Optimality of Four-Threshold Policies in Inventory Systems with Customer Returns and Borrowing/Storage Options

Eugene A. Feinberg
Department of Applied Mathematics and Statistics
State University of New York at Stony Brook
Stony Brook, NY 11794-3600
Eugene.Feinberg@sunysb.edu
631-632-7189

Mark E. Lewis
Department of Industrial and Operations Engineering
University of Michigan
1205 Beal Avenue, Ann Arbor, MI 48109-2117
melewis@engin.umich.edu
734-763-0519

submitted October 29, 2003, revised January 30, 2004

Abstract

Consider a single commodity inventory system in which the demand is modelled by a sequence of i.i.d. random variables that can take negative values. Such problems have been studied in the literature under the name *cash management* and relate to the variations of the on-hand cash balances of financial institutions. The possibility of a negative demand also models product returns in inventory systems. This paper studies a model in which, in addition to standard ordering and scrapping decisions seen in the cash management models, the decision maker can borrow and store some inventory for one period of time. For problems with backorders, zero setup costs, and linear ordering, scrapping, borrowing, and storage costs, we show that an optimal policy has a simple four-threshold structure. These thresholds, in a nondecreasing order, are: order-up-to, borrow-up-to, store-down-to, and scrap-down-to levels. That is, if the inventory position is too low, an optimal policy is to order up to a certain level and then borrow up to a higher level. Analogously, if the inventory position is too high, the optimal decision is to reduce the inventory to a certain point after which one should store some of the inventory down to a lower threshold. This structure holds for the finite and infinite horizon discounted expected cost criteria and for the average cost per unit time criterion. We also provide sufficient conditions when the borrowing and storage options should not be used. In order to prove our results for average costs per unit time, we establish sufficient conditions when the optimality equations hold for a Markov decision process with an uncountable state space, noncompact action sets, and unbounded costs.

1 Introduction

Consider a single commodity inventory system for which we do not assume that the demand is nonnegative. Such problems have been studied in the literature under the name *cash management*, and relate to the variations of the on-hand cash flows of financial institutions. For the cash management problem without fixed ordering costs, an optimal policy is defined by two thresholds; if the inventory level is above the higher threshold it should be reduced to this threshold and, similarly, if the inventory level is below the lower threshold, it should be increased up to this threshold (see Eppen and Fama [16] or Section 8.4 in Heyman and Sobel [27]). This is an extension of S-policies (also called "order-up-to" or "basestock" policies) for models with nonnegative demand.

In cash management problems, ordering and scrapping the inventory is associated with financial transactions. In reality, a number of financial instruments are available to a company. In particular, a firm may be able to borrow money for a short period of time. A natural question to investigate is how should short and long term transactions be linked. A similar dilemma also takes place for production/inventory operations; is it better to buy or lease even when leasing is more expensive per unit time? In fact, in inventory systems with product returns, a company can benefit from temporary increases to its inventory (leasing) for a fixed period of time since, in view of possible returns, after this period ordering may not be needed. In this paper we model short-term transactions by one-step borrowing and storage decisions.

We consider the problem when the manager can borrow or store the inventory at any period of time. The borrowed inventory should be returned at the beginning of the next period and the stored inventory is automatically added to the main inventory in the beginning of the next period. We consider the model with back orders. For example, it is possible that the borrowing option is executed but the inventory level becomes negative in the beginning of the next period. In this case, the borrowed amount is returned anyway and the backorder increases. The manager, though, has an option to borrow again. In the model considered in this paper, the purchasing, scrapping, borrowing, and storing costs are linear. There are no fixed costs associated with ordering for any of these operations. All capacities are unconstrained and all lead times are zero.

We show that when the option to borrow or store for one period is available to the decision-maker, an optimal policy is defined by four thresholds. If the inventory position is too high, the optimal decision is to reduce the inventory via scrapping, but only to a certain point, after which one should also store some of the commodity to meet future demand. Analogously, if the inventory position is too low, the decision-maker should order up to a certain level and then borrow from the secondary source to meet potential demand. We also provide sufficient conditions when optimal policies do not use the borrowing and storage options. That is, the production threshold is equal to the borrowing threshold and the similar equality holds for scrapping and storage. In particular, we show that this holds if borrowing is more expensive than backordering in problems with linear holding and backordering costs or if the demand is nonnegative and borrowing is not cheaper than ordering.

The study of inventory control dates back at least to the work of Arrow, Harris and Marshcak [2] and some would say (in the deterministic case) to the work of Harris [23]. With this in mind, we will not attempt a complete review here except to point the reader to the excellent survey article of Porteus [32] that touches many of the high points in the area until 1990. Cash management has also received a considerable amount of attention although much less in the operations research literature than the inventory models with nonnegative demand. Eppen and Fama [16] considered a model with i.i.d. discrete demands with finite support and showed the existence of order-up-to and down-to levels in the finite horizon case for models without setup costs. Girgis [21] and Neave [30] considered both fixed and variable costs for each transaction. The former shows that when there are fixed costs for increasing or decreasing demand (but not both) an optimal policy analogous to (s, S)-policies in inventory control results. One important difference is that between the order-up-to level (S) and the lower limit (s) where no action is usually taken, the optimal policy in the cash management model may order. Furthermore, it is not immediately obvious what the ordering decision in this case might be. The latter paper shows that when both transactions have fixed costs, this analogy need not hold in both directions and provide conditions under which it does hold. All of these results were collected and simplified in [15]. Other generalizations of the cash management problem appeared in [12, 28]. Other models of product returns were studied in [26, 41, 20]. The most recent results on the cash management problem with fixed costs can be found in Chen and Simchi-Levi [9].

Early models that allow for demand to be met by emergency orders, possibly at the expense of higher costs, include those of Daniel [13], Neuts [31] and Barankin [4]. Aneji and Noori [1] discuss a problem in which unmet demand may be met by a secondary source and show that the ordering policy is an (s, S)-policy. Tagaras and Vlachos [40] discuss a periodic review system with the possibility of emergency replenishments with various lead times. Recently, Huggins and Olsen [29] show that when overtime production is available, an (s, S)-policy is still optimal for regular production and various policies are optimal for overtime. Other related models include those of Chiang and Gutierrez [10, 11] and Arslan, Ayhan and Olsen [3].

Section 2 of this paper provides a formal description of the problem and the optimality criteria considered. Section 3 discusses optimal order-up-to and down-to policies in the finite horizon case and provides conditions under which we need not consider the borrowing or storage options. This is continued in Sections 4 and 5 for the infinite horizon discounted and average cost cases, respectively. Section 6 proves the existence of a solution to the *average cost optimality equations* (ACOE) for our problem. Since our search of the literature did not yield sufficient conditions for the validity of the ACOE for our problem, we provide sufficient conditions for a more general Markov decision process (MDP) and show that they hold in our case. We conclude by discussing avenues of future research in Section 7. For the remainder of this section, we explain what sufficient conditions are available in the literature and what is required for our problem.

For MDPs with Borel state spaces and bounded rewards, the ACOE was introduced by Ross [35] and studied by Gubenko and Statland [22], Dynkin and Yushkevich [14], and Fernández-Gaucherand [18]. For problems with unbounded rewards, according to Schäl [37], Proposition 1.3, a stationary policy that, satisfies the average cost optimality *inequalities*, is optimal. Of course, if the optimality equations hold, a stationary policy that satisfies them is optimal as well. Schäl [37] described two groups of conditions, (W) and (S), and proved that each of these conditions imply the validity of the optimality inequalities. Conditions (W) require weak continuity of transition probabilities. Conditions (S) require setwise continuity of transition probabilities; discussed further below. We remark that if the action sets are finite, as was assumed in Ross [35] and Ritt and Sennott [34], then the setwise convergence conditions (S) from Schäl [37] hold and no specific continuity assumption is needed on transition probabilities.

An important feature of many inventory control models considered in the literature is that the control sets may not be compact and are, in fact, assumed to be unbounded. However, Schäl [37] assumed that the action sets are compact. This assumption prevents direct applications of the results from [37] to classic problems with unlimited ordering/scrapping capacities.

We remark that for inventory control problems, conditions (W) are natural while conditions (S) are too strong. Consider a typical inventory control equation

$$x_{n+1} = x_n + a_n - D_{n+1}, \quad n = 0, 1, 2, \dots, \tag{1.1}$$

where x_n is the inventory at the end of period n, a_n is the decision how much should be ordered, and D_n is the demand during period n. Let $q(dy|x,a)$ be the transition probability for the control problem (1.1). That is, $q(B|x,a)$ represents the probability that a subset B of the state space is visited at the next step, given that action a is chosen in state x. Weak continuity of q in Schäl's [37] condition (W) means that $E_{x_n^k,a_n^k} f(x_{n+1}) \to E_{x_n,a_n} f(x_{n+1})$ for any sequence $\{(x_n^k,a_n^k), k \geq 0\}$ of state-action pairs such that $(x_n^k,a_n^k) \to (x_n,a_n)$ for each bounded, continuous function f. This is true in light of (1.1) and Lebesgue's dominated convergence theorem. On the other hand, recall that setwise continuity in Schäl's [37] condition (S) means that $q(B|x_n,a_n^k) \to q(B|x_n,a_n)$ as $a_n^k \to a_n$ for any Borel subset B of the state space. Suppose the demand is deterministic, $D_n = 1$, and $a_n^k = a_n + \frac{1}{k}$, $x_n^k = x_n$. Then $q(B|x_n,a_n) = 1$ for $B = (-\infty, x_n + a_n - 1]$ and $q(B|x_n,a_n^k) = 0$ for all $k = 1, 2, \dots$.

As discussed above, Schäl [37] established the optimality inequalities under both weak and setwise continuity conditions but assumed the compactness of action sets. Hernández-Lerma and Lasserre [25], Chapter 5, and Fernández-Gaucherand, Arapostathis and Marcus [19] presented results for non-compact action sets but assumed setwise convergence. Moreover, Section 5.7 in Hernández-Lerma and Lasserre [25] provides conditions for the existence of stationary optimal policies for an MDP with weakly continuous transition probabilities but the derivation is done directly; without deriving the optimality equations or inequalities. In this paper, we need not only the existence of optimal policies but also the validity of the ACOE. Hartley [24] established the validity of the ACOE for the class of problems whose dynamics are described by equations that include (1.1). However, he assumed that the demand is a uniformly bounded random variable, i.e. $|D_n| \leq C$ for a finite constant C. Thus, though several sufficient conditions for the validity of the ACOE have been described in the literature, none of them covers our problem. In Section 6 we provide the sufficient conditions that do this.

2 Problem Description

Recall that we consider a single commodity system with positive or negative demand. Assume that items that are returned are immediately available for resale. All lead times are assumed to be zero, the production/scrapping and borrowing/storage costs, are assumed to be linear and the unmet demand is backlogged. The cost of inventory held or backlogged (negative inventory) is modelled as a convex function. The sequence of events is as follows. At the beginning of the period, a decision-maker decides how much of the product to order or to scrap (at a loss) to meet demand. In addition, the decision-maker simultaneously decides how much inventory to borrow from or store to a (potentially external) secondary source. During the period, demand is realized and holding costs are accrued on the surplus or backlogged inventory. At the end of the period, borrowed or stored material is returned and the process continues. Note that since borrowed or stored inventory is returned the next period, holding costs are accrued but the inventory level the next period is not affected. The objective is to minimize the total expected discounted cost over a finite horizon or the discounted or average cost over an infinite horizon. Let

- $\beta \in (0, 1]$ be the discount factor,
- positive numbers c_+ and c_- be the per unit ordering and scrapping costs, respectively,
- positive numbers e_+ and e_- be the per unit borrowing and storage costs, respectively,
- $h(\cdot)$ denote the holding/backordering cost per period; convex, nonnegative function with finite values, and $h(x) \to \infty$ as $|x| \to \infty$,
- $\{D_n, n \geq 0\}$ be a sequence of i.i.d. random variables where D_n represents demand in the n$^{\text{th}}$ period. We assume that $\mathbb{E}\, h(y - D) < \infty$ for all $y \in \mathbb{R} = (-\infty, \infty)$, where D is a random variable with the same distribution as D_n. We notice, that this assumption and the assumed properties of the function h imply that $\mathbb{E}\,|D| < \infty$.

For $a \in \mathbb{R}$ let a^+ and a^- be the positive negative parts of a; $a^+ = \max\{a, 0\}$ and $a^- = \max\{-a, 0\}$. Let X_n be the inventory position at period n and let the ordered pair (Y_n, Z_n) be the amount ordered/scrapped and the amount borrowed/stored in this period. Denote the one-step cost function by $C(x, (a, b))$,

$$C(x, (a, b)) = c_+a^+ + c_-a^- + e_+b^+ + e_-b^- + \mathbb{E}\, h(x + a + b - D). \tag{2.1}$$

We model the decision scenario as a Markov decision process (see e.g. Bertsekas [6, 5], Dynkin and Yushkevich [14], Feinberg and Shwartz [17], Puterman [33], or Ross [36]). A general policy can be randomized and depend on the history of the inventory levels and decisions. A policy is called stationary if decisions are non-randomized and depend only on the current inventory level. That is, a stationary policy is defined by a measurable function that maps the inventory position (the state space, $\mathbb{X}$) to the set of potential actions (the action space). A Markov policy is defined as a sequence $d_0, d_1, \ldots$ where $d_n(x)$ is the decision that should be selected if the inventory level is x at step n. In our model, if an action (a, b) is chosen in state x, the cost $C(x, (a, b))$ is accrued, the system moves to state $x + a - D$ and this process continues. As was alluded to earlier, since the amount that is borrowed or stored is returned the next period, it has no effect on the subsequent inventory position. For a policy π and for an initial inventory level x, we define

$$v^{\pi}_{N,\beta}(x) = \mathbb{E}^{\pi}_{x} \sum_{n=0}^{N-1} [\beta^n C(X_n, d_n(X_n))], \tag{2.2}$$

$$v^{\pi}_{\beta} = \lim_{N\to\infty} v^{\pi}_{N,\beta}(x), \tag{2.3}$$

$$w^{\pi}(x) = \limsup_{N\to\infty} \frac{1}{N} v^{\pi}_{N,1}(x). \tag{2.4}$$

The equations (2.2), (2.3), (2.4) define the $N-$stage expected discounted cost, the infinite horizon expected discounted cost, and the long-run average expected cost, respectively. In the finite horizon problem only the portion of the policy required for the time horizon is used. In each case we define the optimal values

$$v_{N,\beta}(x) = \inf_{\pi\in\Pi} v^{\pi}_{N,\beta}(x), \tag{2.5}$$

$$v_{\beta}(x) = \inf_{\pi\in\Pi} v^{\pi}_{\beta}(x), \tag{2.6}$$

$$w(x) = \inf_{\pi\in\Pi} w^{\pi}(x), \tag{2.7}$$

where Π is the set of all policies. A policy ϕ is called *optimal* for the respective criterion if its value $v^{\phi}_{N,\beta}(x),\ v^{\phi}_{\beta}(x)$, or $w^{\phi}(x)$ corresponds to the value on the right hand side of (2.5), (2.6), or (2.7), respectively for all $x \in \mathbb{X}$.

We remark that our assumptions also imply that $v_\beta(x) < \infty$ for all $x \in \mathbb{R}$. Indeed, the assumptions on the holding cost h imply that there is a point x^* such that $h(x^*) = \min_{x\in\mathbb{R}} h(x) \geq 0$. Without loss of generality, we assume that $x^* = 0$ and $h(0) = 0$. Consider a policy ϕ that never borrows/stores and always orders/scraps in a way that the inventory level before the demand is known is 0. Then

$$v_\beta(x) \leq v^{\phi}_{\beta}(x) = c_- x^+ + c_+ x^- + \mathbb{E}\, h(-D) + \beta(c_+ \mathbb{E}\, D^+ + c_- \mathbb{E}\, D^- + \mathbb{E}\, h(-D))/(1-\beta) < \infty. \tag{2.8}$$

3 Finite Horizon Discounted Cost Optimal Policies

In this section we study the finite horizon problem. Since it is fixed throughout the section, we suppress β whenever possible. It is well-known that if a solution to the following finite horizon optimality equations (FHOE) exists, that solution is equal to v_n (componentwise) as defined in (2.5). Let $v_0 \equiv 0$ and for $n = 1, 2, \ldots$

$$v_n(y) = \inf_{-\infty<a<\infty,\ -\infty<b<\infty} \left\{ c_+ a^+ + c_- a^- + g(y+a, b) + \beta\, \mathbb{E}\, v_{n-1}(y+a-D) \right\}, \tag{3.1}$$

where

$$g(y, b) = e_+ b^+ + e_- b^- + \mathbb{E}\, h(y+b-D). \tag{3.2}$$

We observe that an equivalent system is

$$v_n(y) = \inf_{-\infty<a<\infty} \left\{ c_+ a^+ + c_- a^- + g^*(y+a) + \beta\, \mathbb{E}\, v_{n-1}(y+a-D) \right\}, \tag{3.3}$$

where

$$g^*(y) = \inf_{-\infty<b<\infty} \left\{ e_+ b^+ + e_- b^- + \mathbb{E}\, h(y+b-D) \right\}. \tag{3.4}$$

This observation leads to an algorithm for solving the proposed problem. First, solve (3.4) for g^*. Using this function, find the optimal ordering/scrapping policy by solving (3.3). Finally, using g^* evaluated at $y + a$ find the optimal borrowing/storage decision b. The following lemma provides preliminary results on v_n and g^*.

Lemma 3.1 *(i) The functions $g^*(y)$ and $v_n(y)$ are convex in y for all $n \geq 0$. (ii) $g^*(y) \to \infty$ and $v_n(y) \to \infty$ as $|y| \to \infty$ for all $n \geq 1$. (iii) The* $\inf$ *can be replaced with* $\min$ *in* (3.3) *and* (3.4).

Proof. For any convex function $k(y)$ and random variable X such that $\mathbb{E}\,|k(y-X)| < \infty$, $Ek(y-X)$ is also convex in y (cf. Bertsekas [6, Lemma 4.2.1]). Since h is convex, the function inside the infimum in (3.4) is jointly convex in y and b. Thus, applying Proposition B-4 of Heyman and Sobel [27] we have that $g^*(y)$ is convex in y. The fact that $g^*(y) \to \infty$ as $|y| \to \infty$ follows from the assumption that $h(y) \to \infty$ as $|y| \to \infty$.

We prove the remaining parts of (i) and (ii) by induction. By assumption v_0 is convex. Assume that convexity holds for $n - 1$. Since we have just shown that g^* is convex, the function inside the infimum in (3.3) is jointly convex in y and a. Again applying Proposition B-4 of Heyman and Sobel [27] yields that $v_n(y)$ is convex in y. Moreover, note that $v_n(y) \to \infty$ as $|y| \to \infty$ is implied by the fact that $g^*(y) \to \infty$ as $|y| \to \infty$ and the result is proven.

To verify (iii), we observe that for each y the function of b on the right-hand side of (3.4) is convex in $b \in \mathbb{R}$ and therefore continuous. In addition, this function tends to ∞ as $|b| \to \infty$. This is also true for (3.3) with the variable a instead of b. ∎

For a function $f(y)$, let $d^- f(y)/dy$ denote the left hand derivative of f. This derivative exists if f is convex. In light of the results of Lemma 3.1, the infimums in (3.3) and (3.4) may be replaced by minimums. The minimum may be computed by finding the place at which the left hand derivative changes sign. We, thus, may define the following quantities for $n = 1, 2, \ldots$

$$\underline{L}_n^p = \sup\left\{ z \in \mathbb{R} | \, c_+ + \frac{d^- g^*(z)}{dz} + \beta \frac{d^- [\mathbb{E}\, v_{n-1}(z-D)]}{dz} < 0 \right\}, \tag{3.5}$$

$$\overline{L}_n^p = \sup\left\{ z \in \mathbb{R} | \, c_+ + \frac{d^- g^*(z)}{dz} + \beta \frac{d^- [\mathbb{E}\, v_{n-1}(z-D)]}{dz} \leq 0 \right\}, \tag{3.6}$$

$$\overline{U}_n^p = \inf\left\{ z \in \mathbb{R} | \, -c_- + \frac{d^- g^*(z)}{dz} + \beta \frac{d^- [\mathbb{E}\, v_{n-1}(z-D)]}{dz} > 0 \right\}, \tag{3.7}$$

$$\underline{U}_n^p = \inf\left\{ z \in \mathbb{R} | \, -c_- + \frac{d^- g^*(z)}{dz} + \beta \frac{d^- [\mathbb{E}\, v_{n-1}(z-D)]}{dz} \geq 0 \right\}, \tag{3.8}$$

$$\underline{L}^b = \sup\left\{ z \in \mathbb{R} | \, e_+ + \frac{d^- [\mathbb{E}\, h(z-D)]}{dz} < 0 \right\}, \tag{3.9}$$

$$\overline{L}^b = \sup\left\{ z \in \mathbb{R} | \, e_+ + \frac{d^- [\mathbb{E}\, h(z-D)]}{dz} \leq 0 \right\}, \tag{3.10}$$

$$\overline{U}^b = \inf\left\{ z \in \mathbb{R} | \, -e_- + \frac{d^- [\mathbb{E}\, h(z-D)]}{dz} > 0 \right\}. \tag{3.11}$$

$$\underline{U}^b = \inf\left\{ z \in \mathbb{R} | \, -e_- + \frac{d^- [\mathbb{E}\, h(z-D)]}{dz} \geq 0 \right\}, \tag{3.12}$$

where the supremum (infimum) of the empty set is taken to be $-\infty$ (∞). Although the values defined in (3.5)-(3.8) depend on β, to keep the notation simple, we continue to suppress this dependence in the current section. For example, the full notation for $\underline{L}_n^p$ is $\underline{L}_{n,\beta}^p$.

We remark that equations (8-58a) and (8-58b) of Heyman and Sobel [27] are similar to our (3.5)-(3.12) but only one point where the appropriate convex functions achieve their minimums was considered in [27]. Here we consider the intervals of all possible solutions. Obviously, $\underline{L}_n^p \leq \overline{L}_n^p$, $\underline{U}_n^p \leq \overline{U}_n^p$, $\underline{L}^b \leq \overline{L}^b$, and $\underline{U}^b \leq \overline{U}^b$, where equalities are possible in each case. We say that L_n^p is a *lower actual inventory threshold* if

$$\underline{L}_n^p \leq L_n^p \leq \overline{L}_n^p, \tag{3.13}$$

U_n^p is an *upper actual inventory threshold* if

$$\underline{U}_n^p \leq U_n^p \leq \overline{U}_n^p, \tag{3.14}$$

L^b is a *lower total (actual plus borrowed) inventory threshold* if

$$\underline{L}^b \leq L^b \leq \overline{L}^b, \tag{3.15}$$

and U^b is an *upper total inventory threshold* if

$$\underline{U}^b \leq U^b \leq \overline{U}^b. \tag{3.16}$$

The following lemma clarifies these definitions.

Lemma 3.2 *For any* $n = 1, 2, \ldots$ *consider four threshold levels* L_n^p,, U_n^p, L^b, *and* U^b *satisfying* (3.13) – (3.16)*. Define the following decisions: order up/scrap down to the level*

$$t'_n = \begin{cases} L_n^p & \text{if } y < L_n^p, \\ y & \text{if } L_n^p \leq y \leq U_n^p, \\ U_n^p & \text{if } U_n^p < y, \end{cases} \tag{3.17}$$

where y *is the current inventory level, and borrow up/store down to the level*

$$z'_n = \begin{cases} L^b & \text{if } t'_n < L^b, \\ t'_n & \text{if } L^b \leq t'_n \leq U^b, \\ U^b & \text{if } U^b < t'_n. \end{cases} \tag{3.18}$$

Then the actions $a_n(y) = t'_n - y$ *and* $b_n(y) = z'_n - a_n(y)$ *minimize the right-hand sides of the optimality equations* (3.3) *and* (3.4)*, respectively. Therefore, for any* $N = 1, 2, \ldots,$ *the policy* $\phi = \{d_N, d_{N-1}, \ldots, d_1\}$ *is optimal for the* N*-step problem, where* $d_n(x) = (a_n(x), b_n(x))$*,* $n = 1, \ldots, N$*, and* $-\infty < x < \infty$*.*

Proof. In light of Lemma 3.1, the problem of finding optimal policies via (3.3) and (3.4) is simply the single period cash balance models discussed in [15] and [27, Section 8-4]. Thus, similar to [27, Section 8-4], we have optimal order up-to and down-to levels defined by the minimums of the convex functions defined in (3.3) and (3.4) when the changes of variables $z = y + a$ and $z = y + b$ are applied. For any $L_n^p \in [\underline{L}_n^p, \overline{L}_n^p]$ an optimal action is to order up to level L_n^p if the current inventory $y < L_n^p$. If the inventory level is above any level $U_n^p \in [\underline{U}_n^p, \overline{U}_n^p]$, an optimal action is to reduce the inventory level to U_n^p. If the inventory level is between $\underline{L}_n^p$ and $\overline{U}_n^p$, an optimal action is not to use the ordering/scrapping option. Similarly, the optimal decision obtained from (3.4) is to increase y to $L^b \in [\underline{L}^b, \overline{L}^b]$ if $y < L^b$ and to decrease y to $U^b \in [\underline{U}^b, \overline{U}^b]$ if $y > U^b$. No borrowing/storage action is required if $L^b < y < U^b$. ∎

The following lemma simplifies the structure of the optimal policy in the sense that it displays that we need not produce (scrap) and store (borrow) in the same period.

Lemma 3.3 $\overline{L}_n^p \le \underline{U}^b$ *and* $\overline{L}^b \le \underline{U}_n^p$, $n = 1, 2, \dots$.

Proof. We prove the first inequality. The proof of the second inequality is similar. Suppose $\overline{L}_n^p > \underline{U}^b$ and set $L_n^p = \overline{L}_n^p$ and $U^b = \underline{U}^b$. For $n = 1$ and $y < L_1^p$, (3.1) and Lemma 3.2 imply that for $\bar{a} = L_1^p - y > 0$ and $\bar{b} = L^b - L_1^p < 0$

$$\begin{aligned} v_1(y) &= \min_{-\infty < a < \infty, -\infty < b < \infty} \{c_+ a^+ + c_- a^- + e_+ b^+ + e_- b^- + \mathbb{E}\, h(y + a + b - D)\} \\ &= c_+ \bar{a}^+ + e_- \bar{b}^- + \mathbb{E}\, h(y + \bar{a}^+ - \bar{b}^- - D) \qquad (3.19) \\ &> c_+(\bar{a}^+ - \bar{b}^-) + \mathbb{E}\, h(y + (\bar{a}^+ - \bar{b}^-) - D), \end{aligned}$$

where the last expression corresponds to the policy that orders $(\bar{a}^+ - \bar{b}^-)$ units and borrows nothing. This violation of the optimality equation implies the contradiction. Thus, the case $n = 1$ is proven.

For $n \ge 2$ the inequality $L_n^p > U^b$ is not possible by similar arguments, but we must consider two-step decisions. Again, suppose this inequality holds. According to Lemma 3.2, when the initial state $y < L_n^p$ there is an optimal policy ϕ that prescribes to order up to the level L_n^p ($a = L_n^p - y$) and then to store down to the level U^b ($b = U^b - L_n^p < 0$). Consider now a policy ψ (when the initial state is y) that prescribes to order up to the level U^b at the first step and borrow nothing. At the second step it orders/scraps the amount that the policy ϕ would order/scrap at the second step if the inventory level were $L_n^p - D_1$ plus it orders $-b = L_n^p - U^b$ to make up the difference. It also borrows the same amount as the policy ϕ given that their actual inventory levels are the same. At the following steps, the policies coincide so that the inventory position seen by ψ coincides with ϕ; the processes couple.

Denote by $c(y, d_1(y))$ the ordering/scrapping costs for policy ϕ at the second step. The total ordering costs at the first two steps plus the borrowing/storage cost at stage 2 for policy ϕ are

$$C^{\phi}(y) = c_+ a^+ + e_- b^- + h(y + a^+ - b^- - D_1) + \beta c(y + a^+ - D_1, d_1(y + a^+ - D_1)). \qquad (3.20)$$

Similarly for the policy ψ, we have

$$C^{\psi}(y) \le c_+ a^+ + h(y + a^+ - b^- - D_1) + \beta c(y + a^+ - D_1, d_1(y + a^+ - D)). \qquad (3.21)$$

All other costs for these two policies coincide. Since $C^{\psi}(y) < C^{\phi}(y)$ (almost surely), we have $v_n^{\psi}(y) < v_n^{\psi}(y)$. Thus, ϕ is not an optimal policy. This contradiction completes the proof. ∎

For L_n^p, U_n^p, L^b, and U^b satisfying (3.13) – (3.16), define

$$L_n^b = \max\{L_n^p, L^b\} \qquad (3.22)$$

and

$$U_n^b = \min\{U_n^p, U^b\}. \qquad (3.23)$$

Lemma 3.3 implies

$$L_n^p \le L_n^b \le U_n^b \le U_n^p \qquad (3.24)$$

Combining Lemmas 3.2, 3.3, and (3.24) we arrive at the major result of this section.

Theorem 3.4 *Consider four threshold levels L_n^p, U_n^p, L^b, and U^b satisfying* (3.13) – (3.16) *and consider L_n^b, and U_n^b defined in* (3.22) *and* (3.23). *Moreover, for the current inventory level y, let t'_n be the order up/scrap down to action defined in* (3.17) *and z'_n be the borrow up/store down to action defined by*

$$z'_n = \begin{cases} L_n^b & \text{if } y < L_n^b, \\ t'_n & \text{if } L_n^b \le y \le U_n^b, \\ U_n^b & \text{if } U_n^b < y. \end{cases} \qquad (3.25)$$

Then (3.24) *holds and the actions $a_n(y) = t'_n - y$ and $b_n(y) = z'_n - a_n(y)$ minimize the right-hand sides of the optimality equations* (3.3) *and* (3.4), *respectively. Therefore, for any $N = 1, 2, \ldots,$ the policy $\phi = \{d_N, d_{N-1}, \ldots, d_1\}$ is optimal for the N-step problem, where $d_n(x) = (a_n(x), b_n(x))$, $n = 1, \ldots, N$, and $-\infty < x < \infty$.*

Theorem 3.4 implies the existence of an optimal policy that in each period either produces (scraps) up (down) to the level L_n^p (U_n^p) and then potentially borrows (scraps) up (down) to the level L^b (U^b). These ideas are illustrated in the following example.

Example 3.5 Consider a finite horizon discounted cost problem with the following parameters:

$$N = 10 \qquad \beta = .7$$

$$c_+ = .015 \qquad c_- = .010$$

$$e_+ = .028 \qquad e_- = .025$$

The holding cost is assumed quadratic and when $x > 0$ units are held in inventory the holding cost is $.002x^2$ and when $x \leq 0$ the backordering cost is $.008x^2$. The probability mass function $p(\cdot)$ of demand per period is

$$p(x) = \begin{cases} 5/12 & \text{for } x = 3, \\ 1/12 & \text{for } x = 4, \\ 1/4 & \text{for } x = -5 \text{ and for } x = -7. \end{cases}$$

After a finite state approximation, the optimal policy may be defined by t'_n in (3.17), where

Optimal Production/Scrapping Levels						
n	1	2	3	4	5	≥ 6
L_n^p	$-\infty$	$-\infty$	-6	-6	-6	-6
U_n^p	∞	∞	∞	12	8	7

and $L_n^b = 0, U_n^b = 2$ for all $n = 1, 2, \ldots$ in (3.25).

Aside from the observation that there exists several order-up-to or down-to levels, one should also note that the borrowing and storage options are used as secondary options to meet demand. Thus, the borrow-up-to level is higher than the order-up-to level and the store-down-to level is lower than the scrap-down-to level.

Example 3.5 indicates that it is possible that an optimal policy uses all four options: producing, scrapping, borrowing, and storing. The following two propositions give sufficient conditions when only ordering/scrapping options should be used and managers should not borrow or store. Proposition 3.6 states that borrowing and storage should not be used when they are relatively expensive. Proposition 3.7 indicates that borrowing and storage should not be used when the demand is either nonnegative or nonpositive.

Proposition 3.6 *Suppose the holding and backordering costs are linear,*

$$h(x) = \begin{cases} h_+ x & \textit{for } x \geq 0, \\ h_- x & \textit{for } x < 0. \end{cases} \tag{3.26}$$

If $e_+ > h_-$, then the optimal policy presented in Theorem 3.4 does not borrow. Similarly, if $h_+ < e_-$, this optimal policy does not store. In addition, if $e_+ = h_-$, then $\underline{L}^b = -\infty$ and the optimal policy defined in

Theorem 3.4 with $L^b = -\infty$ does not borrow. Similarly, if $e_- = h_+$, then $\overline{U}^b = \infty$ and the optimal policy defined in Theorem 3.4 with $U^b = \infty$ does not store.

The case when $e_+ = h_-$ ($e_- = h_+$) has the potential that one **could** borrow (store) and still be optimal, but that this option need not be exercised. This is a consequence of the possibility of multiple optimal borrowing (storage) levels.

Proof. Note that

$$\frac{d^-[\mathbb{E}\, h(z-D)]}{dz} = h_+ P(D \le z) - h_- P(D > z) \ge -h_-.$$

Thus, (3.9) implies that $\overline{L}^b = -\infty$ and yields the first result. Now note that $\frac{d^-[\mathbb{E}\, h(z-D)]}{dz} \le h_+$ so that the second result follows by assumption since $h_+ < e_-$ implies $\underline{U}^b = \infty$. The cases $e_+ = h_-$ and $h_+ = e_-$ follow from similar considerations. ∎

Theorem 3.4 implies that

$$g^*(y) = \begin{cases} e_+(L^b - y) + \mathbb{E}\, h(L^b - D) & \text{for } y < L^b, \\ \mathbb{E}\, h(y-D) & \text{for } L^b \le y \le U^b, \\ e_-(y - U^b) + \mathbb{E}\, h(U^b - D) & \text{for } U^b < y. \end{cases} \tag{3.27}$$

Similarly, for $n \ge 1$

$$v_n(y) = \begin{cases} c_+(L_n^p - y) + g^*(L_n^p) + \beta\, \mathbb{E}\, v_{n-1}(L_n^p - D) & \text{for } y < L_n^p, \\ g^*(y) + \beta\, \mathbb{E}\, v_{n-1}(y - D) & \text{for } L_n^p \le y \le U_n^p, \\ c_-(y - U_n^p) + g^*(U_n^p) + \beta\, \mathbb{E}\, v_{n-1}(U_n^p - D) & \text{for } U_n^p < y. \end{cases} \tag{3.28}$$

For standard inventory problems with nonnegative demands and zero set-up costs, so-called S-policies (also called "order-up-to" or "basestock" policies) are optimal: always order up to the level S when the inventory level is smaller than S. For finite-horizon problems these order up to levels may depend on the stage number. The following statement demonstrates that for inventory problems with nonnegative demands, borrowing should not be used unless borrowing costs per unit are less expensive than ordering costs. Unlike Proposition 3.6, we do not assume that the holding costs are linear.

Proposition 3.7 *Suppose $c_+ \le e_+$ and $P(D \ge 0) = 1$. Then $\underline{L}^b \le \underline{L}_n^p$ and by selecting $L^b = \underline{L}^b$ in Theorem 3.4 we have that the optimal policy defined by* (3.17) *and* (3.25) *never borrows.*

Proof. If $\underline{L}^b \le \underline{L}^p_n$ and $L^b = \underline{L}^b$ then $L^b_n = L^p_n$ and the policy defined by (3.17) and (3.25) never borrows. So, we need only prove that $\underline{L}^b \le \underline{L}^p_n$. According to (3.5), this inequality is equivalent to

$$c_+ + \frac{d^- g^*(z)}{dz} + \beta \frac{d^-[\mathbb{E}\, v_n(z-D)]}{dz} < 0 \tag{3.29}$$

for all $z < \underline{L}^b$. We fix $z < \underline{L}^b$. Note that from (3.27)

$$\frac{d^- g^*(z)}{dz} = -e_+ < 0,$$

and (3.29) holds for $n = 0$ as a non-strict inequality. We consider any $n = 1, 2, \ldots$ and make the induction assumption that the left hand side of (3.29) is nonpositive for $n - 1$. Differentiating (3.28) yields (recall $\underline{L}^b \le L^b_n \le U^p_n$)

$$\frac{d^- v_n(z)}{dz} = \begin{cases} -c_+ & \text{for } z < L^p_n, \\ -e_+ + \beta \frac{d^- \,\mathbb{E}\, v_{n-1}(z-D)}{dz} & \text{for } L^p_n \le z \le U^p_n, \end{cases} \tag{3.30}$$

Adding $c_+ + \frac{d^- g^*(z)}{dz} = c_+ - e_+$ to both sides of (3.30) and applying the inductive hypothesis (since $D \ge 0$ almost surely) yields that (3.29) holds and the result is proven. ■

Since the cases when the demand is nonnegative and nonpositive are symmetric, Proposition 3.7 implies the following corollary.

Corollary 3.8 *Suppose* $c_- \le e_-$ *and* $P(D \le 0) = 1$. *Then* $\overline{U}^b \ge \overline{U}^p_n$ *and by selecting* $U^b = \overline{U}^b$ *in Theorem 3.4 we have the optimal policy defined by* (3.17) *and by* (3.25) *never stores.*

Finally, we remark that the assumption that $v_0 = 0$ is simply for convenience; the results of this section hold when v_0 is an arbitrary nonnegative convex function.

4 Infinite Horizon Discounted Cost Optimal Policies

In order to obtain results analogous to Theorem 3.4 for the infinite horizon discounted problem, it is sufficient to justify taking limits as n approaches infinity on each side of the finite horizon optimality equations (3.3). This result is alluded to for the cash balance problem in [21] and [27], but apparently not shown.

Assume $\beta < 1$. Unlike the previous section, we do not suppress β. Since the optimality equations hold for problems with nonnegative costs (see e.g. Theorem 8.2 in [39]), we may write the discounted cost optimality equations (DCOE),

$$v_\beta(y) = \inf_{-\infty<a<\infty,\ -\infty<b<\infty} \left\{ c_+ a^+ + c_- a^- + g(y+a, b) + \beta[\mathbb{E}\, v_\beta(y+a-D)] \right\}, \tag{4.1}$$

where $g(y,b)$ is defined in (3.2). The system (4.1) is equivalent to

$$v_\beta(y) = \inf_{-\infty<a<\infty} \left\{ c_+a^+ + c_-a^- + g^*(y+a) + \beta\, \mathbb{E}\, v_\beta(y+a-D) \right\}, \tag{4.2}$$

where g^* is as defined in (3.4).

The model satisfies the following two conditions: all costs are non-negative and for all $y, \lambda \in \mathbb{R}$ and all $n = 1, 2 \ldots$ the sets

$$B_n(y,\lambda) = \{a \in \mathbb{R} |\; c_+a^+ + c_-a^- + g^*(y+a) + \beta[\mathbb{E}\, v_{n,\beta}(y+a-D) \le \lambda\} \tag{4.3}$$

are compact. Therefore, in view of [5, Proposition 1.7, p. 148] or [7, Proposition 9.17], $v_{n,\beta} \uparrow v_\beta$ as $n \to \infty$. Thus, v_β is a convex function and Lemma 3.1 holds for the objective function v_β and for each of the optimality equations (4.1) and (4.2). Similar to the finite horizon case, we can rewrite the optimality equation (4.2) in the form

$$v_\beta(y) = \min_{-\infty<a<\infty} \left\{ c_+a^+ + c_-a^- + g^*(y+a) + \beta[\mathbb{E}\, v_\beta(y+a-D)] \right\}. \tag{4.4}$$

We define the numbers $\underline{L}^p_\beta$, $\overline{L}^p_\beta$, $\underline{U}^p_\beta$, and $\overline{U}^p_\beta$ defined by formulas (3.5) – (3.8) with v_{n-1} replaced by v_β. As in the finite horizon case, note $\underline{L}^p_\beta \le \overline{L}^p_\beta$ and $\underline{U}^p_\beta \le \overline{U}^p_\beta$ and define lower and upper actual inventory thresholds L^p_β and L^p_β satisfying the inequalities

$$\underline{L}^p_\beta \le L^p_\beta \le \overline{L}^p_\beta, \tag{4.5}$$

$$\underline{U}^p_\beta \le U^p_\beta \le \overline{U}^p_\beta. \tag{4.6}$$

The following lemma is similar to Lemma 3.2 and has a virtually identical proof with the only difference being that (4.2) should be considered instead of (3.3).

Lemma 4.1 *Consider four threshold levels L^p_β, U^p_β, L^b, and U^b satisfying* (4.5), (4.6), (3.15), *and* (3.16), *respectively. The stationary policy that orders up/scraps down to the level*

$$t' = \begin{cases} L^p_\beta & \text{if } y < L^p_\beta, \\ y & \text{if } L^p_\beta \le y \le U^p_\beta, \\ U^p_\beta & \text{if } U^p_\beta < y, \end{cases} \tag{4.7}$$

where y is the current inventory level, and borrows up/stores down to the level

$$z' = \begin{cases} L^b & \text{if } t' < L^b, \\ t' & \text{if } L^b \le t' \le U^b, \\ U^b & \text{if } U^b < t', \end{cases} \tag{4.8}$$

is optimal.

Similar to Lemma 3.3 we have that

$$\overline{L}^p_\beta \leq \underline{U}^b, \tag{4.9}$$

$$\overline{L}^b \leq \underline{U}^p_\beta. \tag{4.10}$$

The proofs of these inequalities coincide with the proof of Lemma 3.3 for $n \geq 2$. We also define $\tilde{L}^b_\beta = \max\{L^p_\beta, L^b\}$ and $\tilde{U}^b_\beta = \min\{U^p_\beta, U^b\}$. Thus, similar to Theorem 3.4 we have the main result for infinite horizon discounted cost problems.

Theorem 4.2 *Consider four threshold levels L^p_β, U^p_β, L^b, and U^b satisfying* (4.5), (4.6), (3.15), *and* (3.16) *respectively. The stationary policy defined by the ordering/scrapping decision* (4.7) *for a current inventory level y and by the borrowing/storage decision to borrow up/store down to the level*

$$z' = \begin{cases} \tilde{L}^b_\beta & \text{if } y < \tilde{L}^b_\beta, \\ t' & \text{if } \tilde{L}^b_\beta \leq y \leq \tilde{U}^b_\beta, \\ \tilde{U}^b_\beta & \text{if } \tilde{U}^b_\beta < y, \end{cases} \tag{4.11}$$

is optimal.

Example 4.3 Consider the infinite horizon analogue of Example 3.5. An optimal policy is defined by $L^p_\beta = -6$, $\tilde{L}^b_\beta = 0$, $\tilde{U}^b_\beta = 2$, and $U^p_\beta = 7$. This is simply the four-threshold policy from Example 3.5 for $n \geq 6$.

For infinite-horizon problems, Propositions 3.6 and 3.7 hold for the stationary policy defined in Theorem 4.2. The proofs are virtually unchanged. However, since Theorem 4.2 describes a stationary policy, a stronger version of Proposition 3.7 holds.

Proposition 4.4 *Suppose $c_+ \leq e_+$ and $P(D \geq 0) = 1$. Then $\underline{L}^b \leq \underline{L}^p_\beta$ and by selecting $L^b = \underline{L}^b$ in Theorem 4.2, we have that the L^p_β-policy that prescribes to order at each step up to the level L^p_β, never borrows, never scraps, and never stores is optimal when the initial inventory level $y \leq L^p_\beta$.*

Similar to Corollary 3.8 we have the following statement.

Corollary 4.5 *Suppose $c_- \leq e_-$ and $P(D \leq 0) = 1$. Then $\overline{U}^b \geq \overline{U}^p_\beta$ and by selecting $U^b = \overline{U}^b$ in Theorem 4.2, we have that the policy that prescribes to scrap at each step down to the level L^p_β, never borrows, never orders, and never stores is optimal when the initial inventory level $y \geq U^p_\beta$.*

We remark that in addition to the convergence of the values $v_{n,\beta} \uparrow v_\beta$, the convergence of the optimal policies takes place. Indeed, if $\hat{L}^p_\beta$ and $\hat{U}^p_\beta$ are limit points of sequences of the optimal ordering/scrapping thresholds for the objective function $v_{n,\beta}$, then for any L^b and U^b satisfying (3.15) and (3.16), we have that the four-threshold policy with $L^p_\beta = \hat{L}^p_\beta$, $U^p_\beta = \hat{U}^p_\beta$, L^b, and U^b defined in Theorem 4.2 is optimal for the infinite horizon problem. This follows, for example, from the remark on p. 149 of [5].

5 Average Cost per Unit Time Optimal Policies

In this section we extend the previous results for the average cost case. We define the constant

$$w = \liminf_{\beta \uparrow 1} (1-\beta) m_\beta, \tag{5.1}$$

where $m_\beta = \inf_{x \in \mathbb{X}} v_\beta(x)$.

As is shown in the next section, for our problem $w < \infty$ and there exists a function $u(x)$ with nonnegative finite values such that

$$w + u(y) = \min_{-\infty < a < \infty,\ -\infty < b < \infty} \left\{ c_+ a^+ + c_- a^- + g(y+a, b) + \mathbb{E}\, u(y+a-D) \right\}, \tag{5.2}$$

where $g(y,b)$ is as defined in (3.2). In addition, there exists a sequence $\beta_n \uparrow 1$ such that $\lim_{\beta_n \uparrow 1} (1-\beta_n) m_{\beta_n} = w$ and $u(x) = \lim_{n\to\infty}\{v_{\beta_n}(x) - m_{\beta_n}\} \geq 0$ for all $x \in \mathbb{X}$. Thus, the function u is nonnegative, convex, and $u(x) \to \infty$ as $|x| \to \infty$.

Analogous to the discounted case, an equivalent system is

$$w + u(y) = \min_{-\infty < a < \infty} \left\{ c_+ a^+ + c_- a^- + g^*(y+a) + \mathbb{E}\, u(y+a-D) \right\}, \tag{5.3}$$

where g^* is defined as in (3.4).

Similar to (3.5) – (3.8), define

$$\underline{L}^p = \sup\left\{ z \in \mathbb{R} |\, c_+ + \frac{d^- g^*(z)}{dz} + \frac{d^-[\mathbb{E}\, u(z-D)]}{dz} < 0 \right\}, \tag{5.4}$$

$$\overline{L}^p = \sup\left\{ z \in \mathbb{R} |\, c_+ + \frac{d^- g^*(z)}{dz} + \frac{d^-[\mathbb{E}\, u(z-D)]}{dz} \leq 0 \right\}, \tag{5.5}$$

$$\underline{U}^p = \inf\left\{ z \in \mathbb{R} |\, -c_- + \frac{d^- g^*(z)}{dz} + \frac{d^-[\mathbb{E}\, u(z-D)]}{dz} > 0 \right\}. \tag{5.6}$$

$$\overline{U}^p = \inf\left\{z \in \mathbb{R} | -c_- + \frac{d^- g^*(z)}{dz} + \frac{d^-[\mathbb{E}\, u(z-D)]}{dz} \geq 0\right\}, \tag{5.7}$$

and consider $\underline{L}^b$, $\overline{L}^b$, $\underline{U}^b$, and $\overline{U}^b$ defined by (3.9), (3.10), (3.11), and (3.12), respectively. Consider L^b and U^b satisfying (3.15) and (3.16) respectively. As analogues to (3.13) and (3.14), consider L^p and U^p satisfying the inequalities

$$\underline{L}^p \leq L^p \leq \overline{L}^p, \tag{5.8}$$

$$\underline{U}^p \leq U^p \leq \overline{U}^p. \tag{5.9}$$

Lemma 5.1 *Consider four threshold levels L^p, U^p, L^b, and U^b satisfying* (5.8), (5.9), (3.15), *and* (3.16) *respectively. The stationary policy that prescribes to order up/scrap down to*

$$t' = \begin{cases} L^p & \text{if } y < L^p, \\ y & \text{if } L^p \leq y < U^p, \\ U^p & \text{if } U^p \leq y, \end{cases} \tag{5.10}$$

where y is the current inventory level, and to borrow up/store down to the level

$$z' = \begin{cases} L^b & \text{if } t' < L^b, \\ t' & \text{if } L^b \leq t' < U^b, \\ U^b & \text{if } U^b \leq t', \end{cases} \tag{5.11}$$

defines the values $a = t' - y$ and $b = z' - t'$ that minimize the right-hand sides of (5.3) *and* (3.4) *respectively. In addition, the infimum of the average costs per unit time, $w(y)$ defined in* (2.7)*, equals the constant w defined in* (5.1)*, and the thresholds t' and z' define a stationary optimal policy.*

Proof. Similar to Lemma 3.2, the convexity of u and h implies that the policy described minimizes the right-hand sides of (5.3) and (3.4). We recall that a policy for an MDP is called stationary if it is defined by a measurable mapping of the state space into an action space. The interpretation of this mapping is that, if the system is at some state, the value of this mapping is the selected action (independent of time). For our problem, a stationary policy ϕ is a measurable mapping from $\mathbb{R}$ to $\mathbb{R} \times \mathbb{R}$, where the first coordinate is how much to order/scrap and the second coordinate is how much to borrow/store. We remark that $t' = t'(y)$ in (5.10) and $z' = z'(t') = z'(t'(y))$ in (5.11) for an inventory level y. We consider the mapping $\phi(y) = (a(y), b(y)) = (t'(y), z'(t'(y)))$. Since the functions t' and z' have simple threshold forms, ϕ is measurable. Thus, we have

$$w + u(y) = c_+(a(y))^+ + c_-(a(y))^- + m(y + a(y), b(y)) + \mathbb{E}\, u(y + a(y) - D). \tag{5.12}$$

According to Schäl [37, Proposition 1.3], if the left-hand side of (5.12) is greater than or equal to the right-hand side, then $w^{\phi}(y) = w(y) = w$ for all $y \in \mathbb{R}$. ∎

Lemma 5.2 $\overline{L}^p \leq \underline{U}^b$ *and* $\overline{L}^b \leq \underline{U}^p$.

Proof. Let ϕ be the (stationary) policy defined by (5.10) and (5.11). Since ϕ defines actions that minimize the right-hand sides of (5.3) and (3.4), the policy ϕ is canonical; see Section 5.2 in [25]. That is, for any $n \geq 1$, ϕ minimizes the criterion $v^{\pi}_{n,1}(x) + \mathbb{E}^{\pi}_{x}\, u(X_n)$. The rest of the proof follows from the coupling arguments in the proof of Lemma 3.3 for $n \geq 2$. ∎

Corresponding to (3.22) and (3.23), define

$$\tilde{L}^b = \max\{L^p, L^b\}, \tag{5.13}$$

$$\tilde{U}^b = \min\{U^p, U^b\}. \tag{5.14}$$

Lemma 5.2 implies $L^p \leq \tilde{L}^b \leq \tilde{U}^b \leq U^p$ which together with Lemma 5.1 imply our main result for undiscounted costs.

Theorem 5.3 *For four thresholds L^p, U^p, L^b, and U^b satisfying* (5.8), (5.9), (3.15), *and* (3.16) *respectively, consider the thresholds $\tilde{L}^b$ and $\tilde{U}^b$ defined in* (5.13) *and* (5.14). *The stationary policy that prescribes for a current inventory level y to order up/scrap down to the level t' defined in* (5.10) *and to borrow up/store down to the level*

$$z' = \begin{cases} \tilde{L}^b & \text{if } y < \tilde{L}^b, \\ t' & \text{if } \tilde{L}^b \leq y \leq \tilde{U}^b, \text{ and} \\ \tilde{U}^b & \text{if } \tilde{U}^b < y, \end{cases} \tag{5.15}$$

where y is the current inventory level and $\tilde{L}^b$ and $\tilde{U}^b$ have been defined in (5.13) *and* (5.14)*, is optimal.*

Example 5.4 Suppose we use the same parameters as Example 3.5, but consider the average cost criterion. An optimal policy is defined by $L^p = -5$, $\tilde{L}^b = 0$, $\tilde{U}^b = 2$, and $U^p = 5$.

The following proposition is similar to Proposition 3.6. Its proof is identical to that of Proposition 3.6 with the major difference being that (5.3) should be considered instead of (3.3).

Proposition 5.5 *Suppose the holding costs are linear, i.e.* (3.26) *holds. If $e_+ > h_-$, then the optimal policy presented in Theorem 5.3 does not borrow. Similarly, if $h_+ < e_-$, this policy does not store. In addition, if $e_+ = h_-$, then $\underline{L}^b = -\infty$ and the optimal policy defined in Theorem 5.3 with $L^b = -\infty$ does not borrow. Similarly, if $e_- = h_+$, then $\overline{U}^b = \infty$ and the optimal policy defined in Theorem 5.3 with $U^b = \infty$ does not store.*

Analogous to Proposition 4.4 and Corollary 4.5, we have the following two results.

Proposition 5.6 *Suppose $c_+ \le e_+$ and $P(D \ge 0) = 1$. Then $\underline{L}^b \le \underline{L}^p$ and by selecting $L^b = \underline{L}^b$ in Theorem 5.3, we have that the L^p-policy that prescribes to order at each step up to the level L^p, never borrows, never scraps, and never stores, is optimal when the initial inventory level $y \le L^p$.*

Corollary 5.7 *Suppose $c_- \le e_-$ and $P(D \le 0) = 1$. Then $\overline{U}^b \ge \overline{U}^p$ and by selecting $U^b = \overline{U}^b$ in Theorem 5.3, we have that the policy that prescribes to scrap at each step down to the level U^p, never borrows, never orders, and never stores, is optimal when the initial inventory level $y \ge U^p$.*

In the end of Section 4 we established that discounted cost finite-horizon optimal thresholds converge to discounted cost infinite-horizon optimal thresholds. Similarly, discounted cost infinite-horizon optimal thresholds converge in some sense to optimal thresholds for the average cost criterion. Indeed, for ordering and scrapping decisions, let $\hat{L}^p$ and $\hat{U}^p$ be limit points of a sequence of optimal ordering/scrapping thresholds $L^p_{\beta_n}$ and $U^p_{\beta_n}$ as $n \to \infty$, where the sequence $\{\beta_n, n \ge 0\}$ is chosen as discussed in the text following (5.2). Appealing to the results in the the next section implies that $\hat{L}^p$ and $\hat{U}^p$ are respectively optimal ordering and scrapping thresholds for the average cost per unit time criterion.

6 The Average Cost Optimality Equations

In this section we prove the existence of a solution to the ACOE; (5.2), (3.2). Instead of restricting attention solely to the problem at hand, we provide sufficient conditions for the existence of a solution to the ACOE for a more general MDP, and then show that these conditions are satisfied for our problem.

Consider an MDP with a standard Borel state space $\mathbb{X}$. Suppose ρ is a metric on $\mathbb{X}$ such that $\rho(x, y) < \infty$ for all $x, y \in \mathbb{X}$ and $(\mathbb{X}, \rho)$ is complete and separable. Assume that the one-step costs $c(x, a)$ are non-negative and that the MDP satisfies the standard Borel measurability conditions; see e.g. [37]. The optimal values in the infinite-horizon discounted cost case, $v_\beta(x)$, are defined for discount factors $\beta \in [0, 1)$. Let $A(x)$ be the set of available actions at state x and q be the transition probability. Consider the following set of assumptions

Assumptions C:

1. There exists a non-decreasing, continuous function η on $R_+ = [0, \infty)$ such that $\eta(r) < \infty$ for all $r \in \mathbb{R}^+$,

$$|v_\beta(x) - v_\beta(y)| \leq \eta(\rho(x, y))$$

   for all $x, y \in \mathbb{X}$, and

$$\int \eta(\rho(x, y))q(dy|x, a) < \infty \tag{6.1}$$

   for all $x, y \in \mathbb{X}$ and $a \in A(x)$.

2. There exists a finite constant w such that

$$\liminf_{\beta \uparrow 1}(1 - \beta)m_\beta = w,$$

   where $m_\beta = \inf_{x \in \mathbb{X}} v_\beta(x)$.

3. For $N \in \mathbb{R}$ there exists a compact subset $K_N \subseteq \mathbb{X}$ such that $c(x) \geq N$ for all $x \in \mathbb{X} \setminus K_N$ where $c(x) = \inf_{a \in A(x)} c(x, a)$.

Note that **C1** implies that $v_\beta(x)$ is a continuous function in x for any $\beta \in [0, 1)$. Since $v_\beta(x) \geq c(x)$, there exist x_β such that $v_\beta(x_\beta) = m_\beta$. Let $M(\beta) = \{x \in X | v_\beta(x) = m_\beta\}$. In particular, $M(\beta) \subseteq K_N$ for $N > m_\beta$. The following lemma is similar to Lemma 4.6 in [37] and to Lemma 4 in [8].

Lemma 6.1 *Let assumptions **C2** and **C3** hold. Suppose $\{\beta(n), n \geq 1\}$ is a subsequence such that $\beta(n) \uparrow 1$ and $w = \lim_{n \to \infty}(1 - \beta(n))m_{\beta(n)}$. Then there exists a compact subset $K \subseteq \mathbb{X}$ and a finite integer ℓ such that $M(\beta(n)) \subseteq K$ for $n \geq \ell$.*

Proof. Consider $N > w + 1$. Then $N > \lim_{n \to \infty}(1 - \beta(n))m_{\beta(n)} + 1$ and there exists an integer ℓ such that for $n \geq \ell$

$$\frac{N}{1 - \beta(n)} > m_{\beta(n)} + \frac{1}{1 - \beta(n)}. \tag{6.2}$$

Set $K = K_N$ as defined in **C3**. We prove by contradiction that K is the set whose existence is postulated by the lemma. Let β satisfy (6.2) and suppose that this is not the case. Then there exists $x \in (\mathbb{X} \setminus K) \cap M(\beta)$. That is, $c(x) \geq N$ and $v_\beta(x) = m_\beta$. Let T be the first hitting time of K; $T = \inf\{n \geq 0 | X_n \in K\}$. Then for any stationary policy π

$$v_\beta^\pi(x) = E_x^\pi \sum_{n=0}^{\infty} c(X_n, d(X_n)) \geq E_x^\pi \sum_{n=0}^{T-1} \beta^n c(X_n, d(X_n)) + E_x^\pi \beta^T v_\beta(X_T)$$

$$\geq E_x^\pi \frac{1-\beta^T}{1-\beta} N + E_x^\pi \beta^T m_\beta > E_x^\pi [(1-\beta^T)(m_\beta + (1-\beta)^{-1}) + \beta^T m_\beta] \geq m_\beta + 1,$$

where the strict inequality follows from (6.2) and the last inequality follows from $T \geq 1$ when $x \notin K$. Since π is arbitrary, $v_\beta(x) > m_\beta$ and the inclusion $x \in M(\beta)$ is not possible. ■

We recall that a real-valued function f defined on a metric space Y is called *inf-compact* if the set $\{y \in \mathbb{Y} \mid f(y) \leq \lambda\}$ is compact for any $\lambda \in \mathbb{R}$. Note that since compact sets are closed, inf-compactness implies lower-semicontinuity. Consider the following two additional conditions:

Assumptions C:(continued)

4. For each fixed $x \in \mathbb{X}$ the function $c(x, a)$ is inf-compact in a.
5. Transition probabilities $q(\cdot|x, a)$ are weakly continuous in a for each $x \in \mathbb{X}$.

Theorem 6.2 *Suppose **C** hold. Then*

1. *there exists a continuous nonnegative function u on $\mathbb{X}$ such that for all $x \in \mathbb{X}$*

$$w + u(x) = \min_{a \in A(x)} \{c(x, a) + \int u(y) q(dy|x, a)\}; \tag{6.3}$$

2. *there exists a sequence $\beta(n) \to 1$ such $u(x) = \lim_{n\to\infty} u_{\beta(n)}(x)$, where $u_\beta(x) = v_\beta(x) - m_\beta$, $x \in \mathbb{X}$;*
3. *if $\mathbb{X}$ is a linear space and the functions v_β are convex for all $\beta \in [0, 1)$ then u is convex;*
4. *for fixed $x \in \mathbb{X}$, let a^* be a limit point of a sequence $\{a_n, n \geq 1\}$ where $v_{\beta(n)}(x) = c(x, a_n) + \beta(n) \int v_{\beta(n)}(y) q(dy|x, a_n)$ (the point a^* exists in view of Lemma 6.1). Then*

$$w + u(x) = c(x, a^*) + \int u(y) q(dy|x, a^*).$$

Proof. Consider the *discount cost optimality equations* (DCOE),

$$v_\beta(x) = \inf_{a \in A(x)} \{c(x, a) + \beta \int v_\beta(y) q(dy|x, a)\}, \text{ for all } x \in \mathbb{X}. \tag{6.4}$$

In particular, the minimum is achieved in the right-hand side of (6.4) since the expression minimized is lower semicontinuous (the sum of two lower semicontinuous functions). The first function is lower semicontinuous because of **C4** and the second function is lower semicontinuous because of **C1** and **C5**. The former implies that $v_\beta(x)$ is continuous. In addition, $v_\beta(x) \geq 0$.

A little algebra in (6.4) reveals

$$u_\beta(x) + (1-\beta)m_\beta = \min_{a\in A(x)}\{c(x,a) + \beta\int u_\beta(y)q(dy|x,a)\}. \tag{6.5}$$

Utilizing **C2** there exists a sequence $\beta(n)\uparrow 1$ as $n\to\infty$ such that $\lim_{n\to\infty}\beta(n)m_{\beta(n)} = w$. Lemma 6.1 implies that we can select this sequence such that $M(\beta(n))\subseteq K$, where K is a compact subset of $\mathbb{X}$. Let $\mathrm{diam}(K)$ denote the maximal distance between two points in K; the diameter of K. Since K is compact, $\mathrm{diam}(K)$ exists and is finite. We fix any point $z_0\in K$. Then according to **C1**

$$u_{\beta(n)}(x) \le \eta(\rho(x,z)) \le \eta(\rho(x,z_0) + \mathrm{diam}(K)), \tag{6.6}$$

where $z\in K$ is such that $v_{\beta(n)}(z) = m_{\beta(n)}$.

Since $|u_\beta(x) - u_\beta(y)| \le \epsilon$ when $\eta(\rho(x,y)) < \epsilon$, the family of functions $u_{\beta(n)}$ is equicontinuous. By the Ascoli theorem, [25], page 96, there exists a subsequence $\{\beta(n_k), k\ge 1\}$ of the sequence $\{\beta(n), n\ge 1\}$ such that $u_{\beta(n_k)}$ converges pointwise to a continuous function u and this convergence is uniform on each compact subset of $\mathbb{X}$. In particular, for each $x\in\mathbb{X}$

$$u(x) = \liminf_{k\to\infty,\ y\to x} u_{\beta(n_k)}(y). \tag{6.7}$$

Fix $x\in\mathbb{X}$. We first show that

$$w + u(x) \ge \min_{a\in A(x)}\{c(x,a) + \int u(y)q(dy|x,a)\}, \tag{6.8}$$

where the minimum in (6.8) is attained by the same reasons as in (6.3). For each $\beta(n_k)$ consider $a(n_k)$ such that

$$(1-\beta(n_k))m_{\beta(n_k)} + u_{\beta(n_k)}(x) = c(x,a(n_k)) + \int u_{\beta(n_k)}(y)q(dy|x,a(n_k)). \tag{6.9}$$

Without loss of generality, select the sequence n_k such that $|(1-\beta(n_k))m_{\beta(n_k)} + u_{\beta(n_k)}(x) - w - u(x)| \le 1$ for all n_k. **C**4 implies that all $a(n_k)$ belong to the compact set $K_N(x)$, where $N = w + u(x) + 1$. Thus, there exists $a^*\in A(x)$ and a subsequence $\{m_k\}$ of $\{n_k\}$ such that $a(m_k)\to a^*$. The result obtained by Serfozo's [38] extension of Fatou's lemma (see also Lemma 2.3 in [37]) implies that

$$\int u(y)q(dy|x,a^*) \le \liminf_{m_k\to\infty}\int u_{\beta(m_k)}(y)q(dy|x,a_{m_k}). \tag{6.10}$$

Since c is lower semicontinuous,

$$
\begin{aligned}
w + u(x) &= \lim_{m_k \to \infty} \{c(x, a(m_k)) + \int u_{\beta(m_k)}(y) q(dy|x, a(m_k))\} \\
&\geq c(x, a^*) + \int u(y) q(dy|x, a^*) \\
&\geq \min_{a \in A(x)} \{c(x, a) + \int u(y) q(dy|x, a)\}.
\end{aligned}
$$

Thus, (6.8) is proved.

From (6.5) we have that for any $a \in A(x)$

$$
(1-\beta)m_\beta + u_\beta(x) \leq c(x, a) + \int u_\beta(y) q(dy|x, a). \tag{6.11}
$$

Consider (6.11) for $\beta = \beta(n)$ defined above and apply sequentially the Ascoli theorem, [25], page 96, and Lebesgue's dominated convergence theorem. The latter is applicable in view of **C1** and (6.6). Indeed (6.6) yields

$$
0 \leq u_\beta(y) = v_\beta(y) - v_\beta(x) + u_\beta(x) \leq \eta(\rho(x, y)) + \eta(\rho(x, z_0) + \mathrm{diam}(K)) \tag{6.12}
$$

and (6.1) implies that

$$
\int (\eta(\rho(x, y)) + \eta(\rho(x, z_0) + \mathrm{diam}(K))) q(dy|x, a) \leq \int \eta(\rho(x, y)) q(dy|x, a) + \eta(\rho(x, z_0) + \mathrm{diam}(K)) < \infty. \tag{6.13}
$$

We, thus, have for any $a \in A(x)$

$$
w + u(x) \leq c(x, a) + \int u(y) q(dy|x, a),
$$

which is equivalent to

$$
w + u(x) \leq \min_{a \in A(x)} \{c(x, a) + \int u(y) q(dy|x, a)\}.
$$

■

The next result shows states the ACOE, (5.2) and (3.2), hold for the inventory problem considered. We remark that Theorem 6.2 also implies the validity of the ACOE for problems without borrowing and storage (let $m(y) = \mathbb{E}\, h(y - D)$ in (5.2)).

Proposition 6.3 *In the inventory problem considered, there exists a solution to the ACOE,* (5.2) *and* (3.2), $(w, u(y))$, *such that the relative value function* $u(y)$ *is nonnegative, convex, and is equal to the limit of functions* $u_{\beta(n)}$ *described in Theorem 6.2.*

Proof. We verify that Assumptions **C** hold so that Theorem 6.2 applies. First, in our case $\mathbb{X} = A = \mathbb{R}$, $\rho(x, y) = |x - y|$, and $\eta(r) = r \max\{c_+, c_-\}$. Consider two inventory levels x and z. In state x suppose the manager orders or scraps inventory to bring the level to z, then implements an optimal policy thereafter. Since this policy may not be optimal $v_\beta(x) \leq v_\beta(z) + \max\{c_+, c_-\}|z - x|$. In other words,

$$v_\beta(x) - v_\beta(z) \leq \max\{c_+, c_-\}|z - x|.$$

Since $\mathbb{E}\, h(x - D) < \infty$ for any $x \in \mathbb{R}$, h is convex and $h(x) \to \infty$ as $x \to \infty$, we have $\mathbb{E}\, |D| < \infty$ and **C1** is verified.

Fix $x \in \mathbb{X}$. Consider the policy ϕ that always orders up to level x. The renewal reward theorem implies that $w^\phi(x) = \lim_{n\to\infty} \frac{1}{n} v^\phi_{n,1}(x) < \infty$. Thus $\liminf_{\beta\uparrow 1}(1 - \beta)m_\beta \leq \liminf_{\beta\uparrow 1}(1 - \beta)v_\beta(x) \leq \liminf_{\beta\uparrow 1}(1 - \beta)v^\phi_\beta(x) = w^\phi(x) < \infty$ and **C2** holds.

To verify **C3**, we must prove that $c(x) \to \infty$ as $|x| \to \infty$. Let $x \to \infty$ and let $d = \min\{c_+, c_-, e_+, e_-\}$. Note that $d > 0$ and for $x > \mathbb{E}\, D$

$$c(x) = \inf_{-\infty<a,b<\infty} \{c_+a^+ + c_-a^- + e_+b^+ + e_-b^- + \mathbb{E}\, h(c+a+b-D)\} \geq \inf_{-\infty<y<\infty} \{d|y| + h(x+y-\mathbb{E}\, D)\}$$

$$\geq \min\{\frac{d(x - \mathbb{E}\, D)}{2}, h\left(\frac{x - \mathbb{E}\, D}{2}\right)\} \to \infty \quad \text{as } x \to \infty,$$

where the first inequality follows by setting $y = a+b$ and applying Jensen's inequality. To verify the second inequality, consider the two possibilities: (a) $y \leq -(x - \mathbb{E}\, D)/2$ and (b) $y \geq -(x - \mathbb{E}\, D)/2$. The case $x \to -\infty$ is similar.

Assumption **C4** holds since for $\lambda \in \mathbb{R}$, $\{(a, b)|C(x, (a, b)) \leq \lambda\}$ is closed and bounded. The fact that it is bounded follows from the fact that $C(x, (a, b)) \to \infty$ as either a or $b \to \infty$. It is closed since the cost function is convex and therefore continuous. Assumption **C5** is equivalent to $\mathbb{E}\, f(D+a_n) \to \mathbb{E}\, f(D+a)$ as $a_n \to a$, for any bounded, continuous f, which follows from Lebesgue's dominated convergence theorem.
∎

7 Conclusions

We have studied an extension of the classic inventory control/cash management models to include one period borrowing and storage. Instead of an optimal policy requiring two thresholds as has been shown for the cash management problem, we have four thresholds. We expect that in the discounted models (both finite and infinite horizon), a fixed cost could be added for either ordering or scrapping (but not both) and our results could be extended without difficulties to analogous results to those shown in [16, 21]. This follows from the fact that the (convex) holding cost in each model would be replaced with the function g^*. On the other hand, we do not believe that allowing for fixed costs to be associated with borrowing/storage and ordering/scrapping would lead to such simple policies. While the borrowing and storage policy would most likely be analogous to (s, S)-policies, it is not immediately clear that the $K-$convexity would carry through. This is left as a potential future research direction. Other potential research directions are to study problems with lost sales, lead times and problems with the borrowing/storage time intervals longer than one period.

8 Acknowledgements

The authors thank Emmanuel Fernandez for the comments regarding the existence of stationary optimal policies for the average cost criterion. We would also like to thank the anonymous referees for several comments that improved the readability of this paper. Research of the first author was partially supported by NSF grant DMI-0300121.

References

[1] Y. Aneja and A. H. Noori. The optimality of (s, s) policies for a stochastic inventory problem with proportional and lump-sum penalty cost. *Management Science*, 33(6):750–755, June 1987.

[2] K. J. Arrow, T. Harris, and J. Marshcak. Optimal inventory policy. *Econometrica*, 53(6):250–272, 1951.

[3] H. Arslan, H. Ayhan, and T. L. Olsen. Analytic models for when and how to expedite in make-to-order systems. *IIE Transactions*, 2003. to appear.

[4] E. Barankin. A delivery-lag inventory model with an emergency provision. *Naval Research Logistics Quarterly*, 8:285–311, 1961.

[5] D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 2. Athena Scientific, Belmont, MA, 1995.

[6] D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific, Belmont, MA, second edition, 2000.

[7] D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control: The Discrete-Time Case*. Athena Scientific, Belmont, MA, 1996.

[8] R. Cavazos-Cadena and L. I. Sennott. Comparing recent assumptions for the existence of average optimal stationary policies. *Operations Research Letters*, 11:33–37, 1992.

[9] X. Chen and D. Simchi-Levi. A new approach for the stochastic cash balance problem with fixed costs. Preprint, 2004.

[10] C. Chiang and G. J. Gutierrez. A periodic review inventory system with two supply modes. *European Journal of Operational Research*, 94(3):527–547, November 1996.

[11] C. Chiang and G. J. Gutierrez. Optimal control policies for a periodic review inventory system with emergency orders. *Naval Research Logistics Quarterly*, 45:187–204, 1998.

[12] G. M. Constantinides and S. F. Richard. Existence of optimal simple policies for discounted cost inventory and cash management in continuous time. *Operations Research*, 26(4):620–636, July-August 1978.

[13] K. Daniel. A delivery-lag inventory model with emergency. In *Multistage Inventory Models and Techniques*, pages 32–46. Stanford University Press, Stanford, CA, 1963.

[14] E. Dynkin and Y. A.A. *Controlled Markov Processes*. Springer-Verlag, New York, 1979.

[15] E. J. Elton and M. J. Gruber. On the cash balance problem. *Operational Research Quarterly*, 25(4):553–572, December 1974.

[16] G. D. Eppen and E. F. Fama. Cash balance and simple dynamic portfolio problems with proportional costs. *International Economic Review*, 10(2):119–133, June 1969.

[17] E. A. Feinberg and A. Shwartz, editors. *Handbook of Markov Decision Processes: Methods and Applications*. Kluwer, Boston, 2002.

[18] E. Fernández-Gaucherand. A note on the Ross-Taylor theorem. *Applied Mathematics and Computation*, 64(2-3):207–212, September 1994.

[19] E. Fernández-Gaucherand, A. Arapostathis, and S. I. Marcus. Convex stochastic control problems. In *Proceedings of the 31st Conference on Decision and Control*, pages 2179–2180. IEEE, December 1992.

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

[21] N. M. Girgis. Optimal cash balance levels. *Management Science*, 15(3):130–140, November 1968.

[22] L. Gubenko and E. Statland. On controlled discrete-time Markov decision processes. *Theory Probability and Mathematical Statistics*, (7):47–61, November 1975.

[23] T. Harris. How many parts to make at once. *Factory, The Magazine of Management*, 10:135–136,152, 1913.

[24] R. Hartley. Dynamic programming and an undiscounted, infinite horizon, convex stochastic control problem. In R. Hartley, L. Thomas, and D. White, editors, *Recent Developments in Markov Decision Processes*, pages 277–300. Academic Press, London, 1980.

[25] O. Hernández-Lerma and J. B. Lasserre. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. Springer, New York, NY, 1996.

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

[27] D. P. Heyman and M. J. Sobel. *Stochastic Models in Operations Research*, volume II. McGraw-Hill, New York, NY, 1984.

[28] K. Hinderer and K.-H. Waldmann. Cash management in a randomly varying environment. *European Journal of Operational Research*, 130:468–485, 2001.

[29] E. L. Huggins and T. L. Olsen. Inventory control with overtime and premium freight. Preprint, 2003.

[30] E. H. Neave. The stochastic cash balance problem with fixed costs for increases and decreases. *Management Science*, 16(7):472–490, March 1970.

[31] M. Neuts. An inventory model with optional lag time. *SIAM Journal of Applied Mathematics*, 12:179–185, 1964.

[32] E. Porteus. Stochastic inventory theory. In D. Heyman and M. Sobel, editors, *Handbooks in Operations Research and Management Science*, volume 2, chapter 12, pages 605–652. Elsevier Science Publishers, Amsterdam, 1990.

[33] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley Series in Probability and Mathematical Statistics. John Wiley and Sons, Inc., New York, 1994.

[34] R. Ritt and L. Sennott. Optimal stationary policies in general state Markov decision chains with finite action sets. *Mathematics of Operations Research*, 17:901–909, 1992.

[35] S. Ross. Arbitrary state Markovian decision processes. *The Annals of Mathematical Statistics*, 39:2118–2122, 1968.

[36] S. Ross. *Introduction to Stochastic Dynamic Programming*. Academic Press, New York, 1983.

[37] M. Schäl. Average optimality in dynamic programming with general state space. *Mathematics of Operations Research*, 18:163–172, 1993.

[38] R. Serfozo. Convergence of Lebesgue integrals with varying measures. *Sankhya Ser. A*, 44:380–402, 1982.

[39] R. E. Strauch. Negative dynamic programming. *Annals of Mathematical Statistics*, 37:871–890, 1966.

[40] G. Tagaras and D. Vlachos. A periodic review inventory system with emergency replenishments. *Management Science*, 47(3):415–429, March 2001.

[41] E. van der Laan and M. Salomon. Production planning and inventory control with remanufacturing and disposal. *European Journal of Operational Research*, 102:264–278, 1997.