Ann Oper Res
DOI 10.1007/s10479-017-2561-9

On the optimality equation for average cost Markov decision processes and its validity for inventory control

Eugene A. Feinberg[1] · **Yan Liang**[1]

Abstract As is well known, average-cost optimality inequalities imply the existence of stationary optimal policies for Markov decision processes with average costs per unit time, and these inequalities hold under broad natural conditions. This paper provides sufficient conditions for the validity of the average-cost optimality equation for an infinite state problem with weakly continuous transition probabilities and with possibly unbounded one-step costs and noncompact action sets. These conditions also imply the convergence of sequences of discounted relative value functions to average-cost relative value functions and the continuity of average-cost relative value functions. As shown in this paper, the classic periodic-review setup-cost inventory control problem with backorders and convex holding/backlog costs satisfies these conditions. Therefore, the optimality inequality holds in the form of an equality with a continuous average-cost relative value function for this problem. In addition, the K-convexity of discounted relative value functions and their convergence to average-cost relative value functions, when the discount factor increases to 1, imply the K-convexity of average-cost relative value functions. This implies that average-cost optimal (s, S) policies for the inventory control problem can be derived from the average-cost optimality equation.

Keywords Dynamic programming · Average-cost optimal equation · Inventory control · (s, S) policies

1 Introduction

For Markov decision processes (MDPs) with average costs per unit time, the existence of stationary optimal policies follows from the validity of the average-cost optimality inequal-

✉ Eugene A. Feinberg
eugene.feinberg@stonybrook.edu

Yan Liang
yan.liang@stonybrook.edu

[1] Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11794, USA

Published online: 22 June 2017

ity (ACOI). Feinberg et al. (2012) established broad sufficient conditions for the validity of ACOIs for MDPs with weakly continuous transition probabilities and possibly noncompact action sets and unbounded one-step costs. In particular, these and even stronger conditions hold for the classic periodic-review inventory control problem with backorders; see Feinberg (2016) or Feinberg and Lewis (2015). Previously, Schäl (1993) established sufficient conditions for the validity of ACOIs for MDPs with compact action sets and possibly unbounded costs. Cavazos-Cadena (1991) provided an example in which the ACOI holds but the average-cost optimality equation (ACOE) does not. This paper presents sufficient conditions for the validity of ACOEs for MDPs with infinite state spaces, weakly continuous transition probabilities and possibly noncompact action sets and unbounded one-step costs and, by showing that the setup-cost inventory control problems with backorders and convex holding/backlog costs satisfy these conditions, establishes the validity of the ACOEs for the inventory control problems.

Sufficient conditions for the validity of ACOEs for discrete-time MDPs with countable and general state spaces with setwise continuous transition probabilities are described in Sennott (1998, Section 7.4; 2002) and Hernández-Lerma and Lasserre (1996, Section 5.5), respectively. Jaśkiewicz and Nowak (2006) considered MDPs with Borel state spaces, compact action sets, weakly continuous transition probabilities and unbounded costs. The geometric ergodicity of transition probabilities is assumed in Jaśkiewicz and Nowak (2006) to ensure the validity of the ACOEs. Costa and Dufour (2012) studied the validity of ACOEs for MDPs with Borel state and action spaces, weakly continuous transition probabilities, which are positive Harris recurrent, and with possibly noncompact action sets and unbounded costs. Neither the geometric ergodicity nor positive Harris recurrent conditions hold for the periodic-review inventory control problems.

As is mentioned above, Hernández-Lerma and Lasserre (1996, Section 5.5) described sufficient conditions for the validity of ACOEs for MDPs with setwise continuous transition probabilities. These condition are based on the equicontinuity property of the value functions for discounted criteria. An attempt, to establish such results for MDPs with weakly continuous transition probabilities, was undertaken in Montes-de-Oca (1994). However, the formulations and proofs in Montes-de-Oca (1994), as well as some proofs in Costa and Dufour (2012), relied on a technically incorrect paper with statements contradicting the counterexample in Luque-Vasques and Hernández-Lerma (1995) relevant to Berge's maximum theorem.

For average-cost periodic-review inventory control problems, Iglehart (1963) and Veinott and Wagner (1965) proved the optimality of (s, S) policies for problems with setup costs and backorders, when demand distributions are either discrete or continuous. Beyer and Sethi (1999) corrected some gaps in these papers. Zheng (1991) proved the optimality of (s, S) policies for problems with discrete demands by constructing a solution to the ACOE. Chen and Simchi-Levi (2004a, b) investigated problems with general demand distributions when prices may depend on current inventory levels. Beyer et al. (2010) studied problems with the demand depending on a state of a Markov chain. All these references used versions of the assumption that the backorder cost is higher than the cost of backordered inventory, if the amount of backordered inventory tends to infinity; see e.g., the assumption in Feinberg and Lewis (2015, Lemma 6.11). Under this assumption, (s, S) policies optimize expected total discounted costs for all values of discount factors. However, without this assumption, (s, S) policies may not be optimal for small values of discount factors, but they are optimal for large values of discount factors; see Feinberg and Liang (2017a) for details. Feinberg and Lewis (2015) proved that (s, S) policies are always optimal for average costs per unit time even if such policies are not optimal for some values of discount factors. However, only the validity of ACOIs was established in Feinberg and Lewis (2015).

Section 2 of this paper describes the general MDPs framework. In particular, it states Assumptions **W*** and **B** from Feinberg et al. (2012), which guarantee the validity of the ACOIs. Section 3 provides the sufficient conditions for the validity of the ACOEs, which extends the sufficient conditions in Hernández-Lerma and Lasserre (1996, Theorem 5.5.4) to weakly continuous transition probabilities. By verifying these conditions, it is shown in Sect. 4, that the ACOE holds for the classic periodic-review inventory control problems with backorders, setup costs, and general demands. It is also shown that optimal average-cost relative value function is the limit inferior of the discounted relative value functions. The paper also establishes K-convexity and continuity of the average-cost relative value function and shows that an optimal (s, S) policy can be derived from the ACOE. It also shows that, if the set of all possible inventory levels is the set of all real numbers $\mathbb{R}$ and any nonnegative amount of inventory can be ordered, then at the appropriately defined level s there are at least two optimal decisions: do not order and order up to the level S.

As shown in our follow-up paper (Feinberg and Liang 2017b), the equicontinuity property holds for the inventory model with more general holding/backlog costs, when the convexity assumption for holding/backlog costs is relaxed to the appropriate quasiconvexity assumption. In addition, there are convergence of discounted relative value functions to average-cost relative value function and convergence of optimal lower thresholds s_α for discounted costs to the the optimal lower threshold s for average costs per unit time as the discount factor α tends to 1. Thus, optimal discounted thresholds with large discount factors approximate optimal average-cost thresholds.

2 Model definition

Consider a discrete-time MDP with a state space $\mathbb{X}$, an action space $\mathbb{A}$, one-step costs c, and transition probabilities q. Assume that $\mathbb{X}$ and $\mathbb{A}$ are Borel subsets of Polish (complete separable metric) spaces.

Let $c(x, a) : \mathbb{X} \times \mathbb{A} \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ be the one-step cost and $q(B|x, a)$ be the transition kernel representing the probability that the next state is in $B \in \mathcal{B}(\mathbb{X})$, given that the action a is chosen at the state x.

We recall that a function $f : \mathbb{U} \to \mathbb{R} \cup \{\infty\}$ defined on a metric space $\mathbb{U}$ is called inf-compact (on $\mathbb{U}$), if for every $\lambda \in \mathbb{R}$ the level set $\{u \in \mathbb{U} : f(u) \le \lambda\}$ is compact. A subset of a metric space is also a metric space with the same metric. For $U \subset \mathbb{U}$, if the domain of f is narrowed to U, then this function is called the restriction of f to U.

Definition 2.1 *(Feinberg et al. 2013, Definition 1.1; Feinberg 2016, Definition 2.1)* A function $f : \mathbb{X} \times \mathbb{A} \to \overline{\mathbb{R}}$ is called $\mathbb{K}$-inf-compact, if for every nonempty compact subset K of $\mathbb{X}$ the restriction f to $K \times \mathbb{A}$ is an inf-compact function.

Let the one-step cost function c and transition probability q satisfy the following condition.

Assumption W* (Feinberg et al. 2012, 2016; Feinberg and Lewis 2015, or Feinberg 2016).

(i) c is $\mathbb{K}$-inf-compact and bounded below, and
(ii) the transition probability $q(\cdot|x, a)$ is weakly continuous in $(x, a) \in \mathbb{X} \times \mathbb{A}$, that is, for every bounded continuous function $f : \mathbb{X} \to \mathbb{R}$, the function $\tilde{f}(x, a) := \int_{\mathbb{X}} f(y)q(dy|x, a)$ is continuous on $\mathbb{X} \times \mathbb{A}$.

The decision process proceeds as follows: at each time epoch $t = 0, 1, \ldots,$ 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 $H_t = (\mathbb{X}\times\mathbb{A})^t \times \mathbb{X}$ be the set of histories for $t = 0, 1, \ldots$. A (randomized) decision rule at period $t = 0, 1, \ldots$ is a regular transition probability $\pi_t : H_t \to \mathbb{A}$, that is, (i) $\pi_t(\cdot|h_t)$ is a probability distribution on $\mathbb{A}$, where $h_t = (x_0, a_0, x_1, \ldots, a_{t-1}, x_t)$, and (ii) for any measurable subset $B \subset \mathbb{A}$, the function $\pi_t(B|\cdot)$ is measurable on H_t. A policy π is a sequence $(\pi_0, \pi_1, \ldots)$ of decision rules. Let Π be the set of all policies. A policy π is called non-randomized if each probability measure $\pi_t(\cdot|h_t)$ is concentrated at one point. A non-randomized policy is called stationary if all decisions depend only on the current state.

The Ionescu Tulcea theorem implies that an initial state x and a policy π define a unique probability P_x^π on the set of all trajectories $\mathbb{H}_\infty = (\mathbb{X} \times \mathbb{A})^\infty$ endowed with the product of σ-field defined by Borel σ-field of $\mathbb{X}$ and $\mathbb{A}$; see Bertsekas and Shreve (1996, pp. 140–141) or Hernández-Lerma and Lasserre (1996, p. 178). Let $\mathbb{E}_x^\pi$ be an expectation with respect to P_x^π.

For a finite-horizon $N = 0, 1, \ldots,$ let us define the expected total discounted costs,

$$v_{N,\alpha}^\pi := \mathbb{E}_x^\pi \sum_{t=0}^{N-1} \alpha^t c(x_t, a_t), \quad x \in \mathbb{X}, \tag{2.1}$$

where $\alpha \in [0, 1)$ is the discount factor and $v_{0,a}^\pi(x) = 0$. When $N = \infty$, Eq. (2.1) defines an infinite-horizon expected total discounted cost denoted by $v_\alpha^\pi(x)$. Let $v_\alpha := \inf_{\pi\in\Pi} v_\alpha^\pi(x)$, $x \in \mathbb{X}$. A policy π is called optimal for the discount factor α if $v_\alpha^\pi(x) = v_\alpha(x)$ for all $x \in \mathbb{X}$.

The *average cost per unit time* is defined as

$$w^\pi(x) := \limsup_{N\to\infty} \frac{1}{N} v_{N,1}^\pi(x), \quad x \in \mathbb{X}. \tag{2.2}$$

Define the optimal value function $w(x) := \inf_{\pi\in\Pi} w^\pi(x)$, $x \in \mathbb{X}$. A policy π is called average-cost optimal if $w^\pi(x) = w(x)$ for all $x \in \mathbb{X}$.

Let

$$\begin{aligned} m_\alpha &:= \inf_{x\in\mathbb{X}} v_\alpha(x), \quad u_\alpha(x) := v_\alpha(x) - m_\alpha, \\ \underline{w} &:= \liminf_{\alpha\uparrow 1}(1-\alpha)m_\alpha, \quad \bar{w} := \limsup_{\alpha\uparrow 1}(1-\alpha)m_\alpha \end{aligned} \tag{2.3}$$

The function u_α is called the discounted relative value function. Assume that the following assumption holds in addition to Assumption **W***.

Assumption B

(i) $w^* := \inf_{x\in\mathbb{X}} w(x) < \infty$, and (ii) $\sup_{\alpha\in[0,1)} u_\alpha(x) < \infty$, $x \in \mathbb{X}$.

As follows from Schäl (1993, Lemma 1.2(a)), Assumption **B**(i) implies that $m_\alpha < \infty$ for all $\alpha \in [0, 1)$. Thus, all the quantities in (2.3) are defined. According to Feinberg et al. (2012, Theorem 4), if Assumptions **W*** and **B** hold, then $\underline{w} = \bar{w}$. In addition, for each sequence $\{\alpha_n\}_{n=1,2,\ldots}$ such that $\alpha_n \uparrow 1$ as $n \to \infty$,

$$\lim_{n\to\infty} (1-\alpha_n)m_{\alpha_n} = \underline{w} = \bar{w}. \tag{2.4}$$

Define the following function on $\mathbb{X}$ for the sequence $\{\alpha_n \uparrow 1\}_{n=1,2,\ldots}$:

$$\tilde{u}(x) := \liminf_{n\to\infty, y\to x} u_{\alpha_n}(y). \tag{2.5}$$

In words, $\tilde{u}(x)$ is the largest number such that $\tilde{u}(x) \leq \liminf_{n\to\infty} u_{\alpha_n}(y_n)$ for all sequences $\{y_n \to x\}$. Since $u_\alpha(x)$ is nonnegative by definition, then $\tilde{u}(x)$ is also nonnegative. The function $\tilde{u}$, defined in (2.5) for a sequence $\{\alpha_n \uparrow 1\}_{n=1,2,\ldots}$ of nonnegative discount factors, is called an average-cost relative value function.

3 Average cost optimality equation

If Assumptions **W*** and **B** hold, then, according to Feinberg et al. (2012, Corollary 2), there exists a stationary policy ϕ satisfying

$$\underline{w} + \tilde{u}(x) \geq c(x, \phi(x)) + \int_{\mathbb{X}} \tilde{u}(y) q(dy|x, \phi(x)), \qquad x \in \mathbb{X}, \tag{3.1}$$

with $\tilde{u}$ defined in (2.5) for an arbitrary sequence $\{\alpha_n \uparrow 1\}_{n=1,2,\ldots}$, and

$$w^\phi(x) = \underline{w} = \lim_{\alpha\uparrow 1}(1-\alpha)v_\alpha(x) = \bar{w} = w^*, \qquad x \in \mathbb{X}. \tag{3.2}$$

These equalities imply that the stationary policy ϕ is average-cost optimal and $w^\phi(x)$ does not depend on x.

Inequality (3.1) is known as the ACOI. We remark that a weaker form of the ACOI with $\underline{w}$ substituted with $\bar{w}$ is also described in Feinberg et al. (2012). If Assumptions **W*** and **B** hold, let us define $w := \underline{w}$; see (3.2) for other equalities for w.

Recall the following definition.

Definition 3.1 *(Hernández-Lerma and Lasserre 1996, Remark 5.5.2)* A family $\mathcal{H}$ of real-valued functions on a metric space X is called equicontinuous at the point $x \in X$ if for each $\epsilon > 0$ there exists an open set G containing x such that

$$|h(y) - h(x)| < \epsilon \quad \text{for all } y \in G \text{ and for all } h \in \mathcal{H}.$$

The family $\mathcal{H}$ is called equicontinuous (on X) if it is equicontinuous at all $x \in X$.

Consider the following equicontinuity condition (EC) on the discounted relative value functions.

Assumption EC There exists a sequence $\{\alpha_n \uparrow 1\}_{n=1,2,\ldots}$ of nonnegative discount factors such that

(i) the family of functions $\{u_{\alpha_n}\}_{n=1,2,\ldots}$ is equicontinuous, and
(ii) there exists a nonnegative measurable function $U(x)$, $x \in \mathbb{X}$, such that $U(x) \geq u_{\alpha_n}(x)$, $n = 1, 2, \ldots$, and $\int_{\mathbb{X}} U(y) q(dy|x, a) < \infty$ for all $x \in \mathbb{X}$ and $a \in \mathbb{A}$.

The following theorem states that Assumption **EC** implies that there exist a stationary policy ϕ and a function $\tilde{u}(\cdot)$ satisfying the ACOE. This theorem is similar to Theorem 5.5.4 in Hernández-Lerma and Lasserre (1996), where MDPs with setwise continuous transition probabilities are considered.

Theorem 3.2 *Let Assumptions* ***W**** *and* ***B*** *hold. Consider a sequence $\{\alpha_n \uparrow 1\}_{n=1,2,\ldots}$ of nonnegative discount factors. If Assumption* ***EC*** *is satisfied for the sequence $\{\alpha_n\}_{n=1,2,\ldots}$, then the following statements hold.*

(i) There exists a subsequence $\{\alpha_{n_k}\}_{k=1,2,\ldots}$ of $\{\alpha_n\}_{n=1,2,\ldots}$ such that $\{u_{\alpha_{n_k}}(x)\}$ converges pointwise to $\tilde{u}(x)$, $x \in \mathbb{X}$, where $\tilde{u}(x)$ is defined in (2.5) for the subsequence $\{\alpha_{n_k}\}_{k=1,2,\ldots}$, and the convergence is uniform on each compact subset of $\mathbb{X}$. In addition, the function $\tilde{u}(x)$ is continuous.

(ii) *There exists a stationary policy* ϕ *satisfying the ACOE with the nonnegative function* $\tilde{u}$ *defined for the sequence* $\{\alpha_{n_k}\}_{k=1,2,\ldots}$ *mentioned in statement (i), that is, for all* $x \in \mathbb{X}$,

$$w+\tilde{u}(x)=c(x,\phi(x))+\int_{\mathbb{X}}\tilde{u}(y)q(dy|x,\phi(x))= v\min_{a\in\mathbb{A}}\left[c(x,a)+\int_{\mathbb{X}}\tilde{u}(y)q(dy|x,a)\right], \tag{3.3}$$

and, since the left equation in (3.3) *implies inequality* (3.1), *every stationary policy satisfying* (3.3) *is average-cost optimal.*

To prove Theorem 3.2, we first state several auxiliary facts.

Lemma 3.3 *Consider a family of nonnegative real-valued functions* $\{f_n\}_{n=1,2,\ldots}$ *on a metric space* X. *If the family of functions* $\{f_n\}_{n=1,2,\ldots}$ *is equicontinuous on* X *and* $\sup_n f_n(x) < \infty$ *for each* $x \in X$, *then*

$$\liminf_{n\to\infty} f_n(x) = \tilde{f}(x) := \liminf_{n\to\infty, y\to x} f_n(y), \qquad x \in X. \tag{3.4}$$

Proof Fix an arbitrary $x \in X$. We first prove that there exists a subsequence $\{f_{\tilde{n}_l}\}_{l=1,2,\ldots}$ of $\{f_n\}_{n=1,2,\ldots}$ such that

$$\lim_{l\to\infty} f_{\tilde{n}_l}(x) = \tilde{f}(x). \tag{3.5}$$

In view of the definition of the function $\tilde{f}(x)$ in (3.4), there exist a subsequence $\{f_{n_k}\}_{k=1,2,\ldots}$ of $\{f_n\}_{n=1,2,\ldots}$ and a sequence $\{y_k\}_{k=1,2,\ldots} \subset X$ such that $n_k \to \infty$, $y_k \to x$ as $k \to \infty$ and

$$\lim_{k\to\infty} f_{n_k}(y_k) = \tilde{f}(x). \tag{3.6}$$

Since the family of functions $\{f_{n_k}\}_{k=1,2,\ldots}$ is equicontinuous on X and $\sup_k f_{n_k}(x) < \infty$, then, according to the Ascoli theorem (Hernández-Lerma and Lasserre 1996, p. 96), there exist a subsequence $\{f_{\tilde{n}_l}\}_{l=1,2,\ldots}$ of $\{f_{n_k}\}_{k=1,2,\ldots}$ and a continuous function $\tilde{f}^*$ such that

$$\lim_{l\to\infty} f_{\tilde{n}_l}(z) = \tilde{f}^*(z), \qquad z \in X. \tag{3.7}$$

Let $\tilde{y}_l := y_{k_l}$, $l = 1, 2, \ldots$. Then (3.6) implies

$$\lim_{l\to\infty} f_{\tilde{n}_l}(\tilde{y}_l) = \tilde{f}(x). \tag{3.8}$$

Let us fix an arbitrary $\epsilon > 0$. Then equality (3.7) implies that there exists a constant $N_1 > 0$ such that for all $l \geq N_1$

$$|\tilde{f}^*(x) - f_{\tilde{n}_l}(x)| < \epsilon/3. \tag{3.9}$$

Since the family of functions $\{f_{\tilde{n}_l}\}_{l=1,2,\ldots}$ is equicontinuous, then there exist a constant $N_2 > 0$ and a neighborhood $B(x)$ of x in X such that, for all $l \geq N_2$ and $\tilde{y}_l \in B(x)$

$$|f_{\tilde{n}_l}(x) - f_{\tilde{n}_l}(\tilde{y}_l)| < \epsilon/3. \tag{3.10}$$

In view of (3.8), there exists $N_3 > 0$ such that for all $l \geq N_3$

$$|f_{\tilde{n}_l}(\tilde{y}_l) - \tilde{f}(x)| < \epsilon/3. \tag{3.11}$$

Then (3.9), (3.10), and (3.11) imply that for all $l \geq \max\{N_1, N_2, N_3\}$

$$\begin{aligned} |\tilde{f}^*(x) - \tilde{f}(x)| &\leq |\tilde{f}^*(x) - f_{\tilde{n}_l}(x)| + |f_{\tilde{n}_l}(x) - f_{\tilde{n}_l}(\tilde{y}_l)| + |f_{\tilde{n}_l}(\tilde{y}_l) - \tilde{f}(x)| \\ &< \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon. \end{aligned} \tag{3.12}$$

Since $\epsilon > 0$ can be chosen arbitrarily, then (3.12) implies that

$$\tilde{f}^*(x) = \tilde{f}(x). \tag{3.13}$$

In view of (3.7) and (3.13),

$$\liminf_{n\to\infty} f_n(x) \le \tilde{f}(x). \tag{3.14}$$

The definition of the function $\tilde{f}$ in (3.4) implies that $\tilde{f}(x) \le \liminf_{n\to\infty} f_n(x)$. This inequality and (3.14) imply (3.4). □

Proof of Theorem 3.2 (i) Since the family of nonnegative functions $\{u_{\alpha_n}\}_{n=1,2,\dots}$ is equicontinuous and Assumption **B**(ii) holds, then, according to the Ascoli theorem (Hernández-Lerma and Lasserre 1996, p. 96), there exist a subsequence $\{\alpha_{n_k}\}_{k=1,2,\dots}$ of $\{\alpha_n\}_{n=1,2,\dots}$ and a continuous function $\tilde{u}^*(\cdot)$ such that

$$\lim_{k\to\infty} u_{\alpha_{n_k}}(x) = \tilde{u}^*(x), \qquad x \in \mathbb{X}, \tag{3.15}$$

and the convergence is uniform on each compact subset of $\mathbb{X}$.

According to Lemma 3.3, since the family of functions $\{u_{\alpha_{n_k}}\}$ is equicontinuous and $\sup_k u_{\alpha_{n_k}}(x) < \infty$ for each $x \in \mathbb{X}$, then $\liminf_{k\to\infty} u_{\alpha_{n_k}}(x) = \tilde{u}(x)$ for $x \in \mathbb{X}$, where $\tilde{u}(x)$ is defined in (2.5) for the subsequence $\{\alpha_{n_k}\}_{k=1,2,\dots}$. Therefore, (3.15) implies that

$$\lim_{k\to\infty} u_{\alpha_{n_k}}(x) = \tilde{u}^*(x) = \tilde{u}(x), \qquad x \in \mathbb{X}, \tag{3.16}$$

and the function $\tilde{u}$ is continuous on $\mathbb{X}$.

(ii) Since Assumptions **W*** and **B** hold, then according to Feinberg et al. (2012, Corollary 2), there exists a stationary policy ϕ satisfying the ACOI with $\tilde{u}$ defined in (2.5) for the sequence $\{\alpha_{n_k}\}_{k=1,2,\dots}$, that is

$$w + \tilde{u}(x) \ge c(x, \phi(x)) + \int_{\mathbb{X}} \tilde{u}(y) q(dy|x, \phi(x)). \tag{3.17}$$

To prove the ACOE, it remains to prove the opposite inequality to (3.17). According to Feinberg et al. (2012, Theorem 2(iv)), the discounted-cost optimality equation is $v_{\alpha_{n_k}}(x) = \min_{a\in\mathbb{A}}[c(x,a) + \alpha \int_{\mathbb{X}} v_{\alpha_{n_k}}(y) q(y|x,a)]$, $x \in \mathbb{X}$, which, by subtracting m_α from both sides, implies that for all $a \in \mathbb{A}$

$$(1-\alpha_{n_k}) m_{\alpha_{n_k}} + u_{\alpha_{n_k}}(x) \le c(x,a) + \alpha \int_{\mathbb{X}} u_{\alpha_{n_k}}(y) q(y|x,a), \qquad x \in \mathbb{X}. \tag{3.18}$$

Let $k \to \infty$. In view of (2.4) and (3.16), and Lebesgue's dominated convergence theorem, (3.18) implies that for all $a \in \mathbb{A}$

$$w + \tilde{u}(x) \le c(x,a) + \int_{\mathbb{X}} \tilde{u}(y) q(dy|x,a), \qquad x \in \mathbb{X},$$

which implies

$$w + \tilde{u}(x) \le \min_{a\in\mathbb{A}} [c(x,a) + \int_{\mathbb{X}} \tilde{u}(y) q(dy|x,a)], \qquad x \in \mathbb{X}. \tag{3.19}$$

Since $\min_{a\in\mathbb{A}}[c(x,a) + \int_{\mathbb{X}} \tilde{u}(y) q(dy|x,a)] \le c(x,\phi(x)) + \int_{\mathbb{X}} \tilde{u}(y) q(y|x,\phi(x))$, then (3.17) and (3.19) imply (3.3). □

4 Inventory control problem

In the rest of this paper, we study the application of Theorem 3.2 to the discrete-time periodic-review inventory control problem with backorders and convex costs. For examples of continuous review inventory control problems, see Bensoussan (2011, Chapter 14), Chen and Simchi-Levi (2006), Katehakis and Smit (2012), Presman and Sethi (2006), and Shi et al. (2013).

Let $\mathbb{R}$ denote the real line, $\mathbb{Z}$ denote the set of all integers, $\mathbb{R}^+ := [0, \infty)$ and $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$. Consider the classic stochastic periodic-review inventory control problem with fixed ordering cost, backorders, and generally distributed demand. At times $t = 0, 1, \ldots,$ a decision-maker views the current inventory of a single commodity 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. The unmet demand is backlogged and the cost of inventory held or backlogged (negative inventory) is modeled as a convex function. The demand and the order quantity are assumed to be non-negative. The state and action spaces are either (i) $\mathbb{X} = \mathbb{R}$ and $\mathbb{A} = \mathbb{R}^+$, or (ii) $\mathbb{X} = \mathbb{Z}$ and $\mathbb{A} = \mathbb{N}_0$. The inventory control problem is defined by the following parameters.

1. $K \geq 0$ is a fixed ordering cost;
2. $\bar{c} > 0$ is the per unit ordering cost;
3. $h(\cdot)$ is the holding/backordering cost per period, which is assumed to be a convex function on $\mathbb{X}$ with real values and $h(x) \to \infty$ as $|x| \to \infty$;
4. $\{D_t, t = 1, 2, \ldots\}$ is a sequence of i.i.d. nonnegative finite random variables representing the demand at periods $0, 1, \ldots$. We assume that $\mathbb{E}[h(x - D)] < \infty$ for all $x \in \mathbb{X}$ and $P(D > 0) > 0$, where D is a random variable with the same distribution as D_1;
5. $\alpha \in [0, 1)$ is the discount factor.

Note that $\mathbb{E}[D] < \infty$ since, in view of Jensen's inequality, $h(x - \mathbb{E}[D]) \leq \mathbb{E}[h(x - D)] < \infty$. Without loss of generality, assume that h is nonnegative and $h(0) = 0$. The assumption $P(D > 0) > 0$ avoids the trivial case when there is no demand. If $P(D = 0) = 1$, then the optimality inequality does not hold because $w(x)$ depends on x; see Feinberg and Lewis (2015) for details.

The dynamic of the system is defined by the equation

$$x_{t+1} = x_t + a_t - D_{t+1}, \quad t = 0, 1, 2, \ldots,$$

where x_t and a_t denote the current inventory level and the ordered amount at period t, respectively. Then the one-step cost is

$$c(x, a) = K I_{\{a>0\}} + \bar{c}a + \mathbb{E}[h(x + a - D)], \quad (x, a) \in \mathbb{X} \times \mathbb{A}, \tag{4.1}$$

where I_B is an indicator of the event B.

According to Feinberg and Lewis (2015, Corollary 6.1, Proposition 6.3), Assumptions **W*** and **B** hold for the MDP corresponding to the described inventory control problem. This implies that the optimality equation for the total discounted costs can be written as

$$v_\alpha(x) = \min\{\min_{a \geq 0}[K + G_\alpha(x + a)], G_\alpha(x)\} - \bar{c}x, \quad x \in \mathbb{X}, \tag{4.2}$$

where

$$G_\alpha(x) := \bar{c}x + \mathbb{E}[h(x - D)] + \alpha\mathbb{E}[v_\alpha(x - D)], \qquad x \in \mathbb{X}. \tag{4.3}$$

According to Feinberg and Liang (2017a, Theorem 5.3), the value function $v_\alpha(x)$ is continuous for all $\alpha \in [0, 1)$. The function $G_\alpha(x)$ is real-valued Feinberg and Liang (Feinberg and Lewis 2015, Corollary 6.4) and continuous (2017a, Theorem 5.3).

The function $c : \mathbb{X} \times \mathbb{A} \to \mathbb{R}$ is inf-compact; see Feinberg and Lewis (2015, Corollary 6.1). This property and the validity of Assumption **W*** imply that for each $\alpha \in [0, 1)$ the function v_α is inf-compact (Feinberg and Lewis 2007, Proposition 3.1(iv)) and therefore the set $X_\alpha := \{x \in \mathbb{X} : v_\alpha(x) = m_\alpha\}$, where m_α is defined in (2.3), is nonempty and compact. The validity of Assumptions **W*** and **B**(i) and the inf-compactness of c imply that there is a compact subset $\mathcal{K}$ of $\mathbb{X}$ such that $\mathbb{X}_\alpha \subseteq \mathcal{K}$ for all $\alpha \in [0, 1)$; Feinberg et al. (2012, Theorem 6). Following Feinberg and Lewis (2015), let us consider a bounded interval $[x_L^*, x_U^*] \subseteq \mathbb{X}$ such that

$$X_\alpha \subseteq [x_L^*, x_U^*] \qquad \text{for all } \alpha \in [0, 1). \tag{4.4}$$

Recall the definitions of K-convex functions and (s, S) policies.

Definition 4.1 A function $f : \mathbb{X} \to \mathbb{R}$ is called K-convex where $K \geq 0$, if for each $x \leq y$ and for each $\lambda \in (0, 1)$,

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda) f(x) + \lambda f(y) + \lambda K.$$

Definition 4.2 Let s and S be real numbers such that $s \leq S$. A stationary policy φ is called an (s, S) policy if

$$\varphi(x) = \begin{cases} S - x & \text{if } x < s; \\ 0 & \text{otherwise.} \end{cases}$$

Suppose $f(x)$ is a continuous K-convex function such that $f(x) \to \infty$ as $|x| \to \infty$. Let

$$S \in \arg\min_{x \in \mathbb{X}} \{f(x)\}, \tag{4.5}$$

$$s = \inf\{x \leq S : f(x) \leq K + f(S)\}. \tag{4.6}$$

Define

$$\alpha^* := 1 + \lim_{x \to -\infty} \frac{h(x)}{\bar{c}x}, \tag{4.7}$$

where the limit exists and $\alpha^* < 1$ since the function h is convex; see Feinberg and Liang (2017a).

Theorem 4.3 (Feinberg and Liang 2017a, Theorem 4.4(i) and Corollary 5.4) *If $\alpha \in (\alpha^*, 1)$ is a nonnegative discount factor, then an (s_α, S_α) policy is optimal for the discount factor α, where the real numbers S_α and s_α satisfy* (4.5) *and are defined in* (4.6) *respectively with $f(x) = G_\alpha(x)$, $x \in \mathbb{X}$. The stationary policy φ coinciding with the (s_α, S_α) policy at all $x \in \mathbb{X}$, except $x = s_\alpha$, where $\varphi(s_\alpha) = S_\alpha - s_\alpha$, is also optimal for the discount factor α.*

As shown in Feinberg and Lewis (2015, Equations (6.20), (6.23)), the optimality inequality can be written as

$$w + \tilde{u}(x) \geq \min\{\min_{a \geq 0}[K + H(x+a)], H(x)\} - \bar{c}x, \tag{4.8}$$

where

$$H(x) := \bar{c}x + \mathbb{E}[h(x - D)] + \mathbb{E}[\tilde{u}(x - D)]. \tag{4.9}$$

The following statement is Theorem 6.10(iii) from Feinberg and Lewis (2015) with the value of α^* is provided in (4.7); see Theorem 4.3.

Theorem 4.4 *For each nonnegative* $\alpha \in (\alpha^*, 1)$, *consider an optimal* (s'_α, S'_α) *policy for the discounted-cost criterion with the discount factor* α. *Let* $\{\alpha_n \uparrow 1\}_{n=1,2,\ldots}$ *be a sequence of nonnegative numbers with* $\alpha_1 > \alpha^*$. *Every sequence* $\{(s'_{\alpha_n}, S'_{\alpha_n})\}_{n=1,2,\ldots}$ *is bounded, and each its limit point* (s^*, S^*) *defines an average-cost optimal* (s^*, S^*) *policy. Furthermore, this policy satisfies the optimality inequality* (4.8), *where the function* $\tilde{u}$ *is defined in* (2.5) *for an arbitrary subsequence* $\{\alpha_{n_k}\}_{k=1,2,\ldots}$ *of* $\{\alpha_n\}_{n=1,2,\ldots}$ *satisfying* $(s^*, S^*) = \lim_{k\to\infty}(s'_{\alpha_{n_k}}, S'_{\alpha_{n_k}})$.

The following theorem states that the conditions and conclusions described in Theorem 3.2 hold for the described inventory control problem. It also states some problem-specific results.

Theorem 4.5 *For each sequence* $\{\alpha_n \uparrow 1\}_{n=1,2,\ldots}$ *of nonnegative discount factors with* $\alpha_1 > \alpha^*$, *the MDP for the described inventory control problem satisfies Assumption* ***EC***. *Therefore, the conclusions of Theorem* 3.2 *hold, that is, there exists a stationary policy* φ *such that for all* $x \in \mathbb{X}$

$$w + \tilde{u}(x) = K I_{\{\varphi(x)>0\}} + H(x + \varphi(x)) - \bar{c}x = \min\{\min_{a\geq 0}[K + H(x + a)], H(x)\} - \bar{c}x, \tag{4.10}$$

where the function H *is defined in* (4.9). *In addition, the functions* $\tilde{u}$ *and* H *are* K*-convex, continuous and inf-compact, and a stationary optimal policy* φ *satisfying* (4.10) *can be selected as an* (s^*, S^*) *policy described in Theorem* 4.4. *It also can be selected as an* (s, S) *policy with the real numbers* S *and* s *satisfying* (4.5) *and defined in* (4.6) *respectively for* $f(x) = H(x)$, $x \in \mathbb{X}$.

To prove Theorem 4.5, we first state several auxiliary facts. Consider the renewal process

$$\mathbf{N}(t) := \sup\{n = 0, 1, \ldots \,|\, \mathbf{S}_n \leq t\},$$

where $t \in \mathbb{R}^+$, $\mathbf{S}_0 = 0$ and $\mathbf{S}_n = \sum_{j=1}^n D_j$ for $n = 1, 2, \ldots$. Observe that since $P(D > 0) > 0$, then $\mathbb{E}[\mathbf{N}(t)] < \infty$, $t \in \mathbb{R}^+$; see Resnick (1992, Theorem 3.3.1).

Consider an arbitrary $\alpha \in [0, 1)$ and a state x_α such that $u_\alpha(x_\alpha) = m_\alpha$. Then, in view of (4.4), the inequalities $x_L^* \leq x_\alpha \leq x_U^*$ take place.

Define $E_y(x) := \mathbb{E}[h(x - \mathbf{S}_{\mathbf{N}(y)+1})]$ for $x \in \mathbb{X}$, $y \geq 0$. In view of Feinberg and Lewis (2015, Lemma 6.2), $E_y(x) < \infty$. According to Feinberg and Lewis (2015, inequalities (6.11), (6.17)), for $x < x^\alpha$

$$u_\alpha(x) \leq K + \bar{c}(x_U^* - x), \tag{4.11}$$

and for $x \geq x^\alpha$

$$u_\alpha(x) \leq K + (E(x) + \bar{c}\mathbb{E}[D])(1 + \mathbb{E}[\mathbf{N}(x - x_L^*)]), \tag{4.12}$$

where $E(x) := h(x) + E_{x-x_L^*}(x)$. Let

$$U(x) := \begin{cases} K + \bar{c}(x_U^* - x), & \text{if } x < x_L^*, \\ K + \bar{c}(x_U^* - x_L^*) + (E(x) + \bar{c}\mathbb{E}[D])(1 + \mathbb{E}[\mathbf{N}(x - x_L^*)]), & \text{if } x \geq x_L^*. \end{cases} \tag{4.13}$$

Lemma 4.6 *The following inequalities hold for* $\alpha \in [0, 1)$:

(i) $u_\alpha(x) \le U(x) < \infty$ *for all* $x \in \mathbb{X}$;
(ii) *If* $x_*, x \in \mathbb{X}$ *and* $x_* \le x$, *then* $C(x_*, x) := \sup_{y \in [x_*, x]} U(y) < \infty$;
(iii) $\mathbb{E}[U(x - D)] < \infty$ *for all* $x \in \mathbb{X}$.

Proof (i) For $x < x_L^*$ the inequality $u_\alpha(x) \le U(x)$ holds because of (4.11). For $x \ge x_L^*$ denote by f the function added to the constant K in the right-hand side of (4.12),

$$f(x) := (E(x) + \bar{c}\mathbb{E}[D]) \left(1 + \mathbb{E}\left[\mathbf{N}(x - x_L^*)\right]\right). \tag{4.14}$$

For $x \ge x_U^*$, inequality (4.12) and the inequality $u_U^* \ge x_L^*$ imply that

$$u_\alpha(x) \le K + f(x) \le K + \bar{c}(x_U^* - x_L^*) + f(x) = U(x),$$

where the first inequality is (4.12), for $x \ge u_U^* \ge x_\alpha$, and the second inequality follows from $u_U^* \ge x_L^*$. Thus, $u_\alpha(x) \le U(x)$ for $x \ge x_U^*$.

For $x_L^* \le x < u_U^*$

$$\begin{aligned} u_\alpha(x) &\le K + \max\{\bar{c}(x_U^* - x), f(x)\} \\ &\le K + \bar{c}(x_U^* - x) + f(x) \le K + \bar{c}(x_U^* - x_L^*) + f(x) = U(x), \end{aligned}$$

where the first inequality follows from (4.11), (4.12), and $x_L^* \le x_\alpha \le x_U^*$, the second inequality holds because the maximum of two nonnegative numbers is not greater than their sum, and the last inequality follows from $x_L^* \le x_\alpha \le x_U^*$. In addition, $U(x) < \infty$ because all the functions in the right-hand side of (4.13) take real values.

(ii) For $x < x_L^*$

$$C(x_*, x) \le \sup_{y \in [x_*, x_L^*)} U(y) \le K + \bar{c}(x_U^* - x_*) < \infty. \tag{4.15}$$

Let $x_L^* \le x_*$. In this case,

$$C(x_*, x) \le C(x_L^*, x) = K + \bar{c}(x_U^* - x_L^*) + \sup_{y \in [x_L^*, x]} f(y),$$

where the function f is defined in (4.14) and $f(y) \le (E(y) + \bar{c}\mathbb{E}[D])(1 + \mathbb{E}[\mathbf{N}(x - x_L^*)])$ for $y \in [x_L^*, x]$. To complete the proof of $C(x_*, x) < \infty$ for $x_L^* \le x_*$, we need to show that $\sup_{y \in [x_L^*, x]} E(y) < \infty$. This is true because of the following reasons. First, by Feinberg and Lewis (2015, inequalities (6.5), (6,6), and the inequality between them), for $z \ge 0$ and $y \in \mathbb{X}$

$$E_z(y) \le (1 + \mathbb{E}[\mathbf{N}(z)])\mathbb{E}[h(y - z - D)] + h(y). \tag{4.16}$$

Therefore, for $y \in [x_L^*, x]$

$$\begin{aligned} E(y) &\le \left(1 + \mathbb{E}\left[\mathbf{N}(y - x_L^*)\right]\right) \mathbb{E}\left[h(x_L^* - D)\right] + 2h(y) \\ &\le \left(1 + \mathbb{E}\left[\mathbf{N}(x - x_L^*)\right]\right) \mathbb{E}\left[h(x_L^* - D)\right] + 2\max\{h(x_L^*), h(x)\} < \infty, \end{aligned}$$

where the first inequality follows from the definition of the function $E(\cdot)$, introduced after (4.12), and from (4.16). The second inequality follows from the convexity of h and from $x_L^* \le y \le x$. Thus, for $x_L^* \le x_*$

$$C(x_*, x) \le C(x_L^*, x) = K + \bar{c}(x_U^* - x_L^*) + \sup_{y \in [x_L^*, x]} f(y) < \infty. \tag{4.17}$$

Now consider arbitrary $x_*, x \in \mathbb{X}$ such that $x_* \leq x$. Choose $z_*, z \in \mathbb{X}$ such that $z_* < \min\{x_*, x_L^*\}$ and $z > \max\{x, x_L^*\}$. Then

$$C(x_*, x) \leq C(z_*, z) \leq \max\left\{ \sup_{y \in [z_*, x_L^*)} \{U(y)\}, C(x_L^*, z) \right\} < \infty,$$

where the first inequality follows from $[x_*, x] \subset [z_*, z]$, the second inequality follows from $[z_*, z] = [z_*, x_L^*) \cup [x_L^*, z]$, and the last one follows from (4.15) and from (4.17).

(iii) Let us define $C(x_*, x) = 0$ for $x_*, x \in \mathbb{X}$ and $x_* > x$. For $x \in \mathbb{X}$

$$\begin{aligned}\mathbb{E}[U(x-D)] &= \mathbb{E}\left[U(x-D)I_{\{x-D<x_L^*\}}\right] + \mathbb{E}\left[U(x-D)I_{\{x_L^* \leq x-D \leq x\}}\right] \\ &\leq \mathbb{E}\left[(K + \bar{c}(x_U^* - x + D))I_{\{x-D<x_L^*\}}\right] + \mathbb{E}\left[C(x_L^*, x)I_{\{x_L^* \leq x-D \leq x\}}\right] \\ &\leq (K + \bar{c}(x_U^* - x))P(D > x - x_L^*) + \bar{c}\mathbb{E}[D] + C(x_L^*, x) < \infty,\end{aligned}$$

where the first equality holds because D is a nonnegative random variable, the first inequality follows from the definitions of the functions U and C, the second inequality holds because an expectation of an indicator of an event is its probability and because the random variable D and the constant $C(x_L^*, x)$ are nonnegative, and the last inequality follows from $\mathbb{E}[D] < \infty$ and from Lemma 4.6(ii). □

The next lemma establishes the equicontinuity on $\mathbb{X}$ of a family of discounted relative value functions.

Lemma 4.7 *For each sequence $\{\alpha_n \uparrow 1\}_{n=1,2,\ldots}$ of nonnegative discount factors with $\alpha_1 > \alpha^*$, the family $\{u_{\alpha_n}\}_{n=1,2,\ldots}$ is equicontinuous on $\mathbb{X}$.*

Proof Before providing the proof, we would like to describe its main idea. It is based on estimating the difference between the total discounted costs incurred when the process starts from two states, z_1 and z_2, when the distance between z_1 and z_2 is small. Let $z_1 < z_2$. This estimation is trivial when $z_2 \leq s_{\alpha_n}$ because the function $u_{\alpha_n}(x)$ is linear on $(-\infty, s_{\alpha_n}]$. By using Lemma (4.6)(ii), it is possible to derive such estimation for $z_1 \leq s_{\alpha_n} < z_2$. For $z_1 > s_{\alpha_n}$, the estimation consists of two parts: (i) the difference between the total holding costs incurred until the process, that starts at z_1, reaches the set $(-\infty, s_{\alpha_n}]$, and this difference is small because of the Lipshitz continuity of the convex function $\mathbb{E}[h(x-D)]$ on a bounded interval and because the average number of jumps is finite; (ii) the difference between the total costs incurred after the process, that starts at z_1, reaches $(-\infty, s_{\alpha_n}]$, and this difference is small because it is bounded by the differences of the total costs for the two cased $z_2 \leq s_{\alpha_n}$ and $z_1 \leq s_{\alpha_n} < z_2$ described above. Now we start the proof.

The discounted-cost optimality equations (4.2) and the optimality of $(s_{\alpha_n}, S_{\alpha_n})$ policies, stated in Theorem 4.3, imply that the function $v_{\alpha_n}(x)$ is linear, when $x \leq s_{\alpha_n}$, and

$$v_{\alpha_n}(x) = \begin{cases} \bar{c}(s_{\alpha_n} - x) + v_{\alpha_n}(s_{\alpha_n}), & \text{if } x \leq s_{\alpha_n}, \\ \tilde{h}(x) + \alpha_n \mathbb{E}[v_{\alpha_n}(x-D)], & \text{if } x \geq s_{\alpha_n}, \end{cases} \tag{4.18}$$

where $\tilde{h}(x) := \mathbb{E}[h(x-D)] < \infty$ is convex in x on $\mathbb{X}$. According to Theorem 4.4, since each sequence $\{(s_{\alpha_n}, S_{\alpha_n})\}_{n=1,2,\ldots}$ is bounded, then there exists a constant $b > 0$ such that

$$s_{\alpha_n} \in (-b, b), \qquad n = 1, 2, \ldots. \tag{4.19}$$

Therefore, there exists a constant $\delta_0 > 0$ such that $-b \leq s_{\alpha_n} - \delta_0 < s_{\alpha_n} + \delta_0 \leq b$, $n = 1, 2, \ldots$.

Consider $z_1, z_2 \geq s_{\alpha_n}$. Without loss of generality, assume that $z_1 < z_2$. According to (4.18), $v_{\alpha_n}(x) = \mathbb{E}[\sum_{j=1}^{\mathbf{N}(x-s_{\alpha_n})+1} \alpha_n^{j-1} \tilde{h}(x - \mathbf{S}_{j-1}) + \alpha_n^{\mathbf{N}(x-s_{\alpha_n})+1} v_{\alpha_n}(x - \mathbf{S}_{\mathbf{N}(x-s_{\alpha_n})+1})]$ for $x \geq s_{\alpha_n}$. Therefore, for $n = 1, 2, \ldots$

$$\begin{aligned}
&|u_{\alpha_n}(z_1) - u_{\alpha_n}(z_2)| = |v_{\alpha_n}(z_1) - v_{\alpha_n}(z_2)| \\
&= \Bigg| \mathbb{E} \Bigg[\sum_{j=1}^{\mathbf{N}(z_1 - s_{\alpha_n})+1} \alpha_n^{j-1} (\tilde{h}(z_1 - \mathbf{S}_{j-1}) - \tilde{h}(z_2 - \mathbf{S}_{j-1})) \\
&\quad + \alpha_n^{\mathbf{N}(z_1 - s_{\alpha_n})+1} \left(v_{\alpha_n}(z_1 - \mathbf{S}_{\mathbf{N}(z_1 - s_{\alpha_n})+1}) - v_{\alpha_n}(z_2 - \mathbf{S}_{\mathbf{N}(z_1 - s_{\alpha_n})+1}) \right) \Bigg] \Bigg| \\
&\leq \mathbb{E} \Bigg[\sum_{j=1}^{\mathbf{N}(z_1 - s_{\alpha_n})+1} |\tilde{h}(z_1 - \mathbf{S}_{j-1}) - \tilde{h}(z_2 - \mathbf{S}_{j-1})| \Bigg] \\
&\quad + \mathbb{E}\left[|u_{\alpha_n}\left(z_1 - \mathbf{S}_{\mathbf{N}(z_1 - s_{\alpha_n})+1}\right) - u_{\alpha_n}\left(z_2 - \mathbf{S}_{\mathbf{N}(z_1 - s_{\alpha_n})+1}\right)| \right],
\end{aligned} \tag{4.20}$$

where the inequality holds because of $\alpha_n < 1$, the change of the expectations and the absolute values, and because the sum of absolute values is greater or equal than the absolute value of the sum.

Consider $\epsilon > 0$. Define a positive number $\bar{N} := \mathbb{E}[\mathbf{N}(z_1 + b)] + 1 < \infty$. Since $b > -s_{\alpha_n}$, then $\mathbb{E}[\mathbf{N}(z_1 - s_{\alpha_n})] + 1 \leq \bar{N}$. Since the function $\tilde{h}(x)$ is convex on $\mathbb{R}$, then it is Lipschitz continuous on $[-b, z_2]$; see Hiriart-Urruty and Lemaréchal (1993, Theorem 3.1.1). Since Lipschitz continuity implies uniformly continuity, then there exists $\delta_1 \in (0, \delta_0)$ such that for $x, y \in [-b, z_2]$ satisfying $|x - y| < \delta_1$, $|\tilde{h}(x) - \tilde{h}(y)| < \frac{\epsilon}{2\bar{N}}$. Therefore, for $s_{\alpha_n} \leq z_1 < z_2$ satisfying $|z_1 - z_2| < \delta_1$

$$|\tilde{h}(z_1 - \mathbf{S}_j) - \tilde{h}(z_2 - \mathbf{S}_j)| < \frac{\epsilon}{2\bar{N}}, \qquad j = 0, 1, \ldots, \mathbf{N}(z_1 - s_{\alpha_n}), \tag{4.21}$$

and

$$\begin{aligned}
&\mathbb{E} \Bigg[\sum_{j=1}^{\mathbf{N}(z_1 - s_{\alpha_n})+1} \left| \tilde{h}(z_1 - \mathbf{S}_{j-1}) - \tilde{h}(z_2 - \mathbf{S}_{j-1}) \right| \Bigg] \leq \mathbb{E} \Bigg[\sum_{j=1}^{\mathbf{N}(z_1 - s_{\alpha_n})+1} \frac{\epsilon}{2\bar{N}} \Bigg] \\
&= \left(\mathbb{E}\left[\mathbf{N}(z_1 - s_{\alpha_n}) \right] + 1 \right) \frac{\epsilon}{2\bar{N}} \leq \frac{\epsilon}{2}.
\end{aligned} \tag{4.22}$$

where the first inequality follows from (4.21) and the last inequality holds because of $\mathbb{E}[\mathbf{N}(z_1 - s_{\alpha_n})] + 1 \leq \bar{N}$.

Additional arguments are needed to estimate the last term in (4.20). Next we prove that there exists $\delta_2 \in (0, \delta_1)$ such that for $x \in [s_{\alpha_n}, s_{\alpha_n} + \delta_2]$,

$$|u_{\alpha_n}(x) - u_{\alpha_n}(s_{\alpha_n})| < \frac{\epsilon}{4}, \qquad n = 1, 2, \ldots . \tag{4.23}$$

Let $x \geq s_\alpha$. Then formula (4.18) implies

$$v_{\alpha_n}(x) = \tilde{h}(x) + \alpha_n \mathbb{E}[v_{\alpha_n}(x - D)] \tag{4.24}$$

and

$$\begin{aligned}
\mathbb{E}\left[v_{\alpha_n}(x - D) \right] &= P(D \geq x - s_{\alpha_n}) \mathbb{E}\left[\bar{c}(s_{\alpha_n} - x + D) | D \geq x - s_{\alpha_n} \right] \\
&\quad + P(0 < D < x - s_{\alpha_n}) \mathbb{E}\left[v_{\alpha_n}(x - D) | 0 < D < x - s_{\alpha_n} \right] \\
&\quad + P(D = 0) v_{\alpha_n}(x)
\end{aligned} \tag{4.25}$$

Formulas (4.24) and (4.25) imply

$$[1-\alpha_n P(D=0)]v_{\alpha_n}(x) = \tilde{h}(x) + \alpha_n \Big(P(D \ge x - s_{\alpha_n})\mathbb{E}\left[\bar{c}(s_{\alpha_n} - x + D)|D \ge x - s_{\alpha_n}\right] + P(0 < D < x - s_{\alpha_n})\mathbb{E}\left[v_{\alpha_n}(x-D)|0 < D < x - s_{\alpha_n}\right]\Big). \tag{4.26}$$

Therefore, since $u_{\alpha_n}(y_1) - u_{\alpha_n}(y_2) = v_{\alpha_n}(y_1) - v_{\alpha_n}(y_2)$ for all $y_1, y_2 \in \mathbb{X}$, for $x \in [s_{\alpha_n}, s_{\alpha_n} + \delta_1]$ and for $n = 1, 2, \ldots$

$$\begin{aligned}
&[1-\alpha_n P(D=0)]|u_{\alpha_n}(x) - u_{\alpha_n}(s_{\alpha_n})| = [1-\alpha_n P(D=0)]|v_{\alpha_n}(x) - v_{\alpha_n}(s_{\alpha_n})| \\
&= \Big|\tilde{h}(x) - \tilde{h}(s_{\alpha_n}) + \alpha_n P(D \ge x - s_{\alpha_n})\bar{c}(s_{\alpha_n} - x) \\
&\quad + \alpha_n P(0 < D < x - s_{\alpha_n})\mathbb{E}[u_{\alpha_n}(x-D) - u_{\alpha_n}(s_{\alpha_n} - D)|0 < D < x - s_{\alpha_n}]\Big| \\
&\le |\tilde{h}(x) - \tilde{h}(s_{\alpha_n})| + \bar{c}(x - s_{\alpha_n}) + 2P(0 < D < x - s_{\alpha_n})C(-b, b),
\end{aligned} \tag{4.27}$$

where the nonnegative function C is defined in Lemma 4.6. Let us define $L := (1 - P(D = 0))^{-1}$, and $Q(x, s_{\alpha_n}) := P(0 < D < x - s_{\alpha_n})$. Recall that $P(D > 0) > 0$, which is equivalent to $P(D = 0) < 1$. Since $(1 - \alpha_n P(D = 0))^{-1} \le L$, Formula (4.27) implies that for $n = 1, 2, \ldots$

$$|u_{\alpha_n}(x) - u_{\alpha_n}(s_{\alpha_n})| \le L(|\tilde{h}(x) - \tilde{h}(s_{\alpha_n})| + \bar{c}(x - s_{\alpha_n}) + 2Q(x, s_{\alpha_n})C(-b, b)). \tag{4.28}$$

Since the function $\tilde{h}$ is convex, it is Lipshitz continuous on $[-b, b]$. Therefore, all three summands in the right-hand side of the last equations converge uniformly in n to 0 as $x \downarrow s_{\alpha_n}$. Therefore, there exists $\delta_2 \in (0, \delta_1)$ such that (4.23) holds for all $x \in [s_{\alpha_n}, s_{\alpha_n} + \delta_2]$.

Since $u_{\alpha_n}(x) = \bar{c}(s_{\alpha_n} - x) + u_{\alpha_n}(s_{\alpha_n})$ for all $x \le s_{\alpha_n}$, then for all $x, y \le s_{\alpha_n}$

$$|u_{\alpha_n}(x) - u_{\alpha_n}(y)| = \bar{c}|x - y| < \frac{\epsilon}{4}, \qquad n = 1, 2, \ldots, \tag{4.29}$$

for $|x - y| < \frac{\epsilon}{4\bar{c}}$. Let $\delta_3 := \min\{\frac{\epsilon}{4\bar{c}}, \delta_2\}$. Then (4.29) holds for $|x - y| < \delta_3$.

For $x \le s_{\alpha_n} \le y$ satisfying $|x - y| < \delta_3$

$$|u_{\alpha_n}(x) - u_{\alpha_n}(y)| \le |u_{\alpha_n}(x) - u_{\alpha_n}(s_{\alpha_n})| + |u_{\alpha_n}(s_{\alpha_n}) - u_{\alpha_n}(y)| < \frac{\epsilon}{2}, \tag{4.30}$$

where the first inequality is the triangle property and the second one follows from (4.23) and (4.29). Therefore, (4.23), (4.29) and (4.30) imply that $|u_{\alpha_n}(x) - u_{\alpha_n}(y)| < \frac{\epsilon}{2}$ for all $x, y \le s_{\alpha_n} + \delta_3$ satisfying $|x - y| < \delta_3$. Then for $|z_1 - z_2| < \delta_3$ with probability 1

$$|u_{\alpha_n}(z_1 - \mathbf{S}_{\mathbf{N}(z_1 - s_{\alpha_n})+1}) - u_{\alpha_n}(z_2 - \mathbf{S}_{\mathbf{N}(z_1 - s_{\alpha_n})+1})| < \frac{\epsilon}{2}, \qquad n = 1, 2, \ldots,$$

and therefore

$$\mathbb{E}[|u_{\alpha_n}(z_1 - \mathbf{S}_{\mathbf{N}(z_1 - s_{\alpha_n})+1}) - u_{\alpha_n}(z_2 - \mathbf{S}_{\mathbf{N}(z_1 - s_{\alpha_n})+1})|] < \frac{\epsilon}{2}, \qquad n = 1, 2, \ldots. \tag{4.31}$$

Formulae (4.20), (4.22) and (4.31) imply that for $z_1, z_2 \ge s_{\alpha_n}$ satisfying $|z_1 - z_2| < \delta_3$

$$|u_{\alpha_n}(z_1) - u_{\alpha_n}(z_2)| < \epsilon, \qquad n = 1, 2, \ldots. \tag{4.32}$$

Therefore, (4.29), (4.30) and (4.32) imply that for each $x \in \mathbb{X}$

$$|u_{\alpha_n}(x) - u_{\alpha_n}(y)| < \epsilon, \qquad n = 1, 2, \ldots,$$

if $|x - y| < \delta_3$, which means that the family $\{u_{\alpha_n}\}_{n=1,2,\ldots}$ is equicontinuous on $\mathbb{X}$. □

Proof of Theorem 4.5 Since $\int_{\mathbb{X}} U(y)q(dy|x,a) = \mathbb{E}[U(x+a-D)]$, where the function U is defined in (4.6), then, in view of Lemma (4.6)(iii), $\int_{\mathbb{X}} U(y)q(dy|x,a) < \infty$ for all $x \in \mathbb{X}$ and $a \in \mathbb{A}$. According to Lemma 4.7, the family $\{u_{\alpha_n}\}_{n=1,2,\ldots}$ is equicontinuous on $\mathbb{X}$. Therefore, Theorem 3.2 implies that there exists a subsequence $\{\alpha_{n_k}\}_{k=1,2,\ldots}$ of $\{\alpha_n\}_{n=1,2,\ldots}$ such that there exists a policy φ satisfying ACOE (3.3) with $\tilde{u}$ defined in (2.5) for the sequence $\{\alpha_{n_k}\}_{k=1,2,\ldots}$, the function $u_{\alpha_{n_k}}$ converges pointwise to $\tilde{u}$, and the function $\tilde{u}$ is continuous.

According to Feinberg and Lewis (2015, Theorem 6.10), the (s^*, S^*) policy satisfies the ACOI with $\tilde{u}$ defined in (2.5) for the sequence $\{\alpha_{n_k}\}_{k=1,2,\ldots}$. Since the ACOE holds with $\tilde{u}$ defined in (2.5) for the sequence $\{\alpha_{n_k}\}_{k=1,2,\ldots}$, then the (s^*, S^*) policy satisfies the ACOE.

Next we show that the functions $\tilde{u}$ and H are K-convex and inf-compact. Since the cost function c is inf-compact, the function $\tilde{u}$ is inf-compact; see Feinberg et al. (2012, Theorem 3 and Corollary 2). According to Feinberg and Lewis (2015, Lemma 6.8), the functions v_{α_n} are K-convex. Therefore the functions u_{α_n} are K-convex. Since $u_{\alpha_{n_k}}$ converges pointwise to $\tilde{u}$, then the function $\tilde{u}$ is K-convex. The function H is K-convex because, in view of (4.9), it is a sum of a linear, convex, and K-convex functions.

Since the (s^*, S^*) policy satisfies the ACOE (4.10) with $\tilde{u}$ defined in (2.5) for the sequence $\{\alpha_{n_k}\}_{k=1,2,\ldots}$, then $\tilde{u}(x) = K + H(S^*) - \bar{c}x - w$, for all $x < s$. Therefore, for $x < s$,

$$\begin{aligned} H(x) &= \bar{c}x + \mathbb{E}[h(x-D)] + \mathbb{E}[\tilde{u}(x-D)] \\ &= \bar{c}x + \mathbb{E}[h(x-D)] + K + H(S^*) - \bar{c}x + \bar{c}\mathbb{E}[D] - w \\ &= \mathbb{E}[h(x-D)] + K + H(S^*) + \bar{c}\mathbb{E}[D] - w. \end{aligned} \tag{4.33}$$

Since $\mathbb{E}[h(x-D)] \to \infty$ as $x \to -\infty$, then (4.33) implies that $H(x)$ tends to ∞ as $x \to -\infty$. Since h and $\tilde{u}$ are nonnegative, then (4.9) implies that $H(x) \geq \bar{c}x \to \infty$ as $x \to \infty$. Therefore, $H(x) \to \infty$ as $|x| \to \infty$.

Since $\tilde{u}$ is continuous and $(y-D)$ converges weakly to $(x-D)$ as $y \to x$, then $\mathbb{E}[\tilde{u}(x-D)]$ is lower semi-continuous. Since $\mathbb{E}[h(x-D)]$ is convex on $\mathbb{X}$ and hence continuous, $\bar{c}x$ is continuous and $\mathbb{E}[\tilde{u}(x-D)]$ is lower semi-continuous, then H is lower semi-continuous. Therefore, since $H(x)$ tends to ∞ as $|x| \to \infty$, then H is inf-compact.

According to the statements following Feinberg and Lewis (2015, Lemma 6.7), since H is K-convex, inf-compact, and tends to ∞ as $|x| \to \infty$, then an (s, S) policy, with the real numbers S and s satisfying (4.5) and defined in (4.6) respectively for $f(x) = H(x)$, $x \in \mathbb{X}$, is optimal.

Now we prove that the function H is continuous. Let us fix an arbitrary $y \in \mathbb{X}$. Define the following function

$$\bar{H}(x) = \begin{cases} \tilde{u}(x) + \bar{c}x, & \text{if } x \leq y+1, \\ \tilde{u}(y+1) + \bar{c}(y+1), & \text{if } x > y+1. \end{cases}$$

Since the functions $\tilde{u}(x)$ and $\bar{c}x$ are continuous, then the function $\bar{H}(x)$ is continuous. In view of (4.10), the function $\bar{H}(x)$ is bounded on $\mathbb{X}$. Therefore,

$$\lim_{z \to y} \{\mathbb{E}[h(z-D)] + \mathbb{E}[\bar{H}(z-D)]\} = \mathbb{E}[h(y-D)] + \mathbb{E}[\bar{H}(y-D)], \tag{4.34}$$

where the equality holds since the function $\mathbb{E}[h(x-D)]$ is convex on $\mathbb{X}$ and hence it is continuous, and $z - D$ converges weakly to $y - D$ as $z \to y$ and the function $\bar{H}(x)$ is continuous and bounded.

Observe that $H(x) = \mathbb{E}[h(x-D)] + \mathbb{E}[\bar{H}(x-D)] + \bar{c}\mathbb{E}[D]$ for all $x \leq y+1$. Therefore, (4.34) implies that $\lim_{z \to y} H(z) = H(y)$. Thus the function $H(x)$ is continuous. □

Corollary 4.8 *Let the state space $\mathbb{X} = \mathbb{R}$ and the action space $\mathbb{A} = \mathbb{R}^+$. For the (s, S) policy defined in Theorems 4.5, consider the stationary policy φ coinciding with this policy at all $x \in \mathbb{X}$, except $x = s$, and with $\varphi(s) = S - s$. Then the stationary policy φ also satisfies the optimality equation (4.10), and therefore the policy φ is average-cost optimal.*

Proof Since the proof of the optimality of (s, S) policies is based on the fact that $K + H(S) < H(x)$, if $x < s$, and $K + H(S) \geq H(x)$, if $x \geq s$. Since the function H is continuous, we have that $K + H(S) = H(s)$. Thus both actions are optimal at the state s. □

Theorem 4.5 states that Assumption **EC** holds. This implies the validity of the ACOE. The following theorem shows that a stronger equicontinuity assumption holds for the inventory control problem studied in this paper.

Theorem 4.9 *For each $\beta \in (\alpha^*, 1)$, the following statements hold:*

(a) the family of functions $\{u_\alpha : \alpha \in [\beta, 1)\}$ is equicontinuous;
(b) there exists a nonnegative measurable function $U(x)$ on $\mathbb{X}$ such that $U(x) \geq u_\alpha(x)$ for each $\alpha \in [\beta, 1)$ and $\int_{\mathbb{X}} U(y)q(dy|x, a) < \infty$ for all $x \in \mathbb{X}$ and $a \in \mathbb{A}$.

Proof Statement (b) follows from Lemma 4.6.

(a) Let $\gamma_1, \gamma_2 \in (\alpha^*, 1)$ be two discount factors such that $\gamma_1 < \gamma_2$. Define

$$s^*(\gamma_1, \gamma_2) := \inf\{x \leq S_{\gamma_1} : G_{\gamma_1}(x) \leq K + 1 + G_{\gamma_2}(S_{\gamma_2})\}. \tag{4.35}$$

Observe that $s^*(\gamma_1, \gamma_2)$ is a finite number because $G_{\gamma_1}(S_{\gamma_1}) \leq G_{\gamma_1}(S_{\gamma_2}) < K+1+G_{\gamma_2}(S_{\gamma_2})$, where the first holds since $G_{\gamma_1}(S_{\gamma_1})$ is the minimum of the function G_{γ_1} and the second one holds since $K + 1 > 0$ and $G_{\gamma_1}(\cdot) \leq G_{\gamma_2}(\cdot)$.

Consider an arbitrary $\gamma \in [\gamma_1, \gamma_2]$ and the quantity x_U^* introduced in (4.4). According to Theorem 4.3, for the discount factor γ there is an optimal (s_γ, S_γ) policy. We first prove that

$$s_\gamma \in [s^*(\gamma_1, \gamma_2), x_U^*]. \tag{4.36}$$

Since $G_{\gamma_2}(\cdot) \geq G_\gamma(\cdot) \geq G_{\gamma_1}(\cdot)$, then for $x \leq s^*(\gamma_1, \gamma_2)$

$$G_\gamma(x) \geq G_{\gamma_1}(x) \geq K + 1 + G_{\gamma_2}(S_{\gamma_2}) > K + G_\gamma(S_\gamma), \tag{4.37}$$

where the second inequality follows from (4.35). Since $G_\gamma(x) \leq K + G_\gamma(S_\gamma)$ for $x \in [s_\gamma, S_\gamma]$, then it follows from (4.37) that $s^*(\gamma_1, \gamma_2) \leq s_\gamma$.

To prove (4.36), it remains to show that $s_\gamma \leq x_U^*$. In view of (4.2), the optimality of (s_γ, S_γ) policies implies that $v_\gamma(x) \geq G_\gamma(S_\gamma) - \bar{c}x \geq G_\gamma(S_\gamma) - \bar{c}S_\gamma = v_\gamma(S_\gamma)$ for $x \leq S_\gamma$. Therefore, $S_\gamma \leq x_\gamma$, where $x_\gamma := \max\{x \in \mathbb{X} : v_\gamma(x) = \min_{z\in\mathbb{X}} v_\gamma(z)\}$. Since $s_\gamma \leq S_\gamma \leq x_\gamma$ and $x_\gamma \leq x_U^*$, then $s_\gamma \leq x_U^*$. Hence, (4.36) holds.

Secondly, we prove that for each $\beta \in (\alpha^*, 1)$ there exists a constant $b > 0$ such that

$$s_\alpha \in (-b, b), \qquad \alpha \in [\beta, 1). \tag{4.38}$$

According to Theorem 4.4, since each sequence $\{(s_{\alpha_n}, S_{\alpha_n})\}_{n=1,2,\dots}$ is bounded, then there exist nonnegative numbers $\gamma^* \in (\alpha^*, 1)$ and $b^* > 0$ such that $s_\alpha \in (-b^*, b^*)$ for all $\alpha \in [\gamma^*, 1)$. Therefore, (4.36) implies that (4.38) holds with b defined as

$$b := \begin{cases} b^* & \text{if } \beta \geq \gamma^*, \\ \max\{b^*, |s^*(\beta, \gamma^*)|, |x_U^*|\} + 1 & \text{otherwise.} \end{cases}$$

The remaining proof of the equicontinuity of the family of functions $\{u_\alpha : \alpha \in [\beta, 1)\}$ coincides with the proof of Lemma 4.7 with (4.19) replaced with (4.38) and α_n replaced with $\alpha \in [\beta, 1)$. □

Acknowledgements This research was partially supported by NSF Grants CMMI-1335296 and CMMI-1636193.

References

Bensoussan, A. (2011). *Dynamic programming and inventory control*. Amsterdam: IOS Press.

Bertsekas, D. P., & Shreve, S. E. (1996). *Stochastic optimal control: The discrete-time case*. Belmont, MA: Athena Scientific.

Beyer, D., Cheng, F., Sethi, S. P., & Taksar, M. (2010). *Markovian demand inventory models*. New York: Springer.

Beyer, D., & Sethi, S. P. (1999). The classical average-cost inventory models of Iglehart and Veinott–Wagner revisited. *Journal of Optimization Theory and Applications*, *101*(3), 523–555.

Cavazos-Cadena, R. (1991). A counterexample on the optimality equation in markov decision chains with the average cost criterion. *System and Control Letters*, *16*(5), 387–392.

Chen, X., & Simchi-Levi, D. (2004a). Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The finite horizon case. *Operations Research*, *52*(6), 887–896.

Chen, X., & Simchi-Levi, D. (2004b). Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The infinite horizon case. *Mathematics of Operations Research*, *29*(3), 698–723.

Chen, X., & Simchi-Levi, D. (2006). Coordinating inventory control and pricing strategies: The continuous review model. *Operations Research Letters*, *34*(3), 323–332.

Costa, O. L. V., & Dufour, F. (2012). Average control of Markov decision processes with Feller transition probabilities and general action spaces. *Journal of Mathematical Analysis and Applications*, *396*(1), 58–69.

Feinberg, E. A. (2016). Optimality conditions for inventory control. In A. Gupta & A. Capponi (Eds.), *Tutorials in operations research. Optimization challenges in complex, networked, and risky systems* (pp. 14–44). Cantonsville, MD: INFORMS.

Feinberg, E. A., Kasyanov, P. O., & Zadoianchuk, N. V. (2012). Average cost Markov decision processes with weakly continuous transition probability. *Mathematics of Operations Research*, *37*(4), 591–607.

Feinberg, E. A., Kasyanov, P. O., & Zadoianchuk, N. V. (2013). Berge's theorem for noncompact image sets. *Journal of Mathematical Analysis and Applications*, *397*(1), 255–259.

Feinberg, E. A., Kasyanov, P. O., & Zgurovsky, M. Z. (2016). Partially observable total-cost Markov decision processes with weakly continuous transition probabilities. *Mathematics of Operations Research*, *41*(2), 656–681.

Feinberg, E. A., & Lewis, M. E. (2007). Optimality inequalities for average cost Markov decision processes and the stochastic cash balance problem. *Mathematics of Operations Research*, *32*(4), 769–783.

Feinberg, E. A., & Lewis, M. E. (2015). On the convergence of optimal actions for Markov decision processes and the optimality of (s, S) policies for inventory control. Preprint arXiv:1507.05125. http://arxiv.org/pdf/1507.05125.pdf.

Feinberg, E. A., & Liang, Y. (2017a). Structure of optimal policies to periodic-review inventory models with convex costs and backorders for all values of discount factors. *Annals of Operations Research*. doi:10.1007/s10479-017-2548-6.

Feinberg, E. A., & Liang, Y. (2017b). Stochastic setup-cost inventory model with backorders and quasiconvex cost functions. Preprint arXiv:1705.06814. http://arxiv.org/pdf/1705.06814.pdf.

Hernández-Lerma, O., & Lasserre, J. B. (1996). *Discrete-time Markov control processes: Basic optimality criteria*. New York: Springer.

Hiriart-Urruty, J.-B., & Lemaréchal, C. (1993). *Convex analysis and minimization algorithms I*. Berlin: Springer.

Iglehart, D. L. (1963). Dynamic programming and stationary analysis of inventory roblems. In H. Scarf, D. Gilford, & M. Shelly (Eds.), *Multistage inventory control models and techniques* (pp. 1–31). Stanford, CA: Stanford University Press.

Jaśkiewicz, A., & Nowak, A. S. (2006). On the optimality equation for average cost Markov control processes with Feller transition probabilities. *Journal of Mathematical Analysis and Applications*, *316*(2), 495–509.

Katehakis, M. N., & Smit, L. C. (2012). On computing optimal (Q, r) replenishment policies under quantity discounts. *Annals of Operations Research*, *200*(1), 279–298.

Luque-Vasques, F., & Hernández-Lerma, O. (1995). A counterexample on the semicontinuity of minima. *Proceedings of the American Mathematical Society*, *123*(10), 3175–3176.

Montes-de-Oca, R. (1994). The average cost optimality equation for Markov control processes on Borel spaces. *Systems and Control Letters*, *22*(5), 251–357.

Presman, E., & Sethi, S. P. (2006). Inventory models with continuous and poisson demands and discounted and average costs. *Production and Operations Management*, *15*(2), 279–293.

Resnick, S. I. (1992). *Adventures in Stochastic Processes*. Boston: Birkhauser.

Schäl, M. (1993). Average optimality in dynamic programming with general state space. *Mathematics of Operations Research*, *18*(1), 163–172.

Sennott, L. I. (1998). *Stochastic dynamic programming and the control of queueing systems*. New York: Wiley.

Sennott, L. I. (2002). Average reward optimization theory for denumerable state systems. In E. A. Feinberg & A. Shwartz (Eds.), *Handbook of Markov decision processes: Methods and applications* (pp. 153–173). Boston, MA: Kluwer.

Shi, J., Katehakis, M. N., & Melamed, B. (2013). Martingale methods for pricing inventory penalties under continuous replenishment and compound renewal demands. *Annals of Operations Research*, *208*(1), 593–612.

Veinott, A. F., & Wagner, H. M. (1965). Computing optimal (*s, S*) policies. *Management Science*, *11*(5), 525–552.

Zheng, Y. (1991). A simple proof for optimality of (*s, S*) policies in infinite-horizon inventory systems. *Journal of Applied Probability*, *28*(4), 802–810.