Dynamic Product Allocation under Non-stationarities

Guanting Chen[†] Xiaocheng Li[◊] Zeyu Zheng[‡]

[†] Institute for Computational and Mathematical Engineering, Stanford University
[◊] Imperial College Business School
[‡]‡Department of Industrial Engineering and Operations Research, University of California, Berkeley

Abstract

In this paper, we first analyze the data set provided by JD.com, a major e-commerce company in China, and discuss the impact of a 9-day promotion event "Butterfly Festival" on the non-stationarities presented in the daily sales data. We formulate a multi-product dynamic product allocation problem under non-stationarities for Front Distribution Centers, with the objective of minimizing the expected overall lost sales, inventory holding and shipping costs. We provide myopic algorithms and identify conditions when the myopic algorithms are asymptotically optimal. When such conditions fail in general, we provide practical data-driven algorithms that can accommodate non-stationarities in the demands and available auxiliary information. Simulation experiments with real demand data show that our proposed data-driven algorithm outperforms a benchmark algorithm based on base stock policy.

1 Introduction

With the advance of the technology and logistics infrastructure, e-commerce enterprises have increasingly taken up the market share from the traditional retailers. To cater the enlarging market and growing number of customer orders, the e-commerce enterprises strive to improve their inventory and fulfillment systems with a major goal to increase customer satisfaction and retention. Companies like Amazon and JD.com (JD) have also developed self-owned logistics system, which enables the companies to offer one-day or even same-day delivery for customers. The success of fast delivery heavily relies on the distribution network infrastructure. However, based on the publicly available data provided by JD, a major e-commerce company in China, we find that there are still a large number of one-day and same-day delivery orders that are delayed. During March 2018, 56% of the same-day delivery orders are delayed. This observation motivates us to investigate possible paths to further improve the already well-built logistics infrastructure and our lens is through JD's two-level distribution system.

JD partitions the entire map into multiple regions to implement the two-level distribution system. Specifically, each region has a regional distribution center (RDC) and multiple front distribution centers (FDCs). An RDC is responsible to receive and manage inventory from vendors to satisfy the demand for that entire region. Once RDC receives inventory from vendors, RDC not only can ship products directly to customers, but also fulfills the duty of allocating inventory to FDCs in that region. Because FDCs are spread across the region, if a customer demands a particular product that has inventory in the nearest FDC, the delivery time from that FDC is generally smaller than the delivery time from the RDC. The availability of a product in FDC's inventory, therefore, is key to enhancing fast delivery guarantees. However, FDCs, as smaller facilities compared to RDC, typically have limited capacities so that it is not possible to store sufficient inventory for all products. Meanwhile, the daily shipping resources from RDC to FDCs may also encounter a constraint. This naturally leads to a research question on how to allocate the inventory in FDCs and utilize the shipping from FDC to RDC to dynamically manage the inventory in FDCs. To our knowledge obtained from public information,[1] the research team of JD has already been formulating the product allocation in FDCs as a stationary stochastic optimization problem and providing wise solutions. In this paper, we hope to provide additional analysis and discussions that may complement JD's existing efforts. Specifically, we find from the data that the daily demand and sales process exhibit significant inter-temporal non-stationarities and irregular uncertainties. Based on this observation, we investigate a dynamic allocation problem in presence of the limited FDC capacity and possibly limited daily shipping capacity with the goal of minimizing an overall cost that considers stock-outs, shipping, and inventory holding.

We investigate the sequence of daily sales quantities for JD self-owned and managed products which exhibits significant non-stationarities. We identify from public information that a 9-day promotion event called "Butterfly Festival" that happened from March 1 to March 9 can largely explain the non-stationarity presented in the sequence of daily sales quantities. The impact of this "Butterfly Festival" on the sales quantities not only exists within the 9-day promotion win-

[1] See https://medium.com/jd-technology-blog/the-secret-behind-jd-coms-super-fast-delivery-71a6c0792405

dow, but also affects the sales quantities in the rest of the month. Specifically, post-promotion, the sales quantities present a gradual increasing trend. Due to the limited size (one month) of the provided data and the lack of multiple representative promotion events (other than the "Butterfly Festival"), our focus is not on developing a tool to quantitatively analyze how a promotion event impacts the non-stationarities in the sales quantities before, during, and after promotion. Instead, we take the non-stationary sales quantities as a modeling element. As promotion events of various sizes are popular on JD, we take the view that the non-stationarities caused by promotion events and other factors should be seriously considered in dynamic product allocation for FDCs.

We first analyze a single-period model to study how the limited capacity for a single FDC impacts the product allocation decision. We assume that RDC has sufficient inventory so that our energy is focused on the capacity constraint of each single FDC. An FDC manages the inventory of a pool of products but in reality the capacity may be far from sufficient to guarantee the availability for all products. For the single-period model, the decision maker knows the current inventory level in the FDC, the total inventory capacity of the FDC, and the probability distribution of the demand for each product in the next period. The decision maker then solves a stochastic optimization problem with the objective of minimizing total expected costs of lost sales, inventory holding, and shipping. When the current inventory level does not exceed the capacity constraint, there always exists an allocation rule that minimizes the single-period cost and binds the capacity constraint. Although this stochastic optimization problem does not adopt a closed-form solution in general, we use a two-product special case to derive the closed-form solution when the demand distribution is uniform. For this special case, we show that the allocation of FDC inventory for each product should be proportional to the expectation of the demand.

We then extend the analysis to a multi-period model to study how the FDC capacity and the shipping constraint jointly impact the product allocation decision. The shipping constraint appeals to scenarios when there is a daily shipping limit from RDC to each FDC. The sequence of demand process is assumed to be stationary, and the decision maker knows the joint probability distribution of the single-period demands for all products. We establish the asymptotic optimality for a myopic policy under strong but practically implementable assumptions. Firstly, the shipping constraint has to be reasonably sufficient compared to the demand distribution. Secondly, if the shipping constraint becomes binding in some time period, the shipping budget can be enhanced to a sufficient level in the next time period.

Next, we analyze the product allocation policy under a non-stationarity environment when the demand distribution changes with time. We assume that the non-stationary demand distribution is known or can be predicted. We prove a competitive ratio for our myopic algorithm under the non-stationary environment.

Further, we relax the assumption on the knowledge of the probability distribution on the demand process. Instead, the decision maker only has access to a stream of previous demand data. We provide an online version of the myopic algorithm that utilizes the up-to-date empirical demand distribution to allocate inventory. We do not prove theoretical guarantee for the algorithm. Instead, we show through numerical experiments that the performance of the proposed online myopic policy is better than the benchmark base stock policy, saving cost up to 5%.

Finally, we consider a setting when the decision maker has additional auxiliary information that is revealed in the beginning of each time period, which can help predict the future demands. We provide a data-driven algorithm that utilizes the dynamically available auxiliary information to solve the product allocation problem under non-stationarities. The availability of auxiliary information can be exploited and carried into a reduction of expected total cost. Our simulation experiments with real data show that when the algorithms leverage the auxiliary information, the total cost can be reduced by 7%, compared to a scenario without the auxiliary information.

1.1 Literature Review

Analyzing optimal policy for inventory/production system. There is a vast literature studying the optimal policies for inventory problems. Under uncapacitated settings where there is no production/warehouse capacity constraint, a stream of research (Karlin and Scarf 1958, Veinott Jr 1965, Morton 1969) studies existence and analytical properties of (near) optimal policy with different conditions on lead time. Karlin and Scarf (1958) introduced the single-item lost-sales inventory system with a linear order cost and a positive order lead time, and the model was expanded in the other two papers listed above. The multi-product capacitated problem was analyzed by Evans (1967) and DeCroix and Arreola-Risa (1998). Evans (1967) showed that a modified base-stock policy is optimal and characterized the condition under which the capacity constraint is binding.

DeCroix and Arreola-Risa (1998) extended the result and derived the optimal policy when the production constraint is active under the condition of homogeneous products; for the case of heterogeneous products, a heuristic algorithm that allocates the order-up to level by scaling down the uncapacitized optimal base-stock level is provided. For results on single-product single-location setting we refer the reader to Federgruen and Zipkin (1986a,b), Zipkin (2008a,b) and Huh et al. (2009b). Federgruen and Zipkin (1986a,b) proved that in a stationary condition with capacity constraint and a convex one-period cost function, the modified base-stock policy is optimal for the single-item, periodic-review inventory model. Zipkin (2008a,b) tested various heuristics and provide new analytical methods for the single location single item inventory problem with lost sale. Huh et al. (2009b) considered the same lost sale problem and derive the bound between the performance of the best order-up-to policy and the optimal policy. There is also a vast literature focused on the capacitized newsvendor problem, which is structurally similar, but methodologically different from our setting. Interested readers are referred to Gallego and Moon (1993), Lau and Hing-Ling Lau (1995), Erlebacher (2000), Mieghem and Rudi (2002), Angelus and Porteus (2002) and Zhang and Du (2010).

Algorithms and heuristics for inventory/production systems: Few inventory policy can be made analytically tractable. Researchers have been developing various methods to solve this problem: approximation algorithms (Nahmias and Schmidt 1984, Federgruen and Zipkin 1984, Aviv and Federgruen 2001, Kunnumkal and Topaloglu 2008, and Levi et al. 2008), simulation methods (Glasserman and Tayur 1995, Mieghem and Rudi 2002), gradient descent based optimization and some other non-parametics statistical methods (Huh and Rusmevichientong 2009, Huh et al. 2009a, and Shi et al. 2016) have been used to develop inventory policies.

E-commerce and Data Driven Machine Learning Algorithms There is abundant literature on different aspect of optimizing the e-commerce business in the big data era. In the stationary environment, Huh and Rusmevichientong (2009) develop an adaptive inventory policy and analyze the regret; another adaptive algorithm for finding the optimal order up-to levels in lost-sales inventory system is provided in Huh et al. (2009a), and later the convergence rate is improved in Zhang et al. (2020). Under capacity constraints, Shi et al. (2016) propose a data-driven, stochastic gradient descent type algorithm for solving the multi-product inventory management problem over a finite horizon. On the demand prediction side, Li et al. (2019) and Ban and Rudin (2019) use machine learning techniques for estimation of the demand distribution. Lastly, Xu et al. (2009), Acimovic and Graves (2015), and Jasin and Sinha (2015) focus on the decision for order fulfillment in order to reduce shipping cost. Interested readers are referred to Qi et al. (2020) for a more comprehensive survey on this topic.

Notice that many theoretical results have the assumption that unsatisfied demand will be back-ordered, while there are much less literature assuming the condition of lost sale. Veinott Jr (1965) consider an uncapacitized multi-product multi-period lost sale model, and an asymptotical optimal policy is derived under stationary and non-stationary (nondecreasing) demand. Mieghem and Rudi (2002) introduce the so-called "newsvendor networks" which has analytical solution under certain conditions, and is aslo applicable to the lost sale condition. Angelus and Porteus (2002) discuss the optimal capacity and inventory plan for a single-item setting. Levi et al. (2008) introduce a 2-approximation algorithm for single-item, single location, periodic-review model with positive lead time. Huh et al. (2009a) design an adaptive algorithm for finding the optimal order up-to levels for a single-item, singe-location inventory system with lost sale. Under the same setup, Huh et al. (2009b) prove that best order up-to policy converges to the optimal policy when the lost sale penalty becomes large, and a bound is also given for any given cost parameters. Zipkin (2008a,b) provide some analytical insights and establishes new bounds on the optimal policy. Goldberg et al. (2016) and Xin and Goldberg (2016) show that the constant-order policy is asymptotically optimal when the lead time goes to infinity, and the optimality gap is also bounded. Feng et al. (2020) analyze the performance of a base stock list price policy for a inventory–pricing problem under lost sale.

Moreover, our model features a capacitized inventory system, which is more challenging to analyze. Nahmias and Schmidt (1984) proposed heuristics for one-period multiple-item inventory system subject to a single constraint. Federgruen and Zipkin (1986a) showed that a base-stock policy is optimal in a single-stage capacitated system. Glasserman and Tayur (1995) use infinitesimal perturbation analysis (IPA) for calculating optimal base-stock policies on multi-item, capacitated inventory systems. Under the condition that demand distributions, capacity levels, and cost parameters vary according to a periodic pattern, Aviv and Federgruen (1997) proved the optimality of a modified base-stock policy for the single-item periodic-review inventory model. Levi et al. (2008) introduce a 2-approximation algorithm for a single-item, single location, periodic-review model with positive lead time under capacity constraint. For the multi-item inventory systems under seasonally fluctuating demands, Aviv and Federgruen

(2001) provide a lower bound and a heuristic solution. Cheung and Simchi-Levi (2019) provide a sample complexity bound for the SAA method on a data-driven capacitized multiperiod newsvendor problem. Perakis et al. (2020) propose an iterative algorithm to solve the one-period multi-item multi-location newsvendor problem, with capacity and depth (limited inventory for each item) constraint. Capacitized problems are also harder to analyze in other fields of operation management. We refer the readers to Chen and Gallego (2018) for capacitized pricing problem and Rusmevichientong et al. (2010) and Li et al. (2018) for capacitized assortment optimization problem.

2 Data Analysis

To motivate our study, we first present an analysis on the transaction level data of JD over the month of March in 2018.

2.1 Impact of a Major 9-Day Promotion Event and Implications

We did a deep search over the publicly available information about this promotion event. The promotion event was called "Butterfly Festival" and lasted for 9 days, starting from March 1st and ending on March 9th. On March 1st, a large number of "buy-100-and-get-100-for-free" type general coupons are released that are eligible to use only within the Festival time window. On the late evening of March 6, additional special promotion rebates are released to users at a total amount around a hundred million Chinese Yuan.

The implicit presence of this major promotion event renders the sequence of daily sales quantity highly non-stationary. We find that the "Butterfly Festival" promotion event may largely explain the non-stationarity in the sequence of daily sales quantity. Figure 1 plots the daily sales quantity over March for two types of products. Type 1 products refer to those whose inventory are owned by JD. Type 2 products refer to those whose inventory are owned by some third party.

We first focus on the sequence of daily sales quantity for Type 1 products. The overall sales quantity levels from March 1 to March 9 are significantly higher than the rest of the month, which exactly matches the time window of the "Butterfly Festival" promotion. The largest daily sales quantity happens on March 1st, which corresponds to the release date of the general coupons. Because this festival was pre-announced, a portion of customers who have been waiting to use the

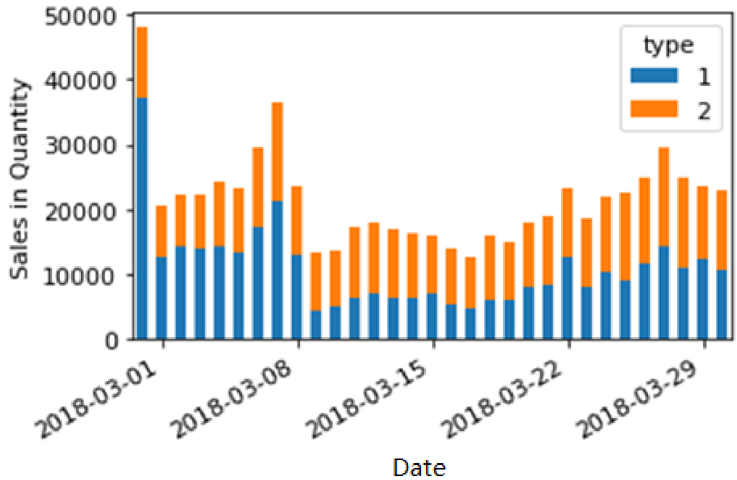


Figure 1: Sales in quantity for type 1 and type 2 products

promotion coupons will likely place their orders on the first day after the coupons are in effect. The Type 1 sales quantity on March 7 - March 8 is higher than March 2 - March 6, likely because of the additional special promotion rebates released on the late evening of March 6.

We next analyze the post-promotion daily sales quantity for Type 1 products. As shown in Figure 1, the daily sales quantity for Type 1 products drops to monthly low on March 10th, which is the exact day after the "Butterfly Festival", indicating a cool-down period after major demands are absorbed into the festival period. From March 10 to the end of March, the sequence of sales quantity presents a moderate increasing trend, likely climbing up to the nominal level with the major impact from "Butterfly Festival" washed out.

We then analyze the sequence of daily sales quantity for Type 2 products. Figure 1 surprisingly shows that the sales quantity sequence of Type 2 products is much more stationary compared to Type 1. There is even no strong statistical evidence that the festival significantly impacted the sales of Type 2 products. While we were not able to figure out a reason based on public information, we speculate the following explanation. A major promotion event typically triggers irregular fluctuations in demands and significantly impacts the coordination of supply chain and logistics systems. Because the inventory of Type 2 products are managed by third party, not by JD, it might be difficult to heavily involve the inventory of Type 2 products in a major promotion event such as the "Butterfly Festival".

The analysis above also provides the following implications on sales quantity forecasting.

- Pre-announced major promotion events should be considered as the key factor in forecasting future sales, for not only the sales during the promotion event, but also for the sales post the promotion event. A major promotion event can have a long-lasting dynamic impact on the sequence of daily sales quantity. A simple inclusion of the promotion event as a covariate for standard machine learning tools likely does not work because it may not appropriately capture the inter-temporal dynamic impact. To forecast the daily sales on a certain future date that is after a major promotion event, careful analysis on the impact of the promotion may be required to render more accurate prediction.
- A mean-reversion model can be considered as an option to model and forecast the inter-temporal dynamics of post-promotion sales.
- Self-owned products and third-party-owned products should receive different treatments in sales forecasting, especially in presence of a major promotion event.

2.2 The Delivery Characteristics for JD's Distribution System

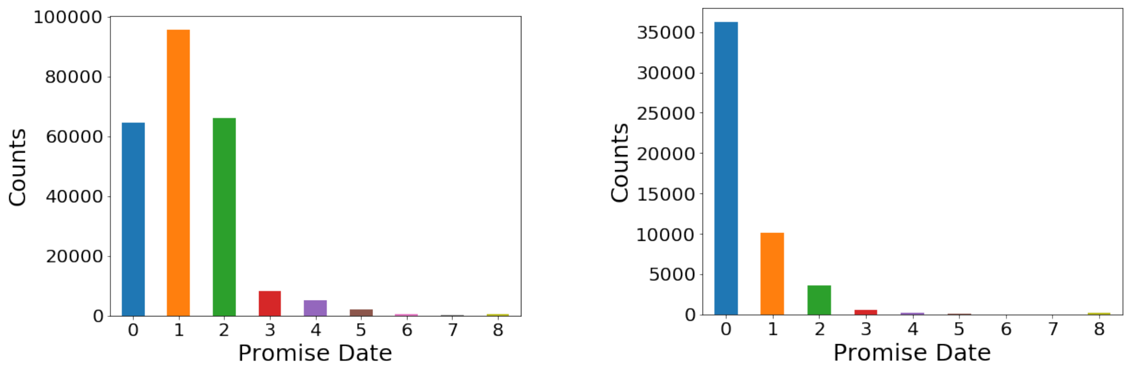
In this subsection, we study the logistics aspect of the data in the hope of identifying the bottleneck of the current outbound distribution system.

The data consists of $457,298$ completed purchases and $31,868$ SKUs in 56 warehouses. Note that these SKUs involved in the data set belong to one product category. The products of JD's logistics system are categorized as either "first party owned" (denoted as 1P) or "third-party owned" (denoted as 3P), depending on the ownership of the inventory. All 1P SKUs (3.7% of all SKUs but it contributes 50.7% of total orders) are managed directly by JD, including product assortment, inventory replenishment, product pricing, order delivery, and after-sales customer service. In general, 1P SKUs are usually top sellers within the category. By owning these 1P products, JD has control over the entire life cycle to provide guaranteed quality, fast delivery, and good customer services. In contrast, all 3P SKUs are managed by third-party merchants on the JD marketplace. In this paper, we take the standpoint of the company of JD and focus on the 1P products.

The product outbound distribution is built upon a two-level distribution center/warehouse system. For a given geographical region, the customers are grouped into multiple districts based on their location. In the bottom level, a "district warehouse" is responsible only for fulfilling orders in a specific district. In the top level, a larger warehouse, usually called the "central warehouse" or "distribution center", has the ability to ship to customers in all districts of a geographical region. A central warehouse also provides the "back-up fulfillment" option when the district warehouses within the same geographical region run out of inventory. We refer to the central warehouse as the regional distribution center (RDC), and each of those district warehouses as front distribution center (FDC). For the distribution system of JD, there are in total 7 RDCs and 49 FDCs serving customers in mainland China. The system works in the following way: if the customer orders certain product and the product is available in the nearest FDC, then the product will be shipped from the FDC to the customer. When the product is unavailable at nearest FDC, the product will then be shipped from the nearest RDC if available; however, this may results in longer shipping time and even lost sales because of the longer estimated shipping time.

A widely-advertised service of JD is the "same-day delivery" service, which enjoys great popularity among the customers. The same-day delivery guarantees that for orders placed by 11 a.m., customers will receive the package by 8 p.m. on the same day. Alternatively, the one-day or k-day delivery guarantees the customers will receive the package within one day or k days after the order placement. In fact, the same-day and one-day orders take the majority (54.5%) of all the customer orders in the data, as shown in Figure 2. For all the delivered orders in March 2018, $51,156$ out of $244,333$ (1P SKU) packages are delayed, and for the same-day delivery, 56% of 64393 orders are delayed. Furthermore, we put deliveries into two categories: local delivery and non-local delivery. Local delivery refers to the type of delivery that the item being ordered is shipped from the nearest FDC to the customer, while non-local delivery refers to that a product in the order is unavailable from the nearest FDC and has to be shipped from the RDC. We find that 88% of non-local orders are delayed; Not surprisingly, the non-local delivery also contributes to 53% of the total delayed orders with same-day or one-day promise. In contrast, the delay rate of the local delivery is 39%.

As suggested by the data, when an order is placed, if all products in the order are available in the nearest FDC, it is more likely for JD to fulfill the same-day or one-day delivery promise. Otherwise if a product in the order is not available in the nearest FDC, that product needs to be shipped from the RDC, which renders a higher chance of failing fulfillment promises. That being said, typically an FDC has a smaller size and limited space compared to a RDC. Therefore, a better utilization of the storage space of an FDC that is tailored to the local demand can improve the fulfillment rate and therefore customer satisfaction.



(a) Delivery promise histogram

(b) Delivery delay histogram

Figure 2: Promise and Delay for 1P item in JD: Items with promised delivery day 1 consists of both same-day and one-day deliver.

Furthermore, Figure 3 and 4 show that the non-local shipment may cause delays in every segment of the delivery procedure. The delivery time can be split into three parts: (i) the processing time: the time length between the order confirmation and the package shipping out of the warehouse; (ii) the inter-station time: the time length between the package shipping out of the warehouse and its arrival at the final logistic station; (iii) the last-mile delivery time: the time length between the package arrival at the final logistic station and delivery to the customer. From the figure, we observe shorter time length for all three segments in the delivery procedure for local orders shipped from the nearest FDC than non-local ones from the nearest RDC.

These observations motivate the main objective of our paper – converting the non-local shipment to local shipment via dynamic product allocation. Empirically, improving the delivery efficiency not only boots sales but also increases customer rating (Fisher et al. 2019, Bray 2020, Deshpande and Pendem 2020, Liu et al. 2020). Ideally, if there are sufficient inventory at local FDCs, then all the shipments will be local ones. In practice, however, it is neither feasible nor economic to hold inventory for all the products at local FDCs. First, it is not feasible because there are a large number of products but only a limited storage for the FDC. Second, the local districts usually have larger demand variation, so

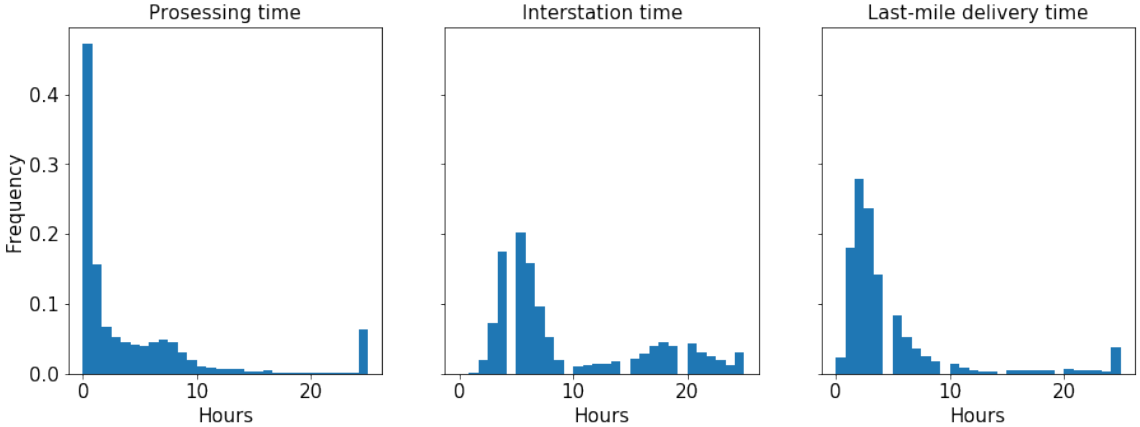


Figure 3: Time distribution for deliveries where the origin and destination is in the same region. In the histograms, the time length greater than 24 hours is denoted by 25.

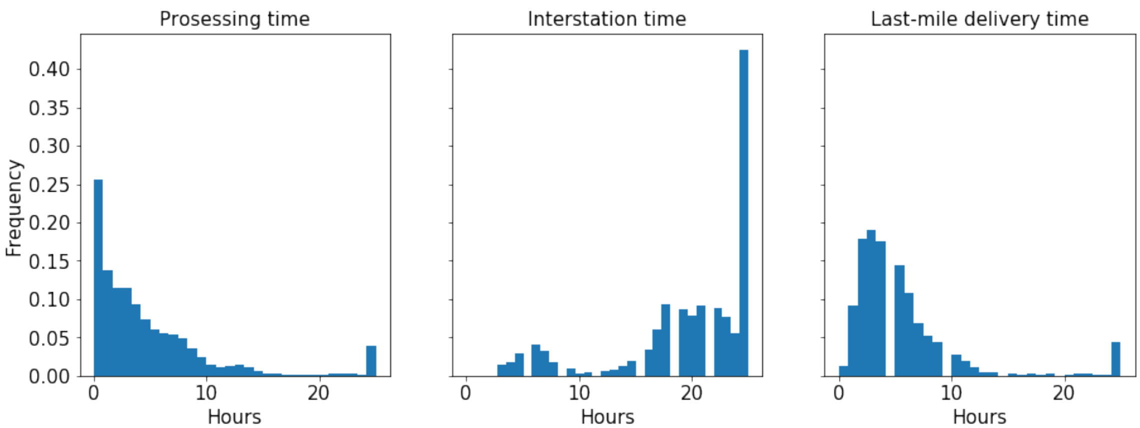


Figure 4: Time distribution for local deliveries where the origin and destination is not in the same region. In the histograms, the time length greater than 24 hours is denoted by 25.

it is more economic to hold the excessive inventory at RDC than FDC to take advantage of risk pooling. In this paper, we address the question of how to allocate the limited storage capacity at the FDCs across products and how to replenish the inventory at FDCs under limited shipping capacity between RDCs and FDCs. An ideal plan would best utilize the storage at the FDCs and maximize the number of local deliveries.

3 Product Allocation: Model and Setup

In this section, we formulate the problem of product allocation and present a few preliminary analytical results. The following assumptions are made on the logistics system.

Assumption 3.1 (Distribution System)

(a) The RDC always has sufficient supply for all the products.

(b) The lead time from RDC to FDC is no more than 1 *day.*

Assumption 3.1 (a) is verified from an empirical observation that the RDCs of JD have large enough capacity to hold sufficient inventory for all the products. Technically, the sufficient supply assumption avoids the contention for the RDC's inventory between different FDCs and it thus disentangles the one-RDC multiple-FDC distribution system to multiple one-RDC one-FDC systems. Note that the RDC-FDC distribution system largely resembles the classic model of one-warehouse multiple-retailer (OWMR) in that both the RDC and the central warehouse hold and replenish inventory for the FDCs and the retailers. In the related literature, the sufficient supply assumption on RDC has also been imposed for analytical tractability (See Veinott Jr 1965, Karmarkar 1981, Cachon 2001, Cheung and Lee 2002, Gürbüz et al. 2007, Huh et al. 2009b). As to Assumption 3.1 (b), it is achieved through the geographical design of the RDC-FDC system. Each RDC is usually in charge of a pre-specified region such that the transshipment from the RDC to every FDC within the region can be done within one day.

Consider a one-RDC one-FDC system that manages the inventory for n products. Let $d_{i,t}$ denote the realized demand of the i-th product at time $t \in [T]$ and denote its distribution as $\mathcal{P}_{i,t}$. We allow demand correlation between different products and different dates; as we will see, it is only the marginal distribution of the product demand that matters in the expected cost. Below, we provide a list of useful notations:

- C: the space capacity/constraint of the FDC.
- $\boldsymbol{c} = (c_1, ..., c_n)^\top \in \mathbb{R}^n$: c_i denotes the size of product i.
- $\boldsymbol{p} = (p_1, ..., p_n)^\top \in \mathbb{R}^n$: p_i denotes the production/procurement/out-bound shipping cost of replenishing unit quantity of product i from the RDC to the FDC.
- $h_i(\cdot)$: the inventory holding cost function for product i.

- $l_i(\cdot)$: the lost sale penalty function for product i.
- M_t: the shipping capacity (maximum total size of products shipped from the RDC to the FDC) on day t.
- $(\cdot)_+$: positive part function.
- $d_{i,t}$: the demand of product t on day t.
- $\mathcal{P}_{i,t}$ or $\mathcal{P}_i$: the distribution of $d_{i,t}$. We denote $\mathcal{P}_t = (\mathcal{P}_{1,t}, ..., \mathcal{P}_{n,t})$ and $\mathcal{P} = (\mathcal{P}_1, ..., \mathcal{P}_n)$.
- $x_{i,t}$: the initial inventory of product i before replenishment items arrive on day t.
- $z_{i,t}$: the replenishment quantity on day t.
- $y_{i,t}$: the inventory level of product i on hand after replenishment in period t, $y_{i,t} = x_{i,t} + z_{i,t}$ and $x_{i,t+1} = (y_{i,t} - d_{i,t})_+$.
- $\wedge$: maximum operator for scalars and vectors; $u \wedge v = \max\{u, v\}$ for $u, v \in \mathbb{R}$; $\boldsymbol{u} \wedge \boldsymbol{v} = (u_1 \wedge v_1, ..., u_m \wedge v_m)^\top$ for $\boldsymbol{u} = (u_1, ..., u_m)^\top, \boldsymbol{v} = (v_1, ..., v_m)^\top \in \mathbb{R}^m$

The below quantity represents all the cost incurred at the FDC and it consists of three parts, production/procurement/out-bound shipping cost, holding cost and lost sale cost:

$$p_i(y_{i,t} - x_{i,t}) + h_i((y_{i,t} - d_{i,t})_+) + l_i((d_{i,t} - y_{i,t})_+),$$

on day t and we define

$$g_i(y, d) \coloneqq h_i((y - d)_+) + l_i((d - y)_+).$$

The quantity $y_{i,t}$ represents the actual inventory on hand, instead of the virtual inventory in the backlogging case (Karlin and Scarf 1958, Ehrhardt 1984, Sethi and Cheng 1997). On day t, after the demand $d_{i,t}$ is realized, the initial inventory on the next day is determined by $x_{i,t+1} = (y_{i,t} - d_{i,t})_+$. In the e-commerce context, the lost sale cost can be interpreted as the sum of the cost for "real" lost sale and a penalty for the slower delivery from the RDC. For example, when there is a shortage for a product at the FDC, the orders from this FDC's district will have to be delivered from the RDC. It will result in a slower (and possibly more costly) delivery or even cause the customer to cancel the order directly. We encapsulate all the costs associated with the shortage of product at the FDC into the lost sale function $l_i(\cdot)$.

With a decision horizon of T, the product allocation problem can be formulated as follows

$$\tilde{L}_T(\boldsymbol{Y}|\boldsymbol{x}_1) := \mathbb{E}\left[\sum_{t=1}^{T} \boldsymbol{p}^\top(\boldsymbol{y}_t - \boldsymbol{x}_t) + \sum_{t=1}^{T}\sum_{i=1}^{n} g_i(y_{i,t}, d_{i,t})\right]$$
$$\text{s.t.} \quad \boldsymbol{c}^\top(\boldsymbol{y}_t - \boldsymbol{x}_t) \le M_t,\ \ \boldsymbol{x}_{t+1} = (\boldsymbol{y}_t - \boldsymbol{d}_t)_+,\ \ \boldsymbol{c}^\top \boldsymbol{y}_t \le C,\ \ \boldsymbol{y}_t \ge \boldsymbol{x}_t,\ \text{for } t \in [T]$$

where the expectation is taken with respect to $\boldsymbol{d}_t = (d_{1,t}, ..., d_{n,t})^\top \sim \mathcal{P}_t$. Here $\boldsymbol{x}_t = (x_{1,t}, ..., x_{n,t})^\top$ and $\boldsymbol{y}_t = (y_{1,t}, ..., y_{n,t})^\top$. In particular, $\boldsymbol{x}_1$ denotes the initial inventory. The decision variables are $\boldsymbol{Y} = (\boldsymbol{y}_1, ..., \boldsymbol{y}_T)$ and the objective is to minimize the total cost associated with the FDC. The first constraint models the shipping capacity, the second constraint captures the dynamics of the inventory level, and the third constraint states that the total size of all the products held in the FDC should not exceed its space capacity. Recall that our goal is to find a tactical way to use the space of the FDC and the shipping capacity between the RDC and the FDC. Thus the objective function restricts its attention to the cost associated with the FDC but ignores that of the RDC.

Consider an alternative cost function $L_T(\boldsymbol{Y}|\boldsymbol{x}_1)$,

$$L_T(\boldsymbol{Y}|\boldsymbol{x}_1) := \mathbb{E}\left[\sum_{t=1}^{T} \boldsymbol{p}^\top(\boldsymbol{y}_t - \boldsymbol{x}_t) + \sum_{t=1}^{T}\sum_{i=1}^{n} g_i(y_{i,t}, d_{i,t}) - \boldsymbol{p}^\top \boldsymbol{x}_{T+1}\right]$$
$$\text{s.t.} \quad \boldsymbol{c}^\top(\boldsymbol{y}_t - \boldsymbol{x}_t) \le M_t,\ \ \boldsymbol{x}_{t+1} = (\boldsymbol{y}_t - \boldsymbol{d}_t)_+,\ \ \boldsymbol{c}^\top \boldsymbol{y}_t \le C,\ \ \boldsymbol{y}_t \ge \boldsymbol{x}_t,\ \text{for } t \in [T]. \tag{1}$$

Compared with the original cost function $\tilde{L}_T(\boldsymbol{Y}|\boldsymbol{x}_1)$, its surrogate $L_T(\boldsymbol{Y}|\boldsymbol{x}_1)$ includes an additional term $\boldsymbol{p}^\top \boldsymbol{x}_{T+1}$ penalizing the production/shipping cost of remaining inventory at the end of the horizon. We will see in the rest of the paper that the additional term creates much analytical convenience and the technique dates back to Veinott Jr (1965). With a slight abuse of the notation, we use L_T and $\tilde{L}_T$ to refer to both the optimization problem and the objective function hereafter.

In practice, both the optimization problems $\tilde{L}_T$ and L_T should be solved in a dynamic fashion. That is, when choosing the value of $\boldsymbol{y}_t$, the decision maker cannot foresee future demands $\{\boldsymbol{d}_{t'}\}_{t'=t}^{T}$. To formalize the decision rule, we define a class of *non-anticipatory and feasible* policies Π. Specifically, we call a product allocation policy *non-anticipatory and feasible* if the decision is a function of history up to time t,

$$\boldsymbol{y}_t = \pi(\boldsymbol{x}_1, \boldsymbol{d}_1, \cdots, \boldsymbol{x}_t)$$

and it conforms to the constraints in (1). We emphasize that the decision $\boldsymbol{y}_t$ has to be made before the demand $\boldsymbol{d}_t$ is realized, i.e., without the knowledge of $\boldsymbol{d}_t$. Also, for different constraint parameters C and $\boldsymbol{M} = (M_1, ..., M_T)$, the set of feasible policy will be different. We denote the set of feasible policies to be $\Pi(\boldsymbol{M}, C, T)$, and we denote the expected cost under a policy π as $L_T^{\pi}(\boldsymbol{Y}|\boldsymbol{x_1})$ and $\tilde{L}_T^{\pi}(\boldsymbol{Y}|\boldsymbol{x_1})$, respectively. Let $\pi^*, \tilde{\pi}^* \in \Pi(\boldsymbol{M}, C, T)$ be the optimal policy for the problem L_T and $\tilde{L}_T$. We state some general assumption on the asymptotic optimality of the cost.

Assumption 3.2 *We assume $\boldsymbol{p} > 0$, $h_i(x) > 0$ and $l_i(x) > 0$ for $x > 0$, and $\mathcal{P}_i(d_i = 0) < 1$ for all $i \in [n]$.*

Proposition 3.3 *Under Assumption 3.2, for all the initial inventory $\boldsymbol{x}_1$,*

$$\frac{\tilde{L}_T^{\tilde{\pi}^*}(\boldsymbol{Y}|\boldsymbol{x}_1)}{\tilde{L}_T^{\pi^*}(\boldsymbol{Y}|\boldsymbol{x}_1)} \geq 1 - \frac{\boldsymbol{p}^\top \mathbb{E}^{\pi^*}[\boldsymbol{x}_{T+1}]}{\tilde{L}_T^{\pi^*}(\boldsymbol{Y}|\boldsymbol{x}_1)}.$$

and

$$\lim \frac{\tilde{L}_T^{\tilde{\pi}^*}(\boldsymbol{Y}|\boldsymbol{x}_1)}{\tilde{L}_T^{\pi^*}(\boldsymbol{Y}|\boldsymbol{x}_1)} \to 1.$$

as T goes to infinity.

The above proposition validates the choice of L_T as a surrogate for the original cost $\tilde{L}_T$. More specifically, the optimal policy on L_T serve as a good approximation on $\tilde{L}_T$. Therefore, for the rest of the paper, we will focus our discussion on the problem L_T.

Remark. We point out that our model shares similar capacitated structure with related literature, though Assumption 3.1 (a) requires a sufficient capacity for the RDC. In fact, the warehouse space capacity for each FDC has the same functionality as the production constraint in Federgruen and Zipkin (1986a) and DeCroix and Arreola-Risa (1998). In addition, the shipping cost and constraint integrate the out-bound shipping aspect into the inventory management decisions. We note that sometimes the shipping cost can be absorbed into the holding and lost sale cost (Levi et al. 2007). However, the cost re-parametrizaton cannot be done in our context, because we deal with a multi-period problem and the remaining inventory at each period can be used as initial inventory in the following period. This is the key distinction of our formulation from the newsvendor setting. As to the sufficient inventory assumption on the RDC, it is crucial because otherwise the optimal solution (even a near-optimal heuristic solution) cannot be obtained tractably. With this assumption, it reduces the problem to a one-RDC one-FDC problem and thus allows us to focus on the allocation of (i) the limited space at the FDC and (ii) the limited daily shipping capacity. To the best of our knowledge, we provide the first study of the one-RDC one-FDC system with the presence of both space and shipping constraints.

3.1 Single-Period Allocation with Space Constraint

For the moment, we ignore the issue of non-stationary demand and study the allocation problem for one single period. Let the demand $d_{i,t}$ follow an identical distribution $\mathcal{P}_i$ for all $t \in [T]$. In next section, we will analyze the multi-period problem and also we will return to address the non-stationary issue later in Section 4.3 and 5.

With the classic transformation of Veinott Jr (1965), we define

$$\begin{aligned} r_i(y, d) &\coloneqq p_i y + g_i(y, d) - p_i(y - d)_+ \\ f_i(y) &\coloneqq \int_{\mathbb{R}^+} r_i(y, \xi) d\mathcal{P}_i(\xi), \end{aligned}$$

where $r_i(y, d)$ represents the single-period cost of product i under decision y and realized demand d. The deterministic quantity $f_i(y)$ takes expectation with respect to the demand d_i. Denote

$$F_{\mathcal{P}}(\boldsymbol{y}) \coloneqq \sum_{i=1}^{n} f_i(y_i), \tag{2}$$

where the $\mathcal{P}$ stands for the demand distribution $\mathcal{P} = (\mathcal{P}_1, ..., \mathcal{P}_n)$. Then the objective function can be rewritten as

$$\begin{aligned} L_T(\boldsymbol{Y}|\boldsymbol{x}_1) &= \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{n} r_i(y_{i,t}, d_{i,t})\right] - \boldsymbol{p}^\top \boldsymbol{x}_1 \\ &= \sum_{t=1}^{T} F_{\mathcal{P}}(\boldsymbol{y}_t) - \boldsymbol{p}^\top \boldsymbol{x}_1 \end{aligned} \tag{3}$$

where $\boldsymbol{y}_t = (y_{1,t}, ..., y_{n,t})$ for $t \in [T]$.

The optimization problem L_T then becomes

$$\begin{aligned}\min \quad & L_T(\boldsymbol{Y}|\boldsymbol{x}_1) = \mathbb{E}\left[\sum_{t=1}^{T} F_{\mathcal{P}}(\boldsymbol{y}_t)\right] - \boldsymbol{p}^\top \boldsymbol{x}_1 \\ \text{s.t.} \quad & \boldsymbol{c}^\top(\boldsymbol{y}_t - \boldsymbol{x}_t) \le M_t, \;\; \boldsymbol{x}_{t+1} = (\boldsymbol{y}_t - \boldsymbol{d}_t)_+, \;\; \boldsymbol{c}^\top \boldsymbol{y}_t \le C, \;\; \boldsymbol{y}_t \ge \boldsymbol{x}_t, \; \text{for } t \in [T]\end{aligned} \tag{4}$$

where the expectation is taken with respect to $\boldsymbol{d}_t \sim \mathcal{P}$ for $t \in [T]$. We note that the function $F_{\mathcal{P}}(\cdot)$ is deterministic as in (3) and the expectation in (4) is due to the randomness of the constraints. In other words, the feasible region for $\boldsymbol{y}_t$ is random and dependent on the random initial inventory level $\boldsymbol{x}_t$. When there is no constraint, the optimization problem (4) is reduced to the deterministic optimization problem (3) which can be separated into independent univariate optimization problems (one for each product). The resultant optimal policy is a base stock (safety stock) policy where the optimal stock level of each product is independently determined by the corresponding optimization problem. The presence of the constraints couples the decisions for different products, and it motivates the definition of the following optimization problem to capture the single-period cost

$$\begin{aligned} OPT_{\mathcal{P}}(W, \boldsymbol{x}) \coloneqq \min_{\boldsymbol{y}} F_{\mathcal{P}}(\boldsymbol{y}) = \sum_{i=1}^{n} f_i(y_i) \\ \text{s.t. } \boldsymbol{c}^\top \boldsymbol{y} \le W, \boldsymbol{y} \ge \boldsymbol{x} \end{aligned} \tag{5}$$

where $\boldsymbol{x} \ge 0$ and $W > \boldsymbol{c}^\top \boldsymbol{x}$. $\mathcal{P} = (\mathcal{P}_1, ..., \mathcal{P}_n)$ encapsulates the distributions of the demand d_i for $i \in [n]$ and it defines the function $F(\cdot)$ as in (2). The parameter $\boldsymbol{x}$ can be interpreted as the initial inventory level, and the decision variables $\boldsymbol{y} = (y_1, ..., y_n)^\top \in \mathbb{R}^n$ represents the inventory level after replenishment.

Assumption 3.4 *We assume f_i is a convex and continuously differentiable function for all $i \in [n]$.*

The convexity assumption is common in relevant literature (Veinott Jr 1965, Evans 1967, Federgruen and Zipkin 1986a, Zipkin 2008b). For example, a common choice is to set the storage and penalty cost to be linear, i.e. $h(x) = h \cdot (x)_+$ and $l(x) = l \cdot (x)_+$. Under the convexity assumption, the above proposition states that if the unconstrained optimization problem outputs an optimal solution $\bar{\boldsymbol{y}}^*$ that takes more space than W, then in the constrained version of the problem, the constraint will be active for the optimal solution $\boldsymbol{y}^*_{W,\boldsymbol{x}}$.

Proposition 3.5 *Under Assumption 3.4, when $W > 0$, $\boldsymbol{x} \geq \mathbf{0}$, and $\boldsymbol{c}^\top \boldsymbol{x} \leq W$, there exists $\boldsymbol{y}^*_{W,\boldsymbol{x}} \in \mathbb{R}^n$ being the optimal solution to the problem $OPT_{\mathcal{P}}(W, \boldsymbol{x})$. In addition, denote $\bar{\boldsymbol{y}}^*$ as the optimal solution to the optimization problem $\min_{\boldsymbol{y} \geq \mathbf{0}} F_{\mathcal{P}}(\boldsymbol{y})$. If $\boldsymbol{c}^\top \boldsymbol{x} \leq W \leq \boldsymbol{c}^\top \bar{\boldsymbol{y}}^*$, then $\boldsymbol{c}^\top \boldsymbol{y}^*_{W,\boldsymbol{x}} = W$.*

3.2 Two-product Example for OPT

To generate some insights for the optimization problem $OPT_{\mathcal{P}}(W, \boldsymbol{x})$, we consider an example with one FDCs and one RDC system with two products. For $i \in \{1, 2\}$, let w_i, p_i, h_i, l_i be the per item size, production/shipping cost, storage cost and lost sale penalty, and let the initial inventory $\boldsymbol{x} = \mathbf{0}$. Assume the demand distribution in the FDC for product i is uniform on $[0, D_i]$. Then, the objective function $F_{\mathcal{P}}(\boldsymbol{y})$ of OPT is specified as

$$\begin{aligned}
F_{\mathcal{P}}(\boldsymbol{y}) &= f_1(y_1) + f_2(y_2) \\
&= \int_0^{D_1} \frac{1}{D_1} \left(p_1 y_1 + (h_1 - p_1)(y_1 - \xi_1)_+ + l_1(\xi_1 - y_1)_+\right) d\xi_1 \\
&\quad + \int_0^{D_2} \frac{1}{D_2} \left(p_2 y_2 + (h_2 - p_2)(y_2 - \xi_2)_+ + l_2(\xi_2 - y_2)_+\right) d\xi_2 \\
&= \sum_{i=1}^{2} \left(p_i y_i + \frac{(h_i - p_i) y_i^2}{2D_i} + \frac{l_i (D_i - y_i)^2}{2D_i} \right)
\end{aligned}$$

At the current time t the optimal policy is obtained by minimizing the following problem

$$\begin{aligned}
&\min_{\boldsymbol{y} = (y_1, y_2)} F_{\mathcal{P}}(\boldsymbol{y}) \\
&\text{s.t. } c_1 y_1 + c_2 y_2 \leq W, \qquad y_1, y_2 \geq 0.
\end{aligned}$$

The optimality condition tells that when the constraint is binding, from Lagrange multiplier we know that

$$\begin{aligned}
y_1^* &= \left(\frac{\left(p_2 - l_2 + \frac{W(h_2 - p_2 + l_2)}{c_2 D_2}\right) c_1 + (l_1 - p_1) c_2}{\frac{(h_2 - p_2 + l_2) c_1}{c_2 D_2} c_1 + \frac{h_1 - p_1 + l_1}{D_1} c_2} \wedge \frac{W}{c_1} \right) \vee 0 \\
y_2^* &= (W - c_1 y_1)/c_2.
\end{aligned} \tag{6}$$

When the cost structure and the size of the two products are the same ($c_1 = c_2, p_1 = p_2, l_1 = l_2$), and the holding cost $h_1 = h_2 = 0$, the optimal solution to the problem is that

$$y_1^* = \frac{D_1 W}{D_1 + D_2}, \quad y_2^* = \frac{D_2 W}{D_1 + D_2},$$

to allocate the capacity W proportional to the demand parameters of the two products. It provides an intuitive and straight-forward heuristic for single-period allocation of two products. In the following sections, we will pursue further along the path and investigate more on the analytical and data-driven aspects of the multi-product multi-period allocation problem (1).

4 Base Stock Policy and Myopic Policy

In this section, we discuss two policies for the multi-period production allocation problem with the help of the preceding result on the single-period problem. For the first two subsections, we assume demand stationarity that the demand $d_{i,t}$ follows the same distribution $\mathcal{P}_i$ for all $t \in [T]$. In the last subsection, we provide a solution for the non-stationary case where the demand $d_{i,t}$ follows a different distribution $\mathcal{P}_{i,t}$ for each $t \in [T]$. Throughout this section, we assume the knowledge of the distribution $\mathcal{P}_i$ and $\mathcal{P}_{i,t}$, and in the next section, we develop data-driven methods that handle the setting with unknown distributions.

4.1 Optimal Policy without Shipping Constraint

We first consider the case when there only exists the FDC space constraint but no shipping constraint. The case can be captured by letting $M_t \geq C$ for $t \in [T]$ in (1). The following proposition states that in this case, a base stock policy is optimal and the optimal base stock level is jointly determined by the cost structure, demand, and the constraint C through the optimization problem $OPT_{\mathcal{P}}(C, \boldsymbol{x}_t)$. In essence, when there is no shipping constraint, the analysis here is parallel to the literature on inventory management with production constraint (Evans 1967, Federgruen and Zipkin 1986a, DeCroix and Arreola-Risa 1998). In the one-RDC one-FDC warehouse system, the production constraint on RDC can be moved to the storage space constraint on FDC without changing the analytical structure.

Proposition 4.1 *Let $\boldsymbol{y}_t^{(bs)}$ be the optimal solution to the optimization problem $OPT_{\mathcal{P}}(C, \boldsymbol{x}_t)$. Under Assumption 3.1 and 3.4, the base policy π_{bs}*

$$\pi_{bs}(\{\boldsymbol{x}_1, \boldsymbol{d}_1, ..., \boldsymbol{x}_t\}) = \boldsymbol{y}_t^{(bs)},$$

is optimal for the problem $L_T(\boldsymbol{Y}|\boldsymbol{x}_1)$ when $M_t \geq C$ for all $t \in [T]$. In addition, denote the optimal solution to $OPT_{\mathcal{P}}(C, \boldsymbol{0})$ as $\boldsymbol{y}^*$. If the initial inventory $\boldsymbol{x}_1 \leq \boldsymbol{y}^*$ (element-wise), then the base stock level $\boldsymbol{y}_t^{(bs)} = \boldsymbol{y}^*$ for $t \in [T]$.

The base stock policy replenishes the inventory at the FDC up to a stock level prescribed by the optimization problem $OPT_{\mathcal{P}}(C, \boldsymbol{x}_t)$. When $\boldsymbol{x}_1 \leq \boldsymbol{y}^*$, the base stock level will remain the same as $\boldsymbol{y}^*$ throughout the horizon. On the contrary, if $x_{i,1} > y_i^*$ for certain product i, the optimal decision awaits the depletion of the inventory of product i and allocates the remaining available space for other product i' that $x_{i',1} < y_{i'}^*$. Also, note that to compute $\boldsymbol{y}_t^{(bs)}$, it entails the knowledge of distributions $\mathcal{P}_i$'s but not that of the realized demand $d_{i,t}$'s to solve $OPT_{\mathcal{P}}(C, \boldsymbol{x}_t)$.

4.2 Myopic Policy With Shipping Constraint

Under the base stock policy π_{bs}, the replenishment quantity on day t is $\boldsymbol{y}_t^{(bs)} - \boldsymbol{x}_t$. When there exists a shipping constraint M_t, it may happen that $\boldsymbol{c}^\top(\boldsymbol{y}_t^{(bs)} - \boldsymbol{x}_t) > M_t$. Then the base stock policy violates the shipping constraint and the replenishment cannot be fully attained. Theoretically, with the shipping constraint $\boldsymbol{M} = (M_1, ..., M_T)$, the optimal policy for $L_T(\boldsymbol{Y}|\boldsymbol{x}_1)$ should be specified via dynamic programming

$$L_{t,T}^*(\boldsymbol{y}_t, ..., \boldsymbol{y}_T|\boldsymbol{x}_t) = \min_{\boldsymbol{y}_t \geq \boldsymbol{x}_t, \boldsymbol{c}^\top(\boldsymbol{y}_t - \boldsymbol{x}_t) \leq M_t, \boldsymbol{c}^\top \boldsymbol{y}_t \leq C} \left\{ F_{\mathcal{P}}(\boldsymbol{y}_t) + \mathbb{E}\left[L_{t+1,T}^*(\boldsymbol{y}_{t+1}, ..., \boldsymbol{y}_T|(\boldsymbol{y}_t - \boldsymbol{d}_t)_+) \right] \right\}, \tag{7}$$

for $t = 1, ..., T$ and the boundary condition $L_{T+1,T}^*(\boldsymbol{Y}_{T+1}|\boldsymbol{x}_{T+1}) \equiv -\boldsymbol{p}^\top \boldsymbol{x}_1$. The expectation is taken with respect to $\boldsymbol{d}_t \sim \mathcal{P} = (\mathcal{P}_1, ..., \mathcal{P}_n)$ for $t \in [T]$. The optimal objective value for the problem L_T is given by $L_{1,T}^*$. Unfortunately, this dynamic programming formulation is in general intractable to solve.

Given the preceding analysis on the base stock policy, it is tempting to design a similar policy by replacing the constraint C with $C \wedge (M_t - \boldsymbol{c}^\top \boldsymbol{x}_t)$. The new constraint capacity takes the minimum of the space constraint and the shipping constraint at time t. In this way, at time t, we solve the following optimization problem

$$\begin{aligned} OPT_{\mathcal{P}}(C \wedge (M_t - \boldsymbol{c}^\top \boldsymbol{x}_t), \boldsymbol{x}_t) = \min_{\boldsymbol{y}} & F_{\mathcal{P}}(\boldsymbol{y}) \\ \text{s.t. } & \boldsymbol{c}^\top \boldsymbol{y} \leq C \wedge (M_t - \boldsymbol{c}^\top \boldsymbol{x}_t), \ \ \boldsymbol{y} \geq \boldsymbol{x}_t, \end{aligned} \tag{8}$$

where $\mathcal{P}$ is the distribution of $\boldsymbol{d}_t$ in defining the function $F_{\mathcal{P}}(\cdot)$ as in (2). Denote the optimal solution of (8) as $\boldsymbol{y}_t^{(mp)}$. First, $\boldsymbol{y}_t^{(mp)}$ is feasible in that it conforms to both the space constraint C and the shipping constraint M_t. Second, it can be interpreted as a myopic heuristic for the dynamic programming formulation (7). At each time t, the optimization problem (8) ignores the cost-to-go function $L^*_{t+1,T}$ in (7) and only focuses on the single-period cost incurred at time t. In this light, we name the policy generated by (8) as *myopic* policy π_{mp}. The myopic policy π_{mp} differs from the base stock policy π_{bs} when the shipping capacity is not enough to close the gap between the warehouse capacity C and the initial inventory level $\boldsymbol{c}^\top \boldsymbol{x}_t$ on day t. It may be caused by a demand surge on day $t-1$ or a shortage of shipping capacity M_t on day t. In this case, the shipping constraint M_t (instead of the space capacity C) becomes the bottleneck of the inventory replenishment.

The myopic policy offers a tractable heuristic to the problem L_T by ignoring the cost-to-go function. Recall that the base stock also focuses on the single-period cost and prescribes an ideal inventory level $\boldsymbol{y}_t^{(bs)}$ in each period. To generate some insights for why the myopic policy is sub-optimal in general but the base stock policy is optimal (without the shipping constraint), suppose the initial inventory level $\boldsymbol{x}_1 \leq \boldsymbol{y}^*$ (see Proposition 4.1 for definition). In this case, Proposition 4.1 tells that the optimal base stock level $\boldsymbol{y}_t^{(bs)} = \boldsymbol{y}^*$ for all $t \in [T]$. When there is no shipping constraint, the replenishment on day t could always recover whatever initial inventory level $\boldsymbol{x}_t (\geq \boldsymbol{0})$ to $\boldsymbol{y}^*$. In other words, the current replenishment decision does not need to account its effect on the future (the initial inventory level $\boldsymbol{x}_{t+1}$ in the next day) because the replenishment on day $t+1$ can always rectify a possibly "bad" status of $\boldsymbol{x}_{t+1}$. In the language of *Markov Decision Process*, we can view $\boldsymbol{x}_t$ as the state in the t-th time period. When there is no shipping constraint, $\boldsymbol{y}^*$ is the action that minimizes the single-period cost of period t; we do not need to worry about the future cost because after a random transition, it launches into the state $\boldsymbol{x}_{t+1}$, and there always exists another action to return to the state $\boldsymbol{y}^*$ again regardless of the value $\boldsymbol{x}_{t+1}$. In this light, all possible $\boldsymbol{x}_{t+1}$ will result in the same cost-to-go value. Then the best practice is to be myopic and to let the decision at each period take care of the single-period cost.

Things become different when there is a shipping constraint. As before, the decision $\boldsymbol{y}_t$ not only determines the cost on day t but also affects the initial inventory $\boldsymbol{x}_{t+1}$ on day $t+1$. However, the shipping constraint may not allow a replenishment to the ideal level $\boldsymbol{y}^*$ on day $t+1$, and consequently the cost-to-go value from day $t+1$ on may be dependent on $\boldsymbol{x}_{t+1}$. The myopic policy focuses only on the single-period cost and it does not capture the different cost-to-go values caused by $\boldsymbol{x}_{t+1}$. This is the reason why the myopic policy may not be optimal when there is a shipping constraint. Figure 5 further illustrates the intuition with a two-product example. $\boldsymbol{x}_t$ denotes the current inventory level and the shaded region is the feasible region for $\boldsymbol{y}_t$ under the shipping constraint M_t. It may happen that the ideal inventory level $\boldsymbol{y}^*$ is outside the feasible region. In this case, the myopic policy may adopt a decision $\boldsymbol{y}_t^{(mp)}$ different from the optimal decision $\boldsymbol{y}_t^{\pi^*}$ which takes into account the future cost-to-go values.

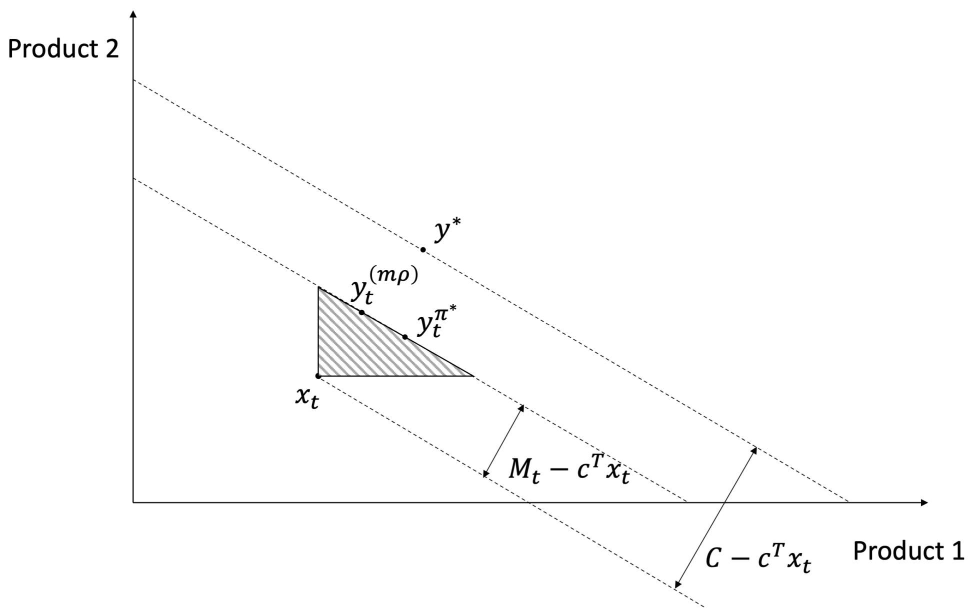


Figure 5: Illustration for the myopic policy

Now we establish that under certain conditions on the shipping constraint, the myopic policy turns out to be optimal. Condition 4.2 a) states that when the myopic policy outputs a different decision $\boldsymbol{y}_t^{(mp)}$ from $\boldsymbol{y}^*$ due to the insufficiency of M_t, then the shipping capacity in the next day will always be able to replenish the inventory to $\boldsymbol{y}^*$. Technically speaking, it guarantees that the cost-to-go value from day $t+1$ is the same and thus being myopic on day t is optimal. Condition 4.2 b) characterize a lower bound $M_{\min}$ for all the M_t. For a two period MDP with any $\boldsymbol{x}_1 \leq \boldsymbol{y}^*$, $\boldsymbol{M}_1 = +\infty$ and $\boldsymbol{M}_2 = \boldsymbol{M}_{\min}$, the $\boldsymbol{M}_{\min}$ needs to be large enough such that the myopic policy will be optimal. Intuitively, this avoids the action at t to affect the optimality of $t+1$. For the most extreme case, if $M_2 = 0$, we have $\boldsymbol{y}_2 = (\boldsymbol{y}_1 - \boldsymbol{d}_1)_+$ and there is no guarantee that taking $\boldsymbol{y}_1 = \boldsymbol{y}^*$ is the optimal solution. Without such conditions the MDP would be computationally infeasible.

Condition 4.2

(a) For $t \in [T]$, if $\boldsymbol{c}^\top(\boldsymbol{y}^* - \boldsymbol{x}_t) > M_t$, the shipping capacity M_{t+1} satisfies $\boldsymbol{c}^\top(\boldsymbol{y}^* - \boldsymbol{x}_{t+1}) \le M_{t+1}$. In addition, the initial inventory $\boldsymbol{x}_1 \le \boldsymbol{y}^*$.

(b) $M_t \ge M_{\min}$, where $M_{\min}$ is the lowest constant that ensures when $T = 2$, $\boldsymbol{x}_1 \le \boldsymbol{y}^*$, $M_1 = +\infty$ and $M_2 = M_{\min}$, the optimal policy is to set $\boldsymbol{y}_t$ to be the optimizer of (8).

Proposition 4.3 *Let $\boldsymbol{y}_t^{(mp)}$ be the optimal solution the optimization problem (8). Under Assumption 3.1, 3.2, 3.4 and Condition 4.2, the myopic policy*

$$\pi_{mp}(\{\boldsymbol{x}_1, \boldsymbol{d}_1, ..., \boldsymbol{x}_t\}) = \boldsymbol{y}_t^{(mp)}$$

is optimal for the problem $L_T(\boldsymbol{Y}|\boldsymbol{x}_1)$.

By definition, the myopic policy π_{mp} extends the base stock policy π_{bs} to the case of shipping constraint. When there is no shipping constraint, the myopic policy π_{mp} naturally reduces to the base stock policy π_{bs}. In this view, Proposition 4.3 with Condition 4.2 extends the optimality of the base stock policy in Proposition 4.1 to a more general context. Specifically, Condition 4.2 requires that if the shipping capacity M_t on day t is insufficient (causing $\boldsymbol{y}_t^{(mp)} \neq \boldsymbol{y}^*$), then the shipping capacity on day $t+1$ should be large enough to recover the ideal stock level $\boldsymbol{y}^*$. In practice, if a demand surge is observed on day t and it renders an insufficient shipping capacity, we can plan one-day ahead and allocate more on the following day; by doing this, the optimality of the myopic policy is ensured.

4.3 Re-allocation Policy under Non-stationary Demand

When analyzing the base stock policy and the myopic policy, we consider a stationary setting where the demand $d_{i,t}$ follows an identical distribution over time. For a non-stationary setting where $d_{i,t}$ follows a different distribution $\mathcal{P}_{i,t}$, the myopic nature of both the two proposed policies may elicit sub-optimal decisions. Consider if the market demand is shrinking over time, the single-day optimal base stock inventory level will decrease over time accordingly. Define $\bar{\boldsymbol{y}}_t$ as the optimal solution to the optimization problem $OPT_{\mathcal{P}_t}(C, \boldsymbol{0})$ where $\mathcal{P}_t = (\mathcal{P}_{1,t}, ..., \mathcal{P}_{n,t})$.

Then $\bar{\boldsymbol{y}}_t$ represents the optimal stock level that minimizes the single-day cost when there is no shipping constraint. A shrinking market will result in a decreasing trend of $\bar{\boldsymbol{y}}_t$'s because in the later periods, we do not need to prepare as much inventory as in early periods. Consequently, at an early period, a myopic decision that only focuses on minimizing single-day cost could result in holding excessive (and unnecessary) inventory in the future. In other words, it may happen frequently that the initial inventory on a day is larger than the ideal inventory level, $\boldsymbol{x}_t > \bar{\boldsymbol{y}}_t$. A natural idea to resolve the issue is to allow "backward" transshipment. Specifically, we allow not only inventory replenishment from the RDC to the FDCs as before but also re-allocation of the excessive inventory at FDCs back to RDCs.

With this additional option of re-allocation, a new optimization problem $\tilde{L}'_T$ can be formulated as below. It modifies the problem $\tilde{L}_T$ by appending an additional term $\boldsymbol{p}'^\top(\boldsymbol{x}_t - \boldsymbol{y}_t)_+$ in the objective function and removing the constraint $\boldsymbol{y}_t \geq \boldsymbol{x}_t$,

$$\tilde{L}'_T(\boldsymbol{Y}|\boldsymbol{x}_1) := \mathbb{E}\left[\sum_{t=1}^{T}\left(\boldsymbol{p}^\top(\boldsymbol{y}_t - \boldsymbol{x}_t)_+ + \boldsymbol{p}'^\top(\boldsymbol{x}_t - \boldsymbol{y}_t)_+ + \sum_{i=1}^{n} g_i(y_{i,t}, d_{i,t})\right)\right]$$
$$\boldsymbol{c}^\top(\boldsymbol{y}_t - \boldsymbol{x}_t)_+ \leq M_t,\ \boldsymbol{c}^\top(\boldsymbol{x}_t - \boldsymbol{y}_t)_+ \leq M_t,$$
$$\boldsymbol{x}_{t+1} = (\boldsymbol{y}_t - \boldsymbol{d}_t)_+,\ \boldsymbol{c}^\top\boldsymbol{y}_t \leq C,\ \text{s.t. for } \text{ for } t \in [T]$$

where the demand $\boldsymbol{d}_t$ follows the distribution $\mathcal{P}_t$ and the expectation is taken with respect to $\boldsymbol{d}_t$'s. The additional term $\boldsymbol{p}'^\top(\boldsymbol{x}_t - \boldsymbol{y}_t)_+$ penalizes the shipping cost of re-allocating the inventory from the FDC to the RDC. The first two constraints characterize the shipping capacity for replenishment and re-allocation, respectively. We also define an alternative objective function

$$L'_T(\boldsymbol{Y}|\boldsymbol{x}_1) := \mathbb{E}\left[\sum_{t=1}^{T}\left(\boldsymbol{p}^\top(\boldsymbol{y}_t - \boldsymbol{x}_t) + \sum_{i=1}^{n} g_i(y_{i,t}, d_{i,t})\right) - \boldsymbol{p}^\top\boldsymbol{x}_{T+1}\right]$$
$$\boldsymbol{c}^\top(\boldsymbol{y}_t - \boldsymbol{x}_t)_+ \leq M_t,\ \boldsymbol{c}^\top(\boldsymbol{x}_t - \boldsymbol{y}_t)_+ \leq M_t,$$
$$\boldsymbol{x}_{t+1} = (\boldsymbol{y}_t - \boldsymbol{d}_t)_+,\ \boldsymbol{c}^\top\boldsymbol{y}_t \leq C,\ \text{s.t. for } \text{ for } t \in [T].$$

Though the problem L'_T can be viewed as a surrogate for the problem $\tilde{L}'_T$, unlike the optimization problem L_T (as the surrogate for the problem $\tilde{L}_T$), the problem L'_T has less tangible meaning and, in fact, it is introduced simply for analytical convenience.

Proposition 4.4 *Denote the optimal policy of the problem $\tilde{L}'_T$ as $\tilde{\pi}'^*$. The following inequality holds for all $T \in \mathbb{N}_+$,*

$$\frac{\tilde{L}'^{\tilde{\pi}'^*}_T(\boldsymbol{Y}|\boldsymbol{x}_1)}{\tilde{L}'^{\pi'^*}_T(\boldsymbol{Y}|\boldsymbol{x}_1)} \geq 1 - \frac{\boldsymbol{p}^\top \mathbb{E}^{\pi'^*}[\boldsymbol{x}_{T+1}] + (\boldsymbol{p}+\boldsymbol{p}')^\top \mathbb{E}^{\pi'^*}\left[\sum_{t=1}^T (\boldsymbol{x}_t - \boldsymbol{y}_t)_+\right]}{\tilde{L}'^{\pi'^*}_T(\boldsymbol{Y}|\boldsymbol{x}_1)} \tag{9}$$

where π'^* is the optimal policy for the problem L'_T as defined earlier.

Proposition 4.4 states the approximation ratio of the value of $\tilde{L}'_T$ under the optimal policy for L'_T against the optimal value of $\tilde{L}'_T$. Compared with Proposition 3.3, there is an additional term on the right-hand-side $(\boldsymbol{p}+\boldsymbol{p}')^\top \mathbb{E}^{\tilde{\pi}'^*}\left[\sum_{t=1}^T (\boldsymbol{x}_t - \boldsymbol{y}_t)_+\right]$ that represents the cumulative shipping cost of inventory re-allocation. Therefore, the approximation ratio of $\tilde{L}'^{\pi^*}_T$ against $\tilde{L}'^{\tilde{\pi}'^*}_T$ depends on the total amount of re-allocated inventory $\sum_{t=1}^T (\boldsymbol{x}_t - \boldsymbol{y}_t)_+$ under the optimal policy for L'_T. When the quantity is small, the problem L'_T works as a good approximation for the problem $\tilde{L}'_T$.

In parallel to the stationary case (Proposition 4.3), we establish optimality result for the myopic policy for the problem L'_T under nonstationary demand. Define an alternative optimization problem

$$\begin{aligned} \mathrm{OPT}'_{\mathcal{P}_t}(C, W, \boldsymbol{x}) \coloneqq \min_{\boldsymbol{y}_t \geq \boldsymbol{x}} \mathbb{E}\left[\boldsymbol{p}^\top \boldsymbol{y}_t + \sum_{i=1}^n g_i(y_{i,t}, d_{i,t}) - \boldsymbol{p}^\top (\boldsymbol{y}_t - \boldsymbol{d}_t)_+\right] \\ \text{s.t. } \boldsymbol{c}^\top (\boldsymbol{y}_t - \boldsymbol{x})_+ \leq W, \quad \boldsymbol{c}^\top (\boldsymbol{x} - \boldsymbol{y}_t)_+ \leq W, \quad \boldsymbol{c}^\top \boldsymbol{y}_t \leq C, \quad \boldsymbol{y}_t \geq \boldsymbol{0} \end{aligned} \tag{10}$$

where the decision variables are $\boldsymbol{y}_t$ and the expectation is taken with respect to $\boldsymbol{d}_t = (d_{1,t}, ..., d_{n,t})^\top \sim \mathcal{P}_t$. We denote the optimal solution to the problem $OPT'_{\mathcal{P}_t}(C, M_t, \boldsymbol{x}_t)$ as $\boldsymbol{y}_t^{(ra)}$ and define the *myopic re-allocation policy* as

$$\pi_{ra}(\{\boldsymbol{x}_1, \boldsymbol{d}_1, ..., \boldsymbol{x}_t\}) = \boldsymbol{y}_t^{(ra)}.$$

Also, we denote

$$\begin{aligned} L_{s,t}(\boldsymbol{y}_s, \cdots, \boldsymbol{y}_t|\boldsymbol{x}_s) \coloneqq \mathbb{E}\left[\sum_{\tau=s}^t \left(\boldsymbol{p}^\top \boldsymbol{y}_\tau + \sum_{i=1}^n g_i(y_{i,\tau}, d_{i,\tau}) - \boldsymbol{p}^\top (\boldsymbol{y}_\tau - \boldsymbol{d}_\tau)_+\right)\right] \\ \boldsymbol{c}^\top (\boldsymbol{y}_\tau - \boldsymbol{x}_\tau)_+ \leq M_\tau, \ \boldsymbol{c}^\top (\boldsymbol{x}_\tau - \boldsymbol{y}_\tau)_+ \leq M_\tau, \\ \boldsymbol{x}_{\tau+1} = (\boldsymbol{y}_\tau - \boldsymbol{d}_\tau)_+, \ \boldsymbol{c}^\top \boldsymbol{y}_\tau \leq C, \ \text{s.t. for } \text{ for } \tau \in [t, s]. \end{aligned}$$

Next, we denote the optimal solution to the optimization problem $OPT'_{\mathcal{P}_t}(C, C, \boldsymbol{x}_t)$

as $\boldsymbol{y}_t^*$, and state the condition for the non-stationary case.

Condition 4.5 *(a) For $t \in [T]$, if $\boldsymbol{y}_t^{(ra)} \neq \boldsymbol{y}_t^*$, then the shipping capacity M_{t+1} is large enough to ensure $\boldsymbol{y}_{t+1}^{(ra)} = \boldsymbol{y}_{t+1}^*$.*

(b) $M_t \geq M_{\min}$, where $M_{\min}$ is the lowest constant that ensures for all $s \in [1, T-1]$, under condition $\boldsymbol{x}_s \leq \boldsymbol{y}_s^*$, $M_s = +\infty$ and $M_{s+1} = M_{\min}$, the optimal policy for minimizing $L_{s,s+1}(\boldsymbol{y}_s, \boldsymbol{y}_{s+1}|\boldsymbol{x}_s)$ is to set $\boldsymbol{y}_t^* = \boldsymbol{y}_t^{(ra)}$.

Proposition 4.6 *Under Assumption 3.1, 3.4 and Condition 4.5 , the myopic re-allocation policy π_{ra} is optimal for the re-allocation problem $L_T'(\boldsymbol{Y}|\boldsymbol{x}_1)$.*

We omit the proof of this proposition because the proof is analogous to Proposition 4.3. Proposition 4.6 states the optimality of the myopic re-allocation policy π_{ra} under Condition 4.5. The condition is similar to the previous one (Condition 4.2) in that it requires a lower bound on the shipping capacity and a one-day look-ahead flexibility on the shipping capacity. To summarize, we provide a solution for the non-stationary demand by allowing re-allocating products backward from the FDC to the RDC. Accordingly, a new optimization problem $\tilde{L}_T'$ is formulated and its surrogate as L_T'. Proposition 4.4 provides the approximation ratio of $\tilde{L}_T'$ and Proposition 4.6 states the optimality of the myopic re-allocation policy for L_T' under certain condition. Note that to implement the policy π_{ra}, it requires the knowledge of $\mathcal{P}_t$ but not the realized demand $\boldsymbol{d}_t$ or the future demand distribution. The same requirements also appears in the stationary setting as discussed earlier. In the next section, we will exploit the available auxiliary information and develop data-driven approaches to relax this requirement.

5 Data-Driven Algorithms and Case Study

In the last section, we analyze the myopic policy under the assumption that the (future) demand distribution is known. Now, based on the myopic policy and our analytical findings, we develop data-driven approaches to relax the known distribution assumption. Moreover, we employ the machine learning tools to better utilize the availability of auxiliary information and further boost the policy performance. We emphasize that all the algorithms developed in this section work for both stationary and non-stationary demand setting.

5.1 Online Myopic Algorithm via Sample Average Approximation

When the distribution $\mathcal{P}_t$ is unknown, a natural solution is to estimate it with the empirical distribution of history data. Algorithm 1 implements the idea and provides a prototype of implementing the myopic policy in practice. On day t, an empirical distribution $\hat{\mathcal{P}}_t$ is constructed based on history observations and it serves as the input for the optimization problem OPT. Thus the optimization problem $OPT_{\hat{\mathcal{P}}_t}$ solved in Algorithm 1 can be viewed as a *sample average approximation* to the (stochastic) optimization problem $OPT_{\mathcal{P}_t}$. From an estimation perspective, Algorithm 1 refines the distribution estimation continually as the new data arrives in an online fashion. We follow the convention in the online learning/online optimization literature and call the algorithm as the *online myopic algorithm*. In addition, the algorithm prototype offers much flexibility in the way of constructing the empirical distribution $\hat{\mathcal{P}}_t$. For example, instead of using all the history observations, we can also construct the empirical distribution with the most recent k observations or assign different weights to each observation.

Algorithm 1: Online Myopic Algorithm

Input: Model parameters $\boldsymbol{c}, C, \boldsymbol{M} = (M_1, ..., M_T), \boldsymbol{x}_1$
Output: $\boldsymbol{y}_1, ..., \boldsymbol{y}_T$
Set the decision $\boldsymbol{y}_1 = \boldsymbol{x}_1$
Set the distribution $\hat{\mathcal{P}}_1 = \mathbb{1}_{\boldsymbol{d}_1}$
for $t = 2, ..., T$ **do**
  Solve the optimization problem

$$OPT_{\hat{\mathcal{P}}_{t-1}}(C \wedge (M_t - \boldsymbol{c}^\top \boldsymbol{x}_t), \boldsymbol{x}_t)$$

  Set $\boldsymbol{y}_t$ to be its optimal solution
  Update the distribution $\hat{\mathcal{P}}_t = \frac{1}{t}\sum_{s=1}^{t} \mathbb{1}_{\boldsymbol{d}_s}$
end

5.2 Online Myopic Algorithm with Predictive Learning

Now we further extends Algorithm 1 with the machine learning tools and the auxiliary data. Intuitively, the demand of a product is dependent on many factors, including the past demand, price, promotion event, etc. This auxiliary information is usually available in advance and can be effectively used to predict the product demand and consequently to reduce the inventory cost. Recall that Algorithm 1 only utilizes the past demand ad hoc to construct an empirical distribution. Algorithm 2 builds upon Algorithm 1 with a predictive learning module. Specifically, in each period, we first compute a tentative inventory level $\boldsymbol{y}_t = (y_{1,t}, ..., y_{n,t})^\top$ as in Algorithm 1 and then employ an expert (machine learning model) which comments on whether the inventory level $\boldsymbol{y}_t$ should be adjusted given the auxiliary information. The expert provides three recommendation: (i) keep and execute with the tentative inventory level $y_{i,t}$; (ii) increase the inventory level by κ_i; (iii) decrease the inventory level by κ_i. The machine learning expert essentially performs a three-way classification and it can be trained by running the online myopic algorithm on history data. In the history data, we can collect and label samples by comparing whether the tentative inventory level $y_{i,t}$ (given by the online myopic algorithm): (i) exceeds the true demand $d_{i,t}$ by ν_i; (ii) goes below $d_{i,t}$ by ν_i; (iii) stays close to $d_{i,t}$ within a range of $[-\nu_i, \nu_i]$. The labelling rule corresponds to the three-way recommendation of the expert and the labelled sample can be used to train a machine learning model. The parameters κ_i and ν_i are treated as hyper-parameters and can also be tuned on the history data. Ideally, the trained machine learning model should be capable of predicting whether the tentative inventory level $\boldsymbol{y}_t$ is insufficient or excessive. With an adjustment of the inventory level $\boldsymbol{y}_t$ accordingly, the machine learning model refines the (static) myopic policy in a data-driven and dynamic fashion.

In Algorithm 2, the adjustment made by the machine learning expert $\mathcal{M}$ may result in an infeasible inventory plan (violating the shipping or warehouse constraint). To resolve the issue, we apply a thresholding rule and allocate the space proportionally to the tentative inventory level. The thresholding rule is inspired by the result in Section 3.2.

The demand/sales forecasting has been a widely studied topic in operations management (Choi et al. 2018). The classic way of using machine learning is to first predict the demand distribution and then feed the prediction in certain decision model. We provide an alternative treatment that first computes the solution to the decision model and then uses the machine learning model to refine the solution. On one hand, the machine learning model can utilize the additional features like promotion information and prices to prepare the inventory level for predictable demand surges or plunges. On the other hand, we can control explicitly the adjustment made by the machine learning model to be subsidiary refinement (such as through the parameter κ_i in Algorithm 2) and let it just serve as an auxiliary for the original solution provided by the (more interpretable) analytical model. In this way, we can avoid the instability issue for

Algorithm 2: Online Myopic Policy with Predictive Learning

Input: Model parameters $\boldsymbol{c}, C, \boldsymbol{M} = (M_1, ..., M_T), \boldsymbol{x}_1$, a machine learning expert $\mathcal{M}$

Output: $\boldsymbol{y}_1, ..., \boldsymbol{y}_T$

Set the decision $\boldsymbol{y}_1 = \boldsymbol{x}_1$

Set the distribution $\hat{\mathcal{P}}_1 = \mathbb{1}_{\boldsymbol{d}_1}$

for $t = 2, ..., T$ **do**

  Solve the optimization problem

$$OPT_{\hat{\mathcal{P}}_{t-1}}(C \wedge (M_t - \boldsymbol{c}^\top \boldsymbol{x}_t), \boldsymbol{x}_t)$$

  Denote $\boldsymbol{y}_t = (y_{1,t}, ..., y_{n,t})^\top$ to be its optimal solution

  for $i = 1, ..., n$ **do**

    Renew $y_{i,t}$ according to the machine learning expert $\mathcal{M}$

$$y_{i,t} \;\leftarrow\; y_{i,t} + \kappa_i \text{ or } y_{i,t} \text{ or } y_{i,t} - \kappa_i$$

  end

  if *the shipping or space constraint is violated* **then**

    Denote $C_t = C \wedge (M_t - \boldsymbol{c}^\top \boldsymbol{x}_t)$

    for $i = 1, ..., n$ **do**

      Apply the thresholding rule:

$$y_{i,t} \;\leftarrow\; C_t \cdot \frac{c_i y_{i,t}}{\sum_{i=1}^n c_i y_{i,t}}$$

    end

  end

  Update the distribution $\hat{\mathcal{P}}_t = \frac{1}{t} \sum_{s=1}^t \mathbb{1}_{\boldsymbol{d}_s}$

end

the machine learning model caused by lack of enough training data or disparity between training and testing data.

5.3 Benchmark Algorithm: Capacitated Base Stock Algorithm

Next we present a base stock algorithm for benchmark purpose. In comparison to the two algorithms presented earlier, the base stock algorithm (Algorithm 3) does not account the shipping constraint when solving the optimization problem OPT. This may result in a violation of the shipping constraint, and the remedy is to apply the thresholding rule and capacitate the tentative inventory level as in Algorithm 2 and Section 3.2. By comparing Algorithm 3 with the myopic algorithms, we will see the effectiveness of considering both space and shipping constraints in the optimization problem OPT. Also, we point out that the input of Algorithm 3 requires an access to history data $\mathcal{D}$. In the following numerical experiments, we will (unrealistically) use the true demand data as the history data $\mathcal{D} = (\boldsymbol{d}_1, ..., \boldsymbol{d}_T)$ for benchmark purpose. Similarly, we can have a predictive learning version of the base stock algorithm via the same method as Algorithm 2.

Algorithm 3: Base Stock Algorithm

Input: Model parameters $\boldsymbol{c}, C, \boldsymbol{M} = (M_1, ..., M_T), \boldsymbol{x}_1$, history data $\mathcal{D}$
Output: $\boldsymbol{y}_1, ..., \boldsymbol{y}_T$
Set the distribution $\hat{\mathcal{P}} = \sum_{\boldsymbol{d} \in \mathcal{D}} \mathbb{1}_{\boldsymbol{d}}$
for $t = 1, ..., T$ **do**
  Solve the optimization problem

$$OPT_{\hat{\mathcal{P}}}(C, \boldsymbol{x}_t)$$

  Denote $\boldsymbol{y}_t = (y_{1,t}, ..., y_{n,t})^\top$ to be its optimal solution
  if *the shipping or space constraint is violated* **then**
    Denote $C_t = C \wedge (M_t - \boldsymbol{c}^\top \boldsymbol{x}_t)$
    for $i = 1, ..., n$ **do**
      Apply the thresholding rule:

$$y_{i,t} \leftarrow C_t \cdot \frac{c_i y_{i,t}}{\sum_{i=1}^{n} c_i y_{i,t}}$$

    end
  end
end

5.4 Case Study

Now we present numerical experiments with data from JD and compare the performances of the proposed algorithms. We aim to compare and interpret different policies, so we will restrict our attention to the original problem $\tilde{L}_T$ which does not permit the backward re-allocation of the products from FDC to RDC. We select the top 100 best-selling products, and the demand data consists of 31 days. As to the cost structure, we assume the shipping, holding, and lost sale cost function to be linear, which allow us to use convex optimization tool

(CVX, for example) to numerically calculate solution for $OPT_{\hat{\mathcal{P}}}$. We randomly sample the product size i.i.d. from geometric distribution with parameter 1 and let the unit shipping cost and unit holding cost proportional to the product size. For the unit lost sale cost, we let it proportional to the product's price. On average, the following relation holds in our numerical experiments

$$\text{unit lost sale cost } > \text{unit shipping cost } > \text{unit holding cost}.$$

Algorithm		C_1, M_1	C_1, M_2	C_1, M_3	C_2, M_2	C_3, M_3
BSA		158988.71	156894.67	155046.70	154658.85	154325.34
OMA		151013.97	149210.09	147879.22	146904.31	145965.72
Train	BSA-PL	156579.27	154670.29	152193.64	151360.36	151171.51
	OMA-PL	140371.42	140691.22	138074.47	135824.38	162185.61
Test	BSA-PL	159800.58	159540.39	159519.79	161565.89	161565.89
	OMA-PL	162491.80	156876.33	161722.30	154572.00	152130.13
True	BSA-PL	153724.24	152453.03	151265.68	151265.69	151086.85
	OMA-PL	144489.82	140531.54	137183.03	135350.42	134167.90
10% Err.	BSA-PL	164397.69	163159.96	162417.90	162172.52	161717.36
	OMA-PL	145563.21	142294.67	141379.94	139711.75	137412.44
20% Err.	BSA-PL	171102.63	170167.88	169770.18	169351.34	169207.50
	OMA-PL	148672.46	145489.40	143571.53	139879.31	141129.88
40% Err.	BSA-PL	192310.64	191414.34	190647.32	190623.88	190124.32
	OMA-PL	152780.52	149761.67	146774.37	146407.81	144785.60
50% Err.	BSA-PL	195975.90	195544.13	195301.85	195267.99	195028.92
	OMA-PL	154658.39	151282.93	147247.74	146938.52	145705.86

Table 1: Total cost of $T = 31$ days for the largest RDC-FDC system: BSA, OMA, BSA-PL, OMA-PL stand for the base stock algorithm, the online myopic algorithm, the base stock algorithm with predictive learning, the online myopic algorithm with predictive learning respectively.

Table 1 reports the cost performance of the algorithms with different constraint levels. The performance is tested on an RDC-FDC system with one RDC and five FDCs, which is the largest regional inventory system of the JD that has the highest total sales. The cost is reported as the aggregated cost of all five FDCs. We also test different constraint combinations: M_2 is set to be the average daily demand (for all 100 products) of the FDC, $M_1 = 0.8M_2$, and $M_3 = 1.2M_2$. C_2 is set to be 70% the total optimal base stock level when there is no space or shipping constraint, and $C_1 = 0.8C_2$. In terms of the algorithms, multiple versions of the base stock algorithm and the online myopic algorithm are tested with different predictive learning models ($\mathcal{M}$): "Train" stands for using the same demand data for training (a Random Forest model is used for training) and testing; "Test" stands for training the model on the demand data from a different region, and testing the performance on the same data that "Train" model uses.

We are also interested in the minimal prediction accuracy that allow the predictive learning models to be effective. We randomly perturb the true label that indicate the correct three-way recommendation, and we take the perturbed label as the output the machine learning expert $\mathcal{M}$ recommends. Therefore, "p% Err." stands for the predictive learning model using the true label that has p% rate of random perturbation.

We denote BSA, OMA, BSA-PL, and OMA-PL for the base stock algorithm, the online myopic algorithm, the base stock algorithm with predictive learning, and the online myopic algorithm with predictive learning. We found that the predictive learning model is beneficial for both OMA and BSA types of algorithms. For OMA type of algorithms, with less than 40% perturbation error on the true label, the OMA PL outperforms OMA consistently under various setting of the capacity constraint C and shipping constraint M. We also found the total cost of OMA PL algorithm on test set being higher than the OMA algorithm without auxiliary information. This is because due to data insufficiency, we have to train the prediction model on the demand from a different region, where demand pattern could be different. In fact, from Table 2 we know for different DCs, the prediction accuracy on test sets ranges from 50% to 68%. We believe that in practice, the company can significantly improve the performance on test set with more data.

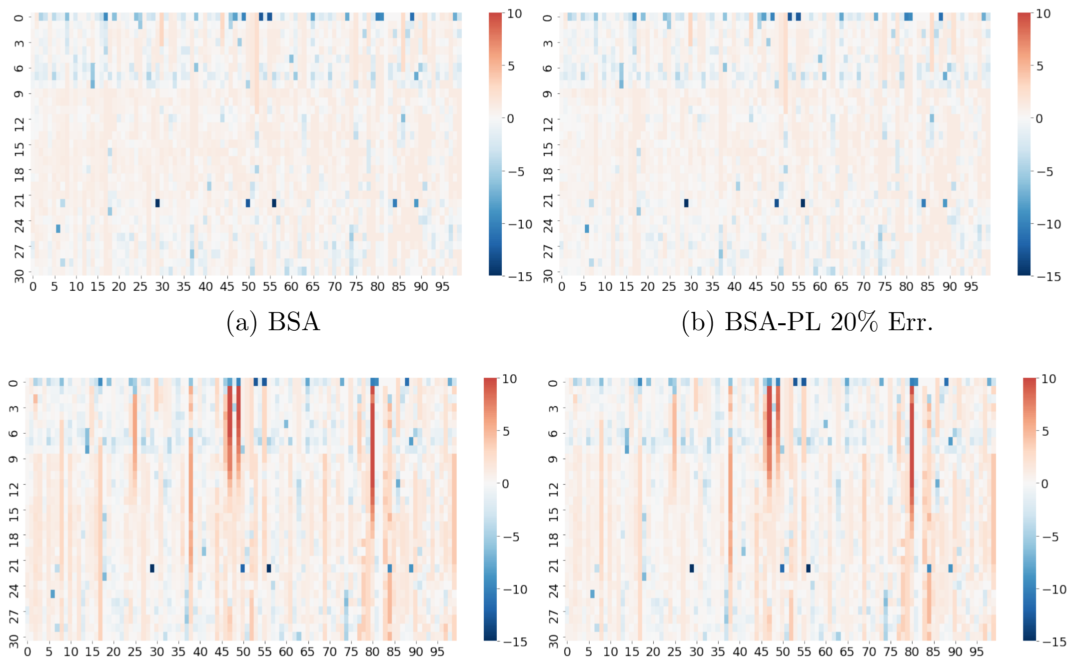
Figure 6 presents a visualization of the policies in terms of the difference between the inventory level and the realized demand. It depicts heatmaps of $\boldsymbol{y}_t - \boldsymbol{d}_t$ for $t \in [T]$. In general, the two right panels have slightly more neutral color than the two left panels where the neutral color indicates smaller absolute values of the difference. More importantly, we find that the predictive learning works effectively in mitigating large differences, which are not reflected on the heatmap due to truncation of large values. Both observations demonstrate the effectiveness of utilizing auxiliary information and the machine learning model. Furthermore, Table 4 breaks down the total cost in terms of shipping cost, holding

Data	Train	Test
DC2	0.964	0.661
DC6	0.967	0.500
DC20	0.957	0.563
DC42	0.983	0.625
DC43	0.957	0.688

Table 2: Predictive learning accuracy: The accuracy are reported for the the task of three-way classification.

Variable	Importance
Promotion length	0.22
Number of days after promotion	0.17
1-day past demand $d_{i,t-1}$	0.13
Change of price	0.13
2-day past demand $d_{i,t-2}$	0.11
3-day past demand $d_{i,t-1}$	0.10

Table 3: Variable importance in the trained random forest model



(a) BSA

(b) BSA-PL 20% Err.

(c) OMA

(d) OMA-PL 20% Err.

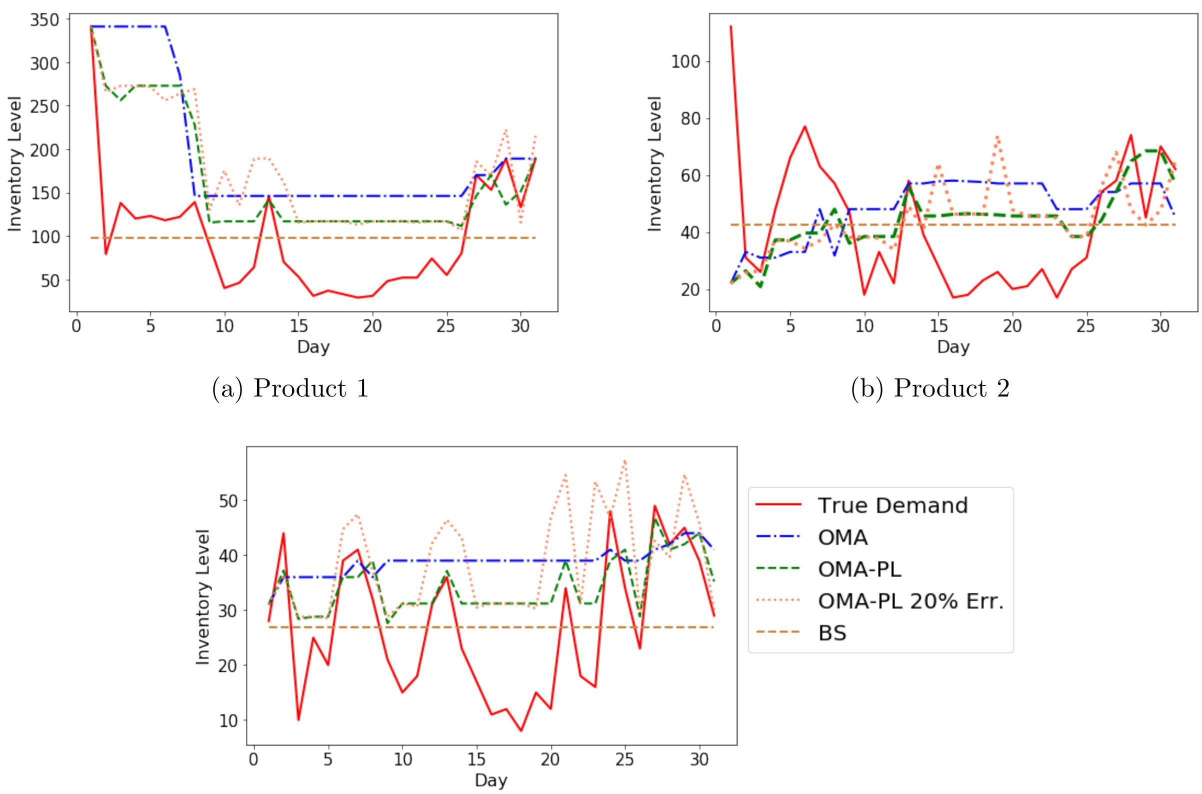
Figure 6: Inventory-demand difference heatmap on a particular DC (as an FDC): We plot the heatmap for $\boldsymbol{y}_t - \boldsymbol{d}_t$ for all $t \in [T]$ under the four algorithms. Each row of the heatmap represents a day vector $\boldsymbol{y}_t - \boldsymbol{d}_t$ and each column represents a product. We truncate the value to 10 or -15 for numbers that are larger than 10 or smaller than -15, respectively.

cost, and lost sale cost. The myopic algorithms have significantly smaller lost sale cost than their base stock counterpart. This phenomenon can also be observed from Figure 6 in that the heatmaps of the myopic algorithms have more balanced coloring than the base stock algorithms.

Lastly, we visualize the decision $\boldsymbol{y}_t$'s under different algorithms in Figure

Algorithm	Shipping	Holding	Lost sale	Total
BSA	14050	23277	119567	156895
OMA	12983	29576	106650	149210
BSA-PL 20% Err.	13937	19860	120873	154670
OMA-PL 20% Err.	12855	24191	103645	140691

Table 4: Cost Break-down for a particular DC (as an FDC)



(a) Product 1

(b) Product 2

(c) Product 3

Figure 7: Inventory level visualizations: The inventory curve $y_{i,t}$ over t under different policies for three different products.

7. The visualization demonstrates the algorithms' ability to adapt the inventory level according to the demand curve. The online myopic algorithm (OMA curve), though incorporate no auxiliary information, slowly adjusts to the changes in demand as more demand data is available. However, with the auxiliary data, the OMA-PL algorithm is able to adapt to the demand fluctuation better: when there is a surge in demand, the OMA-PL replenish the corresponding item more than

OMA does, hence resulting in a lower lost sale cost; when the demand declines, the OMA-PL algorithm maintains less inventory level than OMA, which reduces holding cost. This observation is confirmed with Table 4, where OMA-PL with 20% perturbation rate has a lower holding and lost sale cost.

In conclusion, from our experiment we found that the Online SAA Myopic Algorithm has better performance than its base-stock counterpart, and the auxiliary information can be leveraged to reduce the cost even more. We believe our predictive learning methodology can benefit a lot from a larger size of data and help managers to make more intelligent inventory decisions.

References

Acimovic J, Graves SC (2015) Making better fulfillment decisions on the fly in an online retail environment. *Manufacturing & Service Operations Management* 17(1):34–51.

Angelus A, Porteus EL (2002) Simultaneous capacity and production management of short-life-cycle, produce-to-stock goods under stochastic demand. *Management Science* 48(3):399–413.

Aviv Y, Federgruen A (1997) Stochastic inventory models with limited production capacity and periodically varying parameters. *Probability in the Engineering and Informational Sciences* 11(1):107–135.

Aviv Y, Federgruen A (2001) Capacitated multi-item inventory systems with random and seasonally fluctuating demands: implications for postponement strategies. *Management science* 47(4):512–531.

Ban GY, Rudin C (2019) The big data newsvendor: Practical insights from machine learning. *Operations Research* 67(1):90–108.

Bray RL (2020) Operational transparency: Showing when work gets done. *Manufacturing & Service Operations Management* .

Cachon G (2001) Managing a retailer's shelf space, inventory, and transportation. *Manufacturing & Service Operations Management* 3(3):211–229.

Chen N, Gallego G (2018) A primal-dual learning algorithm for personalized dynamic pricing with an inventory constraint. *Available at SSRN 3301153* .

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

Cheung WC, Simchi-Levi D (2019) Sampling-based approximation schemes for capacitated stochastic inventory control models. *Mathematics of Operations Research* 44(2):668–692.

Choi TM, Wallace SW, Wang Y (2018) Big data analytics in operations management. *Production and Operations Management* 27(10):1868–1883.

DeCroix GA, Arreola-Risa A (1998) Optimal production and inventory policy for multiple products under resource constraints. *Management Science* 44(7):950–961.

Deshpande V, Pendem P (2020) Logistics performance, ratings, and its impact on customer purchasing behavior and sales in e-commerce platforms. *Ratings, and its impact on Customer Purchasing Behavior and Sales in E-commerce Platforms (September 21, 2020)* .

Ehrhardt R (1984) (s, s) policies for a dynamic inventory model with stochastic lead times. *Operations Research* 32(1):121–132.

Erlebacher SJ (2000) Optimal and heuristic solutions for the multi-item newsvendor problem with a single capacity constraint. *Production and Operations Management* 9(3):303–318.

Evans RV (1967) Inventory control of a multiproduct system with a limited production resource. *Naval Research Logistics Quarterly* 14(2):173–184.

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

Federgruen A, Zipkin P (1986a) An inventory model with limited production capacity and uncertain demands i. the average-cost criterion. *Mathematics of Operations Research* 11(2):193–207.

Federgruen A, Zipkin P (1986b) An inventory model with limited production capacity and uncertain demands ii. the discounted-cost criterion. *Mathematics of Operations Research* 11(2):208–215.

Feng Q, Luo S, Shanthikumar JG (2020) Integrating dynamic pricing with inventory decisions under lost sales. *Management Science* 66(5):2232–2247.

Fisher ML, Gallino S, Xu JJ (2019) The value of rapid delivery in omnichannel retailing. *Journal of Marketing Research* 56(5):732–748.

Gallego G, Moon I (1993) The distribution free newsboy problem: review and extensions. *Journal of the Operational Research Society* 44(8):825–834.

Glasserman P, Tayur S (1995) Sensitivity analysis for base-stock levels in multiechelon production-inventory systems. *Management Science* 41(2):263–281.

Goldberg DA, Katz-Rogozhnikov DA, Lu Y, Sharma M, Squillante MS (2016) Asymptotic optimality of constant-order policies for lost sales inventory models with large lead times. *Mathematics of Operations Research* 41(3):898–913.

Gürbüz MÇ, Moinzadeh K, Zhou YP (2007) Coordinated replenishment strategies in inventory/distribution systems. *Management Science* 53(2):293–307.

Huh WT, Janakiraman G, Muckstadt JA, Rusmevichientong P (2009a) An adaptive algorithm for finding the optimal base-stock policy in lost sales inventory systems with censored demand. *Mathematics of Operations Research* 34(2):397–416.

Huh WT, Janakiraman G, Muckstadt JA, Rusmevichientong P (2009b) Asymptotic optimality of order-up-to policies in lost sales inventory systems. *Management Science* 55(3):404–420.

Huh WT, Rusmevichientong P (2009) A nonparametric asymptotic analysis of inventory planning with censored demand. *Mathematics of Operations Research* 34(1):103–123.

Jasin S, Sinha A (2015) An lp-based correlated rounding scheme for multi-item ecommerce order fulfillment. *Operations Research* 63(6):1336–1351.

Karlin S, Scarf H (1958) *Inventory models of the Arrow-Harris-Marschak type with time lag* (Stanford University Press).

Karmarkar US (1981) The multiperiod multilocation inventory problem. *Operations Research* 29(2):215–228.

Kunnumkal S, Topaloglu H (2008) Using stochastic approximation methods to compute optimal base-stock levels in inventory control problems. *Operations Research* 56(3):646–664.

Lau HS, Hing-Ling Lau A (1995) The multi-product multi-constraint newsboy problem: Applications, formulation and solution. *Journal of operations management* 13(2):153–162.

Levi R, Janakiraman G, Nagarajan M (2008) A 2-approximation algorithm for stochastic inventory control models with lost sales. *Mathematics of Operations Research* 33(2):351–374.

Levi R, Roundy RO, Shmoys DB (2007) Provably near-optimal sampling-based policies for stochastic inventory control models. *Mathematics of Operations Research* 32(4):821–839.

Li MM, Liu X, Huang Y, Shi C (2018) Integrating empirical estimation and assortment personalization for e-commerce: A consider-then-choose model. *Available at SSRN 3247323* .

Li X, Zheng Y, Zhou Z, Zheng Z (2019) Demand prediction, predictive shipping, and product allocation for large-scale e-commerce. *Predictive Shipping, and Product Allocation for Large-Scale E-Commerce (March 12, 2019)* .

Liu S, He L, Max Shen ZJ (2020) On-time last-mile delivery: Order assignment with travel-time predictors. *Management Science* .

Mieghem JAV, Rudi N (2002) Newsvendor networks: Inventory management and capacity investment with discretionary activities. *Manufacturing & Service Operations Management* 4(4):313–335.

Morton TE (1969) Bounds on the solution of the lagged optimal inventory equation with no demand backlogging and proportional costs. *SIAM review* 11(4):572–596.

Nahmias S, Schmidt CP (1984) An efficient heuristic for the multi-item newsboy problem with a single constraint. *Naval Research Logistics Quarterly* 31(3):463–474.

Perakis G, Singhvi D, Spantidakis Y (2020) Leveraging the newsvendor for inventory distribution at a large fashion e-retailer with depth and capacity constraints. *Available at SSRN 3632459* .

Qi M, Mak HY, Shen ZJM (2020) Data-driven research in retail operations—a review. *Naval Research Logistics (NRL)* 67(8):595–616.

Rusmevichientong P, Shen ZJM, Shmoys DB (2010) Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Operations research* 58(6):1666–1680.

Sethi SP, Cheng F (1997) Optimality of (s, s) policies in inventory models with markovian demand. *Operations Research* 45(6):931–939.

Shi C, Chen W, Duenyas I (2016) Nonparametric data-driven algorithms for multiproduct inventory systems with censored demand. *Operations Research* 64(2):362–370.

Veinott Jr AF (1965) Optimal policy for a multi-product, dynamic, nonstationary inventory problem. *Management Science* 12(3):206–222.

Xin L, Goldberg DA (2016) Optimality gap of constant-order policies decays exponentially in the lead time for lost sales models. *Operations Research* 64(6):1556–1565.

Xu PJ, Allgor R, Graves SC (2009) Benefits of reevaluating real-time order fulfillment decisions. *Manufacturing & Service Operations Management* 11(2):340–355.

Zhang B, Du S (2010) Multi-product newsboy problem with limited capacity and outsourcing. *European Journal of Operational Research* 202(1):107–113.

Zhang H, Chao X, Shi C (2020) Closing the gap: A learning algorithm for lost-sales inventory systems with lead times. *Management Science* 66(5):1962–1980.

Zipkin P (2008a) Old and new methods for lost-sales inventory systems. *Operations Research* 56(5):1256–1263.

Zipkin P (2008b) On the structure of lost-sales inventory models. *Operations research* 56(4):937–944.

Method		C_1, M_1	C_1, M_2	C_1, M_3
BSA		92	71	42
OMA		43	41	27
Train	BSA-PL	91	67	40
	OMA-PL	69	45	30
Test	BSA-PL	88	60	32
	OMA-PL	58	33	22
True	BSA-PL	87	62	36
	OMA-PL	80	46	28
10% Err.	BSA-PL	89	68	34
	OMA-PL	80	46	28
20% Err.	BSA-PL	86	63	38
	OMA-PL	59	36	25
40% Err.	BSA-PL	87	54	28
	OMA-PL	51	30	22
50% Err.	BSA-PL	84	50	21
	OMA-PL	53	31	21

Table 5: Number of days requiring extra shipment: BSA, OMA, BSA-PL, OMA-PL stand for the base stock algorithm, the online myopic algorithm, the base stock algorithm with predictive learning, the online myopic algorithm with predictive learning respectively.

Appendix

Proof of Proposition 3.3

By definition of L_T and $\tilde{L}_T$, we know

$$\tilde{L}_T^{\pi^*}(\boldsymbol{Y}|\boldsymbol{x}_1) - \boldsymbol{p}^\top \mathbb{E}^{\pi^*}[\boldsymbol{x}_{T+1}] = L_T^{\pi^*}(\boldsymbol{Y}|\boldsymbol{x}_1) \leq \tilde{L}_T^{\tilde{\pi}^*}(\boldsymbol{Y}|\boldsymbol{x}_1) \leq \tilde{L}_T^{\pi^*}(\boldsymbol{Y}|\boldsymbol{x}_1).$$

Hence

$$\frac{\tilde{L}_T^{\tilde{\pi}^*}(\boldsymbol{Y}|\boldsymbol{x}_1)}{\tilde{L}_T^{\pi^*}(\boldsymbol{Y}|\boldsymbol{x}_1)} \geq 1 - \frac{\boldsymbol{p}^\top \mathbb{E}^{\pi^*}[\boldsymbol{x}_{T+1}]}{\tilde{L}_T^{\pi^*}(\boldsymbol{Y}|\boldsymbol{x}_1)}.$$

From Assumption 3.2 we know that the cost at each step t is positive, hence $\tilde{L}_T^{\pi^*}(\boldsymbol{Y}|\boldsymbol{x_1}) \to +\infty$. We conclude that $\lim_{T\to+\infty} \boldsymbol{p}^\top \mathbb{E}^{\pi^*}[\boldsymbol{x}_{T+1}]/\tilde{L}_T^{\pi^*}(\boldsymbol{Y}|\boldsymbol{x}_1) = 0$.

Proof of Proposition 3.5

Without loss of generality we assume that $\bar{\boldsymbol{y}}^* > 0$. Since $F_{\mathcal{P}}$ is continuous, for every compact set $\{\boldsymbol{y} \,|\, \boldsymbol{y} \in \mathbb{R}^n_+, \boldsymbol{x} \le \boldsymbol{y} \le \bar{\boldsymbol{y}}^*, \boldsymbol{c}^\top \boldsymbol{y} \le W\}$ there must be a minimizer $\tilde{\boldsymbol{y}}^*_{\boldsymbol{x}}$. If $\boldsymbol{c}^\top \tilde{\boldsymbol{y}}^*_{\boldsymbol{x}} < W$, and we consider the mapping $P(\lambda, \boldsymbol{x}, \boldsymbol{y}) := \lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}$ and function $g(\lambda) := F_{\mathcal{P}}(\lambda \tilde{\boldsymbol{y}}^*_{\boldsymbol{x}} + (1-\lambda)\bar{\boldsymbol{y}}^*)$. If $g(\cdot)$ achieve a local maximum at $\tilde{\lambda}^* \in (0,1)$, we are able to choose $\lambda_1 < \tilde{\lambda}^* < \lambda_2$ such that for $t = (\lambda_2 - \tilde{\lambda}^*)/(\lambda_2 - \lambda_1)$, $\boldsymbol{y}_1 = P(\lambda_1, \tilde{\boldsymbol{y}}^*_{\boldsymbol{x}}, \bar{\boldsymbol{y}}^*)$, and $\boldsymbol{y}_2 = P(\lambda_2, \tilde{\boldsymbol{y}}^*_{\boldsymbol{x}}, \bar{\boldsymbol{y}}^*)$ we have

$$F_{\mathcal{P}}(P(\tilde{\lambda}^*, \tilde{\boldsymbol{y}}^*_{\boldsymbol{x}}, \bar{\boldsymbol{y}}^*)) = F_{\mathcal{P}}(t\boldsymbol{y}_1 + (1-t)\boldsymbol{y}_2) > tF_{\mathcal{P}}(\boldsymbol{y}_1) + (1-t)F_{\mathcal{P}}(\boldsymbol{y}_2),$$

which contradicts the convexity assumption. Therefore $g(\lambda)$ can only be decreasing in $[0,1]$. Also, we know that there exists λ_3 such that $\boldsymbol{c}^\top P(\lambda_3, \tilde{\boldsymbol{y}}^*_{\boldsymbol{x}}, \boldsymbol{y}^*) = W$, hence $P(\lambda_3, \tilde{\boldsymbol{y}}^*_{\boldsymbol{x}}, \boldsymbol{y}^*)$ will also be a minimizer of (5).

Proof of Proposition 4.1

Since $\boldsymbol{x}_t = (\boldsymbol{y}^* - \boldsymbol{d}_{t-1})_+ \le \boldsymbol{y}^*$, the objective function has the relation that

$$\begin{aligned} L_T(\boldsymbol{Y}|\boldsymbol{x}_1) &= \mathbb{E}\left[\sum_{t=1}^T F_{\mathcal{P}}(\boldsymbol{y}_t)\right] - \boldsymbol{p}^\top \boldsymbol{x}_1 \\ &\ge \mathbb{E}\sum_{t=1}^T F_{\mathcal{P}}(\boldsymbol{y}^*) - \boldsymbol{p}^\top \boldsymbol{x}_1 = \tilde{L}_T^{\pi_{bs}}(\boldsymbol{Y}|\boldsymbol{x}_1). \end{aligned}$$

By taking infimum we know that π_{bs} is the optimal policy.

Proof of Proposition 4.3

We need the following lemmas before starting to prove this proposition.

Lemma 5.1 *For sequence $\{W^{(k)}\}_{k\in\mathbb{N}}$ such that $W^{(k)} \le W^{(k+1)} \le C$, denote the optimal solution to $OPT(W^{(k)}, \boldsymbol{0})$ as $\boldsymbol{y}^*_k$. We have $\boldsymbol{y}_k$ being monotone such that $\boldsymbol{y}^*_k \le \boldsymbol{y}^*_{k+1}$.*

Lemma 5.2 *For $OPT(W^{(k)}, \boldsymbol{0})$, if there exists i such that $y^*_{k,i} = 0$, then there also exists l such that $y^*_{k,l} > 0$ and $\lambda < 0$ such that $\nabla_l F_{\mathcal{P}}(\boldsymbol{y}^*_k) = \lambda c_l$ and $\nabla_i F_{\mathcal{P}}(\boldsymbol{y}^*_k) \ge \lambda c_i$.*

The next lemma state the relationship between $\boldsymbol{y}^*_t$ and $\boldsymbol{y}^{(mp)}_t$ given that the current capacity constraint being $W^{(t)}$. Recall that $\boldsymbol{y}^{(mp)}_t$ is the solution of (8) and $W^{(t)} = C \wedge (M_t - \boldsymbol{c}^\top \boldsymbol{x}_t)$ in its context.

Lemma 5.3 *For a given realization of $\boldsymbol{x}_t$, there exists solution $\boldsymbol{y}_t^{(mp)}$ such that for all i such that $x_{t,i} > y^*_{t,i}$, we have $y_{t,i}^{(mp)} = x_{t,i}$. Also, $y_{t,i}^{(mp)} \le y^*_{t,i}$ for any i such that $x_{t,i} \le y^*_{t,i}$.*

Lemma 5.4 *Under Assumption 3.4, if the myopic policy is adopted, $\boldsymbol{y}_t^{(mp)} \le \boldsymbol{y}^*$ for any $t > 0$.*

Then, we start to prove this proposition. For each realization of $\{\boldsymbol{d}_t\}_{t\le T}$, and $\{M_t\}_{t\le T}$ satisfying Condition 4.2 , recall that the objection function could be written as

$$L_T(\boldsymbol{Y}|\boldsymbol{x}_1) = \sum_{t=1}^{T} \mathbb{E}[F_{\mathcal{P}}(\boldsymbol{x}_t + \boldsymbol{z}_t)] - \boldsymbol{p}^\top \boldsymbol{x}_1 = \sum_{t=1}^{T} \mathbb{E}[F_{\mathcal{P}}((\boldsymbol{y}_{t-1} - \boldsymbol{d}_{t-1})_+ + \boldsymbol{z}_t)] - \boldsymbol{p}^\top \boldsymbol{x}_1 \tag{11}$$

Since $\boldsymbol{p}^\top \boldsymbol{x}_1$ is fixed, with a little abuse of notation we overload the definition $L_T(\boldsymbol{Y}|\boldsymbol{x}_1)$ in the following way

$$L_T(\boldsymbol{Y}|\boldsymbol{x}_1) = \sum_{t=1}^{T} \mathbb{E}[F_{\mathcal{P}}(\boldsymbol{x}_t + \boldsymbol{z}_t)] = \sum_{t=1}^{T} \mathbb{E}[F_{\mathcal{P}}((\boldsymbol{y}_{t-1} - \boldsymbol{d}_{t-1})_+ + \boldsymbol{z}_t)] \tag{12}$$

From Lemma 5.4 we know that the solution always falls into the feasible region ($\boldsymbol{y}_t^{(mp)} \le \boldsymbol{y}^*$), and if the shipping constraint has slacks, we can always order up to $\boldsymbol{y}^*$. Our goal is to prove that the optimal policy features the same allocation rule as the myopic policy, and we are going to prove it by induction. When $T = 1, 2$, we know that the optimal solution is to order to $\boldsymbol{y}^*$ because of the dynamic $\boldsymbol{M}$ ensured by Condition 4.2. We then begin our induction argument. Assume the myopic policy is optimal when $T \le n$. For the case $T = n+1$, at the first round we could be given two scenarios: the first one is that $\mathcal{A}_1 = \{C - \boldsymbol{c}^\top \boldsymbol{x}_1 > M_1\}$, and the second one is that $\bar{\mathcal{A}}_1 = \{C - \boldsymbol{c}^\top \boldsymbol{x}_1 \le M_1\}$. In the first scenario, whichever $\boldsymbol{z}_1$ we choose, it won't affect our ability to choose the optimal solution in $t \in [2, T]$, hence the optimal solution is again myopic. More specifically, for the first scenario, since $\mathcal{A}_1 = \{C - \boldsymbol{c}^\top \boldsymbol{x}_1 > M_1\}$ already taken place, we know that $\bar{\mathcal{A}}_2 = \{C - \boldsymbol{c}^\top \boldsymbol{x}_2 \le M_2\}$ happens almost surely, which means we have the total expected cost being

$$L_{T+1}(\boldsymbol{Y}|\boldsymbol{x}_1, \mathcal{A}_1) = \mathbb{E}[F_{\mathcal{P}}(\boldsymbol{x}_1 + \boldsymbol{z}_1)|\mathcal{A}_1] + \mathbb{E}[L_T(\boldsymbol{Y}|\boldsymbol{x}_2)|\bar{\mathcal{A}}_2].$$

From induction we know that $\mathbb{E}[L_T(\boldsymbol{Y}|\boldsymbol{x}_2)|\bar{\mathcal{A}}_2]$ will attain maximal for $\boldsymbol{z_1}$ such that $\boldsymbol{x}_1 + \boldsymbol{z}_1 \le \boldsymbol{y}^*$, and from Lemmas 5.3 and 5.4 we know under the myopic algorithm the both terms above will achieve maximum hence being optimal.

Under the second scenario $\bar{\mathcal{A}}_1 = \{C - \boldsymbol{c}^\top \boldsymbol{x}_1 \le M_1\}$, if we define event $\mathcal{A}_2 := \{C - \boldsymbol{c}^\top(\boldsymbol{x}_2) > M_2\}$ and $\bar{\mathcal{A}}_2 := \{C - \boldsymbol{c}^\top(\boldsymbol{x}_2) \le M_2\}$, we know that

$$L_{T+1}(\boldsymbol{Y}|\boldsymbol{x}_1, \bar{\mathcal{A}}_1) = \mathbb{E}[L_{T+1}(\boldsymbol{Y}|\boldsymbol{x}_1, \bar{\mathcal{A}}_1)|\bar{\mathcal{A}}_2]P(\bar{\mathcal{A}}_2) + \mathbb{E}[L_{T+1}(\boldsymbol{Y}|\boldsymbol{x}_1, \bar{\mathcal{A}}_1)|\mathcal{A}_2]P(\mathcal{A}_2).$$

For the first term on the RHS, we can use induction argument for the loss from period 2 to $n+1$ to show the optimality. For the second term on the RHS, notice that

$$\mathbb{E}[L_{T+1}(\boldsymbol{Y}|\boldsymbol{x}_1, \bar{\mathcal{A}}_1)|\mathcal{A}_2] = \mathbb{E}[F_{\mathcal{P}}(\boldsymbol{x}_1 + \boldsymbol{z}_1) + F_{\mathcal{P}}(\boldsymbol{x}_2 + \boldsymbol{z}_2)]|\mathcal{A}_2] + \mathbb{E}[L_{T-1}(\boldsymbol{Y}|\boldsymbol{x}_3)|\mathcal{A}_2].$$

Since $\mathcal{A}_2$ implies $\bar{\mathcal{A}}_3 = \{C - \boldsymbol{c}^\top \boldsymbol{x}_3 \le M_3\}$, we know we have enough shipping capacity at $t = 3$. Therefore, from the same analysis above, the feasible decision $\boldsymbol{z}_2$ won't affect $\mathbb{E}[L_{T-1}(\boldsymbol{Y}|\boldsymbol{x}_3)|\mathcal{A}_2]$. From our base condition and induction hypothesis, we know that the both terms above are achieving maximum, hence myopic is optimal.

Proof of Proposition 4.4

Notice that

$$\begin{aligned}
&\tilde{L}'_T(\boldsymbol{Y}|\boldsymbol{x}_1) - \mathbb{E}\left[\boldsymbol{p}^\top \boldsymbol{x}_{T+1}\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T}\left(\boldsymbol{p}^\top(\boldsymbol{y}_t - \boldsymbol{x}_t)_+ + \boldsymbol{p}'^\top(\boldsymbol{x}_t - \boldsymbol{y}_t)_+ + \sum_{i=1}^{n} g_i(y_{i,t}, d_{i,t})\right) - \boldsymbol{p}^\top \boldsymbol{x}_{T+1}\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T}\left(\boldsymbol{p}^\top(\boldsymbol{y}_t - \boldsymbol{x}_t) + (\boldsymbol{p}' + \boldsymbol{p})^\top(\boldsymbol{x}_t - \boldsymbol{y}_t)_+ + \sum_{i=1}^{n} g_i(y_{i,t}, d_{i,t})\right) - \boldsymbol{p}^\top \boldsymbol{x}_{T+1}\right] \\
&= L'_T(\boldsymbol{Y}|\boldsymbol{x}_1) + (\boldsymbol{p}' + \boldsymbol{p})^\top \mathbb{E}\left[\sum_{t=1}^{T}(\boldsymbol{x}_t - \boldsymbol{y}_t)_+\right].
\end{aligned}$$

Consider optimal policy π'^* for $L'_T(\boldsymbol{Y}|\boldsymbol{x}_1)$, we know that

$$\begin{aligned}
&\tilde{L}'^{\pi'^*}_T(\boldsymbol{Y}|\boldsymbol{x}_1) - \boldsymbol{p}^\top \mathbb{E}^{\pi'^*}[\boldsymbol{x}_{T+1}] - (\boldsymbol{p}' + \boldsymbol{p})^\top \mathbb{E}^{\pi'^*}\left[\sum_{t=1}^{T}(\boldsymbol{x}_t - \boldsymbol{y}_t)_+\right] \\
&= L'^{\pi'^*}_T(\boldsymbol{Y}|\boldsymbol{x}_1) \le \tilde{L}'^{\tilde{\pi}'^*}_T(\boldsymbol{Y}|\boldsymbol{x}_1) \le \tilde{L}'^{\pi'^*}_T(\boldsymbol{Y}|\boldsymbol{x}_1).
\end{aligned} \tag{13}$$

Hence we have

$$\frac{\tilde{L}'^{\tilde{\pi}'^*}_T(\boldsymbol{Y}|\boldsymbol{x}_1)}{\tilde{L}'^{\pi'^*}_T(\boldsymbol{Y}|\boldsymbol{x}_1)} \geq 1 - \frac{\boldsymbol{p}^\top \mathbb{E}^{\pi'^*}[\boldsymbol{x}_{T+1}] + (\boldsymbol{p}' + \boldsymbol{p})^\top \mathbb{E}^{\pi'^*}\left[\sum_{t=1}^T (\boldsymbol{x}_t - \boldsymbol{y}_t)_+\right]}{\tilde{L}'^{\pi'^*}_T(\boldsymbol{Y}|\boldsymbol{x}_1)}. \tag{14}$$

Proof of Lemma 5.1

If $y^*_{k,i} = 0$, we have $y^*_{k+1,i} \geq y^*_{k,i}$. Therefore it suffices to focus on the positive entries. From first-order KKT condition we know that for every i such that $y^*_{k,i} > 0$, we have

$$\nabla_i F_{\mathcal{P}}(\boldsymbol{y}^*_k) = f'_i(\boldsymbol{y}^*_{k,i}) = \lambda c_i. \tag{15}$$

Then, consider the same setting for $W^{(k+1)}$, we know the above equation also holds for positive entry of $\boldsymbol{y}^*_{k+1}$, and denote the corresponding multiplier to be $\tilde{\lambda}$. Sicne $W^{(k+1)} > W^{(k)}$ and $f'_i(x)$ is increasing, in order to achieve the KKT condition $f'_i(y^*_{k+1,i}) = \tilde{\lambda} c_i$, we can only have $\boldsymbol{y}^*_k \leq \boldsymbol{y}^*_{k+1}$.

Proof of Lemma 5.2

From Proposition 3.5 we know that the optimal solution satisfy $\boldsymbol{c}^\top \boldsymbol{y}^* = C$ hence there must be an interior point $y^*_j > 0$. From first order KKT condition, we know that there must be a λ such that $\nabla_k F_{\mathcal{P}}(\boldsymbol{y}^*) = \lambda c_k$ for all k such that $y^*_k > 0$. From $\boldsymbol{y}^* \leq \bar{\boldsymbol{y}}^*$ and $F_{\mathcal{P}}$ being convex we know that $\lambda < 0$. If $y^*_i = 0$ and $\nabla_i F_{\mathcal{P}}(\boldsymbol{y}^*) < \lambda c_i$, then we can find a vector $\boldsymbol{v}$ such that $v_i = c_k/\sqrt{c_k^2 + c_i^2}$, $v_k = -c_i/\sqrt{c_k^2 + c_i^2}$, and zero on other entries. By noticing that $\boldsymbol{v}$ is a feasible direction and

$$\nabla F_{\mathcal{P}}(\boldsymbol{y}^*)^T \boldsymbol{v} = (\nabla_i F_{\mathcal{P}}(\boldsymbol{y}^*) c_k - \nabla_k F_{\mathcal{P}}(\boldsymbol{y}^*) c_i)/\sqrt{c_k^2 + c_i^2} < 0,$$

which contradicts the necessary condition of $\boldsymbol{y}^*$. Hence we must have $\nabla_i F_{\mathcal{P}}(\boldsymbol{y}^*) \geq \lambda c_i$.

Proof of Lemma 5.3

Define $\boldsymbol{z}^{(mp)}_t := \boldsymbol{y}^{(mp)}_t - \boldsymbol{x}_t$. The goal is to prove that if for index i we have $x_{t,i} \geq y^*_i$ and $z^{(mp)}_{t,i} > 0$, we can always derive another solution $\tilde{\boldsymbol{z}}^{(mp)}_t$ such that $\tilde{z}^{(mp)}_{t,i} = 0$.

For simplicity of notation, denote $\boldsymbol{x} = \boldsymbol{x}_t$, $\boldsymbol{z} = \boldsymbol{z}_t^{(mp)}$ and $\boldsymbol{y} = \boldsymbol{y}^*$. We reorder the index such that there exist k_1, k_2, and k_3 such that for $i \in (0, k_1]$, we have $x_i \geq y_i$ and $z_i > 0$; for $i \in (k_1, k_2]$, we have $x_i \geq y_i$ and $z_i = 0$; for $i \in (k_2, k_3]$, we have $x_i < y_i$ and $x_i + z_i > y_i$; and for $i \in (k_3, n]$, we have $x_i < y_i$ and $x_i + z_i \leq y_i$. Denote $\boldsymbol{x}^{(1)}$, $\boldsymbol{y}^{(1)}$ and $\boldsymbol{z}^{(1)}$ such that $x_i^{(1)} = x_i$ if $i \in (0, k_1]$ and $x_i^{(1)} = 0$ otherwise, and define $\boldsymbol{y}^{(1)}$ and $\boldsymbol{z}^{(1)}$ accordingly. In the same way we denote $\boldsymbol{x}^{(2)}$, $\boldsymbol{y}^{(2)}$ and $\boldsymbol{z}^{(2)}$ such that $x_i^{(2)} = x_i$ for $i \in (k_1, k_2]$. If we do the same thing for $\boldsymbol{x}^{(3)}$ and $\boldsymbol{x}^{(4)}$ (also for $\boldsymbol{y}$ and $\boldsymbol{z}$), we have $\boldsymbol{x} = \sum_{k=1}^4 \boldsymbol{x}^{(k)}$, $\boldsymbol{y} = \sum_{k=1}^4 \boldsymbol{y}^{(k)}$ and $\boldsymbol{z} = \sum_{k=1}^4 \boldsymbol{z}^{(k)}$. Denote $\boldsymbol{u}^{(4)} = \boldsymbol{y}^{(4)} - (\boldsymbol{x}^{(4)} + \boldsymbol{z}^{(4)})$, $\alpha = |\boldsymbol{c}^\top \boldsymbol{z}^{(1)}|/|\boldsymbol{c}^\top \boldsymbol{u}^{(4)}|$ and $\boldsymbol{u} = -\boldsymbol{z}^{(1)} + \alpha \boldsymbol{u}^{(4)}$. Next, we prove that $\boldsymbol{z} + \boldsymbol{u}$ is also in the feasible region of problem (8), which is obvious by noticing that $\boldsymbol{z} + \boldsymbol{u} \geq 0$ and $\boldsymbol{c}^\top(\boldsymbol{z} + \boldsymbol{u}) = M$. The last thing to do is to prove that $F_{\mathcal{P}}(\boldsymbol{x} + \boldsymbol{z} + \boldsymbol{u}) \leq F_{\mathcal{P}}(\boldsymbol{x} + \boldsymbol{z})$. If we define $g(\gamma) = F_{\mathcal{P}}(\boldsymbol{x} + \boldsymbol{z} + \gamma \boldsymbol{u})$, it suffices to show that $g'(\gamma) = \boldsymbol{u}^\top \nabla F_{\mathcal{P}}(\boldsymbol{x} + \boldsymbol{z} + \gamma \boldsymbol{u}) \leq 0$ for $\gamma \in [0, 1]$. By observing that

$$\begin{aligned}
\sum_{k=1}^{4} \boldsymbol{c}^\top(\boldsymbol{y}^{(k)} - (\boldsymbol{x}^{(k)} + \boldsymbol{z}^{(k)})) &= 0 \\
\boldsymbol{c}^\top(\boldsymbol{y}^{(1)} - (\boldsymbol{x}^{(1)} + \boldsymbol{z}^{(1)})) &< 0 \\
\boldsymbol{c}^\top(\boldsymbol{y}^{(2)} - (\boldsymbol{x}^{(2)} + \boldsymbol{z}^{(2)})) &\leq 0 \\
\boldsymbol{c}^\top(\boldsymbol{y}^{(3)} - (\boldsymbol{x}^{(3)} + \boldsymbol{z}^{(3)})) &< 0 \\
\boldsymbol{c}^\top(\boldsymbol{y}^{(4)} - (\boldsymbol{x}^{(4)} + \boldsymbol{z}^{(4)})) &> 0.
\end{aligned} \tag{16}$$

We know that $\boldsymbol{c}^\top(\boldsymbol{y}^{(4)} - (\boldsymbol{x}^{(4)} + \boldsymbol{z}^{(4)})) + \boldsymbol{c}^\top(\boldsymbol{y}^{(1)} - (\boldsymbol{x}^{(1)} + \boldsymbol{z}^{(1)})) > 0$. Since

$$\boldsymbol{c}^\top(\boldsymbol{y}^{(4)} - (\boldsymbol{x}^{(4)} + \boldsymbol{z}^{(4)})) = \boldsymbol{c}^\top \boldsymbol{u}^{(4)}, \;\; \boldsymbol{c}^\top(\boldsymbol{y}^{(1)} - (\boldsymbol{x}^{(1)} + \boldsymbol{z}^{(1)})) \leq -\boldsymbol{c}^\top \boldsymbol{z}^{(1)}, \tag{17}$$

we know $\boldsymbol{c}^\top \boldsymbol{u}^{(4)} - \boldsymbol{c}^\top \boldsymbol{z}^{(1)} > 0$, hence $\alpha < 1$. Since $\boldsymbol{y}$ is the optimal solution, for all $y_i > 0$ there exist $\lambda \leq 0$ such that $\nabla_i F_{\mathcal{P}}(\boldsymbol{y}) = \lambda c_i$. For the case when $i \in (0, k_1]$, from $(\boldsymbol{x} + \boldsymbol{z} + \gamma \boldsymbol{u})_i > y_i$ and Lemma 5.2 we know that $\nabla_i F_{\mathcal{P}}(\boldsymbol{x} + \boldsymbol{z} + \gamma \boldsymbol{u}) \geq \lambda c_i$ no matter $y_i = 0$ or not. For the case $i \in (k_3, n]$, since $\alpha, \gamma < 1$, we know that component-wisely

$$x_i^{(4)} + z_i^{(4)} + \gamma(-z_i^{(1)} + \alpha u_i^{(4)}) = x_i^{(4)} + z_i^{(4)} + \gamma \alpha u_i^{(4)} < y_i^{(4)}. \tag{18}$$

Also, it is equivalent to say that we have $0 \leq (\boldsymbol{x} + \boldsymbol{z} + \gamma \boldsymbol{u})_i < y_i$, and we have $\nabla_i F_{\mathcal{P}}(\boldsymbol{x} + \boldsymbol{z} + \gamma \boldsymbol{u}) \leq \lambda c_i$ for $i \in (k_3, n]$. To summarize,

$$\begin{aligned}
\nabla_i F_{\mathcal{P}}(\boldsymbol{x} + \boldsymbol{z} + \gamma \boldsymbol{u}) &\geq \lambda c_i, \;\; i \in (0, k_1] \\
\nabla_i F_{\mathcal{P}}(\boldsymbol{x} + \boldsymbol{z} + \gamma \boldsymbol{u}) &\leq \lambda c_i, \;\; i \in (k_3, n].
\end{aligned} \tag{19}$$

Henceforth, $g'(\gamma) = \boldsymbol{u}^\top \nabla F_{\mathcal{P}}(\boldsymbol{x}+\boldsymbol{z}+\gamma\boldsymbol{u}) \leq \boldsymbol{u}^\top \lambda \boldsymbol{c} = \lambda \boldsymbol{c}^\top(\alpha \boldsymbol{u}^{(4)} - \boldsymbol{z}^{(1)}) = 0$ on $\gamma \in (0,1)$. Therefore, $\boldsymbol{z}+\boldsymbol{u}$ is also a optimal solution for (8), and it features the property that for any i such that $x_i \geq y_i$, $(\boldsymbol{z}+\boldsymbol{u})_i = 0$.

Next, we focus on creating the solution $\tilde{\boldsymbol{z}}$ that not only satisfy the property above but also features $(\boldsymbol{x}+\tilde{\boldsymbol{z}})_i \leq y_i$ when $x_i \leq y_i$. The methodology is the same as above. For simplicity, denote $\boldsymbol{v} = \boldsymbol{z}+\boldsymbol{u}$. Based on the solution $\boldsymbol{v}$, we know that

$$\begin{aligned}\sum_{k=2}^{4} \boldsymbol{c}^\top(\boldsymbol{y}^{(k)} - (\boldsymbol{x}^{(k)} + \boldsymbol{v}^{(k)})) &= 0 \\ \boldsymbol{c}^\top(\boldsymbol{y}^{(2)} - (\boldsymbol{x}^{(2)} + \boldsymbol{v}^{(2)})) &\leq 0 \\ \boldsymbol{c}^\top(\boldsymbol{y}^{(3)} - (\boldsymbol{x}^{(3)} + \boldsymbol{v}^{(3)})) &< 0 \\ \boldsymbol{c}^\top(\boldsymbol{y}^{(4)} - (\boldsymbol{x}^{(4)} + \boldsymbol{v}^{(4)})) &> 0.\end{aligned} \tag{20}$$

By letting $\boldsymbol{u}^{(3)} = \boldsymbol{y}^{(3)} - (\boldsymbol{x}^{(3)}+\boldsymbol{v}^{(3)})$, $\boldsymbol{u}^{(4)} = \boldsymbol{y}^{(4)} - (\boldsymbol{x}^{(4)}+\boldsymbol{v}^{(4)})$, $\alpha = |\boldsymbol{c}^\top \boldsymbol{u}^{(3)}|/|\boldsymbol{c}^\top \boldsymbol{u}^{(4)}|$, and $\boldsymbol{u} = -\boldsymbol{u}^{(3)} + \alpha \boldsymbol{u}^{(4)}$. Clearly, $\boldsymbol{v}+\boldsymbol{u}$ is a feasible solution on the boundary because $\boldsymbol{v}+\boldsymbol{u} \geq 0$ and $\boldsymbol{c}^\top \boldsymbol{u} = 0$. Then we want to check the gradient. From (20) we know that

$$\boldsymbol{c}^\top(\boldsymbol{y}^{(4)} - (\boldsymbol{x}^{(4)} + \boldsymbol{v}^{(4)})) + \boldsymbol{c}^\top(\boldsymbol{y}^{(3)} - (\boldsymbol{x}^{(3)} + \boldsymbol{v}^{(3)})) \geq 0, \tag{21}$$

hence $\alpha \leq 1$. Then from the same argument, for $\gamma \in (0,1)$, it suffices to show that $\boldsymbol{u}^\top \nabla F_{\mathcal{P}}(\boldsymbol{x}+\boldsymbol{v}+\gamma\boldsymbol{u}) \leq 0$. For $i \in (k_2, k_3]$ we know that $x_i + \boldsymbol{v}_i + \gamma \boldsymbol{u}_i \geq \boldsymbol{y_i}$, hence from Lemma 5.2 we know that

$$\nabla_i F_{\mathcal{P}}(\boldsymbol{x}+\boldsymbol{v}+\gamma\boldsymbol{u}) \geq \lambda c_i, \ \ i \in (k_2, k_3] \tag{22}$$

Again since $\alpha \leq 1$, we know that $0 \leq (\boldsymbol{x}+\boldsymbol{z}+\gamma\boldsymbol{u})_i \leq y_i$, hence

$$\nabla_i F_{\mathcal{P}}(\boldsymbol{x}+\boldsymbol{v}+\gamma\boldsymbol{u}) \leq \lambda c_i, \ \ i \in (k_3, n] \tag{23}$$

Therefore we have $\boldsymbol{u}^\top \nabla F_{\mathcal{P}}(\boldsymbol{x}+\boldsymbol{v}+\gamma\boldsymbol{u}) \leq \boldsymbol{u}^\top \lambda \boldsymbol{c} = 0$.

Proof of Lemma 5.4

If there exist s and i such that $x_{s,i} > y_i^*$, for this i, we choose $t = \inf\{s | x_{s,i} > y_i^*\}$. Since we assume the initial position $\boldsymbol{x} < \boldsymbol{y}^*$, t has to be strictly positive. Also, since $x_{t,i} > y_i^*$, we know that $y_i^{(mp)} > y_i^*$. From Lemma 5.3 we know that this can happen only when $x_{t-1,i} > y_i^*$, which contradicts the minimal assumption of t.