient with constant step method. This method is very efficient here because we can start the optimization with an approximate value of $z_1^\star$ and $z_2^\star$, resulting from the calculation of the optimal policy.