European Journal of Operational Research 225 (2013) 393–403

Contents lists available at SciVerse ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Production, Manufacturing and Logistics

Optimal inventory management for a retail chain with diverse store demands

Narendra Agrawal *, Stephen A. Smith

OMIS Department, Leavey School of Business, Santa Clara University, Santa Clara, CA 95053, United States

ARTICLE INFO

Article history:
Received 8 February 2012
Accepted 5 October 2012
Available online 23 October 2012

Keywords:
Inventory
Supply chain management
Retailing

ABSTRACT

Item demands at individual retail stores in a chain often differ significantly, due to local economic conditions, cultural and demographic differences and variations in store format. Accounting for these variations appropriately in inventory management can significantly improve retailers' profits. For example, it is shown that having greater differences across the mean store demands leads to a higher expected profit, for a given inventory and total mean demand. If more than one inventory shipment per season is possible, the analysis becomes dynamic by including updated demand forecasts for each store and re-optimizing store inventory policies in midseason. In this paper, we formulate a dynamic stochastic optimization model that determines the total order size and the optimal inventory allocation across nonidentical stores in each period. A generalized Bayesian inference model is used for demands that are partially correlated across the stores and time periods. We also derive a normal approximation for the excess inventory from the previous period, which allows the dynamic programming formulation to be easily solved. We analyze the tradeoffs between obtaining information and profitability, e.g., stocking more stores in period 1 provides more demand information for period 2, but does not necessarily lead to higher total profit. Numerical analyses compare the expected profits of alternative supply chain strategies, as well as the sensitivity to different distributions of demand across the stores. This leads to novel strategic insights that arise from adopting inventory policies that vary by store type.

1. Introduction

The vast majority of multi-location retail inventory models assume that all stores in the chain have identical probability distributions for demand. However, predictable differences often exist in store demands, which, when exploited appropriately, can increase profit significantly. For example, Macy's reported significant profit improvements attributed to varying its inventory policies across stores [Wall Street Journal, August 12, 2010]. Such store differences may arise from local economic conditions, cultural and demographics differences and varying store formats. The retail chain that sponsored this research classifies its stores by volume categories A, B, C, D, and E, where the classifications can vary by product type. The average sales varied by more than a factor of five across these store types. We spoke with two other retailers who use similar store classification systems.

Taking store differences into account in inventory management decisions introduces new strategic challenges for retailers. In addition to setting the various target stock levels, inventory must be allocated optimally across the stores when the total available inventory is constrained. Also, the inventory policy at each store type can be reoptimized in mid-season, based on updated demand information. This can lead to removing an item from certain stores if (1) item's sales are far above expectations, but the available inventory is not sufficient to provide appropriate stock levels at all stores or (2) the item is selling so poorly so that it does not justify the fixed cost of displaying it in some stores. Conversely, it may be advantageous to stock a few stores initially, and then stock more stores later, based on the initial observed sales results. Certain strategic tradeoffs can also be analyzed within this environment, e.g., "is it better to have a fixed supply contract with lower unit costs, or a higher cost supply contract that allows changes in merchandise orders in midseason?" The existing supply chain literature does not address these tradeoffs in the context of nonidentical stores.

We develop a two period inventory optimization model that uses a Bayesian inference model with random, correlated store demands that share a common unknown parameter. An optimization methodology determines the total order quantity, as well as the initial and revised store stocking policies for two periods, taking into account how the choice of stores to stock in the first period impacts the demand information that is collected. Managerial insights are obtained by analyzing various combinations of supply chain features such as *Replenishment policy* (single order or additional mid-season replenishment), *Store adjustment* (whether a different set of stores can be stocked mid-season), *Learning* (whether updated demand distribution based on period 1 sales is used for period 2 decisions), and *Sales distribution* (the variation in expected

* Corresponding author.
E-mail addresses: nagrawal@scu.edu (N. Agrawal), ssmith@scu.edu (S.A. Smith).

0377-2217/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejor.2012.10.006

sales across the store categories). Results are also calculated using different combinations of input parameters for salvage value, shortage cost, fixed cost of stocking a store, and the coefficient of variation of the demand distribution.

This formulation for store inventory allocation and ordering decisions includes several novel concepts – allocating inventory across nonidentical stores in a chain, generalized forecast updating when demands are correlated across the stores, combined with dynamic inventory optimization in the context of nonidentical stores. It is shown that having greater differences across the mean store demands leads to a higher expected profit, for a given inventory and total mean demand. Applying this methodology also generates managerial insights that are unique to retail chains with nonidentical stores. The prevalence of short season fashion goods and long production lead times from "offshore" manufacturers makes this research particularly relevant to apparel and department store retailers. Although the modeling approach extends to more than two periods, the size and complexity of the state space becomes prohibitive in that case. Thus, we confine our analysis to the two period problem, which is sufficient to provide insights regarding the impact of nonidentical stores on inventory strategy.

Previous inventory management models that include forecast updating tend to focus on single locations with multiple periods. For example, Murray and Silver [26] consider a multi-period problem for a single item, where the purchasing decision can be made at several discrete points in time. A stochastic dynamic programming formulation is derived, but it is shown to be computationally prohibitive, and consequently heuristics are developed. Azoury [5] developed a multi-period inventory decision model for a single item, single location problem with Bayesian updating for demand distributions of the exponential type. Assuming that unmet demand is backordered, this problem is solved by applying the scaling results of Scarf [30]. The multi-product case is considered in Hausman and Peterson [21] along with a finite production capacity, and a single, rather than period-by-period delivery. While their formulation allows for multiple forecast revisions, the heuristics they develop do not allow for this option. Bitran et al. [8] and Matsuo [25] formulate the problem as a two-level hierarchical structure characterized by families and items. While they specifically model production complexities such as production changeover costs, they assume that the total demand is fixed and uncertainty is confined to how demand is allocated across the various items. Models with forecast revisions are considered by Bradford and Sugrue [9] and Eppen and Iyer [16]. Both formulate Bayesian models to update demand forecasts based on observed sales in the context of fashion goods, but do not consider nonidentical stores.

Another related stream of research focuses on cases where unmet demand is unobservable in the single product case. Lariviere and Porteus [23] study the dependence of demand information on store inventory policy in a multi-period setting where only actual sales are observed. However, their model is tractable only for a restricted class of distributions and it does not address the decision to stock only certain stores. Ding et al. [14] consider general demand distributions in a two-period model, while Chen and Plambeck [11] and Lu et al. [24] extend the analysis to multiple periods for both perishable as well as non-perishable products. Kok and Shang [22], Bensoussan et al. [6] and DeHoratius et al. [13] model the added complexity that arises from inaccurate inventory information. When no assumptions are made about the form of the demand distribution, Chu et al. [10] derive operational statistics that maximize the performance (e.g., expected profit for the single period newsvendor model) uniformly for all values of the unknown demand parameters using Bayesian analysis. Finally, Petruzzi and Dada [28], Bisi and Dada [7] and the references therein add a pricing decision to the inventory decision problem. Capacity and production constraints are dealt with specifically by Fisher and Raman [19] for a multi-product case, where updated demand information is used for improved ordering and production decisions. Additionally, they also develop a procedure to estimate demand probability parameters for "new" items with no forecast history.

Multi-echelon inventory systems with identical retail stores have been analyzed in a number of papers. We will not review this literature in detail, but refer the reader to Federgruen [17], Axsater [4] and Agrawal and Smith [3]. Erkip et al. [15] consider a multi-echelon model with multiple retail outlets whose demands are correlated with each other and also across time, but do not include revisions. Fisher and Rajaram [20] consider an inventory model with different store types that is closer in spirit to our model. They analyze the optimal set of test stores to stock prior to the beginning of the selling season. Using sales history from a prior season, they specify stores clusters for the chain and then choose one test store from each cluster. This work differs from ours in a number of ways. First, they use linear regression to estimate forecasts for seasonal sales, while we use a Bayesian methodology to update the store demand forecasts based on period 1 sales. Second, their test stores are obtained deterministically by considering only the prior season sales. We determine the optimal set of stores to stock by considering the profit impacts for both periods 1 and 2.

Thus, while a number of papers in the existing literature use Bayesian updating to revise the demand distributions and also develop inventory policies that are optimized across multiple time periods, they consider either just one location, or assume that the multiple stores have identical demand distributions. In this paper, we focus on nonidentical stores, by jointly optimizing their inventory levels and the set of stores to be stocked in each period, taking into account how the store inventory policies in period 1 affect the demand information that is collected. Methodological contributions of our paper include the approximation of the sum of the leftover store inventories using the Central Limit Theorem which leads to manageable state space size, as well as the use of generalized Bayesian updating to obtain sufficient statistics for a set of nonidentical store samples. The resulting model quantifies the benefits of tailoring inventory allocation and store stocking policies to nonidentical stores.

2. The inventory model and the expected profit formulation

2.1. Model formulation

In this section, we formulate the two period dynamic optimization model. For most merchandise sold by retail specialty and department stores, excess demand at any store results in a lost sale. Because of POS (point of sale) scanners and electronic data interchange, store inventory allocation decisions are based on the most recent sales information. Deliveries to stores from regional distribution centers typically take place overnight, so that restocking is completed by the time the store opens the next day. These characteristics make it appropriate to use a "newsvendor" type of inventory model to compute the expected profit at the store level in each period.

Let us define the following notation for two time periods $t = 1, 2$. For parameters that are independent of the time period, the t subscript is omitted.

x	Store index satisfying $0 \leqslant x \leqslant X$
c	Purchase cost per unit
m	Gross margin per unit [m = revenue per unit $-c$]
w	Loss of goodwill per unit short [in addition to the loss of m]
r	Salvage revenue per unit at the end of the final time period

h	Unit cost of holding an item over from the first time period to the second at any store
H	Unit cost of holding an item over from the first time period to the second at a central storage location
$D_t(x)$	Random demand at store x in period t
$p_t(d,x)$	$P\{D_t(x) = d\}$
$P_t(d,x)$	$P\{D_t(x) \leqslant d\}$
$\mu_t(x)$	$E[D_t(x)]$
$\sigma_t^2(x)$	Variance $[D_t(x)]$
K_t	Fixed cost of stocking a store in period t

2.1.1. Decision variables

$s_t(x)$	Base stock level for time period t at store x
Y_t	Set of stores stocked in time period t
W_1	Inventory held at the central storage location during period 1

To simplify notation, we will use boldface type to indicate a vector of values indexed by x, i.e.,

$\mathbf{s_t} = \{s_t(x)\}$ and $\mathbf{D_t} = \{D_t(x)\}$ with particular demand values $\mathbf{d_t} = \{d_t(x)\}$, for $x \in Y_t$ and $t = 1, 2$.

Without loss of generality, the store index x can be defined so that the mean demands $\mu_t(x)$ are *decreasing* in x. In some instances, it may be convenient to treat x as a continuous variable, but our examples deal only with discrete x values.

Stocking a store involves both fixed and variable costs. The variable inventory costs per unit are included in h. The fixed cost K_t per store includes the cost of the space for displaying this particular item and any labor and handling costs that are independent of the inventory level. For example, items displayed on store shelves often have a fixed shelf facing, but a variable amount of inventory can be stored behind the facing. Items displayed in "cubes," such as shirts and casual slacks, also occupy a fixed amount of shelf space regardless of their inventory level. In-store holding cost h is generally much larger than H because space and shrinkage costs are much larger at retail stores than at the warehouse. Also, manufacturers will sometimes agree to hold back the items in W_1 and deliver them later, in which case H would be 0.

Retail price is not a decision variable in our model. This restricts our model to retail chains that do not reduce prices until the end of the season, such as the retailer we studied, or to retailers who follow pre-specified pricing policies during the season. Retail chains that use periodic price promotions to increase store traffic tend to schedule these price changes in advance of the season, coordinating them across products and with planned advertising, as opposed to using price as an inventory management tool. Price reductions that occur during the clearance period are incorporated into the salvage value r.

2.1.2. The expected profit function

The expected profit in each period t is based on a newsvendor model, with a fixed cost K_t if the beginning inventory in period t is greater than zero. The newsvendor expected profit function for any stock level s can be written as

$$\pi_t(s,x) = m\mu_t(x) - c_u E[D_t(x) - s]^+ - c_{ot}E[s - D_t(x)]^+ - K_t I\{s\}, \quad (1)$$

where c_u = the unit understock cost, c_{ot} = the unit overstock cost in period t and $I\{s\} = 1$ if $s > 0$ and 0 otherwise. Integration (or summation) by parts allows us to write the profit function as

$$\pi_t(s,x) = -w\mu_t(x) + c_u s - (c_u + c_{ot})\int_{\{d\leqslant s\}} P_t(d,x) - K_t I\{s\}. \quad (2)$$

Here the notation $\int_{\{d\leqslant s\}}$ represents either an integral or a sum up to s, depending on whether the probability distribution is continuous or discrete. In terms of the symbols defined previously, we have $c_u = m + w$, $c_{o1} = h$ and $c_{o2} = c - r$. That is, excess inventory in period 1 receives a credit of $c - h$ at the end of period 1, and is made available for period 2 at a cost c, while excess inventory in period 2 is salvaged. The special case $s = 0$ gives an expected profit of $\pi_t(0,x) = -w\mu_t(x)$ for period t, corresponding to the expected loss of goodwill.

The accounting above assumes for each x that either $s_2(x) \geqslant s_1(x) - D_1(x)$, or if not, then the inventory $s_1(x) - D_1(x) - s_2(x)$ is allocated across the remaining stores for period 2. This would occur, for example, if store x is stocked in period 1, but not in period 2. In this case, based on our discussions with retailers, the leftover inventory is typically returned to the warehouse and redistributed to other stores in period 2.

For a given subset of stores Y_t to stock, the total expected profit $\Pi_t(\mathbf{s_t}, Y_t)$ in period t can be written as

$$\Pi_t(\mathbf{s_t}, Y_t) = \sum_{x\in Y_t}\left\{c_u s_t(x) - (c_u + c_{ot})\int_{\{d\leqslant s_t(x)\}} P_t(d,x) - K_t\right\} - G_t, \quad (3)$$

where $G_t = w\sum_{\text{all}x}\mu_t(x)$ is independent of the store stocking decisions $s_t(x)$ and Y_t.

Lemmas 1 and 2 state several results for this expected profit function for nonidentical stores, which are important for our analysis and discussions of managerial insights. Some additional details of the proofs of these results are contained in Appendix A.

Lemma 1. [Optimal inventory policies]

(a) *The optimal stock levels in (3) are such that all stores $x \in Y_t$ have the same service level. Thus, the optimal stock levels are of the form $s_t(\alpha_t, x)$, where $P_t(s_t(\alpha_t,x),x) \equiv \alpha_t$ for all $x \in Y_t$ and some service level α_t. If inventory is unconstrained, then $\alpha_t = c_u/(c_u + c_{ot})$. When the total inventory is constrained, the service level α_t is adjusted as needed to satisfy the inventory constraint.*

(b) *The optimal set of stores to stock in (3) is of the form $Y_t = \{x | 0 \leqslant x \leqslant y\}$ for some y, if the probability distributions satisfy the following property*

$$\int_{d\leqslant s}[P_t(d,x_2) - P_t(d,x_1)] > 0 \quad \text{for } t = 1,2, \text{ and } x_2 > x_1. \quad (4)$$

Proof. Part (a) follows from the standard newsvendor analysis, since the individual store profits in (3) are additively separable. Part (b) is proved in Appendix A. It can be shown that the optimal form $Y_t = \{x | 0 \leqslant x \leqslant y\}$ also holds for the normal distribution for all $s > 0$, if the ratio $\sigma(x)/\mu(x)$ is nondecreasing as the mean store demand decreases.

We assume that (4) holds for the demand distributions throughout the paper. Intuitively, (4) means that for a given initial stock level $s > 0$, the expected excess inventory $E[s - D_t(x)]^+ = \int_{d\leqslant s}(s - d)p_t(d,x) = \int_{d\leqslant s} P_t(d,x)$ is strictly increasing in the index x. Condition (4) is sometimes called "second order stochastic dominance," and is satisfied by the Poisson, negative binomial and gamma distributions, which together with the normal distribution include most of the demand distributions commonly used in inventory management. The proofs that (4) holds for these common demand distributions can be obtained from the authors.

Next, we define *steepness*, a property of demand distributions across non-identical stores, in two different ways. Lemma 2 shows that these two definitions are equivalent, and that steeper distributions always lead to higher expected profits, given the same mean demand.

Let $\{P_t(d,x)\}$ and $\{P_t^*(d,x)\}$ be two families of continuous distributions indexed by x, both with the same total mean demand. Then $\{P_t^*(d,x)\}$ is steeper than $\{P_t(d,x)\}$ if:

(i) The corresponding stock levels $s_t(\alpha,x)$ and $s_t^*(\alpha,x)$ needed to achieve a given service level α satisfy $\sum_{x\leqslant y} s_t^*(\alpha,x) \geqslant \sum_{x\leqslant y} s_t(\alpha,x)$ for any y, or

(ii) If the same inventory I is allocated optimally to a set of stores of the form $Y_t = \{x|0 \leqslant x \leqslant y\}$, so that $\sum_{x\leqslant y} s_t^*(\alpha^*,x) = \sum_{x\leqslant y} s_t(\alpha,x) = I$, then the resulting service levels satisfy $\alpha^* \leqslant \alpha$. □

Lemma 2. [Steepness properties of nonidentical demand distributions]

(a) *The two definitions of steepness (i) and (ii) are equivalent.*

(b) *For any total inventory I that is allocated optimally across a set of stores $Y_t = \{x|0 \leqslant x \leqslant y\}$, the expected profits corresponding to $\{P_t(d,x)\}$ and $\{P_t^*(d,x)\}$, respectively, satisfy $\Pi_t(\mathbf{s_t}, Y_t(y)) \leqslant \Pi_t^*(\mathbf{s_t^*}, Y_t(y))$. This is true even when I is the optimal inventory for $\{P_t(d,x)\}$ and Y_t, but is suboptimal for $\{P_t^*(d,x)\}$ and Y_t.*

Proof. Parts (a and b) of Lemma 2 are proved in Appendix A.

These properties imply, when the same set of stores are stocked with the same total inventory, a steeper family of demand distributions always has an expected profit that is at least as large as the expected profit for a less steep family of distributions. This is true even when the resulting service level for the steeper demand distributions is suboptimal. If the set of stores to stock is optimized separately in each case, the steeper distribution would have an even larger expected profit. This result is relevant for retail chains that operate across markets with significant demographic differences. When the resulting demand distributions are more diverse, the total expected profit is higher than when the same expected demand is uniformly distributed across the stores. □

2.2. The inventory optimization problem

The general inventory optimization problem for nonidentical stores is a dynamic programming problem because the inventory and store stocking decisions in period 1 affect both the available inventory for period 2 and the demand information that is obtained from period 1. To simplify this formulation, we make the following informational assumption.

Assumption 1. For period 1, lost sales can be estimated at all stores $x \in Y_1$, so that the total demand at these stores is known, but no demand information is obtained for the stores $x \notin Y_1$.

This assumption is justified because the optimal service level α_1 in period 1 tends to be high for most retail products, e.g., 90–95%, so that lost sales are small relative to actual sales. There are various methods for estimating lost sales that work well for high service levels. For example, Agrawal and Smith [2] developed a lost sales estimation methodology and found that errors were very small when service levels are high.

2.2.1. Single ordering decision

In this case, all inventory must be ordered before period 1. The retailer first chooses $\mathbf{s_1}, Y_1$, and W_1, while the subsequent allocation decisions $\mathbf{s_2}, Y_2$ will be conditional on the outcomes of the first period. Thus, we have the following two period inventory optimization problem

$$\max_{W_1,\mathbf{s_1},Y_1} \left\{ \Pi_1(\mathbf{s_1}, Y_1) + \int_{\mathbf{d_1}} \pi_2(\mathbf{d_1}, \mathbf{s_1}, Y_1, W_1) P\{\mathbf{D_1} = \mathbf{d_1}\} - HW_1 \right\}, \quad (5)$$

where $\pi_2(\mathbf{d_1}, \mathbf{s_1}, Y_1, W_1) = \max_{\mathbf{s_2}, Y_2} \Pi_2(\mathbf{s_2}, Y_2|\mathbf{d_1}, Y_1)$,

$$\text{subject to } \sum_{x\in Y_2} s_2(x) = I_2 = \sum_{x\in Y_1} [s_1(x) - d_1(x)]^+ + W_1. \quad (6)$$

The conditional expected profit $\Pi_2(\mathbf{s_2}, Y_2|\mathbf{d_1}, Y_1)$ is obtained by replacing $P_2(d,x)$ with $P_2(d,x|\mathbf{d_1}) = P\{D_2(x) \leqslant d|\mathbf{D_1} = \mathbf{d_1}\}$ in (3).

2.3. The form of the optimal policies

Let us first consider the form of the optimal policies in period 2. Properties I and II discussed earlier imply that all stores stocked in period 2 should have the same service level α_2, and that Y_2 is of the form $\{x|x \leqslant y_2\}$ for some value y_2. The optimal solutions for period 2 will depend on the results from period 1, i.e., how much inventory is available in (6) and the demand information that was observed in period 1.

For period 1, the service level α_1 can be optimized independently, since the total order quantity is a decision variable at this time. The decision of whether to stock x depends on the profit $\pi_1(s_1(x),x)$ in (2) plus the following variable: $U(x)$ = the incremental expected value contributed to period 2 as a result of stocking store x in period 1, which has the following property.

If the function $U(x)$ is nonincreasing in x, then the optimal set of stores to stock in period 1 is of the form $Y_1 = \{x|x \leqslant y_1\}$ for some value $y_1 \leqslant X$. That is, if we add $U(x)$ to $\pi_1(s_1(x),x)$ in (2) when defining the Period 1 profit in (3), then the policy in Lemma 1b still holds for period 1. This property is needed only for computational efficiency. If $U(x)$ is increasing in x for some x values then Y_1 can be determined by an exhaustive search over all sets of stores that could be stocked in period 1. This exhaustive search is not that difficult computationally when store stocking policies are applied to a few discrete categories of stores, as was the case in our example applications.

The optimal Y_1 and W_1 for period 1 must be determined by searching for the values that maximize (5) and (6) with an optimization for period 2 being performed for each (Y_1,W_1) and each demand outcome $\mathbf{d_1}$. For period 2, the optimal solution in each case is specified by the two variables (α_2,y_2). The period 1 optimization problem has high dimensionality, because (α_2,y_2) will be different for every $(\alpha_1,Y_1,W_1,\mathbf{d_1})$ combination from the first period.

2.3.1. The flexible replenishment case

The optimization problem in (5) and (6) is simpler when any desired amount of replenishment inventory can be purchased for period 2. In this case, $W_1 = 0$ is optimal and $\alpha_2 = c_u/(c_u + c_{o2})$ holds. This solution still requires dynamic optimization, but the calculations are much simpler.

2.4. Bayesian updating of the demand distribution

Bayesian updating greatly reduces the dimensionality of this problem by conditioning the store demand distributions on a common parameter θ. We define $p(\theta)$ as the prior probability and the likelihood functions

$$L_t(d|\theta,x) = P\{\text{demand at store } x \text{ in period } t \text{ is } d|\theta\}, \text{ for each } t,x.$$

The standard Bayesian assumption is that the store demands are independent observations when conditioned on θ. For example, θ could represent the overall market strength of the item. If the prior distribution and likelihood functions are conjugate distributions of the exponential type, the posterior demand distributions are determined from Bayes Formula. Exponential type distributions include the common demand distributions used in inventory management such as the Normal, Poisson, Gamma, Weibull and Negative Binomial.

Store demands that are not identically distributed require generalized formulas for Bayesian updating. (See, e.g., DeGroot [12] for the identical sample case.) Generalized Bayesian updating for non-identical stores can be specified as follows.

Assumption 2. For $t = 1, 2$, there exist functions $g(\theta)$, $h_t(d|x)$, $n_t(x)$ and $u_t(d|x)$ such that

$$p(\theta) = g(\theta)^{\beta} e^{\theta v} \text{ and } L_t(d|\theta, x) = h_t(d|x) g(\theta)^{n_t(x)} e^{\theta u_t(d|x)}. \tag{7}$$

If demands $\mathbf{d_1} = \{d_1(x)\}$ for $x \in Y_1$ have been observed, then

$$P\{\mathbf{D_1} = \mathbf{d_1}|\theta, Y_1\} = \left(\prod_{x\in Y_1} h_1(d_1(x)|x)\right) g(\theta)^{n_1(Y_1)} e^{\theta\tau(\mathbf{d_1})}, \tag{8}$$

$$\text{where } \tau(\mathbf{d_1}) = \sum_{x\in Y_1} u_1(d_1(x)|x) \text{ and } n_1(Y_1) = \sum_{x\in Y_1} n_1(x). \tag{9}$$

The posterior distribution for θ satisfies the proportionality relationship

$$p(\theta|\tau(\mathbf{d_1})) \propto g(\theta)^{n_1(Y_1)+\beta}\ e^{\theta\{\tau(\mathbf{d_1})+v\}}. \tag{10}$$

This implies that the demand distributions defined previously can be written as

$$P_1(d, x) = \int_{\theta} L_1(d|\theta, x) p(\theta) d\theta,$$

$$P_2(d, x|\tau(\mathbf{d_1})) = \int_{\theta} L_2(d|\theta, x) p(\theta|\tau(\mathbf{d_1})) d\theta.$$

In the two period problem, the sufficient statistic $\tau(\mathbf{d_1})$ reduces the dimensionality of the optimization by replacing $\Pi_2(\mathbf{s_2}, Y_2|\mathbf{d_1}, Y_1)$ with $\Pi_2(\mathbf{s_2}, Y_2|\tau(\mathbf{d_1}), Y_1)$.

Appendix A summarizes the updating formulas for the gamma–exponential family of distributions, which are used for the illustrative calculations in this paper. The Bayesian updating formulas have also been developed for the normal and negative binomial distribution families in a Supplementary appendix, which can be obtained from the authors. For the three distributions that we have analyzed, the expected demand at each store conditional on θ can be written in the multiplicatively separable form

$$E[D_t(x)|\theta] = f_t(x)\chi(\theta), \tag{11}$$

where $f_t(x)$ is the expected fraction of period 1 demand occurring at store x and $\chi(\theta)$ is a monotone function of θ. Then (11) implies that

$$E[D_t(x)] = \int_{\theta} E[D_t(x)|\theta] p(\theta) = f_t(x) E[\chi(\theta)].$$

It follows in this case that the expected store demands $\{E[D_t(x)]\}$ uniquely determine the parameters $\{f_t(x)\}$ and $E[\chi(\theta)]$ since $\sum_x f_t(x) = 1$. The previous store demand property $\mu_t'(x) < 0$ is equivalent to $f_t'(x) < 0$ in this case.

2.5. Approximate distribution for the excess inventory

We also reduce the dimensionality of the optimization in (5) and (6) by developing a Normal approximation for the random variable

$$R(\mathbf{D_1}|\mathbf{s_1}) = \sum_{x\in Y_1} [s_1(x) - D_1(x)]^+, \tag{12}$$

which corresponds to the total excess inventory from period 1. It can be shown that a normal approximation can be used to convert the objective function (5) to the following form

$$\max_{W_1, \mathbf{s_1}, Y_1} \left\{ \Pi_1(\mathbf{s_1}, Y_1) + \int_{\tau}\int_{R} \pi_2(Y_1, W_1, R, \tau)\psi(R|\tau, \mathbf{s_1})\ell(\tau) - HW_1 \right\}, \tag{13}$$

$$\text{where } \ell(\tau) = P\{\tau(\mathbf{D_1}) = \tau\} = \int_{\theta} P\{\tau(\mathbf{D_1}) = \tau|\theta\} p(\theta),$$

$$\text{and subject to } P_2(s_2(x), x|\tau, Y_1) = \alpha_2 \text{ and } \sum_{x\leqslant y_2} s_2(x) = R + W_1. \tag{14}$$

The complete details are in the Supplementary appendix. In summary, the normal approximation is derived by conditioning on θ, so that

$$R(\mathbf{D_1}|\mathbf{s_1}, \theta) = \sum_{x\in Y_1} [s_1(x) - D_1(x|\theta)]^+, \tag{15}$$

is a sum of independent random variables, and thus $P\{R(\mathbf{D_1}|\mathbf{s_1}, \theta) = R\}$ can be approximated by a normal random variable $\psi(R|\theta)$ from the Central Limit Theorem. Thus $\psi(R|\tau, \mathbf{s_1})$ is a univariate normal density with mean and variance given respectively by

$$m(\mathbf{s_1}|\tau) = \sum_{x\in Y_1} \int_{\theta} E[[s_1(x) - D_1(x)]^+|\theta]\ p(\theta|\tau), \tag{16}$$

$$v^2(\mathbf{s_1}|\tau) = \sum_{x\in Y_1} \int_{\theta} E[([s_1(x) - D_1(x)]^+)^2|\theta] p(\theta|\tau) - [m(\mathbf{s_1}|\tau)]^2, \tag{17}$$

so that the approximating normal distribution matches the mean and variance of the sum of the actual random variables.

3. Numerical analysis and insights

We have applied the models to the exponential–gamma conjugate family of demand functions to perform the calculations in this section. The Bayesian updating, optimal stock levels, and expected profit formulas are summarized in Appendix A. Derivations have also been done for the normal and negative binomial distributions in Supplementary appendix, which is available from the authors. Wherever possible, we have used numerical inputs obtained from discussions with managers at a major retail chain.

3.1. Assumptions used in the illustrative calculations

The total demand for the product category across all stores and time periods is assumed to have a prior distribution with a mean of 2000 and a standard deviation of 2450 (CV = 1.224). For convenience in specifying inventory policies for different types of stores, the retailer divides all stores into five volume categories (1–5) and manages all stores within a given category identically. For our illustrative calculations, we use 20 stores per category and vary the mean demand, as shown in Table 1. For simplicity, we assume that the two time periods have equal expected demands for the prior demand distribution.

3.2. Financial data

For computational convenience, we set the procurement cost per unit at c = \$10, and the gross margin m = \$7.00, using typical ratios obtained through our discussions with retail managers. Unsold inventory is salvaged at the end of the season at an average value of r = \$6/unit. The in-store inventory carrying cost is set at h = \$0.50/unit, which includes shrinkage and handling costs. The unit holding cost H at the retailer's warehouse is assumed to be

Table 1
Store categories in the retail chain.

Store category	# Stores	% Sales in category	Prior mean demand per store per period
1	20	42.0	21.0
2	20	23.1	11.6
3	20	16.2	8.1
4	20	11.7	5.9
5	20	7.0	3.5

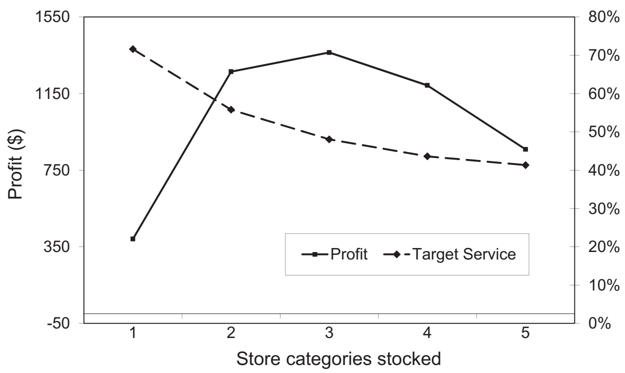

Fig. 1. Effect of inventory availability on profit and service level in period 2 (beginning inventory = 1000).

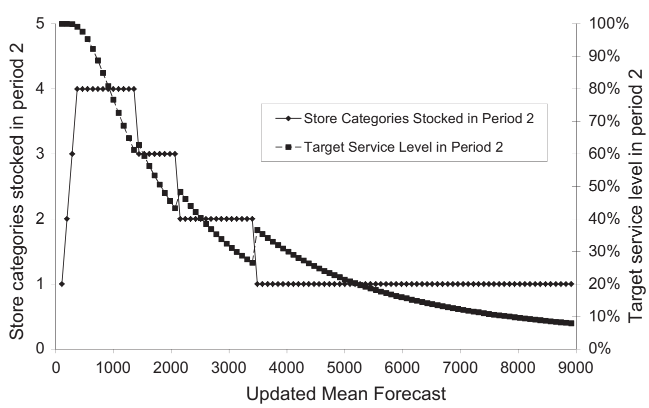

Fig. 2. Effect of period 1 demand on store stocking policies in period 2 (single order case).

negligibly small in this example calculation. Including loss of goodwill, the unit shortage cost is c_u = \$8. Thus, c_{o1} = \$0.50, c_{o2} = \$5.00, which lead to unconstrained newsvendor service levels of $\alpha_1 = 0.94$ and $\alpha_2 = 0.67$, respectively. Finally, the fixed cost K_t of stocking a given SKU at any store is assumed to be \$25 in both periods.

3.3. Illustration of the period 2 tradeoffs when the available inventory is fixed

For this example, suppose that there are 1000 units of inventory available for period 2, but the updated expected demand for period 2 is now 2000 units. This could occur, for example, if period 1 sales were very high. Suppose also that no new orders are permitted. Solving the period 2 problem requires a joint optimization of the set of stores to stock and the common service level for these stores, using the methods discussed in Section 3. Fig. 1 illustrates the tradeoffs that are available, where the dashed line indicates the service level, which decreases monotonically with the number of stores that are stocked. The expected profit, indicated by the solid line, is maximized by stocking the first three categories of stores, which results in a service level of about 46%. This jointly optimized service level is less than the ideal service level of α_2 = 67%. Stocking only the top three categories of stores results in an expected profit that is roughly 70% higher than the profit obtained by stocking all five store categories. This figure illustrates just one possible outcome from period 1. A much lower expected demand in period 2 could result in excess inventory and an optimal service level higher than the ideal of 67%.

Fig. 1 was based on a single outcome for the period 1 demand, as expressed in terms of the sufficient statistic $\tau(\mathbf{d_1})$ in (9). Fig. 2 illustrates the period 2 solutions for all possible values of the updated period 2 expected demand, obtained by varying $\tau(\mathbf{d_1})$. For this chart, we used the optimal period 1 value $y_1 = 2$, which will be obtained in the next subsection. Fig. 2 shows that if the updated mean is too small or to large, it is optimal to stock fewer stores, while for means between 377 and 1267, it is optimal to stock 4 store types. The explanation is that when the mean demand is low, there is not enough operating profit to offset the fixed cost of \$25.00 per store, while when mean demand is very large, it becomes advantageous to stock fewer stores, because there is not enough inventory to allocate. The optimal service level in Fig. 2 also experiences step changes, because when the number of stores steps to a new level, it creates a discrete change in the inventory allocation.

Fig. 3 illustrates the corresponding results when additional inventory can be ordered for period 2. The results in Figs. 2 and 3 match when the posterior mean is low, because there is not sufficient demand to cover the fixed cost. However, as the posterior mean increases, it becomes optimal to stock all stores to a service level of α_2 = 67%, the unconstrained optimum.

3.4. The optimal solution for period 1

Fig. 4 illustrates the optimal solution for period 1 for the cases that correspond to Figs. 2 and 3. In period 1, we need to determine the optimal set of stores to stock, the inventory levels at each store and W_1, the amount of inventory held back at the DC or vendor. The optimal number of stores to stock in period 1 must take into account the value of the information that is obtained and how it affects the expected profit in period 2. If more store categories are stocked in period 1, more information about demand is obtained, which reduces the standard deviation of the demand in period 2. However, this benefit must be compared to the losses that may result from stocking the low volume stores in period 1.

Fig. 4 shows it is optimal to stock the top 2 store categories in period 1. The maximum expected profit is \$1320, and the total order size in 1964 units. The optimal holdback inventory is W_1 = 0 in this example. This is because the leftover store inventory from period 1 provides sufficient safety stock for period 2 which has a service level of 67%, much less than the period 1 service level of 94%. If the target service level in period 2 were higher, the optimal W_1 could have been larger than 0.

3.5. Comparisons of alternative supply chain strategies

The solutions for period 1, illustrated in Fig. 4, assume that Bayesian updating or Learning, takes place before period 2, and that the set of stores to stock is reoptimized for period 2. Tables 2 and 3 compare the expected profits that are obtained, when one or both of the options is unavailable. When there is No Learning, the period 2 decisions are made using the prior demand distribution, but the expected profits in the tables were calculated using the updated demand distribution. When there is "No Store Adjust-

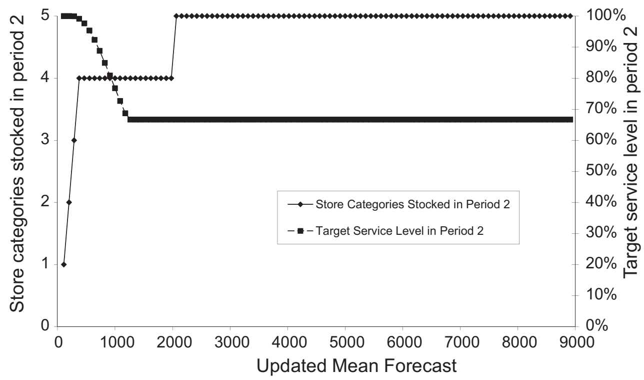

Fig. 3. Effect of period 1 demand on store stocking policies in period 2 (unlimited in-season replenishment).

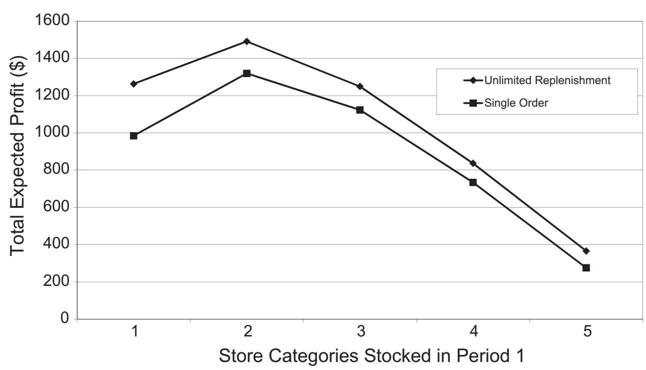

Fig. 4. Total profits as a function of period 1 stocking policy (with learning and Store Adjustment).

Table 2
Profit comparisons in the single order case.

	Store Adjustment	No Store Adjustment
Learning	1320	888
No Learning	1316	888

ment," the same stores are stocked in period 1 and period 2, regardless of the period 1 demand.

Table 2 shows that Store Adjustment provides significant incremental profit improvements-about 49% with Learning and 48% with No Learning. Even with No Learning, Store Adjustment is valuable because it takes into account the leftover inventory from period 1 and jointly reoptimizes the set of stores to stock and the service level, using the prior demand distribution. But in Table 2, Learning alone has very little value in the single order case. Learning has no value at all with No Store Adjustment because the stock level $s_t(x)$ in (20) is proportional to $f_t(x)$, which implies that the same fixed amount of inventory is allocated to each store regardless. With Store Adjustment, Learning improves profit somewhat, but with no possibility to reorder, there is only limited ability to act on the updated demand information.

Table 3 does the analogous comparisons when period 2 replenishment is possible. Here Store Adjustment contributes a 53% increase in profit when there is Learning, and a 48% increase in profit when there is No Learning. However, Learning contributes a 10% increase in profit with No Store Adjustment and a 13% increase in profit with Store Adjustment. This illustrates how the incremental impact of Learning depends crucially on whether or not a second order can be placed. Store Adjustment for period 2 is always beneficial, and more so when there is also Learning.

Table 3
Profit comparisons with replenishments in period 2.

	Store Adjustment	No Store Adjustment
Learning	1491	975
No Learning	1321	890

Comparing the corresponding cells in Tables 2 and 3, we see that period 2 replenishment achieved the greatest numerical increase in profit when there is both Learning and Store Adjustment. Also, when there is No Learning, the incremental benefit of having a second order is minimal.

It is important to note that the expected profits in Tables 2 and 3 also depend on the differences in the store demands, which allow only the more profitable stores to be stocked. When the expected profits for three different distributions of demand across store types are compared later in Table 4, we find that greater differences in store demands lead to significantly higher expected profits, given the same total expected demand.

3.6. Sensitivity analysis of the results

In this section, we analyze the sensitivity of these calculations to a variety of input changes. We also investigate the sensitivity of the results to the distribution of the demand across stores. In Section 2, it was shown that steeper distributions of demand across the stores always leads to higher expected profits for many types of demand probability distributions. In this section, we also determine the impact of steeper or flatter distributions of demand on the benefits associated with Learning, Store Adjustment and period 2 replenishment.

The sales rate curve used in the previous section was based on data obtained from a retail chain, which is referred to as the *Actual* case. The two additional cases are *Flat*, where all stores have the same sales rate, and *Steep*, where the sales curve is much steeper than the Actual case. The number of stores in each category is still fixed at 20, but the percentages of sales by category vary. The three cases are illustrated in Fig. 5. We also analyzed all combinations of the following parameter values: Fixed Cost K_t = \$7 and \$25, Shortage Cost c_u = \$8.00 and \$7.00, Salvage Value r = \$6.00 and \$9.00, and, Demand Variability CV = 1.225 and 1.414.

Tables 4a and 4b display a subset of these calculations, for the policies of Learning and No Learning (L and NL), with and without Store Adjustment (SA and NSA) and for the cases of a single order and period 2 replenishment. The Tables 2 and 3 results appear in the top row of Table 4a. The other entries in Table 4a give the corresponding results for input parameter variations. Table 4b compares the profits for the Flat, Actual and Steep distributions of demand across stores using the base case parameter values. These calculations were also performed for the other combinations of parameter values, and although the magnitudes of the profits changed significantly, the qualitative comparisons and insights that we discuss later in this section remained the same. In both these tables, the number in parentheses below the profit values is the optimal number of store categories to stock in period 1.

For all cases, Learning provides only minimal incremental value with a single order, but provides considerable benefit when a second order is permitted, even without Store Adjustment. However, Store Adjustment provides significant benefit in most of the cases, and the benefit increases when a second order is permitted. The Store Adjustment benefit also increases significantly with greater demand variability (CV = 1.414). An exception to this finding occurs for the case of Salvage Value = \$9.00 in Table 4a, where Store

Table 4a
Sensitivity analysis of profit (using actual sales distribution).

	1 ORDER SA	1 ORDER NSA	2 ORDERS SA	2 ORDERS NSA
BASE CASE				
L	1320 (2)	888 (3)	1491 (2)	975 (3)
NL	1316 (2)	888 (3)	1321 (2)	890 (3)
SHORTAGE COST = $7				
L	2266 (2)	1877 (2)	2388 (2)	1933 (2)
NL	2261 (2)	1877 (2)	2263 (2)	1879 (2)
CV = 1.414				
L	604 (2)	25 (2)	979 (1)	204 (3)
NL	597 (2)	25 (2)	602 (2)	30 (2)
FIXED COST = $7				
L	3912 (3)	3829 (5)	4054 (3)	3940 (5)
NL	3912 (3)	3829 (5)	3916 (4)	3831 (5)
SALVAGE VALUE = $9				
L	4450 (4)	4442 (4)	4962 (4)	4919 (4)
NL	4446 (4)	4442 (4)	4673 (4)	4673 (4)

SA: Store Adjustment. NSA: No Store Adjustment. L: Learning. NL: No Learning. [Numbers in parantheses are the number of stores to stock in period 1.].

Table 4b
Effect of store demand distribution (base case values).

	1 ORDER SA	1 ORDER NSA	2 ORDERS SA	2 ORDERS NSA
FLAT SALES DISTRIBUTION				
L	483 (4)	229 (5)	712 (3)	340 (5)
NL	480 (4)	229 (5)	489 (4)	231 (5)
ACTUAL SALES DISTRIBUTION				
L	1320 (2)	888 (3)	1491 (2)	975 (3)
NL	1316 (2)	888 (3)	1321 (2)	890 (3)
STEEP SALES DISTRIBUTION				
L	2083 (2)	1757 (3)	2219 (2)	1857 (3)
NL	2081 (2)	1757 (3)	2085 (2)	1759 (3)

SA: Store Adjustment. NSA: No Store Adjustment. L: Learning. NL: No Learning. [Numbers in parantheses are the number of stores to stock in period 1.].

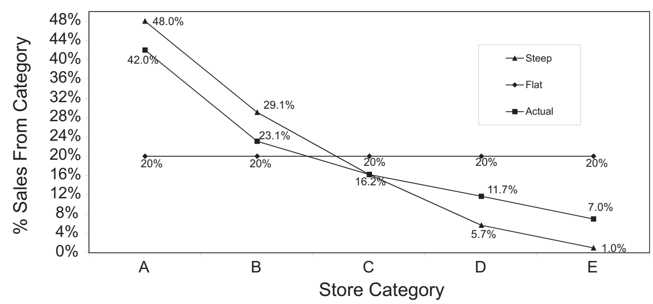

Fig. 5. Sales rate curves.

Adjustment is not very beneficial. The explanation is that with a high salvage value, stocking most of the stores, 4 out of 5 in this case, is an attractive option in both time periods, and thus Store Adjustment is not a very valuable option. But Table 4a contains many other combinations of parameter values for which the Store Adjustment option provides significant benefits to the retail chain.

3.6.1. Effect of sales distribution across stores

Major differences arise in Table 4b as the distribution of demand across stores changes. When the distribution of demand is steeper, i.e., concentrated in fewer stores, expected profits increase because stocking a few stores captures most of the potential sales, and these same stores also provide most of the information about demand for period 2. For example, for the one order case with Learning and Store Adjustment, it is optimal to stock the top four store categories in period 1 in the Flat sales distribution case, but only the top two categories are stocked in the Normal and Steep cases. Table 4b also shows that on percentage basis, the incremental benefit of Store Adjustment tends to be greater with flatter distributions of sales.

With Learning, a second order for period 2 provides significant benefit in every case in Tables 4a and 4b. On the other hand, with No Learning, the second order for period 2 provides little incremental benefit for almost all cases. The exception is Salvage Value = $9.00 in Table 4a. Since store inventories are high in both periods in this case, there is significant benefit in being able to replenish for period 2, even with No Learning. Tables 4a and 4b illustrate that in many instances Store Adjustment can serve as an attractive substitute for period 2 replenishment when it is impossible or expensive to obtain. And when period 2 replenishment is possible, it should be complemented with Learning.

The numerical analysis found that fewer store are stocked in period 1 with the steeper distributions of demand across stores, in both the one- and two-order cases. We also note that the number of store categories stocked in period 1 in the 2 Orders case is always less than or equal to the number stocked in the 1 Order case, i.e., having the second order makes carrying the additional inventory at these stores in period 1 less beneficial. Thus, in general, more stores are stocked in period 1 with the Flat sales distribution, and in the 1 Order case. Additional inventory may also be purchased and held back in the 1 Order case when the expected profit per store is higher in period 2. In our illustrative calculations, this happens when salvage value is high.

4. Discussion of results and conclusions

The models developed in this paper provided a number of new analytical results for retail inventory management with nonidentical stores. First, sufficient conditions were developed that imply an optimal rank order for the stores that should stock a given item. Under these conditions, it is always preferable to stock the stores with higher expected demands. Second, it was shown that if the total chain-wide demand is equal, a steeper distribution of demand across stores, i.e., more store variety, leads to greater profits. Bayesian updating was formulated for the case of nonidentical stores for probability distributions of the exponential type, which includes the commonly used demand distributions in inventory management. Finally, a normal approximation for the total leftover inventory was derived that allowed the dynamic programming problem for inventory optimization to be solved in the two period case.

The numerical analyses also lead to some interesting results. In allocating a fixed amount of inventory in period 2, it was noted that if period 1 demands are either much higher or much lower than expected, then the number of stores to stock should be reduced. Also, a tradeoff arises between the collection of information in period 1 and the expected profit in period 1. Stocking more stores in period 1 provides more information, which in turn allows for more accurate forecasts for period 2, but it may require stocking some unprofitable stores and may result in excess inventory for period 2. The value of the information collected in period 1 also depends on how much flexibility the retailer has to respond to the new information. If there is no period 2 replenishment, for example, the value of the period 1 information is much less. Our methodology provides a way to calculate these tradeoffs in an operational setting.

We also analyzed how store variety affects the benefits associated with various forms of supply chain flexibility. Broadly speaking, we considered two sources of supply chain flexibility that may be available to retailers. The first can be viewed as *external* supply flexibility, i.e., the ability to revise replenishment orders after the selling season has begun. Although the incremental value of this flexibility can be significant in certain cases, our numerical analysis found that this only occurs if it is accompanied by Learning, i.e., the ability to update demand forecasts based on observed sales. Examples of industry initiatives that rely on external flexibility include Vendor Managed Inventory (VMI) and Vendor Owned Inventory (VOI), which evolved originally in the high technology industry and are offered by some brand name suppliers for retail department store items. (See, e.g., Achabal et al. [1].) Examples of retailers that have benefited significantly from external flexibility include the World Co. Raman et al. [29] and Zara [18]. A highly efficient and responsive supply chain allows Zara to operate with significant flexibility in store ordering and low markdowns. For instance, while typical stores might be allowed to change orders by up to 20% once the season has begun, Zara allows changes of up to 50%. Zara's unsold inventory is less than 10% of stock as compared to an average of nearly 20% for the industry.

For fashion and seasonal basic items that are manufactured as private label merchandise, our discussions with retailers indicate that this external flexibility is rarely available because of the typical long production, procurement and shipment lead times. In these cases, *internal* flexibility, which we have referred to as Store Adjustment, can be quite attractive. Internal flexibility uses updated forecasts to make midseason adjustments in the set of stores that carry a particular item. While store level sales data are typically available to update forecasts, certain organizational barriers may prevent midseason store adjustments. For example, some retailers may wish to present a common set of merchandise at all their stores, or store managers may resist centralized decision making with regard to their assortments. Our analysis quantifies the potential benefit of allowing this additional inventory flexibility. Our numerical studies found that the value of store inventory flexibility can be significant, even without learning. On the other hand, the value of learning alone was small without either external supply flexibility or Store Adjustment capability. Thus, internal flexibility increases the value of external ordering flexibility. Our numerical analyses indicate that these results hold over a wide range of costs and demand distribution assumptions.

Two possible generalizations of the models developed in this paper are to multiple products, and more than two time periods. Generalized Bayesian updating can be applied to multiple products that all share a common parameter θ, e.g., the products are the sizes S, M, L of a given style, and their inventory decisions can be made independently. However, a linkage occurs in period 1 when considering the shared value of the information. Thus, the optimal set of stores to stock in period 1 must be determined by a search across all nested combinations of stores to stock for each of the items. Unfortunately, this makes the optimization significantly more complex.

The optimization problem for more than two time periods can be formulated but the solution becomes very complex when only a single order is placed, because the probability distribution of leftover inventory is different for every period. For the case of replenishment in each period where the service level resets to the optimal value each time, the state space of the dynamic programming problem is still quite large, because the period 1 optimization must take into account all possible demand outcomes in periods 2, 3, ..., etc. Practically speaking, a multi-period generalization appears likely to have little incremental value for seasonal and fashion products, because the potential benefit of re-optimizing more than once is limited due to their short seasons.

We have also assumed that interstore shipments at the beginning of the second period are either costless, or that the cost is fixed. This assumption is reasonable when the likelihood of such shipments is minimal, or when the shipment schedule is predetermined (and hence the costs are sunk). However, when this is not the case, the actual ending inventory levels at each store are needed to calculate the optimal interstore transfers and the resulting shipping costs, which leads to a much more complex dynamic programming problem than the one considered here (see Paterson et al. [27] and the references therein for literature on the tradeoffs associated with such lateral shipments). Finally, combining optimal inventory allocation and price reoptimization in midseason is a challenging and unsolved problem, which could have a potential payoffs for retailers. We hope that our research encourages further developments in these areas.

Appendix A

Proof of Lemma 1b. The proof is by contradiction. Suppose that there are two store indices $x > x^*$ such that $x \in Y_2$ but $x^* \notin Y_2$. If we simply move the inventory $s_2(x)$ from store x to x^*, the change in total profit can be calculated based on (3) by replacing $P_t(d,x)$ with $P_2(d,x|\mathbf{d_1})$ to obtain

$$(c_u + c_{o2}) \int_{d \leqslant s_2(x)} [P_2(d,x|\mathbf{d_1}) - P_2(d,x^*|\mathbf{d_1})]. \tag{18}$$

This change is positive by (4). This contradicts the optimality of any Y_2 in (3) that contains "gaps." □

Proof of Lemma 2a. Dropping the subscript t, let $s(x,\alpha)$ and $s^*(x,\alpha)$ for each α be chosen so that $P(s(x,\alpha),x) \equiv P^*(s^*(x,\alpha),x) \equiv \alpha$. (Equality holds exactly only for continuous distributions.) Since optimal store sets do not contain gaps, we can write the total inventory allocated across $x \leqslant y$ as

$$S(\alpha, y) = \sum_{x \leqslant y} s(\alpha, x) dx, \quad S^*(\alpha, y) = \sum_{x \leqslant y} s^*(\alpha, x) dx \text{ for all } \alpha, y.$$

Suppose that $S^*(\alpha,y) \geqslant S(\alpha,y)$ for all α,y. Then let inventory I be allocated optimally to $\{x|0 \leqslant x \leqslant y\}$ which results in $S(\alpha,y) = S^*(\alpha^*,y) = I$. Thus, we have

$$0 = [S^*(\alpha^*, y) - S(\alpha^*, y)] + [S(\alpha^*, y) - S(\alpha, y)].$$

The first term is positive by assumption, so the second term must be negative. Since $\partial S(\alpha,y)/\partial\alpha > 0$, this implies that $\alpha^* < \alpha$.

Conversely, suppose that $\alpha^* < \alpha$ with the same inventory $I = S(\alpha,y) = S^*(\alpha^*,y)$. Since $\partial\ S^*(\alpha,y)/\partial\alpha > 0$, this means that $S^*(\alpha,y) > S(\alpha,y)$. □

Proof of Lemma 2b. Let $\{P^*(d,x)\}$ be steeper than $\{P(d,x)\}$ so that $S^*(\alpha,y) \geqslant S(\alpha,y)$ for all α,y. Based on (3) for continuous distributions without the t subscripts, the expected profit from stocking the stores $0 \leqslant x \leqslant y$ with service level α can be written as:

$$\Pi(\alpha, y) = c_u S(\alpha, y) - (c_u + c_o) \sum_{x \leqslant y} \int_{-\infty}^{s(\alpha,x)} P(\xi, x) d\xi - Ky - G,$$

with a similar expression for $\Pi^*(\alpha, y)$. First, integrate the integral by parts to obtain

$$\int_{-\infty}^{s(\alpha,x)} P(\xi, x) d\xi = s(\alpha, x) P(s(\alpha, x), x) - \int_{-\infty}^{s(\alpha,x)} \xi p(\xi, x) d\xi.$$

Now, introduce the change of variable $\beta = P(\xi, x)$. Since$P(s(\beta, x), x) = \beta$ for all β, it follows that $\xi = s(\beta, x)$ and $d\beta = p(\xi, x)d\xi$. Thus, the expected profit can be written as

$$\begin{aligned}\Pi(\alpha, y) &= c_u S(\alpha, y) - (c_u + c_o) \sum_{x \leqslant y} \left\{ s(\alpha, x)\alpha - \int_{-\infty}^{s(\alpha,x)} P(\xi, x) d\xi \right\} \\ &\quad - Ky - G, = \{c_u - (c_u + c_o)\alpha\} S(\alpha, y) + (c_u + c_o) \\ &\quad \times \int_0^\alpha S(\beta, y) d\beta - Ky - G,\end{aligned}$$

with a similar expression for $\Pi^*(\alpha, y)$. Now suppose that inventory $I = S(\alpha, y)$ is allocated optimally to the stores with the distribution family $\{P^*(d,x)\}$. By the steepness property, the resulting service level α^* must be less than α. Then consider

$$\begin{aligned}\Pi^*(\alpha^*, y) - \Pi(\alpha, y) &= (c_u + c_o) \left\{ (\alpha - \alpha^*) I + \int_0^{\alpha^*} S^*(\beta, y) d\beta - \int_0^\alpha S(\beta, y) d\beta \right\} \\ &= (c_u + c_o) \left\{ \int_{\alpha^*}^{\alpha} [I - S(\beta, y)] d\beta + \int_0^{\alpha^*} [S^*(\beta, y) - S(\beta, y)] d\beta \right\}.\end{aligned}$$

The last integral is non negative by the steepness property. Since $I = S(\alpha,y) \geqslant S(\beta,y)$ for all β such that $\alpha^* \leqslant \beta \leqslant \alpha$, the first integral is non- negative as well. Thus, when a fixed inventory I is optimally allocated, the expected profit is always at least as large for steeper demand distributions. □

This also implies, if the optimal inventory $S^*(\alpha, y) > S(\alpha, y)$ is allocated across the steeper family $\{P^*(d,x)\}$, that the profit will be even larger.

A.1. Formulas for the gamma–exponential distributions used in Section 4 calculations

This table summarizes the results for the gamma–exponential conjugate distributions used in the calculations. Detailed derivations available from the authors. Prior Distribution:

$$p(\theta) = \frac{(\theta v)^{k-1}}{\Gamma(k)} e^{-\theta v} v, \quad \text{for } \theta > 0, \quad k \geqslant 1, \quad E[\theta] = \frac{v}{k-1}$$

Likelihood function:

$$P\{D_t(x) = d|\theta\} = \frac{\theta}{f_t(x)} e^{-\theta d/f_t(x)}, \ d \geqslant 0.$$

Unconditional demand distribution:

$$P_t(d, x) = 1 - \left(1 + \frac{d}{v f_t(x)}\right)^{-k}, \ \text{for } d \geqslant 0. \tag{19}$$

For any service level α, stock level:

$$s_t(x) = f_t(x) v \{(1 - \alpha)^{-1/k} - 1\}. \tag{20}$$

and

$$\begin{aligned}\Pi_t(\alpha, y) &= \frac{v F_t(y)}{k-1} \{c_u + k c_{ot} - [c_u + k c_{ot} - \alpha(c_u + c_{ot})](1 - \alpha)^{-1/k}\} \\ &\quad - \frac{v w}{k-1} - K_t y.\end{aligned}$$

Optimal service level for constrained inventory I:

$$\alpha = 1 - \left(1 + \frac{I}{F_t(y) v}\right)^{-k}.$$

Updated parameter values for period 2:

$$k^* = k + y_1 \text{ and } v^* = v + \sum_{x \leqslant y_1} \frac{d_1(x)}{f_1(x)}. \tag{21}$$

Posterior distribution of θ:

$$p(\theta|\tau(\mathbf{d_1})) = \frac{(\theta v^*)^{k^*-1}}{\Gamma(k^*)} e^{-\theta v^*} v^*.$$

Appendix B. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ejor.2012.10.006.

References

[1] D.D. Achabal, S. McIntyre, S.A. Smith, K. Kalyanam, A decision support system for vendor managed inventory, Journal of Retailing 77 (4) (2000) 429–452.
[2] N. Agrawal, S.A. Smith, Estimating negative binomial demand for retail inventory management with unobservable lost sales, Naval Research Logistics 43 (4) (1996) 839–861.
[3] N. Agrawal, S.A. Smith, Retail Supply Chain Management: Quantitative Models and Empirical Studies, Springer, New York, 2009.
[4] S. Axsater, Continuous review policies for multi-level inventory systems with stochastic demand, in: S.C. Graves, A.H.G. Rinnooy Kan, P.H. Zipkin (Eds.), Handbooks in OR and MS, Logistics of Production and Inventory, vol. 4, Elsevier Science Publishers, 1993, pp. 175–197.
[5] K.S. Azoury, Bayes solution to dynamic inventory models under unknown demand distribution, Management Science 31 (9) (1985) 1150–1160.
[6] A. Bensoussan, M. Cakanyildirim, S.P. Sethi, Partially observed inventory systems: the case of zero-balance walk, SIAM Journal of Control Optimization 46 (1) (2007) 176–209.
[7] A. Bisi, M. Dada, Dynamic learning, pricing and ordering by a censored newsvendor, Naval Research Logistics 54 (4) (2007) 448–461.
[8] G.R. Bitran, E. Haas, H. Matsuo, Production planning of style goods with high setup costs and forecast revisions, Operations Research 34 (2) (1986) 226–236.
[9] J.W. Bradford, P.K. Sugrue, A Bayesian approach to the two-period style goods inventory problem with single replenishment and heterogenous poisson demands, Journal of Operations Research Society 41 (3) (1990) 211–218.
[10] L.Y. Chu, J.G. Shanthikumar, Z.M. Shen, Solving operational statistics via a Bayesian analysis, Operations Research Letters 36 (1) (2008) 110–116.
[11] L. Chen, E.L. Plambeck, Dynamic inventory management with learning about the demand distribution and substitution probability, Manufacturing and Service Operations Management 10 (2) (2007) 236–256.
[12] M.H. DeGroot, Optimal Statistical Decisions, McGraw-Hill, 1970.
[13] N. DeHoratius, A.J. Mersereau, L. Schrage, Retail inventory management when records are inaccurate, Manufacturing and Service Operations Management 10 (2) (2008) 257–277.
[14] X. Ding, M.L. Puterman, A. Bisi, The censored newsvendor and the optimal acquisition of information, Operations Research 50 (3) (2002) 528–537.
[15] N. Erkip, W.H. Hausman, S. Nahmias, Optimal centralized ordering policies in multi-echelon inventory systems with correlated demands, Management Science 36 (3) (1990) 381–392.

[16] G.D. Eppen, A.V. Iyer, Improved fashion buying with Bayesian updates, Operations Research 45 (6) (1997) 805–819.
[17] A. Federgruen, Centralized planning models for multi-echelon inventory systems under uncertainty, in: S.C. Graves, A.H.G. Rinnooy Kan, P.H. Zipkin (Eds.), Handbooks in OR and MS, Logistics of Production and Inventory, vol. 4, Elsevier Science Publishers, 1993, pp. 133–173.
[18] K. Ferdows, M.A. Lewis, J.A.D. Machuca, Rapid-fire fulfillment, Harvard Business Review (11) (2004).
[19] M.L. Fisher, A. Raman, Reducing the cost of demand uncertainty through accurate response to early sales, Operations Research 44 (1) (1996) 87–99.
[20] M.L. Fisher, K. Rajaram, Accurate retail testing of fashion merchandise: methodology and application, Marketing Science 19 (3) (2000) 266–278.
[21] W.H. Hausman, R. Peterson, Multi-product production scheduling for style goods with limited capacity, forecast revisions and terminal delivery, Management Science 18 (7) (1972) 370–383.
[22] A.G. Kok, K. Shang, Inspection and replenishment policies for systems with inventory record inaccuracy, Manufacturing and Service Operations Management 9 (2) (2007) 185–215.
[23] M. Lariviere, E.L. Porteus, Stalking information: Bayesian inference with unobservable lost sales, Management Science 45 (3) (1999) 346–363.
[24] X. Lu, J.S. Song, K. Zhu, Analysis of perishable inventory systems with censored demand data, Operations Research 56 (4) (2008) 1034–1038.
[25] H. Matsuo, A stochastic sequencing problem for style goods with forecast revisions and hierarchical structure, Management Science 36 (3) (2008) 332–347.
[26] G.R. Murray, E.A. Silver, A Bayesian analysis of the style goods inventory problem, Management Science 12 (11) (1966) 785–797.
[27] C. Paterson, R. Teunter, K. Glazerbrook, Enhanced lateral transshipments in a multi-location inventory system, European Journal of Operational Research 221 (1) (2012) 317–327.
[28] N.C. Petruzzi, M. Dada, Dynamic pricing and inventory control with learning, Naval Research Logistics 49 (3) (2002) 303–325.
[29] A. Raman, M.L. Fisher, A.S. McClelland, Supply Chain Management at World Co., Ltd. Harvard Business School Case 601-072, 2001.
[30] H. Scarf, Some remark on Bayes solutions to the inventory problem, Naval Research Logistics Quarterly 7 (1960) 59–596.