PARTIALLY OBSERVED INVENTORY SYSTEMS: THE CASE OF RAIN CHECKS

ALAIN BENSOUSAN∗, METIN ÇAKANYILDıRıM †, J. ADOLFO MINJÁREZ-SOSA‡, SURESH P. SETHI §, AND RUIXIA SHI ¶

Abstract. In many inventory control contexts, inventory levels are only partially (i.e., not fully) observed. This may be due to non-observation of demand, spoilage, misplacement, or theft of inventory. We study a discrete-time periodic-review inventory system where the unmet demand is backordered. When inventory level is nonnegative, the inventory manager does not know the exact inventory level. Otherwise, inventory shortages occur and the inventory manager issues rain checks to customers. The shortages are fully observed via the rain checks. The inventory manager determines the order quantity based on the partial information on the inventory level. The objective is to minimize the expected total discounted cost over an infinite horizon. The dynamic programming formulation of this problem has an infinite dimensional state space. We use the methodology of the unnormalized probability to establish the existence of an optimal feedback policy when the periodic cost has linear growth. Moreover, uniqueness and continuity of the solution to dynamic programming equations are proved when the discount factor is sufficiently small.

Key words. stochastic inventory problem, partial observations, the Zakai equation, rain check

AMS subject classifications. 90B05, 93E20, 93C41, 90C39

1. Introduction. Most inventory models assume that the inventory levels at any given time are fully observed. Under this assumption, some of the most celebrated results, such as the optimality of the base stock policy [1], have been obtained. However, in reality, the inventory levels are often only partially observed. In such cases, most of the well-known inventory policies are not even admissible, let alone optimal.

There are a number of reasons for partial observability of inventory levels. We only list some, as they have been discussed in detail in Bensoussan et al. [3] (referred to as BCS hereafter) and references therein. Demand may be incorrectly observed or observed with some delay. Inventory may be misplaced or stolen. Some products such as groceries and drugs deteriorate over time, and are no longer fit for sale after a period of time. In addition, the saleable quantities received may be different from those ordered because of uncertain yields.

Even though partial observations in inventory systems are very common, there has not been much research activity in this area. The main reason may be the mathematical difficulty. While a finite dimensional space suffices to accommodate the system state in the full observation case, an infinite dimensional state space is required in the partial observation setting. More specifically, the inventory level at a given time is no longer a state in \( \mathbb{R}^n \), it must now be represented by its

∗International Center for Decision and Risk Analysis, School of Management, P.O. Box 830688, SM 30, University of Texas at Dallas, Richardson, TX 75083-0688 (alain.bensoussan@utdallas.edu).
†School of Management, P.O. Box 830688, SM 30, University of Texas at Dallas, Richardson, TX 75083-0688 (metin@utdallas.edu).
‡Departamento de Matemáticas, Universidad de Sonora, Rosales s/n, Centro, 83000 Hermosillo, Sonora, Mexico (aminjare@gauss.mat.uson.mx).
§Center for Intelligent Supply Networks, School of Management, P.O. Box 830688, SM 30, University of Texas at Dallas, Richardson, TX 75083-0688 (sethi@utdallas.edu).
¶School of Management, P.O. Box 830688, SM 30, University of Texas at Dallas, Richardson, TX 75083-0688 (ruixia.shi@utdallas.edu).
conditional probability given the partial observations available at that time. Thus, the analysis takes place in the space of probability distributions.

When there is no physical inventory, i.e., the inventory is zero or negative, then none of the following would happen: transaction errors, misplaced inventories, spoilage, or theft. Most companies pay utmost attention to an item when its inventory reaches zero. In these companies, employees walk around the shelves to see if the item is stocked out. This process is implemented at the office supply store Staples, and is called “zero-balance walk” in [9] and [12]. BCS study this situation, in which the positive inventory is represented by a conditional distribution, no backordering is allowed, and an empty shelf is noticed when the inventory reaches zero. BCS were the first to introduce a methodology to handle the difficulty of dealing with partially observed discrete-time inventory systems.

This paper extends BCS to the case when backordering is allowed, i.e. when sales are not lost. It is very common in reality that a store issues rain checks to a customer when there is a stockout. A rain check is an assurance to a customer that a sold out item can be purchased later at the same price. Usually, the inventory manager (IM) monitors the issuing of the rain checks. Thus, in the inventory systems with rain checks, the negative inventory level is fully observed by the IM. It is the purpose of this paper to formulate and analyze the partially observed inventory system with rain checks. However, compared to the work of BCS, where the inventory level is observed when it is zero, the full observation of negative inventory in the present paper makes the analysis much more difficult. Indeed, in the model of BCS, the optimality equation is represented by a pair formed by a constant and a function of the conditional distribution of the inventory level. The constant corresponds the case of zero inventory (observable term), while the function represents the optimality equation with positive inventory (non-observable term). However, in the present model, instead of a constant, the observable term via rain checks is represented by a function of the backordered amount. Therefore, the present study must consider suitable functional spaces to define appropriate dynamic programming operators, which necessitate a more complicated mathematical analysis.

It is worth emphasizing that even though the study of the rain check model is more involved than that of the zero-balance walk model, we are able to develop an alternative approach that allows us to improve the results in BCS. BCS show the existence of a solution to the dynamic programming equations for the problem by obtaining a decreasing sequence of solutions from a value iteration scheme that converge to the value function. Furthermore, the existence of an optimal ordering policy in BCS is obtained under the condition of bounded ordering quantities or with a sufficiently small discount factor. We, on the other hand, do not require these conditions to get the existence of an optimal feedback ordering policy. This is accomplished by first proving the convergence of an increasing sequence of solutions, obtained from a value iteration algorithm, to the value function, and then using a selection theorem to establish the existence of an optimal feedback policy. Furthermore, for a sufficiently small discount factor, and using a decreasing approximation scheme as in BCS, we show in addition that the value function is continuous and is the unique solution of the optimality equation. Moreover, the value iteration procedure developed in this paper is a valid approximation scheme for any discount factor, whereas the BCS scheme is restricted to a sufficiently small discount factor.

The mathematical treatment for our problem is similar to those applying to the standard partially observed stochastic control problems. This consists of introducing
a filtering process for the partially observed variable, and then the partially observed problem is transformed into an equivalent fully observed problem where the filter becomes the new system state. In our case, the filtering process is defined by the conditional density of the inventory level given the observed history. The filters are obtained recursively in Theorem 2. Besides the difficulties associated with the infinite dimensional state space of the new problem, the equation defining the filtering process is highly nonlinear, which leads to a nontrivial problem. However, by introducing what is known as the unnormalized probability, we transform this equation to a linear one. Therefore, our approach finally consists of using an unnormalized filtering process, which allows us to easily prove our main results.

The rest of the paper is organized as follows. In the next section we describe our model and obtain the evolution equation for the inventory distribution, which is linearized by introducing the unnormalized probability. In Section 3, we formulate the problem and provide the dynamic programming equations for both normalized and unnormalized probabilities. In Section 4 we present our results, and prove them in Section 5.

2. Model Development. We study an infinite horizon, single-item, discrete-time periodic-review inventory system in which the inventory level is partially observed due to unobservable demand. We assume that the unmet demand is backordered. Because of the rain checks, the IM knows exactly the backordered quantity. But when the inventory is nonnegative, he knows only the probability distribution of the inventory level. Thus, the IM only partially observes the inventory level.

We divide the time line into intervals of equal length. Each interval is called a period. In practice, a period can be one week, one month, etc. The sequence of events in any given period \( t \) is as follows: At the beginning of the period (before the demand occurs), the IM observes if the inventory level is negative or non-negative. If the inventory level is nonnegative, he has the probability distribution of the inventory level based on prior observations. Otherwise, the IM knows exactly the backordered quantity. Then, the IM determines how much to order, and the order is delivered immediately. Next the demand occurs, the IM’s order can be used to meet the demand of period \( t \). But the demand is not observed by the IM unless the inventory level falls below zero. In each period, the IM incurs inventory related costs.

We let \( I_t \) denote the inventory level at the beginning of period \( t \), \( q_t \) denote the order quantity determined by the IM at that time taking values in \( \mathbb{R}^+ := [0, \infty) \), and \( D_t \) denote the random demand occurring in period \( t \). We assume that \( D_t, t = 1,2,3\ldots \) are nonnegative, independently and identically distributed random variables, and let the generic demand be denoted by \( D \geq 0 \) with the density and distribution functions \( f \) and \( F \), respectively. We assume that \( E(D) < \infty \). Unsatisfied demand is fully backordered, so that the evolution of inventory is given by

\[
I_{t+1} = I_t + q_t - D_t \quad \text{for } t \geq 1. \tag{1}
\]

\( I_t \) can be either negative or nonnegative. If \( I_t \) is negative, the demand in period \( t-1 \) is not fully met and the negative part of \( I_t \) represents the backordered amount. Let \( z_t \) denote the backordered amount in period \( t \), then \( z_t = (I_t)^- \). If \( I_t \) is nonnegative, obviously it is the ending inventory carried over from period \( t-1 \). But the IM cannot tell exactly how much it is. So what he can do is to evolve the conditional distribution of the inventory level.

In our model, for period \( t \geq 1 \), the IM can only observe

\[
z_t = (I_t)^- \quad \text{for } t \geq 1, \tag{2}
\]
where \((x)^- = \max\{0, -x\}\). If \(z_t > 0\), it is the backordered quantity. If \(z_t = 0\), the inventory level at the beginning of period \(t\) is nonnegative and a distribution of the inventory level can be derived based on prior observations. The IM determines the order quantity \(q_t\) for each period \(t\). The order quantity \(q_t\) is adapted to the sigma field \(\mathcal{Z}_t := \sigma(\{z_j : 1 \leq j \leq t\})\). Thus, \(\mathcal{Z}_t\) is the history available to the IM in period \(t\). By an ordering policy (or simply a policy), we mean a sequence \(\tilde{q}\) of \(\mathbb{R}^+\)-valued random variables such that \(q_t\) is \(\mathcal{Z}_t\)-measurable for each \(t \geq 1\). Let \(\Gamma\) be the set of such policies.

When the demand is met entirely in period \(t - 1\), inventory holding costs are incurred on the remaining inventory carried from period \(t - 1\) to period \(t\). Otherwise, there are backorder costs. Then, given the cost function \(c(I_t, q_t)\) depending only on the inventory level \(I_t\) and the order size \(q_t\) in period \(t, \ t = 1, 2, 3\ldots\) and given the policy \(\tilde{q} \in \Gamma\), the total expected discounted cost can be written as

\[
J(\zeta, \pi, \tilde{q}) := E \sum_{t=1}^{\infty} \alpha^{t-1} c(I_t, q_t),
\]

where \(0 < \alpha < 1\) is the discount factor. The pair \((\zeta, \pi(.))\) is the inventory state at the beginning of period 1. When \(\zeta > 0\), \(\zeta\) is the backordered quantity at the beginning of period 1. If \(\zeta = 0\), the inventory level \(I_1\) is nonnegative and \(\pi \in \Pi\) is the probability density of \(I_1\), where \(\Pi\) is the set of density function such that \(\int x \pi(x) dx < \infty\). Observe that \(f \in \Pi\).

We want to find a policy \(\tilde{q} \in \Gamma\) that minimizes \(J(\zeta, \pi, \tilde{q})\). When \(\zeta > 0\), \(\pi\) is of no consequence and \(J(\zeta, \pi, \tilde{q})\) does not depend on \(\pi\), which will be noted explicitly below.

### 2.1. Evolution of State Probabilities

At the beginning of the first period, a prior probability density \(\pi \in \Pi\) of nonnegative inventory is given. From the second period on, the IM needs to derive the probability distribution function for nonnegative inventory to use in his decision process. In what follows, we show how these distribution functions evolve over time.

It is convenient to introduce the indicator random variables

\[
\mathbb{I}_{z_t = 0} \equiv \mathbb{I}_{I_t^- = 0} = \mathbb{I}_{I_t \geq 0} \quad \text{for} \quad t \geq 1.
\]  \tag{3}

When there is no backorder, i.e., \(z_t = I_t^- = 0\) or \(I_t \geq 0\), \(\mathbb{I}_{z_t = 0} = 1\). Otherwise, \(\mathbb{I}_{z_t = 0} = 0\). The indicator random variable \(\mathbb{I}_{z_t = 0}\) is a discrete-time Markov chain with the state space \(\{0, 1\}\): 1 means there is no backorder and 0 means there is.

Let \(\pi_t(.)\) be the conditional density of \(I_t\) given \(\mathcal{Z}_{t-1}\) and \(I_t \geq 0\). That is,

\[
P(I_t \leq x | \mathcal{Z}_{t-1}, I_t \geq 0) = \int_0^x \pi_t(y) dy.
\]

For any bounded real test function \(\varphi(.)\), the conditional Bayes Theorem gives

\[
\int_0^\infty \varphi(x) \pi_t(x) dx = E[\varphi(I_t)|\mathcal{Z}_{t-1}, I_t \geq 0] = \frac{E[\varphi(I_t) \mathbb{I}_{I_t \geq 0} | \mathcal{Z}_{t-1}]}{P(I_t \geq 0 | \mathcal{Z}_{t-1})} = \frac{E[\varphi(I_t) \mathbb{I}_{I_t \geq 0} | \mathcal{Z}_{t-1}]}{P(I_t \geq 0 | \mathcal{Z}_{t-1})}
\]  \tag{4}

\footnote{See Remark 2.3 in [7] for inventory holding cost accounting based on ending or beginning inventory levels.}
In order to obtain a recursive expression for \( \pi_t \) in terms of \( \pi_{t-1} \), we express \( E(\varphi(I_t) | Z_t) \) in terms of the conditional density \( \pi_t \) in the following lemma.

**Lemma 1.**

\[
E(\varphi(I_t) | Z_t) = \mathbb{1}_{z_t > 0} \varphi(-z_t) + \mathbb{1}_{z_t = 0} \int_0^\infty \varphi(\eta) \pi_t(\eta) d\eta.
\]  

(5)

From this lemma we can derive the density function \( \pi_t \) as stated in the following theorem.

**Theorem 2.** For \( t \geq 2 \), the conditional density \( \pi_t \) can be expressed recursively as follows:

\[
\pi_t(x) = \mathbb{1}_{z_{t-1} > 0} \left\{ \frac{f(-z_{t-1} + q_{t-1} - x) \mathbb{1}_{z_{t-1} + q_{t-1} - x} \geq 0}{F(q_{t-1} - z_{t-1})} \right\}
+ \mathbb{1}_{z_{t-1} = 0} \int_{(x-q_{t-1})^+}^\infty f(y + q_{t-1} - x) \pi_{t-1}(y) dy.
\]  

(6)

For \( t = 1 \), \( \pi_1(x) = \pi(x) \).

So, \( \pi_t \) evolves according to a highly nonlinear equation which corresponds to the Kushner equation [11] in our inventory context.

We can use the following method to linearize (6). We define the sequence of functions \( \{p_t\} \) by the recursive linear equation

\[
p_t(x) = \mathbb{1}_{z_{t-1} > 0} f(-z_{t-1} + q_{t-1} - x) \mathbb{1}_{z_{t-1} + q_{t-1} - x} \geq 0
+ \mathbb{1}_{z_{t-1} = 0} \int_{(x-q_{t-1})^+}^\infty f(y + q_{t-1} - x) p_{t-1}(y) dy.
\]  

(7)

\[
p_1(x) = \pi(x),
\]

which corresponds to the Zakai equation obtained in the context of systems with diffusions in [2] and [13]. Also,

\[
\lambda_t = \int_0^\infty p_t(x) dx.
\]

Then, we have \( \lambda_1 = 1 \), and for \( t \geq 2 \) from (7),

\[
\lambda_t = \mathbb{1}_{z_{t-1} > 0} F(-z_{t-1} + q_{t-1}) + \mathbb{1}_{z_{t-1} = 0} \int_0^\infty F(q_{t-1} + y) p_{t-1}(y) dy.
\]  

(8)

Moreover,

\[
p_t(x) = \lambda_t \pi_t(x).
\]  

(9)

Clearly, (9) holds for \( t = 1 \). Assuming (9) to hold for any \( t \), we proceed to \( t + 1 \) by multiplying (6) side-by-side by (8) to obtain

\[
\lambda_{t+1} \pi_{t+1}(x) = \mathbb{1}_{z_{t+1} > 0} f(-z_{t+1} + q_{t+1} - x) \mathbb{1}_{z_{t+1} + q_{t+1} - x} \geq 0
+ \mathbb{1}_{z_{t+1} = 0} \int_{(x-q_{t+1})^+}^\infty f(y + q_{t+1} - x) \pi_t(y) dy \int_0^\infty F(q_{t+1} + y) p_t(y) dy
\]

\[
\int_0^\infty F(y + q_{t+1}) \pi_t(y) dy.
\]

By multiplying the numerator and the denominator of the second term on the right-hand side by \( \lambda_t \), we establish (9) for \( t + 1 \).
On account of the weighting factor $\lambda_t$ in (9), $p_t(x)$ can be viewed as an unnormalized probability; see [10] or [13]. We can easily get the normalized probability $\pi_t(x)$ by

$$\pi_t(x) = \frac{p_t(x)}{\int_0^\infty p_t(x)dx}.$$  

To write equations (6) and (7) in the operator form, we define the spaces $\mathcal{H} := \left\{ p \in L^1(\mathbb{R}^+) : \int_0^\infty x|p(x)|dx < \infty \right\}$ and $\mathcal{H}^+ := \left\{ p \in \mathcal{H} : p(x) \geq 0, x \in \mathbb{R}^+ \right\}$, where $L^1(\mathbb{R}^+)$ is the space of integrable functions whose domain is the set of nonnegative real numbers. Note that $\Pi \subseteq \mathcal{H}^+$. Observe that $\mathcal{H}^+$ is not a subspace of $\mathcal{H}$, for it does not include $-p$ for some $p$. In the remainder, we identify $\mathcal{H}^+$ as the set of all unnormalized probabilities with the norm $||p|| := \int_0^\infty |p(x)|dx + \int_0^\infty x|p(x)|dx$.

Since $\Pi \subseteq \mathcal{H}^+$, the norm applies to the normalized probability $\pi$ to say that the inventory level must have a finite mean. For a sequence $\{p_n\}$ in $\mathcal{H}^+$ and $p \in \mathcal{H}^+$, “$p_n \to p$” means $\|p_n - p\| \to 0$ as $n \to \infty$, which is equivalent to

$$\int_0^\infty |p_n(x) - p(x)|dx \to 0 \quad \text{and} \quad \int_0^\infty x|p_n(x) - p(x)|dx \to 0. \quad (10)$$

Let $L$ be the space of functions $\phi$ with linear growth, i.e.,

$$L := \left\{ \phi : \sup_{x \geq 0} \frac{|\phi(x)|}{1 + x} < \infty \right\}$$

with the norm $||\phi||_L := \sup_{x \geq 0} \frac{|\phi(x)|}{1 + x}$.

Furthermore, we define the product

$$\langle p, \phi \rangle := \int_0^\infty p(x)\phi(x)dx \quad \text{for } p \in \mathcal{H}, \; \phi \in \mathcal{L}.$$

Similarly, we define the space $L^+ := \left\{ \phi \in \mathcal{L} : \phi(x) \geq 0, x \in \mathbb{R}^+ \right\}$. We assume that for every fixed $q_t$, the one-period cost $c(I_t, q_t)$ is in $L^+$.

For any order quantity $q$ and the inventory level $y$ before the receipt of the order, we let $w = y + q$ be the inventory level after the receipt. We also let $L(\mathcal{H}, \mathcal{H})$ denote the space of bounded linear maps from $\mathcal{H}$ to $\mathcal{H}$. Then for $y \geq 0$, we can define the linear operator $\varrho(q) \in L(\mathcal{H}, \mathcal{H})$ as

$$\varrho(q)p(x) := \int_{(x-q)^+}^\infty f(y + q - x)p(y)dy.$$  

For $y < 0$, we define

$$\varrho_0(w)(x) := \left\{ \begin{array}{ll} f(w - x) \mathbb{1}_{x \leq w} & \text{if } w \geq 0, \\ 0 & \text{if } w < 0. \end{array} \right.$$
After defining the operators above, we can define the corresponding nonlinear operator \( \theta(q, \cdot) \) and the function \( \theta_0(w) \), respectively, as

\[
\theta(q, p) := \frac{\rho(q)p}{\int_0^\infty F(y + q)p(y)dy},
\]

\[
\theta_0(w) := \frac{\rho_0(w)}{F(w)}.
\]

From the definition of \( \theta(q, p) \), we can see that the operator \( \theta \) is homogenous of degree 0 in \( p \), and it is well defined if \( \langle \rho(q)p, 1 \rangle > 0 \), i.e.,

\[
\langle \rho(q)p, 1 \rangle = \int_0^\infty p(y)F(y + q)dy \neq 0.
\] (11)

With these operators, we can write (6) and (7) in the operator form:

\[
\pi_t = \mathbb{I}_{z_t-1>0} \theta_0(q_{t-1} - z_{t-1}) + \mathbb{I}_{z_t-1=0} \theta(q_{t-1}, \pi_{t-1}),
\] (12)

\[
p_t = \mathbb{I}_{z_t-1>0} \rho_0(q_{t-1} - z_{t-1}) + \mathbb{I}_{z_t-1=0} \rho(q_{t-1})p_{t-1},
\] (13)

with the initial conditions \( \pi_1 = p_1 = \pi \). Once again, we emphasize that (13) is a linear equation, while (12) is not.

For the linear operator \( \rho(q) \), the norm is defined as

\[
||\rho(q)p|| := \sup_{p \in \mathcal{H}} \frac{||\rho(q)p||}{||p||},
\]

where we adopt the convention that \( 0/0 = 0 \), here and throughout. For this norm, we have the following lemma.

**Lemma 3.** \( ||\rho(q)p|| \leq ||p|| + q \int_0^\infty |p(y)| dy \) and \( ||\rho(q)||_\mathcal{L} \leq 1 + q \).

From this lemma we can see that when \( q = 0 \), the operator \( \rho(q)p \) is a contraction mapping. Additional properties of operators \( \rho \) and \( \theta \) are given in BCS.

### 3. Dynamic Programming Formulation

We assume that \( c(I, q) \) is a continuous function of linear growth in \( I \) for every fixed \( q \), i.e., \( c(\cdot, q) \in L^+ \). A cost function with linear growth satisfies the following inequality:

**Linear growth assumption:** \( c(y, q) \leq c_0 + c_1q + hy^+ + sy^- \), \( y \in \mathbb{R} \), (14)

where \( c_1, h, \) and \( s \) are positive constants. The constant \( c_0 \) can be interpreted as the maximum expected backorder cost that can be incurred in a period if the beginning inventory level is zero. Indeed, we set \( c_0 = c(0, 0) \). \( c_0 \) will be bounded by the cost of backordering \( E(D) \) units. If the unit order cost is \( c \), obviously the ordering cost \( cq \) constitutes a lower bound on \( c(I, q) \). Under the linear growth assumption, we have the following lemma.

**Lemma 4.** If \( q_n \to q \) in the Euclidean norm and \( p_n \to p \) in the norm \( ||\cdot|| \) as \( n \to \infty \), then

\[
\int_0^\infty c(y, q_n)p_n(y)dy \to \int_0^\infty c(y, q)p(y)dy.
\]

From this lemma, we can see that the function

\[
(q, p) \to \int_0^\infty c(y, q)p(y)dy
\] (15)
For a fixed $E$, by conditioning on the event $E$, then the cost-to-go can be written as
\[ J(\zeta, \pi, \bar{q}) = \sum_{t=1}^{\infty} \alpha^{t-1} E[c(I_t, q_t) | Z_t] \]
\[ = \sum_{t=1}^{\infty} \alpha^{t-1} E\{ \mathbb{I}_{z_t > 0} c(z_t, q_t) + \mathbb{I}_{z_t = 0} (c(I_t, q_t), \pi_t) \}, \] (16)
where $\pi_t$ is the solution of (6) and $z_1 = \zeta \in \mathbb{R}^+$ is the initial condition. When $\zeta = 0$, the inventory level at the beginning of the first period is nonnegative and the given density function is $\pi_1 = \pi$. When $\zeta > 0$, it is the observed backordered quantity at the beginning of period one.

The value function is defined as
\[ V(\zeta, \pi) := \inf_{\bar{q} \in F} J(\zeta, \pi, \bar{q}). \] (17)

By looking one period ahead from period one, the value function can be written as
\[ V(\zeta, \pi) = \inf_{q} \left\{ \mathbb{I}_{\zeta > 0} c(-\zeta, q) + \mathbb{I}_{\zeta = 0} \int_{0}^{\infty} c(y, q) \pi(y) dy + \alpha E[V(z_2, \pi_2) | \zeta, \pi] \right\}, \] (18)
which is the inventory related cost in period one plus the cost-to-go.

By (1) and (2), we can express
\[ z_2 = (I_1 + q - D_1)^{-} = \mathbb{I}_{\zeta > 0} (-\zeta + q - D_1)^{-} + \mathbb{I}_{\zeta = 0} (I_1 + q - D_1)^{-}. \]

Then, the cost-to-go can be written as
\[ E[V(z_2, \pi_2) | \zeta, \pi] = \mathbb{I}_{\zeta > 0} E[V((-\zeta + q - D)^{-}, \theta_0(q-\zeta))] + \mathbb{I}_{\zeta = 0} E[V((I_1 + q - D)^{-}, \theta(q, \pi))]. \] (19)

By conditioning on the event $[q - \zeta < D]$, we can express the first term on the right-hand side of (19) as
\[ E[V((-\zeta + q - D)^{-}, \theta_0(q-\zeta))] = V(0, \theta_0(q-\zeta)) F(q-\zeta) + \int_{q-\zeta}^{\infty} V(\eta + \zeta - q, \theta_0(q-\zeta)) f(\eta) d\eta. \] (20)

For a fixed $I_1$, we obtain
\[ E[V((I_1 + q - D)^{-}, \theta(q, \pi)) | I_1] = V(0, \theta(q, \pi)) F(I_1 + q) + \int_{q+I_1}^{\infty} V(\eta - I_1 - q, \theta(q, \pi)) f(\eta) d\eta. \] (21)
Using (21), we can express the second term on the right-hand side of (19) explicitly as
\[ I_{\zeta=0} E[V((I_1 + q - D)^-, \theta(q, \pi))] = I_{\zeta=0} V(0, \theta(q, \pi)) \int_0^\infty F(y + q) \pi(y) dy \]
\[ + I_{\zeta=0} \int_0^\infty \pi(y) \left[ \int_{q+y}^\infty V(\eta - y, q, \theta(q, \pi)) f(\eta) d\eta \right] dy \]
\[ = I_{\zeta=0} V(0, \theta(q, \pi)) \int_0^\infty F(y + q) \pi(y) dy \]
\[ + I_{\zeta=0} \int_0^\infty \pi(y) \int_0^\infty V(\eta, \theta(q, \pi)) f(q + y + \eta) d\eta dy. \]

From (20) and the above equation, we obtain
\[ E[V(z_2, \pi_2) | \zeta, \pi] = I_{\zeta>0} \left\{ V(0, \theta_0(q - \zeta)) F(q - \zeta) + \int_0^\infty V(\eta, \theta_0(q - \zeta)) f(\eta + q - \zeta) d\eta \right\} \]
\[ + I_{\zeta=0} V(0, \theta(q, \pi)) \int_0^\infty F(y + q) \pi(y) dy \]
\[ + I_{\zeta=0} \int_0^\infty \pi(y) \int_0^\infty V(\eta, \theta(q, \pi)) f(q + y + \eta) d\eta dy. \]

From the above expression, we can see that when \( \zeta > 0 \), we have
\[ V(\zeta, \pi) = \inf_q \left\{ c(-\zeta, q) + \alpha V(0, \theta_0(q - \zeta)) F(q - \zeta) + \alpha \int_0^\infty V(\eta, \theta_0(q - \zeta)) f(\eta + q - \zeta) d\eta \right\}. \]

Since \( V(\zeta, \pi) \) is independent of \( \pi \) when \( \zeta > 0 \), we can let \( v(\zeta) := V(\zeta, \pi) \).

On the other hand, when \( \zeta = 0 \),
\[ V(0, \pi) = \inf_q \left\{ \int_0^\infty c(y, q) \pi(y) dy \right\} \]
\[ + \alpha V(0, \theta(q, \pi)) \int_0^\infty F(y + q) \pi(y) dy \]
\[ + \alpha \int_0^\infty \pi(y) \int_0^\infty V(\eta, \theta(q, \pi)) f(q + y + \eta) d\eta dy. \]

If we denote \( W(\pi) := V(0, \pi) \), then together with \( v(\zeta) \) we have the system
\[ v(\zeta) = \inf_q \left\{ c(-\zeta, q) + \alpha W(\theta_0(q - \zeta)) F(q - \zeta) + \alpha \int_0^\infty v(\eta) f(\eta + q - \zeta) d\eta \right\}, \text{ for } \zeta \neq 0, \tag{22} \]
\[ W(\pi) = \inf_q \left\{ \int_0^\infty c(y, q) \pi(y) dy \right\} \]
\[ + \alpha \left[ W(\theta(q, \pi)) \int_0^\infty F(y + q) \pi(y) dy + \int_0^\infty \pi(y) \int_0^\infty v(\eta) f(q + y + \eta) d\eta dy \right]. \tag{23} \]

Observe that
\[ V(\zeta, \pi) = I_{\zeta>0} v(\zeta) + I_{\zeta=0} W(\pi), \quad (\zeta, \pi) \in \mathbb{R}^+ \times \Pi. \tag{24} \]
Since $\pi$ evolves according to the nonlinear operator $\theta$, a direct study of the system (22)-(23) is not easy. If we can use $p \in H^+$ as the system state instead of $\pi$, then the analysis becomes much easier, because $p$ evolves with the linear operator $\varrho$. To make ideas concrete, we define a new value function $Y(\cdot)$ as follows:

$$Y(p) := W\left(\frac{p}{\lambda}\right) \lambda, \quad \lambda := \int_0^\infty p(x)dx.$$  \hspace{1cm} (25)

Then, from (23) we have

$$Y(p) = \lambda \inf_q \left\{ \int_0^\infty c(y, q)(p(y)/\lambda)dy + \alpha \left[ W(\theta(q, p/\lambda)) \int_0^\infty F(y + q)(p(y)/\lambda)dy + \int_0^\infty p(y)/\lambda \int_0^\infty v(\eta)f(q + y + \eta)d\eta dy \right] \right\}$$

$$= \inf_q \left\{ \int_0^\infty c(y, q)p(y)dy + \alpha \left[ W(\theta(q, p)) \int_0^\infty F(y + q)p(y)dy + \int_0^\infty p(y)\int_0^\infty v(\eta)f(q + y + \eta)d\eta dy \right] \right\}.$$  

The above equation follows from the fact that the operator $\theta$ is homogenous of degree 0. Next, we simplify the term $W(\theta(q, \pi))$ and $W(\theta_0(q - \zeta))$ on the right-hand sides of the above equation and equation (22), respectively. From the definition of $\theta$ and (11), we get

$$Y(q(p)) = \left\{ \int_0^\infty q(p(x))dx \right\} \left\{ W\left(\frac{q(p)}{\pi}\right) \right\} \left\{ \int_0^\infty F(y + q)p(y)dy \right\} \{W(\theta(q, p))\}.$$  \hspace{1cm} (26)

By the definition of $\theta_0$, we get

$$Y(q_0(q - \zeta)) = W\left(\frac{q_0(q - \zeta)}{F(q - \zeta)}\right) F(q - \zeta) = W(\theta_0(q - \zeta))F(q - \zeta).$$

Using these results in the expression of $Y(p)$, we obtain the following new system of equations:

$$v(\zeta) = \inf_q \left\{ c(-\zeta, q) + \alpha \int_0^\infty v(\eta)f(\eta + q - \zeta)d\eta + \alpha Y(q_0(q - \zeta)) \right\}, \quad \zeta \in \mathbb{R}^+,$$  \hspace{1cm} (27)

$$Y(p) = \inf_q \left\{ \int_0^\infty c(y, q)p(y)dy + \alpha Y(q(p)) + \alpha \int_0^\infty p(y)\int_0^\infty v(\eta)f(y + \eta)d\eta dy \right\}, \quad p \in \mathbb{P}_T.$$  \hspace{1cm} (28)

As in (24), observe that the corresponding unnormalized value function $Z : \mathbb{R}^+ \times H^+ \rightarrow \mathbb{R}$ takes the form

$$Z(\zeta, p) = \mathbb{1}_{\zeta > 0}v(\zeta) + \mathbb{1}_{\zeta = 0}Y(p), \quad (\zeta, p) \in \mathbb{R}^+ \times H^+.$$  \hspace{1cm} (29)

Moreover,

$$Y(\mu p) = \mu Y(p) \quad \text{for every } \mu > 0.$$  \hspace{1cm} (30)

Thus, $Y(0) = 0.$
Compared with the value functions in BCS, the left-hand side of (26) is an unknown number $v$, but in the present paper it is a function $v(\cdot)$. This makes the analysis of the system more difficult. In what follows, we will show the details of the methodology we use to prove the existence of an optimal control. Important costs and spaces are summarized in table 3. Sequences of cost functions and some of the spaces will be formally defined soon. In this paper, we interpret decreasing and increasing sequences in a non-strict way.

<table>
<thead>
<tr>
<th>Definition</th>
<th>Normalized System</th>
<th>Unnormalized System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>$\pi$</td>
<td>$p$</td>
</tr>
<tr>
<td>Value function–backorder case</td>
<td>$v$</td>
<td>$L^+$</td>
</tr>
<tr>
<td>Value function–no backorder case</td>
<td>$W$</td>
<td>$B^+$</td>
</tr>
<tr>
<td>Value function–combined case</td>
<td>$V$</td>
<td>$G$</td>
</tr>
<tr>
<td>Increasing value iteration sequence</td>
<td>${V_n}$</td>
<td>$G$</td>
</tr>
<tr>
<td>Decreasing value iteration sequence</td>
<td>${V_n^+}$</td>
<td>$G$</td>
</tr>
</tbody>
</table>

Table 1
Frequently used notation.

4. Main Results. When the cost functions are bounded, we can apply commonplace arguments of the standard theory (see, e.g., [6]) to obtain a contraction property for an appropriate operator. Then by using the Banach’s Fixed Point Theorem, we prove directly the existence and uniqueness of the solution to the optimality equation as well as the value iteration algorithm. However, consideration of the unbounded one-period costs as in (14) brings in some challenges. For instance, the nice contraction property does not work for the discounted cost criterion, and so we need to consider ceiling functions for the optimal cost. These functions, in turn, will define the appropriate functional space where the solution to the system (22)-(23) is unique and it is the value function. Moreover, besides the result of existence of the optimal ordering policy, ceiling functions allow us to prove that the value iteration algorithm converges to the value function.

Before proving our main results, we give some preliminaries.

4.1. Preliminaries. We define the functional space
\[
\mathcal{B} := \left\{ \phi(p) : \mathcal{H} \to \mathbb{R} : \sup_{p \in \mathcal{H}} \frac{\|\phi(p)\|}{\|p\|} < \infty \right\}
\]
with the norm
\[
\|\phi\|_\mathcal{B} := \sup_{p \in \mathcal{H}} \frac{\|\phi(p)\|}{\|p\|} .
\]
From the definition, we can see that $\mathcal{B}$ is a Banach space. For any $\phi \in \mathcal{B}$, we must have $\phi(0) = 0$.

We define the space
\[
\mathcal{B}^+ := \left\{ \phi : \mathcal{H}^+ \to \mathbb{R}^+ : \sup_{p \in \mathcal{H}^+} \frac{\phi(p)}{\|p\|} < \infty \right\} \subseteq \mathcal{B},
\]
and denote by $B^\Pi$, the subset of $B^+$ restricted to $\Pi$. That is,
\[
\mathcal{B}^\Pi := \left\{ \phi : \Pi \to \mathbb{R}^+ : \sup_{\pi \in \Pi} \frac{\phi(\pi)}{||\pi||} < \infty \right\}.
\]

Similar to $L^+$, $B^+$ is also used to accommodate the dynamic programming cost.

We consider $v^0 \in L^+$ and $Y^0 \in B^+$ defined as
\[
v^0(\zeta) := \frac{a s E(D)}{(1 - \alpha)^2} + \frac{c_0 + s \zeta}{1 - \alpha}, \quad \zeta \in \mathbb{R}^+.
\]
\[
Y^0(p) := \frac{a_0}{1 - \alpha} ||p||, \quad p \in H^+,
\]
where
\[
a_0 = \max \left\{ h, \frac{a s E(D)}{(1 - \alpha)^2} + \frac{c_0 + s E(D)}{1 - \alpha} \right\}.
\]

For any $\phi_1 \in L^+$ and $\phi_2 \in B^+$ such that $\phi_1(\cdot) \leq v^0(\cdot)$ and $\phi_2(\cdot) \leq Y^0(\cdot)$, we define the set
\[
\mathcal{G} = \left\{ \phi : \mathbb{R}^+ \times H^+ \to \mathbb{R} \mid \phi(\zeta, p) = I_{\zeta > 0} \phi_1(\zeta) + I_{\zeta = 0} \phi_2(0) \text{ and } \phi(\zeta, p) \leq I_{\zeta > 0} v^0(\zeta) + I_{\zeta = 0} Y^0(p) \right\}.
\]

So the domain of the members of the set $\mathcal{G}$ is $H^+ \times \mathbb{R}^+$, where $H^+$ is the space of unnormalized probabilities. If $\phi_1(\zeta)$ and $\phi_2(p)$ are solutions of the system (26)-(27), then they correspond to $v(\zeta)$ and $Y(p)$, respectively. As a result, $\phi(\zeta, p)$ corresponds to $Z(\zeta, p)$. The counterpart of the normalized probabilities $\Pi \times \mathbb{R}^+$ is denoted by $\tilde{\mathcal{G}}$ and defined for functions $\tilde{\phi}_1 \in L^+$ and $\tilde{\phi}_2 \in B^\Pi$ such that $\tilde{\phi}_1(\cdot) \leq v^0(\cdot)$ and $\tilde{\phi}_2(\cdot) \leq Y^0(\cdot)$.

If $\tilde{\phi}_1$ and $\tilde{\phi}_2$ are solutions of the system (22)-(23), then $\tilde{\phi}_1$ corresponds to $v(\zeta)$ and $\tilde{\phi}_2$ corresponds to $W(\pi)$. As a result $\tilde{\phi}(\zeta, \pi) = V(\zeta, \pi)$. In the next subsection, we will show that the value functions $Z$ and $V$ are in the set $\mathcal{G}$ and $\tilde{\mathcal{G}}$, respectively.

For $\phi \in \mathcal{G}$, we define the operators
\[
T_q^{(1)}(\phi_1, \phi_2) = c(-\zeta, q) + \alpha \int_0^\infty \phi_1(\eta) f(\eta + q - \zeta) d\eta + \alpha \phi_2(q - \zeta),
\]
\[
T_q^{(2)}(\phi_1, \phi_2) = \int_0^\infty c(y, q) p(y) dy + \alpha \int_0^\infty p(y) \int_0^\infty \phi_1(\eta) f(\eta + q + \eta) d\eta dy + \alpha \phi_2(p(q)p).
\]

We can see that if $\phi_1$ and $\phi_2$ are solutions of (26)-(27), then $T_q^{(1)}(\phi_1, \phi_2)$ and $T_q^{(2)}(\phi_1, \phi_2)$ represent the cost when the ordering quantity is $q$ and when there are backorders and no backorders, respectively. By combining these two cases, we define the operator $T_q$ as follows:
\[
T_q(\zeta, p) := I_{\zeta > 0} T_q^{(1)}(\phi_1, \phi_2) + I_{\zeta = 0} T_q^{(2)}(\phi_1, \phi_2).
\]

Let
\[
T_q^{(1)}(\phi_1, \phi_2) := \inf_{q \geq 0} T_q^{(1)}(\phi_1, \phi_2), \quad T_q^{(2)}(\phi_1, \phi_2) := \inf_{q \geq 0} T_q^{(2)}(\phi_1, \phi_2), \quad T_0(\zeta, p) := \inf_{q \geq 0} T_q(\zeta, p).
\]
The operators $T^1$, $T^2$, and $T$ represent the corresponding optimal costs. After defining the above operators, we have the following lemma.

**Lemma 5.** When $\zeta \geq 0$, for any $\tilde{q} \in \Gamma$, $J(\zeta, \pi, \tilde{q}) \leq v^0(\zeta)$. Moreover, the operator $T$ maps $G$ into itself, and for $\phi_1 \leq v^0$ and $\phi_2 \leq Y^0$, $T^{(1)}(\phi_1, \phi_2) \leq v^0(\zeta)$ and $T^{(2)}(\phi_1, \phi_2) \leq Y^0(p)$.

From this lemma, we can see that $v^0$ and $Y^0$ are ceiling functions for the optimal cost. The solution $(v, Y)$ of the system (26)-(27) satisfies $v \leq v^0$ and $Y \leq Y^0$.

### 4.2. Existence of a Solution of the Bellman Equation and an Optimal Feedback Policy

We define a value iteration procedure as follows. Let $\{v_n\}$ and $\{Y_n\}$ be sequences of functions defined by $v_1 = Y_1 = 0$, and for $n \geq 1$,

$$v_{n+1} = T^{(1)}(v_n, Y_n), \quad Y_{n+1} = T^{(2)}(v_n, Y_n).$$  \hspace{1cm} (33)

From the definitions of operators $T^{(1)}$ and $T^{(2)}$, $\{v_n\}$ and $\{Y_n\}$ are increasing sequences. This can be established by induction. Clearly, $v_1 \leq v_2$ and $Y_1 \leq Y_2$. Let us assume that $v_n \leq v_{n+1}$ and $Y_n \leq Y_{n+1}$. Then,

$$v_{n+1}(\zeta) = T^{(1)}(v_n, Y_n) = \inf_{q \geq 0} \left\{c(-\zeta, q) + \alpha \int_0^\infty v_n(\eta)f(\eta + q - \zeta)d\eta + \alpha Y_n(\hat{q}_0(\eta - \zeta))\right\}$$

$$\leq \inf_{q \geq 0} \left\{c(-\zeta, q) + \alpha \int_0^\infty v_{n+1}(\eta)f(\eta + q - \zeta)d\eta + \alpha Y_{n+1}(\hat{q}_0(\eta - \zeta))\right\}$$

$$= T^{(1)}(v_{n+1}, Y_{n+1}) = v_{n+2}(\zeta),$$

and similarly we prove that $Y_{n+1}(p) \leq Y_{n+2}(p)$.

Let

$$Z_n(\zeta, p) := 1_{\zeta > 0}v_n(\zeta) + 1_{\zeta = 0}Y_n(p), \quad (\zeta, p) \in \mathbb{R}^+ \times \mathcal{H}^+. \hspace{1cm} (34)$$

Observe that

$$Z_{n+1}(\zeta, p) = 1_{\zeta > 0}T^{(1)}(v_n(\zeta), Y_n(p)) + 1_{\zeta = 0}T^{(2)}(v_n(\zeta), Y_n(p)) = TZ_n(\zeta, p). \hspace{1cm} (35)$$

Then we have the following theorem.

**Theorem 6.**

a) There exist lower semi-continuous (l.s.c.) functions $\overline{v} \leq Y^0$ and $\underline{v} \leq v^0$ such that $Z_n \nearrow Z$, and $Z = T\overline{v}$, where

$$Z(\zeta, p) := 1_{\zeta > 0}\overline{v}(\zeta) + 1_{\zeta = 0}\overline{v}(p), \quad (\zeta, p) \in \mathbb{R}^+ \times \mathcal{H}^+.$$

b) For each $(\zeta, p) \in \mathbb{R}^+ \times \mathcal{H}^+$, there exists a measurable function $g_Z : \mathbb{R}^+ \times \mathcal{H}^+ \rightarrow \mathbb{R}^+$ such that $T_Z(\zeta, p) = T(g_Z(\zeta, p)Z(\zeta, p))$.

From this theorem we can see that $\overline{\pi} \in \mathcal{H}^+$ and $\overline{\pi} \in B^+$. So far we have analyzed the unnormalized system (26)-(27). We show that there exists a solution of this system. Next, we will show that there is a solution of system (22)-(23). This result immediately follows from Theorem 6. Let $\{W_n, v_n\}$ be the normalized value iteration functions corresponding to $\{Y_n, v_n\}$. That is, $W_n$ is a function on $\Pi$ satisfying (see (25))

$$Y_n(p) = W_n\left(\frac{p}{\int p(x)dx}\right) \int_0^\infty p(x)dx. \hspace{1cm} (36)$$
Let $V_n(\zeta, \pi) := \mathbb{1}_{\zeta > a} v_n(\zeta) + \mathbb{1}_{\zeta = a} W_n(\pi)$, $\zeta, \pi \in \mathbb{R}^+ \times \Pi$. Then, as $v_1 = W_1 = 0$, we have $V_1 = 0$. Moreover,

$$V_{n+1}(\zeta, \pi) = \tilde{T} V_n(\zeta, \pi) := \inf_{q \geq 0} \tilde{T}_q V_n(\zeta, \pi), \quad (\zeta, \pi) \in \mathbb{R}^+ \times \Pi,$$

where, for a function $\tilde{\phi} \in \tilde{G}$,

$$\tilde{T}_q \tilde{\phi}(\zeta, \pi) := \mathbb{1}_{\zeta > a} \tilde{T}_q^{(1)} (\tilde{\phi}_1(\zeta), \tilde{\phi}_2(\pi)) + \mathbb{1}_{\zeta = a} \tilde{T}_q^{(2)} (\tilde{\phi}_1(\zeta), \tilde{\phi}_2(\pi))$$

and

$$\tilde{T}_q^{(1)} (\tilde{\phi}_1, \tilde{\phi}_2) = c(-\zeta, q) + \alpha \tilde{\phi}_2(\theta_0(q-\zeta)) F(q-\zeta) + \alpha \int_0^\infty \tilde{\phi}_1(\eta) f(\eta + q - \zeta) d\eta,$$

$$\tilde{T}_q^{(2)} (\tilde{\phi}_1, \tilde{\phi}_2) = \int_0^\infty c(y, q) \pi(y) dy + \alpha \int_0^\infty \pi(y) \int_0^\infty \tilde{\phi}_1(\eta) f(q + y + \eta) d\eta d\eta d\eta.$$

The above two equations correspond to equations (22) and (23), which yield the value functions for the normalized probabilities.

We know that $Y_n(\pi) = W_n(\pi)$ for all $\pi \in \Pi$ from (36). From (29), Theorem 6(a) yields the existence of an l.s.c. function $\tilde{W}(\pi) \leq Y^0(\pi)$ such that $W_n \nearrow \tilde{W}$. After defining $\tilde{V}(\zeta, \pi) := \mathbb{1}_{\zeta > a} \tilde{W}(\zeta) + \mathbb{1}_{\zeta = a} \tilde{W}(\pi)$, $(\zeta, \pi) \in \mathbb{R}^+ \times \Pi$, we have

$$\tilde{V}_n \nearrow \tilde{V}, \text{ as } n \to \infty,$$

(37)

and $\tilde{V}(\zeta, \pi) = \tilde{T} \tilde{V}(\zeta, \pi)$ for $(\zeta, \pi) \in \mathbb{R}^+ \times \Pi$. Hence, similar to (18), for each $(\zeta, \pi) \in \mathbb{R}^+ \times \Pi$,

$$\tilde{V}(\zeta, \pi) = \inf_{q} \left \{ \mathbb{1}_{\zeta > a} c(-\zeta, q) + \mathbb{1}_{\zeta = a} \int_0^\infty c(y, q) \pi(y) dy + \alpha \mathbb{E}[\tilde{V}(z_2, \pi_2) | \zeta, \pi] \right \}.$$

(38)

Furthermore, from Theorem 6(b), there exists a map $\tilde{g} : \mathbb{R}^+ \times \Pi \to \mathbb{R}^+$ such that

$$\tilde{V}(\zeta, \pi) = \tilde{T} \tilde{V}(\zeta, \pi) = \tilde{T}_{\tilde{g}(\zeta, \pi)} \tilde{V}(\zeta, \pi), \quad (\zeta, \pi) \in \mathbb{R}^+ \times \Pi.$$

(39)

So we can see that there is a solution of the system (22)-(23). In the next theorem, we establish that this solution is actually the value function.

**Theorem 7.** a) For each $(\zeta, \pi) \in \mathbb{R}^+ \times \Pi$, we have $\tilde{V}(\zeta, \pi) = V(\zeta, \pi)$, where $V$ is the optimal value function defined in (17). Hence, $Z(\zeta, \pi) = \tilde{Z}(\zeta, \pi)$ for each $(\zeta, \pi) \in \mathbb{R}^+ \times \mathbb{H}^+$. b) The functions $\tilde{V}$ and $\tilde{Z}$ are the minimal solutions in $\tilde{G}$ and $\tilde{G}$ of the optimality equations (24) and (28), respectively. c) There exists an optimal feedback policy $\tilde{\varphi}^* \in \Gamma$ for the partially observed inventory problem. That is,

$$V(\zeta, \pi) := \inf_{\varphi \in \Gamma} J(\zeta, \pi, \varphi) = J(\zeta, \pi, \tilde{\varphi}^*), \quad \forall (\zeta, \pi) \in \mathbb{R}^+ \times \Pi.$$

In this subsection, we have shown that starting from $v^0$ and $Y^0$, sequences $\{v_n\}$ and $\{Y_n\}$ are increasing and converge to $\tilde{v}$ and $\tilde{Y}$, respectively, and that the pair $(\tilde{v}, \tilde{Y})$ is a solution of the system (26)-(27). Since $Y_n(\pi) = W_n(\pi)$, sequences $v_n$ and $W_n$ converge to $\tilde{v}$ and $\tilde{W}$, respectively. As a result the pair $(\tilde{v}, \tilde{W})$ is a solution of (22)-(23). Theorem 7 states that $\tilde{V}(\zeta, \pi)$ is the value function defined in (17). In the next subsection, we will show that the solution of the system (26)-(27) is unique. As a result, the solution of the system (22)-(23) is also unique.
4.3. Uniqueness and Continuity of the Solution of the Bellman Equation. We consider the unnormalized value iteration procedure (33)-(35), but starting with the functions $v^0$ and $Y^0$. That is, for $(\zeta, p) \in \mathbb{R}^+ \times \mathcal{H}^+$,
\[
\begin{align*}
v^1(\zeta) &= v^0(\zeta), \\
v^{n+1}(\zeta) &= T^{(1)}(v^n(\zeta), Y^n(p)),
\end{align*}
\]
and define $Y^1(p) = Y^0(p)$, $Y^{n+1}(p) = T^{(2)}(v^n(\zeta), Y^n(p))$. (40)

Using induction arguments, it is easy to see that $\{v^n\}$ and $\{Y^n\}$ are nonnegative decreasing sequences. From the previous subsection we have $v^2 \leq v^1$ and $Y^2 \leq Y^1$.

To establish the decreasing property for $n \geq 1$, we assume that it holds for $n = k$, that is, $v^k \leq v^{k-1}$ and $Y^k \leq Y^{k-1}$, and then establish it for $n = k + 1$. It follows that
\[
v^{k+1}(\zeta) = T^{(1)}(v^k, Y^k) = \inf_{q \geq 0} \left\{ c(-\zeta, q) + \alpha \int_0^\infty v^k(q)f(\eta + q - \zeta)d\eta + \alpha Y^k(q_0(q - \zeta)) \right\}
\]
\[
\leq \inf_{q \geq 0} \left\{ c(-\zeta, q) + \alpha \int_0^\infty v^{k-1}(q)f(\eta + q - \zeta)d\eta + \alpha Y^{k-1}(q_0(q - \zeta)) \right\}
\]
\[
= T^{(1)}(v^{k-1}, Y^{k-1}) = v^k(\zeta),
\]
and similarly we prove that $Y^{k+1}(p) \leq Y^k(p)$.

Since $\{v^n\}$ and $\{Y^n\}$ are nonnegative decreasing sequences and both have lower bounds 0, there exist functions $\underline{v} \leq v^0$ and $\underline{Y} \leq Y^0$ such that
\[
v^n \downarrow \underline{v} \quad \text{and} \quad Y^n \downarrow \underline{Y}, \quad \text{as} \quad n \to \infty. \quad (41)
\]

In other words, if we denote
\[
Z^n(\zeta, p) = \mathbb{I}_{\zeta > 0}v^n(\zeta) + \mathbb{I}_{\zeta = 0}Y^n(p),
\]
\[
\underline{Z}(\zeta, p) = \mathbb{I}_{\zeta > 0}\underline{v}(\zeta) + \mathbb{I}_{\zeta = 0}\underline{Y}(p), \quad (\zeta, p) \in \mathbb{R}^+ \times \mathcal{H}^+,
\]
we have
\[
Z^{n+1} = T\underline{Z} \quad \text{and} \quad \underline{Z} \in \mathcal{G}, \quad Z^n \downarrow \underline{Z}, \quad \text{as} \quad n \to \infty. \quad (42)
\]

Since $Z^n$ is a decreasing sequence, we have $Z^n \geq \underline{Z}$ for a finite $n$. Applying $T$ infinitely many times on both sides, we obtain
\[
\underline{Z} \geq T\underline{Z}. \quad (43)
\]

Moreover, by applying the monotone convergence theorem together with (40) and (41), we can prove that for all $q \in \mathbb{R}^+$, $g(\zeta) \leq T^1_g(q(\zeta), Y(p))$ and $Y(p) \leq T^2_g(q(\zeta), Y(p))$, which implies $Z(\zeta, p) \leq T\underline{Z}(\zeta, p)$. This inequality combined with (43) yields
\[
\underline{Z}(\zeta, p) = T\underline{Z}(\zeta, p). \quad (44)
\]

Now, because $\underline{Z} \in \mathcal{G}$ (see (28) and Theorem 7(b)), we have $v(\zeta) \leq v^0(\zeta)$ and $Y(p) \leq Y^0(p)$.

Then, as $Z(\zeta, p)$ is a fixed point of the operator $T$, we obtain $Z(\zeta, p) \leq Z^n(\zeta, p)$ for all $n$, which implies from (42) that
\[
Z(\zeta, p) \leq \underline{Z}(\zeta, p), \quad \forall (\zeta, p) \in \mathbb{R}^+ \times \mathcal{H}^+. \quad (45)
\]

On the other hand, let $W^n : \Pi \to \mathbb{R}$ be the corresponding normalized value iteration function defined in the same way as (36), and define $V^n(\zeta, \pi) = \mathbb{I}_{\zeta > 0}v^n(\zeta) +
\( I_{c=0} W(n) \). Observe that \( V_{n+1} = \tilde{T} V_n \). Then, there exists a function \( V \in \tilde{G} \) such that \( V_n \downarrow V \). Furthermore, from (44) and (45), we can prove that for each \((\zeta, \pi) \in \mathbb{R}^+ \times \Pi\),

\[
V(\zeta, \pi) = \tilde{T} V(\zeta, \pi)
\]

and

\[
V(\zeta, \pi) \leq \tilde{V}(\zeta, \pi).
\]

**Theorem 8.** Suppose \( \alpha \left( 1 + \frac{\alpha}{\alpha + \alpha_j} \right) < 1 \). Then we have the following.

a) For each \((\zeta, \pi) \in \mathbb{R}^+ \times \Pi\), \( V(\zeta, \pi) = \tilde{V}(\zeta, \pi) \). Moreover, \( V \) is the maximal solution in \( \tilde{G} \) of the optimality equation (24). Hence, \( Z(\zeta, \pi) = \tilde{Z}(\zeta, \pi) \) for each \((\zeta, \pi) \in \mathbb{R}^+ \times \mathcal{H}^+\), and \( Z \) is the maximal solution of (28) in \( G \). Moreover, the solutions of (24) and (28) are unique.

b) The value functions \( V \) and \( Z \) are continuous on \( \mathbb{R}^+ \times \Pi \) and \( \mathbb{R}^+ \times \mathcal{H}^+ \), respectively.

### 4.4. Conclusions

We have studied an infinite-horizon single-stage periodic-review inventory control system with backorders. In this system, the inventory level is partially observed when it is nonnegative. Partial inventory observations eventually lead to a dynamic program in the space of probability distributions. This dynamic program is highly nonlinear. We apply the methodology of unnormalized probability to linearize the dynamic programming equations. This linearization facilitates the analysis.

As we discussed in the introduction, partial observations in inventory systems are very common. But there is not much research done on this topic. This motivated us to analyze this problem with rain checks. In Theorems 6 and 7, we prove the convergence of a value iteration algorithm to the optimal value function, as well as the existence of an optimal feedback ordering policy. In addition, the value function is the minimal solution of the optimality equation. Furthermore, from Theorem 8, for a sufficiently small discount factor, the solution of the optimality equation is unique and continuous, and it is the value function.

We conclude the paper by describing an extension of the model and discussing some computational aspects related to the rain check model.

Sometimes, when the customer comes to the store during a stock out, she does not ask for a rain check, leaves the store, and comes back later. Some stores do not have a practice of offering rain checks. In the absence of rain checks, the inventory manager cannot fully observe the backordered quantity. This setting gives rise to an interesting model where even less information is available to IM than the current model. That is, in this new context, the IM cannot fully observe the negative inventory, and the available information to make his decisions in each period would be the sign of the inventory. The authors currently are studying this extension [4].

Although the unnormalized probability simplifies the analysis, it does not lead to a dynamic program in finite dimensional spaces. Finiteness of these spaces, which accommodate the conditional distribution \( p \) of the inventory level, can be imposed to develop approximations. The challenge here is to find an appropriate finite family of functions to represent \( p \). This is a computational issue worth exploring. We provide some numerical techniques in [5] to solve the dynamic programming equation.
5. Proofs.

5.1. Preliminary results. **Lemma 9.** The function

\[
(q, p) \to \int p(y) \int \phi_1(\eta)f(q + y + \eta)d\eta dy
\]  

(48)

is continuous for every bounded function \( \phi_1 : \mathbb{R}^+ \to \mathbb{R} \).

**Proof.** Let \( \phi_1 \) be a bounded function and \( \{(q_n, p_n)\} \) be a sequence in \( \mathbb{R}^+ \times \mathcal{H}^+ \) converging to \((q, p) \in \mathbb{R}^+ \times \mathcal{H}^+ \). Then, by adding and subtracting the term \( \int \phi_1(\eta)f(q_n + y + \eta)p(y)d\eta dy \), we have

\[
\left| \int \phi_1(\eta)f(q_n + y + \eta)p_n(y)d\eta dy - \int \phi_1(\eta)f(q + y + \eta)p(y)d\eta dy \right|
\]

\[
\leq \int \phi_1(\eta)f(q_n + y + \eta)p_n(y) - p(y)\cdot d\eta dy
\]

\[
+ \int \phi_1(\eta)p(y)\cdot |f(q_n + y + \eta) - f(q + y + \eta)|d\eta dy,
\]

which, by the Dominated Convergence Theorem, converges to zero as \( n \to \infty \). This proves the continuity of the function defined in (48).

The next lemma is a key result to prove that the operators \( T_q \) and \( T \) map continuous functions into continuous functions (see Remark 1(b) and Lemma 11). So, in each iteration of the monotone approximations introduced in Section 4, we get continuous functions when we use both increasing and decreasing processes.

**Lemma 10.** For each nonnegative measurable function \( \phi_1 \) on \( \mathbb{R}^+ \), the function

\[
(q, p) \to \int p(y) \int \phi_1(\eta)f(q + y + \eta)d\eta dy
\]

(49)

is continuous.

**Proof.** Let \( \{(q_n, p_n)\} \) be a sequence in \( \mathbb{R}^+ \times \mathcal{H}^+ \) converging to \((q, p) \in \mathbb{R}^+ \times \mathcal{H}^+ \), and let \( \phi_1 \) be a measurable function on \( \mathbb{R}^+ \). Then, there exists a sequence \( \{\phi_k\} \) of bounded functions on \( \mathbb{R}^+ \) such that \( \phi_k \to \phi_1 \), as \( k \to \infty \). Therefore from Lemma 9, for each \( k \in \mathbb{N} \),

\[
\lim_{n \to \infty} \int \phi_k(\eta)f(q_n + y + \eta)p_n(y)d\eta dy = \int \phi_k(\eta)f(q + y + \eta)p(y)d\eta dy.
\]

Thus,

\[
\lim_{n \to \infty} \int \phi_1(\eta)f(q_n + y + \eta)p_n(y)d\eta dy \geq \lim_{n \to \infty} \int \phi_k(\eta)f(q_n + y + \eta)p_n(y)d\eta dy
\]

\[
= \int \phi_k(\eta)f(q + y + \eta)p(y)d\eta dy.
\]

Letting \( k \to \infty \) and using Fatou’s Lemma, we obtain

\[
\lim_{n \to \infty} \int \phi_1(\eta)f(q_n + y + \eta)p_n(y)d\eta dy \geq \int \phi_1(\eta)f(q + y + \eta)p(y)d\eta dy.
\]

That is, the function (49) is l.s.c. Now, applying the same arguments to the function \(-\phi_1\), we have that \(-\int \phi_1(\eta)f(q + y + \eta)p(y)d\eta dy \) is l.s.c. in \((q, p)\). Hence, the function (49) is upper semi-continuous (u.s.c.), which yields the desired result.

\[ \square \]
**Remark 1.**  
(a) Applying similar arguments as in the proofs of Lemmas 9 and 10, we can show the continuity of the functions 

\[(x, q, p) \rightarrow \int f(y + q - x)p(y)dy, \quad (x, q, p) \in \mathbb{R}^+ \times \mathbb{R}^+ \times \mathcal{H}^+\]

and 

\[(x, q) \rightarrow \int \phi_1(\eta)f(q - x + \eta)d\eta, \quad (x, q) \in \mathbb{R}^+ \times \mathbb{R}^+\]

for each measurable function \(\phi_1\) on \(\mathbb{R}^+\).

(b) Let \(\phi \in \mathcal{G}\) be a continuous function on \(\mathbb{R}^+ \times \mathcal{H}^+\). That is, \(\phi\) is a function of the form \(\phi(\zeta, p) = \mathbb{I}_{\zeta > 0}\phi_1(\zeta) + \mathbb{I}_{\zeta = 0}\phi_2(p)\), where \(\phi_1\) and \(\phi_2\) are continuous functions on \(\mathbb{R}^+\) and \(\mathcal{H}^+\) respectively. Then, in view of (15), Lemmas 9 and 10, and Remark 1(a), we have that 

\[T_q\phi(\zeta, p)\]

is continuous in \((\zeta, q, p) \in \mathbb{R}^+ \times \mathbb{R}^+ \times \mathcal{H}^+\).

**Remark 2.**  
(a) Observe that by defining the operator \(T\) for each \(\phi \in \mathcal{G}\), we can obtain bounds for the range of \(q\) for which 

\[cq_\pm \leq T_q^{(1)}(\phi_1(\zeta), \phi_2(p)) \leq v^0(\zeta) = \frac{\alpha s E(D)}{(1 - \alpha)^2} + \frac{c_0 + s\zeta}{1 - \alpha},\]

Hence, 

\[q_- \leq \frac{\alpha s E(D)}{(1 - \alpha)^2 c} + \frac{c_0 + s\zeta}{(1 - \alpha)c}.\]  

(50)

Similarly, we have 

\[cq_\pm \int_0^\infty p(y)dy \leq T_q^{(2)}(\phi_1(\zeta), \phi_2(p)) \leq Y^0(p) = \frac{a_0}{1 - \alpha}||p||,\]

which implies 

\[q_+ \leq \frac{a_0}{c(1 - \alpha)} \int_0^\infty p(x)dx ||p||.\]  

(51)

Then from (50) and (51), for each \((\zeta, p) \in \mathbb{R}^+ \times \mathcal{H}^+, q\) must belong \(Q^*(\zeta, p) := Q_1^*(\zeta) \cup Q_2^*(p)\), where 

\[Q_1^*(\zeta) := \left\{ q \in \mathbb{R}^+ : q \leq \frac{\alpha s E(D)}{(1 - \alpha)^2 c} + \frac{c_0 + s\zeta}{(1 - \alpha)c} \right\}\]

and 

\[Q_2^*(p) := \left\{ q \in \mathbb{R}^+ : q \leq \frac{a_0}{c(1 - \alpha)} \int_0^\infty p(x)dx ||p|| \right\}.\]

So, for a fixed \((\zeta, p) \in \mathbb{R}^+ \times \mathcal{H}^+, q\) remains bounded.
b) Taking into account the latter, since for each $\phi \in G$, the map $(\zeta, p, q) \mapsto T_q \phi(\zeta, p)$ is continuous, there exists a measurable function $g_\phi : \mathbb{R}^+ \times H^+ \to \mathbb{R}^+$ that attains the minimum in (32). That is,

$$T\phi(\zeta, p) = T_{g_\phi}(\zeta, p), \ \forall(\zeta, p) \in \mathbb{R}^+ \times H^+. \tag{52}$$

**Lemma 11.** For each continuous function $\phi \in G$, $T\phi(\zeta, p)$ is continuous in $(\zeta, p) \in \mathbb{R}^+ \times H^+$.  

**Proof.** Let $\{(\zeta_k, p_k)\}$ be a sequence in $\mathbb{R}^+ \times H^+$ such that $(\zeta_k, p_k) \to (\zeta, p) \in \mathbb{R}^+ \times H^+$, $p \neq 0$, and $q_k = g_\phi(\zeta_k, p_k) \in \mathcal{Q}^*(\zeta_k, p_k)$ satisfies (See Remark 2(b))

$$T\phi(\zeta, p_k) = T_{q_k} \phi(\zeta, p_k). \tag{53}$$

Clearly $q_k$ remains in a compact set. Then, we can extract a subsequence $\{(\zeta_{k_l}, p_{k_l})\}$ of $\{(\zeta_k, p_k)\}$ such that $(\zeta_{k_l}, p_{k_l}) \to (\zeta, p, q')$ for some $q' \in \mathcal{Q}^*(\zeta, p)$. Now from the continuity of $T_{q'} \phi(\zeta, p)$ (see Remark 1(b)), we have

$$\lim_{l \to \infty} T_{q_{k_l}} \phi(\zeta_{k_l}, p_{k_l}) = T_{q'} \phi(\zeta, p).$$

Hence, from (53) and (32),

$$\liminf_{k \to \infty} T\phi(\zeta_k, p_k) = T_{q'} \phi(\zeta, p) \geq T\phi(\zeta, p). \tag{54}$$

On the other hand, we have

$$T\phi(\zeta_k, p_k) \leq T_{q'} \phi(\zeta_k, p_k), \ \forall q \in \mathbb{R}^+.$$

Then, again from the continuity of $T_{q'} \phi(\zeta, p)$,

$$\limsup_{k \to \infty} T\phi(\zeta_k, p_k) \leq T_{q'} \phi(\zeta, p), \ \forall q \in \mathbb{R}^+, \tag{55}$$

which yields

$$\limsup_{k \to \infty} T\phi(\zeta_k, p_k) \leq T\phi(\zeta, p).$$

Therefore, by combining (54) and (55), we obtain

$$\lim_{k \to \infty} T\phi(\zeta_k, p_k) = T\phi(\zeta, p).$$

**Remark 3.** a) If $\phi \in G$ is only l.s.c., we can follow similar arguments as in the proof of Lemma 11 (see also Remark 1(b)) to show that the functions $T_q \phi(\zeta, p)$ and $T\phi(\zeta, p)$ are l.s.c. in $(\zeta, p, q)$ and $(\zeta, p)$, respectively. In addition, the selection theorem ensures the existence of a measurable function $g_\phi : \mathbb{R}^+ \times H^+ \to \mathbb{R}^+$ satisfying (52).

b) It is worth noting that if $\phi_1 : \mathbb{R}^+ \to \mathbb{R}^+$ and $\phi_2 : H^+ \to \mathbb{R}^+$ are continuous, then $T^{(1)}(\phi_1(\zeta), \phi_2(p))$ and $T^{(2)}(\phi_1(\zeta), \phi_2(p))$ are continuous (see (64) and (65) for definitions of $T^{(1)}$ and $T^{(2)}$).
5.2. Proof of Lemma 1. Proof. We can express the left-hand side of (5) as
\[ E(\varphi(I_t) \mid Z_t) = E[\varphi(I_t)(1_{z_t > 0} + 1_{z_t = 0}) \mid Z_t] = \varphi(-z_t)1_{z_t > 0} + E[\varphi(I_t)1_{z_t = 0} \mid Z_t]. \] (56)
From the last term of (56), we have
\[ E(\varphi(I_t)1_{z_t = 0} \mid Z_t) = 1_{z_t = 0}E(\varphi(I_t) \mid Z_t) = 1_{z_t = 0}\psi(z_1, \ldots, z_{t-1}, z_t) = 1_{z_t = 0}\psi(z_1, \ldots, z_{t-1}, 0), \] (57)
where the first equality follows from the fact that \( 1_{z_t = 0} \) is \( Z_t \) measurable. The second equality introduces a measurable function \( \psi \) to express \( E(\varphi(I_t) \mid Z_t) \) as \( \psi(z_1, \ldots, z_{t-1}, z_t) \). If we set \( z_t = 0 \) into the second equality, we can obtain the last equality.

Taking conditional expectation of (57) with respect to \( Z_{t-1} \) and observing that \( Z_{t-1} \subseteq Z_t \) and \( \psi(z_1, \ldots, z_{t-1}, 0) \) is \( Z_{t-1} \)-measurable, we obtain
\[ E(\varphi(I_t)1_{z_t = 0} \mid Z_{t-1}) = \psi(z_1, \ldots, z_{t-1}, 0)E[1_{I_t \geq 0} \mid Z_{t-1}] = \psi(z_1, \ldots, z_{t-1}, 0)P(I_t \geq 0 \mid Z_{t-1}), \]
or
\[ \psi(z_1, \ldots, z_{t-1}, 0) = \frac{E(\varphi(I_t)1_{I_t \geq 0} \mid Z_{t-1})}{P(I_t \geq 0 \mid Z_{t-1})}. \] (58)
We insert (58) into (57) and then the result into (56) to get
\[ E(\varphi(I_t) \mid Z_t) = 1_{z_t > 0}\varphi(-z_t) + 1_{z_t = 0}\frac{E(\varphi(I_t)1_{z_t = 0} \mid Z_{t-1})}{P(I_t \geq 0 \mid Z_{t-1})}. \]
By using (3) and (4), we have
\[ E(\varphi(I_t) \mid Z_t) = 1_{z_t > 0}\varphi(-z_t) + 1_{z_t = 0}\int_0^\infty \varphi(\eta)\pi_t(\eta)\,d\eta. \]

D

5.3. Proof of Theorem 2. Proof. We start to prove this theorem from the right-hand side of (4). Take the numerator and obtain
\[ E(\varphi(I_t)1_{z_t = 0} \mid Z_{t-1}) = E(\varphi(I_{t-1} + q_{t-1} - D_{t-1})1_{I_{t-1} + q_{t-1} - D_{t-1} \geq 0} \mid Z_{t-1}) \]
\[ = E\left(E(\varphi(I_{t-1} + q_{t-1} - D_{t-1})1_{I_{t-1} + q_{t-1} - D_{t-1} \geq 0} \mid Z_{t-1}, I_{t-1}) \mid Z_{t-1}\right) \]
\[ = E\left(\int_0^\infty \varphi(I_{t-1} + q_{t-1} - y)f(I_{t-1} + q_{t-1} - y \geq 0)\,dy \mid Z_{t-1}\right) \]
\[ \text{set } x := I_{t-1} + q_{t-1} - y \]
\[ = E\left(\int_0^\infty \varphi(x)f(I_{t-1} + q_{t-1} - x)1_{I_{t-1} + q_{t-1} - x \geq 0}\,dx \mid Z_{t-1}\right) \]
\[ = \int_0^\infty \varphi(x)E\left(f(I_{t-1} + q_{t-1} - x)1_{I_{t-1} + q_{t-1} - x \geq 0} \mid Z_{t-1}\right)\,dx. \] (59)
Letting the time index in the first equality of (5) be \( t - 1 \) instead of \( t \) and replacing \( \varphi(I_{t-1}) \) with \( f(I_{t-1} + q_{t-1} - x)1_{I_{t-1} + q_{t-1} - x \geq 0} \), we obtain
\[ E(f(I_{t-1} + q_{t-1} - x)1_{I_{t-1} + q_{t-1} - x \geq 0} \mid Z_{t-1}) = 1_{z_{t-1} > 0}f(-z_{t-1} + q_{t-1} - x)1_{z_{t-1} + q_{t-1} - x \geq 0} \]
\[ + 1_{z_{t-1} = 0}\int_{(x-q_{t-1})+}^\infty f(\eta + q_{t-1} - x)\pi_{t-1}(\eta)\,d\eta. \]
By inserting (60) into (59), we get

\[ E(\varphi(I_t) \mathbb{1}_{z_{t-1}} | Z_{t-1}) = \mathbb{1}_{z_{t-1} > 0} \int_{0}^{\infty} \varphi(x) f(-z_{t-1} + q_{t-1} - x) \mathbb{1}_{-z_{t-1} + q_{t-1} - x \geq 0} dx \]

\[ + \mathbb{1}_{z_{t-1} = 0} \int_{0}^{\infty} \varphi(x) \left( \int_{(x-q_{t-1})^{+}}^{\infty} f(\eta + q_{t-1} - x) \pi_{t-1}(\eta) d\eta \right) dx. \]

Similarly, we can express explicitly the denominator of the right-hand side of (4) as

\[ P(I_t \geq 0 | Z_{t-1}) = E(\mathbb{1}_{I_t} + q_{t-1} - D_{t-1} \geq 0 | Z_{t-1}) \]

\[ = E(E(\mathbb{1}_{I_t} + q_{t-1} - D_{t-1} \geq 0 | Z_{t-1}, I_{t-1}) | Z_{t-1}) \]

\[ = \mathbb{1}_{z_{t-1} > 0} F(-z_{t-1} + q_{t-1}) + \mathbb{1}_{z_{t-1} = 0} \int_{0}^{\infty} F(y + q_{t-1} \pi_{t-1}(y) dy. \]

Because we have already obtained explicit expressions of the numerator and denominator of the right-hand side of (4), we can express the fraction, say the right-hand side of (4), as follows:

\[ \frac{E(\varphi(I_t) \mathbb{1}_{z_{t-1}} | Z_{t-1})}{P(I_t \geq 0 | Z_{t-1})} = \mathbb{1}_{z_{t-1} > 0} \int_{0}^{\infty} \varphi(x) f(-z_{t-1} + q_{t-1} - x) \mathbb{1}_{-z_{t-1} + q_{t-1} - x \geq 0} \]

\[ \int_{0}^{\infty} \frac{f(y + q_{t-1} \pi_{t-1}(y) dy} \]

\[ + \mathbb{1}_{z_{t-1} = 0} \int_{0}^{\infty} \varphi(x) \left( \int_{(x-q_{t-1})^{+}}^{\infty} f(\eta + q_{t-1} - x) \pi_{t-1}(\eta) d\eta \right) dx \]

Then, we have

\[ \int_{0}^{\infty} \varphi(\eta) \pi_{t}(\eta) d\eta = \mathbb{1}_{z_{t-1} > 0} \int_{0}^{\infty} \varphi(x) f(-z_{t-1} + q_{t-1} - x) \mathbb{1}_{-z_{t-1} + q_{t-1} - x \geq 0} \]

\[ + \mathbb{1}_{z_{t-1} = 0} \int_{0}^{\infty} \varphi(x) \left( \int_{(x-q_{t-1})^{+}}^{\infty} f(\eta + q_{t-1} - x) \pi_{t-1}(\eta) d\eta \right) dx \]

Since the above equation is satisfied for all test functions \( \varphi(x) \), we have

\[ \pi_{t}(x) = \mathbb{1}_{z_{t-1} > 0} \left\{ \frac{f(-z_{t-1} + q_{t-1} - x) \mathbb{1}_{-z_{t-1} + q_{t-1} - x \geq 0}}{F(q_{t-1} - z_{t-1})} \right\} \]

\[ + \mathbb{1}_{z_{t-1} = 0} \left\{ \frac{\int_{(x-q_{t-1})^{+}}^{\infty} f(\eta + q_{t-1} - x) \pi_{t-1}(\eta) dy}{\int_{0}^{\infty} F(y + q_{t-1} \pi_{t-1}(y) dy} \right\} \]

This equality is exactly (6), which completes the proof. \( \Box \)

5.4. Proof of Lemma 3. Proof. By changing the order of integration, we obtain

\[ \int_{0}^{\infty} |\varphi(q)p(x)| dx \leq \int_{0}^{\infty} \int_{(x-q)^{+}}^{\infty} f(y + q - x) |p(y)| dy dx \]

\[ = \int_{0}^{\infty} \int_{0}^{y+q} f(y + q - x) |p(y)| dx dy \]

\[ = \int_{0}^{\infty} |p(y)| \int_{0}^{y+q} f(y + q - x) dx dy \]

\[ \leq \int_{0}^{\infty} |p(y)| F(y + q) dy. \] (61)
Using similar operations, we see that

$$\int_0^\infty x |\varrho(q)p(x)|dx \leq \int_0^\infty \int_0^\infty xf(y + q - x)|p(y)|dydx$$

$$= \int_0^\infty \int_0^{y+q} xf(y + q - x)|p(y)|dx dy$$

$$= \int_0^\infty |p(y)| \int_0^{y+q} (y + q - z)f(z)dz dy$$

$$\leq \int_0^\infty (y + q)|p(y)|F(y + q)dy. \quad (62)$$

Using the two inequalities above, we get

$$||\varrho(q)p|| = \int_0^\infty |\varrho(q)p(x)|dx + \int_0^\infty x |\varrho(q)p(x)|dx$$

$$\leq \int_0^\infty |p(y)|F(y + q)dy + \int_0^\infty |p(y)|(y + q)F(y + q)dy$$

$$\leq ||p|| + q \int_0^\infty |p(y)|dy. \quad (63)$$

From (63) we can easily obtain $||\varrho(q)||_L \leq 1 + q$. □

5.5. Proof of Lemma 4. Proof. Since $c(y, q) \geq 0$, $p_n \in H^+$ and $p \in H^+$, we have

$$\lim_{n \to \infty} \left| \int_0^\infty c(y, q_n)p_n(y)dy - \int_0^\infty c(y, q)p(y)dy \right|$$

$$= \lim_{n \to \infty} \left| \int_0^\infty c(y, q_n)p_n(y)dy - \int_0^\infty c(y, q)p_n(y)dy + \int_0^\infty c(y, q)p_n(y)dy - \int_0^\infty c(y, q)p(y)dy \right|$$

$$\leq \lim_{n \to \infty} \left| \int_0^\infty c(y, q_n)p_n(y)dy - \int_0^\infty c(y, q)p_n(y)dy \right|$$

$$+ \lim_{n \to \infty} \left| \int_0^\infty c(y, q)p_n(y)dy - \int_0^\infty c(y, q)p(y)dy \right|$$

$$= \lim_{n \to \infty} \left| \int_0^\infty [c(y, q_n) - c(y, q)]p_n(y)dy \right| + \lim_{n \to \infty} \int_0^\infty c(y, q)|p_n(y) - p(y)|dy$$

$$\leq \lim_{n \to \infty} \int_0^\infty |c(y, q_n) - c(y, q)|p_n(y)dy + \lim_{n \to \infty} \int_0^\infty [c_0 + c_1 q + bz]p_n(y) - p(y)|dy = 0,$$

where both limits approach zero, respectively, on account of continuity of $c(y, q)$ in $q$ and equation (10). □

5.6. Proof of Lemma 5. Proof. The function $v^0$ is an upper bound for the discounted cost of backordering. Indeed, let $\bar{q}_0 = \{0, 0, \ldots\} \in \Gamma$ and $\bar{q} \in \Gamma$ be an
arbitrary policy. Then,

\[
J(\zeta, \pi, \tilde{q}) \leq J(\zeta, \pi, \tilde{q}_0) = \sum_{t=1}^{\infty} \alpha^{t-1} E\{1_{z_t > 0} \zeta(-z_t, q_t)\} \leq \sum_{t=1}^{\infty} \alpha^{t-1} E\{1_{z_t > 0}(\zeta_0 + z_t)\}
\]

\[
\leq (c_0 + s\zeta + \alpha (c_0 + s(\zeta + E(D))) + \alpha^2 (c_0 + s(\zeta + 2E(D))) + \cdots
= c_0 + s\zeta + sE(D) + (c_0 + s\zeta)(\alpha + \alpha^2 + \alpha^3 + \cdots)
+ sE(D)(2\alpha + 3\alpha^2 + 4\alpha^3 + 5\alpha^4 + \cdots)
\]

\[
= \frac{\alpha sE(D)}{(1-\alpha)^2} + \frac{c_0 + s\zeta}{1-\alpha} = v^0(\zeta), \quad \zeta \in \mathbb{R}^+,
\]

where we use the upper bound of the cost in (14) to obtain the second inequality. For \(\phi_1 \leq v^0\) and \(\phi_2 \leq Y^0\), since \(Y^0(0) = 0\),

\[
T^{(1)}(\phi_1, \phi_2) = \inf_q T^{(1)}_q(\phi_1, \phi_2) \leq c(-\zeta, 0) + \alpha \int_0^\infty v^0(\eta)f(\eta - \zeta)d\eta + \alpha Y^0(0)
\]

\[
= c(-\zeta, 0) + \alpha \int_0^\infty v^0(\eta)f(\eta - \zeta)d\eta
\]

\[
= c_0 + s\zeta + \alpha \int_0^\infty \left[ \frac{\alpha sE(D)}{(1-\alpha)^2} + \frac{c_0 + s\eta}{1-\alpha} \right] f(\eta - \zeta)d\eta.
\]

We know that \(\int_0^\infty \eta f(\eta - \zeta)d\eta \leq E(D) + \zeta\), for \(\zeta \in \mathbb{R}^+\). Then,

\[
T^{(1)}(\phi_1, \phi_2) \leq c_0 + s\zeta + \frac{\alpha^2 sE(D)}{(1-\alpha)^2} + \frac{\alpha c_0}{1-\alpha} + \alpha (E(D) + \zeta) \frac{s}{1-\alpha}
\]

\[= \alpha^2 sE(D) \frac{(1-\alpha)^2}{(1-\alpha)^2} + \frac{c_0 + s\zeta}{1-\alpha} \leq v^0(\zeta).
\]

So, \(T^{(1)}(\phi_1, \phi_2) \in L^0\).

For \(\phi_1 \leq v^0\) and \(\phi_2 \leq Y^0\),

\[
T^{(2)}(\phi_1, \phi_2) = \inf_q T^{(2)}_q(\phi_1, \phi_2) \leq T^{(2)}_0(\phi_1, \phi_2) \leq T^{(2)}(v^0, Y^0)
\]

\[
= \int_0^\infty c(y, 0)p(y)dy + \alpha \int_0^\infty p(y) \int_0^\infty v^0(\eta)f(\eta + \zeta)d\eta dy + \alpha Y^0(p(0)p)
\]

\[
\leq \int_0^\infty (c_0 + h\eta)p(y)dy + \alpha \int_0^\infty p(y) \int_0^\infty \left[ \frac{\alpha sE(D)}{(1-\alpha)^2} + \frac{c_0 + s\eta}{1-\alpha} \right] f(\eta + \zeta)d\eta dy
\]

\[+ \alpha \frac{a_0}{1-\alpha} ||p(0)p||
\]

\[
\leq \left( c_0 + \frac{sE(D)}{(1-\alpha)^2} + \frac{c_0 + sE(D)}{1-\alpha} \right) \int_0^\infty p(y)dy + h \int_0^\infty yp(y)dy + \alpha \frac{a_0}{1-\alpha} ||p||
\]

\[\leq a_0 ||p|| + \alpha \frac{a_0}{1-\alpha} ||p|| = \frac{a_0}{1-\alpha} ||p|| = Y^0(p).
\]

Obviously, \(T^{(2)}(\phi_1, \phi_2) \in B^+\). Therefore, The operator \(T\) maps \(G\) into itself. For \(\phi_1 \leq v^0\) and \(\phi_2 \leq Y^0\), \(T^{(1)}(\phi_1, \phi_2) \leq v^0(\zeta)\), \(T^{(2)}(\phi_1, \phi_2) \leq Y^0(p)\).
5.7. **Proof of Theorem 6.** Proof. From Remark 3(a), part (b) is a consequence of part (a).

To prove part (a), observe that since \( Y_1 = v_1 = 0 \), we have \( v_1 \leq v^0 \) and \( Y_1 \leq Y^0 \). Hence, we have that for each \( (\zeta, \pi) \), by applying induction arguments, we can prove that combined with (67), yields

\[ v_n \not\rightarrow \overline{\nu} \quad \text{and} \quad Y_n \not\rightarrow \overline{\nu} \quad \text{as} \quad n \to \infty. \]

Hence, we have that for each \( (\zeta, \pi) \in \mathbb{R}^+ \times \mathcal{H}^+ \),

\[ Z_n(\zeta, \pi) \not\rightarrow \overline{Z}(\zeta, \pi) \]  \hspace{1cm} (66)

and \( \overline{Z} \in \mathcal{G} \). Now, using the fact that the operators \( T^{(1)} \) and \( T^{(2)} \) are monotone, we have \( Z_n = T Z_{n-1} \leq T \overline{Z} \). Thus, from (66),

\[ \overline{Z}(\zeta, \pi) = T \overline{Z}(\zeta, \pi), \quad (\zeta, \pi) \in \mathbb{R}^+ \times \mathcal{H}^+. \]  \hspace{1cm} (67)

To obtain the reverse inequality, let \( q_n = \overline{g}_n(\zeta, \pi) \) be such that \( Z_{n+1}(\zeta, \pi) = T q_n Z_n(\zeta, \pi) \). Thus, observe that for any \( N \),

\[ Z_{n+1}(\zeta, \pi) \geq T q_n Z_N(\zeta, \pi), \quad \forall n \geq N. \]

Then from (66),

\[ \overline{Z}(\zeta, \pi) \geq T q_n Z_N(\zeta, \pi). \]  \hspace{1cm} (68)

In addition, we can extract a subsequence \( \{q_{n_k}\} \) of \( \{q_n\} \) such that \( q_{n_k} \to \overline{q} \in \mathcal{Q}^*(\zeta, \pi) \) as \( k \to \infty \). Then, by the continuity of the function \( q \to T q Z_N(\zeta, \pi) \), we have that \( T q_{n_k} Z_N(\zeta, \pi) \to T \overline{q} Z_N(\zeta, \pi) \) as \( k \to \infty \). Therefore, from (68), we have \( \overline{Z}(\zeta, \pi) \geq T \overline{q} Z_N(\zeta, \pi) \). Letting \( N \to \infty \), we obtain \( \overline{Z}(\zeta, \pi) \geq T \overline{q} \overline{Z}(\zeta, \pi) \geq T \overline{Z}(\zeta, \pi) \), which, combined with (67), yields \( \overline{Z}(\zeta, \pi) = T \overline{Z}(\zeta, \pi) \) for each \( (\zeta, \pi) \in \mathbb{R}^+ \times \mathcal{H}^+ \). \( \square \)

5.8. **Proof of Theorem 7.** Proof. a) Because \( V_1 = 0 \), we have \( V_1 \leq V \). Then, by applying induction arguments, we can prove that \( V_n(\zeta, \pi) \leq V(\zeta, \pi) \) for all \( n \) and \( (\zeta, \pi) \in \mathbb{R}^+ \times \Pi \). Therefore (see (37)), since \( V_n \not\rightarrow \overline{V} \) as \( n \to \infty \),

\[ \overline{V}(\zeta, \pi) \leq V(\zeta, \pi) \quad \forall (\zeta, \pi) \in \mathbb{R}^+ \times \Pi. \]  \hspace{1cm} (69)

To prove the reverse inequality, let \( \hat{q}_t = \hat{g}(z_t, \pi_t) \) be the map satisfying (see (38) and (39))

\[ \nabla(z_t, \pi_t) = E[\mathbb{I}_{z_t>0} c(-z_t, \hat{q}_t) + \mathbb{I}_{z_t=0} c(I_t, \hat{q}_t, \pi_t) \mid Z_t] + \alpha E[\nabla(z_{t+1}, \pi_{t+1}) \mid Z_t]. \]  \hspace{1cm} (70)

Hence,

\[ E[\alpha^{t-1} \nabla(z_t, \pi_t)] - E[\alpha^{t} \nabla(z_{t+1}, \pi_{t+1})] = \alpha^{t-1} E[\mathbb{I}_{z_t>0} c(-z_t, \hat{q}_t) + \mathbb{I}_{z_t=0} c(I_t, \hat{q}_t, \pi_t)]. \]

Summing up for \( t = 1, 2, \ldots, M \) yields

\[ \nabla(\zeta, \pi) = \sum_{t=1}^{M} \alpha^{t-1} E[\mathbb{I}_{z_t>0} c(-z_t, \hat{q}_t) + \mathbb{I}_{z_t=0} c(I_t, \hat{q}_t, \pi_t)] + \alpha^{M} E[\nabla(z_{M+1}, \pi_{M+1})] \]

\[ \geq \sum_{t=1}^{M} \alpha^{t-1} E[\mathbb{I}_{z_t>0} c(-z_t, \hat{q}_t) + \mathbb{I}_{z_t=0} c(I_t, \hat{q}_t, \pi_t)]. \]  \hspace{1cm} (71)
Letting $M \to \infty$ and denoting $\hat{q} = \{\hat{g}, \hat{g}, \ldots\}$, from (16) and (17), we get
\[
\nabla(\zeta, \pi) \geq J(\zeta, \pi, \hat{q}) \geq V(\zeta, \pi), \quad \forall (\zeta, \pi) \in \mathbb{R}^+ \times \Pi,
\]
(72)
which, from (69), proves part (a).

b) Let $\phi \in \hat{G}$ be an arbitrary function such that $\tilde{\phi}(\zeta, \pi) = \tilde{T}\phi(\zeta, \pi)$. That is, $\phi$ is of the form $\phi(\zeta, \pi) = \mathbb{I}_{\zeta>0} \phi_1(\zeta) + \mathbb{I}_{\zeta=0} \phi_2(\pi)$, and $\phi_1$ and $\phi_2$ satisfy the system (22)-(23). Then, applying the arguments in the proof of part (a) with $\tilde{\phi}$ instead of $\nabla$ (see (72)), we conclude that $\phi \geq V$. Since $V = \nabla$, then $\nabla \leq \phi$. Hence, $\nabla$ is minimal in $\hat{G}$. In addition, the corresponding unnormalized value function $Z$ is minimal in $\hat{G}$.

c) Let $q^* = \{g^*, g^*, \ldots\} \in \Gamma$ be the policy determined by the map $g^* : \mathbb{R}^+ \times \Pi \to \mathbb{Q}$. By denoting $q_t^* = g^*(z_t, \pi_t)$ (see (38) and (39)), we write
\[
V(z_t, \pi_t) = E[\mathbb{I}_{z_t>0} c(-z_t, q_t^*) + \mathbb{I}_{z_t=0} \langle c(I_t, q_t^*), \pi_t \rangle | Z_t] + aE[V(z_{t+1}, \pi_{t+1}) | Z_t].
\]
Then the first inequality in (72) implies
\[
V(\zeta, \pi) \geq J(\zeta, \pi, q^*) \quad \forall (\zeta, \pi) \in \mathbb{R}^+ \times \Pi.
\]
Therefore, from (17), $q^*$ is optimal. \(\square\)

5.9. Proof of Theorem 8. Proof. a) Let $\tilde{q} \in \hat{\Gamma}$ be an arbitrary policy and $\{q_1, q_2, \ldots\}$ be the decisions corresponding to application of $\tilde{q}$. Then, from (46), we can proceed as in (70) and (71) to obtain
\[
V(\zeta, \pi) = \sum_{t=1}^{M} \alpha^{t-1} E[\mathbb{I}_{z_t>0} c(-z_t, \tilde{q}_t) + \mathbb{I}_{z_t=0} \langle c(I_t, \tilde{q}_t), \pi_t \rangle | Z_t] + \alpha^M E[V(z_{M+1}, \pi_{M+1})].
\]
(73)
Now, as $V \in \hat{G}$, it follows that
\[
V(z_{M+1}, \pi_{M+1}) \leq \mathbb{I}_{z_{M+1}>0} v^0(z_{M+1}) + \mathbb{I}_{z_{M+1}=0} Y^0(\pi_{M+1}).
\]
(74)
On the other hand, observe that from (2) $z_t > 0$ if, and only if, $I_t < 0$, for each $t \geq 1$. Then, if $z_{M+1} > 0$, we have $z_{M+1} = -I_{M+1} - D_{M} - I_M$. Hence, if $z_t > 0$ $\forall t$, iterating this inequality we have,
\[
-E[I_{M+1}] \leq (M - 1) E(D) - \zeta,
\]
where $I_1 = \zeta$. Therefore, from (30),
\[
\alpha^M E[\mathbb{I}_{z_{M+1}>0} v^0(z_{M+1})] = \alpha^M E \left[ \mathbb{I}_{z_{M+1}>0} \left[ \frac{\alpha s E(D)}{(1 - \alpha)^2} + \frac{c_0 + sz_{M+1}}{1 - \alpha} \right] \right] \leq \alpha^M E \left[ \frac{\alpha s E(D)}{(1 - \alpha)^2} + \frac{c_0 + s(M - 1) E(D) - s\zeta}{1 - \alpha} \right] \to 0, \quad \text{as} \quad M \to \infty.
\]
(75)
On the other hand, from (31),
\[
Y^0(\pi_{M+1}) = \frac{a_0}{1 - \alpha} ||\pi_{M+1}|| = \frac{a_0}{1 - \alpha} \left( 1 + \int_0^\infty x \pi_{M+1}(x) dx \right) = \frac{a_0}{1 - \alpha} (1 + E[I_{M+1} | Z_M]).
\]
Then,
\[ E \left[ Y^0(\pi_{M+1}) \right] = \frac{a_0}{1-\alpha} + \frac{a_0}{1-\alpha} E \left[ I_{M+1} \right]. \] (76)

If \( z_{M+1} = 0 \), we have from (51),
\[ I_{M+1} = I_M + q_M - D_M = I_M + q_M \leq I_M + \frac{a_0}{c(1-\alpha)} ||\pi_M|| \]
\[ = I_M + \frac{a_0}{c(1-\alpha)} + \frac{a_0}{c(1-\alpha)} E \left[ I_M \mid Z_{M-1} \right]. \]

Hence,
\[ E \left[ I_{M+1} \right] \leq \left( 1 + \frac{a_0}{c(1-\alpha)} \right) E \left[ I_M \right] + \frac{a_0}{c(1-\alpha)}. \]

Therefore, if \( z_t = 0 \ \forall \ t \), iteration of this inequality yields
\[ E \left[ I_{M+1} \right] \leq \left( 1 + \frac{a_0}{c(1-\alpha)} \right)^M (E(I_1) + 1) - 1. \]

Thus, from (76), as \( M \to \infty \),
\[ \alpha^M E \left[ 1_{z_{M+1}=0} Y^0(\pi_{M+1}) \right] \leq \alpha^M \left[ \frac{a_0}{1-\alpha} (1 + E(I_1)) \left( 1 + \frac{a_0}{c(1-\alpha)} \right)^M \right] \to 0. \] (77)

Combining (74), (75), and (77), we have
\[ \alpha^M [V(z_{M+1}, \pi_{M+1})] \to 0, \text{ as } M \to \infty. \]

Letting \( M \to \infty \) in (73), we get \( V(\zeta, \pi) \leq J(\zeta, \pi, \tilde{\eta}) \), and as \( \tilde{\eta} \in \Gamma \) was arbitrary, we have
\[ V(\zeta, \pi) \leq V(\zeta, \pi). \] (78)

This combined with (47) yields \( V(\zeta, \pi) = \tilde{V}(\zeta, \pi) \).

Finally, let \( \tilde{\phi} \in \tilde{G} \) be an arbitrary function such that \( \tilde{\phi}(\zeta, \pi) = \tilde{T} \tilde{\phi}(\zeta, \pi) \). Then, by using \( \tilde{\phi} \) instead of \( V \), we obtain \( V \geq \tilde{\phi} \) (see (78)). Since \( V = V \), \( V \geq \tilde{\phi} \). Then, \( V \) is maximal in \( \tilde{G} \). Similarly, \( Z \) is maximal in \( G \). Furthermore, in view of the results in Theorem 7, we can conclude that the solutions of (24) and (28) are unique.

b) According to part (a), we see that the value iteration functions \( Z^n \) converge decreasingly to the value function \( Z \), that is,
\[ Z^n \searrow Z, \text{ as } n \to \infty. \] (79)

On the other hand, since \( v^t(\zeta) = v^0(\zeta) \) and \( Y^t(p) = Y^0(p) \) are continuous functions, we have from Lemma 11 that \( Z^n \) is continuous on \( R^+ \times H^+ \) for all \( n \geq 1 \). Therefore, from (79), we can ensure that the value function \( Z \) is u.s.c. Hence, from the lower semi-continuity of \( Z \) given in Theorems 6 and 7, we conclude that \( Z \) is continuous on \( R^+ \times H^+ \). This also yields the continuity of \( V \) on \( R^+ \times \Pi \). This completes the proof.

\[ \Box \]

REFERENCES