POMS PRODUCTION AND OPERATIONS MANAGEMENT

PRODUCTION AND OPERATIONS MANAGEMENT POMS

Vol. 25, No. 1, January 2016, pp. 9–21
ISSN 1059-1478 | EISSN 1937-5956 | 16 | 2501 | 0009
DOI 10.1111/poms.12511

Managing Inventory with Cash Register Information: Sales Recorded but Not Demands

Alain Bensoussan
Naveen Jindal School of Management, University of Texas at Dallas, Richardson, TX;
College of School of Engineering, Department SEEM, The City University of Hong Kong, Hong Kong, alain.bensoussan@utdallas.edu

Metin Çakanyildirim
Naveen Jindal School of Management, University of Texas at Dallas, 800 West Campbell Road, SM 30, Richardson, Texas 75080-3021, USA, metin@utdallas.edu

Meng Li
Charlton College of Business, University of Massachusetts Dartmouth, 285 Old Westport Road, North Dartmouth, Massachusetts 02747-2300, USA, mli1@umassd.edu

Suresh P. Sethi
Naveen Jindal School of Management, University of Texas at Dallas, 800 West Campbell Road, SM 30, Richardson, Texas 75080-3021, USA, sethi@utdallas.edu

Inventory inaccuracy is common in many businesses. While retailers employ cash registers to enter incoming orders and outgoing sales, inaccuracy arises because they do not record invisible demand such as spoilage, damage, pilferage, or returns. This setting results in incomplete inventory and demand information. An important inventory control problem therefore is to maximize the total expected discounted profit under this setting. Allowing for dependence between demand and invisible demand, we obtain the associated dynamic programming equation with an infinite-dimensional state space, and reduce it to a simpler form by employing the concept of unnormalized probability. We develop an analytical upper bound on the optimal profit as well as an iterative algorithm for an approximate solution of the problem. We compare profits of the iterative solution and the myopic solution, and then to the upper bound. We see that the iterative solution performs better than the myopic solution, and significantly so in many cases. Furthermore, it gives a profit not far from the upper bound, and is therefore close to optimal. Using our results, we also discuss meeting inventory service levels.

Key words: inventory uncertainty; Markov dynamic programming; invisible demand; censored demand
History: Received: August 2014; Accepted: March 2015 by Kalyan Singhal, after 1 revision.

1. Introduction

Almost every retail store is equipped with a *cash register* (or a similar device) that collects point-of-sale data. In a typical sales transaction for an item, a customer comes to a store and buys the quantity equal to his demand if there is sufficient inventory. The store clerk scans the (barcodes of) purchase, which enables the cash register to record the transaction. If the inventory is insufficient, the customer buys the entire inventory, and leaves the store with a portion of his demand unmet. The register updates the inventory record by deducting the quantity sold from the inventory level. The same (or a similar) system also processes order receipts from suppliers.

Such a cash register records outgoing sales and incoming ordered quantities, thus provides what we refer to as the *recorded inventory* levels of the items in a store. When these levels differ from their *actual inventories* that are available for sale, the inventory records are said to be incomplete. Inventory inaccuracy is sometimes represented as the proportion of items that have inaccurate records. Kang and Gershwin (2005) and DeHoratius and Raman (2008) report that 49% and 65% of stock-keeping units have inaccuracy in the respective retail chains they respectively investigated. Inventory inaccuracy can be measured as the percentage of sales. According to Miller and Allen (2005), the average inaccuracy rate for US supermarkets increased from 2.32% of sales in 2002 to 2.77% in 2005. Hollinger (2009) reports that US retail losses from inventory inaccuracy is 1.52% of sales, or 702 million a week, in 2008. Inventory inaccuracy for an item is often measured as the extent of discrepancy between the recorded inventory and the actual inventory at a given point of time. The chain in DeHoratius and Raman (2008) has more than 37 stores and uses advanced technologies, and yet "approximately 17% of the records … have … discrepancies that were 50% or more of the average on-hand quantity present in the stores." Reasons for inventory inaccuracy include spoilage, damage, pilferage, returns, and information delay; see Kang and Gershwin (2005), Lee and Özer (2007), Bensoussan et al. (2007a), Huh et al. (2010), Camdereli and Swaminathan (2010), and Sethi (2010). Depletion of inventory caused by these reasons are referred to as *invisible demand*.

An IM (inventory manager) can deal with incomplete inventory information (partially observed inventory) in three ways: prevention, correction, and integration (DeHoratius et al. 2008). As prevention, the IM can eliminate some of the root causes of inventory inaccuracy by implementing technologies such as RFID tags (Dutta et al. 2007, Heese 2007, Langer et al. 2007, De Kok et al. 2008). Using Walmart's data, Hardgrave et al. (2009) illustrate that RFID tags reduce inventory inaccuracy. But RFID technology cannot eliminate inaccuracy as it suffers from inaccurate read (item detection) rates. As correction, the IM can reconcile inventory records with actual inventories through an inventory count, inspection, or audit (Kang and Gershwin 2005). Fleisch and Tellkamp (2005) demonstrate that a correction of inventory inaccuracy can decrease supply chain costs as well as stockouts. As integration, the IM can incorporate the uncertainty (distribution) of inventory levels into the decision-making framework (Bensoussan et al. 2007c, 2008 2010, DeHoratius et al. 2008). Such incorporation can be used to evaluate the benefits of an RFID implementation (prevention) or frequent counts (correction), so these three approaches are not mutually exclusive. This study mainly focuses on integration but also alludes to correction.

We consider a discrete-time cash register inventory problem for a single item with lost sales, where a customer finding an empty shelf (a stockout) simply leaves without revealing his demand. Therefore, inventory and demand are both partially observed. In a retail store, events such as demand, theft and spoilage happen randomly without any prescribed sequence. "In reality, we might expect that visible and invisible demand arrivals would be interwoven throughout a day" (DeHoratius et al. 2008). Lee and Özer (2007) also discuss this interspersed nature of (visible) demand and invisible demand events. Given the available inventory, the sequence of events of both demand types determines the sales, which in turn influences the distribution of the inventory level. The simplest two of the sequences are demand-first (Bensoussan et al. 2013) and demand-last in which all the demand in a period events occur, respectively, before and after all invisible demand events. In another (demand-between) sequence, all demand events can be considered as a lump demand occurring between two separate sets of invisible demand events. While the demand-first and demand-last sequences provide two extreme cases, many more sequences are possible in real life, giving rise to intermediate sales. The sequence-based models are too specific, cumbersome to set up, and difficult to analyze. Moreover, the main reason for invoking a sequence is for obtaining sales, which could possibly be determined in alternative ways. So, as long as there is a process to obtain sales, sequences can be avoided to build models that are parsimonious and, at the same time, represent the real sales better than the extreme demand-first and demand-last sequences. The *allocation model* presented in this study produces such a process.

The allocation model works with two random variables, *demand sum* and *demand ratio*, defined as demand plus invisible demand and demand divided by demand sum, respectively. Our model splits demand sum into demand and invisible demand by using the random demand ratio, and gives us sales as a multiplication of demand ratio by the minimum of demand sum and available inventory. This mechanism can yield each and every sales that is sandwiched (inclusively) between the two extreme sales. A numerical example is provided in section 2 to illustrate the computation of sales with demand sum and demand ratio. Furthermore, every realized sales in the model arises from a set of sequences representing interspersed demand and invisible demand events that can occur in real life. Thus, our allocation model is formulated to provide a closer approximation of the reality than any specific sequence model. Finally, the allocation idea has been used by academics in similar contexts and is not unfamiliar to practitioners (Beck et al. 2003, Miller and Allen 2005, Hollinger 2009) who measure inventory inaccuracy in relative terms as percentages.

Given the sales, we can derive the inventory evolution and the Bellman equation whose state is a probability distribution and whose analysis must be carried out in the infinite dimensional space of probability distributions. Using the concept of unnormalized probability, we simplify the analysis and computations for implementation. We present two special cases of the Bellman equation for deterministic demand sum and deterministic demand ratio, respectively, and establish some monotonicity properties. We develop an iterative algorithm that can be converted into software for use by practitioners. Under various parameter settings deduced from the setting of a large retail chain (DeHoratius et al. 2008), the performances of the myopic policy and the iterative algorithm are compared with each other and also with an analytical upper bound. These comparisons reveal that our iterative algorithm beats the myopic policy always and comfortably in many settings. Only when the discount factor, the lost sales penalty cost, and the product price are all low, does the myopic policy perform nearly as well as the iterative algorithm. Moreover, when demand ratio is relatively large, the iterative algorithm produces a nearly optimal solution as inferred from the narrow gap between its profit and the upper bound.

Our study is closely related to the literature on inventory replenishment under inventory inaccuracy: Heese (2007), Kök and Shang (2007), DeHoratius et al. (2008), Atali et al. (2009), Camdereli and Swaminathan (2010), Huh et al. (2010), Mersereau (2013), Chen (2015), and Nathan et al. (2013). Among the recent inventory inaccuracy models, Heese (2007) and Camdereli and Swaminathan (2010) study single-period models with inventory shrinkage and inventory misplacement, respectively, in a supply chain. DeHoratius et al. (2008) focus on myopic (single-period) ordering policies and suggest a heuristic for inventory counting in a model with discrete inventories and demands. Mersereau (2013) analyzes in detail the model of DeHoratius et al. (2008), and provides an analytical comparison of the orders obtained from single-period and two-period settings. When the most recent inventory status known to the IM is dated due to information delays, then the current inventory is incompletely observed. In such cases, Bensoussan et al. (2007b, 2009) show a state-dependent basestock policy to be optimal. Kök and Shang (2007) and Huh et al. (2010) investigate the inventory counting policy in different models. The only information available to the IM in Bensoussan et al. (2007c) is a binary indicator that signals a stockout, whereas the IM in Bensoussan et al. (2008, 2011) observes the amount of backorder. Most papers in this literature stream assume an invisible demand that contributes to unobserved inventory inaccuracy. Modeling of this invisible demand is a key distinguishing feature among these papers. For example, the invisible demand is decoupled with the inventory (Kök and Shang 2007, Atali et al. 2009) or coupled with the inventory (DeHoratius et al. 2008, Bensoussan et al. 2007c). Also, the invisible demand can be modeled as additive (Kök and Shang 2007) or multiplicative (Heese 2007) relative to the inventory level, whereas invisible demand can be modeled as occurring before (DeHoratius et al. 2008, Huh et al. 2010), after, or between demands within a period.

This study strives to give a thorough analysis of profit maximization in an inventory system with only sales observations. The system is subject to censored demand and inventory inaccuracy. Anchoring on the idea that the sales information can be used to obtain the conditional probability of actual inventory, we develop and study a novel multi-period lost-sales model that uses a Bayesian updating mechanism for the distribution of the inventory level from the information recorded by the cash register. Under this mechanism, invisible demand affects the sales, which in turn affects the inventory distribution. Different from most of the previous works, we develop the allocation process for obtaining sales *without* focusing on the invisible demand.

Our contributions to the literature include the following: (i) We formulate and analyze an allocation model based on demand sum and demand ratio parametrization, and compare it with a model obtained from demand and invisible demand parametrization. (ii) We allow for *dependence* between demand sum and demand ratio (or between demand and invisible demand) for the first time in the literature. Demand and invisible demand can be dependent in practice. "For example, when a product is sought after by paying customers it will also be attractive for thieves" (Lee and Özer 2007). Furthermore, higher consumer traffic in a store leads to more consumer touches to and trials of products, and in turn to more damaged products. Both of these statements potentially imply correlation between the demand and invisible demand. The relationship between different demands is captured by the dependence of the demand sum and the demand ratio. (iii) We use unnormalized probability to simplify the inventory evolution and the Bellman equation, and to devise an iterative algorithm. We show that the value iteration in the algorithm converges to an upper bound on the profit. (iv) We develop an analytical upper bound to assess the performances of the iterative algorithm and the myopic policy. Numerical experiments based on real data are used to identify parameter ranges where these performances are significantly different, and to provide insights and guidelines to the IM.

For implementation, the IM can first compare his profit and demand parameters to those in our numerical experiments (subsection 3.2) in order to assess the benefit of using the iterative solution rather than the myopic solution before obtaining these solutions (subsection 3.3). This benefit is significant in many settings and likely to warrant implementation of the iterative solution. Because of the unnormalized probability approach (section 2), the iterative solution (subsection 3.1) is relatively easy to obtain. With his profit and demand parameters, the IM can also compute our upper bound and compare it to the profit resulting from the iterative solution to assess its proximity to the optimal solution.

2. The Model and Analysis

Our model is a single-item, lost-sales, periodic-review inventory model in which only sales are recorded, leaving the demands to be not fully observed. Invisible demands occur and cause inventory inaccuracy. Although IM does not observe the exact actual inventory without stocktaking (DeHoratius and Raman 2008), he can infer its distribution from the orders and the sales. We use the following notations for period t: I_t is beginning actual inventory, R_t is beginning recorded inventory, q_t is amount ordered, X_t is the demand sum (sum of the demand and invisible demand), z_{t+1} is observed sales, and Y_t is the demand ratio (the portion of demand sum attributable to the demand). Figure 1 shows the sequence of events in period t: Period t starts with actual and recorded inventories I_t and R_t, and the IM places order q_t and immediately receives it with no lead time. Then demand sum X_t occurs and reduces the actual inventory by $\min\{X_t, I_t + q_t\}$. Although demands and invisible demands occur throughout the period, we only illustrate the demand sum at a specific moment in Figure 1 for simplicity. Sales $z_{t+1} = Y_t \min\{X_t, I_t + q_t\}$ are recorded by the cash register, and the recorded inventory is reduced by the sales. This sales formulation is parsimonious in the sense that it does not require the detailed sequence of demand and invisible demand events during a period. Although we describe the model using the random variables demand sum and demand ratio for the sake of (sale and inventory formulations) simplicity, the analysis in this model applies as well to that using directly demand and invisible demand as the random variables.

We assume pairs (X_t, Y_t), $t = 1, 2, \ldots$, are independent and identically distributed. X_t and Y_t can depend on each other, and f_{xy} denotes the joint probability density of X_t and Y_t. For generality, we keep the supports of X_t and Y_t as the real line. These supports can be restricted to $[X_t \geq 0, 0 \leq Y_t \leq 1]$, for example, when the demand and invisible demand are both nonnegative, as in the case in most real-life problems.

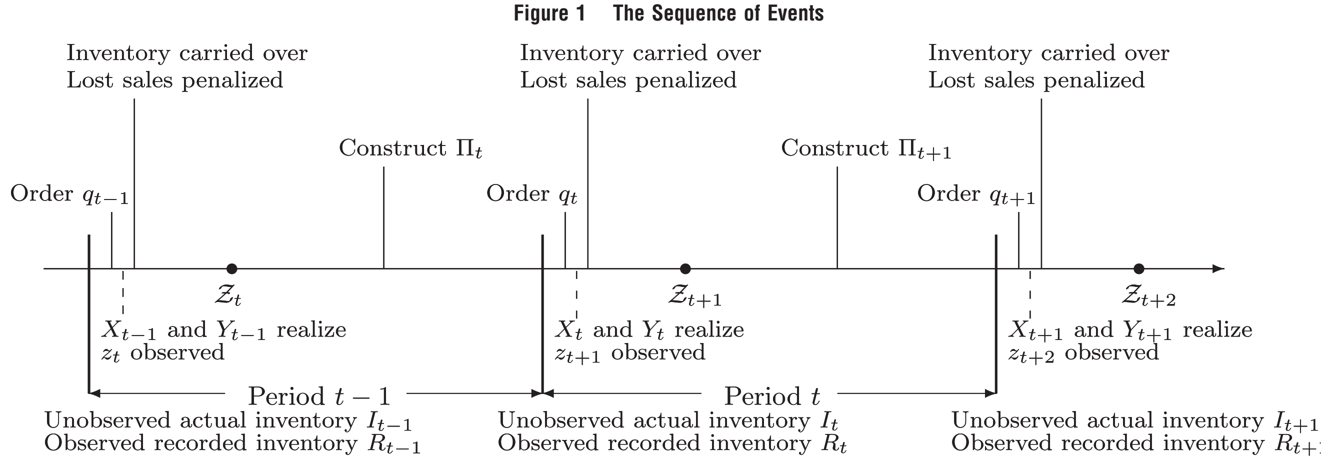
The actual inventory level in period t is $I_t + q_t$ immediately after the order is delivered; it is $I_t + q_t - \min\{X_t, I_t + q_t\}$ after the demand sum occurs. Combining these, we obtain $I_{t+1} = I_t + q_t - \min\{X_t, I_t + q_t\}$. This equation can be rewritten as follows using the notation $x^+ = \max\{0, x\}$:

$$I_{t+1} = (I_t + q_t - X_t)^+. \tag{1}$$

Refer to Kang and Gershwin (2005), Lee and Özer (2007), and Huh et al. (2010) for similar inventory formulas. Evolving as $R_{t+1} = (R_t + q_t - z_{t+1})^+$, the recorded inventory differs from the actual inventory because of the invisible demand. Orders based on the recorded inventory levels can be severely inaccurate (DeHoratius et al. 2008). In this study, both orders and profit are based on the actual inventory (or its distribution to be precise) rather than the recorded inventory. The recorded inventory levels can be computed with known orders and observed sales, hence they do not provide more information than sales. So only the actual inventory is considered and inventory without a qualifier refers to the actual inventory.

Based on prior observations, the IM considers the conditional distribution of the (actual) inventory. He observes only sales $z_{t+1} = Y_t \min\{I_t + q_t, X_t\}$ for $t \geq 1$. The term $\min\{I_t + q_t, X_t\}$ is the drop in the inventory level in period t as a result of the demand and invisible demand. If there is enough inventory to meet these demands, that is, $I_t + q_t \geq X_t$, then the sales observation is the demand, that is, $z_{t+1} = Y_t X_t$. Otherwise, $I_t + q_t < X_t$ and the inventory is insufficient to meet the demand sum. $Y_t X_t$ and $(1 - Y_t)X_t$ are two demand streams representing the demand and invisible demand, respectively. Inspired from the

Figure 1 The Sequence of Events



demand-rationing literature (Lee et al. 1997, Cachon and Lariviere 1999), which allocates the scarce inventory based on the amount of demands from different sources, we allocate $Y_t(I_t + q_t)$ of inventory to the demand and the remaining inventory to the invisible demand. Hence, the sales turns out to be $z_{t+1} = Y_t(I_t + q_t)$ if $I_t + q_t < X_t$. Combining the cases $I_t + q_t \geq X_t$ and $I_t + q_t < X_t$, we obtain $z_{t+1} = Y_t \min\{I_t + q_t, X_t\}$; for this sales formula, see Kang and Gershwin (2005) and Lee and Özer (2007). Similarly, Atali et al. (2009) obtain the sales as a sum of as many as $\min\{I_t + q_t, X_t\}$ Bernoulli trials in a discrete demand model.

In each period, there are two random demands: demand and invisible demand. In a discrete time model, strictly speaking, everything happens at a time epoch instantaneously and there is no concept of time between any two successive epochs. Thus, demand and invisible demand arrive instantaneously and they deplete the inventory in that proportion instantaneously, which is a reasonable way to look at the situation. Alternatively, we can think of the interval between the two successive epochs as a unit of continuous time (say, a day) and the visible and invisible demands to arrive at their rates (say 10 and 90) per day throughout the day. In this case, if the available inventory is $I_t + q_t = 10$ on day t, then it will be depleted during the first tenth of the day resulting in the sale of 1. With $X_t = 100$, the sales will vary from 0 to 10 depending on the realized value of Y_t; for example, if Y_t were to realize as 0.2, the sales would be 2. Needless to say, the sales also vary with the available inventory; thus with $X_t = 100, Y_t = 0.2$, and the available inventory 1, the sales will be 1. In summary, X_t and Y_t together determine the rates of two demand types that will materialize during the day, and that with the available inventory, in turn, will determine the amount sold. Thinking in this way avoids the complications of introducing sequences of discrete demands along with their amounts during the day. These complications would require a model with many components to estimate and far too complex to be tractable.

The IM generates revenue from the sales and incurs an ordering cost when placing orders. Furthermore, an inventory holding cost is incurred when there is excess inventory. Otherwise, there is lost sales resulting in a shortage cost. The sales revenue, inventory ordering, holding and shortage costs are captured by the single-period profit function $\tau(I_t, q_t)$ that depends on inventory level I_t and order size q_t. The profit function can be as general as

$$\tau(I_t, q_t) = \mathrm{E}[r \min\{X_t Y_t, I_t + q_t\} - h(I_t + q_t - X_t)^+ - p(X_t Y_t - I_t - q_t)^+] - c q_t, \tag{2}$$

where c, h, p, r, and s are unit order cost, holding cost, shortage penalty cost, and retail price, respectively. The function $\tau(I_t, q_t)$ is the profit for the known inventory level I_t. When I_t is a random variable with distribution Π_t, the single-period profit is found by taking expectation over I_t as well: $\mathrm{E}\tau(I_t, q_t) = \int \tau(x, q_t) d\Pi(x)$. As a convention, when a lower limit or an upper limit is not shown in an integral, it is understood to be zero or infinity, respectively.

The IM decides q_t based on the information $\mathcal{Z}_t$ available from the cash register at the beginning of period $t \geq 2$, which is generated by the sales $\{z_2, z_3, \ldots, z_t\}$. Without the exact knowledge of the inventory, the IM evaluates the profit as the expected value based on the inventory distribution. That is, the (conditionally) expected profit in period t is $\alpha^{t-1}\mathrm{E}[\tau(I_t, q_t)|\mathcal{Z}_t]$, where $\alpha \in [0, 1)$ is the discount factor. Given a policy $\tilde{q}$, the cumulative probability law $\Pi(\cdot)$ on the initial inventory level, the total expected discounted profit can be written as $J(\Pi, \tilde{q}) := \mathrm{E}\sum_{t=1}^{\infty} \alpha^{t-1}\tau(I_t, q_t) = \mathrm{E}\sum_{t=1}^{\infty} \alpha^{t-1}\mathrm{E}[\tau(I_t, q_t) \quad |\mathbb{Z}_t]$. The optimal profit $V(\Pi) = \sup_{\tilde{q}} J(\Pi, \tilde{q})$. Note that the state Π in the optimal profit function is a probability law. Our analysis starts with an infinite horizon problem, but it applies to finite horizon settings.

The IM can derive the conditional probability $\Pi_t(\cdot)$ of I_t given $\mathcal{Z}_t$, that is, $\mathrm{P}(I_t \leq \eta | \mathcal{Z}_t) = \Pi_t(\eta)$.

Theorem 1. *The conditional probability of the inventory evolves as*

$$\Pi_{t+1}(\eta) = \frac{\displaystyle\int \theta(\lambda + q_t, z_{t+1}) d\Pi_t(\lambda) + \int^{\eta} \int \vartheta(\lambda + q_t - x, z_{t+1}) d\Pi_t(\lambda) dx}{\displaystyle\int \theta(\lambda + q_t, z_{t+1}) d\Pi_t(\lambda) + \int \int \vartheta(\lambda + q_t - x, z_{t+1}) d\Pi_t(\lambda) dx}, \tag{3}$$

$\Pi_0(\cdot)$ *given,*

where $\theta(u, v) := \int_u f_{xy}(x, v/u) dx/u$ *and* $\vartheta(u, v) := f_{xy}(u, v/u)/u$.

From Equation (3) and the proof of Theorem 1, we observe that the inventory distribution has a mass at zero. Recognizing this mass, we can write $\Pi_t(\eta) = \pi_t(0) + \int^{\eta} \pi_t(x) dx$ for a function π_t that captures the mass and the continuous part. The derivation of Equation (3) is technically interesting. When inventories and demands are discrete, the derivation of inventory evolution generally follows an intuitive approach. That is, the event for the inventory level taking a certain value has a probability, and it can be intuitively mapped by Equation (1) to events causing this inventory level. However, this intuitive mapping does not exist in the continuous case because the probability of a specific inventory value is zero in general. This somewhat explains the long-time prevalence of partially observed discrete models in the operations literature (Smallwood and Sondik 1973, Monahan 1982, Lovejoy 1991, DeHoratius et al. 2008, Huh et al. 2010, Saghafian 2015). Here, we derive the inventory evolution for continuous inventories/demands to complement the current literature, because the continuous inventories/demands have been as common as their discrete counterparts in the inventory control literature. Furthermore, the framework we use in Appendix S1 is quite general and with a potential to be adopted in other applications with partially observed (continuous) states.

To provide insight into Equation (3), we consider two events $[X_t \geq I_t + q_t]$ and $[X_t < I_t + q_t]$. Under $[X_t \geq I_t + q_t]$, $I_{t+1} = 0$ and $z_{t+1} = Y_t(I_t + q_t)$, that is, $Y_t = z_{t+1}/(I_t + q_t)$. This event corresponds to the first terms in the numerator and denominator of Equation (3). Under $[X_t < I_t + q_t]$, the next period's inventory is $I_{t+1} = x = I_t + q_t - X_t \leq \eta$ and $z_{t+1} = Y_t X_t$. Consequently, $X_t = I_t + q_t - x$ and $Y_t = z_{t+1}/(I_t + q_t - x)$. This event leads to the second terms in the numerator and denominator of Equation (3). Π_t captures all of the sales observations in the previous periods, and it is as sufficient as observed sales for inferring I_t. Hence, Π_t can serve as the system state.

Theorem 1 can also help the IM to meet an inventory service level. When the shortage cost, being intangible, is difficult to estimate, then the IM cannot construct $\tau(\cdot, \cdot)$, and can instead opt for satisfying a specified service level. Also when the IM wants to benchmark his inventory management performance against the others, he is more likely to compare the service levels than the profits. This is because the service level at a company does not depend on the size of that company. In these cases, the IM can adopt a type-I service level and require that the demand is met with some probability, say 90%. This translates to $\int \mathrm{P}(X_t Y_t \leq x + q_t)\, d\Pi_t(x) \geq 0.9$ in period t. Since the left-hand side is increasing in q_t, the IM can easily search for the unique q_t that satisfies this service level. The conditional probability Π_t allows us to assess and maintain the specified service level.

In Equation (3), Π_t evolves according to a highly nonlinear equation akin to the Kushner equation (Kushner 1967). The Zakai (1969) equation provides a linear recursion of the conditional probability for models whose uncertainty is represented by a Brownian motion. The Zakai equation uses unnormalized probability, which is obtained from the likelihood functions and the initial prior probability law (without performing the normalization step). Analogous to Π_t, the unnormalized probability Φ_t is obtained as follows:

$$\begin{aligned}\Phi_{t+1}(\eta) &= \int \theta(\lambda + q_t, z_{t+1})d\Phi_t(\lambda) \\ &\quad + \int^{\eta}\int \vartheta(\lambda + q_t - x, z_{t+1})d\Phi_t(\lambda)dx, \\ \Phi_0(\cdot) &= \Pi_0(\cdot).\end{aligned} \tag{4}$$

More importantly, as justified in Appendix S1,

$$\Pi_t(x) = \frac{\Phi_t(x)}{\int d\Phi_t(x)}. \tag{5}$$

Φ_t inherits the mass from Π_t, so $\Phi_t(\eta) = \phi_t(0) + \int^{\eta}\phi_t(x)dx$ for a function ϕ_t that captures the mass and the continuous part. Since Π_t integrates to 1 while Φ_t does not, Φ_t is called the unnormalized probability. Equation (4) is linear in Φ_t. We mainly use the unnormalized probability in the remainder and justify this below.

We now proceed to derive the dynamic programming equation for the infinite horizon cash register problem. With the current inventory level distribution $\Pi(\cdot)$, the profit in the current period is $\int\tau(x, q)\Pi(dx)$. Adding future profits to this, which is formalized in Appendix S1, leads to the Bellman equation

$$\begin{aligned}V(\Pi(\cdot)) &= \sup_q \Bigg\{ \int \tau(x,q)d\Pi(x) \\ &+ \alpha \int \left[\int \theta(\lambda+q,z)d\Pi(\lambda) + \int\int \vartheta(\lambda+q-x,z)d\Pi(\lambda)dx \right] \\ &\times V\left(\frac{\int \theta(\lambda+q,z)d\Pi(\lambda) + \int^{(\cdot)}\int \vartheta(\lambda+q-x,z)d\Pi(\lambda)dx}{\int \theta(\lambda+q,z)d\Pi(\lambda) + \int\int \vartheta(\lambda+q-x,z)d\Pi(\lambda)dx} \right) dz \Bigg\}.\end{aligned} \tag{6}$$

The nonlinear evolution of Π in Equation (3) makes a direct study of Equation (6) difficult. If we instead use the unnormalized probability Φ as the system state, then the analysis becomes easier, because Φ evolves linearly. To make the ideas concrete, we define

$$W(\Phi) := V\left(\frac{\Phi(x)}{\int d\Phi(x)}\right) \int d\Phi(x). \tag{7}$$

Using the Bellman Equation (6) and some algebra, we can show that W satisfies the Bellman equation

$$\begin{aligned}W(\Phi(\cdot)) &= \sup_q \Bigg\{ \int \tau(\eta,q)d\Phi(\eta) + \alpha \int W\Bigg(\int \theta(\lambda+q,z)d\Phi(\lambda) \\ &\quad + \int^{(\cdot)}\int \vartheta(\lambda+q-x,z)d\Phi(\lambda)dx \Bigg) dz \Bigg\}.\end{aligned} \tag{8}$$

The update of Φ_t in Equation (4) is linear, while the update of Π_t in Equation (3) is nonlinear. That is, Equation (8) does not have a denominator that involves the control variable q. After obtaining the Bellman equation, one needs to show that the Bellman equation has a solution and that the solution is indeed the value function. This analysis is facilitated by the linearity of Equation (8); see for example, Bensoussan et al. (2012). From Equation (7), W exactly matches V when the argument Φ is a probability distribution. To calculate the optimal policy for our model, we only need to find W. The solution of Equation (8) consists of the functionals $W(\Phi)$ and $q(\Phi)$, where $q(\Phi)$ provides the optimal feedback policy for our model. We can now specify the optimal order quantity in any given period t as follows.

THEOREM 2. *The IM's optimal order quantity is $q_t = q(\Phi_t)$, where Φ_t is the unnormalized distribution of inventory in period $t \geq 1$.*

Equation (8) provides $q(\Phi)$ in terms of the unnormalized probability Φ. In subsection 3.1, we develop an iterative solution approach based on Equation (8) and Equation (3). Alternatively, since the normalized probability Π also belongs in the set of unnormalized probabilities, the functional $q(\cdot)$ also applies to Π. As a result, if IM can assess the current distribution of the inventory by any means such as through expert opinion, then he can insert that assessment in $q(\cdot)$ to obtain the order quantity in a heuristic fashion.

The value function W has some useful properties and simplifies when demand sum or ratio is deterministic. Despite these simplifications, the argument of W remains to be a function. These are provided in the next theorem.

THEOREM 3.

(a) The value function W has the following properties: (i) $W(0) = 0$, $W(\kappa\Phi) = \kappa W(\Phi)$ for $\kappa \geq 0$, (ii) $W(\kappa_1\Phi_1 + \kappa_2\Phi_2) \leq \kappa_1 W(\Phi_1) + \kappa_2 W(\Phi_2)$ for $\kappa_1, \kappa_2 \geq 0$, and (iii) $W(\Phi_1) \leq W(\Phi_2)$ if $d\Phi_1(\eta) \leq d\Phi_2(\eta)$ for all η.

(b) When the demand ratio is deterministic at y_0, the value function is given by

$$\begin{aligned} W(\Phi(\cdot); X, y_0) = \sup_q \bigg\{ &\int \tau(\eta, q; X, y_0) d\Phi(\eta) \\ &+ \alpha y_0 \int \frac{1}{z} W\big(\bar{F}_x(z)\phi(z-q) \\ &+ f_x(z) \int^{(\cdot)} \phi(z-q+x)dx; X, y_0\big)dz \bigg\}. \end{aligned} \tag{9}$$

(c) When the demand sum is deterministic at x_0, the value function for the scaled ratio κY is given by

$$\begin{aligned} W(\Phi(\cdot); x_0, \kappa Y) = \sup_q \bigg\{ &\int \tau(\eta, q; x_0, \kappa Y) d\Phi(\eta) \\ &+ \alpha \int W\bigg(\int_0^{(x_0-q)^+} \frac{f_y(\frac{z}{\lambda+q})}{\lambda+q} d\Phi(\lambda) \\ &+ \frac{f_y(\frac{z}{x_0})}{x_0} \int_{(x_0-q)^+}^{(x_0-q+\cdot)^+} d\Phi(\lambda); x_0, \kappa Y\bigg)dz \bigg\}. \end{aligned} \tag{10}$$

So the scaling of random demand ratio does not affect the evolution of the unnormalized inventory distribution.

(d) If the shortage cost $p = 0$, the single-period profit $\tau(\eta, q; X, Y)$ increases stochastically in X or Y. In particular, $\tau(\eta, q; X, y_0)$ and $\tau(\eta, q; X, \kappa Y)$, respectively, increase in y_0 and κ. Consequently, $W(\Phi(\cdot); X, y_0)$ and $W(\Phi(\cdot); x_0, \kappa Y)$ increase in y_0 and κ, resepctively.

Theorem 3a shows that the value function W is homogeneous of degree 1 and subadditive, which together imply the convexity of W. Theorem 3b–c provide the value functions in the special cases and lead to Theorem 3d that establishes the sensitivity of the profit. The sensitivity discussion requires $p = 0$ to guarantee that $\tau(\eta, q; X, Y)$ increases stochastically in X or Y. If $p > 0$, even the deterministic profit $\tau(\eta, q; x_0, y_0) = r \min\{x_0 y_0, \eta + q\} - h(\eta + q - x_0)^+ - p(x_0 y_0 - \eta - q)^+ - cq$ can decrease in x_0 and y_0 when $x_0 \min\{1, y_0\} > \eta + q$. Note that only in Theorem 3b-d, we use the notations $W(\Phi; x, y)$ and $\tau(I, q; x, y)$ rather than $W(\Phi)$ and $\tau(I, q)$ to highlight the demand sum x and demand ratio y. In the rest of the study, we continue to use $W(\Phi)$ and $\tau(I, q)$.

3. Computation, Insights and Concluding Remarks

For an implementation of our results, we present an iterative algorithm. The second subsection includes a comprehensive computational study whose objectives are twofold: (i) to compare our iterative solution and the myopic policy; (ii) to provide insights to the IM as to when the iterative algorithm can be used with confidence. The last subsection has a summary of the insights. We choose to work with an infinite horizon problem to keep the notation simple.

3.1. Iterative Algorithm

The iterative algorithm is designed to account for the profits in all of the future periods, which are ignored by the myopic model. A myopic model maximizes only the profit incurred in the current period $L(q, \Phi) := \int \tau(x, q) d\Phi(x)$, that is, if the current inventory has the distribution Φ, then the myopic ordering quantity is $\arg\max_q L(q, \Phi)$. The associated myopic policy can be optimal for a fully observed multi-period problem (Theorem 6.2 of Porteus 2002). When the inventory is partially observed, the computational complexity increases significantly. Then, the myopic policy is generally regarded as a useful heuristic (DeHoratius et al. 2008). To the extent that future profits constitute a larger portion of the total profit (e.g., a larger α corresponding to a smaller discounting rate), the solution of the iterative algorithm will differ more from the myopic solution.

3.1.1. Discretization. Following the standard solution approach in the literature, we first obtain a discrete version of the system state. In this version, demand sum, inventory and demand ratio can all be discrete. For example, inventory and demand sum can be integers, and demand ratio is a rational number. Then the inventory remains an integer by Equation (1) and sales are rational numbers. As justified in Appendix S1, the inventory evolution for the discrete model is

$$\Pi_{t+1}(\eta) = \frac{\sum_{\lambda\geq 0}\theta_d(\lambda+q_t, z_{t+1})\mathrm{p}_t(\lambda) + \sum_{0<x\leq\eta}\sum_{\lambda\geq 0}\vartheta_d(\lambda+q_t-x, z_{t+1})\mathrm{p}_t(\lambda)}{\sum_{\lambda\geq 0}\theta_d(\lambda+q_t, z_{t+1})\mathrm{p}_t(\lambda) + \sum_{x>0}\sum_{\lambda\geq 0}\vartheta_d(\lambda+q_t-x, z_{t+1})\mathrm{p}_t(\lambda)}, \quad (11)$$

where $\mathrm{p}_t(x) = \mathrm{P}(I_t = x)$, $\vartheta_d(u, v) = f_{xy}(u, v/u)$, $\theta_d(u, v) = \sum_{x\geq u}\vartheta_d(x, v)$ for $u \neq 0$, $\vartheta_d(0, v) = \mathbb{1}_{v=0}\mathrm{P}(X_t = v)$, $\theta_d(0, v) = \mathbb{1}_{v=0}\mathrm{P}(X_t \geq v)$ and f_{xy} is joint probability mass of X_t and Y_t. Note that $\vartheta_d(0, v) = 0$ for $v \neq 0$ and $v_d(0, 0) = \sum_y f_{xy}(0, y)$. The discrete evolution is different from the continuous evolution as the terms in the integrals of Equation (3) and in the sums of Equation (11) differ by u, that is, $\theta(u, v) = \theta_d(u, v)/u =$ and $\vartheta(u, v) = \vartheta_d(u, v)/u$.

Using the discrete version of the system state and the unnormalized probability corresponding to Equation (11), we set

$$\begin{aligned}\Phi_{t+1}(\eta) &= p(q_t, \Phi_t, z_{t+1})\\ &= \sum_{\lambda\geq 0}\theta_d(\lambda+q_t, z_{t+1})\rho_t(\lambda)\\ &\quad + \sum_{0<x\leq\eta}\sum_{\lambda\geq 0}\vartheta_d(\lambda+q_t-x, z_{t+1})\rho_t(\lambda),\end{aligned} \quad (12)$$

where $\rho_t(\lambda) = \Phi_t(\lambda) - \Phi_t(\lambda - 1)$, and $p(\cdot,\cdot,\cdot)$ is an operator. Let $L(q, \Phi) = \sum_{\lambda\geq 0}\tau(\lambda, q)\rho(\lambda)$, we have the Bellman equation

$$W(\Phi) = \max_q\left\{L(q, \Phi) + \alpha\sum_{z\geq 0}W(p(q, \Phi, z))\right\}. \quad (13)$$

To improve on the myopic policy, we develop a value iteration algorithm for solving system Equation (13). We set the initial solution as $W_0(\Phi) = 0$ and define the following iteration:

$$W_{n+1}(\Phi) = \max_q\left\{L(q, \Phi) + \alpha\sum_{z\geq 0}W_n(p(q, \Phi, z))\right\}, \quad (14)$$

where W_n represents the value function in the n-th iteration. The iteration generates an increasing sequence of functions $\{W_0, W_1, \ldots\}$ that converges to a solution of Equation (13). Obtaining this solution is not straightforward since $p(q, \cdot, z)$ is an infinite-dimensional object. So an iterative algorithm is developed and its steps are discussed below.

3.1.2. Initialization. In the first iteration of the value iteration process, we set $n = 1$ and $W_0(\cdot) = 0$, and calculate $W_1(\Phi) = \max_q L(q, \Phi)$. When the first iteration is completed, we have two numbers, the order quantity $\hat{q}_1 = \arg\max_q L(q, \Phi)$ and the associated profit $W_1(\Phi)$. Then, we go to the next iteration.

3.1.3. Iteration. To calculate $W_2(\Phi)$ in the second iteration $n = 2$, we need to insert $W_1(p(q, \Phi, z))$ into Equation (14) to obtain

$$W_2(\Phi) = \max_q\left\{L(q, \Phi) + \alpha\sum_{z\geq 0}W_1(p(q, \Phi, z))\right\}. \quad (15)$$

Using the values obtained in iteration $n - 1$ to solve for the values in iteration n is a standard approach to solve dynamic programming. But while calculating $W_1(p(q, \Phi, z))$, the standard approach does not work for the following reason: In the first iteration, we calculate $W_1(\cdot)$ for a fixed inventory law Φ only, so $W_1(p(q, \Phi, z))$ is not available in general. Alternatively, $W_1(\cdot)$ can be computed for each probability law Φ, which is impossible to do for all Φ because of the uncountability of the set containing Φ.

The computational issues can be bypassed by using approximations (Powell 2007). Our approximation uses the previously computed $\hat{q}_1, \ldots, \hat{q}_{n-1}$, when evaluating $W_n(p(q, \cdot, z))$. We specify the corresponding profit $\hat{W}_n$ as follows:

$$\hat{W}_n(\cdot) = \max_{q_n}\left\{L(q_n, \cdot) + \alpha\sum_{z_{n+1}\geq 0}\hat{W}_{n-1}(p(q_n, \cdot, z_{n+1}))\right\}, \quad (16)$$

where $\hat{W}_{n-1}$ is defined by starting with $\hat{W}_0 = 0$ and by recursively proceeding via

$$\begin{aligned}\hat{W}_{n-1}(p(q_n,\cdot,z_{n+1})) &= L(\hat{q}_{n-1},p(q_n,\cdot,z_{n+1}))\\ &+\alpha\sum_{z_n\geq 0}\hat{W}_{n-2}(p(\hat{q}_{n-1},p(q_n,\cdot,z_{n+1}),z_n))\\ &=\sum_{i=0}^{n-2}\sum_{z_{n-i+1},,z_n\geq 0}\alpha^i L(\hat{q}_{n-1-i},p(\hat{q}_{n-i},\ldots,p(\hat{q}_{n-1},\\ &\quad p(q_n,\cdot,z_{n+1}),z_n),\ldots,z_{n-i+1})),\end{aligned} \tag{17}$$

with the understanding that the last sum above has no z to sum over when $i = 0$. $\hat{W}_{n-1}(p(q_n, \cdot, z_{n+1})$ is the discounted profit of implementing the sequence of previous computed orders $\{\hat{q}_{n-1},\ldots,\hat{q}_1\}$ when the current state is $p(q_n, \cdot, z_{n+1})$. Equations (16–17) illustrate how the profit and the order quantities are computed in the iterative algorithm, whose pseudocode is given in Table 1.

As an illustration of computations in the second iteration, we approximate the value function $W_2(\Phi)$ with

$$\begin{aligned}\hat{W}_2(\Phi) &= \max_{q_2}\left\{L(q_2,\Phi)+\alpha\sum_{z_3\geq 0}\hat{W}_1(p(q_2,\Phi,z_3))\right\}\\ &=\max_{q_2}\left\{L(q_2,\Phi)+\alpha\sum_{z_3\geq 0}L(\hat{q}_1,p(q_2,\Phi,z_3))\right\},\end{aligned} \tag{18}$$

Table 1 Iterative Algorithm

Inputs: Distribution Φ, profit function $L(\cdot,\cdot)$, operator $p(\cdot,\cdot,\cdot)$ and accuracy parameter ε (≥ 0).
Initialize $n = 1$; $\hat{q}_1 = \arg\max_q L(q,\Phi)$, $\hat{W}_1(\Phi) = L(\hat{q}_1,\Phi)$. Save $\hat{q}_1$ and $\hat{W}_1(\Phi)$.
Repeat $n = n + 1$;

$$\hat{q}_n = \arg\max_q\left\{L(q,\Phi)+\sum_{i=0}^{n-2}\sum_{z_{n-i+1},\ldots,z_{n+1}\geq 0}\alpha^{i+1}L(\hat{q}_{n-1-i},p(\hat{q}_{n-i},\ldots,p(q,\Phi,z_{n+1}),\ldots,z_{n-i+1}))\right\}.$$

Save $\hat{q}_n$ and $\hat{W}_n(\Phi)$, which is the value of the terms inside the curly brackets above at $\hat{q}_n$.
Until value functions satisfy $|\hat{W}_n(\Phi)-\hat{W}_{n-1}(\Phi)|\leq\epsilon$.
Outputs: Sequences of $\{\hat{q}_n(\Phi)\}$ and $\{\hat{W}_n(\Phi)\}$.

where $\hat{q}_1$ is obtained in the first iteration. Now, rather than calculating the order quantity q_1 that maximizes $L(q_1, p(q_2, \Phi, z_3))$, we use the order quantity $\hat{q}_1$ to calculate $L(\hat{q}_1, p(q_2, \Phi, z_3))$. This only approximates $\max_{q_1} L(q_1, p(q_2, \Phi, z_3)) = W_1(p(q_2, \Phi, z_3))$ because $\hat{q}_1 \neq \arg\max_q L(q, p(q_2, \Phi, z_3))$ in general. Then according to Equation (18), $\hat{W}_2(\Phi)$ is evaluated by implementing $\hat{q}_2$, which maximizes the profit in the brackets in Equation (18), in period 1 and $\hat{q}_1$ in period 2.

For a more detailed illustration of the second iteration, let $I_1 = 1$ and $f_{xy}(0,0) = f_{xy}(1,0) = f_{xy}(0,1) = f_{xy}(1,1) = 1/4$. To find the maximizer $\hat{q}_2$ of the profit in the brackets in Equation (18), we evaluate the profit at various q_2 values. That is,

$$\begin{aligned}&L(q_2,\Phi)+\alpha\sum_{z_3\geq 0}L(\hat{q}_1,p(q_2,\Phi,z_3))\\ &=\begin{cases}L(0,\Phi)+\alpha\sum_{z_3\geq 0}L(\hat{q}_1,p(0,\Phi,z_3)), & q_2=0,\\ L(1,\Phi)+\alpha\sum_{z_3\geq 0}L(\hat{q}_1,p(1,\Phi,z_3)), & q_2=1.\end{cases}\end{aligned}$$

Suppose $q_2 = 0$, then

$$\begin{aligned}&L(0,\Phi)+\alpha\sum_{z_3\geq 0}L(\hat{q}_1,p(0,\Phi,z_3))\\ &=\tau(1,0)+\alpha[L(\hat{q}_1,p(0,\Phi,0))+L(\hat{q}_1,p(0,\Phi,1))],\end{aligned} \tag{19}$$

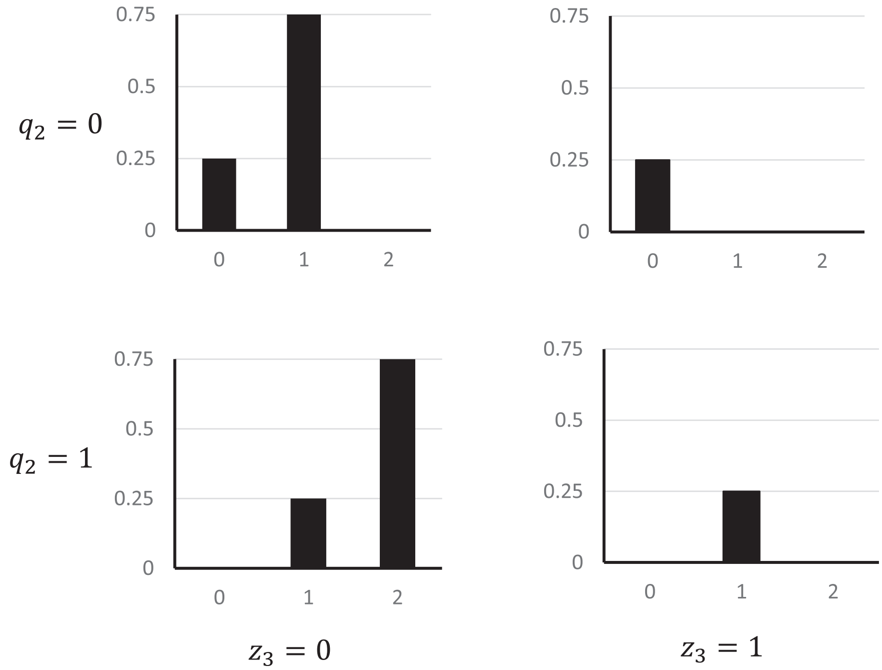
where $\hat{q}_1 = \arg\max_q \tau(1,q)$, and $\tau(1,0)$ can be obtained by Equation (2). Thus, we are left to evaluate only $p(0, \Phi, 0)$ and $p(0, \Phi, 1)$ to compute Equation (19). From Equation (12), $p(0, \Phi, 0) = \theta_d(1,0) + \sum_{0<x\leq\eta}\vartheta_d(1-x,0)$. Hence, $p(0, \Phi, 0) = 1/4$ when $\eta = 0$ and $p(0, \Phi, 0) = 3/4$ when $\eta = 1$; see the top left subfigure in Figure 2. Furthermore, $p(0, \Phi, 1) = \theta_d(1,1) + \sum_{0<x\leq\eta}\vartheta_d(1-x,1)$. As a result, $p(0, \Phi, 1) = 1/4$ when $\eta = 0$ and $p(0, \Phi, 1) = 1/4$ when $\eta = 1$; see the top right subfigure in Figure 2. With $p(0, \Phi, 0)$ and $p(0, \Phi, 1)$ in hand, we can evaluate Equation (19). Similarly, we can do above computations for $q_2 = 1$,

$$\begin{aligned}&L(q_2=1,\Phi)+\alpha\sum_{z_3\geq 0}L(\hat{q}_1,p(q_2=1,\Phi,z_3))\\ &=L(1,\Phi)+\alpha\sum_{z_3\geq 0}L(\hat{q}_1,p(1,\Phi,z_3))\\ &=\tau(1,1)+\alpha[L(\hat{q}_1,p(1,\Phi,0))+L(\hat{q}_1,p(1,\Phi,1))].\end{aligned}$$

From Equation (12), $p(1, \Phi, 0) = \theta_d(2,0) + \sum_{0<x\leq\eta}\vartheta_d(2-x,0)$. Hence, $p(1, \Phi, 0) = 1/4$ when $\eta = 1$ and $p(1, \Phi, 0) = 3/4$ when $\eta = 2$; see the bottom left subfigure in Figure 2. Furthermore, $p(1, \Phi, 1) = \theta_d(2,1) + \sum_{0<x\leq\eta}\vartheta_d(2-x,1)$. As a result, $p(1, \Phi, 1) = 1/4$ when $\eta = 1$ and $p(1, \Phi, 1) = 1/4$ when $\eta = 2$; see the bottom right subfigure in Figure 2. Repeating this process for other q_2 values, we can find the solution $\hat{q}_2$ for Equation (12). At the end of the second iteration, we have $\hat{q}_2$ from the second iteration, and $\hat{q}_1$ from the first iteration. Repeating the computations in this paragraph for each $n \geq 3$ yields $\hat{W}_n$ in Equations (16–17) and $\hat{q}_n$. In the n-th iteration, we determine $\hat{q}_n$ by maximizing the total profit resulting from orders $q_n, \hat{q}_{n-1}, \hat{q}_{n-2}, \ldots, \hat{q}_1$.

We note that the approximation of the value function using the available order quantities is new and not adopted from the general approaches available in the literature. This approximation is motivated by the

Figure 2 Evolution of Unnormalized Probability for I_2 Starting with $I_1 = 1$ and for $q_2, z_3 \in \{0, 1\}$



intuition that accounting future profits obtained with suboptimal profits is better than ignoring these profits. An interesting feature of our solution procedure is that it uses the Bellman equation associated with the unnormalized probability, which along with our approximation reduces the computational burden. The running time of computing the nested applications of the operator p in Equation (17) is linear in n and the number of sums to compute in Equation (17) is at most n^2, so the running time for each iteration (computation of q_n and $\hat{W}_n$) is n^3. The algorithm runs in n^4 time, if it finishes in n iterations. The memory required to store q_n and $\hat{W}_n(\Phi)$ is linear in n. We highlight these features here because they have potential for use in other applications with incomplete information.

3.1.4. Termination. Theorem 4 below validates the convergence of the iterative algorithm. This gives us the stopping rule specified in Table 1 for terminating our algorithm when the difference between the two consecutive computations of the value function becomes small.

THEOREM 4. *The sequence* $\{\hat{W}_n\}$ *obtained by Equations* (16–17) *with* $\hat{W}_0(\cdot) = 0$ *is increasing and it converges pointwise, that is,* $\hat{W}_n(\cdot) \nearrow \hat{W}(\cdot)$. *Moreover,* $\hat{W}(\cdot) \leq W(\cdot)$, *where* $W(\cdot)$ *is the solution of the Bellman equation* (13).

When the initial inventory law is Φ and accuracy parameter is ϵ , let $N(\Phi, \epsilon)$ be the number of iterations performed by the iterative algorithm. The algorithm outputs order quantities $\{\hat{q}_1(\Phi), \hat{q}_2(\Phi), \ldots, \hat{q}_{N(\Phi,\epsilon)}(\Phi)\}$ and approximate value functions $\{\hat{W}_1(\Phi), \hat{W}_2(\Phi), \ldots, \hat{W}_{N(\Phi,\epsilon)}(\Phi)\}$. The limit of these value functions (as $\epsilon \to 0$) forms a lower bound on the optimal profit.

3.1.5. Evaluation. Since the lower bound provided by our algorithm and the upper bound obtained in Theorem 5 below sandwich the optimal profit, they are used in the numerical evaluations of the profit associated with the myopic policy and that obtained from our iterative algorithm.

THEOREM 5. $V(\Pi(\cdot)) \leq \int v(\eta) d\Pi(\eta)$, *where* $v(x) = \sup_q \{\tau(x, q) + \alpha E v((x + q - X_t)^+)\}$.

The upper bound in Theorem 5 is calculated by averaging the optimal profits of the fully observed systems with different initial inventories weighted by their probabilities. It corresponds to the situation in which the inventory is initially unknown, but becomes known afterwards, and so it is not the optimal profit of any given fully observed model.

3.1.6. Rolling Horizon Implementation. For a problem with N time periods and the initial inventory law Φ_1, the iterative algorithm outputs $\{\hat{q}_1(\Phi_1), \hat{q}_2(\Phi_1), \ldots, \hat{q}_N(\Phi_1)\}$, and $\hat{q}_N(\Phi_1)$ is implemented in period 1 in a rolling horizon implementation. At the beginning of period 2, Φ_2 is obtained by inserting the sales observation of period 1 into Equation (4). From period 2 onwards, IM has a problem with $N - 1$ time periods and Φ_2. For this problem, the iterative algorithm outputs $\{\hat{q}_1(\Phi_2), \hat{q}_2(\Phi_2), \ldots, \hat{q}_{N-1}(\Phi_2)\}$, and $\hat{q}_{N-1}(\Phi_2)$ is implemented in period 2. Repeating this process, IM orders $\hat{q}_{N-i+1}(\Phi_i)$ in period i for $1 \le i \le N$ in a rolling horizon implementation.

3.2. Numerical Experiments

In the experiments, the base case has the penalty cost of $p = 0.75$ per unit of lost sales, the holding cost of $h = 0.03$ per unit per period, the ordering cost of $c = 0.3$ per unit, and the retail price of $r = 1.0$ per unit. The profit function is taken to be $\tau(I, q) = r\mathrm{E}[\min\{XY, I+q\}] - h\mathrm{E}[(I+q-X)^+] - p\mathrm{E}[(XY-I-q)^+] - cq$. The discount factor is $\alpha = 0.9$. For demand sum and demand ratio parameters, we refer to subsections 7.2–3 of DeHoratius et al. (2008), which has a demand-first model. There the demand is a negative binomial random variable with parameters 2.27 and 0.47, which imply a mean demand of 2.013. The invisible demand is the difference of two independent Poisson random variables with means 0.577 and 0.580. Starting from these distributions and using the mapping between (demand, invisible demand) and (demand sum, demand ratio), we obtain the distribution of (X_t, Y_t). In the base case, $\mathrm{E}(X_t) = 2.1$ and $\mathrm{E}(Y_t) = 1.001$.

The inventory level is initially set to zero, that is, the prior of the inventory is $\delta_0(\cdot)$, and is later assumed to remain in the finite interval $[0, \mathcal{U}]$. When $\mathcal{U}$ is greater than 50, the total profit in our experiments is not sensitive to the value of $\mathcal{U}$, so we set $\mathcal{U} = 50$.

We evaluate the upper bound and the profits of the myopic policy and the policy obtained by the iterative algorithm. The iterative algorithm stops in most cases under $n = 20$ iterations with an accuracy parameter of $\epsilon = 0.001$. When running on an Intel Pentium Dual CPU (1.80 GHz), the duration of each iteration varies and is often less than 160 seconds. The profit resulting from using the myopic order quantities is computed as in Equation (16), except that max is not needed when the order quantities are given. When max is included, we obtain the profit of implementing the order quantities obtained from the iterative algorithm. The profit associated with the myopic policy, the profit resulting from the iterative algorithm, and the upper bound (Theorem 5) are all evaluated for the infinite horizon setting.

3.2.1. Comparison of Solution Procedures.

We generate test cases by altering the values of α, p, r, and $\mathrm{E}(Y_t)$ from their values in the base case; see Table 2. As the discount factor α increases in the left-hand side of the table, the gap between the iterative solution and the myopic policy grows. For realistic values of α, the gap is significant, and therefore we recommend the iterative algorithm. In the right-hand side of the table, we vary the unit shortage penalty p. We see that the iterative solution is better than the myopic policy, especially when p is high. A low p implies low inventory levels, which weakens the coupling between the periods. This, in turn, reduces the gap between the profits of the myopic and the iterative solutions. In the bottom portion of the table, we vary price r and the expected value $\mathrm{E}(Y_t)$ of the demand ratio. The profit obtained from the iterative algorithm grows faster than the myopic profit as either r or $\mathrm{E}(Y_t)$ increases. So the iterative algorithm outperforms the myopic policy for higher prices and demand ratios. Besides, the gap between the iterative solution and the upper bound becomes smaller as r or $\mathrm{E}(Y_t)$ increases. This means that the iterative algorithm yields a solution that is fairly close to being optimal when r or $\mathrm{E}(Y_t)$ is large. In all of our experiments, the iterative algorithm finds a better solution than the myopic solution.

Table 2 The Total Profit as α, p, r or $\mathrm{E}(Y_t)$ Varies

Varying parameter	Myopic policy	Iterative algorithm	Upper bound	Varying parameter	Myopic policy	Iterative algorithm	Upper bound
α				p			
0.70	1.319	1.517	1.665	0.50	6.239	6.301	6.312
0.75	1.652	1.859	2.002	0.55	5.803	5.913	6.121
0.80	2.132	2.260	2.595	0.60	5.322	5.676	5.862
0.85	2.618	3.015	3.451	0.65	4.419	5.028	5.673
0.90	3.370	4.617	5.191	0.70	3.892	4.897	5.415
0.95	7.988	9.314	10.384	0.75	3.370	4.617	5.192
r				$\mathrm{E}(Y_t)$			
0.96	2.751	3.722	4.347	0.88	2.697	2.833	4.304
0.98	3.024	4.189	4.543	0.90	2.765	2.979	4.471
1.00	3.370	4.617	5.191	0.92	2.827	3.139	4.681
1.02	3.716	4.920	5.614	0.94	2.913	3.432	4.581
1.04	4.197	5.431	6.037	0.96	2.969	3.751	4.721
1.06	4.651	6.079	6.459	0.98	3.019	4.029	4.964

3.3. Insights and Concluding Remarks

Our numerical study helps us to gain the following insights:

- The iterative solution is better than the myopic solution when the discount factor, the loss sales penalty or the price is relatively high. The iterative solution always beats the myopic policy and the difference between the two profits is 16% on average. Only when the discount factor, the shortage penalty and the price are all low, the myopic policy provides a reasonable alternative to the iterative algorithm.
- When the invisible demand is less than 10% of the demand, the iterative solution is not only better than the myopic solution, but also it is closer to the upper bound. In many real-life cases, the invisible demand can be less than 10%. In these cases, the IM can confidently use the iterative algorithm for maximizing the profit.
- Even when the invisible demand is more than 10%, the iterative solution comfortably beats the myopic solution but it moves apart from the upper bound. In these cases, the IM can use the iterative algorithm but with limited confidence. He can also take measures, such as RFID, employee training, or security guards to reduce the invisible demand or count inventory (frequently).

In summary, we study a periodic-review inventory system with observed sales, unobserved lost sales and invisible demand. Demand and invisible demand events are interspersed in a period and can depend on each other. We introduce the demand sum-demand ratio parameterization to describe the (inventory) probability evolution. The probability evolution can be used to attain inventory service levels. The probability evolution is highly nonlinear and leads to a non-trivial and a novel dynamic program. For the purpose of optimizing the profit, we apply the device of the unnormalized probability to linearize the evolution equation. This simplifies computations in the iterative algorithm devised to find order quantities. We compare the profits obtained with the iterative algorithm and myopic policy with each other and with an upper bound. This analytical upper bound is developed to get a sense of the proximity to optimality when the problem is solved with the iterative algorithm or the myopic policy. In view of our comparisons, solving the iterative algorithm turns out to be an effective approach in a majority of settings.

In this focused study, we have assumed that the probabilities associated with demand sum and demand ratio are known and that the inventory is not counted. Relaxing these assumptions would lead to the emergence of two related research topics: estimation of the demand distributions (with or without inventory counting (Mersereau 2015)) and optimization of the counting frequency. The models and methods presented here provide a foundation for these important future research pursuits.

Acknowledgments

We are heavily indebted to Prof. Kalyan Singhal and two anonymous referees for their very useful comments that have improved the presentation and exposition of this study. Alain Bensoussan's research is supported by the Research Grants Council of HKSAR (CityU 500111).

References

Atali, A., H. Lee, Ö. Özer. 2009. If the inventory manager knew: Value of visibility and RFID under imperfect inventory information. Working paper, Stanford University, Stanford, CA.

Beck, A., P. G. Chapman, C. Peacock. 2003. Shrinkage: A collaborative approach to reducing stock loss in the supply chain. ECR Europe. Available at http://ecr-all.org (accessed date July 15, 2014).

Bensoussan, A., M. Çakanyildirim, Q. Feng, S. P. Sethi. 2009. Optimal ordering policies for stochastic inventory problems with observed information delays. *Prod. Oper. Manag.* **18**(5): 546–559.

Bensoussan, A., M. Çakanyildirim, M. Li, S. P. Sethi. 2012. Existence and uniqueness of solutions for a partially observed stochastic control problem. S. N. Cohen, D. Madan, T. K. Siu, and H. Yang, eds. *Stochastic Processes, Finance and Control, A Festschrift in Honor of Robert J. Elliot*, World Scientific, Singapore, 393–413.

Bensoussan, A., M. Çakanyildirim, M. Li, S. P. Sethi. 2013. Optimal inventory control with shrinkage and observed sales. *Stochastics Int. J. Probab. Stochastic Process.* **85**(4): 589–603.

Bensoussan, A., M. Çakanyildirim, J. A. Minjárez-Sosa, S. P. Sethi, R. Shi. 2008. Partially observed inventory systems: The case of rain checks. *SIAM J. Cont. Optimiz.* **47**(5): 2490–2519.

Bensoussan, A., M. Cakanyildirim, S. P. Sethi. 2007a. Economic evaluation of systems that expedite inventory information. *Prod. Oper. Manag.* **16**(3): 360–368.

Bensoussan, A., M. Çakanyildirim, S. P. Sethi. 2007b. Optimal ordering policies for inventory problems with dynamic information delays. *Prod. Oper. Manag.* **16**(2): 241–256.

Bensoussan, A., M. Çakanyildirim, S. P. Sethi. 2007c. Partially observed inventory systems: The case of zero-balance walk. *SIAM J. Control Optim.* **46**(1): 176–209.

Bensoussan, A., M. Çakanyildirim, S. P. Sethi. 2010. Filtering for discrete-Time Markov processes and applications to inventory control with incomplete information. D. Crisan and B. Rozovsky, eds. *Handbook of Nonlinear Filtering*, Oxford University Press, United Kingdom, 500–525.

Bensoussan, A., M. Çakanyildirim, S. P. Sethi, R. Shi. 2011. Computation of approximate optimal policies in a partially observed inventory model with rain checks. *Automatica* **47**(8): 1589–1604.

Cachon, G. P., M. A. Lariviere. 1999. Capacity choice and allocation: Strategic behavior and supply chain performance. *Management Sci.* **45**(8): 1091–1108.

Camdereli, A. Z., J. M. Swaminathan. 2010. Misplaced inventory and radio-frequency identification (RFID) technology: Information and coordination. *Prod. Oper. Manag.* **19**(1): 1–18.

Chen, L. 2015. Fixing phantom stockouts: a POS-based shelf inspection model.Working paper, Cornell University, Ithaca, NY.

De Kok, A. G., K. H. Van Donselaar, T. Van Woensel. 2008. A break-even analysis of RFID technology for inventory sensitive to shrinkage. *Int. J. Prod. Econ.* **112**(2): 521–531.

DeHoratius, N., A. J. Mersereau, L. Schrage. 2008. Retail inventory management when records are inaccurate. *Manuf. Serv. Oper. Manag.* **10**(2): 257–277.

DeHoratius, N., A. Raman. 2008. Inventory record inaccuracy: An empirical analysis. *Management Sci.* **54**(4): 627–641.

Dutta, A., H. L. Lee, S. Whang. 2007. RFID and operations management: Technology, value, and incentives. *Prod. Oper. Manag.* **16**(5): 646–655.

Fleisch, E., C. Tellkamp. 2005. Inventory inaccuracy and supply chain performance: A simulation study of a retail supply chain. *Int. J. Prod. Econ.* **95**(3): 373–385.

Hardgrave, B. C., J. Aloysius, S. Goyal. 2009. Does RFID improve inventory accuracy? A preliminary analysis. *Int. J. RF Technol.* **1**(1): 44–56.

Heese, H.S. 2007. Inventory record inaccuracy, double marginalization, and RFID adoption. *Prod. Oper. Manag.* **16**(5): 542–553.

Hollinger, R.C. 2009. National retail security survey final report. University of Florida, Gainesville, FL.

Huh, W. T., M. Olvera-Cravioto, Ö. Özer. 2010. Joint audit and replenishment decisions for an inventory system with unrecorded demands. Working paper, Columbia University.

Kang, Y., S. Gershwin. 2005. Information inaccuracy in inventory systems: Stock loss and stockout. *IIE Trans.* **37**(9): 843–859.

Kök, A.G., K.H. Shang. 2007. Inspection and replenishment policies for systems with inventory record inaccuracy. *Manuf. Serv. Oper. Manag.* **9**(2): 185–205.

Kushner, H. J. 1967. Dynamical equations for optimal nonlinear filtering. *J. Differ. Eq.* **3**(2): 179–190.

Langer, N., C. Forman, S. Kekre, A. Scheller-Wolf. 2007. Assessing the impact of RFID on return center logistics. *Interfaces* **37**(6): 501–514.

Lee, H., Ö. Özer. 2007. Unlocking the value of RFID. *Prod. Oper. Manag.* **16**(1): 40–64.

Lee, H. L., V. Padmanabhan, S. Whang. 1997. Information distortion in a supply chain: The bullwhip effect. *Management Sci.* **43**(4): 546–558.

Lovejoy, W.S. 1991. Computationally feasible bounds for partially observed Markov decision processes. *Oper. Res.* **39**(1): 162–175.

Mersereau, A. J. 2013. Information-sensitive replenishment when inventory records are inaccurate. *Prod. Oper. Manag.* **22**(4): 792–810.

Mersereau, A. J. 2015. Demand estimation from censored observations with inventory record inaccuracy. *Manuf. Serv. Oper. Manag.* **17**(3): 335–349.

Miller, L., J. Allen. 2005. The national supermarket research group. Available at http://www.securestoreadvantage.com/pdfs/Supermarket_Survey_Executive_Summary.pdf (accessed date July 15, 2014).

Monahan, G. E. 1982. A survey of partially observable Markov decision processes: Theory, models, and algorithms. *Management Sci.* **28**(1): 1–16.

Nathan, C., N. DeHoratius, Y. Jiang, D. Klabjan. 2013. Inventory management with purchase order errors and rework. Working paper, The University of Chicago, Chicago, IL.

Porteus, E. L. 2002. *Foundations of Stochastic Inventory Theory*. Stanford Business Books, Stanford, CA.

Powell, W. B. 2007. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. John Wiley & Sons, New York.

Saghafian, S. 2015. Ambiguous partially observable markov decision processes: Structural results and applications, Working paper, Harvard University, Cambridge, MA.

Sethi, S. 2010. i3: Incomplete inventory information models. *Decisi. Line* **41**(5): 16–19.

Smallwood, R. D., E. J. Sondik. 1973. The optimal control of partially observable Markov processes over a finite horizon. *Oper. Res.* **21(5)**: 1071–1088.

Zakai, M. 1969. On the optimal filtering of diffusion processes. *Probab. Theory Rel. Fields* **11**(3): 230–243.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1: Proofs.