Dynamic Inventory Allocation with Demand Learning for Seasonal Goods

Citation: Nambiar, Mila, Simchi#Levi, David and Wang, He. 2020. "Dynamic Inventory Allocation with Demand Learning for Seasonal Goods." Production and Operations Management, 30 (3).

As Published: http://dx.doi.org/10.1111/poms.13315

Publisher: Wiley

Persistent URL: https://hdl.handle.net/1721.1/140503

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Dynamic Inventory Allocation with Demand Learning for Seasonal Goods

Mila Nambiar

Agency for Science, Technology and Research, Singapore. milanambiar@gmail.com

David Simchi-Levi*

Institute for Data, Systems and Society, Department of Civil and Environmental Engineering, and Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA. dslevi@mit.edu

He Wang

H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA. he.wang@isye.gatech.edu

October 20, 2020

Abstract

We study a multi-period inventory allocation problem in a one-warehouse multiple-retailer setting with lost sales. At the start of a finite selling season, a fixed amount of inventory is available at the warehouse. Inventory can be allocated to the retailers over the course of the selling horizon (transshipment is not allowed). The objective is to minimize the total expected lost sales and holding costs. In each period, the decision maker can use the realized and possibly censored demand observations to dynamically update demand forecast and consequently make allocation decisions. Our model allows a general demand updating framework, which includes ARMA models or Bayesian methods as special cases. We propose a computationally tractable algorithm to solve the inventory allocation problem under demand learning using a Lagrangian relaxation technique, and show that the algorithm is asymptotically optimal. We further use this technique to investigate how demand learning would affect inventory allocation decisions in a two-period setting. Using a combination of theoretical and numerical analysis, we show that demand learning provides an incentive for the decision maker to withhold inventory at the warehouse rather than allocating it in early periods.

Keywords: multi-echelon inventory; demand learning; dynamic programming

History: received: September 2020; accepted: October 2020 by Christopher Tang after one revision.

*Corresponding author.

1 Introduction

We study an inventory allocation problem for seasonal goods in a one-warehouse multiple-retailer system. At the start of the selling season, a fixed amount of inventory is available at the warehouse. The decision maker must decide how to allocate this inventory from the warehouse to the retailers. We assume that lateral transshipment between the retailers is not allowed. The decision maker's objective is to minimize the total expected lost sales costs and holding costs.

Our study is motivated by the challenges faced by a large US retail chain to manage its seasonal merchandise inventory. The retail chain's nationwide distribution network is divided into several regions, where each region has around 30 to 50 retail locations serviced by a regional distribution center. Seasonal goods offered by this retailer usually have selling horizons that last from 12 to 16 weeks. Demand uncertainty is high, and there is little salvage value associated with seasonal goods at the end of their life cycles. Due to lengthy supply lead times, production orders must be placed well before the selling season begins, and in most cases, inventory cannot be replenished during the selling season after the initial order quantity is determined. There is no transshipment between the store locations. Therefore, the decision to allocate warehouse inventory among different retail locations can have a significant impact on sales.

For one-warehouse multiple-retailer systems, it has been widely observed in the literature (Jackson 1988, Jackson and Muckstadt 1989) that the warehouse can reduce inventory cost by making *multiple* allocations throughout the selling season, rather than allocate all its inventory to the retailers at once. The benefit of postponement and delayed allocation has also been studied by Lee (1996), Lee and Tang (1997), and Cheung and Lee (2002). The intuition is that if some retailers experience high demands in earlier periods, reserving inventory at the warehouse allows the retailer to re-balance inventory levels by allocating more inventory to these retailers (and less inventory to retailers who experience low demands) in later periods. This ability to mitigate demand fluctuations by storing inventory in a central location and delaying allocation is known as the "risk-pooling" effect or the "postponement" strategy.

In this work, our chief contribution is to study a second motive – namely, demand learning – for the decision maker to delay inventory allocations to later periods. When the customer demand distribution is unknown, learning demand allows the decision maker to continuously improve her demand forecasts during the season, by observing historical demand and updating her demand beliefs in response to these observations. Intuitively, one would expect that by reserving inventory at the warehouse, the decision maker can make more informed allocation decisions later on in the time horizon, and allocate more inventory to retailers who have been observed to experience high demands, and less inventory to retailers who have been observed to experience low demands. In this paper, we provide evidence to support this intuition by showing that we can expect the decision maker's first-period allocations from the warehouse to the retailers to *decrease* with the extent of learning.

To prove this property, we first address the question of how the decision maker can optimally allocate inventory from the warehouse to the retailers in a computationally tractable way. Since the exact solution to the allocation problem becomes computationally intractable as the number of retailers grows large, we propose a heuristic method. Of the many heuristics proposed in the literature (Federgruen and Zipkin 1984, Jackson 1988, McGavin et al. 1993) for the two-echelon inventory allocation problem, we base our work on the heuristic developed by Marklund and Rosling (2012) for a backordering, independent demand setting. The idea behind their method is that the source of the computational complexity in the allocation problem is that the fixed warehouse inventory couples the allocation decisions across the retailers. If each retailer had to pay some ordering cost for each unit of inventory, instead of satisfying the fixed warehouse inventory constraint exactly, the allocation problem would decouple, and each retailer could solve its inventory ordering problem separately.

We propose a heuristic that relaxes the fixed inventory constraint to allow for this problem decomposition. We prove an optimality gap between the expected costs of applying the heuristic and the optimal value of the original allocation problem. A key difference between our result and that of Marklund and Rosling (2012), besides the fact that they study a backordering setting and we study a lost sales setting, is that our result applies to a setting with correlated demands, which includes demand forecasting settings as well as settings with Bayesian learning. In Marklund and Rosling (2012), however, demands at each retailer are assumed to be i.i.d. across time.

Then, using the heuristic as a proxy for the exact optimal solution, we investigate how the decision maker's allocation policy depends on demand learning. We consider a two-period demand forecasting model with identical retailers that parametrizes the extent of demand uncertainty in the second period. We show analytically, by further approximating our lost sales setting with a no carryover inventory setting (where no remaining inventory at the retailers can be carried over to the next period), that the dependence of the first-period allocation on the level of uncertainty is consistent with the property that early allocations should decrease as the amount of demand learning increases. This allows us to conclude that demand learning complements risk pooling in incentivizing the decision maker to reserve inventory at the warehouse, and to delay inventory allocations to later periods.

We also separately use the heuristic to prove an additional structural result on the allocation decisions when the retailers are non-identical. In particular, we look at a setting with independent but non-identical retailers experiencing truncated normal demands, and who share the same demand means but different variances. We ask how the retailer should prioritize among these retailers when allocating inventory. We show that the decision maker's strategy will depend on the amount of available warehouse inventory: When the warehouse inventory is small, the decision maker should favor a conservative policy, and allocate more inventory to retailers with lower demand variances, since these retailers have a lower chance of experiencing low demands. On the other hand, when the warehouse inventory is large, the decision maker should take a risk on retailers with higher demand variances, and allocate more inventory to these retailers, since they have a higher chance of experiencing high demands.

1.1 Literature Review

Inventory allocation in two-echelon systems was studied in the seminal work of Clark and Scarf (1960), who observed that the complexity of this problem relative to a serial system stems from the fact that that the retailers' inventory positions cannot be lowered through transshipments or returns to the warehouse. This causes the optimal allocations to depend not only on the echelon inventory level, but also on the inventory positions at *all* locations. Tan (1974) analyzed the structure of optimal policies for a special case with two retailers. Since the inventory allocation problem becomes computationally intractable as soon as either the number of retailers or the number time periods grows large, several papers have proposed effective but computationally tractable policies that rely on approximations of the original problem. Jackson (1988) considers a class of "order-up-to-S" policies, where the warehouse stocks each retailer (all of whom are assumed to experience i.i.d. demand) up to S every period until it runs out of inventory. He proposes that when the warehouse runs out of inventory to allocate all the retailers' inventory levels up to S, it should solve a "run out allocation" problem, and develops approximations to efficiently solve this optimization problem. Jackson and Muckstadt (1989) study a two-period model with backordering. They approximate the cost function by analyzing the case where the number of retailers tends to infinity, and use this approximation to develop an efficient optimization procedure. McGavin et al. (1993) and McGavin et al. (1997), like us, study the lost sales setting, but with only two periods and identical retailers. They show that the optimal policy takes the form of balancing policy, and also propose a heuristic, known as the infinite retailer heuristic, which estimates the first-period optimal allocations and second-period order-up-to levels by approximating their set up with a deterministic setting with infinitely many retailers.

The above papers demonstrate the effectiveness of their proposed heuristics through numerical simulations. Marklund and Rosling (2012) is the first that we know of that establishes the optimality gap between the expected cost of their heuristic and that of the exact optimal solution. They study a backordering setting with N non-identical retailers, each of whom experiences demand that is i.i.d. across time. They propose a Lagrangian relaxation heuristic and prove that the ratio of the optimality gap between the heuristic and the original problem to the value of the optimal problem is bounded by $O(\sqrt{N})$, implying that this optimality gap goes to 0 as the number of retailers grows large. In this work, we adapt the result in Marklund and Rosling (2012) to our lost sales, correlated demand setting to get the same optimality gap bound of $O(\sqrt{N})$. The main technical contribution we have made in adapting their proof lies in showing the convexity of the relaxed optimization problem in the lost sales setting. We give a sufficient condition on the relationship between prices and holding costs that guarantees the convexity of this optimization problem, which ensures that their result holds in our setting.

Another stream of literature that our paper is closely related to is the literature on inventory allocation in a two-echelon setting with demand learning. Many of these papers are motivated by real examples of allocation decision problems faced by fashion retail companies, where demand in later periods of a season can only be learned by observing early sales. Eppen and Iyer (1997) consider a catalog merchandiser that uses early (uncensored) sales to learn whether a fashion item will be popular or not. Agrawal and Smith (2013) study a two-period multi-location inventory allocation problem, but they assume that retailers' inventory can be sent back to the warehouse with no additional cost. Fisher and Raman (1996) study a two-period model where the first-period allocation is unconstrained but the total second-period allocations are limited. They approximate the decision maker's optimization problem in order to solve it in a more computationally tractable way, and test their algorithm on data from the fashion retail company Sport Obermeyer. In addition, they look at the special case that demand is bivariate normal, and give a closed-form solution of the optimal first-period allocations to the retailers in terms of their demand means and variances. This is related to our result in Section 4.2, where we compare the allocations to retailers with different demand variances. However, the result in Fisher and Raman (1996) does not show how the warehouse should prioritize among retailers with low and high variances. In this work, we show that the warehouse's strategy should depend on the amount of available inventory at the warehouse.

Besides Fisher and Raman (1996), Fisher and Rajaram (2000) study the problem of merchandise testing, i.e. of allocating small amounts of inventory to a small number of selected retailers before the season starts to learn demand. They develop an algorithm to determine which retailers testing inventory should be allocated to in order to maximize learning during the testing period, and test their algorithm on data from a real fashion retailer. Gallien et al. (2017) work with Zara to study the problem of determining inventory allocations to retailers early in the season. They propose an algorithm that approximately solves this problem, prove an asymptotic optimality bound on the proposed algorithm, and run field experiments to validate its performance. We note that all three of these demand learning papers study two-period models. This shows that many real-world settings faced by fashion retailers can be formulated as two-period models, and suggests that the model used in our structural analysis in Section 4 to a two-period setting is not only simpler to analyze, but also practical.

It is also worth comparing our work with several papers on demand learning for the single retailer setting. This classical problem has been studied for several decades, see Clark and Scarf (1960), Iglehart (1964), Azoury (1985), Miller (1986), Lovejoy (1990), Lariviere and Porteus (1999), Lu et al. (2008), Chen and Plambeck (2008), Chen (2010), Jain et al. (2014), and references therein. Our Lagrangian relaxation heuristic introduced in Section 3 aims to reduce the multi-retailer problem to N separate single retailer problems, for which the computation methods developed by these papers can be readily applied. From this stream of literature, Ding et al. (2002) study a two-period newsvendor model where demand learning can take place through Bayesian updates of unknown parameters based on observed sales data. They compare the optimal first-period allocations when demand is censored (i.e. demand at each retailer can only be observed up to its inventory position) with the optimal allocations when demand is fully observable. They analytically derive the intuitive result that when demand is censored, the decision maker should allocate more inventory in the first-period so as to obtain more accurate demand information. Azoury (1988) and Azoury and Miller (1984) also study a two-period model, though with backordering, and like us, they compare the optimal first-period allocations with and without learning. In their setting, however, demand learning takes place through Bayesian updates of the unknown parameters, whereas our structural analysis in Section 4 is based on a demand forecasting model. They prove that under specific assumptions on demand (such as the fact that it belongs to a family of distributions that satisfies what is known as the single crossing property), the first-period allocations are greater without learning (when the parameter is updated) than when it is updated. The intuition behind this result is that in the learning setting, a parameter update could reveal that demand is on average lower than anticipated; Then, allocating too much inventory in the first-period puts the decision maker at risk of having a higher inventory position than is optimal. In this work, we derive a similar structural result that says that the first-period allocations are decreasing with the extent of learning. However, because we assume a multiple-retailer setting with fixed inventory (whereas in Azoury (1988) and Azoury and Miller (1984) inventory is unlimited), the interpretation of our result is different. In our case, allocating less inventory when there is learning has to do with saving inventory for the second period, when demand uncertainty is lower, and it is clearer which retailers will experience high demands, and which retailers will experience low demands.

1.2 Notation

In the paper, (column) vectors are denoted by boldface letters. Let $\mathbf{1}_n$ denote a vector of all ones of length n, and let $\mathbf{0}_n$ denote a vector of all zeros of length n. We will omit the subscript n when the dimension of these vectors is evident. For any m-dimensional vector $\mathbf{x}$, let $[\mathbf{x}]^+ := \max\{0, \mathbf{x}\}$ where max is the element-wise maximum.

2 Model

We study an inventory allocation problem for seasonal goods in a two-echelon network with one warehouse and N retailers. At the start of the selling season of length T, a fixed amount of inventory w_0 is available at the warehouse, while the N retailers have no starting inventory. In each time period $t = 1, 2, \ldots, T$, the warehouse can choose to allocate some amount of inventory $a_{i,t} \geq 0$ to each retailer i. Let $\mathbf{a}_t$ denote the N dimensional vector with i-th entry being $a_{i,t}$. Transshipments (moving inventory between retailers) or returning inventory from the retailers to the warehouse are not allowed, and we assume that there is no additional replenishment to the warehouse during the season. This implies that the total allocation to all retailers over the selling horizon satisfies $\sum_{i=1}^{N} \sum_{t=1}^{T} a_{i,t} \leq w_0$ almost surely.

We denote the retailer inventory positions at the beginning of period t by a N-dimensional vector $\mathbf{x}_t$. The retailer inventory positions after the allocations are denoted by the N-dimensional vector $\mathbf{y}_t$, where $\mathbf{y}_t = \mathbf{x}_t + \mathbf{a}_t$. After the warehouse allocates inventory to the retailers, each retailer i then sells the product at price $p_{i,t}$ and correspondingly observes demand $D_{i,t}$. We assume that for each i and t, $D_{i,t}$ is a random variable that is discrete and bounded uniformly by some constant $D_{\max}$. We also assume that demand is independent from retailer to retailer. However, for each retailer, we allow the demands $D_{i,t}$ to be correlated across time. In particular, the distribution of $D_{i,t}$ can depend on the demand history $\{D_{i,1}, \ldots, D_{i,t-1}\}$ as well as other exogenous variables.

We study a lost sales setting with no lead times. Any demand that is not met is lost and not backordered, giving a starting inventory position at the retailers at time t of $\mathbf{x}_{t+1} = [\mathbf{y}_t - \mathbf{D}_t]^+$. Any demand that is not met by retailer i incurs a per-unit lost sales cost of $p_{i,t}$, where $p_{i,t}$ is the price of the product and $p_{i,t} \geq 0$. Any demand that remains at retailer i at time t incurs a per-unit holding cost of $h_{i,t}$, with $h_{i,t} \geq 0$. We ignore transportation costs.

The decision maker's objective is to minimize the expected discounted costs for a given discount factor is $0 < \gamma \leq 1$ incurred by the different retailers over the course of the selling horizon. At time t, given that retailer i has a post allocation inventory position of $y_{i,t}$, the cost incurred by this retailer is this is the sum of a lost sales component of $p_{i,t}[D_{i,t} - y_{i,t}]^+$ and a holding cost component of $h_{i,t}[y_{i,t} - D_{i,t}]^+$. We define the cost

$$L_{i,t}(y_{i,t}) := p_{i,t}[D_{i,t} - y_{i,t}]^+ + h_{i,t}[y_{i,t} - D_{i,t}]^+.$$

The decision maker then determines inventory allocations to the retailers based on the current warehouse inventory, the history of demand observations, and its belief of future demand. At each time t, we assume she has knowledge of the history of past demand realizations, which can either be the full demand history $\mathcal{H}_t = \{D_{i,s}, \forall i = 1, \ldots, N, s = 1, \ldots, t-1\}$ or the censored information $\mathcal{H}_t = \{y_{i,s} \wedge D_{i,s}, \forall i = 1, \ldots, N, s = 1, \ldots, t-1\}$. With this knowledge, she can derive the conditional distribution of current demand $\mathbb{P}[D_{i,t} \mid \mathcal{H}_t]$ for $i = 1, \dots, N$ on these past demand realizations. However, she does not know the current demand realizations beforehand.

Using the notation defined above, the inventory allocation optimization problem at time t can be formulated with the dynamic program given by equations (1)-(3). The cost-to-go function V_t, is simply 0 at time $T+1$, and can otherwise be recursively computed by miniminizing (2).

$$G_t(\mathbf{y}, w_t, \mathbf{x}_t, \mathcal{H}_t) = \sum_{i=1}^{N} \mathrm{E}\left[L_{i,t}(\mathbf{y}) \mid \mathcal{H}_t\right] + \gamma \mathrm{E}\left[V_{t+1}(w_t - \mathbf{1}^{\mathsf{T}}(\mathbf{y} - \mathbf{x}_t), [\mathbf{y} - \mathbf{D}_t]^+, \mathcal{H}_{t+1}) \mid \mathcal{H}_t\right] \tag{1}$$

$$V_t(w_t, \mathbf{x}_t, \mathcal{H}_t) = \min_{\substack{\mathbf{y} \geq \mathbf{x}_t \\ \mathbf{1}^{\mathsf{T}}\mathbf{y} \leq \mathbf{1}^{\mathsf{T}}\mathbf{x}_t + w_t}} G_t(\mathbf{y}, w_t, \mathbf{x}_t, \mathcal{H}_t) \tag{2}$$

$$V_{T+1}(w_{0,T+1}, \mathbf{x}_{T+1}, \mathcal{H}_{T+1}) = 0. \tag{3}$$

2.1 Demand Models with Learning/Forecasting

In the above model formulation, we describe a general demand learning framework where demand forecasts are updated in each period using realized demand/sales information from the last period. Two types of demand learning models that this framework can capture are Bayesian methods, where the decision maker updates her beliefs on the unknown demand model parameter distributions with time, and time series models (e.g., ARMA or ARIMA). These models are described below.

Bayesian methods. Suppose the demand distributions are parameterized by a vector $\boldsymbol{\theta} \in \Theta$, and the PMF of demand at retailer i in period t is given by $p_{i,t}(d \mid \boldsymbol{\theta})$ (recall that we assume demand is discrete). If the decision maker has a prior on the parameters, she can update the posterior on these parameters based on the demand realization history, and thus knows the posterior distribution of demand at any retailer and time period conditional on this history. In particular, for all $i = 1, \dots, N$, $t = 1, \dots, T$, the demand distribution of $D_{i,t}$ is given by

$$\mathbb{P}[D_{i,t} = d \mid \mathcal{H}_t] = \int_{\boldsymbol{\theta} \in \Theta} p_{i,t}(d \mid \boldsymbol{\theta}) f(\boldsymbol{\theta} \mid \mathcal{H}_t) \mathrm{d}\boldsymbol{\theta},$$

where $f(\boldsymbol{\theta} \mid \mathcal{H}_t)$ is the posterior distribution of the parameters $\boldsymbol{\theta}$ given history $\mathcal{H}_t$. Details of Bayesian methods for demand learning in inventory models can be found in, e.g., Scarf (1960), Azoury (1985), Lovejoy (1990), Lariviere and Porteus (1999), Chen (2010), among others.

ARMA/ARIMA models. Our model framework also allows another type of demand learning approach based on time series models, such as the autoregressive moving average (ARMA) model and the autoregressive integrated moving average (ARIMA) model. In this case, the demand at each retailer and time period is a weighted sum of the previous period demand at that retailer as well as some demand noise. For example, consider the following ARMA(p, q) model: the demand for period t and retailer i is given by

$$D_{i,t} = \mu_i + \alpha_1 D_{i,t-1} + \cdots + \alpha_p D_{i,t-p} + \epsilon_{i,t} + \beta_1 \epsilon_{i,t-1} + \cdots + \beta_q \epsilon_{i,t-q},$$

where μ_i is a constant, $\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q$ are parameters, and $\epsilon_{i,t}, \epsilon_{i,t-1}, \dots$ are error terms with mean zero. Assuming the full demand information is observed (i.e., no demand censoring), the history $\mathcal{H}_t$ includes information about previous demand at the retailer $D_{i,t-1}, D_{i,t-2}, \ldots$ and the noise terms $\epsilon_{i,t-1}, \epsilon_{i,t-2}, \ldots$. Given the parameters of this ARMA model, the decision maker can update demand forecasts based on this history. See Aviv (2003), Gilbert (2005) for examples that apply time series models to inventory management.

3 Computating Inventory Allocation Decisions

In this section, we discuss computation issues of the dynamic programming model. The dynamic program (1)-(3) is difficult to solve explicitly, as it suffers from the curse of dimensionality, where its state space grows exponentially in the number of retailers N. In this section, we consider a heuristic that is computationally much less expensive to implement. This heuristic is based on the observation that (1)-(3) is weakly coupled, meaning that if the constraint $\mathbf{1}^\mathsf{T}\mathbf{y} \leq \mathbf{1}^\mathsf{T}\mathbf{x}_t + w_t$ is relaxed, the problem decouples into N separate dynamic programs, reducing the computational complexity to linear in N. The idea is to achieve this problem decomposition by approximating the original problem with its Lagrangian relaxation. We dualize the coupling constraint and add an associated Lagrangian term to the objective function, thus decomposing the problem into N separable optimization problems. This Lagrangian relaxation technique has been used to approximate weakly coupled optimization problems for a variety of applications. A survey is given in Adelman and Mersereau (2008). In the literature on inventory allocation and risk pooling, Marklund and Rosling (2012), who also study an inventory allocation problem for single warehouse multiple retailer setting, proposed a heuristic based on this technique. They show that this relaxation technique gives a lower bound on the optimal cost-to-go function, and that the performance of the heuristic converges to the lower bound as N goes to infinity, implying that the heuristic is asymptotically optimal. However, while Marklund and Rosling (2012) assumed that each retailer's demands are i.i.d. across time, we allow the demands experienced by each retailer to be correlated across time, such as in the demand learning setting described in the previous section. Another difference is that Marklund and Rosling (2012) assumed a backordering setting, while we consider a lost sales setting. Below, we show how the Lagrangian relaxation approach can be adapted to our setting.

3.1 Lagrangian Relaxation Heuristic

We start by rewriting (1)-(3) as a stochastic program as follows:

$$\begin{aligned} \min \quad & \sum_{t=1}^{T}\sum_{i=1}^{N}\gamma^{t-1}\mathrm{E}\left[\mathrm{E}[L_{i,t}(y_{i,t})|\mathcal{H}_t]\right] \\ \text{subject to} \quad & y_{i,t} = x_{i,t} + a_{i,t}\ \forall i,t,\mathcal{H}_t \\ & x_{i,t+1} = [y_{i,t} - D_{i,t}]^+,\ \forall i,t,\mathcal{H}_t \\ & \sum_{i=1}^{N}\sum_{t=1}^{T} a_{i,t} \leq w_0 \text{ a.s.} \\ & a_{i,t} \geq 0\ \forall i,t,\mathcal{H}_t\,. \end{aligned} \tag{P1}$$

Here, the variables $a_{i,t}, x_{i,t}, y_{i,t}$ are non-anticipative and are functions of the demand history $\mathcal{H}_t$.

To achieve the desired problem decomposition, we will relax the inventory constraint

$$\sum_{i=1}^{N}\sum_{t=1}^{T} a_{i,t} \le w_0$$

by requiring that it is only satisfied in expectation over all demand realizations, rather than almost surely. This gives the constraint $\sum_{i=1}^{N}\sum_{t=1}^{T} \mathrm{E}[a_{i,t}] \le w_0$. We will also eliminate the state variables $x_{i,t}$ and $y_{i,t}$ from the formulation and express the problem in terms of the decision variables $a_{i,t}$. This can be achieved by using the equations

$$\begin{aligned} x_{i,t} &= \max\{0, \sum_{s=u}^{t-1}(a_{i,s} - D_{i,s}),\ \forall 1 \le u \le t-1\}\,, \\ y_{i,t} &= x_{i,t} + a_{i,t}\,. \end{aligned}$$

This gives

$$\begin{aligned} \min\quad & \sum_{t=1}^{T}\sum_{i=1}^{N}\gamma^{t-1}\mathrm{E}\left[\mathrm{E}\big[L_{i,t}(a_{i,t} + \max\{0, \sum_{s=u}^{t-1}(a_{i,s} - D_{i,s}),\ \forall 1 \le u \le t-1\})|\mathcal{H}_t\big]\right] \\ \text{subject to}\quad & \sum_{i=1}^{N}\sum_{t=1}^{T}\mathrm{E}[a_{i,t}] \le w_0 \\ & a_{i,t} \ge 0 \qquad \forall i, t, \mathcal{H}_t\,. \end{aligned} \tag{P2}$$

We then consider the Lagrangian relaxation for the constraint $\sum_{i=1}^{N}\sum_{t=1}^{T}\mathrm{E}[a_{i,t}] \le w_0$. This gives the following decomposition:

$$\begin{aligned} \max\quad & -\lambda w_0 + \sum_{i=1}^{N}\mathrm{SUB}_i(\lambda) \\ \text{subject to}\quad & \lambda \ge 0\,, \end{aligned} \tag{D1}$$

where $\mathrm{SUB}_i(\lambda)$ is the optimal objective value to the following problem

$$\begin{aligned} \mathrm{SUB}_i(\lambda) := \min\quad & \sum_{t=1}^{T}\gamma^{t-1}\mathrm{E}\left[\mathrm{E}\big[L_{i,t}(a_{i,t} + \max\{0, \sum_{s=u}^{t-1}(a_{i,s} - D_{i,s}),\ \forall 1 \le u \le t-1\})|\mathcal{H}_t\big]\right] \\ & + \sum_{t=1}^{T}\gamma^{t-1}\mathrm{E}\left[\lambda\gamma^{1-t}a_{i,t}\right] \\ \text{subject to}\quad & a_{i,t} \ge 0 \qquad \forall t, \mathcal{H}_t\,. \end{aligned} \tag{D2}$$

For a given value of $\lambda \ge 0$, the subproblem (D2) is a single retailer inventory problem, where the unit ordering cost at period t is $\lambda\gamma^{1-t}$. Existing algorithms for single echelon systems can be readily applied to solve (D2) using either Bayesian methods or time series methods (see Section 2.1). Also, the N retailers' subproblems can be solved independently and in parallel. The master problem (D1) tries to find the optimal value of the Lagrangian multiplier λ.

Before we show how to find the optimal λ^* that maximizes (D1), we discuss the convexity of the allocation problem (P2). The optimization problem (P2) is convex if $1 \le i \le N$, $p_{i,t} + h_{i,t} \ge$

$\gamma p_{i,t+1}$, that is, the discounted price at period $t+1$ is no more than the sum of the price and holding cost at period t. The convexity result is stated in Lemma 1 and proven in Appendix A.1.

Lemma 1. *Suppose that $p_{i,t} + h_{i,t} \geq \gamma p_{i,t+1}$ for all $1 \leq t < T$ and $1 \leq i \leq N$. Then* (P2) *is a convex optimization problem. By strong duality, the optimal values of* (P2) *and* (D1) *are equal.*

An intuitive interpretation of the condition in Lemma 1 is that there should be no incentive for a retailer to reject any customer demand: if the retailer fulfills one unit of demand in period t, it receives revenue $p_{i,t}$; if the retailer rejects the current demand and withholds this unit for the next period, it pays holding cost $h_{i,t}$ and receives at most $\gamma p_{i,t+1}$. The latter results in lower profit since $\gamma p_{i,t+1} - h_{i,t} \leq p_{i,t}$. A sufficient condition for $p_{i,t} + h_{i,t} \geq \gamma p_{i,t+1}$ is that the prices at each retailer are nonincreasing with time, a widely adopted practice known as *markdown pricing*. In this case, we simply have $p_{i,t} \geq p_{i,t+1} \geq \gamma p_{i,t+1}$.

Given the convexity result in Lemma 1, we can solve for the optimal value λ^* that maximizes (D1) as follows. Let $a^*_{i,t}(\lambda^*)$ be the optimal solution to (D2) associated with λ^*. By Lemma 1, they must satisfy the complementary slackness condition of (P2), namely

$$\lambda^* \left(w_0 - \sum_{t=1}^{T} \sum_{i=1}^{N} \mathrm{E}[a^*_{i,t}(\lambda^*)] \right) = 0.$$

Define $p_{\max} = \max_{1 \leq i \leq N, 1 \leq t \leq T} \{p_{i,t}\}$. We know that λ^* must always fall within the range $[0, p_{\max}]$, since for all λ such that $\lambda > p^*$, the ordering cost $\lambda \gamma^{1-t}$ in subproblem (D2) is always greater than the retail price $p_{i,t}$. Therefore, the retailers will not order any inventory from the warehouse, i.e., $a^*_{i,t}(\lambda) = 0$, which violates complementary slackness condition. Therefore, the optimal solution λ^* can be found by searching the interval $[0, p_{\max}]$ (e.g., using the bisection method) until the complementary slackness condition is satisfied. These steps are summarized in the following inventory allocation heuristic.

0. Set the initial condition $\underline{\lambda} = 0, \overline{\lambda} = p_{\max}$.
1. Set $\lambda = (\underline{\lambda} + \overline{\lambda})/2$. For each retailer i, solve (D2) with λ and calculate the corresponding optimal allocations $a^*_{i,t}(\lambda)$ for all demand scenarios. (This step can be done for each retailer in parallel.)
2. Calculate the sum $\sum_{t=1}^{T} \sum_{i=1}^{N} \mathrm{E}[a^*_{i,t}(\lambda)]$. If the sum is greater than w_0, update $\underline{\lambda}$ to be the current value of λ; otherwise, update $\overline{\lambda}$ to be the current value of λ. If $\overline{\lambda} - \underline{\lambda}$ is bigger than the predetermined tolerance, go back to Step 1.
3. Set $\lambda^* = \lambda$. After the season starts, allocate inventory according to $a^*_{i,t}(\lambda^*)$ until the warehouse runs out inventory or until the season ends.

3.2 Optimality Bound

We will now analyze the performance of the Lagrangian relaxation heuristic. This result is stated in Theorem 1 below.

Theorem 1. *Suppose $p_{i,t} + h_{i,t} \geq \gamma p_{i,t+1}$ for all $1 \leq t < T$. Denote the expected cost of the Lagrangian relaxation heuristic by* UB*, and denote the value of* (P1) *by* OPT*. Then we have* $\mathsf{UB} - \mathsf{OPT} = O(\sqrt{N}T)$.

If we assume further that the value of each retailer subproblem (D2) is always lower bounded by a constant for $\lambda = \lambda^*$, it is easy to see that OPT is lower bounded by N times this constant. Then, Theorem 1 implies that $\frac{\mathsf{UB}-\mathsf{OPT}}{\mathsf{OPT}}$ is $O(1+\frac{1}{\sqrt{N}})$. The heuristic would thus be asymptotically optimal in the number of retailers N, in that its expected costs converge to the costs of the optimal policy as N grows large. However, the heuristic is not necessarily asymptotically optimal in the length of the selling horizon T, because we have assumed a fairly general condition of demands: they can be correlated over the selling horizon, and their distribution can be unknown.

The proof of Theorem 1 is deferred to Appendix A.2. The idea behind the proof is that the value of (D1) is a lower bound on (P1). This follows from the fact that the convexity of (P2) implies strong duality. Thus the relaxed optimization problem (P2), whose optimal value is a lower bound on (P1), is equal to its dual (D1). Now the cost of the heuristic is clearly an upper bound on the optimal value of the problem (thus we denoted it by UB). The cost of applying the heuristic is greater than the optimal value of (D1) exactly when the sum of the recommended allocations (i.e., $\sum a_{i,t}$) is greater than its expected value. Since the demand of the N retailers are independent, the difference between the sum of the recommended allocations and its expected values is of order $\sqrt{N}$ rather than N, thus allowing us to bound the optimality gap as a factor of $\sqrt{N}$.

We would like to make two remarks comparing the analysis of Theorem 1 with the analysis in Marklund and Rosling (2012). Firstly, in the backordering setting studied in Marklund and Rosling (2012), the convexity of the relaxed inventory allocation problem is guaranteed regardless of the assumptions on prices and holding costs. However, in our lost sales setting, we prove a sufficient condition on prices and holding costs to guarantee the convexity of (P2). Secondly, our heuristic and optimality bound, unlike Marklund and Rosling (2012), can be applied to settings with correlated demand, and hence with demand learning frameworks described in Section 2.1. However, it is important to note that the optimality bound 1 is not a regret bound and that it does not specify the rate of learning any underlying demand model parameters.

4 The Impact of Demand Learning on Inventory Allocations

The heuristic presented in Section 3 allows us to solve the two-echelon inventory allocation problem in a computationally tractable way by expressing this problem in terms of the simpler single retailer inventory ordering problem. In this section, we will also use this connection to shed light on the structure of inventory allocation policies. Using the Lagrangian relaxation heuristic as a proxy for the exact optimal solution, we investigate how the decision maker's inventory allocation policies depend on the effect of demand learning.

To address this question, we limit our analysis in this section to a two-period setting ($T = 2$), but continue to allow the number of retailers to be arbitrary. As discussed in the literature review in Section 1.1, two-period models are widely used in the literature on inventory allocation in one-warehouse multiple-retailer systems with demand learning, including in papers that study practical settings faced by fashion retailers (Fisher and Raman 1996, Fisher and Rajaram 2000, Gallien et al. 2017). Thus we expect the two-period setting studied in this section to be not only simpler to analyze, but to also be of practical relevance.

4.1 Inventory Allocation with Symmetric Retailer Demands

We first investigate the impact of demand learning on optimal allocation policies with symmetric retailers. We assume that the N retailers face identical demand distributions, and experience the same price p_t and holding cost h_t in time period $t = 1, 2$. We then model demand learning using the following demand forecasting model: In the first period, each retailer i experiences demand $D_{i,1}$, where $\{D_{i,1}, i = 1, \dots, N\}$ are IID. At the end of the first period, the decision maker observes uncensored demands $\{D_{i,1}, i = 1, \dots, N\}$. We assume the second-period demand follows an AR(1) time series model; that is, the demand experienced by retailer i in the second period is then given by

$$D_{i,2} = D_{i,1} + \rho\epsilon_i \text{ for some } \rho > 0 \tag{4}$$

where the demand noises $\{\epsilon_i, i = 1, \dots, N\}$ are i.i.d. with mean 0. The decision maker thus knows a component of the second-period demand beforehand (i.e., $D_{i,1}$) and can adjust inventory allocation to the retailers in the second period based on the first-period demand observations. The component of demand that is not learned is $\rho\epsilon_i$. The parameter ρ is a measure of the amount of learning, because when ρ increases, the second-period demand forecast accuracy decreases. We are interested in how the first-period allocations to the retailers depend on this parameter ρ, or, equivalently, on the extent of learning.

Due to complications from the lost sales setting and demand learning, analyzing the exact cost of inventory policies becomes prohibitive, even for two-period models (see Fisher and Raman (1996)). We thus approximate the inventory allocation problem (P1) with (P2), which relaxes the hard inventory constraint to one in expectation. (By Theorem 1, (P2) is a good approximation of (P1) when the number of retailers N is large.) We then make a further simplification that inventory cannot be carried over from the first period to the second period. We refer to this simplification as the *multiperiod newsvendor* setting (see, e.g., Ding et al. 2002, Bensoussan et al. 2007), which allows us to isolate the effect of demand learning.

The dynamic program representing this approximation is given below in (P3). Similar to the analysis in Section 3, we can show that the inventory constraint can be dualized, causing the problem to separate into N independent single-retailer inventory ordering problems.

$$\begin{aligned} \min \quad & \sum_{t=1}^{T}\sum_{i=1}^{N} \gamma^{t-1}\mathrm{E}\left[\mathrm{E}[L_{i,t}(a_{i,t})|\mathcal{H}_t]\right] \\ \text{subject to} \quad & \sum_{i=1}^{N}\sum_{t=1}^{T} \mathrm{E}[a_{i,t}] \le w_0 \\ & a_{i,t} \ge 0 \quad \forall i, t, \mathcal{H}_t. \end{aligned} \tag{P3}$$

The dual problem is given below in (D3).

$$\begin{aligned} \max \quad & -\lambda w_0 + \mathrm{SUB}_i(\lambda) \\ \text{subject to} \quad & \lambda \ge 0, \end{aligned} \tag{D3}$$

where $\mathrm{SUB}_i(\lambda)$ is the solution to the single retailer inventory ordering problem

$$\begin{aligned} \min \quad & \sum_{t=1}^{T}\sum_{i=1}^{N}\gamma^{t-1}\mathrm{E}[\mathrm{E}[L_{i,t}(a_{i,t})|\mathcal{H}_t]] \\ & + \gamma^{t-1}\sum_{t=1}^{T}\sum_{i=1}^{N}\mathrm{E}[\lambda\gamma^{1-t}a_{i,t}] \\ \text{subject to} \quad & a_{i,t} \geq 0 \quad \forall i, t, \mathcal{H}_t. \end{aligned} \tag{D4}$$

Since the retailers have identically distributed demands, the first-period allocations that optimize (D4) are by symmetry the same across all retailers. This allocation $a^*_{i,1}(\rho)$ depends on ρ, the standard deviation of the demand forecasting error in the second period. In fact, we can show that $a^*_{i,1}(\rho)$ is strictly increasing in ρ, i.e. the first-period allocations are strictly decreasing with the extent of learning. This result is stated in Theorem 2 below and is proven in Appendix A.3.

Theorem 2. *In the multiperiod newsvendor setting* (P3) *with inventory relaxation , there exists a maximum warehouse inventory level* $w_{\max}$*, such that for any* $0 \leq w_0 \leq w_{\max}$*, the optimal first-period allocation* $a^*_{i,1}(\rho)$ *to each retailer* i *is strictly increasing in* ρ*.*

An intuitive explanation of Theorem 2 is that when the decision maker is able to forecast the second-period demand more accurately (i.e., smaller ρ), she should save more of the available warehouse inventory for the second period, as she will derive more value from deploying this inventory in the second period rather than in the first period. Equivalently, this implies that with a smaller ρ, she should allocate less inventory in the first period. Theorem 2 thus suggests that demand learning has an effect that incentivizes the decision maker to reserve inventory at the warehouse and to delay inventory allocations to later periods. Note that this intuition only holds when the amount of inventory is limited; if there is sufficient inventory at the warehouse, there is no need for the decision maker to reserve inventory. This explains why Theorem 2 requires the initial inventory level to be less than some constant $w_{\max}$.

4.2 Inventory Allocation with Asymmetric Retailer Demands

We now look at a different setting where the retailers have asymmetric demand distributions. Our objective is to investigate the impact of different levels of demand uncertainty among the retailers on the inventory allocation decisions. If the mean demand varies from retailer to retailer, intuition says that it is generally optimal to allocate more inventory to retailers with higher demand means. However, the structure of the optimal policy is less clear when retailers have the same means but different variances: should the decision maker allocate more inventory to retailers with higher or lower demand variance? By allocating more inventory to retailers with *higher* demand variance, the decision maker can potentially satisfy more customer demand. On the one hand, if the decision maker allocates more inventory to retailers with *lower* demand variance, she will experience a smaller chance of misallocating inventory.

To understand the above tradeoff, we analyze a two-period model where the prices and holding costs are the same across retailers and time periods. Denote the price by p and the holding cost by h, and assume further that $p > h$. We model the demands at the retailers using (discrete) truncated normal demand distributions that are symmetrical about their means. For each retailer i, we let demand $D_{i,t}$ have mean μ and range $[\mu - b,\ \mu + b]$, i.e. the means and ranges are kept constant across retailers. Since we are interested in isolating the impact of different levels of demand uncertainties among the retailers, only the variances $\{\sigma_i, i = 1, \ldots, N\}$ corresponding to demands $\{D_{i,t}, i = 1, \ldots, N\}$ may be non-identical across retailers.

We once again analyze the allocations chosen by the Lagrangian relaxation heuristic (D1). The heuristic's first-period allocations are denoted by $\{a^*_{i,1}, i = 1, \ldots, N\}$. We find that the decision maker's allocations under the heuristic depend on the initial inventory at the warehouse. If the initial inventory level is high, then the decision maker should take higher risks and allocate more inventory to retailers with higher demand variances. However, if the initial inventory level at the warehouse is low, the decision maker should be more conservative and allocate more inventory to retailers with lower demand variances. This result is stated in Theorem 3 and proven in Appendix A.4.

Theorem 3. *There exists a positive constant $\overline{w}$ such that for any initial warehouse inventory level $w_0 \geq \overline{w}$, the heuristic's first-period allocations satisfy $a^*_{i,1} < a^*_{j,1}$ whenever $\sigma_i < \sigma_j$. The second-period order-up-to levels, denoted by $\{x^*_{i,2}, i = 1, \ldots, N\}$, also satisfy $x^*_{i,1} < x^*_{j,1}$ whenever $\sigma_i < \sigma_j$.*

There also exists a positive constant $\underline{w}$ such that for any initial warehouse inventory level $0 \leq w_0 \leq \underline{w}$, the heuristic's first-period allocations satisfy $a^*_{i,1} > a^*_{j,1}$ whenever $\sigma_i < \sigma_j$. Similarly, the second-period order-up-to levels satisfy $x^*_{i,1} > x^*_{j,1}$ whenever $\sigma_i < \sigma_j$.

5 Numerical Experiments

The structural properties in Section 4 were analyzed under several model approximations, including relaxing the inventory constraint and (sometimes) ignoring inventory carryover. In this section, we use numerical experiments to verify if these results hold for the original model. Section 5.1 presents our simulations for the symmetric demand setting, and Section 5.2 presents simulations for the asymmetric demand setting.

5.1 Demand Learning with Symmetric Retailer Demands

We first ran a set of simulations to empirically investigate the impact of demand learning on inventory allocations. We consider a two-period setting with one warehouse and two retailers. As in Section 4, we assume the demand follows the AR(1) model in Equation (4). For these simulations, we fixed the prices at \$1, the holding costs at \$0.20, and the starting warehouse inventory at $w_0 = 12$, but varied the distributions of demand at the retailers.

We simulated for two types of demand distributions. First we simulated a setting where the first-period demand is drawn from a truncated discrete normal distribution, i.e., $D_{i,1}$ is normally distributed according to $\mathcal{N}(\mu, \sigma)$ conditional on $D_{i,1}$ belonging to the interval $[a, b]$. We discretized the interval $[a, b]$ with stepsize 0.01, and forcing demand to take values from this discretized set. We then set $\sigma = 1$, varied the mean demand μ in the set $\{2, 2.5, 3, 3.5, 4\}$, and for each μ set $a = \mu - 1, b = \mu + 1$. The second-period noise is drawn from a similar truncated normal distribution from $[-1, 1]$.

Next, we also simulated a setting where each retailer i's first-period demand that is drawn from a discrete uniform distribution. For a given mean demand μ, period 1 demand $D_{i,1}$ is drawn from the set $\{\mu-1, \mu-1+0.01, \ldots, \mu+1-0.01, \mu+1\}$ with equal probability. We varied the mean demand μ within the set $\{2, 2.5, 3, 3.5, 4\}$. We use the same second-period noise as in the first setting.

For each choice of demand distribution, we solved for the optimal allocations corresponding to different values of ρ in the interval $[0, 1]$. We solved the retailers' optimal allocations as follows: First, we computed the optimal second-period allocations for each retailer corresponding to each possible tuple of the first-period allocation y, and realized demands $D_{1,1}$ and $D_{2,1}$. Then, using the property that the optimal first-period allocations must by symmetry be the same for both retailers, we recursively computed the expected cost of each possible first-period allocation from the set $\{0, 0.01, 0.02, \ldots, 6\}$, and selected the allocation minimizing this cost. We calculated the optimal allocations for both the original lost-sales model and the multiperiod newsvendor setting.

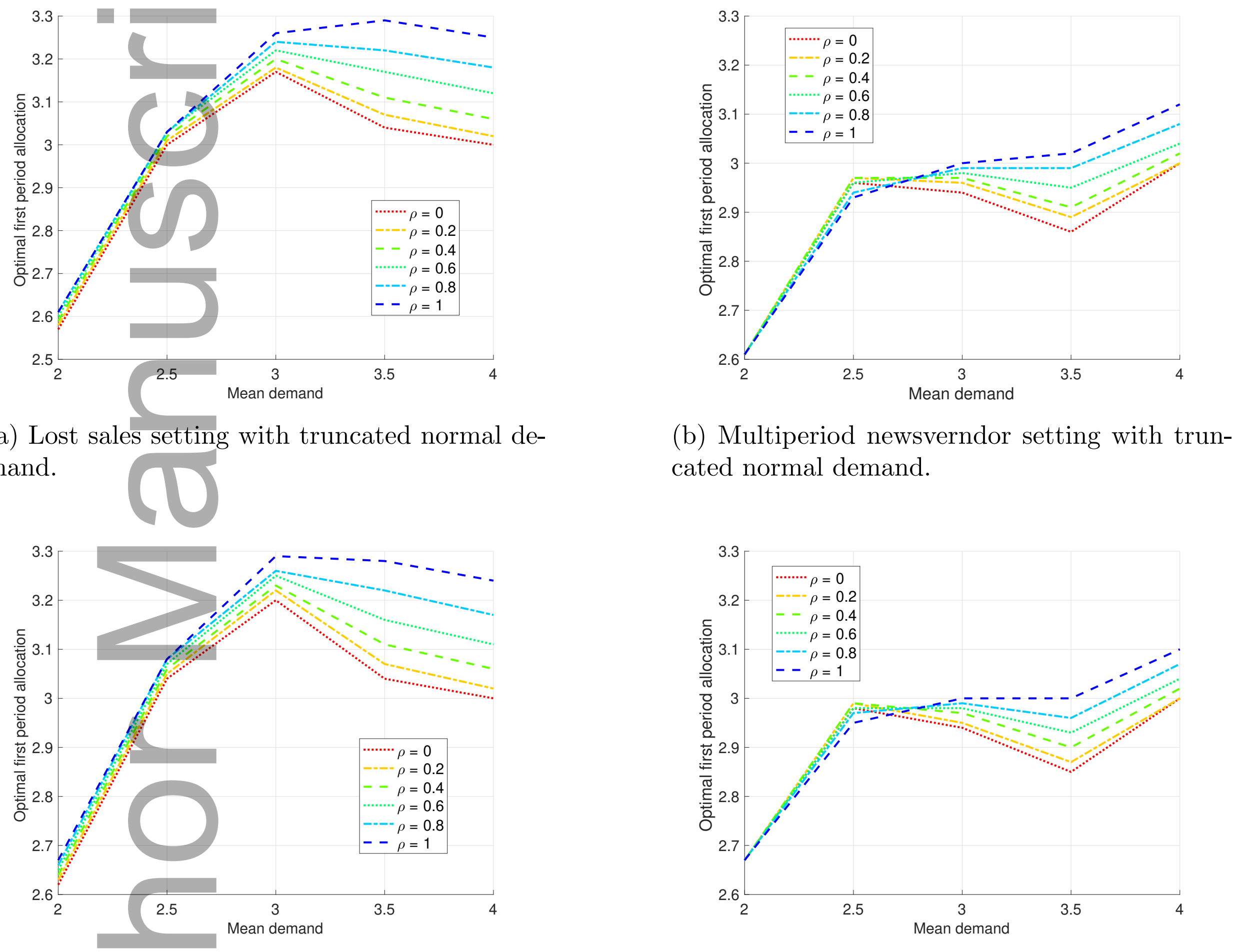
The results of these simulations are given in Figures 1a-1d. Figures 1a and 1b correspond to the setting with truncated normal demand. In Figure 1a, we plot the optimal allocations for the lost sales setting for ρ in $\{0, 0.2, 0.4, 0.6, 0.8, 1.0\}$, and in Figure 1b we plot the optimal allocations for the multiperiod newsvendor setting (i.e., no inventory carryover) for these values of ρ. The results for uniformly distributed demand are plotted in Figures 1c and 1d.

For all these settings, we see that the optimal first-period allocations are indeed increasing in ρ. This agrees with our finding in Theorem 2 that the first-period allocations are decreasing as the extent of learning increases, which is in turn consistent with the property that demand learning, like the risk pooling effect, incentivizes reserving inventory at the warehouse for later periods. However, unlike Theorem 2, which makes several approximations of the inventory allocation problem (P1) and analyzes the allocation decisions corresponding to the approximate optimization problem, Figures 1a and 1c give the allocations that exactly solve the inventory allocation problem in its original form. These results thus suggest that the demand learning effect is a property of our original allocation problem, and not just of the approximate problem studied in Theorem 2.

Finally, it is also interesting to note that our plots of the exact optimal solutions in Figures 1a and 1c show that for our demand forecasting model, the optimal first-period allocations are *not* necessarily monotonically increasing in the mean demand at the retailers. This is contrary to what we would observe if the demands at each retailer were i.i.d. across time periods, and indeed, this effect becomes less pronounced as ρ increases from 0 to 1 (i.e. as the periods 1 and 2 demands become less strongly correlated). Although this result may seem counterintuitive, we interpret it as being related to the property that demand learning can incentivize saving inventory for the second period: As the demand means increase, and the warehouse inventory becomes more limited with respect to demand, the decision maker derives more value from deploying this limited inventory in the second period, when an improved demand forecast is available.

5.2 Non-identical Retailers with Different Levels of Demand Uncertainty

We also conducted a second set of simulations on a setting with non-identical retailers in order to verify Theorem 3. For these simulations, we considered a set up with three retailers, all of whom experience demand drawn from truncated normal distributions with parameters μ, σ_i, a, b,



(a) Lost sales setting with truncated normal demand.

(b) Multiperiod newsverndor setting with truncated normal demand.

(c) Lost sales setting with uniformly distributed demand.

(d) Multiperiod newsvendor setting with uniformly distributed demand.

Figure 1: Optimal first-period allocations with truncated normal and uniform demand distribution.

i.e. retailer i's demands $D_{i,t}$ are normally distributed according to $\mathcal{N}(\mu, \sigma^2)$, conditional on $D_{i,t}$ belonging to the range $[a, b]$. As in the simulations on demand learning, we discretized these truncated normal distributions by discretizing the interval $[a, b]$ with stepsize 0.01, and forcing demand to take values from the discretized set.

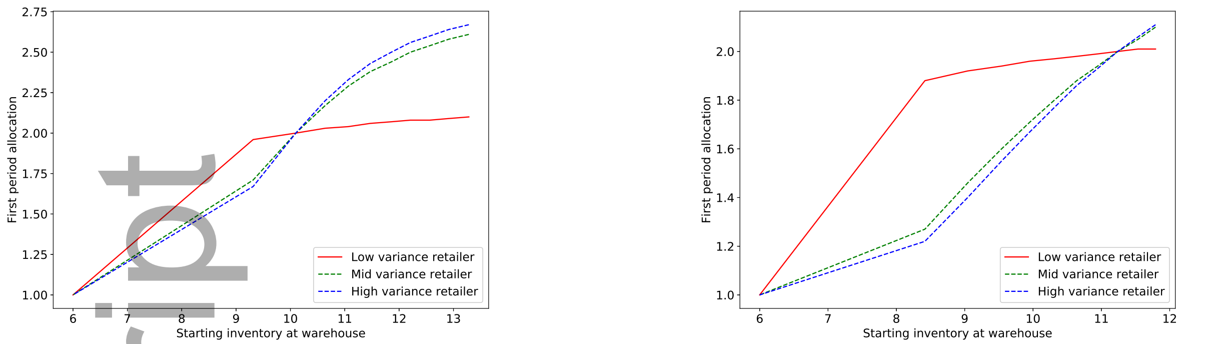
For the three different retailers, we kept the demand means and ranges fixed at $\mu = 2, a = 1, b = 3$, but varied the demand variances. For Retailer 1, or the 'low variance retailer', we set $\sigma = 0.1$. For Retailer 2, or the 'mid variance retailer', we set $\sigma = 1$. Finally, for Retailer 3, or the 'high variance retailer', we set $\sigma = 5$.

We then computed the first-period heuristic allocations to the different retailers. Instead of solving for the λ that optimizes the dual problem (D1) for some given level of the starting warehouse inventory w_0, we varied λ between 0 and price p (since, as we have observed in Section 3, the optimal λ always lies within this range), and solved for the first-period allocations that optimize (D2) given this λ. This was done recursively: First, using the well known result that the optimal allocation policy for the single retailer inventory ordering problem is an order-up-to policy, we computed the second-period order-up-to levels by discretizing the interval $[1, 3]$ with a stepsize of 0.01 (i.e. for each retailer, we searched the set $[1, 1.001, 1.002, \ldots, 3]$ for the second-period newsvendor levels). We then recursively computed the optimal first-period allocation by once again discretizing the interval [1 3], and selecting the allocation from this set with the lowest expected cost. By also recursively computing the total expected allocations across all retailers and time periods, we were able to obtain the warehouse inventory w_0 corresponding to our chosen λ.

Figure 2a plots the heuristic's first-period allocations to the three retailers when prices at all retailers are set to \$1 and holding costs are set to \$0.20, and Figure 2b plots the heuristic's first-period allocations to the three retailers when prices at all retailers are set to \$1 and holding costs are set to \$0.80. For both configurations of the price and holding cost, we see that the results directly verify Theorem 3: For sufficiently small starting warehouse inventory, the first-period heuristic allocations are decreasing in the variance of the retailers. However, when the starting warehouse inventory is sufficiently large, the first-period heuristic allocations are increasing in the variance of the retailers.

6 Conclusion

We have studied a two-echelon inventory allocation problem for a lost sales, correlated demand setting. We have shown that a Lagrangian relaxation heuristic can reduce the computationally intractable two-echelon inventory allocation problem to a set of separate single retailer inventory ordering problems. Under some general assumptions on the prices and holding costs, we show that the heuristic is asymptotically optimal in the number of retailers N. Using the heuristic as a proxy for the optimal solution, we prove that under in a two-period newsvendor setting, that demand learning incentivizes the decision maker to reserve more inventory at the warehouse for later periods. Demand learning should thus have a similar effect on allocation decisions as risk pooling. Much remains to be said on the subject of demand learning in the two-echelon inventory setting. Since we have only been able to show that the optimality gap between the heuristic and the exact solution to the allocation problem is sublinear in the number of retailers, and not in the length of the selling horizon, proving a non-trivial regret upper bound in terms



(a) Heuristic allocations with prices = \$1, Holding costs = \$0.20

(b) Heuristic allocations with prices = \$1, Holding costs = \$0.80

Figure 2: First period allocations under the heuristic in a setting with 3 retailers, all off whom experience demand drawn from truncated normal distributions with parameters μ, σ_i, a, b (retailer i's demands $D_{i,t}$ are normally distributed according to $\mathcal{N}(\mu, \sigma^2)$, conditional on $D_{i,t}$ belonging to the range $[a\ b]$.) For Retailer 1, the 'low variance retailer', $\mu = 2, \sigma = 0.1, a = 1, b = 3$. For Retailer 2, the 'mid variance retailer', $\mu = 2, \sigma = 1, a = 1, b = 3$, and for Retailer 3, the 'high variance retailer', $\mu = 2, \sigma = 5, a = 1, b = 3$.

of the length of the selling horizon would be difficult. We leave such questions to future work.

References

Adelman, D. and Mersereau, A. J. (2008). Relaxations of weakly coupled stochastic dynamic programs. *Operations Research*, 56(3):712–727.

Agrawal, N. and Smith, S. A. (2013). Optimal inventory management for a retail chain with diverse store demands. *European Journal of Operational Research*, 225(3):393–403.

Aviv, Y. (2003). A time-series framework for supply-chain inventory management. *Operations Research*, 51(2):210–227.

Azoury, K. S. (1985). Bayes solution to dynamic inventory models under unknown demand distribution. *Management science*, 31(9):1150–1160.

Azoury, K. S. (1988). *Bayesian Policies for Dynamic Inventory Models*. PhD thesis, UCLA.

Azoury, K. S. and Miller, B. L. (1984). A comparison of the optimal ordering levels of Bayesian and non-Bayesian inventory models. *Operations Research*, 30(8):993–1003.

Bensoussan, A., Çakanyıldırım, M., and Sethi, S. P. (2007). A multiperiod newsvendor problem with partially observed demand. *Mathematics of Operations Research*, 32(2):322–344.

Chen, L. (2010). Bounds and heuristics for optimal Bayesian inventory control with unobserved lost sales. *Operations research*, 58(2):396–413.

Chen, L. and Plambeck, E. L. (2008). Dynamic inventory management with learning about the demand distribution and substitution probability. *Manufacturing & Service Operations Management*, 10(2):236–256.

Cheung, K. L. and Lee, H. L. (2002). The inventory benefit of shipment coordination and stock rebalancing in a supply chain. *Management science*, 48(2):300–306.

Clark, A. J. and Scarf, H. (1960). Optimal policies for a multi-echelon inventory problem. *Management Science*, 6(4):475–490.

Ding, X., Puterman, M., and Bisi, A. (2002). The censored newsvendor and optimal acquisition of information. *Operations Research*, 50(3):517–527.

Eppen, G. and Iyer, A. (1997). Improved fashion buying with Bayesian updates. *Operations research*, 45(6):805–819.

Federgruen, A. and Zipkin, P. (1984). Approximations of dynamic, multilocation production and inventory problems. *Management Science*, 30(1):69–84.

Fisher, M. and Rajaram, K. (2000). Accurate retail testing of fashion merchandise: Methodology and application. *Management Science*, 19(3):266–278.

Fisher, M. and Raman, A. (1996). Reducing the cost of demand uncertainty through accurate response to early sales. *Operations Research*, 44(1):87–99.

Gallien, J., Mersereau, A. J., Garro, A., Mora, A. D., and Vidal, M. N. (2017). Initial shipment decisions for new products at zara. *Operations Research*, 63(2):269–286.

Gilbert, K. (2005). An arima supply chain model. *Management Science*, 51(2):305–310.

Iglehart, D. L. (1964). The dynamic inventory problem with unknown demand distribution. *Management Science*, 10(3):429–440.

Jackson, P. L. (1988). Stock allocation in a two-echelon distribution system or "what to do until your ship comes in". *Management Science*, 34(7):1–17.

Jackson, P. L. and Muckstadt, J. A. (1989). Risk pooling in a two-period, two-echelon inventory stocking and allocation problem. *Naval Research Logistics*, 36:1–26.

Jain, A., Rudi, N., and Wang, T. (2014). Demand estimation and ordering under censoring: Stock-out timing is (almost) all you need. *Operations Research*, 63(1):134–150.

Lariviere, M. A. and Porteus, E. L. (1999). Stalking information: Bayesian inventory management with unobserved lost sales. *Management Science*, 45(3):346–363.

Lee, H. L. (1996). Effective inventory and service management through product and process redesign. *Operations Research*, 44(1):151–159.

Lee, H. L. and Tang, C. S. (1997). Modelling the costs and benefits of delayed product differentiation. *Management science*, 43(1):40–53.

Lovejoy, W. S. (1990). Myopic policies for some inventory models with uncertain demand distributions. *Management Science*, 36(6):724–738.

Lu, X., Song, J.-S., and Zhu, K. (2008). Analysis of perishable-inventory systems with censored demand data. *Operations Research*, 56(4):1034–1038.

Marklund, J. M. and Rosling, K. (2012). Lower bounds and heuristics for supply chain stock allocation. *Operations Research*, 60(1):iii–248.

McGavin, E. J., Schwarz, L. B., and Ward, J. E. (1993). Two-interval inventory-allocation policies in a one-warehouse n-identical-retailer distribution system. *Management Science*, 39(9):1092–1107.

McGavin, E. J., Schwarz, L. B., and Ward, J. E. (1997). Balancing retailer inventories. *Operations Research*, 45(6):780–989.

Miller, B. L. (1986). Scarf's state reduction method, flexibility, and a dependent demand inventory model. *Operations Research*, 34(1):83–90.

Scarf, H. E. (1960). Some remarks on bayes solutions to the inventory problem. *Naval Research Logistics Quarterly*, 7(4):591–596.

Tan, F. K. (1974). Optimal policies for a multi-echelon inventory problem with periodic ordering. *Management Science*, 20(7):1104–1111.

A Appendix

A.1 Proof of Lemma 1

Proof. Since we have assumed that the demands $D_{i,t}$ are discrete and bounded for all i, t, there are finitely many scenarios for the stochastic program (P2). Defining a different allocation decision $a_{i,t}$ for each scenario in $\mathcal{H}_t$ (recall that the allocation must be non-anticipative), the *deterministic equivalent* of (P2) is a deterministic optimization problem with finitely many variables and constraints. The constraints are affine in $a_{i,t}$ and therefore convex. It remains to show that the objective function of (P2) is convex. We will show that this is true given the assumption $p_{i,t} + h_{i,t} \geq \gamma p_{i,t+1}$.

Given a demand realization $D_{i,t}$, define $\tilde{L}_{i,t}(a_{i,t}) := \gamma^{t-1} h_{i,t}[a_{i,t}+x_{i,t}-D_{i,t}]^+ + \gamma^{t-1} p_{i,t}[D_{i,t} - a_{i,t} - x_{i,t}]^+$ and define $\tilde{L}_{i,0}(a_{i,0}) := 0$ for each i. We will prove that $\sum_{t=1}^{T} \tilde{L}_{i,t}$ is convex, which is sufficient to prove the theorem. First, we will expand the expression for $\tilde{L}_{i,t}(a_{i,t})$ as follows:

$$\begin{aligned}
\tilde{L}_{i,t}(a_{i,t}) &= \gamma^{t-1} h_{i,t}[a_{i,t} + x_{i,t} - D_{i,t}]^+ + \gamma^{t-1} p_{i,t}[D_{i,t} - a_{i,t} - x_{i,t}]^+ \\
&= \gamma^{t-1} h_{i,t} \max\{0, \sum_{s=u}^{t} a_{i,s} - D_{i,s},\ \forall 1 \leq u \leq t\} \\
&\quad + \gamma^{t-1} p_{i,t} \max\{0, D_{i,t} - a_{i,t} - \max\{0, \sum_{s=u}^{t-1} a_{i,s} - D_{i,s},\ \forall 1 \leq u \leq t-1\}\} \\
&= \gamma^{t-1} h_{i,t} \max\{0, \sum_{s=u}^{t} a_{i,s} - D_{i,s},\ \forall 1 \leq u \leq t\} \\
&\quad + \gamma^{t-1} p_{i,t} \max\{D_{i,t} - a_{i,t}, \max\{0, \sum_{s=u}^{t-1} a_{i,s} - D_{i,s},\ \forall 1 \leq u \leq t-1\}\} && (5) \\
&\quad - \gamma^{t-1} p_{i,t} \max\{0, \sum_{s=u}^{t-1} a_{i,s} - D_{i,s},\ \forall 1 \leq u \leq t-1\}. && (6)
\end{aligned}$$

Note that the sum of the two addends in (5) is convex, since the pointwise maximum of convex functions is convex, and $\gamma, h_{i,t}, p_t \geq 0$. We will now show by induction that

$$-\gamma^{t-1} p_{i,t} \max\{0, \sum_{s=u}^{t-1} a_{i,s} - D_{i,s},\ \forall 1 \leq u \leq t-1\} + \sum_{s=1}^{t-1} \tilde{L}_{i,s}(a_{i,s}) \qquad (7)$$

is convex. This will prove that $\sum_{t=1}^{T} \tilde{L}_{i,t}$, as the sum of convex functions, is convex in $\{a_{i,s}, s = 1, \ldots, t-1\}$. For $t = 0$, (7) is 0 and therefore clearly convex. Suppose now that (7) is convex for some $1 \leq t < T$. We will show that it must be convex for $t+1$ as well. By using the same expansion in (5) we have

$$
\begin{aligned}
&-\gamma^t p_{i,t+1}\max\{0,\sum_{s=u}^{t} a_{i,s}-D_{i,s},\ \forall 1\le u\le t\}+\sum_{s=1}^{t}\tilde{L}_{i,s}(a_{i,s})\\
&=-\gamma^t p_{i,t+1}\max\{0,\sum_{s=u}^{t} a_{i,s}-D_{i,s},\ \forall 1\le u\le t\}+\gamma^{t-1}h_{i,t}\max\{0,\sum_{s=u}^{t-1} a_{i,s}-D_{i,s},\ \forall 1\le u\le t\}\\
&+\gamma^{t-1}p_{i,t}\max\{D_{i,t}-a_{i,t},\max\{0,\sum_{s=u}^{t-1} a_{i,s}-D_{i,s},\ \forall 1\le u\le t-1\}\}\\
&-\gamma^{t-1}p_{i,t}\max\{0,\sum_{s=u}^{t-1} a_{i,s}-D_{i,s},\ \forall 1\le u\le t-1\}\}+\sum_{s=1}^{t-1}\tilde{L}_{i,s}(a_{i,s}).
\end{aligned}
$$

By the induction hypothesis, the term in the final line is convex. The sum of the remaining terms is equal to

$$
\begin{aligned}
&-\gamma^t p_{i,t+1}\max\{0,a_{i,t}-D_{i,t},a_{i,t}-D_{i,t}+\sum_{s=u}^{t-1} a_{i,s}-D_{i,s},\ \forall 1\le u\le t\}\\
&+\gamma^{t-1}h_{i,t}\max\{0,\sum_{s=u}^{t-1} a_{i,s}-D_{i,s},1\le u\le t\}\\
&+\gamma^{t-1}p_{i,t}\max\{0,D_{i,t}-a_{i,t},\max\{0,\sum_{s=u}^{t-1} a_{i,s}-D_{i,s},\ \forall 1\le u\le t-1\}\\
&=\gamma^{t-1}(h_{i,t}-\gamma p_{i,t+1})(a_{i,t}-D_{i,t})\\
&+\gamma^{t-1}(h_{i,t}-\gamma p_{i,t+1},+p_{i,t})\max\{0,D_{i,t}-a_{i,t},\max\{\sum_{s=u}^{t-1} a_{i,s}-D_{i,s},\ \forall 1\le u\le t-1\}\}.
\end{aligned}
$$

The first addend in the final line is linear and therefore convex. The term $\max\{0,\dots,\}$ in the second addend is also convex as it is the pointwise maximum of convex functions. Finally, since we have assumed $h_{i,t}+p_{i,t}\ge\gamma p_{i,t+1}$, the second addend in the final line is convex, proving the induction hypothesis and completing the proof that (P2) is a convex optimization problem. Finally, this problem clearly satisfies Slater's condition (e.g., setting $a_{i,t}=w_0/(2NT)$), so strong duality holds. □

A.2 Proof of Theorem 1

Proof. Let LB denote the value of the dual (D1). Since the Lagrangian relaxation heuristic allocates inventory according to the solution of (D1), $a^*_{i,t}(\lambda^*)$, until the warehouse runs out of inventory, UB is only different from LB when $\sum_{i=1}^{N}\sum_{t=1}^{T}a_{i,t}>w_0$. If the warehouse runs out of inventory, the heuristic incurs additional lost sales costs not accounted for in the dual problem. Denoting $p_{\max}:=\max\{p_{i,t},i=1,\dots,N,t=1,\dots,T\}$, then each unit of lost sales accounts for at most $p_{\max}$. So we have

$$\begin{aligned}\mathsf{UB} &\leq \mathsf{LB} + p_{\max}\mathrm{E}\left[[\sum_{i=1}^{N}\sum_{t=1}^{T} a^*_{i,t}(\lambda^*) - w_0]^+\right] \\ &\overset{(*)}{\leq} \mathsf{LB} + p_{\max}\mathrm{E}\left[[\sum_{i=1}^{N}\sum_{t=1}^{T} a^*_{i,t}(\lambda^*) - \mathrm{E}[\sum_{i=1}^{N}\sum_{t=1}^{T} a^*_{i,t}(\lambda^*)]]^+\right] \\ &\leq \mathsf{LB} + p_{\max}\mathrm{E}\left[|\sum_{i=1}^{N}\sum_{t=1}^{T} a^*_{i,t}(\lambda^*) - \mathrm{E}[\sum_{i=1}^{N}\sum_{t=1}^{T} a^*_{i,t}(\lambda^*)]|\right] \\ &\leq \mathsf{LB} + p_{\max}\sqrt{\mathrm{Var}\left[\sum_{i=1}^{N}\sum_{t=1}^{T} a^*_{i,t}(\lambda^*)\right]}. \end{aligned} \tag{8}$$

The step $(*)$ above follows because we have strong duality between (P2) and its dual (D1), implying that the primal feasibility constraint

$$\mathrm{E}[\sum_{t=1}^{T}\sum_{i=1}^{N} a^*_{i,t}(\lambda^*)] \leq w_0$$

is satisfied for the optimal allocation policy $\{a^*_{i,t}(\lambda^*), i = 1, \ldots, N, t = 1, \ldots, T\}$. The final line follows from Jensen's inequality, which applies because of the convexity of the square function.

We will now bound $\mathrm{Var}[\sum_{i=1}^{N}\sum_{t=1}^{T} a^*_{i,t}(\lambda^*)]$. (We will do so using a different argument from Marklund and Rosling (2012), since the latter assumes that demand is independent from time period to time period. We show that even without this assumption, a bound of the same order in terms of N can be achieved, and that our bound is in fact an improvement in terms of T.) We have

$$\begin{aligned}\mathrm{Var}\left[\sum_{i=1}^{N}\sum_{t=1}^{T} a^*_{i,t}(\lambda^*)\right] &\overset{(*)}{=} \sum_{i=1}^{N}\mathrm{Var}\left[\sum_{t=1}^{T} a^*_{i,t}(\lambda^*)\right] \\ &\overset{(**)}{\leq} \sum_{i=1}^{N}\frac{1}{4}D^2_{\max}T^2 \\ &= \frac{1}{4}D^2_{\max}NT^2 .\end{aligned}$$

The equality $(*)$ follows from the assumption that different retailers experience independent demands. Since the allocations $a^*_{i,t}$ are solutions to a decoupled optimization problem, they are independent. The inequality $(**)$ follows from Popoviciu's inequality on variances, using the assumption that $D_{i,t}$ is upper bounded by $D_{\max}$ for each i, t. Then the optimal $a^*_{i,t}$ must satisfy $0 \leq a^*_{i,t}(\lambda^*) \leq D_{\max}$, because if $a^*_{i,t}(\lambda^*)$ exceeds $D_{\max}$, the amount of allocation $a^*_{i,t}(\lambda^*) - D_{\max}$ can be deferred to the next period, resulting in strictly lower cost. The gap between UB and LB in (8) is thus $O(T\sqrt{N})$, proving the theorem.

□

A.3 Proof of Theorem 2

Proof. Given a dual variable λ, the optimal second-period allocation to each retailer is the solution to the minimization problem

$$\min_{a_{i,2}\geq 0} \ \lambda a_{i,2} + p_2 \mathrm{E}[[D_{i,1} + \rho\epsilon_i - a_{i,2}]^+] + h_2 \mathrm{E}[[a_{i,2} - (D_{i,2} + \rho\epsilon_i)]^+] \,.$$

It is easy to see that this is a convex optimization problem, and that the optimal allocation $a^*_{i,2}$ conditional on the first-period demand is given by $a^*_{i,2} = D_{i,1} + \rho F^{-1}_{\epsilon_i}(\frac{p_2-\lambda}{p_2+h_2})$, where $F^{-1}_{\epsilon_i}$ represents the inverse CDF of the demand noise ϵ_i. The optimal first-period allocation to each retailer is then the solution to the minimization problem

$$\min_{a_{i,1}\ \geq 0} \lambda a_{i,1} + p_1 \mathrm{E}[[D_{i,1} - a_{i,1}]^+] + h_1 \mathrm{E}[[a_{i,1} - D_{i,1}]^+] \,.$$

Again, this is a convex optimization problem, and the optimal $a^*_{i,1}$ is given by the newsvendor level $a^*_{i,1} = F^{-1}_{D_{i,1}}(\frac{p_1-\lambda}{p_1+h_1})$, where $F^{-1}_{D_{i,1}}$ represents the inverse CDF of the first-period demand at retailer i.

For a given $\lambda > 0$, by complementary slackness, the total expected allocation across retailers and time periods is equal to the initial warehouse inventory, namely

$$\sum_{i=1}^{N} \left(F^{-1}_{D_{i,1}}(\frac{p_1-\lambda}{p_1+h_1}) + \mathrm{E}[D_{i,1}] + \rho F^{-1}_{\epsilon_i}(\frac{p_2-\lambda}{p_2+h_2}) \right) = w_0 \ .$$

Because we assumed that demand for retailers is symmetric, for each retailer i, we have

$$F^{-1}_{D_{i,1}}\left(\frac{p_1-\lambda}{p_1+h_1}\right) + \mathrm{E}[D_{i,1}] + \rho F^{-1}_{\epsilon_i}\left(\frac{p_2-\lambda}{p_2+h_2}\right) = \frac{w_0}{N}.$$

For each ρ, denote the corresponding dual variable λ that solves the above equation be $\lambda^*(\rho)$. Choose a positive constant $w_{\max}$ such that for any $w_0 \leq w_{\max}$, the solution $\lambda^*(\rho)$ satisfies

$$F^{-1}_{\epsilon_i}\left(\frac{p_2-\lambda^*(\rho)}{p_2+h_2}\right) < 0 \ .$$

Set $w_{\max} = N\mathrm{E}[D_{i,1}]$.

(Since ϵ_i has median 0, $\lambda^*(\rho)$ must be sufficiently large that $(p_2 - \lambda^*(\rho))/(p_2+h_2) < 1/2$). Now for a fixed ρ, suppose we increase ρ by some $\Delta\rho > 0$. For any λ such that $\lambda \geq \lambda^*(\rho)$

$$F^{-1}_{D_{i,1}}\left(\frac{p_1-\lambda}{p_1+h_1}\right) \leq F^{-1}_{D_{i,1}}\left(\frac{p_1-\lambda^*(\rho)}{p_1+h_1}\right) \ ,$$

and

$$0 > \rho F^{-1}_{\epsilon_i}\left(\frac{p_2-\lambda^*(\rho)}{p_2+h_2}\right) > (\rho+\Delta\rho) F^{-1}_{\epsilon_i}\left(\frac{p_2-\lambda}{p_2+h_2}\right) \ ,$$

where both inequalities follow from the fact that when $\lambda \geq \lambda^*(\rho)$, we have $(p-\lambda)/(p+h) < 1/2$. Thus

$$F^{-1}_{D_{i,1}}\left(\frac{p_1-\lambda}{p_1+h_1}\right) + \mathrm{E}[D_{i,1}] + (\rho+\Delta\rho) F^{-1}_{\epsilon_i}\left(\frac{p_2-\lambda}{p_2+h_2}\right) \neq \frac{w_0}{N} \ ,$$

and we must have $\lambda^*(\rho) > \lambda^*(\rho+\Delta\rho)$. Since the optimal first-period allocation $a^*_{i,1}(\rho)$ is given by $a^*_{i,1}(\rho) = F^{-1}_{D_{i,1}}(\frac{p_1-\lambda^*(\rho)}{p_1+h_1})$, it is hence strictly increasing in ρ. □

A.4 Proof of Theorem 3

Proof. For a given starting warehouse inventory level w_0 in (P1), let the associated dual variable of (D1) be λ. Each retailer i's subproblem (D2) is

$$\min_{a_{i,2}\geq 0} \lambda a_{i,2} + p\mathrm{E}[[D_{i,2} - a_{i,2} - [a_{i,1} - D_{i,1}]^+]^+] + h\mathrm{E}[[a_{i,2} + [a_{i,1} - D_{i,1}]^+ - D_{i,2}]^+] \,.$$

If we reformulate this problem as

$$\begin{aligned}\min_{a_{i,2}\geq 0} \; & \lambda a_{i,2} + p\mathrm{E}[[D_{i,1} - a_{i,1}]^+] + p\mathrm{E}[[D_{i,2} - a_{i,2} - [a_{i,1} - D_{i,1}]^+]^+] \\ & + h\mathrm{E}[[a_{i,2} + [a_{i,1} - D_{i,1}]^+ - D_{i,2}]^+] \,,\end{aligned} \tag{9}$$

we can see that this reformulated optimization problem is jointly convex in $a_{i,1}$ and $a_{i,2}$. Indeed, the first term in (9) is linear in $a_{i,2}$, the last term in (9) (associated with holding costs) is convex since it is the composition of a convex increasing function and a convex function, and the sum of the second and third terms (associated with lost sales) can be expanded as

$$\begin{aligned} V_2(a_{i,1}, D_{i,1}) &= p\mathrm{E}[[D_{i,1} - a_{i,1}]^+] + p\mathrm{E}[[D_{i,2} - a_{i,2} - [a_{i,1} - D_{i,1}]^+]^+] \\ &= p\mathrm{E}[[D_{i,1} - a_{i,1}]^+] + p\mathrm{E}[\max\{D_{i,2} - a_{i,2}, [a_{i,1} - D_{i,1}]^+\}] - p\mathrm{E}[[a_{i,1} - D_{i,1}]^+] \\ &= p\mathrm{E}[\max\{D_{i,2} - a_{i,2}, a_{i,1} - D_{i,1}, 0\}] - (D_{i,1} - a_{i,1}) \,. \end{aligned}$$

This is the sum of the pointwise maximum of affine functions, which is convex, and an affine function. Thus the reformulated optimization problem (9) is jointly convex in $a_{i,1}$ and $a_{i,2}$.

We can then differentiate the objective function with respect to $a_{i,2}$, and get the first order condition

$$\begin{aligned} (h+\lambda)\mathbb{P}[D_{i,2} - [a_{i,1} - D_{i,1}]^+ \leq a_{i,2}] &= (p-\lambda)\mathbb{P}[D_{i,2} - [a_{i,1} - D_{i,1}]^+ > a_{i,2}] \\ a^*_{i,2} &= \max\{F_i^{-1}(\frac{p-\lambda}{p+h}), [a_{i,1} - D_{i,1}]^+\} - [a_{i,1} - D_{i,1}]^+ \,, \end{aligned} \tag{10}$$

where F_i denotes the CDF of retailer i's demand, and F_i^{-1} denotes the inverse CDF of this demand. Then by the convexity of the period 2 allocation problem in $a_{i,2}$ (convexity in $a_{i,1}$ will be used later), the period 2 optimal order up to level is thus $F_i^{-1}((p-\lambda)/(p+h))$.

The first-period allocation is then the solution to the optimization problem

$$\min_{a_{i,1}} \lambda a_{i,1} + h\mathrm{E}[[a_{i,1} - D_{i,1}]^+] + V_2(a_{i,1}, D_{i,1}) \,.$$

Using the well known property that $\pi(x) = \min_{g(y,x)\leq 0} f(y,x)$ is convex given that f, g are jointly convex in x, y, we know that V_2 is convex in $a_{i,1}$. Then the objective function of the first-period optimization problem, as the sum of convex functions, is convex in $a_{i,1}$, and any solution of the first order conditions will be a global optimum. Using the expression for the period 2 optimal order up to level, $F_i^{-1}((p-\lambda)/(p+h))$, we can expand V_2 and write

$$\min_{a_{i,1}} \lambda a_{i,1} + p\mathrm{E}[[D_{i,1}-a_{i,1}]^+] + h\mathrm{E}[[a_{i,1}-D_{i,1}]^+] + \lambda\mathrm{E}[[F_i^{-1}(\frac{p-\lambda}{p+h}) - [a_{i,1}-D_{i,1}]^+]^+]$$
$$+ p\mathrm{E}[[D_{i,2} - \max\{F_i^{-1}(\frac{p-\lambda}{p+h}), [a_{i,1}-D_{i,1}]^+\}]^+]$$
$$+ h\mathrm{E}[[\max\{F_i^{-1}(\frac{p-\lambda}{p+h}), [a_{i,1}-D_{i,1}]^+\} - D_{i,2}]^+]\,,$$

which is equivalently

$$\min_{a_{i,1}} \lambda a_{i,1} + p\mathrm{E}[[D_{i,1}-a_{i,1}]^+] + h\mathrm{E}[[a_{i,1}-D_{i,1}]^+] + \lambda\mathrm{E}[D_{i,2} - [a_{i,1}-D_{i,1}]^+]$$
$$+ (p-\lambda)\mathrm{E}[[D_{i,2} - \max\{F_i^{-1}(\frac{p-\lambda}{p+h}), [a_{i,1}-D_{i,1}]^+\}]^+]$$
$$+ (h+\lambda)\mathrm{E}[[\max\{F_i^{-1}(\frac{p-\lambda}{p+h}), [a_{i,1}-D_{i,1}]^+\} - D_{i,2}]^+]\,. \tag{11}$$

Now suppose $\underline{w}$ is the minimum inventory level such that for $w_{0,1} \geq \underline{w}$, $\lambda = 0$. Such a $\underline{w}$ exists since demand at all the retailers is almost surely upper bounded by the parameter b. Then if $\lambda = 0$, the second-period order-up-to level is $F_i^{-1}(p/(p+h))$. Since we have assumed that $p > h$, we have $p/(p+h) < 1/2$. Then, since demand follows the truncated normal distribution, for $i > j$, we have $\mathbb{P}[D_{i,t} > d] > \mathbb{P}[D_{j,t} > d]$ when $d > \mu$, implying that the second-period order-up-to level is increasing in i.

As for the optimal first-period allocation when $\lambda = 0$, suppose we know that $a^*_{i,1} \leq F_i^{-1}(p/(p+h))$. Then, differentiating the objective function of (11) evaluated on any $a_{i,1} \leq F_i^{-1}(p/(p+h))$ gives

$$\lambda - p\mathbb{P}[D_{0,i} > a_{0,i}] + h\mathbb{P}[D_{0,i} \leq a_{0,i}]\,.$$

Then the left derivative of (11) is 0 when $a^*_{0,i} = F_i^{-1}(p/(p+h))$. By the differentiability of (11), the derivative of (11) is also 0 when $a^*_{0,i} = F_i^{-1}(p/(p+h))$, and by the convexity of (11), the optimal first-period allocation is $F_i^{-1}(p/(p+h))$, i.e. the same as the second-period order-up-to level. This implies that the optimal first-period allocation is also increasing in i.

To analyze the case when the starting warehouse inventory $w_{0,1}$ is small, and λ is large, we will construct $\overline{w}$ by noting that demand is truncated normal, which implies that there exists some $t_{\max}$, $0 < t_{\max} < 1/2$ such that the demand probability density functions $f_i(t)$ is increasing in i for all $t \leq t_{\max}$. Then set $\overline{w}$ as the starting warehouse inventory that corresponds to dual variable $\lambda = p - (p+h)F_i(t_{\max}/2)$. For $w_{0,1}$, $w_{0,1} \leq \overline{w}$, the associated dual variable λ satisfies $\lambda \geq p - (p+h)F_i(t_{\max}/2)$.

For such $w_{0,1}$ that satisfies $w_{0,1} \leq \overline{w}$, since the associated λ satisfies $\lambda \geq p-(p+h)F_i(t_{\max}/2) \geq (p-h)/2$, we have $(p-\lambda)/(p+h) \leq 1/2$ and thus, $F_i^{-1}((p-\lambda)/(p+h))$ is strictly decreasing in i by our assumption that demand is truncated normal, which implies that for $i > j$, $\mathbb{P}[D_{i,t} > d] < \mathbb{P}[D_{j,t}] > d]$ when $d < \mu$.

We now complete the proof by showing that the optimal first-period allocation $a_{i,1}$ is also decreasing in i for $w_{0,1}$ that satisfies $w_{0,1} \leq \overline{w}$. We do so by writing the first order conditions of (11):

$$p-\lambda = (p+h-\lambda)\mathbb{P}[D_{i,1} \le a_{i,1}] + (p-\lambda)\mathbb{P}[D_{i,1} \le a_{i,1} - F_i^{-1}(\frac{p-\lambda}{p+h})] + (h-p+2\lambda)\mathbb{P}[D_{i,1}+D_{2,i} \le a_{i,1} \mid D_{i,1} \le a_{i,1} - F_i^{-1}(\frac{p-\lambda}{p+h})]\mathbb{P}[D_{i,1} \le a_{i,1} - F_i^{-1}(\frac{p-\lambda}{p+h})]\,. \tag{12}$$

We claim that $2F_i^{-1}((p-\lambda)/(p+h))$ is an upper bound on the optimal $a_{i,1}$ when $w_{0,i} \le \overline{w}$. Suppose instead that $a_{i,1} > 2F_i^{-1}((p-\lambda)/(p+h))$. Then the right hand side of (12) is at least

$$\begin{aligned}
&(p+h-\lambda)\mathbb{P}[D_{i,1} \le 2F_i^{-1}(\frac{p-\lambda}{p+h})] \\
&+ (p-\lambda)\mathbb{P}[D_{i,1} \le F_i^{-1}(\frac{p-\lambda}{p+h})] + (h-p+2\lambda)\mathbb{P}[D_{i,1} \le F_i^{-1}(\frac{p-\lambda}{p+h})] \\
\le& 2(p+h-\lambda)\mathbb{P}[D_{i,1} \le F_i^{-1}(\frac{p-\lambda}{p+h})] + (p-\lambda)\mathbb{P}[D_{i,1} \le F_i^{-1}(\frac{p-\lambda}{p+h})] \\
&+ (h-p+2\lambda)\mathbb{P}[D_{i,1} \le F_i^{-1}(\frac{p-\lambda}{p+h})] \\
=& \mathbb{P}[D_{i,1} \le F_i^{-1}(\frac{p-\lambda}{p+h})](2p+3h-\lambda) \\
=& \frac{p-\lambda}{p+h}(p+h+p+2h-\lambda) \\
>& p-\lambda.
\end{aligned}$$

The last line follows from the fact that $w_{0,i} \le \overline{w}$ implies that the associated dual variable λ satisfies $\lambda > (p-h)/2$, and the third line follows from our assumption of truncated normal demand that is symmetrical about the mean, which implies that

$$\begin{aligned}
\mathbb{P}[D_{i,1} \le 2F_i^{-1}(\frac{p-\lambda}{p+h})] &= \mathbb{P}[D_{i,1} \le F_i^{-1}(\frac{p-\lambda}{p+h})] + \mathbb{P}[F_i^{-1}(\frac{p-\lambda}{p+h}) \le D_{i,1} \le 2F_i^{-1}(\frac{p-\lambda}{p+h})] \\
&\le 2\mathbb{P}[D_{i,1} \le F_i^{-1}(\frac{p-\lambda}{p+h})]
\end{aligned}$$

when $(p-\lambda)/(p+h) < 1/2$. Thus the right hand side of (12) is strictly greater than the left hand side, leading to a contradiction.

For each $a_{i,1}$ such that $a_{i,1} < 2F_i^{-1}((p-\lambda)/(p+h))$, we will show that each of the summands on the right hand side of (12) is strictly increasing in both $a_{i,1}$ and i. This will imply that the heuristic first-period allocations $a^*_{i,1}$ are strictly decreasing in i, proving the theorem.

Consider the first summand of (12). Since $\lambda \le p$, this term is strictly increasing in $a_{i,1}$. Further, since $a_{i,1} \le 2F_i^{-1}((p-\lambda)/(p+h)) \le t_{\max}$, by our definition of $t_{\max}$, the summand is strictly increasing in $a_{i,1}$.

Now consider the second summand of (12). Again, since $\lambda \le p$, this term is increasing in $a_{i,1}$. And since $a_{i,1} - F_i^{-1}((p-\lambda)/(p+h)) \le 2F_i^{-1}((p-\lambda)/(p+h)) \le t_{\max}$, this summand is also increasing in $a_{i,1}$.

Finally, consider the third summand of (12). We have set $w_{0,i}$ such that the associated dual variable λ satisfies $\lambda > (p-h)/2$. Thus the coefficient of the probability term, $h-p+2\lambda$, is positive, and the summand is increasing in $a_{i,1}$. We can write the probability term as

$$\mathbb{P}[D_{i,1} + D_{2,i} \leq a_{i,1} \mid D_{i,1} \leq a_{i,1} - F_i^{-1}(\frac{p-\lambda}{p+h})]\mathbb{P}[D_{i,1} \leq a_{i,1} - F_i^{-1}(\frac{p-\lambda}{p+h})]$$
$$= \int_a^{a_{i,1}} \int_a^{s-a} f_i(t) f_i(s-t) dt ds.$$

For t such that $a \leq t \leq s-a$, and s such that $s \leq a_{i,1}$, we also have $a \leq t \leq a_{i,1} \leq 2F_i^{-1}((p-\lambda)/(p+h)) \leq t_{\max}$. Thus $f_i(t)$ is increasing in i for each t within the bounds of the integral. Similarly, for each t and s within the bounds of the integral, we have $a \leq s-t \leq s-a$. By the same argument, $f_i(s-t)$ is increasing in i. Thus the probability term in the third summand of (12) is increasing in i, proving the claim and the theorem. □