Chapter 3

DISCOUNT COST MODELS WITH POLYNOMIALLY GROWING SURPLUS COST

3.1. Introduction

This chapter studies stochastic inventory problems with unbounded Markovian demands and more general costs than those considered in Chapter 2. Finite horizon problems, as well as stationary and nonstationary discounted cost infinite horizon problems, are addressed. Existence of optimal Markov or feedback policies is established with Markovian demand: unbounded, ordering costs that are l.s.c., and surplus costs that are l.s.c. with polynomial growth. Furthermore, optimality of (s, S)-type policies is proved when the ordering cost consists of fixed and proportional cost components and the surplus cost is convex.

The literature on infinite horizon inventory models involving a fixed ordering cost assumes surplus cost to be of linear growth and uniformly continuous as in Karlin (1958c), Scarf (1960), Bensoussan *et al.* (1983), and others. Even quadratic surplus costs that are popular in the production planning literature dating back to the classical HMMS model of Holt *et al.* (1960) have not been considered in infinite horizon inventory models.

As for demand, Karlin and Fabens (1960), Song and Zipkin (1993), Sethi and Cheng (1997), and Beyer and Sethi (1997) have all considered Markovian demands. Karlin and Fabens consider only the class of state-independent (s, S) policies, which does not in general include optimal policies. Song and Zipkin consider Markov-modulated Poisson demand in their analysis. Sethi and Cheng consider general Markovian demands in their treatment of discounted cost problems.

In this chapter, we consider unbounded Markovian demands but require that a certain number (depending on the growth rate of the surplus

D. Beyer et al., *Markovian Demand Inventory Models*, International Series in Operations Research & Management Science 108,
DOI 10.1007/978-0-387-71604-6_3,

cost function) of moments be finite. This is an essential requirement and yet not very restrictive. As both cost and demand are generalized, the chapter represents a significant extension of the infinite horizon inventory problems (involving a fixed ordering cost component) that have appeared in the literature.

In this chapter, we conduct a detailed analysis of the discounted cost problem. The problem is carefully formulated in Section 3.2. In Section 3.3, we use the dynamic programming equation to prove the existence of an optimal Markov control for the finite horizon problem. We also provide a verification theorem for the solution of the dynamic programming equation to be the value function. We prove the value function to be continuous when the surplus cost is continuous and the ordering cost is l.s.c. As we will remark later, this has some implications for *whether or not to order at the level* s in an optimal (s, S)-type policy. The nonstationary infinite horizon problem is treated in Section 3.4. With further assumptions on costs, the optimality of (s, S)-type policies is established in Section 3.5. The stationary infinite horizon problem is briefly discussed in Section 3.6. The chapter concludes with end notes in Section 3.7.

3.2. Formulation of the Model

Let us consider an inventory problem over a finite number of periods $\langle n, N\rangle = \{n, n+1, \ldots, N\}$, and an initial inventory of x units at the beginning of period n, where n and N are any given integers satisfying $0 \leq n \leq N < \infty$. The demand in each period is assumed to be a random variable defined on a given probability space $(\Omega, \mathcal{F}, \mathsf{P})$, and not necessarily identically distributed. More specifically, the demand distributions in successive periods are defined as below.

Consider a finite collection of demand states $\mathbb{I} = \{1, 2, \ldots, L\}$, and let i_k denote the demand state in the k^{th} period. We assume that i_k, $k \in \langle n, N\rangle$, with known initial demand state i_n, is a Markov chain over $\mathbb{I}$ with the transition matrix $P = \{p_{ij}\}$. Thus,

$$0 \leq p_{ij} \leq 1, \; i \in \mathbb{I}, j \in \mathbb{I}, \;\; \text{and} \;\; \sum_{j=1}^{L} p_{ij} = 1, \; i \in \mathbb{I}.$$

Let a nonnegative random variable ξ_k denote the demand in a given period k, $k = 0, \ldots, N-1$. Demand ξ_k depends only on period k and the demand state in that period, by which we mean that it is independent of past demand states and past demands. We denote its cumulative probability distribution by $\Phi_{i,k}(x)$, when the demand state $i_k = i$. In the following period, if the state changes to state j, which happens with probability p_{ij}, then the demand distribution is $\Phi_{j,k+1}$ in that period. We further assume that for some $\gamma \geq 1$ and a positive constant D,

$$\mathsf{E}(\xi_k^\gamma | i_k = i) = \int_0^\infty x^\gamma d\Phi_{i,k}(x) \leq D < \infty,\ k = 0, \dots, N-1,\ i \in \mathbb{I}. \quad (3.1)$$

This is not a very restrictive assumption from an applied perspective.

We denote by

$$\begin{aligned} &\mathcal{F}_l^k, \text{ the } \sigma\text{-algebra generated by } \{i_l, \dots, i_{k-1}, i_k; \xi_l, \dots, \xi_{k-1}\}, \\ &\quad 0 \leq l \leq k \leq N, \\ &\mathcal{F}^k = \mathcal{F}_0^k. \end{aligned} \quad (3.2)$$

Since $i_k,\ k = 1, \dots, N$, is a Markov chain and ξ_k depends only on i_k, we have

$$\begin{aligned} \mathsf{E}(\xi_k | \mathcal{F}^k) &= \mathsf{E}(\xi_k | i_0, i_1, \dots, i_k; \xi_0, \xi_1, \dots, \xi_{k-1}) \\ &= \mathsf{E}(\xi_k | i_k). \end{aligned} \quad (3.3)$$

An admissible decision (ordering quantities) for the problem on the interval $\langle n, N \rangle$ with initial state $i_n = i$ can be denoted as

$$U = (u_n, \dots, u_{N-1}), \quad (3.4)$$

where u_k is a nonnegative $\mathcal{F}_n^k$-measurable random variable. In simpler terms, this means that decision u_k depends only on the past information. Note that since i_n is known in period n, $\mathcal{F}_n^n = (\Omega, \emptyset)$; hence u_n is deterministic. Moreover, it should be emphasized that this class of admissible decisions is larger than the class of admissible feedback policies.

Ordering quantities are decided upon at the beginning of each period. Demand in each period is supposed to occur at the end of the period after the order has been delivered. Unsatisfied demand is carried forward as backlog. The inventory balance equations are defined by

$$\begin{cases} x_{k+1} = x_k + u_k - \xi_k,\ k = n, \dots, N-1, \\ x_n = x, \ \text{ initial inventory level}, \\ i_k, k = n, \dots, N,\ \text{Markov chain with transition matrix } P, \\ i_n = i,\ \text{initial state}, \end{cases}$$

where x_k is the surplus level at the beginning of period k, u_k is the quantity ordered at the beginning of period k, i_k is the demand state in period k, and ξ_k is the demand in period k. Note that $x_k > 0$ represents an inventory of x_k and $x_k < 0$ represents a backlog (or shortage) of $-x_k$.

Next, we specify the relevant costs and the assumptions they satisfy.

(i) The production cost function $c_k(i,u) : \mathbb{I} \times \mathbb{R}^+ \to \mathbb{R}^+$ is l.s.c., $c_k(i,0) = 0$, $k = 0,1,\ldots,N-1$.

(ii) The surplus cost function $f_k(i,x) : \mathbb{I} \times \mathbb{R}^+ \to \mathbb{R}^+$ is l.s.c., with $f_k(i,x) \le \bar{f}(1+|x|^\gamma)$, $i = 1,2\ldots,L$, $k = 0,1,\ldots,N-1$, where $\bar{f}$ is a nonnegative constant. When $x < 0$, $f_k(i,x)$ is the cost of backlogged sales x, and when $x > 0$, $f_k(i,x)$ is the carrying cost of holding inventory x during that period.

(iii) $f_N(i,x) : \mathbb{I} \times \mathbb{R}^+ \to \mathbb{R}^+$, the penalty cost/disposal cost for the terminal surplus, is l.s.c. with $f_N(i,x) \le \bar{f}(1+|x|^\gamma)$. When $x < 0$, $f_N(i,x)$ represents the penalty cost of unsatisfied demand x, and when $x > 0$, $f_N(i,x)$ represents the disposal cost of the inventory level x.

The objective function to be minimized is the expected present value of all the costs incurred during the interval $\langle n,N\rangle$, i.e.,

$$\begin{aligned} J_n(i,x;U) \quad = \quad & \mathsf{E}\Bigg\{\sum_{k=n}^{N-1} \alpha^{k-n}[c_k(i_k,u_k) + f_k(i_k,x_k)] \\ & +\alpha^{N-n} f_N(i_N,x_N)\Bigg\}, \end{aligned} \tag{3.5}$$

which is always defined, provided we allow this quantity to be infinite.

In (3.5), α denotes the discount factor with $0 < \alpha < 1$. While introducing α in this finite horizon model does not add to generality, in view of the already time-dependent nature of the cost functions, we do so for convenience of exposition in dealing with the average cost optimality criterion studied in Chapter 6.

3.3. Dynamic Programming and Optimal Feedback Policy

Let us first introduce the following definitions.

$\mathbb{B}^\gamma =$ Banach space of Borel functions $b : \mathbb{I} \times \mathbb{R} \to \mathbb{R}$ with polynomial growth with power γ or less. More specifically, if $b \in \mathbb{B}^\gamma$, then $|b(i,x)| \le \| b \|_\gamma (1+|x|^\gamma)$, where the norm

$$\| b \|_\gamma = \max_i \sup_x \frac{|b(i,x)|}{1+|x|^\gamma} < \infty. \tag{3.6}$$

$\mathbb{L}^\gamma =$ the subspace of l.s.c. functions in $\mathbb{B}^\gamma$. The space $\mathbb{L}^\gamma$ is closed in $\mathbb{B}^\gamma$.

$\mathbb{L}^{\gamma^-}$ = the class of all l.s.c. functions which are of polynomial growth with power γ or less on $(-\infty, 0]$.

In view of (3.3), we can write

$$\mathsf{E}[b(i_{k+1}, \xi_k)|\mathcal{F}^k] = \mathsf{E}[b(i_{k+1}, \xi_k)|i_k], \tag{3.7}$$

for any $b \in \mathbb{B}^\gamma$.

To write the dynamic programming equations more concisely, we define the operator F_{n+1} on $\mathbb{B}^\gamma$ as follows:

$$\begin{aligned} F_{k+1}b(i,y) &= \mathsf{E}[b(i_{k+1}, y-\xi_k)|i_k = i] \\ &= \sum_{j=1}^{L} \{\mathsf{P}(i_{k+1} = j|i_k = i)\mathsf{E}[b(j, y-\xi_k)|i_k = i]\} \\ &= \sum_{j=1}^{L} p_{ij} \int_0^\infty b(j, y-\xi) d\Phi_{i,k}(\xi). \end{aligned} \tag{3.8}$$

In addition to Assumptions (i)-(iii) on costs, we also require that for $k = 0, 1, \ldots, N-1$,

$$0 \le c_k(i,u) + \alpha F_{k+1}(f_{k+1})(i,u) \to \infty \text{ for } u \to \infty. \tag{3.9}$$

REMARK 3.1 Condition (3.9) implies that both the purchase cost and the inventory (or salvage) cost associated with a decision in any given period cannot both be zero. The conditions rule out the trivial and unrealistic situation of ordering an infinite amount as the optimal policy. See Remark 2.2 for further elaboration.

Let $v_n(i,x)$ represent the optimal value of the expected costs during the time horizon $\langle n, N\rangle$ with demand state i in period n, i.e.,

$$v_n(i,x) = \inf_U J_n(i,x;U).$$

Then, $v_n(i,x)$ satisfies the dynamic programming equations

$$\begin{aligned} v_n(i,x) &= f_n(i,x) + \inf_{u\ge 0} \{c_n(i,u) + \alpha\mathsf{E}[v_{n+1}(i_{n+1}, \\ &\qquad x+u-\xi_n)|i_n = i]\} \\ &= f_n(i,x) + \inf_{u \ge 0} \{c_n(i,u) + \alpha F_{n+1}(v_{n+1})(i, x+u)\}, \\ &\qquad n = 0, 1, \ldots, N-1, \end{aligned} \tag{3.10}$$

$$v_N(i,x) = f_N(i,x). \tag{3.11}$$

We can now state our existence results in the following two theorems.

THEOREM 3.1 *The dynamic programming equations* (3.10) *and* (3.11) *define a sequence of functions in* $\mathbb{L}^\gamma$. *Moreover, there exists a Borel function* $\hat{u}_n(i,x)$ *such that the infimum in* (3.10) *is attained at* $u = \hat{u}_n(i,x)$ *for any* x. *Furthermore, if the functions* $f_n(i,\cdot)$, $n = 0, 1, \ldots, N$, $i = 1, 2 \ldots, L$, *are continuous, then the functions defined by* (3.10) *and* (3.11) *are continuous.*

Proof. We proceed by induction. Because of Assumption (iii) and (3.11), it follows that $v_N(i,x) \in \mathbb{L}^\gamma$. Assume $v_{n+1}(i,x)$ belongs to $\mathbb{L}^\gamma$. Consider points x such that $|x| \leq M$, where M is an arbitrary nonnegative integer. Let

$$B^M_{n,i} = \sup_{|x|\leq M} \{\alpha F_{n+1}(v_{n+1})(i,x)\}. \tag{3.12}$$

The constant $B^M_{n,i}$ is finite since $v_{n+1}(i,x)$ is in $\mathbb{B}^\gamma$ and therefore bounded on $|x| \leq M$, and F_{n+1} is a continuous linear operator; (see Lemma A.4.1). Because of (3.9), we know that the set

$$N^M_{n,i} := \{u \geq 0 : \inf_{|x|\leq M} \{c_n(i,u) + \alpha F_{n+1}(f_{n+1})(i, x+u) \leq B^M_{n,i}\} \tag{3.13}$$

is bounded, i.e., there is a $\bar{u}^M_{n,i}$ such that

$$N^M_{n,i} \subseteq [0, \bar{u}^M_{n,i}]. \tag{3.14}$$

Because of $v_{n+1} \geq f_{n+1}$, we conclude that

$$\begin{aligned}\{u \geq 0 : \inf_{|x|\leq M} \{c_n(i,u) + \alpha F_{n+1}(v_{n+1})(i, x+u) \leq B^M_{n,i}\}\\ \subseteq N^M_{n,i} \subseteq [0, \bar{u}^M_{n,i}],\end{aligned} \tag{3.15}$$

and, therefore without loss of optimality, we can restrict our attention to $0 \leq u \leq \bar{u}^M_{n,i}$ for all x satisfying $|x| \leq M$. This is because for any $u > \bar{u}^M_{n,i}$,

$$\begin{aligned} c_n(i,u) &+ \alpha F_{n+1}(v_{n+1})(i, x+u) \\ &\geq \inf_{|x|\leq M} \{c_n(i,u) + \alpha F_{n+1}(v_{n+1})(i, x+u)\} \\ &> B^M_{n,i} = \sup_{|x|\leq M} \{\alpha F_{n+1}(v_{n+1})(i,x)\} \\ &\geq \alpha F_{n+1}(v_{n+1})(i,x) \\ &= c_n(i,0) + \alpha F_{n+1}(v_{n+1})(i,x), \end{aligned}$$

and thus u cannot be the point where the infimum is attained.

Since the function

$$\psi_n(i,x,u) = c_n(i,u) + \alpha F_{n+1}(v_{n+1})(i,x+u)$$

is l.s.c. and bounded from below, its minimum over a compact set is attained; (see Theorem A.1.5). Moreover, from the classical Selection Theorem A.1.7, we know that there exists a Borel function $\hat{u}_n^M(i,x)$ such that

$$\psi_n(i,x,\hat{u}_n^M(i,x)) = \inf_{0\le u\le \bar{u}_{n,i}^M} \psi_n(i,x,u),\ |x| \le M. \tag{3.16}$$

Upon defining

$$\hat{u}_n(i,x) = \hat{u}_n^M(i,x) \ \text{ for } M-1 < |x| \le M,$$

we obtain a Borel function such that

$$\psi_n(i,x,\hat{u}_n(i,x)) = \inf_{u\ge 0} \psi_n(i,x,u),\ \forall x. \tag{3.17}$$

Since

$$\inf_{u\ge 0} \psi_n(i,x,u) \le \psi_n(i,x,0) \le c_n(i,0) + \alpha \parallel F_{n+1} \parallel \parallel v_{n+1} \parallel_\gamma (1+|x|^\gamma),$$

we can use (3.10), Assumption (ii), and Lemma A.4.1 to conclude that $v_n(i,x) \in \mathbb{B}^\gamma$. Furthermore, because $\psi_n(i,\cdot,\cdot)$ is l.s.c., it follows from equation (3.16) that $v_n(i,\cdot)$ is l.s.c. for each i (see Theorem A.1.6), and therefore $v_n \in \mathbb{L}^\gamma$.

To prove the last part of the theorem, we begin with the fact that for each i the function $f_n(i,\cdot)$ is continuous, $n = 0,1,\ldots,N$. Then $v_N(i,\cdot) = f_N(i,\cdot)$ is continuous for all i, and the continuity of v_n can be proved by induction as follows. Assume v_{n+1} to be continuous. From (3.10) and (3.17) we derive

$$v_n(i,x_0) = f_n(i,x_0) + c_n(i,\hat{u}(i,x_0)) + F_{n+1}(v_{n+1})(i,x_0+\hat{u}(i,x_0))$$

and, because for $x = x_1$ the infimum in (3.10) is not necessarily attained at $\hat{u}(i,x_0)$, we have

$$v_n(i,x_1) \le f_n(i,x_1) + c_n(i,\hat{u}(i,x_0)) + F_{n+1}(v_{n+1})(i,x_1+\hat{u}(i,x_0)).$$

Thus,

$$\begin{aligned} v_n(i,x_1) - v_n(i,x_0) &\le f_n(i,x_1) - f_n(i,x_0) \\ &\quad + F_{n+1}(v_{n+1})(i,x_1+\hat{u}(i,x_0)) - F_{n+1}(v_{n+1})(i,x_0+\hat{u}(i,x_0)), \end{aligned}$$

which, in view of the continuity of f_n and $v_{n+1} \in \mathbb{L}^\gamma$ and (3.1), yields

$$\limsup_{x_1\to x_0}\ v_n(i,x_1) - v_n(i,x_0) \le 0.$$

On the other hand, we have already proved that v_n is l.s.c., which means

$$\liminf_{x_1 \to x_0} \; v_n(i, x_1) - v_n(i, x_0) \geq 0.$$

Therefore, v_n is continuous. □

To solve the problem of minimizing $J_0(i, x; U)$, let us define

$$\begin{cases} \hat{x}_0 = x, \\ \hat{u}_n = \hat{u}_n(i_n, \hat{x}_n),\ n = 0, \dots, N-1, \\ \hat{x}_n + 1 = \hat{x}_n + \hat{u}_n - \xi_n,\ n = 0, \dots, N-1, \\ i_n,\ n = 0, \dots, N,\ \text{Markov chain with transition matrix } P, \\ i_0 = i, \end{cases}$$

where $\hat{u}_n(i, x)$ is a Borel function for which the infimum in (3.10) is attained for any i and x.

THEOREM 3.2 (Verification Theorem) *The policy* $\hat{U} = (\hat{u}_0, \hat{u}_1, \dots, \hat{u}_{N-1})$ *minimizes* $J_0(i, x; U)$ *over the class* $\mathcal{U}$ *of all admissible decisions. Moreover,*

$$v_0(i, x) = \min_{U \in \mathcal{U}} J_0(i, x; U). \tag{3.18}$$

Proof. Let $U = (u_0, \dots, u_{N-1})$ be any admissible decision with the corresponding trajectory $(x_0, \dots, x_{N-1})$. Without loss of generality, we may assume that $\mathsf{E}c_n(i_n, u_n) < \infty$, $\mathsf{E}f_n(i_n, x_n) < \infty$, $n \in \langle 0, N-1 \rangle$, and $\mathsf{E}f_N(i_N, x_N) < \infty$. Otherwise, $J_0(i, x; U) = \infty$ and U cannot be optimal since $J_0(i, x; \mathbf{0}) < \infty$ in view of (3.1) and Assumptions (i)-(iii).

Because $v_n(i_N, x_N) = f_N(i_N, x_N)$, it follows that $\mathsf{E}v_N(i_N, x_N) < \infty$. Using arguments analogous to those in the proof of Theorem 2.2, we proceed by induction. Assume that $\mathsf{E}v_{n+1}(i_{n+1}, x_{n+1}) < \infty$. Using property (vii) in Section B.2, (3.7), and (3.8), we obtain (see details leading to (2.15))

$$\mathsf{E}\{v_{n+1}(i_{n+1}, x_{n+1}) | \mathcal{F}^n\} = F_{n+1}(v_{n+1})(i_n, x_n + u_n) \text{ a.s.} \tag{3.19}$$

Since U is admissible but not necessarily optimal, we can use (3.10) to assert that

$$v_n(i_n, x_n) \leq f_n(i_n, x_n) + c_n(i_n, u_n) + \alpha F_{n+1}(v_{n+1})(i_n, x_n + u_n) \text{ a.s.}$$

Then from the relation (3.19), we can derive

$$v_n(i_n, x_n) \leq f_n(i_n, x_n) + c_n(i_n, u_n) + \alpha \mathsf{E}\{v_{n+1}(i_{n+1}, x_{n+1}) | \mathcal{F}^n\}.$$

Taking expectation of both sides of the above inequality, we obtain

$$\begin{aligned}\alpha^n \mathsf{E} v_n(i_n, x_n) \quad &\leq \quad \alpha^n \mathsf{E}(f_n(i_n, x_n) + c_n(i_n, x_n)) \\ &\qquad + \alpha^{n+1} \mathsf{E}(v_{n+1}(i_{n+1}, x_{n+1})). \qquad (3.20)\end{aligned}$$

We can conclude recursively that $\mathsf{E} v_n(i_n, x_n) < \infty$ for all $n \in \langle 0, N \rangle$ and that (3.20) holds. Then, by summing (3.20) from 0 to $N-1$ and canceling identical terms on both sides, we obtain

$$v_0(i, x) \leq J_0(i, x; U). \qquad (3.21)$$

Consider now the decision $\hat{U}$. Using the definition of $\hat{u}_n(i_n, x)$ as the Borel function for which the infimum in (3.10) is attained, and proceeding as above, we can obtain

$$\begin{aligned}&\alpha^n \mathsf{E} v_n(i_n, \hat{x}_n) \\ &\quad = \quad \alpha^n \mathsf{E}(f_n(i_n, \hat{x}_n) + c_n(i_n, \hat{u}_n)) + \alpha^{n+1} \mathsf{E}(v_{n+1}(i_{n+1}, \hat{x}_{n+1})).\end{aligned}$$

Note that $\hat{x}_0 = x$ is deterministic and $v_0(i, x) \in \mathbb{L}^\gamma$. Therefore, $\mathsf{E} v_0(i_0, \hat{x}_0) = v_0(i, x) < \infty$, and furthermore, it can be shown recursively that $\mathsf{E} c_n(i_n, \hat{u}_n) < \infty$, $n \in \langle 0, N{-}1 \rangle$ and $\mathsf{E} f_n(i_n, \hat{x}_n) < \infty$, $\mathsf{E} v_n(i_n, \hat{x}_n) < \infty$, $n \in \langle 0, N \rangle$. Adding up for n from 0 to $N{-}1$ and canceling terms, we get

$$v_0(i, x) = J_0(i, x; \hat{U}).$$

This and the inequality (3.21) complete the proof. □

3.4. Nonstationary Discounted Infinite Horizon Problem

In this section, we consider an infinite horizon version of the model formulated in Section 3.2. We require that the Assumptions (i) and (ii) hold with $N = \infty$, and that $i_0, i_1, \ldots$ is a Markov chain with the same transition matrix P. We set $N = \infty$, replace $\langle n, N \rangle$ by $\langle n, \infty \rangle$, replace the admissible decision in (3.4) by

$$U = (u_n, u_{n+1}, \ldots), \qquad (3.22)$$

and replace (3.5) by the objective function

$$J_n(i, x; U) = \sum_{k=n}^{\infty} \alpha^{k-n} \mathsf{E}[c_k(i_k, u_k) + f_k(i_k, x_k)], \qquad (3.23)$$

where α is the given discount factor, $0 < \alpha < 1$. The dynamic programming equations for each $i \in \mathbb{I}$ and $n = 0, 1, 2, \ldots$ can be written as

$$v_n(i,x) = f_n(i,x) + \inf_{u \geq 0}\{c_n(i,u) + \alpha F_{n+1}(v_{n+1})(i, x+u)\}. \tag{3.24}$$

In what follows, we will show that there exists an $\mathbb{L}^\gamma$-solution of (3.24), which is the value function of the infinite horizon problem. Moreover, the decision, for which the infimum in (3.24) is attained, is an optimal feedback policy.

First, let us examine the finite horizon approximation $J_{n,m}(i,x;U)$, $m \geq 1$, of (3.23). The approximation is obtained by the first m-period truncation of the infinite horizon problem of minimizing $J_n(i,x;U)$, i.e.,

$$J_{n,m}(i,x;U) = \sum_{k=n}^{n+m-1} \mathsf{E}[c_l(i_k,u_k) + f_k(i_k,x_k)]\alpha^{k-n}. \tag{3.25}$$

Let $v_{n,m}(i,x)$ be the value function of the truncated problem with no penalty cost in the last period, i.e.,

$$v_{n,m}(i,x) = \inf_U J_{n,m}(i,x;U). \tag{3.26}$$

Since the truncated problem is a finite horizon problem defined on the interval $\langle n, n+m \rangle$, Theorems 3.1 and 3.2 apply. Therefore, its value function can be obtained by solving the corresponding dynamic programming equations

$$\begin{aligned} v_{n,m+1}(i,x) &= f_n(i,x) + \inf_{u \geq 0}\{c_n(i,u) \qquad (3.27) \\ &\qquad + \alpha F_{n+1}(v_{n+1,m})(i,x+u)\}, \\ v_{n+m,0}(i,x) &= 0. \qquad (3.28) \end{aligned}$$

Moreover, $v_{n,0}(i,x) = 0$ and

$$v_{n,m}(i,x) = \min_U J_{n,m}(i,x;U).$$

Next we will show that an *a priori* upper bound on $\inf J_n(i,x;U)$ can be easily constructed. Let us define

$$w_n(i,x) = J_n(i,x;\mathbf{0}), \tag{3.29}$$

where $\mathbf{0} = \{0, 0, \ldots\}$ is the policy that never orders anything. Then, since no production costs are incurred in view of Assumption (i), we have

$$w_n(i,x) = f_n(i,x) + \mathsf{E}\Big[\sum_{k=n+1}^{\infty} \alpha^{k-n} f_k(i_k, x-(\xi_n+\xi_{n+1}+\cdots+\xi_{k-1})\Big| i_n = i\Big]. \quad (3.30)$$

LEMMA 3.1 *$w_n(i,x)$ is well defined and $w_n(i,x) \in \mathbb{L}^\gamma$.*

Proof. On account of $f_n(i,x) \in \mathbb{L}^\gamma$, it is sufficient to show that

$$\mathsf{E}\Big[\sum_{k=n+1}^{\infty} \alpha^{k-n} f_k(i_k, x-(\xi_n+\xi_{n+1}+\cdots+\xi_{k-1}))\Big| i_n = i\Big] \in \mathbb{L}^\gamma.$$

Assumption (ii) yields

$$\begin{aligned}
&\mathsf{E}\Big[\sum_{k=n+1}^{\infty} \alpha^{k-n} f_k(i_k, x-(\xi_n+\xi_{n+1}+\cdots+\xi_{k-1}))\Big| i_n = i\Big] \\
&\le \bar{f} \sum_{k=n+1}^{\infty} \alpha^{k-n}(1+\mathsf{E}\Big[|x-(\xi_n+\xi_{n+1}+\cdots+\xi_{k-1})|^\gamma\Big| i_n = i\Big]) \\
&= \frac{\bar{f}\alpha}{1-\alpha} + \bar{f}\sum_{k=n+1}^{\infty} \alpha^{k-n}\mathsf{E}\Big[|x-(\xi_n+\xi_{n+1}+\cdots+\xi_{k-1})|^\gamma \,\Big|\, i_n = i\Big].
\end{aligned}$$

Note that it follows from (3.1) that $\mathsf{E}(\xi_k^\gamma | i_k = i) \le D+1$ for all k and i. Let $\overline{M}_x = \max\{|x|^\gamma, D+1\}$. Now let us consider the argument of the sum for a fixed $k \ge n+1$. Let $\Pi_k := \{(i_n,\ldots,i_k) : i_n,\ldots,i_k \in \mathbb{I}$ and $i_n = i\}$ be the set of all combinations of demand states in periods n through k for a given initial state $i_n = i$. Note that for a given sequence of demand states, the one-period demands are independent. Then we have,

$$\begin{aligned}
&\mathsf{E}\Big[|x-(\xi_n+\cdots+\xi_{k-1})|^\gamma \,\Big|\, i_n = i\Big] \\
&= \sum_{\pi\in\Pi_k} \mathsf{E}\Big[|x-(\xi_n+\cdots+\xi_{k-1})|^\gamma \,\Big|\, (i_n,\ldots,i_k) = \pi\Big] \\
&\qquad \times \mathsf{P}((i_n,\ldots,i_k) = \pi) \\
&\le \sum_{\pi\in\Pi_k} (k-n+1)^\gamma \mathsf{E}\Big[|x|^\gamma + |\xi_n|^\gamma + \cdots + |\xi_{k-1})|^\gamma \,\Big|\, (i_n,\ldots,i_k) = \pi\Big] \\
&\qquad \times \mathsf{P}((i_n,\ldots,i_k) = \pi) \\
&\le \sum_{\pi\in\Pi_k} (k-n+1)^\gamma (n-k)\bar{M}_x \mathsf{P}((i_n,\ldots,i_k) = \pi) \\
&= (k-n+1)^\gamma (k-n)\bar{M}_x.
\end{aligned}$$

Therefore,

$$\begin{aligned}&\mathsf{E}\Big[\sum_{k=n+1}^{\infty}\alpha^{k-n}f_k(i_k,x-(\xi_n+\cdots+\xi_{k-1}))\Big|i_n=i\Big]\\&\le\ \bar{f}\Big(\frac{\alpha}{1-\alpha}+\bar{M}_x\sum_{k=n+1}^{\infty}\alpha^{k-n}(k-n+1)^{\gamma}(k-n)\Big)<\infty.\end{aligned}$$

Consequently, $w_n(i,x)<\infty$. In view of Theorem A.1.8, $w_n(i,x)$ is also l.s.c. as the sum of nonnegative l.s.c. functions. Moreover, because $\bar{M}_x = |x|^{\gamma}$ for $|x|^{\gamma}\ge D+1$, $w_n(i,x)$ is at most of polynomial growth with power γ. Thus, we have $w_n(i,x)\in\mathbb{L}^{\gamma}$. □

We can now state the following result for the infinite horizon problem.

THEOREM 3.3 *Let Assumptions* (i)-(ii) *and* (3.1) *hold. Then, we have*

$$0=v_{n,0}\le v_{n,1}\le\ldots\le v_{n,m}\le w_n \tag{3.31}$$

and

$$v_{n,m}\uparrow v_n\in\mathbb{B}^{\gamma}, \tag{3.32}$$

where v_n is a solution of (3.24) *in $\mathbb{L}^{\gamma}$. Furthermore, there exists $\hat{U}=\{\hat{u}_n,\hat{u}_{n+1},\ldots\}$ for which the infimum in* (3.24) *is attained, and $\hat{U}$ is an optimal feedback policy, i.e.,*

$$v_n(i,x)=\min_U J_n(i,x;U)=J_n(i,x;\hat{U}). \tag{3.33}$$

Proof. By definition, $v_{n,0}=0$. Let $\tilde{U}_{n,m}=\{\tilde{u}_n,\tilde{u}_{n+1},\ldots,\tilde{u}_{n+m-1}\}$ be a minimizer of (3.25). Thus,

$$\begin{aligned}w_n(i,x)&=J_n(i,x;\mathbf{0})\ge J_{n,m}(i,x;\mathbf{0})\\&\ge v_{n,m}(i,x)=J_{n,m}(i,x;\tilde{U}_{n,m})\ge J_{n,m-1}(i,x;\tilde{U}_{n,m-1})\\&\ge \min_U J_{n,m-1}(i,x;U)=v_{n,m-1}(i,x).\end{aligned}$$

This proves (3.31). Moreover, it follows from (3.31) that there is a function $v_n(i,x)$ such that

$$v_{n,m}(i,x)\uparrow v_n(i,x)\le w_n(i,x). \tag{3.34}$$

Next, we will show that the functions v_n satisfy the dynamic programming equations (3.24). Observe from (3.27) and (3.31) that for each m, we have

$$v_{n,m}(i,x)\le f_n(i,x)+\inf_{u\ge 0}\{c_n(i,u)+\alpha F_{n+1}(v_{n+1,m})(i,x+u)\}.$$

Thus, in view of (3.34), we can replace $v_{n+1,m}$ by v_n on the RHS of the above inequality and then pass to the limit on the LHS as $m \to \infty$ to obtain

$$v_n(i,x) \le f_n(i,x) + \inf_{u\ge 0}\{c_n(i,u) + \alpha F_{n+1}(v_{n+1})(i,x+u)\}. \tag{3.35}$$

Let the infimum in (3.27) be attained at $\hat{u}_{n,m}$. In order to obtain the reverse inequality, we first prove that $\hat{u}_{n,m}(i,x)$ is uniformly bounded with respect to m.

In the proof of Theorem 3.1 we showed that for $|x| \le M$, there is an $\bar{u}^M_{n,i}$ such that $\hat{u}_{n,N-n}(i,x) \le \bar{u}^M_{n,i}$. Furthermore, if we replace v_{n+1} by w_{n+1} in (3.12) and follow the same line of arguments as in the proof of Theorem 3.1, we obtain an upper bound $\bar{\bar{u}}^M_{n,i}$, which does not depend on the horizon N. Therefore, we can conclude that

$$\bar{u}_n(i,x) := \bar{\bar{u}}^M_{n,i} \text{ is an upper bound for } \hat{u}_{n,m}(i,x) \text{ if } M-1 < |x| \le M, \tag{3.36}$$

independent of m.

For $l > m$, we see from (3.27)that

$$\begin{aligned} v_{n,l+1}(i,x) &= f_n(i,x) + c_n(i,\hat{u}_{n,l}(x)) \\ &\qquad + \alpha F_{n+1}(v_{n+1,l})(i,x+\hat{u}_{n,l}(x)) \\ &\ge f_n(i,x) + c_n(i,\hat{u}_{n,l}(x)) \\ &\qquad + \alpha F_{n+1}(v_{n+1,m})(i,x+\hat{u}_{n,l}(x)). \end{aligned} \tag{3.37}$$

Fix m and let $l \to \infty$. In view of (3.36), we can choose a sequence of periods l' such that

$$\hat{u}_{n,l'}(i,x) \to \tilde{u}_n(i,x) \text{ for } l' \to \infty. \tag{3.38}$$

We then conclude from (3.37) and Fatou's Lemma (Lemma B.1.1) that

$$\begin{aligned} v_n(i,x) &\ge f_n(i,x) + \lim_{l'\to\infty} c_n(i,\hat{u}_{n,l'}(i,x)) \\ &\qquad + \alpha \liminf_{l'\to\infty} F_{n+1}(v_{n+1,m})(i,x+\hat{u}_{n,l'}(i,x)), \\ &\ge f_n(i,x) + \lim_{l'\to\infty} c_n(i,\hat{u}_{n,l'}(i,x)) \\ &\qquad + \alpha \sum_{j=1}^{L} p_{ij} \int_0^\infty (\liminf_{l'\to\infty} v_{n+1,m}(i,x+\hat{u}_{n,l'}(i,x)-\xi))d\Phi_{i,n}(\xi). \end{aligned}$$

Since $v_{n+1,m}$ and c_n are l.s.c., we can, in view of (3.38), pass to the limit in the argument of these functions to obtain

$$\begin{aligned} v_n(i,x) &\ge f_n(i,x) + c_n(i,\tilde{u}_n(i,x)) + \alpha F_{n+1}(v_{n+1,m})(i,x+\tilde{u}_n(i,x)), \\ &\ge f_n(i,x) + \inf_{u\ge 0}\{c_n(i,u) + \alpha F_{n+1}(v_{n+1,m})(i,x+u)\}. \end{aligned} \tag{3.39}$$

This, along with (3.35) and (3.34), proves (3.32).

From Theorem A.1.8, it follows that v_n is l.s.c., as the monotone limit of l.s.c. functions $v_{n,m}$ as $m \to \infty$. Also, since v_n is bounded by a function w_n of polynomial growth, we have $v_n \in \mathbb{L}^\gamma$.

Because $w_n \geq v_n \geq f_n$, it is clear that $\bar{u}_n(i,x)$ defined in (3.36) is also an upper bound for the minimizer in (3.24). Therefore, there exists a Borel map $\hat{u}_n(i,x)$ such that

$$\begin{aligned} & c_n(i, \hat{u}_n(i,x)) + \alpha F_{n+1}(v_{n+1})(i, x + \hat{u}_n(i,x)) \\ & \quad = \inf_{u \geq 0} \{c_n(i,u) + \alpha F_{n+1}(v_{n+1})(i, x+u)\}. \end{aligned} \tag{3.40}$$

With that in mind, we can use (3.24) to obtain

$$\begin{aligned} & \mathsf{E}[v_k(i_k, x_k)] \\ & \quad = \mathsf{E}[f_k(i_k, x_k) + c_k(i_k, \hat{u}_k)] + \alpha \mathsf{E}[F_{k+1}(v_{k+1})(i_k, x_k + \hat{u}_k)] \\ & \quad = \mathsf{E}[f_k(i_k, x_k) + c_k(i_k, \hat{u}_k] + \alpha \mathsf{E}[v_{k+1}(i_{k+1}, x_{k+1})], k = 0, 1, 2, \ldots. \end{aligned}$$

Multiplying by α^{k-n}, summing from n to $N-1$, and canceling terms yield

$$v_n(i,x) \geq \mathsf{E}\Big[\sum_{k=n}^{N-1} \alpha^{k-n}(c_k(i_k, \hat{u}_k) + f_k(i_k, \hat{x}_k))\Big] + \alpha^{N-n}\mathsf{E}[v_N(i_N, \hat{x}_N)].$$

Letting $N \to \infty$, we conclude

$$v_n(i,x) \geq J_n(i,x;\hat{U}). \tag{3.41}$$

From Theorem 3.2 we know that $v_{n,k}(i,x) \leq J_{n,k}(i,x;U)$ for any policy U, and we let $k \to \infty$ to obtain

$$v_n(i,x) \leq J_n(i,x;U) \text{ for any admissible } U. \tag{3.42}$$

Together, inequalities (3.41) and (3.42) imply

$$v_n(i,x) = J_n(i,x;\hat{U}) = \min_U J_n(i,x;U),$$

which completes the proof. □

Before we prove the optimality of an (s,S)-type policy for the nonstationary finite and infinite horizon problems, we should note that Theorem 3.3 does not imply uniqueness of the solution to the dynamic programming equations (3.27) and (3.28). There may be other solutions. Moreover, one can show that the value function is the minimal positive solution of (3.27) and (3.28). It is also possible to obtain a uniqueness proof under additional assumptions. For our purpose, however, it is sufficient to have the results of Theorem 3.3.

3.5. Optimality of (s, S)-type Ordering Policies

The existence and optimality of a *feedback* (or Markov) policy $\hat{u}_n(i,x)$ was proved in Theorems 3.1 and 3.2. We now make additional assumptions to further characterize the optimal feedback policy.

Let us assume that for any demand state i,

$$f_n(i,x) \qquad \text{is convex with respect to } x, \quad n = 0,1,\ldots,N, \tag{3.43}$$

$$c_n(i,u) = \begin{cases} 0, & u = 0, \\ K_n^i + c_n^i \cdot u, & u > 0, \end{cases} \tag{3.44}$$

where $c_n^i \geq 0$ and $K_n^i \geq 0$, $n = 0,1,\ldots,N-1$, and

$$K_n^i \geq \alpha \bar{K}_{n+1}^i \equiv \alpha \sum_{j=1}^{L} p_{ij} K_{n+1}^j, \quad n = 0,1,\ldots,N. \tag{3.45}$$

It should be noted that (3.43) implies that $f_n(i,\cdot)$, for any i and $n = 0,1,\ldots,N$, is continuous on $\mathbb{R}$.

Remark 3.2 Assumptions (3.43)–(3.45) reflect the usual structure of costs to prove optimality of an (s,S)-type policy.

Theorem 3.4 *Let N be finite. Let Assumptions* (i)-(iii), (3.1), (3.9), *and* (3.43)-(3.45) *hold. Then, there exists a sequence of numbers $s_{n,i}$, $S_{n,i}$, $n = 0,\ldots,N-1, i = 1,\ldots,L$, with $s_{n,i} \leq S_{n,i}$, such that the optimal feedback policy is*

$$\hat{u}_n(i,x) = \begin{cases} S_{n,i} - x, & x \leq s_{n,i}, \\ 0, & x > s_{n,i}. \end{cases} \tag{3.46}$$

Proof. The dynamic programming equations (3.10) and (3.11) can be written as

$$\begin{aligned} v_n(i,x) &= f_n(i,x) - c_n^i x + h_n(i,x), & 0 \leq n \leq N-1,\ i = 1,\ldots,L, \\ v_N(i,x) &= f_N(i,x), & i = 1,\ldots,L, \end{aligned}$$

where

$$h_n(i,x) = \inf_{y \geq x} [K_n^i \mathbb{1}_{y>x} + z_n(i,y)], \tag{3.47}$$

$$z_n(i,y) = c_n^i y + \alpha F_{n+1}(v_{n+1})(i,y). \tag{3.48}$$

From (3.10), we have $v_n(i,x) \geq f_n(i,x)$. This inequality, along with (3.9), ensures for $n = 1,2,\ldots,N-1$ and $i = 1,2,\ldots,L$ that

$$z_n(i,x) \to +\infty \text{ as } x \to \infty. \tag{3.49}$$

Furthermore, from the last part of Theorem 3.1, it follows that v_{n+1} is continuous; therefore $z_n(i,x)$ is continuous.

In order to obtain (3.46), we need to prove that $z_n(i,x)$ is K_n^i-convex. This is done by induction. First, $v_N(i,x)$ is convex by definition and therefore, K-convex for any $K \geq 0$. Let us now assume that for a given $n \leq N-1$ and i, $v_{n+1}(i,x)$ is K_{n+1}^i-convex. By Assumption (3.45), it is easy to see that $z_n(i,x)$ is $\alpha\bar{K}_{n+1}^i$-convex, hence also K_n^i-convex. Then, Theorem C.2.3 implies that $h_n(i,x)$ is K_n^i-convex. Therefore, $v_n(i,x)$ is K_n^i-convex. This completes the induction argument.

Thus, it follows that $z_n(i,x)$ is K_n^i-convex for each n and i. In view of (3.49), we apply Theorem C.2.3 to obtain the desired $s_{n,i}$ and $S_{n,i}$. According to Theorem 3.2 and the continuity of z_n, the optimal feedback policy defined in (3.46) is optimal. □

THEOREM 3.5 *Let Assumptions* (i)-(ii), (3.1), (3.9), *and* (3.43)–(3.45) *hold for the cost functions for the infinite horizon problem. Then, there exists a sequence of numbers $s_{n,i}, S_{n,i}$, $n = 0, 1, \ldots$, with $s_{n,i} \leq S_{n,i}$ for each $i \in \mathbb{I}$, such that the feedback policy*

$$\hat{u}_n(i,x) = \begin{cases} S_{n,i} - x, & x < s_{n,i}, \\ 0, & x \geq s_{n,i}, \end{cases} \tag{3.50}$$

is optimal.

Proof. Let v_n denote the value function. Define the functions z_n and h_n as above. We know that $z_n(i,x) \to \infty$ as $x \to +\infty$ and $z_n(i,x) \in \mathbb{L}^\gamma$ for all n and $i = 1, 2, \ldots, L$.

We now prove that v_n is K_n-convex. Using the same induction as in the proof of Theorem 3.4, we can show that $v_{n,m}(i,x)$, defined in (3.26), is K_n^i-convex. This induction is possible since we know that $v_{n,m}(i,x)$ satisfies the dynamic programming equations (3.27) and (3.28). It is clear from the definition of K-convexity that this property is preserved under monotone limit procedures. Thus, the value function $v_n(i,x)$, which is the limit of $v_{n,m}(i,x)$ as $m \to \infty$, is K_n^i-convex.

From Theorem 3.3 we know that v_n satisfies the dynamic programming equations (3.27) and (3.28). Therefore, we can obtain an optimal feedback policy $\hat{U} = \{\hat{u}_n, \hat{u}_{n+1}, \ldots\}$, for which the infimum in (3.27) is attained. Because z_n is K_n^i-convex and l.s.c., $\hat{u}_n$ can be expressed as in (3.50). □

REMARK 3.3 It is important to emphasize the difference between the (s,S) policies defined in (3.46) and (3.50). In (3.46), an order is placed when the inventory level is s or below, whereas in (3.50) an order is placed only when the inventory is strictly below s. Most of the literature uses the policy type (3.46). While (3.46) in Theorem 3.4 can be replaced by (3.50) on account of the continuity of z_n, it is not possible to replace (3.50) in Theorem 3.5 by (3.46), since z_n is proved only to be l.s.c.

REMARK 3.4 In the stationary infinite horizon discounted cost case discussed in the next section, we are able to prove that the value function is locally Lipschitz, and therefore continuous. The proof is provided in Chapter 5, Lemma 5.3. Thus, in this case, policies of both types (3.50) and (3.46) are optimal.

3.6. Stationary Infinite Horizon Problem

If the cost functions, as well as the distributions of the demands, do not explicitly depend on time, i.e., for each k

$$c_k(i,u) = c(i,u), \quad f_k(i,x) = f(i,x), \quad \text{and} \quad \Phi_{i,k} = \Phi_i,$$

then it can be easily shown that the value function $v_n(i,x)$ does not depend on n. In what follows, we will denote the value function of the stationary discounted cost problem by $v^\alpha(\cdot,\cdot)$, in order to emphasize the dependence on the discount factor α. In the same manner as in Section 3.4, it can be proved that the function v^α satisfies the dynamic programming equation

$$v^\alpha(i,x) = f(i,x) + \inf_{u\geq 0}\{c(i,u) + \alpha F(v^\alpha)(i,x+u)\}, \tag{3.51}$$

where F is the same as F_{n+1}, defined in (3.8), i.e.,

$$Fb(i,y) = \sum_{j=1}^{L} p_{ij}\int_0^\infty b(j,y-\xi)d\Phi_i(\xi),$$

for $b \in \mathbb{B}^\gamma$.

Furthermore, for any α, $0 < \alpha < 1$, there is a *stationary optimal feedback policy* $U^\alpha = (u^\alpha(i,x), u^\alpha(i,x), \ldots)$, where $u^\alpha(i,x)$ is the minimizer on the RHS of (3.51). Moreover, if the cost functions also satisfy the Assumptions (3.43)–(3.45) introduced in Section 3.5, then we can obtain pairs (s_i^α, S_i^α) such that either of the (s_i^α, S_i^α)-policies of types (3.50) and (3.46) is optimal; (see Remark 3.4).

3.7. Concluding Remarks and Notes

This chapter, based on Beyer and Sethi (1997) and Beyer *et al.* (1998), generates the discounted infinite horizon inventory model involving fixed costs that have appeared in the literature, to allow for unbounded demand and costs with polynomial growth. We have shown the existence of an optimal Markov policy, and that this can be a state-dependent (s, S) policy.

This chapter makes several specific contributions. It extends the proofs of existence and verification of optimality in the discounted cost case given in Chapter 2, to allow for more general costs including l.s.c. surplus cost with polynomial growth.

Some problems of theoretical interest remain open. One might want to show that the value function in the discounted nonstationary infinite horizon case is continuous if the surplus cost function is continuous.