informs
https://pubsonline.informs.org/journal/opre

OPERATIONS RESEARCH
Vol. 73, No. 6, November–December 2025, pp. 2914–2932
ISSN 0030-364X (print), ISSN 1526-5463 (online)

Contextual Areas

E-Commerce Order Fulfillment Problem with Limited Time Window

Quan Zhou,[a,*] Mehmet Gümüş,[a] Sentao Miao[b]

[a] Desautels Faculty of Management, McGill University, Montreal, Quebec H3A 1G5, Canada; [b] Leeds School of Business, University of Colorado Boulder, Boulder, Colorado 80309

*Corresponding author

Contact: quan.zhou3@mail.mcgill.ca, https://orcid.org/0009-0007-2737-6293 (QZ); mehmet.gumus@mcgill.ca, https://orcid.org/0000-0003-3814-896X (MG); sentao.miao@colorado.edu, https://orcid.org/0000-0002-0380-0797 (SM)

Received: August 22, 2023
Revised: June 4, 2024; January 7, 2025; May 20, 2025
Accepted: May 30, 2025
Published Online in Articles in Advance: July 11, 2025

Area of Review: Operations and Supply Chains

https://doi.org/10.1287/opre.2023.0453

Abstract. We explore the optimization of the middle-mile fulfillment process in the context of e-commerce. In collaboration with a prominent e-commerce retailer in North America specializing in electronics and computer products, we develop a stochastic optimization problem to demonstrate how an efficient middle mile can alleviate strain on the critical last mile, leading to cost reduction and improved performance. First, we prove that the optimal policy is of a state-dependent threshold type. However, computing such thresholds is notoriously difficult because of the curse of dimensionality. We then introduce a Lagrangian relaxation-based policy (referred to as threshold Lagrangian relaxation (tLR)) as a heuristic approach for fulfillment decisions and prove its performance guarantee. To validate our findings, we utilize both synthetically generated data sets and data provided by our partner e-commerce retailer, and we conduct two numerical studies comparing the tLR policy with benchmark policies. These studies validate our findings on the performance guarantees of the proposed heuristic approach, highlighting the benefits of multiperiod fulfillment windows and the cost-reducing capabilities of the tLR policy. We conclude by emphasizing the importance of dynamic fulfillment strategies and the considerations that e-commerce companies should take into account when selecting their fulfillment policies.

Funding: The work of M. Gümüş was supported in part by the Natural Sciences and Engineering Research Council of Canada [Grant 217601, NSERC RGPIN-2025-05765], Fonds de Recherche du Québec [Grant FRQ 266584], Institute for Data Valorization [Grant IVADO G254088], and Desautels Chair-2 [Grant 265284].

Supplemental Material: All supplemental materials, including the code, data, and files required to reproduce the results, are available at https://doi.org/10.1287/opre.2023.0453.

Keywords: e-commerce • middle mile • fulfillment optimization

1. Introduction

E-commerce has consistently gained a larger portion of the retail market because of various factors, such as shifting consumer buying patterns and technological advancements. According to eMarketer (2023), in 2015, e-commerce accounted for 7.5% of global retail sales, generating over $1.5 trillion in revenue. Fast forward to 2021, this percentage has risen to 20%, equivalent to slightly over $4.9 trillion. As the demand for e-commerce continues to surge, the decision regarding how to fulfill online orders becomes crucial as it can greatly impact the profit margins of e-commerce companies. To illustrate, Amazon experienced a significant boost in annual revenue, reaching approximately $470 billion in 2021, a growth of 21.7% compared with the previous year. However, its fulfillment costs rose at a faster rate, surpassing $75 billion in 2021, an increase of over 28% of the prior year (Amazon 2024).

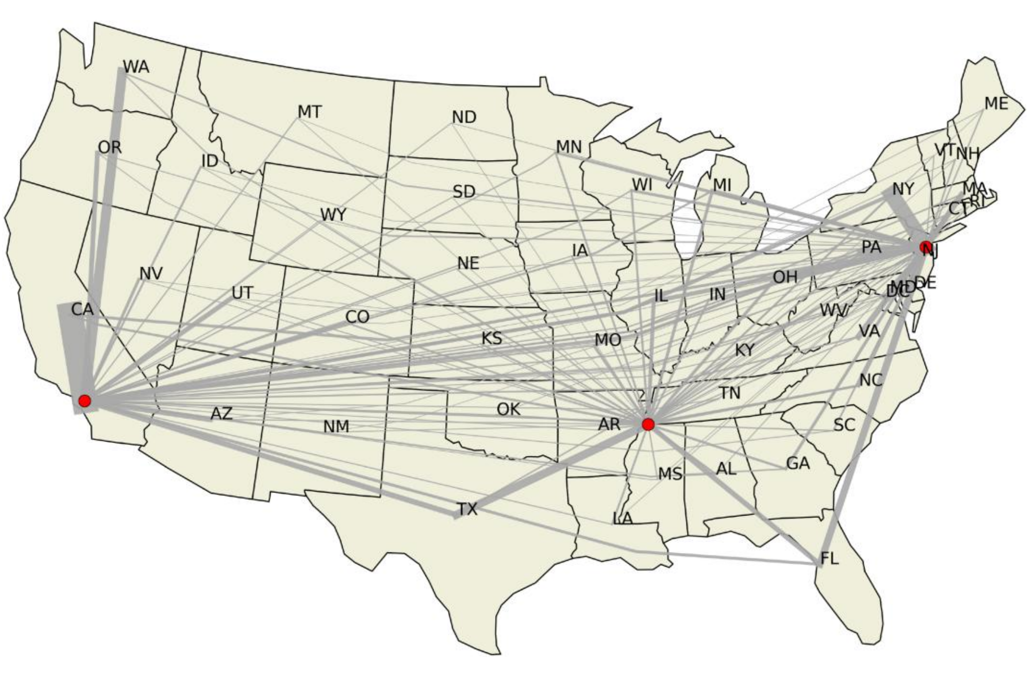
In recent years, there has been a notable shift in discussions surrounding the optimization of e-commerce fulfillment (Fuller 2021). Initially, much of the focus was on the last mile, the most expensive part of the fulfillment process. However, there is now a growing trend among e-commerce companies to pay more attention to the middle mile, which falls between the first and last legs of the typical fulfillment chain (Newton 2023). The first mile primarily involves moving products from suppliers to a fulfillment center, whereas the middle mile involves transporting goods from the fulfillment center to local distribution hubs. This middle-mile segment presents significant opportunities for enhancing the overall order fulfillment process.

There are several key differences between the first mile and the middle mile. To begin with, the first mile often employs standard-sized containers and typically relies on full-truckload (FTL) shipments. In contrast, the middle mile utilizes less-than-truckload (LTL) shipments as they involve smaller quantities and more frequent deliveries. Consequently, although FTL employment has remained relatively stable, there has been a substantial increase in LTL employment (BLS 2022). However, the logistics industry is grappling with challenges in meeting the growing demand, with a shortfall of 80,000 truck drivers in 2021, projected to rise to 160,000 by 2030 (Kriewaldt 2022). Moreover, the high shipping demand surpasses the limited capacity in the freight sector, resulting in carriers having greater pricing power. As a result, LTL prices have significantly outpaced FTL prices (JMR 2022).

The primary objective of this paper is to address the aforementioned challenges by optimizing the middle-mile fulfillment process within an e-commerce context. To achieve this goal, we collaborate with a prominent e-commerce retailer specializing in electronics and computer products. We develop a stochastic optimization model demonstrating how a more efficient middle mile can alleviate the strain on the critical last mile for e-commerce businesses. Our e-commerce partner operates multiple fulfillment centers and local distribution hubs to ensure timely delivery of orders throughout North America. Figure 1 illustrates the geographical distribution of these facilities, with two fulfillment centers on the East and West Coasts, respectively, and a third center located centrally.

Thanks to its network structure, our e-commerce partner efficiently brings merchandise from overseas suppliers to fulfillment centers using standardized containers, which helps keep costs low. However, they encounter various challenges in their middle-mile operations, specifically between fulfillment centers and local distribution hubs. First, like many e-commerce companies, they lack the capacity to maintain their in-house truck fleet. Consequently, they heavily rely on LTL carriers, such as FedEx, UPS, DHL, and other third-party trucking companies, for their middle-mile transportation. This reliance means that they have limited control over their expenses at this stage. Second, they negotiate different rates and LTL capacities (e.g., weights, volumes, or units shipped daily) with various carriers. If the total LTL capacity falls short during certain periods, they turn to the spot market, where costs are typically higher than the negotiated rates. This leads to an increasingly expensive cost structure as shipment quantities accumulate. Third, they offer options like next-day, two-day, or one-week delivery to meet diverse customer expectations. Although this provides flexibility in aligning various delivery options with varying shipping rates, poor management of these options can result in higher fulfillment costs. For instance, let us consider the two-day delivery option for simplicity. If the total demand surpasses the contracted LTL capacity, the e-commerce company can delay

Figure 1. (Color online) The Fulfillment Network of Our Partner E-Commerce Company Is Illustrated on This Map



Notes. Red dots represent distribution centers, whereas lines depict the flow of fulfillment operations connecting these distribution centers with local hubs within each state. The thickness of these lines corresponds to the average daily fulfillment volume originating from each distribution center.

transporting a portion of the demand with the two-day delivery option. This delayed portion can later be fulfilled at a lower rate using future LTL capacity. However, excessive delays risk fulfilling the delayed orders at even higher rates.

This example highlights the significance of adopting a dynamic fulfillment strategy when logistical resources are increasingly expensive. It showcases how an e-tailer can effectively reduce costs by leveraging its fulfillment network's flexibility or strategically delaying fulfillment. Determining the optimal decision, however, is complicated by factors such as the distribution of demand, the structure of the fulfillment network, and heterogeneous delivery options. In what follows, we summarize our main contribution and then, review the relevant literature.

1.1. Main Contributions

In our study, we first demonstrate that an optimal policy prioritizes order fulfillment and takes the form of a state-dependent threshold policy tailored to each warehouse. This finding implies that for an e-commerce company, it is optimal to satisfy demand up to a certain threshold whereas postponing the remaining orders when the cost of exceeding capacity becomes prohibitively high.

Second, despite demonstrating the optimality of a threshold-type policy, computing the state-dependent thresholds poses significant challenges because of the curse of dimensionality. As a result, we introduce a simple heuristic approach for fulfillment decisions based on the concept of Lagrangian relaxation. Our method involves decomposing the original problem both across warehouses and over time. To account for state dependency, we scale the Lagrangian optimal fulfillment decision based on the observed state, prioritize the orders based on their discounted penalty costs, and limit the number of orders fulfilled from each warehouse. Furthermore, we provide a performance guarantee for the Lagrangian relaxation-based policy (threshold Lagrangian relaxation (tLR))). Our analysis reveals that the difference in optimality between the tLR policy and an optimal policy grows at a rate no worse than $\mathcal{O}(TMK\sqrt{N})$, where T is the length of the horizon, M is the number of distribution centers (warehouses), N is the number of local fulfillment hubs (locations), and K is the largest length of a fulfillment window. Indeed, the sublinear growth with respect to N is particularly crucial for many real-world online retail settings, such as our e-commerce partner, which operates only a few large distribution centers catering to a vast number of local fulfillment hubs (i.e., small M and large N). Additionally, we extended the tLR policy to an infinite horizon and demonstrated that a similar performance guarantee holds for the infinite horizon version.

Finally, we use both synthetic and real data to investigate the performance of the proposed policy. Our results indicate that incorporating multiperiod fulfillment windows can effectively alleviate logistical capacity constraints. Moreover, we discover that integrating the information about remaining fulfillment windows into the decision-making process can yield additional benefits. Our findings suggest that online retailers can gain a competitive advantage by offering customers options such as a "two-day fulfillment" service instead of a "same-day fulfillment" service, particularly when faced with limited logistical capacities.

1.2. Related Literature

1.2.1. Order Fulfillment. Our paper is closely related to research on online order fulfillment problems. Xu et al. (2009) introduce one of the first models in this area, focusing on order consolidation. Other studies concentrate on minimizing operational costs under different cost structures and inventory constraints. For instance, Acimovic and Graves (2015) and Andrews et al. (2019) both adopt linear costs and examine single-product fulfillment problems, with stochastic and adversarial demand arrivals, respectively. In contrast, Jasin and Sinha (2015) address a multi-item fulfillment problem, where shipping costs consist of both variable and fixed components. However, their variable shipping cost is still product specific. More recently, Balseiro et al. (2024) introduce a unified model and apply it to the online order fulfillment problem with inventory constraints. They demonstrate that the average optimality gap of their heuristic is independent of the remaining horizon.

Although previous research focuses on fulfillment policies with limited inventory capacity, our paper addresses optimal strategies when shipping capacity is constrained. We model the challenges faced by online retailers using LTL services for middle-mile transportation, where total fulfillment costs depend on the weight (or volume) of shipped products rather than the number of units. We assume nonlinear fulfillment cost functions, where marginal costs increase with the shipped weight. Instead of making fulfillment decisions as orders arrive, we require that all orders be fulfilled within specified time windows. Our contribution lies in demonstrating the benefits of multiperiod fulfillment windows under limited logistics resources and proposing a practical heuristic policy that integrates remaining fulfillment window information into the decision-making process.

Three other papers that explore the benefit of delaying fulfillment decisions are Mahar and Wright (2009), Wei et al. (2021), and Xie et al. (2022). In Mahar and Wright (2009), researchers investigate a quasidynamic assignment policy and demonstrate it reduces inventory costs and backorder costs. Wei et al. (2021) study optimal policies for consolidating orders with varying fulfillment windows, assuming fixed and variable costs that increase as the fulfillment window shortens. Although we also assume that each order has a limited fulfillment window, our focus is on delayed fulfillment to manage logistics capacity constraints. We employ convex cost functions that depend on the number of orders in progress and consider a general fulfillment network with multiple warehouses, unlike Wei et al. (2021), who consider at most two warehouses. Xie et al. (2022) examine an online resource allocation problem where decisions are delayed by K periods to leverage actual demand data. They show that the gap between their delayed online algorithm and the offline optimal policy diminishes exponentially in K. In contrast, our study requires fulfillment decisions at the end of each period, without the benefit of future demand knowledge. Instead, we assume that orders have varying fulfillment windows, with unfulfilled orders carried over. Our numerical analysis indicates that the total fulfillment costs decrease as the fulfillment window lengthens, suggesting that multiperiod fulfillment windows help e-commerce companies mitigate logistics capacity constraints.

Previous researchers have also extended this stream of research to joint optimizations. For instance, both Lei et al. (2018) and Harsha et al. (2019) examine the joint optimization problem of pricing and order fulfillment for an omnichannel retailer, whereas Lei et al. (2022) consider joint pricing, display, and fulfillment problem. A detailed literature review can be found in Acimovic and Farias (2019) and Qi et al. (2020).

1.2.2. Multistage Stochastic Programming. Our study intersects with a comprehensive body of literature on multistage stochastic programming, where the focus has predominantly been on linear objective functions as evidenced in domains such as inventory management (Jasin and Kumar 2012) and Markovian bandits (Brown and Smith 2020). We direct readers to the seminal work by Shapiro et al. (2021) for a thorough overview. Unlike most existing studies, our formulation introduces a convex cost function, which emerges organically within our specific study context. The work most closely aligned with ours is that of Yan and Reiffers-Masson (2023), which centered around certainty equivalence control, develops two heuristics (an update policy and a projection policy) for tackling general multistage stochastic convex optimization challenges along with providing a theoretical foundation for these strategies.

Our research diverges from Yan and Reiffers-Masson (2023) in several significant ways. First, the convex cost structure in our model is a natural consequence of employing LTL contracts within middle-mile logistics operations. Our model does not impose strict conditions on the convexity of the cost function; it is flexible enough to accommodate piecewise linear and convex functions and linear functions within a bounded support. Second, our approach does not require independent and identically distributed demand so that it effectively addresses demand with temporal correlations, a critical factor given the prevalent strong time-based correlations observed in real sales data. Third, although our methodology shares similarities with the projection heuristic policy described by Yan and Reiffers-Masson (2023), we employ a dual formulation via Lagrangian relaxation. Using a scaling technique and adjustments tailored to optimize policy characteristics, our method offers computational efficiencies and good performance.

1.2.3. Weakly Coupled Stochastic Dynamic Program. From the methodology perspective, our paper is related to the literature on weakly coupled stochastic dynamic programs (DP), in which the problem consists of a finite number of subproblems and their actions are linked with a set of constraints. Hawkins (2003) describes many practical problems that fit this structure and provided Lagrangian decomposition to both finite and infinite horizon problems. Adelman and Mersereau (2008) study two commonly used techniques (linear program-based approximate DP and Lagrangian relaxation) for a discounted infinite horizon stochastic DP with finite state and action spaces. Brown and Zhang (2023) further study the relation between the two methods and provided the conditions under which both methods have identical performance. Brown and Smith (2020) adopt the Lagrangian relaxation to a finite horizon dynamic assortment problem and showed that the heuristic policy admits an optimality gap of $\mathcal{O}(2^T\sqrt{N})$. Brown and Zhang (2022) propose a novel dynamic fluid relaxation for the weakly coupled stochastic DP problem, which obtains a performance loss of $\mathcal{O}(2^T\sqrt{N})$. This performance is further improved to $\mathcal{O}(T^3\sqrt{N})$ by reoptimization (Brown and Zhang 2025).

Our problem is also a finite horizon weakly coupled stochastic DP, but the state space is not the Cartesian product of the state spaces of subproblems. Despite that, we show in the later section that the problem can be decomposed into optimization of subproblems after relaxing the linking constraints. We propose two heuristic policies based on *restricted Lagrangian relaxations*, where the value of Lagrangian multipliers is only period dependent. We then study the performance of the Lagrangian relaxation policies theoretically and numerically, and we show that it admits an optimality gap of $\mathcal{O}(TMK\sqrt{N})$. This result implies the asymptotic optimality of the policies in terms of average fulfillment cost per location when the firm serves a large number of locations.

1.2.4. Inventory Control. Our paper is also related to the literature on multiwarehouse, multistore (MWMS) problems, where products are periodically shipped from warehouses to stores and where local demands are fulfilled using store inventories. Marklund and Rosling (2012) study the inventory control problem with one-warehouse multistores. Nambiar et al. (2021) extend to correlated demands and lost sales. A more general class of MWMS problems is handled by Miao et al. (2022), whose heuristic is also based on Lagrangian relaxation but yields a better performance bound. Their heuristic is asymptotically optimal for both the horizon and the number of retailers. In contrast, we consider a middle-mile logistic problem where products are transported in batches from distribution centers (warehouses) to local fulfillment hubs (locations) and then, delivered to customers. The key differences include that inventory cannot be held in the local fulfillment hubs and that demand realizations are observed before orders are shipped from the distribution centers. The trade-off in our study is to utilize future fulfillment capacity whereas avoiding the risk of paying an even higher cost.

1.3. Organization

The remainder of this paper is organized as follows. We formulate the problem as a finite horizon stochastic DP in Section 2 and present the structure of an optimal policy in Section 3. Based on Lagrangian relaxation of the linking constraints, our heuristic is developed in Section 4 together with its theoretical performance bound. We then extend the formulation and results to an infinite horizon in Section 5. In Section 6, we conduct extensive numerical analysis of our heuristic policy with synthetic and real data. Finally, we conclude the paper in Section 7 with discussions for future research.

2. Problem Formulation

Motivated by our industry partner, we model the order fulfillment problem as follows. An online retailing company (DM) receives orders from the customers in each period over a T-period horizon. The customers place these orders from N distinct locations, and the DM has M warehouses to fulfill these orders. There is no inventory constraint for the warehouses (which is relaxed in Online Appendix EC.2), and we cannot store inventory in each location. We assume that the fulfillment network from the warehouses to locations is fully connected. However, our model and results can be easily extended to a partially connected fulfillment network. We discuss partially connected cases after Proposition 3 in Section 4. Each order has a limited time frame known as the fulfillment window, which consists of k periods ($k = 1, \dots, K$), within which the order should be shipped from one of the M warehouses. If the DM cannot fulfill an order within its fulfillment window, then she incurs a penalty cost, which may represent either lost sales cost or the expedited delivery cost for the online retailer so that the order delivery promise is fulfilled. We use penalty cost and lost sales cost interchangeably throughout the paper. To minimize the expected fulfillment cost, the online retailer decides when and from which warehouse to fulfill each order within the fulfillment window.

Before presenting the stochastic DP formulation of this problem, we first define the notation. Let $\mathbb{R}_+$ denote $[0, +\infty)$. Let $[n]$ denote the set $\{1, 2, \dots, n\}$. Let $i \in [M]$ denote the ith warehouse, $j \in [N]$ denote the jth location, $t \in [T]$ denote the tth period, and $k \in [K]$ denote the remaining number of periods left within the fulfillment window.

We assume that a customer may purchase more than one item per order, and our focus is on the fulfillment-related attributes, such as the total weight or volume, of the order. Thus, we define a demand vector for period t as $D_t = (D_{k,j,t} \in \mathbb{R}_+ : k \in [K], j \in [N])$, where each element $D_{k,j,t}$ is a random variable and represents the *weights or volume* of goods from location j scheduled for fulfillment within k periods. This formulation effectively captures customer preferences regarding shipping speed. For instance, $D_{1,j,t} > 0$ corresponds to orders requiring immediate fulfillment services (e.g., "same-day" delivery), whereas $D_{K,j,t} > 0$ refers to orders that can be processed within longer fulfillment windows, such as those fulfilled via ground transportation. We assume to know the demand expectation $\mathbb{E}\{D_{k,j,t}\}$ for all k, j, and t. Nonetheless, these demands may adhere to distinct and unknown distributions.

In practice, customers place orders for various products through the online store, and the retailer schedules fulfillment periodically. To model this, we assume that demand for the current period D_t becomes known before the fulfillment assignment, whereas future demand remains uncertain to the DM. Additionally, we assume that demands across different locations and demand from the same location across different periods are independent and bounded. We discuss how our analyses can be extended to the scenarios where demands at the same location exhibit temporal correlation (see Section 4.2) and unbounded (see Online Appendix EC.5). With this context, we introduce the following assumption regarding the demand.

Assumption 1. *The support for demand distribution is bounded such that $0 \le D_{k,j,t} \le \overline{D}$ almost surely for some constant $\overline{D}$ for all k, j, and t.*

2.1. State and Action

With this, we can now define the state $x_{k,j,t}$ as the weight of cumulative unfulfilled orders (i.e., outstanding orders) received from customer location j that have k remaining periods left in the fulfillment window at period t. Denote by $x_t = (x_{k,j,t} \in \mathbb{R}_+ : k \in [K], j \in [N])$ the vector of the weights of unfulfilled orders at period t. After observing the state vector x_t, the online retailer makes a fulfillment decision $u_t = (u_{k,j,i,t} \in \mathbb{R}_+ : k \in [K], j \in [N], i \in [M])$, where each element $u_{k,j,i,t}$ corresponds to the weights of the orders shipped in period t from warehouse i to demand location j to satisfy the unfulfilled orders with k periods left in fulfillment window.

The DM cannot ship products in advance to locations for potential future demand realizations. That is, the total weights of products shipped from M warehouses to location j cannot exceed that of unfulfilled orders placed from location j. Consequently, we impose the following feasibility constraints $\sum_{i=1}^{M} u_{k,j,i,t} \leq x_{k,j,t}$ for all $j \in [N]$, $k \in [K]$, and $t \in [T]$. For notational brevity, we can define an M-block matrix $G = [I, \ldots, I]$, where $I \in \mathbb{R}^{KN \times KN}$ is an identity matrix. Hence, the feasibility constraint can be written as

$$Gu_t \leq x_t.$$

Note that this constraint connects the fulfillment decisions made from different warehouses together and is also referred to as the *linking constraint* in the weakly coupled stochastic DP literature (e.g., Adelman and Mersereau 2008). We now define feasible action vector $u_t \in \mathcal{U}_t(x_t)$, where

$$\mathcal{U}(x_t) = \{u_t \in \mathbb{R}_+^{KNM} : Gu_t \leq x_t\}.$$

In addition to linking constraints, our methodology can also incorporate inventory and warehouse capacity constraints. We refer to Online Appendices EC.1 and EC.2 for these extensions.

A *policy* $\pi = \{\pi_1, \pi_2, \ldots, \pi_T\}$ is a sequence of functions that map from state space to action space. A policy is *feasible* if $\pi_t(x_t) \in \mathcal{U}(x_t)$ for all periods and states. We denote by $\mathbf{\Pi}$ the set of all feasible policies.

2.2. System Dynamics

Next, we define the state update equations. Once the orders are shipped, unfulfilled orders will stay in the system, with the number of remaining periods left in the fulfillment window reduced by one period. New demands are then realized for the next period $t+1$. The state vector of the system is updated according to the following system dynamics:

$$x_{k,j,t+1} = \begin{cases} x_{k+1,j,t} - \sum_i u_{k+1,j,i,t} + D_{k,j,t+1} & k < K \\ D_{k,j,t+1} & k = K \end{cases}.$$

We can consolidate the above equation in a matrix form and define the one-step state transition function as

$$x_{t+1} = A(x_t - Gu_t) + D_{t+1}, \tag{1}$$

where A represents an N-block diagonal matrix and each block is a $K \times K$ one-off upper diagonal matrix with ones: that is,

$$A = \mathrm{diag}([A_K, \ldots, A_K]),$$
$$\text{where } A_K = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} \in \mathbb{R}^{K \times K}.$$

2.3. Cost Structure

Our partner online retailer depends on transportation contracts for outsourcing its logistics operations, structured to have escalating costs with increased fulfillment capacity. To accurately represent this contractual framework, we assert that the fulfillment cost function demonstrates convex behavior, indicating that costs increase at an accelerating rate with increased fulfillment quantity. Specifically, we denote the warehouse-specific fulfillment cost function as $\mathcal{C}_i : \mathbb{R}_+ \to \mathbb{R}_+$ for warehouse $i \in [M]$ under the following assumption.

Assumption 2. *$\mathcal{C}_i(y)$ is convex increasing in y, and its derivative $\mathcal{C}_i'(y) < \overline{C}$ for all $y \in \mathbb{R}_+$, where $0 < \overline{C} < \infty$.*

The fulfillment cost for warehouse i, given its fulfillment vector $u_{i,t}$, can be obtained as

$$\mathcal{C}_i\left(\sum_{k\in[K]} \sum_{j\in[N]} u_{k,j,i,t}\right) = \mathcal{C}_i(e^T u_{i,t}),$$

where e is a vector of ones whose dimension can be easily determined in the context.

The foundation of Assumption 2 is built upon three key observations drawn from industry practices. (1) Shippers (i.e., online retailers) strategically engage with multiple LTL carriers; (2) LTL contracts are generally negotiated based on clusters of demand locations, whereby shipments originating from the same warehouse to any destination within the same cluster share identical pricing and collectively utilize a common logistics capacity; and (3) each carrier offers finite logistics capacity to shippers. Specifically, engaging multiple carriers allows shippers to access diverse competitive pricing structures and logistical capacities, providing critical flexibility in responding to varying demand levels and mitigating supply chain risks. Cluster-based negotiation simplifies the process and offers significant advantages to both shippers and logistics carriers. Shippers benefit from more predictable and stable regional shipping costs, whereas carriers achieve simplified operational management and improved efficiency, significantly reducing the complexity associated with managing capacity and pricing at an individual lane level. Finally, capacity constraints defined by carriers prevent overcommitment, promoting reliable and sustainable operations.

Whenever the DM cannot fulfill the orders from location j within its fulfillment period k because of the insufficient contracted capacity, each order leaves the system with a unit penalty cost b_j, which can be considered as the spot market rates. The penalty does not depend on the warehouse because the DM can always choose to fulfill from the warehouse closest to location j. We define the highest penalty cost as $\overline{b} := \max_j b_j$ and denote $b = (b_j : j \in [N])$ the vector of penalty costs.

We now make remarks regarding the cost function. First, for the sake of technical simplicity, we assume that the fulfillment cost function is continuously differentiable. This assumption can be relaxed without loss of generality. In practice, online retailers often allocate shipments across multiple carriers. This leads to a piecewise increasing convex cost function that is nondifferentiable. In such cases, our analysis remains valid by replacing gradients with subgradients.

Second, we assume that the marginal cost is bounded. This assumption can be relaxed. For instance, if the marginal cost increases unboundedly, there exists a threshold above which the marginal cost exceeds the cost of fulfillment via the spot market defined by the shipping rate of an expedited service. Thus, without loss of generality, we can always define a bound on the marginal cost. Indeed, defining a very high marginal cost allows us to model scenarios with strict fulfillment capacity constraints. Let $U_i > 0$ represent the fulfillment capacity for warehouse i. In those instances, we can define $\mathcal{C}_i$ such that $\mathcal{C}_i(e^\top u_{i,t}) > \overline{b}$ whenever $e^\top u_{i,t} \geq U_i$, ensuring that in the optimal solution, the fulfillment from warehouse i does not exceed U_i. A detailed discussion on this is included in Online Appendix EC.1.

Third, our formulation of the fulfillment cost function, $\mathcal{C}_i(e^\top u_{i,t})$, assumes a uniform unit shipment cost across all demand locations. In practice, however, carriers, such as UPS, group demand locations into clusters—commonly referred to as shipping zones—and apply different unit costs based on the destination zone (ShipBob 2022). This clustering approach is also used by our industry partner. For instance, shipments from the West Coast warehouse are divided into five distinct zones. To accommodate such settings, we extend our model by defining the shipping cost function with two indices, $\mathcal{C}_{i,z}(\cdot)$, where i denotes the warehouse and z denotes the destination zone. A detailed discussion of this extension is provided in Online Appendix EC.3. Finally, we assume that marginal costs are time invariant, an assumption that we relax and examine further in Online Appendix EC.4.

Using the state and action defined above, we can calculate the total number of outstanding orders from location j in period t that are not fulfilled within its fulfillment window by $x_{1,j,t} - \sum_{i\in[M]} u_{1,j,i,t}$. Multiplying this by location-specific unit penalty cost b_j and summing over all locations $j \in [N]$, we obtain the total penalty cost as follows: $\sum_{j\in[N]} b_j(x_{1,j,t} - \sum_{i\in[M]} u_{1,j,i,t})$. Adding the total fulfillment and penalty cost, we define the *immediate cost function* at period t as follows:

$$g(x_t, u_t) = \sum_{i\in[M]} \mathcal{C}_i(e^\top u_{i,t}) + \sum_{j\in[N]} b_j \left(x_{1,j,t} - \sum_{i\in[M]} u_{1,j,i,t} \right). \tag{2}$$

To ensure that our setup is nontrivial, we make the following assumption for all warehouses $i \in [M]$.

Assumption 3. $\mathcal{C}_i'(0) < \overline{b}$.

Note that if Assumption 3 does not hold for some warehouse i, we can remove that warehouse without loss of generality. This is because no order will be fulfilled from that warehouse i within its fulfillment window because incurring penalty costs would be less costly than fulfilling orders within their fulfillment windows.

2.4. Putting Everything Together

Let $V_t^\pi(x_t)$, $t = 1, \ldots, T$, be the expected discounted cost-to-go function under a given feasible policy π starting from any admissible state x_t. The total expected cost can be expressed as follows:

$$V_1^\pi(x) = \mathbb{E}\left\{ \sum_{t\in[T]} \gamma^{t-1} g(x_t, \pi_t(x_t)) + \gamma^T V_{T+1}(x_{T+1}) \,|\, x_1 = x + D_1 \right\},$$

where γ, $0 < \gamma \leq 1$, is the discount factor; x represents the vector of unfulfilled orders before the process and is known to the DM; and $\mathbb{E}\{\cdot\}$ denotes expectation with respect to demand distribution. We set the terminal value function $V_{T+1}(x) = \sum_{i,j,k} b_j x_{k,j,T+1}$, meaning that all unfulfilled orders at the end of the horizon are handled with expedited shipping services. This is standard assumption in many dynamic inventory models (Porteus 2002) and will facilitate our assumption. Therefore, the stochastic optimization problem can be expressed as

$$V_1^*(x) = \min_{\pi\in\mathbf{\Pi}} V_1^\pi(x). \tag{3}$$

Following a standard approach in the stochastic optimization literature (e.g., Puterman 2014), the online retailer's problem can be expressed as a stochastic DP satisfying the following recursive relation:

$$V_t^*(x_t) = \min_{u_t\in\mathcal{U}(x_t)} g(x_t, u_t) + \gamma \mathbb{E}\{V_{t+1}^*(x_{t+1}) \,|\, x_t, u_t\}.$$

We characterize the order fulfillment problem by the number of warehouses, the number of locations, and the length of the fulfillment window, and we use the notation "MW-NL-KP" for the problem with M warehouses, N locations, and K periods of fulfillment window.

3. Structure of an Optimal Fulfillment Policy

To determine the structure of optimal policy in a stochastic DP model, the typical strategy involves proving the value function's convexity (or concavity under maximization) in the state variable, preserving this property under the DP operator, and using induction to demonstrate the desired threshold-based structure of the optimal policy. Although this optimal structure is quite intuitive, unfortunately, because of the curse of dimensionality, it is not very helpful to solve the fulfillment problem exactly. Therefore, in the next section, we will use the structural policy to derive practical heuristics with theoretical guarantees. To establish the desired threshold-based structure of the optimal policy, we need to first characterize the convexity of the optimal value function $V_t^*(x)$ with respect to x.

Proposition 1. *$V_t^*(x)$ is convex, and $\frac{\partial V_t^*(x)}{\partial x_{k,j,t}} \le \gamma^{\min\{k-1,T+1-t\}} b_j$ for all k, j, and t.*

The proof can be seen in Online Appendix EC.9. At period t, any unfilled order has three options. It can be fulfilled immediately in period t; if there are remaining periods in its fulfillment window (i.e., $k > 1$), it can be deferred to the next period; or it leaves the system with a penalty if it is in the last period of its fulfillment window (i.e., $k = 1$). Which of these options is optimal in period t is determined by comparing the marginal cost of fulfilling now, the marginal cost of deferring to the next period, and the penalty cost. Note that the marginal cost of deferring an order from location j to the next period can be expressed as follows:

$$-\gamma \frac{\partial \mathbb{E}\{V_{t+1}^*(x_{t+1})\}}{\partial u_{k,j,i,t}} = -\gamma \frac{\partial \mathbb{E}\{V_{t+1}^*(x_{t+1})\}}{\partial x_{k-1,j,t+1}} \frac{\partial x_{k-1,j,t+1}}{\partial u_{k,j,i,t}} = \gamma \frac{\partial \mathbb{E}\{V_{t+1}^*(x_{t+1})\}}{\partial x_{k-1,j,t+1}}$$

for $k > 1$, where the last equality holds by the system dynamics $x_{k-1,j,t+1} = x_{k,j,t} - \sum_{i\in[M]} u_{k,j,i,t} + D_{k-1,j,t+1}$; the negative sign in the leftmost term indicates that a postponement is achieved by reducing one fulfillment from some warehouse i. We can deduce that an optimal policy would allocate the order from location j to warehouse i as long as the marginal fulfillment cost $C_i'(e^\top u_{i,t})$ at warehouse i is smaller than $\gamma \frac{\partial \mathbb{E}\{V_{t+1}^*(x_{t+1})\}}{\partial x_{k-1,j,t+1}}$ for $k > 1$ and b_j for $k = 1$. To unify both, we denote $\delta_{1,j,t} = b_j$ and $\delta_{k,j,t} = \gamma \frac{\partial \mathbb{E}\{V_{t+1}^*(x_{t+1})\}}{\partial x_{k-1,j,t+1}}$ for $k > 1$ for all j and t. We have $\delta_{k,j,t} \le \gamma^{\min\{k-1,T+1-t\}} b_j$ by the definition and Proposition 1. This implies that the total number of orders fulfilled by warehouse i should also be bounded. Given that the total number of orders fulfilled from each warehouse i does not exceed a threshold, orders must be prioritized based on their locations j and how many periods are left in their fulfillment windows k. The following proposition formalizes this intuition.

Proposition 2. *In an "MW-NL-KP" problem,*

1. *each warehouse has a state-dependent threshold optimal fulfillment policy, ensuring that the total number of orders fulfilled from any warehouse does not exceed its specific threshold. In particular, if the warehouse i is ever used to fulfill any orders from location j with remaining fulfillment window k, then the total fulfillment from that warehouse must satisfy $C_i'(e^\top u_{i,t}^*) \le \gamma^{\min\{k-1,T+1-t\}} b_j$.*
2. *The orders to be fulfilled are prioritized according to values $\delta_{k,j,t}$ assigned to each demand location j and fulfillment window k in the sense that orders with higher $\delta_{k,j,t}$ receive priority.*

The proof is also available in Online Appendix EC.9. We have several takeaways from Proposition 2. First, for those orders whose marginal cost of postponing, $\delta_{k,j,t}$, is greater than $\overline{C}$, we fulfill them immediately now and allocate these orders fully to the warehouses. This is intuitive because the marginal fulfillment cost at each warehouse $i \in [M]$ is bounded by $\overline{C}$ according to Assumption 2. Therefore, if an order's marginal cost of postponing is more than $\overline{C}$, it is cheaper for the online retailer to fulfill this order now rather than postpone its fulfillment to later periods. Second, if there is an order (k, j) that satisfies $\delta_{k,j,t} < \overline{C}$, then whether such an order would be fulfilled now or later depends on the sequence in which the other orders are fulfilled. Because the fulfillment sequence determines the marginal cost at each warehouse i, determining the optimal fulfillment policy requires us to solve the optimal DP in Equation (3).

As discussed above, although we have demonstrated the existence of an optimal threshold-type fulfillment policy, determining which and how many orders (indexed by k and j) to fulfill at the current period and how to allocate them to warehouses suffers from the curse of dimensionality. In addition, the characterization of an optimal policy does not mention how to allocate the orders among the warehouses. Therefore, we are motivated to consider efficient heuristics with provable performance bounds. Along these lines, Propositions 1 and 2 provide us with two valuable insights. First, as suggested by the existence of thresholds characterized in Proposition 2, there is a cap on the number of orders fulfilled in each period. That means that we have to prioritize which order to fulfill immediately and postpone to the next period. Second, in the scheme of prioritization among the unfulfilled orders, Proposition 2 suggests that we take into account the discounted penalty cost associated with each order. Building on these insights, we propose efficient heuristics in the subsequent section. This approach allows us to make tractable decisions on whether to fulfill an order immediately or postpone it for later as well as how to allocate orders from each location j to each warehouse i, enabling us to manage the fulfillment process effectively.

4. Heuristic Policies and Their Performances

As discussed in the previous section, the DP problem in Equation (3) is difficult to solve in part because of the linking constraints. To overcome this challenge, we relaxed the problem with Lagrangian relaxation. In the literature on weakly coupled stochastic DP, it is well known that this technique can decompose the original problem into subproblems in both finite and infinite horizons (e.g., Hawkins 2003, Adelman and Mersereau 2008, Brown and Zhang 2022). Heuristics can be developed based on this decomposition if the subproblems are more manageable. The fact that our state is location specific suggests the possibility of decomposing the problem into single-location problems. However, the fulfillment cost is incurred at each warehouse level, and the total fulfillment cost for each location depends not only on the amount shipped to that location but also, on the amounts shipped to other locations. Therefore, location-specific decomposition of the problem does not yield a viable option. Instead, we opt to decompose the problem along the warehouse dimension. In what follows, we provide the details.

4.1. Properties of the Lagrangian Relaxation

We develop Lagrangian relaxation for a T-period "*MW-NL-KP*" problem starting from state x_t. Let $\lambda_\tau : \mathcal{S} \to \mathbb{R}^{KN}$, for $\tau \in [t, T]$, be an arbitrary function that generates Lagrangian multipliers associated with the linking constraint $\sum_i u_{i,\tau} \le x_\tau$. Hence, the corresponding Lagrangian DP value function L_t can be defined recursively as follows:

$$\begin{aligned} &L_t(x_t; \lambda_t(x_t), \lambda_{t+1}(x_{t+1}), \ldots, \lambda_T(x_T)) \\ &= \min_{u_t \ge 0} g(x_t, u_t) - (x_t - Gu_t)^\top \lambda_t(x_t) \\ &\quad + \gamma \mathbb{E}\{L_{t+1}(x_{t+1}; \lambda_{t+1}(x_{t+1}), \ldots, \lambda_T(x_T)) | x_t, u_t\}, \end{aligned} \tag{4}$$

where $L_{T+1}(x) = V_{T+1}(x)$. The following results demonstrate that even though the state cannot be decomposed into subproblems, the Lagrangian relaxation exhibits well-established properties characterized in the literature on weakly coupled stochastic DP (see, e.g., Hawkins 2003, section 2.3).

Proposition 3 (Finite Horizon Lagrangian Relaxation). *For any $t \in [1, T]$, $x_t \in \mathcal{S}$, and $\lambda_\tau : \mathcal{S} \to \mathbb{R}_+^{KN}$ for $\tau \in [t, T]$, L_t satisfies the following.*

1. (*Weak duality*) $V_t^*(x_t) \ge L_t(x_t; \lambda_t(x_t), \lambda_{t+1}(x_{t+1}), \ldots, \lambda_T(x_T))$.

2. (*Decomposition*) *If $\lambda_\tau(x) = \lambda_\tau$ for all $x \in \mathcal{S}$ for all $\tau \in [t, T]$, then*

$$L_t(x_t; \boldsymbol{\lambda}_{t,T}) = B_t^\top \bar{\boldsymbol{x}}_t - \boldsymbol{\lambda}_{t,T}^\top A_t \bar{\boldsymbol{x}}_t + \sum_{i=1}^{M} \sum_{\tau=t}^{T} \gamma^{\tau-t} L_{i,\tau}(\boldsymbol{\lambda}_{\tau,T}) \tag{5}$$

and

$$L_{i,\tau}(\boldsymbol{\lambda}_{\tau,T}) = \min_{u_i \ge 0} \mathcal{C}_i(e^\top u_i) - b^\top Q_{T+1-\tau} u_i + \sum_{\tau'=\tau}^{T} \lambda_{\tau'}^\top (\gamma A)^{\tau'-\tau} u_i, \tag{6}$$

where

$$A_t = \begin{bmatrix} \gamma^0 A^0 & 0 & \cdots & 0 \\ \gamma^1 A^1 & \gamma^1 A^0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \gamma^{T-t} A^{T-t} & \gamma^{T-t} A^{T-t-1} & \cdots & \gamma^{T-t} A^0 \end{bmatrix} \in \mathbb{R}^{(T-t+1)NK \times (T-t+1)NK}, \tag{7}$$

$$B_t^\top = [\gamma^0 b^\top Q_{T+1-t},\ \gamma^1 b^\top Q_{T-t}, \ldots,\ \gamma^{T-t} b^\top Q_1] \in \mathbb{R}^{(T-t+1)NK}, \tag{8}$$

$$\bar{\boldsymbol{x}}_t^\top = [x_t^\top,\ \mathbb{E}\{D_{t+1}^\top\}, \ldots, \mathbb{E}\{D_T^\top\}] \in \mathbb{R}^{(T-t+1)NK}, \tag{9}$$

$$\boldsymbol{\lambda}_{t,T}^\top = [\lambda_t^\top,\ \lambda_{t+1}^\top, \ldots, \lambda_T^\top] \in \mathbb{R}^{(T-t+1)NK}, \tag{10}$$

A is the transition matrix in (1), $Q_t = diag([\boldsymbol{\gamma}_t, \ldots, \boldsymbol{\gamma}_t])$ is an N-block diagonal matrix, and $\boldsymbol{\gamma}_t$ is a row vector with length K whose ith *element is defined as $\gamma^{\min(t,i)}$, where $i = 0, \ldots, K-1$.*

3. (*Concavity*) *$L_{\tau,i}(\boldsymbol{\lambda}_{\tau,T})$ is piecewise linear and concave in λ_τ for $\tau \in [t, T]$. So is $L_t(x_t; \boldsymbol{\lambda}_{t,T})$.*

We defer the proof to Online Appendix EC.10. One way to interpret $\bar{\boldsymbol{x}}_t$ is the expected extended state, taking into account the expected demand for the remaining periods. The matrices A_t and B_t represent the dynamics of the system and the cost vector associated with the penalty, respectively, for the extended state. The matrix Q_t takes into account the discounting factor.

Note that future states x_τ are random, and our notation $\lambda_\tau(x_\tau)$ implies that the Lagrangian multipliers depend on the state realizations. We need to introduce a Lagrangian variable for each state in order to fully describe the Lagrangian function, which is intractable because of the curse of dimensionality. Therefore, we impose a state-independent constraint ($\lambda_\tau(x) = \lambda_\tau$ for all x) for each Lagrangian multiplier in order to obtain a tractable decomposition.

Remark 1. For a partially connected fulfillment network in which some locations cannot be fulfilled with certain warehouses, we introduce an index set $\mathcal{N}_i$ for each warehouse. Location j is "connected" with warehouse i if $j \in \mathcal{N}_i$ such that i can fulfill orders from j. For nonconnected warehouse-location pairs, we cannot make fulfillment. Therefore, we add a nonconnectivity constraint $u_{k,j,i,t} = 0$ for all k, i, $j \notin \mathcal{N}_i$, and t to the feasible action sets $\mathcal{U}(x_t)$, which can be handled by adding this constraint to Subproblem (6).

Unlike a classic weakly coupled stochastic DP problem (e.g., Hawkins 2003, Adelman and Mersereau 2008), whose decomposition leads to a series of DP subproblems, we decompose the original problem both in warehouses and in periods. Consequently, the decomposed Subproblems (6) are easier to solve. This is because it reduces the dimension of the convex optimization problem. Specifically, for a given set of $\boldsymbol{\lambda}_{t,T}$, it only requires solving a series single-period convex optimization for each warehouse i and each period $\tau \geq t$. We denote a solution to (6) with a specific index i and τ as $\psi_{i,\tau}$ and denote ψ_τ, for $\tau \in [t,T]$, as the set of the fulfillment decision $\psi_{i,\tau}$ for all $i \in [M]$. Consequently, $\psi = \{\psi_t, \psi_{t+1}, \ldots, \psi_T\}$ can be considered as an optimal policy for $L_t(x;\boldsymbol{\lambda}_{t,T})$.

The weak duality and the concavity of L_t suggest that we can obtain the tightest lower bound (LB) by solving the *Lagrangian dual problem*:

$$\mathcal{D}(t,x_t) := \max_{\boldsymbol{\lambda}_{t,T}} L_t(x_t;\boldsymbol{\lambda}_{t,T}).$$

In particular, we only solve $\mathcal{D}(1,x_1)$ at the first period. The following proposition provides the supergradient of L_t with respect to $\boldsymbol{\lambda}_{t,T}$ so that numerical methods (e.g., the subgradient method (Topaloglu 2009) and the cutting-plane method (Brown and Smith 2020)) can be implemented to solve $\mathcal{D}(t,x_t)$.

Proposition 4 (Supergradient). *For a given set of $\boldsymbol{\lambda}_{t,T}$ and state x_t, let ψ be an optimal policy for $L_t(x;\boldsymbol{\lambda}_{t,T})$. For any $\tau \in [t,T]$, the supergradient can be calculated by*

$$\nabla_{\lambda_\tau} L_t(x_t;\boldsymbol{\lambda}_{t,T}) = -(\gamma A)^{\tau-t}x_t - \gamma^{\tau-t}\sum_{\tau'=t+1}^{\tau} A^{\tau-\tau'}\mathbb{E}\{D_{\tau'}\} + \sum_{\tau'=t}^{\tau}\gamma^{\tau'-t}A^{\tau-\tau'}\sum_{i=1}^{M}\psi_{i,\tau'}.$$

Hence,

$$\nabla_{\boldsymbol{\lambda}_{t,T}} L_t(x_t;\boldsymbol{\lambda}_{t,T}) = A_t\left(\sum_{i=1}^{M}\boldsymbol{\psi}_i - \bar{x}_t\right), \tag{11}$$

where $\boldsymbol{\psi}_i^\top = [\psi_{i,t}^\top, \ldots, \psi_{i,T}^\top]$.

The proof can be seen in Online Appendix EC.10. Let $\boldsymbol{\lambda}^*_{1,T}$ be the optimal solution to $\mathcal{D}(1,x_1)$. Then, we can generate a Lagrangian optimal policy ψ^* by solving $L_1(x_1;\boldsymbol{\lambda}^*_{1,T})$ with respect to u_i. If we apply optimal policy ψ^* starting from state x_1 ($\tilde{x}_1 = x_1$) and update the state $\tilde{x}_t$ for $t = 2,\ldots,T$ by iterating over system dynamics $\tilde{x}_{t+1} = f(\tilde{x}_t, \psi^*_t, \mathbb{E}\{D_{t+1}\})$, we can show that the following complementary slackness conditions would hold on the induced state path in expectation.

Proposition 5. *Let $\{\tilde{x}_t\}_{t=1}^T$ be defined as above; then,*

$$\tilde{x}_{k,j,t} - \sum_{i=1}^{M}\psi^*_{k,j,i,t} = 0 \text{ for all } k,j \text{ such that } \lambda^*_{k,j,t} > 0,$$

$$\tilde{x}_{k,j,t} - \sum_{i=1}^{M}\psi^*_{k,j,i,t} \geq 0 \text{ for all } k,j \text{ such that } \lambda^*_{k,j,t} = 0.$$

Note that the above complementary slackness conditions imply that the policy ψ^* derived from Lagrangian relaxation satisfies the linking constraint in expectation from one to T, and the constraints are tight if the corresponding $\lambda^*_{k,j,t} > 0$. We conceptualize the Lagrangian multipliers $\lambda^*_{k,j,t}$ as the shadow price associated with meeting an extra unit of demand from location j in period t rather than postponing it to the next period. When $\lambda^*_{k,j,t}$ is positive, the fulfillment constraint is binding, meaning $\sum_{i=1}^{M}\psi^*_{k,j,i,t} = \tilde{x}_{k,j,t}$. This implies that deferring demand to the next period incurs a higher cost. Therefore, fulfilling an additional unit from location j with fulfillment window k in period t comes with a cost equivalent to $\lambda^*_{k,j,t}$. Conversely, if the fulfillment constraint is nonbinding (i.e., $\sum_{i=1}^{M}\psi^*_{k,j,i,\tau} < \tilde{x}_{k,j,t}$), delaying an additional unit from this location is cost free (i.e., $\lambda^*_{k,j,t} = 0$) for the system.

However, it does not mean that we can directly apply ψ^* into the real system as our heuristic. This is because ψ^* makes the fulfillment decisions to allocate the orders from location j to warehouse i in expectation. If the actual demand deviates from the expected demand, the allocation of orders from j to warehouse i induced by ψ^* may not be sufficient to fulfill the realized demand. To resolve this issue, we develop our heuristics based on ψ^* and prove their performance bounds.

4.2. Scaled Lagrangian Relaxation Policy

We refer to the first heuristic policy as the scaled Lagrangian relaxation (sLR) policy, which handles the tractability and allocation issue. The idea behind sLR is to convert the Lagrangian optimal policy ψ^* to a feasible policy as defined in Algorithm 1.

Algorithm 1 (sLR (Scaled Lagrangian Relaxation Policy))

Input: Demand vector $D_{k,j,t}$, $\forall k \in [K], j \in [N], t \in [T]$, fulfillment cost vector C_i, $\forall i \in [M]$
Output: Policy $\pi^{\text{sLR}} = \{u^{\text{sLR}}_{k,j,i,t}\}$

```
Solve D(1, x_1) to obtain λ*_{1,T}.
Generate an optimal policy ψ*_{k,j,i,t} for all k, j, i, t by
    solving Lagrangian L_1(x_1; λ*_{1,T}) with respect to u_i.
for (t,k,j,i) ∈ [T] × [K] × [N] × [M] do
    Allocate unfulfilled order x_{k,j,t} to warehouse i as
    follows:
    if x̃_{k,j,t} > 0 then
        Set u^{sLR}_{k,j,i,t} = (ψ*_{k,j,i,t} / x̃_{k,j,t}) · x_{k,j,t}
    else
        Set u^{sLR}_{k,j,i,t} = 0
    end
end
```

Recall that Lagrangian optimal policy ψ^* satisfies the complementary slackness conditions in expectation.

This implies that the allocation of the unfulfilled orders from location j in period t with k periods left in the fulfillment window to each warehouse i satisfies the linking constraints in expectation (i.e., $\sum_{i=1}^{M}\psi^*_{k,j,i,t} \le \tilde{x}_{k,j,t}$). Calculating the allocation ratio between $\tilde{x}_{k,j,t}$ and $\psi^*_{k,j,i,t}$ and applying this ratio to allocate the realized unfilled orders, the sLR policy ensures that the allocation of unfulfilled orders to each warehouse adheres to the linking constraints, thereby making $\pi^{\mathrm{sLR}}_{k,j,i,t}$ feasible.

The following results characterize the feasibility of the sLR policy and establish a theoretical upper bound on its gap from the optimal policy. Specifically, we explain the source of the suboptimality of the sLR policy. As illustrated in line 6 in Algorithm 1, the fulfillment decision $u^{\mathrm{sLR}}_{k,j,i,t}$ is entirely correlated with the actual demand realization $x_{k,j,t}$. Thus, if a large realized demand $x_{k,j,t}$ is encountered for some k, j, and t, it is possible that the marginal fulfillment cost from a warehouse i, $\mathcal{C}_i'(e^\top u^{\mathrm{sLR}}_{i,t})$, may exceed the marginal cost of postponing for one period, $\delta_{k,j,t}$. Therefore, the sLR policy may end up shipping the orders from that warehouse even when it is cheaper to postpone them.

Theorem 1. π^{sLR} *is a feasible policy to the original problem. Moreover, the optimality gap is bounded by*

$$V_1^{\mathrm{sLR}}(x) - V_1^*(x) \le \sum_{t=1}^{T}\gamma^{t-1}\cdot\frac{1}{2}\overline{C}\overline{D}MK\sqrt{N}.$$

The proof of Theorem 1 follows several key steps to establish the performance guarantee of the sLR policy π^{sLR}. First, we show that the sLR policy π^{sLR} is an unbiased estimator of the Lagrangian optimal policy ψ^*. This, in turn, implies that the states generated by π^{sLR} are also unbiased estimators of those generated by ψ^*. Next, by utilizing Popoviciu's inequality on variance, we bound the expected difference between the total fulfillment under π^{sLR} and ψ^* in terms of the square root of the variation in the total quantity of fulfillment represented by $\sqrt{Var(\sum_{k,j}u^{\mathrm{sLR}}_{k,j,i,t})}$. Finally, by linking the variation in $u^{\mathrm{sLR}}_{k,j,i,t}$ to the variation in demand realizations ($Var(u^{\mathrm{sLR}}_{k,j,i,t}) \le Var(D_{k,j,t}) + \cdots + Var(D_{K,j,t+k-K})$), we can obtain a performance guarantee that increases at a rate of $\mathcal{O}(MK\sqrt{N})$. The formal proof is provided in Online Appendix EC.6.

Remark 2. When all orders have a K-period fulfillment window upon arrival, the performance bound for the sLR policy can be further improved as $\mathcal{O}(M\sqrt{KN})$. This is because $Var(u^{\mathrm{sLR}}_{k,j,i,t}) \le Var(D_{K,j,t+k-K}) \le \frac{1}{4}\overline{D}^2$. On the other hand, if the demand $D_{k,j,t}$ is not independent across time for each location j, the performance bound for the sLR policy would need to be revised as $\mathcal{O}(MK^2\sqrt{N})$. This is because $\sqrt{Var(\sum_{k,j}u^{\mathrm{sLR}}_{k,j,i,t})} = \sqrt{\sum_j Var(\sum_k u^{\mathrm{sLR}}_{k,j,i,t})}$ and $\sum_k u^{\mathrm{sLR}}_{k,j,i,t} \le \sum_k\sum_{k'=k}^{K} D_{k',j,t+k'-K} \le K^2\overline{D}$.

Remark 3. In the cases where there are Z shipping zones for each warehouse, the theoretical guarantee of the worst-case optimality gap of sLR becomes $\sum_{t=1}^{T}\gamma^{t-1}\cdot\frac{1}{2}\overline{C}\,\overline{D}MK\sqrt{NZ}$. We refer the readers to Online Appendix EC.3 for detailed discussion.

The performance bound characterized in Theorem 1 shows that the optimality gap of the sLR policy is growing linearly with respect to the number of warehouses M and sublinearly with respect to the number of demand locations N. This has the following implications. Given that demand from each location is relatively similar, the total amount of fulfillment increases linearly as the number of locations N increases. Because $\mathcal{C}_i$ is a convex and increasing function of fulfillment, this implies that the total cost grows at least linearly in N. Because the optimality gap grows sublinearly in N, whereas the total cost grows linearly in N, the relative optimality gap converges to zero as N increases. Consequently, our policy becomes asymptotically optimal in large N. This implication is particularly crucial in online retail scenarios, where there are typically only a few warehouses (small M) catering to a vast number of demand locations (large N). Indeed, the asymptotic optimal of our policy regarding N is highly relevant for real-world applications, such as our e-commerce partner that operates three warehouses (i.e., $M = 3$) and serves over 10,000 demand locations (i.e., $N > 10{,}000$).

In contrast, our bound does not imply asymptotic optimality in K, not even in the special cases mentioned in Remark 2. This is because the optimal value function does not necessarily increase as K grows. Despite the absence of a theoretical guarantee, the numerical experiments conducted in Section 6.1 indicate that increasing K helps reduce fulfillment costs, suggesting that multiperiod fulfillment windows can effectively alleviate fulfillment capacity constraints. Additionally, this bound may highlight that sLR might not fully leverage the advantage of the fulfillment window. Indeed, although the total cost in sLR decreases as K increases (see Section 6.1), the reduction rate is slower than that of the optimal policy, causing the gap to widen slightly.

Finally, regarding the horizon T, our performance bound reveals only a linear rate for the optimality gap. Specifically, the optimality gap grows at most at a rate of $\mathcal{O}(T)$ because $\sum_{t=1}^{T}\gamma^{t-1} \le T$ when $0 < \gamma \le 1$.

We also point out that to implement the sLR policy, it is only required to solve for the optimal Lagrangian multipliers $\boldsymbol{\lambda}^*_{1,T}$ and the corresponding Lagrangian optimal policy ψ^* once at the beginning of the horizon and store both ψ^* and its generated states $\{\tilde{x}_t\}$. After that, we can directly obtain feasible actions based on the stored information. This single-shot optimization guarantees computational efficiency when implementing the sLR policy.

Although the sLR policy possesses a good theoretical performance guarantee, our numerical experiments in Section 6.1 reveal one important drawback. More specifically, the sLR policy allocates the current unfulfilled orders to M warehouses based on the percentages computed from the Lagrangian policy. However, this allocation rule does not take into account the existence of thresholds for each warehouse as characterized in Proposition 2. Therefore, following the sLR policy blindly without considering the thresholds may eventually allocate too many orders prematurely before the end of their fulfillment windows. This drawback serves as motivation to propose a second and improved heuristic policy in the following section.

4.3. Threshold Lagrangian Relaxation Policy

Proposition 2 provides two insights for the development of an improved heuristic, which we refer to as the tLR policy. First, the total number of orders allocated to each warehouse i should not exceed a certain threshold. Second, the orders to be fulfilled should be prioritized according to a specific criterion. Proposition 2 also provides an easy-to-compute metric denoted by discounted penalty cost $\gamma^{\min\{k-1,T+1-t\}}b_j$ to determine both the threshold for each warehouse and the prioritization scheme for the orders.

Building on these insights, we develop the tLR policy as follows. First, we prioritize orders to be fulfilled based on the descending order of the discounted penalty cost $\gamma^{\min\{k-1,T+1-t\}}b_j$ (i.e., orders with the largest penalty get fulfilled first). Second, we allocate the orders indexed by (k,j) to a warehouse i based on the sLR allocation rule in an iterative fashion. After each allocation, we check whether the marginal fulfillment cost at each warehouse i is smaller than $\gamma^{\min\{k-1,T+1-t\}}b_j$. If yes, we keep warehouse i eligible for the next allocation. Otherwise, we close warehouse i to ensure that the total number of orders fulfilled from that warehouse does not exceed a threshold. Under the tLR policy, we guarantee $\mathcal{C}'_i(e^\top u^{\text{tLR}}_{it}) \le \gamma^{\min\{k-1,T+1-t\}}b_j$ if an order (k,j) is allocated to a warehouse i, which is consistent with the first optimal structure in Proposition 2. The detailed algorithm tLR is given in Algorithm 2.

Algorithm 2 (tLR (Threshold Lagrangian Relaxation Policy))

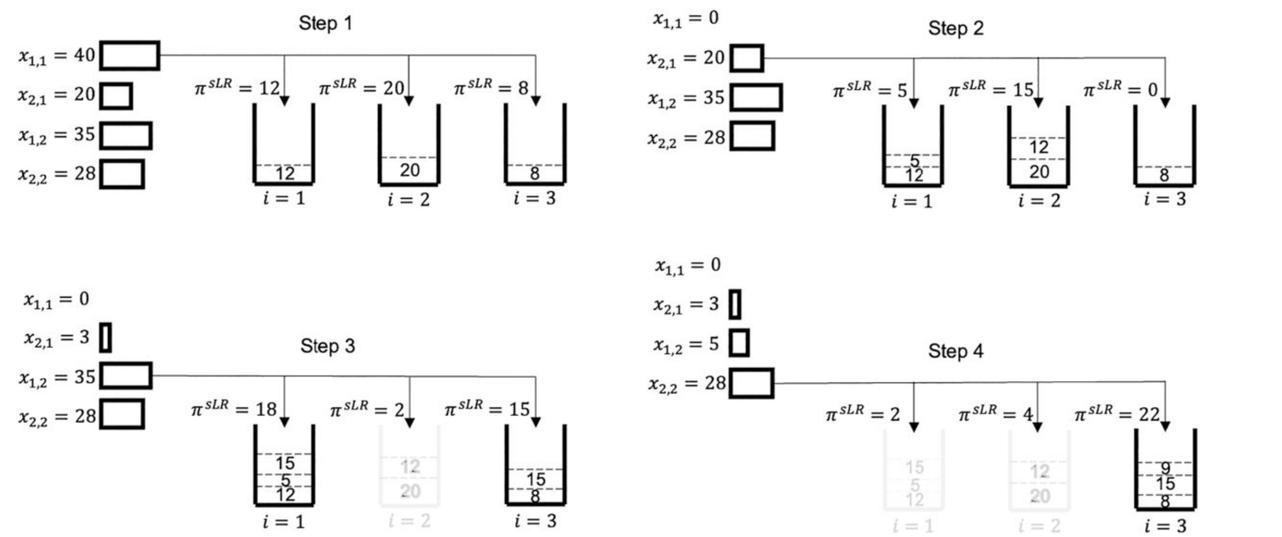
Input: Demand vector $D_{k,j,t}$, $\forall k \in [K], j \in [N], t \in [T]$, fulfillment cost vector C_i, $\forall i \in [M]$
Output: Policy $\pi^{\text{tLR}} = \{u^{\text{tLR}}_{k,j,i,t}\}$
Initialize u^{tLR} with the output from Algorithm 1.
for $t \in [T]$ **do**
  Sort orders indexed by (k,j) by $\gamma^{\min\{k-1,T+1-t\}}b_j$ in descending order.
  Let $\Pi(k,j) : [K] \times [N] \to Z^+$ be the rank of (k,j).
  Set $U^{tLR}_{i,t} = 0$ for all I, and set $\mathcal{I}_t = [M]$.
  for $(k,j) \in [K] \times [N]$ *in descending order* $\Pi(k,j)$ **do**
    for $i \in \mathcal{I}_t$ **do**
      Modify fulfillment decisions for warehouse i as follows:
      if $\mathcal{C}'_i(U^{\text{tLR}}_{i,t} + u^{tLR}_{k,j,i,t}) > \gamma^{\min\{k-1,T+1-t\}}b_j$ **then**
        if $\mathcal{C}'_i(U^{\text{tLR}}_{i,t}) > \gamma^{\min\{k-1,T+1-t\}}b_j$ **then**
          Set $u^{tLR}_{k,j,i,t} = 0$.
        else
          Set $u^{tLR}_{k,j,i,t} = x$ that solves $\mathcal{C}'_i(U^{\text{tLR}}_{i,t} + x) = \gamma^{\min\{k-1,T+1-t\}}b_j$.
        end
        Remove warehouse i from set $\mathcal{I}_t$.
      end
      Set $U^{\text{tLR}}_{i,t} = U^{\text{tLR}}_{i,t} + u^{tLR}_{k,j,i,t}$.
    end
  end
end

We illustrate the tLR policy as shown in Figure 2. For illustration, we consider two locations, two fulfillment windows, and three warehouses (i.e., $N = 2, K = 2$, and $M = 3$). Let us assume that $b_1 = 1$ and $b_2 = 0.5$, $\gamma = 0.8$, and the marginal fulfillment cost satisfies the following conditions. For all $i = 1,2,3$, $\mathcal{C}'_i(x) = 0$ if $x \le 32$, and $\mathcal{C}'_i(x) = 2$ if $x > 32$. We restrict our attention to the allocations in period t, where $t \le T - 1$. Suppose that we compute the allocations under the sLR policy and find out that the allocations are as follows.

According the tLR policy, first, we sort indices (k,j) by $\gamma^{\min\{k-1,T+1-t\}}b_j$ in descending order. Using the given parameters, the resulting criteria generate the following ordering: $(1,1) \succ (2,1) \succ (1,2) \succ (2,2)$. As illustrated in Figure 2, we then iterate over each order index (k,j) and allocate the unfulfilled orders to three warehouses in four steps as follows.

- Step 1. Because $\mathcal{C}'_i(0) = 0 < b_1 = 1$ for all $i = 1,2,3$, we complete all of the allocations from $x_{1,1,t}$ according to the sLR policy shown in Table 1.
- Step 2. The sLR policy stipulates the allocation of $x_{2,1,t} = 20$ as 5 units to warehouse 1 and 15 to warehouse 2. Because $\mathcal{C}'_1(17) = 0 < \gamma b_1 = 0.8$, we allocate five units to warehouse 1. However, because $\mathcal{C}'_2(35) = 2 > \gamma b_1 = 0.8$ but $\mathcal{C}'_2(32) = 0 < \gamma b_1 = 0.8$, we allocate only 12 units to warehouse 2 and set that warehouse

Figure 2. Illustration of the tLR Policy for a System with $M = 3$, $N = 2$, and $K = 2$



ineligible for further fulfillment. The remaining three units are deferred to the next period.

- Step 3. The sLR policy stipulates 18 units to warehouse 1, 2 units to warehouse 2, and 15 units to warehouse 3. Because $\mathcal{C}_1'(35) = 2 > b_1 = 0.5$, we cannot allocate 18 units to warehouse 1. Instead, only 15 units are allocated to warehouse 1, and warehouse 1 is then set ineligible. Because warehouse 2 is already ineligible, no unit is allocated to that warehouse. Finally, we allocate all 15 units to warehouse 3. After the fulfillment decisions, we defer the remaining five units (three units from warehouse 1 and two units from warehouse 2) to the next period.
- Step 4. The sLR policy stipulates 2 units to warehouse 1, 4 units to warehouse 2, and 22 units to warehouse 3. However, both warehouses 1 and 2 are ineligible, and only nine units can be allocated to warehouse 3. The remaining 19 units are deferred to the next period.

To demonstrate the performance bound for the tLR policy, we first establish that the tLR policy always outperforms the sLR policy.

Proposition 6. *The sLR and tLR policies satisfy the following relation*:

$$V_1^{\mathrm{tLR}}(x) \le V_1^{\mathrm{sLR}}(x).$$

Table 1. Fulfillment Decisions Under the sLR Policy

Orders	Allocation to Warehouses		
	$i = 1$	$i = 2$	$i = 3$
$x_{1,1,t} = 40$	12	20	8
$x_{2,1,t} = 20$	5	15	0
$x_{1,2,t} = 35$	18	2	15
$x_{2,2,t} = 28$	2	4	22

We defer the formal proof to Online Appendix EC.7 and only describe the main ideas below. The difference of $V_1^{\mathrm{tLR}}(x)$ and $V_1^{\mathrm{sLR}}(x)$ can be obtained by a sum of cost differences $\{e_t\}_{t=1}^{T}$. Each e_t consists of three parts: the difference in fulfillment cost and the penalty cost between π^{tLR} and π^{sLR} and the difference in the expected cost to go by following π^{tLR} with different states. According to Algorithm 2, when one warehouse i is set ineligible after fulfilling some order indexed by (k_i, j_i), we have $\mathcal{C}_i'(\sum_{k,j} u_{k,j,i,t}^{\mathrm{tLR}}) \le \gamma^{\min\{k_i-1, T+1-t\}} b_{j_i}$ and $u_{k,j,i,t}^{\mathrm{tLR}} = 0$ for all orders (k,j) ranked after (k_i, j_i). The difference of the quantity of fulfillment under two policies because of removing warehouse i equals $\sum_{(k,j)\in\underline{\Theta}_i:} u_{k,j,i,t}^{\mathrm{sLR}} + u_{k_i,j_i,i,t}^{\mathrm{sLR}} - u_{k_i,j_i,i,t}^{\mathrm{tLR}}$, where $\underline{\Theta}_i$ is the index set of orders that ranked after (k_i, j_i). Therefore, π^{sLR} has an additional fulfillment cost larger than $\sum_{i\in\mathcal{I}_t} \gamma^{\min\{k_i-1, T+1-t\}} b_{j_i}(\sum_{(k,j)\in\underline{\Theta}_i} u_{k,j,i,t}^{\mathrm{sLR}} + u_{k_i,j_i,i,t}^{\mathrm{sLR}} - u_{k_i,j_i,i,t}^{\mathrm{tLR}})$, where $\mathcal{I}_t$ is the index set of removed warehouses at period t. On the other hand, these $\sum_{(k,j)\in\underline{\Theta}_i:} u_{k,j,i,t}^{\mathrm{sLR}} + u_{k_i,j_i,i,t}^{\mathrm{sLR}} - u_{k_i,j_i,i,t}^{\mathrm{tLR}}$ units are either lost or delayed under π^{tLR} with marginal cost no larger than $\gamma^{\min\{k_i-1, T+1-t\}} b_{j_i}$ for all $i \in \mathcal{I}_t$ by the definition of Algorithm 2. Therefore, we must have $e_t \le 0$ for all t, which implies that π^{tLR} always performs better than π^{sLR}.

Combining Proposition 6 with Theorem 1, we have the same performance bound for π^{tLR}.

Corollary 1. *The optimality gap of the tLR policy is bounded by*

$$V_1^{\mathrm{tLR}}(x) - V_1^*(x) \le \sum_{t=1}^{T} \gamma^{t-1} \cdot \frac{1}{2} \overline{CD} M K \sqrt{N}.$$

We conclude this section by discussing the conditions under which the sLR and tLR policies defer an order to a later period, noting that both policies employ different strategies for determining postponed orders. For the sLR policy, an order is postponed only if the Lagrangian optimal policy dictates it (i.e., $\sum_i \psi^*_{k,j,i,t} < \tilde{x}_{k,j,t}$ for some k and j). This is because, according to the scaling law in line 6 in Algorithm 1, $\sum_i u^{\mathrm{sLR}}_{k,j,i,t} < x_{k,j,t}$ only when $\sum_i \psi^*_{k,j,i,t} < \tilde{x}_{k,j,t}$. However, this type of postponement does not necessarily satisfy the structural properties of an optimal fulfillment policy outlined in Proposition 2. In contrast, the tLR policy *actively* postpones orders with lower discounted penalty costs to preserve the structural properties of an optimal policy. As illustrated in Figure 2, the tLR policy ranks orders (k,j) in descending order and fulfills them sequentially. Orders are deferred only when all warehouses reach their capacity determined by the criterion $\mathcal{C}'_i(e^\top u^{\mathrm{tLR}}_{it}) \le \gamma^{\min\{k-1, T+1-t\}} b_j$.

5. Extension to Infinite Horizon

In this section, we extend the tLR policy to a problem with infinite horizon and stationary demand over time at each location. Specifically, we denote the value function under some policy π for some state x by

$$J^\pi(x) = \lim_{T\to\infty} \mathbb{E}\left\{ \sum_{t=1}^{T} \gamma^t g(x_t, \pi_t(x_t)) | x_1 = x \right\},$$

where $\gamma \in (0,1)$. The DM wants to minimize the total discounted expected cost in the long run: that is,

$$J^*(x) = \min_{\pi \in \mathbf{\Pi}} J^\pi(x).$$

Standard DP arguments (e.g., Puterman 2014) imply that there is an optimal solution such that

$$J^*(x) = \min_{u \in \mathcal{U}(x)} g(x,u) + \gamma \mathbb{E}\{J^*(x') | x, u\},$$
$$J^*(x) = \lim_{T\to\infty} V^*_1(x),$$

where $V^*_1(x)$ is the optimal value function for a T-period finite horizon problem starting with x and ending with any bounded terminating function. Following Proposition 1, we have the following characteristics of J^* and an optimal policy that supports us to apply the tLR policy to an infinite horizon problem.

Corollary 2 (Optimal Value Function and Optimal Policy for Infinite Horizon).

1. *J^* is real valued, bounded, convex, and nondecreasing in $x \in \mathcal{S}$.*
2. *The "MW-NL-KP" problem with infinite horizon has a warehouse-specific state-dependent threshold-type stationary optimal fulfillment policy.*

Because of the curse of dimensionality, it is practically impossible to compute J^*. Hence, we adopt the tLR heuristic policy in Algorithm 2 such that if the Lagrangian optimal policy $\psi^* = (\psi^*_t : t \ge 1)$ is already available, we can obtain the performance guarantee as

$$\frac{\overline{C}\,\overline{D}MK\sqrt{N}}{2(1-\gamma)}$$

because Corollary 1 can be extended to an infinite horizon. Unfortunately, ψ^* is not easily obtainable in the infinite horizon as the dimension of the optimal Lagrangian multipliers in the infinite horizon case is unbounded. Similar to Brown and Zhang (2025), we consider a truncated tLR policy.

Definition 1 (Truncated Threshold Lagrangian Relaxation Policy). Given a starting state x, a truncated tLR policy $\tilde{\pi}^{\mathrm{tLR}}$ generates actions according to π^{tLR} for a T-period problem if $t \le T$ and generates actions according to an arbitrary feasible policy if $t > T$.

We can still have a similar performance guarantee if we select T properly. The following theorem formally states the results, whose proof can be found in Online Appendix EC.8. We point out that the value of T determined as in the following theorem makes the truncated tLR policy tractable in practice. For example, for a system with $N = 100$, $K = 2$, and $\gamma = 0.99$, we only need to calculate the Lagrangian optimal policy ψ^* up to $T = 265$ periods.

Theorem 2. *The truncated tLR policy $\tilde{\pi}^{\mathrm{tLR}}$ satisfies*

$$\tilde{J}^{\mathrm{tLR}}(x) - J^*(x) \le \frac{1}{2(1-\gamma)} \overline{C}\overline{D}M\sqrt{N} + \frac{(2-\gamma)\gamma^T}{1-\gamma} \max\{\overline{b}, \overline{C}\}\overline{D}KN.$$

In particular, if $T = \log_{\frac{1}{\gamma}} \sqrt{N}$,

$$\tilde{J}^{\mathrm{tLR}}(x) - J^*(x) \le \frac{1}{2(1-\gamma)} \overline{C}\overline{D}M\sqrt{N} + \frac{2-\gamma}{1-\gamma} \max\{\overline{b}, \overline{C}\}\overline{D}K\sqrt{N}.$$

6. Numerical Study

This section provides the performance evaluation of two heuristic policies through two experiments. Using synthetic data, experiment 1 involves a sensitivity analysis concerning various model parameters. Experiment 2 evaluates the performance using a unique data set from our partner e-commerce company. We report the average performance and standard error based on 50 simulations conducted in each experiment. Experiment 1 investigated both sLR and tLR policies, whereas experiment 2 tested only the tLR policy.

In addition to sLR and tLR, we included two benchmark policies in our study: the *fulfill-all* (FA) policy and the *myopic* (MYO) policy. The FA policy opts for the optimal action to fulfill all received orders within the current period, explicitly avoiding order deferral.

Under the MYO policy, fulfillment decisions are made short sighted, prioritizing the minimization of immediate costs plus penalties for any orders not assigned. The decision-making process involves solving a convex optimization problem for each period t to determine u_t^*. The DM assigns fulfillment according to u_t^*. Orders not assigned $(x_t - Gu_t^*)$ and having $k=1$ are subjected to penalties and exit the system, whereas others are deferred with their fulfillment window decremented by one. There is a significant difference between the FA and MYO policies. FA optimizes the shipping cost for the current period whereas enforcing a strict constraint that prohibits any carryover of remaining orders from one period to the next. On the other hand, MYO also optimizes the shipping cost for the current period but does not enforce the same strict constraint as FA regarding carryover orders. This means that MYO may not fulfill all orders and may incur penalty costs for orders that are not satisfied within their respective fulfillment periods. In addition, we also consider a Lagrangian approximation (LA) policy, which use the Lagrangian function $L_{t+1}(x_{t+1}; \boldsymbol{\lambda}^*_{t+1,T})$ to approximate the optimal value function $V^*_{t+1}(x_{t+1})$. The LA policy extends the MYO policy by leveraging knowledge of future fulfillment expenses through the use of optimal Lagrangian dual variables. Although LA performs well with a small number of locations, its performance deteriorates similarly to that of the MYO policy as the number of locations increases. Detailed implementation and further discussion are provided in Online Appendix EC.11.

The effectiveness of each policy is assessed by comparing the gap between its value function and an LB of the optimal value function. The LB is derived from calculating $L_1(x; \lambda^*_{1,T})$. Detailed explanations of the implementation for all benchmark policies are delineated in Online Appendix EC.11.

6.1. Experiment 1: Synthetic Data

In this section, we examine the performance of sLR and tLR under various scenarios by analyzing different parameters, including the size of the fulfillment network, the length of the fulfillment window, and the mean and coefficient of variation (CoV) of the demand distribution.

For the baseline setup, we assume that the e-commerce retailer has a fulfillment network of two warehouses ($M=2$) and 50 demand locations ($N=50$). Each demand location represents orders received from one region. The demand in each period follows a stationary negative binomial distribution with a mean of $\mu = 80$ and a standard deviation of $\sigma = 120$, resulting in a CoV of 1.5. A negative binomial distribution is chosen to independently control the mean and standard deviation. Each order has two periods ($K=2$) to be fulfilled, and the fulfillment cost function for each warehouse takes a quadratic form (i.e., $\mathcal{C}_i(x) = \alpha_i x^2$ with $\alpha_i = 2 \times 10^{-4}$). The unit lost sale cost is set to $b_j = 1$ for all locations. According to Proposition 2, each warehouse can at most fulfill $(\mathcal{C}_i')^{-1}(\bar{b}) = \frac{1}{2\times 2\times 10^{-4}} = 2500$ units as any extra unit of fulfillment is more expensive than lost sales. As a result, the total periodic fulfillment capacity is 5,000 units. The average demand is 4,000 units ($80 \times 50 = 4000$). We set the discounting factor to $\gamma = 0.99$ and the simulated horizon to $T = 265$. We measure performance using the *weighted average optimality gap* defined as

$$\Delta^\pi = \frac{V^\pi(x) - V^{\mathrm{LB}}(x)}{\sum_{t=0}^{T} \gamma^t}.$$

We conduct a sensitivity analysis by changing one parameter at a time during the numerical analysis. The results are reported below.

6.1.1. Impact of the Number of Locations.

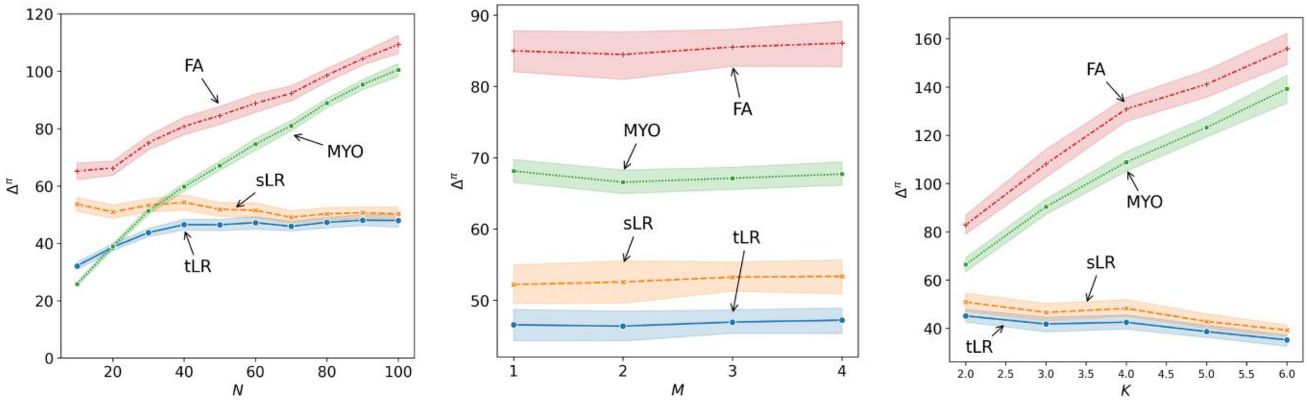
We investigate the impact of the number of locations on different policies by changing N from 10 to 100. We assume that demand from each location follows the same negative binomial distribution. In the meantime, we also scale α_i by $50/N$ to guarantee the constant ratio between demand and fulfillment capacity. The left panel of Figure 3 presents the results.

We observe that the optimality gap of the tLR policy increases sublinearly with N, whereas that of the sLR policy shows little change. Moreover, the tLR policy consistently outperforms the sLR policy, but this advantage diminishes as N increases. Comparing the performance of the tLR policy with the benchmark policies, we find that the tLR policy outperforms both benchmark policies when N is greater than 20. When $N=10$, the MYO policy achieves the smallest optimality gap. The FA policy performs the worst across all cases, indicating that fulfilling all orders under the logistic resource capacity constraint may not be a viable policy. As N increases, the MYO policy's performance becomes similar to the FA policy's performance. This can be attributed to the fact that when the number of locations is high, the realized demand per location tends to converge to its expected value, and the system has sufficient capacity for fulfillment on average. However, the performance difference between the heuristic and benchmark policies grows as N increases. This implies that even when the system has sufficient capacity on average, the tLR policy benefits from balancing the demand from period to period by actively delaying fulfillment.

6.1.2. Impact of the Number of Warehouses.

Similar to the previous study, we vary the number of warehouses but maintain a constant ratio between demand and fulfillment capacity to investigate the impact of M

Figure 3. (Color online) Change in the Optimality Gap with Respect to the Number of Locations N (Left Panel), the Number of Warehouses M (Center Panel), and the Number of Fulfillment Windows K (Right Panel)



Note. The shades of the curves represent standard errors.

on the policies. As shown in the center panel of Figure 3, we change M from one to four, but the performances for all policies are barely affected. Such results together with the previous results demonstrate that the tLR policy can be adopted by online retailers with various-sized fulfillment networks.

6.1.3. Impact of the Length of the Fulfillment Window. We evaluate the impact of the length of the fulfillment window by varying K from one to six. When $K = 1$, orders cannot be delayed. Both the tLR policy and the MYO policy are optimal, and the sLR policy performs just like the FA policy. When $K > 1$, delay fulfillment is allowed, and the right panel of Figure 3 presents the results. Consistent with the previous analysis, the tLR policy performs best, whereas the FA policy performs the worst. The optimality gap of both sLR and tLR decreases slightly as K increases, whereas that of the benchmark policies increases.

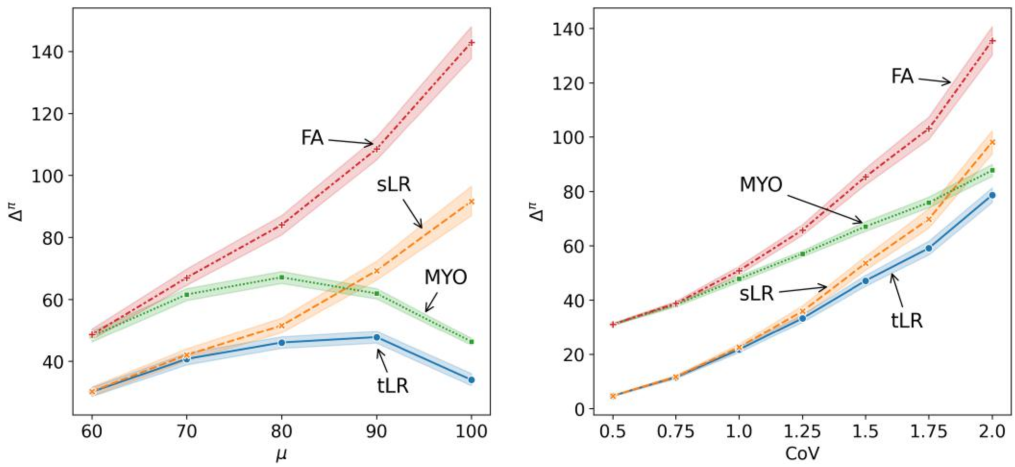
Besides comparing different fulfillment policies, it may be more important to understand the value of having multiperiod fulfillment windows and allowing orders to be postponed if necessary. Therefore, we compare the weighted average costs of different policies when $K > 1$ with that of the lower bound when $K = 1$. Our simulation shows that the total fulfillment costs exhibit a decreasing relationship with K for all policies, including LB, and the effect is diminishing (Figure EC.7 in the online appendix). This implies that having longer windows is beneficial. Specifically, allowing orders to be postponed by just one period (i.e., $K = 2$) can reduce the total cost by 3.3% for sLR, 3.6% for tLR, and 2.4% for MYO. This reduction further improves to 5.2% for sLR, 5.5% for tLR, and 2.5% for MYO when two-period postponement is allowed (i.e., $K = 3$). However, the benefit diminishes as K increases further. The cost reductions become 5.8% for sLR, 6.1% for tLR, and 2.5% for MYO when $K = 4$.

6.1.4. Impact of Demand Distribution. We study the impact of the expectation and CoV of the demand distribution on the performance of the policies. To accomplish the first task, we vary the expectation μ from 60 to 100 without changing the fulfillment capacity (5,000 units per period). The increase in μ implies that the fulfillment capacity becomes less sufficient. In particular, $\mu = 60$ implies that the average demand is 3,000 units, whereas $\mu = 100$ implies that the average demand is 5,000 units. Consequently, we expect all policies except for FA to exhibit improved performance because even an optimal policy cannot do much to reduce the costs. The results are presented in the left panel of Figure 4.

We observe that Δ^{FA} increases as μ grows. This suggests that fulfilling all orders at each period when the fulfillment capacity is insufficient can only harm the company. The sLR policy performs similarly to the FA policy because there is no limitation on fulfillment in the sLR policy. The tLR and MYO policies exhibit distinct performances relative to the FA and sLR policies; the optimality gap reduces as μ approaches 60 or 100. Although the two policies exhibit similarities, the tLR policy outperforms the MYO policy because of its ability to foresee future fulfillment requirements and balance the load accordingly.

For the second task, we vary the CoV from 0.5 to 2. A larger CoV indicates more volatile demand, and hence, we expect the tLR policies to perform better because they actively balance demand from one period to another. The right panel of Figure 4 displays the results. We observe that the optimality gaps of all policies increase as CoV increases, which is not surprising. One interesting observation is that tLR and MYO share similar performances when CoV is large. Note that this could be because when demand is volatile, too much demand will likely lead to penalties in both policies.

Figure 4. (Color online) Change in the Optimality Gap with Respect to Expected Demand μ (Left Panel) and Coefficient of Variation *COV* (Right Panel)



Note. The shades of the curves represent standard errors.

6.2. Experiment 2: Real Data

This section examines the performance of the tLR policy with real data obtained from our partner e-commerce company that sells computer hardware and consumer electronics in North America. The retailer operates three warehouses ($M = 3$) and fulfills orders using third-party carriers. The data set comprises transaction-level data for over 39,000 stock-keeping units (SKUs) from July to December 2012. It includes over 5 million transaction records (i.e., orders), and each record provides comprehensive details about the online order, such as SKU, product weight and quantity, shipped warehouse, customer location (five-digit zip code), carrier, shipping cost, and product price.

To align the data with our model, we first calculated each order's total weight (respectively, value) by summing all included products' weights (respectively, price). We excluded orders that utilized multiple shipping services as our model does not address order consolidation. Additionally, we filtered out orders not handled by the three primary LTL carriers. The resulting data set comprises 4.84 million orders over 161 sales days ($T = 161$). We aggregated the daily demand at the state level from all remaining orders, resulting in an average daily fulfillment requirement of 118,533 pounds.

Assuming that the demand from each state is stationary and independent, we estimated the demand expectations using the state-wise empirical average. This approach is consistent with our heuristic, which does not require detailed knowledge of the demand distribution. We further validated these estimates using parametric methods to ensure their reliability.

Additionally, we developed three supplementary scenarios to rigorously test our heuristic performance across varied conditions. In the first scenario, we randomly selected 25% of the orders to simulate the operations of a smaller company. In the second and third scenarios, we altered the demand distribution by categorizing orders based on their total value: low-value orders defined as those between the minimum and the 75th percentile of order values and high-value orders defined as those between the 75th percentile and the maximum value. Daily demands for both low-value and high-value orders were also aggregated at the state level following the same procedure. The primary goal of these adjustments is to utilize the available data set fully and generate diverse testing environments to robustly demonstrate the efficacy of our heuristics.

Next, we calibrate the fulfillment cost function with the selected data set. We assume that all states belong to the same shipping zone (we provide another simulation with multiple shipping zones in Online Appendix EC.12.1) and that the fulfillment cost functions are identical for all warehouses. We focus on three primary carriers (UPS, FedEx, and DHL) and assume that each carrier guarantees some capacity at a fixed unit rate with an LTL contract. Specifically, the unit fulfillment rate for each carrier is determined based on the median ground shipping cost, and the mean daily fulfillment weights calibrate the capacity commitment. Therefore, the fulfillment costs are calibrated as a piecewise linear and convex function. In contrast, the rate in the spot market is determined by the median cost of the fastest fulfillment services (e.g., "UPS Next Day Air" and "FedEx 1Day"). Further details on the selected data set, the expectation estimation, and cost calibration can be found in Online Appendix EC.12.

We employ bootstrapping to generate sample paths for the simulation. We resample the daily demand for all states from the selected data set. Hence, each sample path consists of a subset of the randomized demand history. We then apply tLR and benchmark policies to

Table 2. Relative Optimality Gap (Percentage) of Different Policies with Real Data

	All orders		25% Orders		Low-value orders		High-value orders	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
sLR	27.53	5.22	29.76	5.53	27.97	5.44	32.08	5.66
tLR	**2.23**	0.48	**2.13**	0.45	**2.05**	0.54	**2.36**	0.6
MYO	4.85	0.79	5.08	1.08	4.73	0.89	5.34	1.14
FA	39.61	4.54	41.59	5.25	38.32	4.87	45.86	5.33

Note. The bold font highlights the best performance among all policies.

each sample path and investigate their performance in terms of the relative optimality gap, which is calculated as $(V^{\pi} - V^{\text{LB}})/V^{\text{LB}}$. The results can be found in Table 2.

The tLR policy exhibits the best performance and outperforms the benchmark policies in all four study cases, which is consistent with the results obtained from the synthetic data. The FA policy has the worst performance, implying that executing all orders every period is unsustainable in scenarios where cheaper fulfillment capacity is scarce. Surprisingly, although sLR performs comparable with tLR with synthetic data, it performs poorly with real data in all four cases. This highlights the necessity of checking fulfillment thresholds to guarantee good performance. A comparison between the MYO and FA policies reveals that an online retailer can diminish the optimality gap by over 85% by allowing carryover of remaining orders from one period to the next. This observation underscores the potential for e-commerce firms to effectively mitigate logistical capacity limitations by applying a multiperiod fulfillment window. Moreover, if a policy in practice, such as the tLR policy, incorporates information regarding deferred fulfillment, the optimality gap could be decreased by as much as 95%. This emphasizes the significance of an efficient fulfillment strategy for e-commerce firms operating under restricted logistic capacity.

7. Conclusion

This paper provides insights into how e-commerce companies should manage fulfillment when their capacity is limited because of using LTL services. The proposed heuristic policy based on Lagrangian relaxation approximates the optimal policy by emphasizing order prioritization based on the discounted penalty cost and dynamically managing warehouse allocations to ensure that the total number of fulfilled orders from each warehouse does not surpass a predetermined limit. The performance analysis of the proposed heuristics proves that it is asymptotically optimal in various scenarios. Through numerical studies conducted with synthetically generated data sets and data provided by our partner e-commerce retailer, the importance of considering logistic capacity constraints, allowing multiperiod fulfillment windows, and incorporating remaining fulfillment window information into the decision-making process is demonstrated. These findings have practical implications for online retailers in terms of offering different fulfillment options to customers and managing their logistic capacity effectively. For example, online retailers are able to alleviate the logistic capacity constraint to some extent by offering their customers a "two-day fulfillment" option. Customers who wish to have faster fulfillment may choose to pay an additional fee.

One direction to extend the current research is to study the proper LTL capacity of online retailers. As observed in the paper, the fulfillment cost is highly related to the LTL rate and capacity. An online retailer may negotiate a contract with a third-party logistic service supplier to secure fulfillment capacity at a cost as a proper amount of capacity can lead to a lower unit cost and guarantee the retailer's service rate.

Another research direction is to find a simpler and stationary policy for the infinite horizon. Although our heuristic is already simple and implementable, we conjecture that there may be an even simpler and stationary policy worth investigating. Investigating the independence of the Lagrangian optimal policy to the initial state and exploring alternative approaches to finding a stationary policy without solving the Lagrangian dual problem are promising avenues for further study.

References

Acimovic J, Farias VF (2019) The fulfillment-optimization problem. *Oper. Res. Management Sci. Age Analytics* 2019(October):218–237.

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

Adelman D, Mersereau AJ (2008) Relaxations of weakly coupled stochastic dynamic programs. *Oper. Res.* 56(3):712–727.

Amazon (2024) Annual fulfillment expenses of Amazon from 2001 to 2023 (in million U.S. dollars). Accessed February 20, 2024, https://www.statista.com/statistics/236456/amazons-annual-fulfillment-expenses/.

Andrews JM, Farias VF, Khojandi AI, Yan CM (2019) Primal–dual algorithms for order fulfillment at Urban Outfitters, Inc. *INFORMS J. Appl. Analytics* 49(5):355–370.

Balseiro SR, Besbes O, Pizarro D (2024) Survey of dynamic resource-constrained reward collection problems: Unified model and analysis. *Oper. Res.* 72(5):2168–2189.

BLS (2022) Heavy and tractor-trailer truck drivers. Accessed February 20, 2024, https://www.bls.gov/ooh/transportation-and-material-moving/heavy-and-tractor-trailer-truck-drivers.htm.

Brown DB, Smith JE (2020) Index policies and performance bounds for dynamic selection problems. *Management Sci.* 66(7):3029–3050.

Brown DB, Zhang J (2022) Dynamic programs with shared resources and signals: Dynamic fluid policies and asymptotic optimality. *Oper. Res.* 70(5):3015–3033.

Brown DB, Zhang J (2023) Technical note—On the strength of relaxations of weakly coupled stochastic dynamic programs. *Oper. Res.* 71(6):2374–2389.

Brown DB, Zhang J (2025) Fluid policies, reoptimization, and performance guarantees in dynamic resource allocation. *Oper. Res.* 73(2):1029–1045.

eMarketer (2023) Retail e-commerce sales worldwide from 2014 to 2027 (in billion U.S. dollars). Accessed March 16, 2023, https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/.

Fuller S (2021) The rise of LTL: How e-commerce changed the middle mile. Accessed March 16, 2023, https://www.supplychaindive.com/news/logistics-ltl-trucking-middle-mile-e-commerce/604842/.

Harsha P, Subramanian S, Uichanco J (2019) Dynamic pricing of omnichannel inventories. *Manufacturing Service Oper. Management* 21(1):47–65.

Hawkins JT (2003) A Langrangian decomposition approach to weakly coupled dynamic optimization problems and its applications. PhD thesis, Massachusetts Institute of Technology, Cambridge.

Jasin S, Kumar S (2012) A re-solving heuristic with bounded revenue loss for network revenue management with customer choice. *Math. Oper. Res.* 37(2):313–345.

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

JMR (2022) [January market report] Transpacific rates and space situation updates. Accessed February 20, 2024, https://www.jmrodgers.com/2022-january-market-report-transpacific-rates-and-space-situation-updates/.

Kriewaldt R (2022) What does the labor shortage mean for your supply chain? Accessed February 20, 2024, https://www.inboundlogistics.com/articles/what-does-the-labor-shortage-mean-for-your-supply-chain/.

Lei Y, Jasin S, Sinha A (2018) Joint dynamic pricing and order fulfillment for e-commerce retailers. *Manufacturing Service Oper. Management* 20(2):269–284.

Lei Y, Jasin S, Uichanco J, Vakhutinsky A (2022) Joint product framing (display, ranking, pricing) and order fulfillment under the multinomial logit model for e-commerce retailers. *Manufacturing Service Oper. Management* 24(3):1529–1546.

Mahar S, Wright PD (2009) The value of postponing online fulfillment decisions in multi-channel retail/e-tail organizations. *Comput. Oper. Res.* 36(11):3061–3072.

Marklund J, Rosling K (2012) Lower bounds and heuristics for supply chain stock allocation. *Oper. Res.* 60(1):92–105.

Miao S, Jasin S, Chao X (2022) Asymptotically optimal Lagrangian policies for multi-warehouse, multi-store systems with lost sales. *Oper. Res.* 70(1):141–159.

Nambiar M, Simchi-Levi D, Wang H (2021) Dynamic inventory allocation with demand learning for seasonal goods. *Production Oper. Management* 30(3):750–765.

Newton E (2023) How did ecommerce change the middle mile. Accessed March 16, 2023, https://www.fieldtechnologiesonline.com/doc/how-did-ecommerce-change-the-middle-mile-0001.

Porteus EL (2002) *Foundations of Stochastic Inventory Theory* (Stanford University Press, Stanford, CA).

Puterman ML (2014) *Markov Decision Processes: Discrete Stochastic Dynamic Programming* (John Wiley & Sons, Hoboken, NJ).

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

Shapiro A, Dentcheva D, Ruszczynski A (2021) *Lectures on Stochastic Programming: Modeling and Theory* (SIAM, Philadelphia).

ShipBob (2022) Shipping zones: Guide to shipping costs and transit times for ecommerce. Accessed June 6, 2024, https://www.shipbob.com/ecommerce-shipping/shipping-%20zones/.

Topaloglu H (2009) Using Lagrangian relaxation to compute capacity-dependent bid prices in network revenue management. *Oper. Res.* 57(3):637–649.

Wei L, Kapuscinski R, Jasin S (2021) Shipping consolidation across two warehouses with delivery deadline and expedited options for e-commerce and omni-channel retailers. *Manufacturing Service Oper. Management* 23(6):1634–1650.

Xie Y, Ma W, Xin L (2022) The benefits of delay to online decision-making. Preprint, submitted October 20, http://dx.doi.org/10.2139/ssrn.4248326.

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

Yan C, Reiffers-Masson A (2023) Certainty equivalence control-based heuristics in multi-stage convex stochastic optimization problems. Preprint, submitted August 25, https://doi.org/10.48550/arXiv.2308.13166.

Quan Zhou is a PhD student of operations management at the Desautels Faculty of Management at McGill University. His research focuses on optimization problems in supply chain management and revenue management and designing efficient algorithms for decision making with uncertainty.

Mehmet Gümüş is a professor of operations management and the Desautels Chair in Supply Chain Management and Business Analytics at McGill University. In his research, he focuses on supply chain management, dynamic pricing, and risk management. He is the founding academic director for the Masters of Management in Analytics Program at the Desautels Faculty of Management at McGill University.

Sentao Miao is an assistant professor of operations management at the University of Colorado Boulder. His research focuses on developing efficient learning and optimization algorithms, particularly in online learning, bandit problems, reinforcement learning, and approximation algorithms. He applies these methods to operations management problems, like dynamic pricing, assortment selection, and inventory control.