

Optimality Inequalities for Average Cost Markov Decision Processes and the Stochastic Cash Balance Problem

Eugene A. Feinberg

Department of Applied Mathematics & Statistics, State University of New York at Stony Brook, Stony Brook, NY 11794-3600

email: efeinberg@notes.cc.sunysb.edu <http://www.ams.sunysb.edu/~feinberg/>

Mark E. Lewis

School of Operations Research & Industrial Engineering, Cornell University, 226 Rhodes Hall, Ithaca, NY 14853

email: mark.lewis@cornell.edu

<http://www.orie.cornell.edu/orie/people/faculty/profile.cfm?netid=mel47>

For general state and action space Markov decision processes, we present sufficient conditions for the existence of solutions of the average cost optimality inequalities. These conditions also imply the convergence of both the optimal discounted cost value function and policies to the corresponding objects for the average costs per unit time case. Inventory models are natural applications of our results. We describe structural properties of average cost optimal policies for the cash balance problem; an inventory control problem where the demand may be negative and the decision-maker can produce or scrap inventory. We also show the convergence of optimal thresholds in the finite horizon case to those under the expected discounted cost criterion and those under the expected discounted costs to those under the average costs per unit time criterion. 1. Markov decision process, average cost per unit time, optimality inequality, optimal policy, inventory control

2. MSC Subject Classification, iPrimary: 90C40 (Markov and semi-Markov decision processes); Secondary: 90B05 (Inventory, storage, reservoirs)

3. OR/MS classification, Primary: Dynamic Programming/optimal control/Markov/Infinite state; Secondary: Inventory/production/Uncertainty/Stochastic).

Key words: Markov decision process, average cost per unit time, optimality inequality, optimal policy, inventory control

MSC2000 Subject Classification: Primary: 90C40 (Markov and semi-Markov decision processes); Secondary: 90B05 (Inventory, storage, reservoirs)

OR/MS subject classification: Primary: Dynamic Programming/optimal control/Markov/Infinite state; Secondary: Inventory/production/Uncertainty/Stochastic

History: Received: XXXX xx, xxxx; Revised: Yyyyyy yy, yyyy and Zzzzzz zz, zzzz.

1. Introduction In a discrete-time Markov decision process (MDP) the usual method to study the average cost criterion is to find a solution to the average cost optimality equations. A policy that achieves the minimum in this system of equations is then average cost optimal. When the state and action spaces are infinite, one may be required to replace the equations with inequalities, yet the conclusions are the same; a policy that achieves the minimum in the inequalities is average cost optimal. Schäl [27] provides two groups of general conditions that imply the existence of a solution to the average cost optimality *inequalities* (ACOI). The first group, referred to as Assumptions (**W**) in Schäl [27], require weak continuity of the transition probabilities. The second group, Assumptions (**S**), require setwise continuity of the transition probabilities. In either case, for each state a compact action set was assumed in [27]. The purpose of this paper is to adapt Schäl's [27] conditions to problems with noncompact action sets; in particular to those related to inventory control. As was noted in [12], typical inventory control models (with general demand distributions) require weak continuity; setwise continuity is not enough to yield the conclusions that the ACOI have a solution. On the other hand, when the demand distribution is restricted to be continuous or the inventory is restricted to be integer, we show that setwise continuity does suffice.

The books by Sennott [28] and Hernández-Lerma and Lasserre [19] deal with countable and general

state MDPs, respectively. Hernández-Lerma and Lasserre [19, Chapter 5], Hernández-Lerma [18], and Fernández-Gaucherand [14] present results for non-compact action sets but assume setwise continuity. The results of Hernández-Lerma [18] extend Schäl’s [27] results on MDP’s with setwise continuous transition probabilities from compact action sets to noncompact action sets.

In this paper we study MDPs with weakly continuous transition probabilities. The major motivation for this study is their relevance to inventory control problems. Section 5.7 in Hernández-Lerma and Lasserre [19] 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. We are interested not only in the existence of optimal policies but in the validity of the optimality inequalities. This is an important step since these inequalities can be used to prove structural properties of optimal stationary policies and to prove convergence of discounted cost optimal policies to average cost optimal policies. We recall that, according to the example constructed by Cavazos-Cadena [4], optimality inequalities may hold for an MDP for which optimality equalities do not hold. In addition, optimality inequalities imply the existence of optimal policies [27, Proposition 1.3].

In this paper we consider a class of problems with noncompact action sets and introduce an additional condition, Assumption (LB), that states that a certain function, relevant to the relative value functions, is locally bounded from above. This assumption is important for the following reasons: (i) if it is satisfied, noncompact action sets can be reduced to compact action subsets in a way that the value functions remain unchanged and Schäl’s assumptions [27] hold, (ii) it can be verified easily, and (iii) it typically holds for inventory control problems. Thus, this paper provides straightforward tools to analyze inventory control problems with the average costs per unit time criterion and without the assumption that the demand is either discrete or continuous.

Though the optimality of (s, S) policies for periodic review inventory control problems is a well-known fact, for average cost problems its rigorous proof without the assumption that the demand is either discrete or continuous was established not long ago by Chen and Simchi-Levi [7]. Even so, the proof provided in [7] is nontrivial and problem-specific. Feinberg and Lewis [13] provided a straightforward proof of this fact by using the results described here. In this paper we apply our results to the cash balance problem; an inventory control problem with possibly negative demand where the controller can increase and decrease the inventory levels.

Eppen and Fama [10] considered a cash balance model with i.i.d. discrete demands with finite support and without setup costs. They showed the existence of order-up-to and down-to levels in the finite and infinite horizon models with the discounted cost criterion; see Heyman and Sobel [21, Section 8.4], where the discrete nature of the demand was not assumed. Feinberg and Lewis [12] justified the average cost case with the general demand distribution and studied the problems with borrowing and lending options and without fixed costs, for which they established the optimality of simple four-threshold policies.

Girgis [17], Neave [25], and Chen and Simchi-Levi [6] studied problems with both fixed and variable costs for each transaction. Girgis [17] investigated finite and infinite horizon discounted cost problems when the demand is continuous and when there are fixed costs for increasing or decreasing demand (but not both); see also [9] where this and other simplified models were considered. Neave [25] studied finite horizon problems with continuous demand when both transactions have fixed costs. However, Ye and Duenyas [31] and Chen and Simchi-Levi [6] noticed that some of the claims in [25] were not proved. By using the notion of a (K, Q) -convex function introduced by Ye and Duenyas [31], Chen and Simchi-Levi [6] described structural properties of optimal solutions of finite horizon cash balance problems when both transactions have fixed costs. They assumed that certain functions relevant to the value functions (the functions $G_{n,\alpha}^+$ and $G_{n,\alpha}^-$ in the notation of this paper) are continuous and tend to ∞ as the inventory level tends to ∞ or $-\infty$. Other versions of the cash balance problem and models with product returns can be found in [20, 8, 30, 23, 15] and in the references therein.

We show that the structural results stated by Chen and Simchi-Levi [6] indeed hold, with a minor difference at the four threshold points, for finite horizon cash balance problems with the discounted and average cost per unit time criteria. Instead of continuity of the functions $G_{n,\alpha}^+$ and $G_{n,\alpha}^-$ assumed in [6], we use their lower semicontinuity that follows from the general MDP theory developed here. We also establish that these functions indeed tend to ∞ as the inventory level tends to ∞ or to $-\infty$.

The rest of the paper is organized as follows. In Section 2 we discuss the general Markov decision

process framework. Section 3 contains the main theoretical contributions of the paper, Theorem 3.1, that provides the assumptions that lead to: (i) the validity of the optimality inequalities for average cost MDPs, (ii) the existence of optimal policies for average cost MDPs, and (iii) the convergence of discounted cost optimal values and policies to average cost per unit time optimal values and policies. In Section 4 we discuss the relevance of the setwise and weak continuity assumptions to inventory control. In Section 5, we formulate the stochastic cash balance problem, show that it satisfies the assumptions of Theorem 3.1, and prove several structural properties of optimal policies under each criterion.

2. Model Definition Consider a discrete-time Markov decision process with state space \mathbb{X} and action space \mathbb{A} . Assume that both \mathbb{X} and \mathbb{A} are Borel subsets of Polish (complete separable metric) spaces. For each $x \in \mathbb{X}$ a nonempty Borel subset $A(x)$ of \mathbb{A} represents the set of actions available at x . Define the graph of \mathbb{A} by

$$Gr(\mathbb{A}) := \{(x, a) \mid x \in \mathbb{X}, a \in A(x)\},$$

and assume that: (i) $Gr(\mathbb{A})$ is a measurable subset of $\mathbb{X} \times \mathbb{A}$, and (ii) there exists a measurable mapping ϕ from \mathbb{X} to \mathbb{A} such that $\phi(x) \in A(x)$ for all $x \in \mathbb{X}$. The one step cost, $c(x, a)$, for choosing action $a \in A(x)$ in state x is a non-negative (or equivalently, bounded below) measurable function on $Gr(\mathbb{A})$. Let $q(B|x, a)$, also measurable on $Gr(\mathbb{A})$, be the transition kernel representing the probability that $B \subseteq \mathbb{X}$ is entered next, given that action a is chosen in state x . This means that $q(\cdot|x, a)$ is a probability measure on \mathbb{X} for each pair $(x, a) \in \mathbb{X} \times \mathbb{A}$, and $q(B|\cdot, \cdot)$ is a Borel function on $\mathbb{X} \times \mathbb{A}$ for any Borel subset $B \subseteq \mathbb{X}$.

The decision process proceeds as follows: at each time epoch $n = 0, 1, \dots$ the current state of the system, x , is observed. A decision-maker chooses an action a , the cost $c(x, a)$ is accrued, and the system moves to the next state according to $q(\cdot|x, a)$. Let $H_n = (\mathbb{X} \times \mathbb{A})^n \times \mathbb{X}$ be the set of histories for $n = 0, 1, \dots$. A (randomized) decision rule at epoch $n = 0, 1, \dots$ is a regular transition probability π_n from H_n to \mathbb{A} concentrated on $A(x_n)$. In other words, (i) $\pi_n(\cdot|h_n)$ is a probability distribution on \mathbb{A} such that $\pi_n(A(x_n)|h_n) = 1$, where $h_n = (x_0, a_0, x_1, \dots, a_{n-1}, x_n)$, and (ii) for any measurable subset $B \subseteq \mathbb{A}$, the function $\pi_n(B|\cdot)$ is measurable on H_n . A policy π is a sequence (π_0, π_1, \dots) of decision rules. Moreover, π is called non-randomized if each probability measure $\pi_n(\cdot|h_n)$ is concentrated at one point. A non-randomized policy is called Markov if all of the decisions depend only on the current state and time. A Markov policy is called stationary if all of the decisions depend only on the current state. Thus, a Markov policy ϕ is defined by a sequence ϕ_0, ϕ_1, \dots of measurable mappings $\phi_n : \mathbb{X} \rightarrow \mathbb{A}$ such that $\phi_n(x) \in A(x)$ for all $x \in \mathbb{X}$. A stationary policy ϕ is defined by a measurable mapping $\phi : \mathbb{X} \rightarrow \mathbb{A}$ such that $\phi(x) \in A(x)$ for all $x \in \mathbb{X}$.

The Ionescu–Tulcea theorem, [3, p. 140-141] or [19, p. 178], yields that an initial state x and a policy π define a unique probability distribution \mathbb{P}_x^π on the set of all trajectories $H_\infty = (\mathbb{X} \times \mathbb{A})^\infty$ endowed with the product σ -field defined by Borel σ -fields of \mathbb{X} and \mathbb{A} . Let \mathbb{E}_x^π be the expectation with respect to this distribution. For a finite horizon $N = 0, 1, \dots$, define the expected total discounted costs

$$v_{N,\alpha}^\pi(x) := \mathbb{E}_x^\pi \sum_{n=0}^{N-1} \alpha^n c(x_n, a_n), \quad (1)$$

where $\alpha \in [0, 1)$ and $v_{0,\alpha}^\pi = 0$. When $N = \infty$, (1) defines the infinite horizon expected total discounted cost of π denoted $v_\alpha^\pi(x)$. The average costs per unit time are

$$w^\pi(x) := \limsup_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}_x^\pi \sum_{n=0}^{N-1} c(x_n, a_n). \quad (2)$$

For each function $G^\pi(x) = v_{N,\alpha}^\pi(x)$, $v_\alpha^\pi(x)$, or $w^\pi(x)$, define the optimal cost $G(x) := \inf_{\pi \in \Pi} G^\pi(x)$, where Π is the set of all policies. A policy π is called *optimal* for the respective criterion if $G^\pi(x) = G(x)$ for all $x \in \mathbb{X}$.

It is well-known (see e.g. [3, Proposition 8.2]) that $v_{n,\alpha}(x)$ satisfies the following *optimality equations*,

$$v_{n+1,\alpha}(x) = \inf_{a \in A(x)} \{c(x, a) + \alpha \int v_{n,\alpha}(y)q(dy|x, a)\}, \quad x \in \mathbb{X}, n = 0, 1, \dots \quad (3)$$

In addition, a Markov policy ϕ , defined at the first N steps by the mappings $\phi_0, \dots, \phi_{N-1}$, that satisfy for all $n = 1, \dots, N$ the equations

$$v_{n,\alpha}(x) = c(x, \phi_{N-n}(x)) + \alpha \int v_{n-1,\alpha}(y)q(dy|x, \phi_{N-n}(x)), \quad x \in \mathbb{X}, \quad (4)$$

is optimal for the horizon N ; see e.g. [3, Lemma 8.7].

It is also well-known (see e.g. [3, Propositions 9.8 and 9.12]) that $v_\alpha(x)$ satisfies the following *discounted cost optimality equations* (DCOE),

$$v_\alpha(x) = \inf_{a \in A(x)} \{c(x, a) + \alpha \int v_\alpha(y)q(dy|x, a)\}, \quad x \in \mathbb{X}, \quad (5)$$

and a stationary policy ϕ_α is optimal if and only if

$$v_\alpha(x) = c(x, \phi_\alpha(x)) + \alpha \int v_\alpha(y)q(dy|x, \phi_\alpha(x)), \quad x \in \mathbb{X}. \quad (6)$$

Since $c \geq 0$, for each $\pi \in \Pi$ and $x \in \mathbb{X}$, $v_{n,\alpha}^\pi(x)$ is non-decreasing in n and bounded above by $v_\alpha^\pi(x)$. Thus, $v_{n,\alpha}(x)$ is also non-decreasing in n and bounded above by $v_\alpha(x)$. Therefore, $v_{\infty,\alpha}(x) \leq v_\alpha(x)$, $x \in \mathbb{X}$, where $v_{\infty,\alpha}(x)$ is the limit of $v_{n,\alpha}(x)$ as $n \rightarrow \infty$. In general, it is possible that $v_{\infty,\alpha}(x) < v_\alpha(x)$; [11, Example 6.6]. However, Assumption **(Wu)** below is sufficient for $v_{\infty,\alpha} = v_\alpha$; see Proposition 3.1.

A little more subtle is the average cost case. For the remainder of the paper, assume that the following condition holds.

Assumption (G): There exists a policy π and an initial state x such that $w^\pi(x) < \infty$.

Note that Assumption **(G)** is equivalent to the General Assumption of Schäl [27]: $\inf_{x \in \mathbb{X}} \inf_{\phi \in \Pi} w^\phi(x) < \infty$. Define the following quantities

$$m_\alpha := \inf_{x \in \mathbb{X}} v_\alpha(x), \quad u_\alpha(x) := v_\alpha(x) - m_\alpha,$$

and $w^* = \liminf_{\alpha \rightarrow 1} (1 - \alpha)m_\alpha$.

Assumption **(G)** implies that $w^* < \infty$; Schäl [27, Lemma 1.2]. According to Schäl [27, Proposition 1.3], if there exists a measurable function $u : \mathbb{X} \rightarrow [0, \infty)$ and a stationary policy ϕ such that

$$w^* + u(x) \geq c(x, \phi(x)) + \int u(y)q(dy|x, \phi(x)), \quad x \in \mathbb{X}, \quad (7)$$

then ϕ is average cost optimal and $w(x) = w^*$ for all $x \in \mathbb{X}$. The following condition plays an important role for the validity of (7).

Assumption (B): Assumption **(G)** holds and $\sup_{\alpha < 1} u_\alpha(x) < \infty$ for all $x \in \mathbb{X}$.

We note that the second part of Assumption **(B)** is Condition **(B)** in Schäl [27]. Thus, under Assumption **(G)**, which is assumed throughout [27], Assumption **(B)** is equivalent to Condition **(B)** in [27].

3. Main Structural Result Weak continuity (or continuity with respect to weak convergence) of q in (x, a) means that

$$\int f(z)q(dz|x^k, a^k) \rightarrow \int f(z)q(dz|x, a) \quad (8)$$

for any sequence $\{(x^k, a^k), k \geq 0\}$ converging to (x, a) , where $(x^k, a^k), (x, a) \in Gr(\mathbb{A})$, and for any bounded continuous function f . For a real-valued function f defined on a metric space \mathbb{Y} , consider the level sets $D(\lambda) = \{y \in \mathbb{Y} \mid f(y) \leq \lambda\}$, $-\infty < \lambda < \infty$. We recall that the function f is *lower semi-continuous* if all the level sets $D(\lambda)$ are closed and the function is *inf-compact* if all these sets are compact; see [19, page 170]. Let $\mathbb{P}(\mathbb{X})$ be the set of all probability measures on \mathbb{X} . We next state Assumption **(Wu)**, which is similar to [27, Condition **(W)**], but does not require that action sets are compact. The letter “u” signifies unbounded action sets.

Assumption (Wu): \mathbb{X} is locally compact with a countable base, and

- (i) c is inf-compact on $Gr(\mathbb{A})$.
- (ii) $q : Gr(\mathbb{A}) \rightarrow \mathbb{P}(\mathbb{X})$ is continuous with respect to weak convergence on $\mathbb{P}(\mathbb{X})$.

We remark that if (7) holds for a stationary policy ϕ and for a non-negative function u then obviously

$$w^* + u(x) \geq \inf_{a \in A(x)} \{c(x, a) + \int u(y)q(dy|x, a)\}, \quad x \in \mathbb{X}. \quad (9)$$

and Assumption **(Wu)** and the Arsenin-Kunugui theorem [24, Theorem 35.46] imply the existence of a policy ϕ such that the minimum in (9) is achieved when $a = \phi(x)$. Thus, infimum can be replaced with minimum in (9). The following proposition states that Assumption **(Wu)** implies the existence of stationary optimal policies for the expected total discounted cost criterion. When transition probabilities are weakly continuous, this proposition is similar to [27, Proposition 2.1] for models with compact action sets.

PROPOSITION 3.1 *Let Assumption **(Wu)** hold. Then*

- (i) *For any $N = 1, 2, \dots$ there exists a Markov optimal N -horizon policy $(\phi_0, \dots, \phi_{N-1})$ satisfying (4).*
- (ii) *For infinite horizon expected total discounted costs there exists a stationary optimal policy ϕ_α satisfying (6).*
- (iii) *$v_{n,\alpha}(x) \uparrow v_\alpha(x)$ as $n \rightarrow \infty$ for all $x \in \mathbb{X}$.*
- (iv) *The functions $v_{n,\alpha}$, $n = 1, 2, \dots$, and v_α are inf-compact on \mathbb{X} .*

PROOF. Note that the weak continuity Assumption **(Wu(ii))** is equivalent to lower semi-continuity in (x, a) of the function $p(x, a) = \int f(y)q(dy|x, a)$ for all lower semi-continuous, non-negative functions f ; [19, Proposition C.4]. Since $v_{0,\alpha}(x) = 0$ for all x , it is lower semi-continuous. Let $v_{n,\alpha}$ be lower semi-continuous for some n . Note that for all $\lambda \in \mathbb{R}$,

$$\begin{aligned} B_n(\lambda) &:= \{(x, a) \in Gr(\mathbb{A}) \mid c(x, a) + \alpha \int v_{n,\alpha}(y)q(dy|x, a) \leq \lambda\} \\ &\subseteq \{(x, a) \in Gr(\mathbb{A}) \mid c(x, a) \leq \lambda\} =: D(\lambda). \end{aligned}$$

Since $c(x, a)$ and $v_{n,\alpha}(x)$ are lower semi-continuous, $B_n(\lambda)$ is a closed set. Thus, since the definition of inf-compactness implies that $D(\lambda)$ is compact, $B_n(\lambda)$ is compact because it is a closed subset of a compact set. Thus, $J_n(x, a) = c(x, a) + \alpha \int v_{n,\alpha}(y)q(dy|x, a)$ is inf-compact on $Gr(\mathbb{A})$. In addition, $v_{n+1,\alpha}(x) = \min_{a \in A(x)} J_n(x, a)$ is inf-compact because $\{x \in \mathbb{X} \mid v_{n+1,\alpha}(x) \leq \lambda\}$ is the projection of the compact set $B_n(\lambda)$.

The compactness of $B_n(\lambda)$ implies that $v_{n,\alpha} \uparrow v_\alpha$ (see [3, Proposition 9.17]). Therefore, the sets $\{x \in \mathbb{X} \mid v_\alpha(x) \leq \lambda\} = \bigcap_{n=1}^{\infty} \{x \in \mathbb{X} \mid v_{n,\alpha}(x) \leq \lambda\}$ are compact; v_α is inf-compact. This coupled with the weak continuity of q implies that $J(x, a) := c(x, a) + \alpha \int v_\alpha(y)q(dy|x, a)$ is lower semi-continuous. Repeating the same argument as above yields the inf-compactness of $J(x, a)$ and (5) holds with infimum replaced by minimum. Moreover, since $J_n(x, a)$ and $J(x, a)$ are inf compact for all n , the Arsenin-Kunugui theorem [24, Theorem 35.46] implies the existence of a Markov policy satisfying (4) and the existence of a stationary policy satisfying (6). \square

We remark that Proposition 9.17 in Bertsekas and Shreve [3] states the existence of a stationary optimal policy in addition to the convergence of $v_{n,\alpha}$ to v_α . In the proof of Proposition 3.1 we used the convergence result but not the existence result because [3] used a more general concept of measurability than is assumed in this paper. In fact, the above proof implies that the assumptions of [3, Proposition 9.17] imply the existence of a Borel measurable stationary optimal policy. Since this observation is not stated in [3], we provided the proof here.

The next definition is a natural extension of the definition of a locally bounded function [16, p. 113].

DEFINITION 3.1 *A real-valued function f defined on a metric space \mathbb{Y} is called **locally bounded above at** $x \in \mathbb{Y}$ if there exists an open set $B(x)$ containing x such that $\sup_{y \in B(x)} f(y) < \infty$. The function f is called **locally bounded above** if it is locally bounded for each $x \in \mathbb{Y}$.*

For a point $x \in \mathbb{Y}$, let $\mathcal{O}(x) := \{B \subseteq \mathbb{Y} \mid x \in B, B \text{ is open}\}$ be the set of open sets containing x . Consider the function

$$\bar{f}(x) := \inf_{B \in \mathcal{O}(x)} \sup_{y \in B} f(y). \tag{10}$$

Obviously, $f(x) \leq \bar{f}(x)$ and, if f is locally bounded at x , $\bar{f}(x) < \infty$.

LEMMA 3.1 *Suppose f is a locally bounded above function on a complete separable metric space \mathbb{Y} . Then the function $\bar{f}(x)$ is upper semi-continuous.*

PROOF. Fix $\epsilon > 0$ and let $B_\epsilon(x)$ be an open set containing x such that $\bar{f}(x) \geq \sup_{y \in B_\epsilon(x)} f(y) - \epsilon$. Let $x_n \rightarrow x$. Choose $k(\epsilon)$ such that $x_n \in B_\epsilon(x)$ for all $n \geq k(\epsilon)$. Then $\bar{f}(x_n) \leq \sup_{y \in B_\epsilon} f(y) \leq \bar{f}(x) + \epsilon$ for $n \geq k(\epsilon)$. Thus, $\bar{f}(x) \geq \limsup_{n \rightarrow \infty} \bar{f}(x_n) - \epsilon$. Since $\epsilon > 0$ is arbitrary, $\bar{f}(x) \geq \limsup_{n \rightarrow \infty} \bar{f}(x_n)$. \square

For $\alpha \in [0, 1)$ define $r_\alpha(x) = \sup_{\alpha \leq \beta < 1} u_\beta(x)$. Assumption **(B)** is equivalent to the validity of the following pair of assumptions: (i) Assumption **(G)**, and (ii) for any $x \in \mathbb{X}$ there exists $\alpha \in [0, 1)$ with $r_\alpha(x) < \infty$. Thus, Assumption **(B)** is equivalent to [19, Condition 5.4.5] and implies that $r_\alpha(x) < \infty$ for all $\alpha \in [0, 1)$ and for $x \in \mathbb{X}$.

Let $\Gamma_\alpha(x) := \sup_{\alpha \leq \beta < 1} \{v_\beta(x) - \beta m_\beta\}$, where $\alpha \in [0, 1)$. According to [27, Lemma 1.2], Assumption **(G)** implies that $\limsup_{\alpha \rightarrow 1} (1 - \alpha)m_\alpha < \infty$. Thus, if Assumption **(B)** holds, there exists $\alpha^* \in [0, 1)$ such that for all $\alpha \in [\alpha^*, 1)$

$$\Gamma_\alpha(x) \leq r_\alpha(x) + \sup_{\alpha \leq \beta < 1} \{(1 - \beta)m_\beta\} < \infty. \quad (11)$$

The following condition strengthens **(B)**.

Assumption (LB). Assumption **(G)** holds and there exists $\alpha_0 \in [0, 1)$ such that the function $r_{\alpha_0}(x)$ is locally bounded above on \mathbb{X} .

We observe that the function $\Gamma_\alpha(x)$ decreases in α . Thus, if **(LB)** holds then the function Γ_α is locally bounded above for any $\alpha \in [\alpha_0, 1)$. We select an arbitrary $\alpha^* \in [\alpha_0, 1)$, such that (11) holds when $\alpha = \alpha^*$, and denote $\Gamma = \Gamma_{\alpha^*}$.

We say that an MDP is a *submodel* of another MDP if the only difference between these MDPs is that the sets of available actions of the former MDP are subsets of available action sets of the latter. For $\bar{\Gamma}$ defined by (10) with $\Gamma = f$, consider the following subsets of the sets of available actions in states $x \in \mathbb{X}$

$$\tilde{A}(x) := \{a \in A(x) \mid c(x, a) \leq \bar{\Gamma}(x)\}.$$

Let Assumption **(LB)** hold and $\alpha \in [\alpha_0, 1)$. According to Lemma 3.1, the function $\bar{\Gamma}$ is upper semi-continuous and is therefore measurable. This implies that the graph of the mapping $\tilde{A} : \mathbb{X} \rightarrow \mathbb{A}$ is Borel, and the set $\{\mathbb{X}, \mathbb{A}, \tilde{A}(x), q, c\}$ defines a submodel of the original MDP $\{\mathbb{X}, \mathbb{A}, A(x), q, c\}$. Recall that a set-valued mapping $F : \mathbb{X} \rightarrow \mathbb{A}$ is called *closed* at x if $a_n \in F(x_n)$ and $(x_n, a_n) \rightarrow (x, a)$ imply $a \in F(x)$. A set-valued mapping $F : \mathbb{X} \rightarrow \mathbb{A}$ is called *upper semi-continuous* at x if for any neighborhood G of the set $F(x)$, there is a neighborhood of x , say $U(x)$, such that $F(y) \subseteq G$ for all $y \in U(x)$. A set-valued mapping is called closed (upper semi-continuous) if it is closed (upper semi-continuous) at all $x \in \mathbb{X}$; see Definitions 4.4 and 4.5 in [26] or Definitions 1 and 2 in [22, pp. 21-23]. Moreover, a closed set-valued mapping F is upper semi-continuous if all sets, $F(x)$, are subsets of a compact set; see [26, Lemma 4.4] or the comments prior to Proposition 2 of [22, p. 23]). We make use of this fact in the proof of the following lemma.

LEMMA 3.2 *If Assumptions **(LB)** and **(Wu)** hold then the mapping $\tilde{A}(x)$ is upper semi-continuous.*

PROOF. We show first that the mapping \tilde{A} is closed. Suppose $a_n \in \tilde{A}(x_n)$ and $(a_n, x_n) \rightarrow (a, x)$. As previously noted, Lemma 3.1 yields that the function $\bar{\Gamma}(x)$ is upper semi-continuous. The definition of $\tilde{A}(x_n)$ implies $c(x_n, a_n) \leq \bar{\Gamma}(x_n)$. Thus,

$$c(x, a) \leq \liminf_{n \rightarrow \infty} c(x_n, a_n) \leq \limsup_{n \rightarrow \infty} c(x_n, a_n) \leq \limsup_{n \rightarrow \infty} \bar{\Gamma}(x_n) \leq \bar{\Gamma}(x),$$

where the first and fourth inequalities follow from the lower semi-continuity of c and from the upper semi-continuity of $\bar{\Gamma}$, respectively. So, $a \in \tilde{A}(x)$ as desired, and the mapping $\tilde{A}(x)$ is closed.

To show that $\tilde{A}(x)$ is upper semi-continuous, it is sufficient to show that any $x \in \mathbb{X}$ has a neighborhood U such that $\tilde{A}(U) := \{a \in \tilde{A}(y) \mid y \in U\}$ is a subset of a compact set. Fix $x \in \mathbb{X}$. Since $\bar{\Gamma}$ is an upper semi-continuous function, there exists a neighborhood U of x such that $\bar{\Gamma}(y) \leq \bar{\Gamma}(x) + 1$ for all $y \in U$.

The set $B = \{(y, a) \in \mathbb{X} \times \mathbb{A} \mid c(y, a) \leq \bar{\Gamma}(x) + 1\}$ is compact. Therefore, its projection on \mathbb{A} , say $B_{\mathbb{A}}$, is also compact. For each $y \in U$ we have

$$\tilde{A}(y) = \{a \in A(y) \mid c(y, a) \leq \bar{\Gamma}(y)\} \subseteq \{a \in A(y) \mid c(y, a) \leq \bar{\Gamma}(x) + 1\} \subseteq B_{\mathbb{A}},$$

and the proof is complete. \square

DEFINITION 3.2 *We say that a submodel with action sets $A'(x)$, $x \in \mathbb{X}$, **represents** the original MDP if there exists $\tilde{\alpha} \in [0, 1)$ such that, under the α -discounted cost criterion with $\alpha \in [\tilde{\alpha}, 1)$, any stationary optimal policy, say $\phi_{\alpha}(x)$, for the original MDP belongs to this submodel, i.e. $\phi_{\alpha}(x) \in A'(x)$ for all $x \in \mathbb{X}$.*

PROPOSITION 3.2 *If Assumptions **(Wu)** and **(LB)** hold for the original MDP then the submodel $\{\mathbb{X}, \mathbb{A}, \tilde{A}(x), q, c\}$ represents the original MDP. Moreover, this submodel satisfies Conditions **(B)** and **(W)** from Schäl [27, page 165].*

PROOF. In view of Proposition 3.1, there exist stationary α -discounted cost optimal policies. According to (5) and (6), a stationary policy ϕ_{α} is α -discounted cost optimal if and only if

$$(1 - \alpha)m_{\alpha} + u_{\alpha}(x) = \min_{a \in A(x)} \{c(x, a) + \alpha \int u_{\alpha}(y)q(dy|x, a)\} = c(x, \phi_{\alpha}(x)) + \alpha \int u_{\alpha}(y)q(dy|x, \phi_{\alpha}(x)). \quad (12)$$

Let $\alpha \in [\alpha^*, 1)$, where α^* is defined in the paragraph following Assumption **(LB)**. Since $u_{\alpha} \geq 0$, we have that $c(x, \phi_{\alpha}(x)) \leq (1 - \alpha)m_{\alpha} + u_{\alpha}(x) = v_{\alpha}(x) - \alpha m_{\alpha} \leq \Gamma_{\alpha^*}(x) \leq \bar{\Gamma}(x)$. Therefore, $\phi_{\alpha}(x) \in \tilde{A}(x)$ for all $x \in \mathbb{X}$. The stationary policy ϕ_{α} belongs to the submodel and, thus, is optimal for the submodel. In light of this fact, the value functions v_{α} for the original MDP and for the submodel are equal. Thus, the submodel $\{\mathbb{X}, \mathbb{A}, \tilde{A}(x), q, c\}$ satisfies Assumption **(LB)** which implies that it satisfies Assumption **(B)**.

Since the function $c(x, a)$ is inf-compact in a , the sets $\tilde{A}(x)$ are compact. Lemma 3.2 implies that the mapping $\tilde{A}(x)$ is upper semi-continuous under **(LB)** and **(Wu)**. The remaining conditions in **(W)** from [27] follow directly from the corresponding conditions in **(Wu)**. \square

Proposition 3.2 coupled with Schäl [27, Proposition 3.5 and Theorem 3.8] imply the main result of this section.

THEOREM 3.1 *Suppose Assumptions **(Wu)** and **(LB)** hold. There exists a function $u : \mathbb{X} \rightarrow [0, \infty)$ and a stationary policy ϕ satisfying (7). Thus, $w^{\phi}(x) = w(x) = w^*$. Furthermore $w^* = \lim_{\alpha \rightarrow 1} (1 - \alpha)m_{\alpha} = \lim_{\alpha \rightarrow 1} (1 - \alpha)v_{\alpha}(x)$, $x \in \mathbb{X}$. Also, for any discount factor α fix an optimal policy ϕ_{α} . Then for any sequence of discount factors $\alpha(k) \rightarrow 1$ the following statements hold:*

(a) *u can be defined as the inf-compact function*

$$u(x) = \inf \left\{ \liminf_{n \rightarrow \infty} u_{\alpha(n)}(x_n) : x_n \rightarrow x \right\}; \quad (13)$$

(b) *for each $x \in \mathbb{X}$ there exists a sequence $x_m \rightarrow x$ and a subsequence α_m of the sequence $\alpha(k)$ such that the mapping $\phi(x) = \lim_{m \rightarrow \infty} \phi_{\alpha_m}(x_m)$ defines a stationary policy that satisfies (7) with u defined in (13).*

PROOF. All of the statements of this theorem, except the inf-compactness of u stated in (a), follow from Schäl [27, Proposition 3.5 and Theorem 3.8] and Proposition 3.2 above. Let Assumption **(Wu)** hold. Consider a sequence $x_n \rightarrow x$. For each $n = 1, 2, \dots$ consider an integer $k(n) \geq n$ and a point $y_{k(n)}$ in the neighborhood of x_n with the radius n^{-1} such that $u_{\alpha(k(n))}(y_{k(n)}) \leq u(x_n) + n^{-1}$. Then

$$\liminf_{n \rightarrow \infty} u(x_n) \geq \liminf_{n \rightarrow \infty} u_{\alpha(k(n))}(y_{k(n)}) \geq u(x),$$

where the last inequality follows from (13). Thus, u is lower semi-continuous. We remark that this proof also implies that u is lower semi-continuous if Assumption **(Wu)** is substituted with Condition **(W)** from Schäl [27].

Since the function u is lower semi-continuous, for any finite constant λ , the set $D_u(\lambda) = \{x \in \mathbb{X} \mid u(x) \leq \lambda\}$ is closed. Since c is inf-compact under **(Wu)**, the set $D_c(\lambda^*) = \{(x, a) \in \mathbb{X} \times \mathbb{A} \mid c(x, a) \leq \lambda^*\}$, where

$\lambda^* = \lambda + w^*$, is compact. Therefore, its projection $D_c^{\mathbb{X}}(\lambda^*)$ on \mathbb{X} is compact as well. Since $D_u(\lambda) \subseteq D_c^{\mathbb{X}}(\lambda^*)$, the set $D_u(\lambda)$ is compact and u is inf-compact under **(Wu)**. \square

In the next section, we discuss when setwise continuity holds in inventory control/cash balance problems.

4. Relevance of Weak and Setwise Continuity to Inventory Control We recall that setwise continuity of the transition probability q in a means that $q(B|x, a^k) \rightarrow q(B|x, a)$ as $a^k \rightarrow a$, where $a^k, a \in A(x)$, for any Borel subset B of \mathbb{X} and for any state $x \in \mathbb{X}$. Setwise continuity is a stronger assumption than weak continuity; see e.g. Hernández-Lerma and Lasserre [19, page 186].

Consider the typical dynamic state equation for inventory control/cash balance models

$$x_{n+1} = x_n + a_n - D_{n+1}, \quad n = 0, 1, 2, \dots, \quad (14)$$

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 . The demand is assumed to be i.i.d. Let $q(dx_{n+1}|x_n, a_n)$ be the probability distribution of x_{n+1} for given x_n and a_n . As was shown in Feinberg and Lewis [12, page 48], the weak continuity Assumption **(Wu(ii))** holds when the dynamics of the controlled process are defined by (14) and the state space is the real line, while the setwise continuity assumption does not hold.

In this section, we study the following natural question: when are the transition probabilities setwise continuous for inventory control/cash balance problems? The answer is that it holds for two particular cases often considered in the inventory control literature: (i) when the demand may take only integer values and the state space is the set of (or some subset of) the integers, and (ii) when demand D is a continuous random variable and the state space is the real line or some subset of the real line.

For (i) we observe that any function on a countable space with the discrete topology is continuous and therefore the notions of weak and setwise continuity coincide. In order to analyze the continuous demand case (ii), consider the notion of the distance in variation (also often called in total variation) between two probability measures P and Q on \mathbb{X} ,

$$\|P - Q\| = \sup_{A \in \mathcal{B}(\mathbb{X})} |P(A) - Q(A)|,$$

where $\mathcal{B}(\mathbb{X})$ is the Borel σ -field on \mathbb{X} . Since $|P(A) - Q(A)| \leq \|P - Q\|$, convergence in total variation implies setwise convergence.

The following lemma applied to (14) implies that the setwise continuity of transition probabilities holds when D_n are continuous random variables.

LEMMA 4.1 *Let a random variable ξ have a density $f(x)$ with respect to Lebesgue integration. Consider random variables $\eta^k = y^k + \xi$, where y^k is a convergent sequence of real numbers, $y^k \rightarrow y^0$. Then the probability distributions of the random variables η^k converge in variation to the probability distribution of η^0 .*

PROOF. See Appendix.

We end this section by mentioning that the above conclusions that setwise continuity holds in inventory control when either the demand distribution is continuous or the inventory is integer extends to other classic models. Consider, for example, the dynamic equations for inventory control problems with lost sales,

$$x_{n+1} = (x_n + a_n - D_{n+1})^+, \quad n = 0, 1, 2, \dots, \quad (15)$$

where $c^+ = \max\{0, c\}$ for a number c . Observe that the counter-example [12, page 48], where weak continuity holds and setwise continuity does not, still applies. Moreover, Lemma 4.1 also remains valid.

5. The Cash Balance Problem In this section we consider the stochastic cash balance problem with fixed ordering costs. We show that Assumptions **(LB)** and **(Wu)** hold so that the results of Proposition 3.1 and Theorem 3.1 yield the existence of stationary optimal policies for the infinite-horizon discounted and average costs per unit time criterion, respectively. We then use these results to prove that the structure of the optimal policy, established by Chen and Simchi-Levi [6] for the finite horizon

case, extends to the infinite horizon expected discounted cost and average cost cases. Before describing the model we state a technical lemma and a definition that will be used in the analysis. Consider the following assumptions.

Assumption (C):

- (i) The constant $\bar{w} := \limsup_{\alpha \rightarrow 1} (1 - \alpha)m_\alpha$ is finite.
- (ii) For any finite real number N 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 (C(i)) actually follows from Assumption (G) (see [27, Lemma 1.2]), but is included here for completeness. Let $M(\alpha) := \{x \in \mathbb{X} | v_\alpha(x) = m_\alpha\}$. The next result is similar to [12, Lemma 6.1] and, as mentioned in [12], to [5, Lemma 4] and [27, Lemma 6.6].

LEMMA 5.1 *Let Assumption (C) hold. Then there exists $\alpha_0 < 1$ and a compact subset $\mathcal{K} \subseteq \mathbb{X}$ such that $M(\alpha) \subseteq \mathcal{K}$ for all $\alpha \in [\alpha_0, 1)$.*

PROOF. Consider $N > \bar{w} + 1 = \limsup_{\alpha \rightarrow 1} (1 - \alpha)m_\alpha + 1$. Therefore, there exists $\alpha_0 < 1$ such that $N/(1 - \alpha) > m_\alpha + (1 - \alpha)^{-1}$ for all $\alpha \in [\alpha_0, 1)$. This formula is identical to (6.2) in [12] and the rest of the proof coincides with the proof of Lemma 6.1 in [12] following (6.2) there. \square

The model has the following decision-making scenario: a decision-maker views the current inventory and makes an ordering decision. Assuming zero lead times, the products are immediately available to meet demand. Demand is then realized, the decision-maker views the remaining inventory, and the process continues. Assume the unmet demand is backlogged and the cost of inventory held or backlogged (negative inventory) is modeled as a convex function. The difference between this model and the classic inventory control model is that the demand and the order quantity are not assumed to be non-negative. That is customer returns are modeled as negative demand and a decision-maker may decide to order if inventory is too low or to scrap inventory if the level is too high. The major parameters of the model follow:

- $\alpha \in (0, 1)$ is the discount factor,
- $K \geq 0$ is the fixed ordering cost,
- $Q \geq 0$ is the fixed scrapping cost,
- $c_+ > 0$ is the per unit ordering cost,
- $c_- > 0$ is the per unit scrapping cost,
- $h(\cdot)$ denotes the holding/backordering cost per period; convex, non-negative, takes finite values, and $h(x) \rightarrow \infty$ as $|x| \rightarrow \infty$, and
- $\{D_n, n \geq 0\}$ is a sequence of i.i.d. random variables where D_n (not necessarily non-negative) represents demand in the n^{th} period. We assume that $\mathbb{E}h(x - D) < \infty$ for all $x \in \mathbb{R}$ and $P(D \neq 0) > 0$, where D is a random variable with the same distribution as D_n .

Without loss of generality, assume that $h(0) = 0$ and that $K \geq Q$. The fact that $P(D \neq 0) > 0$ avoids the trivial case. For example, if $D = 0$ almost surely then the policy that orders or scraps to zero initially and then does not order or scrap at zero is optimal under the average cost criterion. Note that the finiteness of $\mathbb{E}h(x - D)$ and the assumed properties of the function h imply that $\mathbb{E}|D| < \infty$. Also note that the problem is posed with $\mathbb{X} = \mathbb{R}$. However, if the demand and action sets are integer or with probability 1 on a lattice, the model can be restated with $\mathbb{X} = \mathbb{Z}$; see Remark 5.1.

The cost function for the model is

$$c(x, a) = K1_{\{a > 0\}} + Q1_{\{a < 0\}} + c_+a^+ + c_-a^- + \mathbb{E}h(x + a - D). \quad (16)$$

For $n = 0, 1, \dots$ define the functions

$$\tilde{J}_n(x, a) := K1_{\{a > 0\}} + Q1_{\{a < 0\}} + c_+a^+ + c_-a^- + \mathbb{E}h(x + a - D) + \alpha \mathbb{E}v_{n, \alpha}(x + a - D), \quad (17)$$

$$\tilde{J}(x, a) := K1_{\{a > 0\}} + Q1_{\{a < 0\}} + c_+a^+ + c_-a^- + \mathbb{E}h(x + a - D) + \alpha \mathbb{E}v_\alpha(x + a - D). \quad (18)$$

The finite-horizon optimality equations are

$$v_{n+1, \alpha}(x) = \inf_{a \in \mathbb{R}} \{\tilde{J}_n(x, a)\}, \quad x \in \mathbb{X}, \quad n = 0, 1, \dots \quad (19)$$

with $v_{0,\alpha}(x) = 0$. The DCOE are

$$v_\alpha(x) = \inf_{a \in \mathbb{R}} \{\tilde{J}(x, a)\}, \quad x \in \mathbb{X}, \quad (20)$$

and the ACOI are

$$\begin{aligned} w^* + u(x) &\geq \inf_{a \in \mathbb{R}} \{K1_{\{a>0\}} + Q1_{\{a<0\}} + c_+a^+ + c_-a^- \\ &\quad + \mathbb{E}[h(x+a-D) + u(x+a-D)]\}, \quad x \in \mathbb{X}. \end{aligned} \quad (21)$$

The next result states that the weak continuity and local boundedness assumptions hold for the cash balance problem.

PROPOSITION 5.1 *In the stochastic cash balance problem, Assumptions **(Wu)** and **(LB)** hold and, therefore, Proposition 3.1 and Theorem 3.1 hold.*

PROOF. Let ϕ be the policy that always orders up or scraps down to the level 0. Then $w^\phi(x) \leq K + \max\{c_+, c_-\} \mathbb{E}|D| + \mathbb{E}h(-D) < \infty$. Note that this implies **(C(i))**. Moreover, Jensen's inequality implies that $c(x, a) \geq c_-a^- + c_+a^+ + h(x+a - \mathbb{E}D)$. For any $N > 0$ consider $z_N > 0$ such that $\min\{h(z_N), h(-z_N)\} \geq N$. Then $c(x, a) \geq N$ when $x \notin [-z_N - N/c_+ + \mathbb{E}D, z_N + N/c_- + \mathbb{E}D]$. Indeed, either the action takes the inventory within $[-z_N, z_N]$ at a cost higher than N or the action does not take the inventory in this interval and the holding cost is greater than N . Thus, **(C(ii))** holds so that the results of Lemma 5.1 hold.

To complete the proof that Assumption **(LB)** holds, let α_0 be such that the statement of Lemma 5.1 holds and let $\mathcal{K} = [x_L, x_U] \supseteq M(\alpha)$ for all $\alpha \in [\alpha_0, 1)$. We fix $x_\alpha \in M(\alpha)$ for each $\alpha \in [\alpha_0, 1)$. Suppose ϕ_α is a policy that at the first step immediately produces up to level x_α if the initial state $x \leq x_\alpha$ and it immediately scraps down to x_α if $x > x_\alpha$ plus at the first step ϕ_α orders or scraps whatever amount a stationary, discounted cost optimal policy starting in x_α would. Starting at step 2, this policy coincides with a stationary optimal policy. Thus, for all $\alpha \in [\alpha_0, 1)$,

$$u_\alpha(x) \leq v^{\phi_\alpha}(x) - m_\alpha \leq K + \max\{c_+, c_-\} \max\{|x_L - x|, |x_U - x|\} < \infty. \quad (22)$$

Since the right hand side of (22) is independent of α we have $r_{\alpha_0} \leq K + \max\{c_+, c_-\} \max\{|x_L - x|, |x_U - x|\}$. Note that the maximum of two convex functions is convex and therefore continuous. Since continuous functions are locally bounded, Assumption **(LB)** holds.

The cost function c in (16) is the sum of an inf-compact function in a and an inf-compact function on $Gr(\mathbb{A})$. Therefore, c is inf-compact on $Gr(\mathbb{A})$. The weak continuity of the transition kernel is discussed in detail in the beginning of Section 4. Thus, Assumption **(Wu)** holds. \square

Proposition 5.1 implies that the infimums in (19)–(21) can be replaced with minimums. Thus, these formulas can be rewritten as

$$\begin{aligned} v_{n+1,\alpha}(x) &= \min\{\min_{b>x}\{K + c_+(b-x) + L_{n,\alpha}(b)\}, \min_{b<x}\{Q + c_-(x-b) + L_{n,\alpha}(b)\}, L_{n,\alpha}(x)\}, \\ v_\alpha(x) &= \min\{\min_{b>x}\{K + c_+(b-x) + L_\alpha(b)\}, \min_{b<x}\{Q + c_-(x-b) + L_\alpha(b)\}, L_\alpha(x)\}, \\ w + u(x) &\geq \min\{\min_{b>x}\{K + c_+(b-x) + W(b)\}, \min_{b<x}\{Q + c_-(x-b) + W(b)\}, W(x)\}, \end{aligned}$$

where $L_{n,\alpha}(x) := \mathbb{E}h(x-D) + \alpha \mathbb{E}v_{n,\alpha}(x-D)$, $L_\alpha(x) := \mathbb{E}h(x-D) + \alpha \mathbb{E}v_\alpha(x-D)$, and $W(x) := \mathbb{E}h(x-D) + \alpha \mathbb{E}u(x-D)$. Consider the following definition.

DEFINITION 5.1 (Ye and Duenyas [31]). *A measurable real-valued function f is called (K, Q) -convex if for any $x \leq y$, $K, Q \geq 0$ and $\lambda \in [0, 1]$,*

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y) + \lambda K + (1-\lambda)Q - \min\{\lambda, 1-\lambda\} \min\{K, Q\}.$$

We remark that the notions of $(K, 0)$ -convex and K -convex functions coincide.

Define the following functions:

$$\begin{aligned} G_{n,\alpha}^+(x) &= c_+x + L_{n,\alpha}(x), & G_{n,\alpha}^-(x) &= -c_-x + L_{n,\alpha}(x), \\ G_\alpha^+(x) &= c_+x + L_\alpha(x), & G_\alpha^-(x) &= -c_-x + L_\alpha(x). \end{aligned}$$

For an inf-compact function g , fix any $S \in \operatorname{argmin}_{x \in \mathbb{R}} \{g(x)\}$. Let

$$s = \min\{x \mid g(x) \leq K + g(S)\}, \quad (23)$$

$$s' = \min\{x \mid Q + g(x) \leq g(S) + K\}, \quad (24)$$

$$u = \max\{x \mid g(x) \leq g(S) + Q\}. \quad (25)$$

These finite numbers are well-defined because g is inf-compact. Obviously, $s \leq s' \leq S \leq u$.

The following lemma is a restatement of Lemma 1 in Chen and Simchi-Levi [6] for lower semi-continuous functions and is analogous to Bertsekas [1, Lemma 4.2.1(c,d)] for K -convexity.

LEMMA 5.2 (1) Suppose $g_1(y)$ and $g_2(y)$ are (K_1, Q_1) -convex and (K_2, Q_2) -convex, respectively and that $(K_1 - Q_1)(K_2 - Q_2) \geq 0$. Then for $\beta_1, \beta_2 \geq 0$, $\beta_1 g_1 + \beta_2 g_2$ is $(\beta_1 K_1 + \beta_2 K_2, \beta_1 Q_1 + \beta_2 Q_2)$ -convex.

(2) If D is a random variable, then $\mathbb{E}g(y - D)$ is also (K, Q) -convex provided $\mathbb{E}|g(y - D)| < \infty$.

(3) Suppose g is a lower semi-continuous, (K, Q) -convex function with $K \geq Q$ such that $g(x) \rightarrow \infty$ as $|x| \rightarrow \infty$. Then

(i) $g(s) \leq g(S) + K$ and $g(x) > g(S) + K$ for all $x < s$,

(ii) $g(u) \leq g(S) + Q$ and $g(x) > g(S) + Q$ for all $x > u$,

(iii) $g(x) \leq g(z) + Q$ for all x, z with $z \leq x \leq s'$,

(iv) $g(x) \leq g(z) + K$ for all x, z with $s' \leq x \leq z$,

(v) $g(x) \leq g(z) + K$ for all x, z with $\frac{s+S}{2} \leq x \leq z$.

Following Chen and Simchi-Levi [6], we shall apply Lemma 5.2 to the functions $g = G_{N,\alpha}^+$, $g = G_{N,\alpha}^-$, $g = G_\alpha^+$, and $g = G_\alpha^-$. In order to do this, we must show that these functions are lower semi-continuous, (K, Q) -convex, and approach infinity in both directions. The next result takes the first steps in that direction.

LEMMA 5.3 For each $x \in \mathbb{R}$ and $\alpha \in [0, 1)$,

(i) $\mathbb{E}v_\alpha(x - D)$ and $\mathbb{E}v_{n,\alpha}(x - D)$ are finite and lower semicontinuous in x for all $n \geq 0$, and

(ii) G_α^+ , G_α^- , $G_{n,\alpha}^+$ and $G_{n,\alpha}^-$ are (K, Q) -convex and lower semicontinuous in x for all $n \geq 0$.

PROOF. We prove the results in the infinite horizon case. The finite horizon case is similar. Recall from Proposition 3.1 that v_α is inf-compact and therefore, is lower semicontinuous. In addition, v_α is nonnegative. The lower semicontinuity of $\mathbb{E}v_\alpha(x - D)$ now follows from the fact that $(x_n - D)$ converges in distribution to $(x - D)$ as $x_n \rightarrow x$; see Hernández-Lerma and Lasserre [19, p. 185]). To show that it is finite for each x , suppose that the policy ϕ always orders up or scraps down to zero. Then $v_\alpha(x) \leq v_\alpha^\phi(x) \leq K + \max\{c_+, c_-\}|x| + B_\alpha$, where $B_\alpha = \frac{\alpha(K + \max\{c_+, c_-\}\mathbb{E}|D|) + \mathbb{E}h(-D)}{1-\alpha}$. Thus, $\mathbb{E}v_\alpha(x - D) \leq K + \max\{c_+, c_-\}\mathbb{E}|x - D| + B_\alpha < \infty$.

Consider the second assertion. Chen and Simchi-Levi [6] have shown that $v_{n,\alpha}$ is (K, Q) -convex. Since $v_{n,\alpha} \uparrow v_\alpha$, v_α is (K, Q) -convex. Thus, the finiteness of $\mathbb{E}v_\alpha(x - D)$ and the first assertion of Lemma 5.2 yield that $\mathbb{E}v_\alpha(x - D)$ is (K, Q) -convex. Applying the second result in Lemma 5.2 (twice) yields that G_α^+ and G_α^- are (K, Q) -convex. Since $\mathbb{E}v_\alpha(x - D)$ is lower semicontinuous and $\mathbb{E}h(x - D)$ is convex, and therefore continuous, the lower semicontinuity of G_α^+ and G_α^- follows. \square

PROPOSITION 5.2 The following hold

(i) Under the assumption that there exist finite numbers u and y such that $u < y$ and

$$\frac{\mathbb{E}[h(y - D) - h(u - D)]}{y - u} < -c_+, \quad (26)$$

$G_{N,\alpha}^+(x) \rightarrow \infty$ and $G_\alpha^+(x) \rightarrow \infty$ as $|x| \rightarrow \infty$ for all $N \geq 0$. Similarly, if there exist finite numbers w and z such that $w < z$ and

$$\frac{\mathbb{E}[h(z-D) - h(w-D)]}{z-w} > c_-, \quad (27)$$

then $G_{N,\alpha}^-(x) \rightarrow \infty$ and $G_\alpha^-(x) \rightarrow \infty$ as $|x| \rightarrow \infty$ for all $N \geq 0$.

- (ii) There exists α^* such that $G_\alpha^+(x) \rightarrow \infty$ and $G_\alpha^-(x) \rightarrow \infty$ as $|x| \rightarrow \infty$ for all $\alpha \in [\alpha^*, 1)$. In addition, there exists $N \geq 0$ such that for $n \geq N$, $G_{n,\alpha}^+(x) \rightarrow \infty$ and $G_{n,\alpha}^-(x) \rightarrow \infty$ as $|x| \rightarrow \infty$ for all $\alpha \in [\alpha^*, 1)$.

PROOF. It is obvious that $G_\alpha^+(x) \rightarrow \infty$ as $x \rightarrow \infty$. To prove the first statement, we need only to show that it holds when $x \rightarrow -\infty$. We prove that $G_{N,\alpha}^+(x) \rightarrow \infty$ as $x \rightarrow -\infty$ for all N by induction. For $N = 0$ let x and y be such that (26) holds (with u replaced with x). Since $G_{0,\alpha}^+(x) = c_+x + \mathbb{E}h(x-D)$ is convex, $G_{0,\alpha}^+(x)$ is strictly decreasing for all $x \leq y$; the result holds for $N = 0$. Since $v_\alpha(x) \geq v_{N+1,\alpha}(x) \geq v_{N,\alpha}(x)$, we have that $G_\alpha^+(x) \geq G_{N,\alpha}^+(x) \geq G_{0,\alpha}^+(x) \rightarrow \infty$ as $x \rightarrow -\infty$.

Since $G_{0,\alpha}^+(x)$ may not tend to ∞ as $x \rightarrow -\infty$, when assumption (26) does not hold, additional arguments are needed to prove the second statement. To prove that statement, consider the model with $K = Q = 0$ and note that, since the discounted cost value function is non-decreasing in both K and Q , it suffices to prove the result for this case. Using a superscript zero to denote the case with zero fixed costs, let $G_{N,\alpha}^{0,+}(x)$ and $G_{N,\alpha}^{0,-}(x)$ be defined analogously to $G_\alpha^{0,+}(x)$ and $G_\alpha^{0,-}(x)$ with v_α replaced with $v_{N,\alpha}^0$. When $K = Q = 0$ the optimality equation for the finite-horizon problem becomes

$$v_{N,\alpha}^0(x) = \min\{\min_{b \geq x}\{-c_+x + G_{N-1,\alpha}^{0,+}(b)\}, \min_{b < x}\{c_-x + G_{N-1,\alpha}^{0,-}(b)\}\}. \quad (28)$$

Applying [21, Proposition B-4] iteratively implies that $v_{N,\alpha}^0(x)$, $G_{N,\alpha}^{0,+}$ and $G_{N,\alpha}^{0,-}$ are convex. Let $N \rightarrow \infty$ in (28) to obtain the infinite horizon optimality equations

$$v_\alpha^0(x) = \min\{\min_{b \geq x}\{-c_+x + G_\alpha^{0,+}(b)\}, \min_{b < x}\{c_-x + G_\alpha^{0,-}(b)\}\}.$$

Since Proposition 3.1(v) yields $v_{N,\alpha}^0 \uparrow v_\alpha^0$, we have that v_α^0 , $G_\alpha^{0,+}$ and $G_\alpha^{0,-}$ are also convex.

It is obvious that $G_\alpha^{0,+}(x) \rightarrow \infty$ as $x \rightarrow \infty$. We show that $G_\alpha^{0,+}(x) \rightarrow \infty$ as $x \rightarrow -\infty$ by contradiction. Since $G_\alpha^{0,+}(x)$ is convex, it has a finite or infinite limit as $x \rightarrow -\infty$. Suppose this limit is finite. Then the convex function $G_\alpha^{0,+}(x)$ is nondecreasing. Therefore, $G_\alpha^{0,+}(b)$ for $b \geq x$ is minimized at x and there is an optimal policy that never orders. Let ϕ denote such a policy.

Recall that $h(x)$ is nonnegative, convex, $h(0) = 0$, and $h(x) \rightarrow \infty$ as $|x| \rightarrow \infty$. Thus, for any x there exist finite constants γ and λ such that $h(x) \geq -\gamma x + \lambda$ and $\gamma > 0$. Suppose that there exists $\alpha^* > 0$ such that for all $\alpha \in [\alpha^*, 1)$ there exists $x^* < 0$ such that for $x \leq x^*$,

$$L_\alpha^0(x) = \mathbb{E}h(x-D) + \mathbb{E}v_\alpha^0(x-D) \geq \left(\frac{2c_+}{\gamma} + 1\right)h\left(\frac{x}{2} + \Delta\right), \quad (29)$$

where $\Delta \geq 0$ is finite. Straightforward calculations imply for $x \leq x^*$, $G_\alpha^{0,+}(x) = c_+x + L_\alpha^0(x) \geq -\frac{\gamma}{2}x + (\lambda - \gamma\Delta)\left(\frac{2c_+}{\gamma} + 1\right)$. Taking limits on both sides of this inequality yields that $G_\alpha^{0,+}(x) \rightarrow \infty$ as $x \rightarrow -\infty$. It remains to show that the postulated α^* exists.

Let $\tilde{D}_n = D_{n-1}$, $n \geq 1$, where $\{D_i, i \geq 0\}$ is the sequence of independent demands. Define $Z_n = -\sum_{i=1}^n \tilde{D}_i$, $n \geq 1$. Consider the stopping time (with respect to \tilde{D}_n and Z_n) $M(t) = \inf\{n \geq 1 : Z_n > t\}$, where $M(t)$ represents the first epoch by which the cumulative demand is smaller than $-t$. Recall that ϕ does not order. For $x < 0$

$$v_\alpha^0(x) = v_\alpha^{0,\phi}(x) \geq \mathbb{E} \sum_{n=1}^{M(\lfloor \frac{x}{2} \rfloor)-1} \alpha^n h(x + Z_n) \geq h\left(\frac{x}{2}\right) \mathbb{E} \sum_{n=1}^{M(\lfloor \frac{x}{2} \rfloor)-1} \alpha^n, \quad (30)$$

where the last inequality follows from the fact that $h(x)$ is decreasing on $(-\infty, 0)$. According to Wald's identity, if $\mathbb{E}M(\lfloor \frac{x}{2} \rfloor) < \infty$ then $\mathbb{E} \sum_{i=1}^{M(\lfloor \frac{x}{2} \rfloor)} \tilde{D}_i = \mathbb{E}D \cdot \mathbb{E}M(\lfloor \frac{x}{2} \rfloor)$. Moreover note that the definition of $M(\lfloor \frac{x}{2} \rfloor)$ implies that $-\mathbb{E} \sum_{i=1}^{M(\lfloor \frac{x}{2} \rfloor)} \tilde{D}_i \geq \lfloor \frac{x}{2} \rfloor$. Combining these results yields that $\mathbb{E}M(\lfloor \frac{x}{2} \rfloor) < \infty$ implies

$$- \mathbb{E} D \cdot \mathbb{E} M\left(\left|\frac{x}{2}\right|\right) \geq \left|\frac{x}{2}\right|. \quad (31)$$

Define $F(\alpha, x) = \mathbb{E} \sum_{n=1}^{M(|\frac{x}{2}|)-1} \alpha^n$. Then $F(\alpha, x) \uparrow M(|\frac{x}{2}|) - 1$ as $\alpha \uparrow 1$. Consider two cases.

Case 1: $\mathbb{E} D \geq 0$. In this case inequality (31) cannot hold and therefore $\mathbb{E} M(|\frac{x}{2}|) = \infty$. Thus, $F(\alpha, x) \uparrow \infty$ as $\alpha \uparrow 1$. Select $\alpha^* \in [0, 1)$ such that $\alpha F(\alpha, x) \geq (\frac{2c_+}{\gamma} + 1)$ for all $\alpha \in [\alpha^*, 1)$. Since $v_\alpha(x) \geq h(x - \mathbb{E} D) \rightarrow \infty$ as $|x| \rightarrow \infty$, choose x small enough for the convex function $v_\alpha^0(x)$ to be decreasing. Then for $\alpha \in [\alpha^*, 1)$

$$L_\alpha^0(x) \geq \alpha \mathbb{E} v_\alpha^0(x - D) \geq \alpha v_\alpha^0(x - \mathbb{E} D) \geq \alpha v_\alpha^0(x) \geq \alpha h\left(\frac{x}{2}\right) F(\alpha, x) \geq \left(\frac{2c_+}{\gamma} + 1\right) h\left(\frac{x}{2}\right),$$

where the second inequality follows from Jensen's inequality and the fourth inequality is due to (30). Thus, (29) holds with $\Delta = 0$.

Case 2: $\mathbb{E} D < 0$. Wald's identity and (31) imply that $\mathbb{E} M(|\frac{x}{2}|) \geq |\frac{x}{2\mathbb{E} D}|$. This coupled with the fact that $F(\alpha, x) \uparrow \mathbb{E} M(|\frac{x}{2}|)$ as $\alpha \uparrow 1$ implies that we can select $\alpha^* \in [0, 1)$ and $x^* < 0$ such that $\alpha F(\alpha, x) \geq (\frac{2c_+}{\gamma} + 1)$ when $\alpha \in [\alpha^*, 1)$ and $x \leq x^*$. For $\alpha \in [\alpha^*, 1)$ and $x \leq x^*$ we have

$$L_\alpha^0(x) \geq \alpha \mathbb{E} v_\alpha^0(x - D) \geq \alpha v_\alpha^0(x - \mathbb{E} D) \geq \alpha h\left(\frac{x - \mathbb{E} D}{2}\right) F(\alpha, x - \mathbb{E} D) \geq \left(\frac{2c_+}{\gamma} + 1\right) h\left(\frac{x}{2} - \frac{\mathbb{E} D}{2}\right),$$

where the last inequality follows since $F(\alpha, x)$ is non-decreasing in x . Thus, (29) holds with $\Delta = -\frac{\mathbb{E} D}{2}$.

Since $G_\alpha^+(x) \geq G_{\alpha^*}^{0,+}(x)$ and $G_\alpha^{0,+}(x) \rightarrow \infty$ as $x \rightarrow -\infty$, $G_\alpha^+(x) \rightarrow \infty$ as $x \rightarrow -\infty$. Now consider the functions $G_{n,\alpha}^+$. Since $G_{n+1,\alpha}^+(x) \geq G_{n,\alpha}^+(x)$ and $G_{n,\alpha}^+(x) \geq G_{n,\alpha}^{0,+}(x)$, it is sufficient to show that $G_{n,\alpha}^{0,+}(x) \rightarrow \infty$, as $|x| \rightarrow \infty$ for some n . If this is not true, the convex functions $G_{n,\alpha}^{0,+}(x)$ are monotone. Since $G_{n,\alpha}^{0,+}(x) \rightarrow G_{\alpha^*}^{0,+}(x)$, this would imply that $G_{\alpha^*}^{0,+}(x)$ is monotone. This contradicts $G_{\alpha^*}^{0,+}(x) \rightarrow \infty$ as $|x| \rightarrow \infty$. The statements for $G_{n,\alpha}^-(x)$ and $G_\alpha^-(x)$ are symmetric. \square

Recall that Proposition 5.2 provides sufficient conditions for $g(x) \rightarrow \infty$ as $|x| \rightarrow \infty$, where $g = G_\alpha^+$, G_α^- , $G_{n,\alpha}^+$, or $G_{n,\alpha}^-$. When such conditions hold, following Chen and Simchi-Levi [6], we define certain thresholds. For $n = 1, 2, \dots$, fix any $S_{n,\alpha} \in \operatorname{argmin}_{x \in \mathbb{R}} \{G_{n,\alpha}^+(x)\}$ and $U_{n,\alpha} \in \operatorname{argmin}_{x \in \mathbb{R}} \{G_{n,\alpha}^-(x)\}$. Let

$$s_{n,\alpha} = \min\{x \mid G_{n,\alpha}^+(x) \leq G_{n,\alpha}^+(S_{n,\alpha}) + K\}, \quad (32)$$

$$s'_{n,\alpha} = \min\{x \mid Q + G_{n,\alpha}^+(x) \leq G_{n,\alpha}^+(S_{n,\alpha}) + K\}, \quad (33)$$

$$u'_{n,\alpha} = \min\{x \mid Q + G_{n,\alpha}^-(x) \leq G_{n,\alpha}^-(U_{n,\alpha}) + K\}, \quad (34)$$

$$u_{n,\alpha} = \max\{x \mid G_{n,\alpha}^-(x) \leq G_{n,\alpha}^-(U_{n,\alpha}) + Q\}. \quad (35)$$

For the infinite horizon expected discounted costs define the vector $\mathbf{S}_\alpha = (s_\alpha, s'_\alpha, S_\alpha, U_\alpha, u'_\alpha, u_\alpha)$ similar to the vector $\mathbf{S}_{n,\alpha} = (s_{n,\alpha}, s'_{n,\alpha}, S_{n,\alpha}, U_{n,\alpha}, u'_{n,\alpha}, u_{n,\alpha})$ when the functions $G_{n,\alpha}^+$ and $G_{n,\alpha}^-$ are replaced respectively with the functions G_α^+ and G_α^- . By Lemma 3 in Chen and Simchi-Levi [6], $S_{n,\alpha} \leq U_{n,\alpha}$, $S_\alpha \leq U_\alpha$, $s'_{n,\alpha} \leq u'_{n,\alpha}$ and $s'_\alpha \leq u'_\alpha$. Since we have assumed $K \geq Q$, $u'_{n,\alpha} \leq U_{n,\alpha}$ and $u'_\alpha \leq U_\alpha$. Thus, $s_{n,\alpha} \leq s'_{n,\alpha} \leq S_{n,\alpha} \leq U_{n,\alpha} \leq u_{n,\alpha}$, $s'_{n,\alpha} \leq u'_{n,\alpha} \leq U_{n,\alpha} \leq u_{n,\alpha}$, $s_\alpha \leq s'_\alpha \leq S_\alpha \leq U_\alpha \leq u_\alpha$, and $s'_\alpha \leq u'_\alpha \leq U_\alpha \leq u_\alpha$.

Consider an N -horizon problem. The following theorem follows from Chen and Simchi-Levi [6], where the continuity of $G_n^+(x)$ and $G_n^-(x)$ and the property, that these functions tend to infinity as $|x|$ tends to infinity, were explicitly assumed. The continuity of $G_n^+(x)$ and $G_n^-(x)$ can be substituted by their lower semi-continuity established in Lemma 5.3 and the required properties that $G_n^+(x)$ and $G_n^-(x)$ tend to infinity were established in Proposition 5.2.

THEOREM 5.1 *Consider the N -horizon expected discounted cost criterion with $\alpha \in (0, 1]$. Suppose that either (26) and (27) hold for $\alpha \in (0, 1)$ or $\alpha \in [\alpha^*, 1)$, where α^* is as described in Proposition 5.2. For any $S_{n,\alpha} \in \operatorname{argmin}_{x \in \mathbb{R}} \{G_{n,\alpha}^+(x)\}$ and any $U_{n,\alpha} \in \operatorname{argmin}_{x \in \mathbb{R}} \{G_{n,\alpha}^-(x)\}$ there exists a Markov optimal policy $\{t'_N, t'_{N-1}, \dots, t'_1\}$ such that $v_{n+1,\alpha}(x) = \tilde{J}_n(x, t'_{n+1}(x))$, where $n = 0, 1, \dots, N-1$ and $x \in \mathbb{R}$, and the following properties hold. If $K > Q$ then*

$$t'_{n+1}(x) = \begin{cases} S_{n,\alpha} & \text{if } x < s_{n,\alpha}, \\ x & \text{if } x \in (s'_{n,\alpha}, u'_{n,\alpha}), \\ U_{n,\alpha} & \text{if } x > u_{n,\alpha}, \end{cases} \quad (36)$$

and

$$t'_{n+1}(x) \in \begin{cases} \{x, S_{n,\alpha}\} & \text{if } x \in [s_{n,\alpha}, s'_{n,\alpha}], \\ [s'_{n,\alpha}, x] & \text{if } x \in [u'_{n,\alpha}, u_{n,\alpha}]. \end{cases} \quad (37)$$

If $K = Q$ then

$$t'_{n+1}(x) = \begin{cases} S_{n,\alpha} & \text{if } x < s_{n,\alpha}, \\ x & \text{if } x \in (\frac{s_{n,\alpha} + S_{n,\alpha}}{2}, \frac{u_{n,\alpha} + U_{n,\alpha}}{2}), \\ U_{n,\alpha} & \text{if } x > u_{n,\alpha}, \end{cases} \quad (38)$$

and

$$t'_{n+1}(x) \in \begin{cases} \{x, S_{n,\alpha}\} & \text{if } x \in [s_{n,\alpha}, \frac{s_{n,\alpha} + S_{n,\alpha}}{2}], \\ \{x, U_{n,\alpha}\} & \text{if } x \in [\frac{u_{n,\alpha} + U_{n,\alpha}}{2}, u_{n,\alpha}]. \end{cases} \quad (39)$$

Let N be as described in Proposition 5.2(ii). We observe that there are finite constants A_α and B_α such that $A_\alpha \leq s_{n,\alpha} \leq u_{n,\alpha} \leq B_\alpha$ when $n \geq N$. Indeed, let $A_\alpha = \min\{x \mid G_{N,\alpha}^+(x) \leq G_\alpha^+(S_\alpha) + K\}$. Since for $n \geq N$, $G_{N,\alpha}^+(x) \leq G_{n,\alpha}^+(x) \leq G_\alpha^+(x)$, we have $A_\alpha \leq s_{n,\alpha} \leq s_\alpha$. The arguments for B_α are symmetric. Therefore, all the limit points of the vector $\mathbf{S}_{n,\alpha}$ are finite.

THEOREM 5.2 *Consider the infinite horizon discounted cost criterion. Suppose that either (26) and (27) hold for $\alpha \in (0, 1)$ or $\alpha \in [\alpha^*, 1)$, where α^* is as described in Proposition 5.2. Then for any $S_\alpha \in \operatorname{argmin}_{x \in \mathbb{R}} \{G_\alpha^+(x)\}$ and any $U_\alpha \in \operatorname{argmin}_{x \in \mathbb{R}} \{G_\alpha^-(x)\}$ there exists a stationary optimal policy t'_α satisfying (36) and (37) when $K > Q$ and satisfying (38) and (39) when $K = Q$ with $\mathbf{S}_{n,\alpha}$ replaced by \mathbf{S}_α .*

Let $\tilde{\mathbf{S}}_\alpha = (\tilde{s}_\alpha, \tilde{s}'_\alpha, \tilde{S}_\alpha, \tilde{S}'_\alpha, \tilde{U}_\alpha, \tilde{u}'_\alpha, \tilde{u}_\alpha)$ be a limit point of the sequence $\{\mathbf{S}_{n,\alpha}, n \geq 1\}$. Then there exists a stationary optimal policy t'_α that satisfies (36) and (37), when $K > Q$, and it satisfies (38) and (39), when $K = Q$, with $\mathbf{S}_{n,\alpha}$ replaced by $\tilde{\mathbf{S}}_\alpha$.

PROOF. For \mathbf{S}_α , the proofs are identical to Theorem 5.1 because G_α^+ and G_α^- are inf-compact (K, Q) -convex functions. For the vector of thresholds $\tilde{\mathbf{S}}_\alpha$ consider a sequence m_k such that $\mathbf{S}_{m_k, \alpha} \rightarrow \tilde{\mathbf{S}}_\alpha$. In addition, we select m_k in a way that $t'_{m_k}(x)$ converges as $k \rightarrow \infty$ when $x \in \{\tilde{s}_\alpha, \tilde{s}'_\alpha, \tilde{u}'_\alpha, \tilde{u}_\alpha\}$ for the case $K > Q$ and when $x \in \{\tilde{s}_\alpha, \frac{(\tilde{s}_\alpha + \tilde{S}_\alpha)}{2}, \frac{(\tilde{u}_\alpha + \tilde{U}_\alpha)}{2}, \tilde{u}_\alpha\}$ for the case $K = Q$. For $x \in \mathbb{R}$, let $t_\alpha(x) = \lim_{k \rightarrow \infty} t'_{m_k}(x)$. Then $t_\alpha(x)$ satisfies (36)-(37) and (38)-(39) in each of the corresponding cases $K > Q$ and $K = Q$ with the respective right-hand sides. In addition, t_α is a stationary optimal policy; see the proof of Proposition 9.17 in Bertsekas and Shreve [3] or Bertsekas [2, page 157]. \square

THEOREM 5.3 *Consider the average cost criterion. For each $\alpha \in (0, 1)$ consider $S_\alpha \in \operatorname{argmin}_{x \in \mathbb{R}} \{G_\alpha^+(x)\}$ and $U_\alpha \in \operatorname{argmin}_{x \in \mathbb{R}} \{G_\alpha^-(x)\}$. Then there exists a sequence of discount factors $\{\alpha_k, k \geq 0\}$ converging to 1 such that there exists a stationary optimal policy t' satisfying (36) and (37), when $K > Q$, and satisfying (38) and (39), when $K = Q$, with $\mathbf{S}_{n,\alpha}$ replaced by $\mathbf{S}_1 = (s_1, s'_1, S_1, U_1, u'_1, u_1)$, where \mathbf{S}_1 is the limit of $\{\mathbf{S}_{\alpha_k}, k \geq 0\}$ as $\alpha_k \rightarrow 1$, and some of the coordinates of \mathbf{S}_1 are allowed to be equal to $+\infty$ or $-\infty$.*

PROOF. Similar to the proof of Theorem 5.2, consider the sequence $\alpha_k \rightarrow 1$ such that $\mathbf{S}_{\alpha_k} \rightarrow \mathbf{S}_1$. In addition, select α_k in a way that $t'_{\alpha_k}(x)$ has a (possibly infinite) limit as $k \rightarrow \infty$ when $x \in \{s_1, s'_1, u'_1, u_1\}$ for the case $K > Q$ and when $x \in \{s_1, \frac{(s_1 + S_1)}{2}, \frac{(u_1 + U_1)}{2}, u_1\}$ for the case $K = Q$. As per the comments following Theorem 5.1 we note that there is an ambiguity about which action to choose at these points. This ambiguity is covered by the possibilities in (37) and (39). Let $t'(x) = \lim_{k \rightarrow \infty} t'_{\alpha_k}(x)$. Then $t'(x)$ satisfies (36)-(37) ((38)-(39)) when $K > Q$ ($K = Q$). In addition, t' is a stationary optimal policy in view of Theorem 3.1. \square

REMARK 5.1 *For the stochastic cash balance problem, we have considered an MDP with $\mathbb{X} = A(x) = \mathbb{R}$ for each $x \in \mathbb{X}$. However, if the demand takes only integer values, for many problems it is natural to consider $\mathbb{X} = A(x) = \mathbb{Z}$. Therefore, if the demand is integer, we have two MDPs for the inventory control*

problems: an MDP with $\mathbb{X} = \mathbb{R}$ and an MDP with $\mathbb{X} = \mathbb{Z}$. Though the first MDP yields potentially lower costs, its implementation may not be reasonable for some applications because it may prescribe to order up to a non-integer inventory level. However, all of the results of this paper hold for the second representation, when the state space is integer, with a minor modification that the action sets are integer as well. In addition, in this case, for the optimal policies of the form described in Theorems 5.1, 5.2, and 5.3 it is possible to fix a version whether we order at s ($s_{n,\alpha}$) up to the level S ($S_{n,\alpha}$) or do not order at s ($s_{n,\alpha}$). Similarly for u ($u_{n,\alpha}$) and U ($U_{n,\alpha}$). One more note is that the case when the possible demand is proportional to some number $d \in \mathbb{R}$ is similar to the integer demand case. One can consider an MDP with an integer state space for this case as well.

REMARK 5.2 *The definitions of optimal finite-horizon policies in [6] are slightly different than in (36)-(39). The difference is that the intervals in (36) and (38) are closed and the intervals in (37) and (39) are open in [6]. The slightly stronger formulation in [6] is probably correct but it is based on the explicit assumption in [6] that the functions $G_{n,\alpha}^+$ and $G_{n,\alpha}^-$ are continuous. The validity of the following properties are open questions for the cash balance problem: (i) the continuity of the functions $G_{n,\alpha}^+$, $G_{n,\alpha}^-$, G_α^+ , G_α^- , and W , (ii) the correctness of $c_+x + W(x) \rightarrow \infty$ and $-c_-x + W(x) \rightarrow \infty$ as $|x| \rightarrow \infty$, (iii) the correctness of $s_1 > -\infty$ and $u_1 < \infty$, (iv) the characterization of optimal average cost policies via thresholds defined through the function W , (v) the validity of Theorems 5.1, 5.2, and 5.3 with closed intervals in (36) and (38) and open intervals in (37) and (39), and (vi) the validity of the optimality equality for the average costs criterion.*

Appendix A. Proof of Lemma 4.1 Without loss of generality set $y^0 = 0$. Then $\eta^0 = \xi$ and η^k has the density $f(x - y^k)$. Let P^k be the probability distribution of η^k on $(-\infty, \infty)$. According to Shiryaev [29, page 362],

$$\|P^k - P^0\| = \int_{-\infty}^{\infty} |f(x) - f(x - y^k)| dx. \quad (\text{A.40})$$

Fix $\epsilon > 0$ and consider $K > 0$ such that $\int_{-K}^K f(x) dx \geq 1 - \frac{\epsilon}{8}$. Consider integers k large enough for $|y^k| \leq 1$ and select any number $K^* \geq K + 1$. Then

$$0 \leq \int_{-\infty}^{\infty} |f(x) - f(x - y^k)| dx - \int_{-K^*}^{K^*} |f(x) - f(x - y^k)| dx \leq \frac{\epsilon}{4}. \quad (\text{A.41})$$

Fix $G > 0$ such that

$$\int_{-\infty}^{\infty} f(x) I\{f(x) > G\} dx \leq \frac{\epsilon}{8} \quad (\text{A.42})$$

and define the bounded function $f_G(x) = f(x) I\{f(x) \leq G\} + GI\{f(x) > G\}$. Then, in view of (A.42),

$$\int_{-K^*}^{K^*} |f(x) - f(x - y^k)| dx \leq \int_{-K^*}^{K^*} |f_G(x) - f_G(x - y^k)| dx + \frac{\epsilon}{4}. \quad (\text{A.43})$$

From (A.40), (A.41), and (A.43), we have that

$$\|P^k - P^0\| \leq \int_{-K^*}^{K^*} |f_G(x) - f_G(x - y^k)| dx + \frac{\epsilon}{2}. \quad (\text{A.44})$$

Lusin's theorem [24, page 108] state that for any real-valued measurable function f and for any $\epsilon > 0$ there exists a continuous function g such that the Lebesgue measure of the set $\{f(x) \neq g(x)\}$ is not greater than ϵ . Consider the density function f and select a continuous function g such that Lusin's theorem holds for $\epsilon = \frac{\epsilon}{(8G)}$.

Define the non-negative bounded continuous function $g_G(x) = g(x)I\{0 \leq g(x) \leq G\} + GI\{g(x) > G\}$. Since $\{f_G(x) \neq g_G(x)\} \subseteq \{f(x) \neq g(x)\}$,

$$\begin{aligned} \int_{-K^*}^{K^*} |f_G(x) - f_G(x - y^k)| dx &\leq \int_{-K^*}^{K^*} |g_G(x) - g_G(x - y^k)| dx + \int_{-K^*}^{K^*} |g_G(x) - f_G(x)| dx \\ &+ \int_{-K^*}^{K^*} |g_G(x - y^k) - f_G(x - y^k)| dx \leq \int_{-K^*}^{K^*} |g_G(x) - g_G(x - y^k)| dx + \frac{\epsilon}{4}. \end{aligned} \quad (\text{A.45})$$

Since the function g_G is continuous, it is uniformly continuous on the interval $[-(K^* + 1), (K^* + 1)]$. Thus, there exists an integer N such that $|g_G(x) - g_G(x - y^k)| \leq \frac{\epsilon}{8K^*}$ when $k \geq N$ and $-K^* \leq x \leq K^*$.

This bound and (A.45) imply that $\int_{-K^*}^{K^*} |f_G(x) - f_G(x - y^k)| dx \leq \frac{\epsilon}{2}$ for $k \geq N$. The last inequality and (A.44) imply that $\|P^k - P^0\| \leq \epsilon$ when $k \geq N$. \square

Acknowledgements. The authors thank Irwin Schocetman for providing several references on upper semi-continuity of a set-valued function. Research of the first author was partially supported by NSF grants DMI-0300121 and DMI-0600538. The work of the second author is partially supported by NSF grant CMMI-0540808.

References

- [1] Dimitri P. Bertsekas, *Dynamic programming and optimal control*, second ed., vol. 1, Athena Scientific, Belmont, MA, 2000.
- [2] _____, *Dynamic programming and optimal control*, second ed., vol. 2, Athena Scientific, Belmont, MA, 2001.
- [3] Dimitri P. Bertsekas and Steven E. Shreve, *Stochastic optimal control: The discrete-time case*, Athena Scientific, Belmont, MA, 1996.
- [4] Rolando Cavazos-Cadena, *A counterexample on the optimality equation in Markov decision chains with the average cost criterion*, *Systems & Control Letters* **16** (1991), 387–392.
- [5] Rolando Cavazos-Cadena and Linn I. Sennott, *Comparing recent assumptions for the existence of average optimal stationary policies*, *Operations Research Letters* **11** (1992), 33–37.
- [6] Xin Chen and David Simchi-Levi, *A new approach for the stochastic cash balance problem with fixed costs*, Preprint, Operations Research Center, Massachusetts Institute of Technology, 2003, <https://netfiles.uiuc.edu/xinchen/www/papers/CBP7.pdf>.
- [7] _____, *Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The infinite horizon case*, *Math. Oper. Res.* **29** (2004), 698–723.
- [8] George M. Constantinides and Scott F. Richard, *Existence of optimal simple policies for discounted cost inventory and cash management in continuous time*, *Operations Research* **26** (1978), no. 4, 620–636.
- [9] Edwin J. Elton and Martin J. Gruber, *On the cash balance problem*, *Operational Research Quarterly* **25** (1974), no. 4, 553–572.
- [10] Gary D. Eppen and Eugene F. Fama, *Cash balance and simple dynamic portfolio problems with proportional costs*, *International Economic Review* **10** (1969), no. 2, 119–133.
- [11] Eugene A. Feinberg, *Total reward criteria*, *Handbook of Markov Decision Processes* (Eugene A. Feinberg and Adam Shwartz, eds.), Kluwer, Boston, 2002, pp. 173–207.
- [12] Eugene A. Feinberg and Mark E. Lewis, *Optimality of four-threshold policies in inventory systems with customer returns and borrowing/storage options*, *Probability in the Engineering and Informational Sciences* **19** (2005), no. 1, 45–71.
- [13] _____, *Optimality inequalities for average cost Markov decision processes and the optimality of (s,S) policies*, Technical Report TR1442, Cornell University, 2006, <http://legacy.orie.cornell.edu/techreports/TR1442.pdf>.

- [14] Emmanuel Fernández-Gaucherand, Aristotle Arapostathis, and Steven I. Marcus, *Convex stochastic control problems*, Proceedings of the 31st Conference on Decision and Control, IEEE, December 1992, pp. 2179–2180.
- [15] Moritz Fleischmann, Roelof Kuik, and Rommert Dekker, *Controlling inventories with stochastic item retruns: A basic model*, European Journal of Operational Research **138** (2002), 63–75.
- [16] Herbert S. Gaskill and P.P. Narayanaswami, *Elements of real analysis*, Prentice-Hall, Upper Saddle River, New Jersey, 1998.
- [17] Nadia M. Girgis, *Optimal cash balance levels*, Management Science **15** (1968), no. 3, 130–140.
- [18] Onésimo Hernández-Lerma, *Average optimality in dynamic programming on Borel spaces*, Systems and Control Letters **17** (1991), no. 3, 237–242.
- [19] Onésimo Hernández-Lerma and Jean Bernard Lasserre, *Discrete-time Markov control processes: Basic optimality criteria*, Springer, New York, NY, 1996.
- [20] Daniel P. Heyman, *Optimal disposal policies for a single-item inventory system with returns*, Naval Research Logistics Quarterly **24** (1977), 385–405.
- [21] Daniel P. Heyman and Matthew J. Sobel, *Stochastic models in operations research*, vol. II, McGraw-Hill, New York, NY, 1984.
- [22] Werner Hildenbrand, *Core and equilibria of a large economy*, Princeton Studies in Mathematical Economics, Princeton University Press, Princeton, 1974.
- [23] Karl Hinderer and Karl-Heinz Waldmann, *Cash management in a randomly varying environment*, European Journal of Operational Research **130** (2001), 468–485.
- [24] Alexander S. Kechris, *Classical descriptive set theory*, Springer-Verlag, New York, 1994.
- [25] Edwin H. Neave, *The stochastic cash balance problem with fixed costs for increases and decreases*, Management Science **16** (1970), no. 7, 472–490.
- [26] Hukukane Nikaido, *Convex structures and economic theory*, Mathematics in Science and Engineering, vol. 51, Academic Press, New York, 1968.
- [27] Manfred Schäl, *Average optimality in dynamic programming with general state space*, Mathematics of Operations Research **18** (1993), no. 1, 163–172.
- [28] Linn I. Sennott, *Stochastic dynamic programming and the control of queueing systems*, Wiley Series in Probability and Statistics, John Wiley and Sons, Inc., New York, 1999.
- [29] Albert N. Shiryaev, *Probability. second edition*, Springer, New York, 1996.
- [30] Erwin van der Laan and Marc Salomon, *Production planning and inventory control with remanufacturing and disposal*, European Journal of Operational Research **102** (1997), 264–278.
- [31] Qing Ye and Izak Duenyas, *Optimal capacity investment decisions with two-sided fixed capacity adjustment costs*, Operations Research **55** (2007), no. 2, 272–283.