Fulfillment Optimization for Online Retailers under Stochastic Demands

Hedayat Alibeiki*
California State University, San Marcos, San Marcos, California, USA

Maryam Hassanlou
California State University, San Marcos, San Marcos, California, USA

Soheila Jorjani
California State University, San Marcos, San Marcos, California, USA

Online retailers have usually three options for fulfilling the customers' demands. They may fulfill an order by either the local fulfillment center (closest to the demand point), neighbor centers (any center other than the local), or directly from the suppliers or manufacturers (aka drop-shipping). Given a finite set of products with varying prices and costs, we consider the problem of determining the optimal fulfillment plan for an online retailer who sells a finite set of products to customers from different regions in a single period. The retailer fulfills orders using multiple fulfillment centers and/or by diverting the orders to the suppliers with the objective of maximizing the profit. Two Mixed Integer Linear Programing models are proposed for deterministic and stochastic demands. We applied the models to a numerical example and conducted sensitivity analyses to test the robustness of the models.

* Corresponding Author. E-mail address: halibeiki@csusm.edu

I. INTRODUCTION

Due to the rapidly increasing population of Internet users, online retailing has become very common and there are many retailers today selling their products only through online channels (e.g., Amazon, eBags, Blulight.com, Netgrocer.com, onlinefoodgrocery.com, etc.). According to U.S. Census Bureau (2019), in the year of 2018, e-commerce sales in the United States accounted for 9.63% of the total retail sales compared to 8.90% in 2017. The overall retail sales had a growth rate of 4.74% in 2018, while the sales of e-commerce retail grew at a significant rate of 13.38% accounting for 25% of the total growth of the sector. As a purely online retailer, Amazon's sales of 2018 accounted for 5% of all retail sales and 49.1% of all online retail sales throughout the country. Despite this impressive growth, online retailing is a very challenging business to run, for instance, Amazon's operating margin was below 4% in the entire period of 2011-2018 (Macrotrends. 2019), which is a significantly small margin compared to an average of 6% to 10% for brick-and-mortar stores (Rigby, 2014).

An online retailer functions very differently from a traditional retailer. One of the main differences is in the shopping experience and the delivery time. A consumer who orders online typically waits for the delivery of the goods, whereas in brick-and-mortar stores, they can immediately receive their purchased items. This difference makes the delivery more costly for the online retailer. Expediting the delivery through any mechanism results in higher delivery and/or inventory costs, which makes it necessary for the retailer to streamline its inventory and fulfillment plans.

Furthermore, there are many additional costly, logistics-related activities involved in online retailing compared to in-store shopping such as packaging, outbound shipping, return handling, among which outbound shipping is often quoted as the main source of fulfillment cost (Dinlersoz and Li 2006). These extra activities significantly increase the fulfillment costs of the online retailers, as they are not likely to be completely covered by the customers (Howland 2016). However, due to the high competition and consumers' increasing expectations, many retailers have no choice but to allow flexible and less expensive shipping options. These include unconditional free shipping (e.g., Nordstrom and Zappos), conditional free shipping (e.g., Amazon.com, Jet.com, Walmart), or free in-store pickup (e.g., Macy's and Walmart). The no-cost or fixed-cost shipping options imply that the retailer should incur the remainder of the shipping cost. Therefore, online retailers are highly incentivized to find the fastest and cheapest fulfillment plans for any order received.

Retailers' fulfillment decisions are tightly connected to the storage and distribution of their inventories and can have a significant impact on their overall operating costs. Online retailers often have multiple fulfillment centers (FCs). Amazon, for example, has currently 175 warehouses throughout the country (Amazon, 2019), and for fulfilling an order, they ship the product from the closest warehouse to the demand point (local fulfillment center), or in case of a stock-out, the inventory in a neighbor fulfillment center can be used to meet the demand. In addition to these two options, many online retailers (e.g., Dell) send the orders received from their customers to the manufacturers/suppliers so that the products are shipped directly from the manufacturer/supplier to the customer. This fulfillment option is also known as drop-shipping, and requires a high level of coordination among players of the supply chain (Newsgram, 1963). Indeed, the success of retailers such as Amazon in selling their products through online channels is mainly attributed to their efficient and low-cost supply chain (Nitin Chaturvedi and Ulker, 2013).

In this paper, we model and solve a fulfillment planning problem for an online retailer in a single period. The retailer offers a set of n products with given prices and unit costs. The objective of the retailer is to maximize the total profit that is the total sales revenue minus the total cost, which includes the unit cost of items plus the fulfillment and holding costs. We first consider a deterministic case where all products have fixed, but different demands. Then, the fixed demand assumption is relaxed by considering stochastic demands from known normal distributions. The means and standard deviations of demand vary for products. The retailer makes decisions about the type and quantity of products that are stored in each FC as well as the fulfillment assignment using one or a combination of the three options: local FC, neighbor FCs, or suppliers. We propose tractable Mixed Integer Linear Programming (MILP) models for the fulfillment optimization problem under deterministic and stochastic demand assumptions. The tractability of our models makes it possible for the retailers to solve this problem in large scales within reasonable time spans.

The remainder of this paper is organized as follows. Section 2 reviews the literature. Section 3 presents the notations and formulation of the fulfillment optimization problem under deterministic and stochastic demand assumptions for the online retailer. Section 4 presents a set of numerical analyses along with sensitivity analyses of the results. Section 5 concludes the research.

II. RELATED LITERATURE

The problem studied in this paper is closely related to two streams of literature: e-fulfillment and dynamic assortment planning. The literature in e-fulfillment can be divided up to three areas of distribution network design, inventory and capacity management, and delivery planning and execution. We specifically review the distribution network design and inventory and capacity management literature as they are more related to this research.

In the area of distribution network design when online channels are added, the designing of the e-fulfillment center and the level of automation at these centers are of strategic importance. Furthermore, whether these channels are operated as separate entities, or integrated and shared entities in fulfilling in-store or online demands are reviewed extensively in the literature. Bendloy et al. (2007) argue that inventory pooling increases the service level and/or decreases operating costs and suggest that only if as a percentage of the online demand (base level) is met it makes sense to introduce a dedicated warehouse. Liu et al. (2010) discuss the trade-offs between the risk of inventory pooling and the transportation costs. They showed that both demand variability and transport costs are important parameters in warehouse selection. Hubner et al. (2015) empirically investigate how and why retailers with multiple channels develop their logistics activities into Omni-channel systems in the retail industry. They conclude that Omni-channel retailers need to create new logistics models that reviews the trade-off between the process of integration and separation between different channels. Bretthauer et al. (2010) discuss where and how much inventory should be allocated and held at each site for a retailer that satisfies both in-store and online demand. Specifically, they recommend how many and which of a firm's capacitated locations should handle online sales to minimize total cost (holding, backorder, fixed operating, transportation, and handling costs).

Distribution network design modelling concentrates on how online orders are assigned within the supply chains and mainly propose a dynamic allocation policy that uses real time information to reevaluate e-fulfillment decisions in view of reducing costs.

In the stream of dynamic assortment planning, Lei Jasin and Sinha (2019) discuss the use of dynamic pricing and order fulfillment and develop two heuristic methods to find an optimal pricing policy. They first solve a deterministic approximation of the pricing policy and use the derived solution to solve a second heuristic that is adjusted by the realized demand. Rodriguez and Aydin (2015) develop a model that optimizes pricing and assortment decisions in a dual channel where the retailer carries a subset of the products that the manufacturer carries. Their pricing strategies is influenced by inventory related costs. Randall, Netessine and Rudi (2005) use data from 50 publicly traded Internet retailers to understand the role of inventory ownership and fulfillment capabilities. They propose the role of factors such as product variety, demand uncertainly, and the firm's age on the decision of de-coupling storefronts from inventory. Their model rely on drop-shipping to fulfill customer orders. They prove that the firm's survival is directly related to aligning its supply chain structure to its inventory ownership decisions.

Van Ryzin and Mahajan in their 1999 paper used a newsvendor framework where the demand process is based on the multinomial logit model of consumer choice. They developed a structure for the optimal assortment planning and provided insights on how the demand and price factors affect the optimal level of assortment variety. Smith and Agrawal (2000) model the consumer choice mechanism as a static probabilistic choice process and study the optimal assortment and stocking decisions under different substitution mechanisms. Mahajan and Van Ryzin (2001a, b) and Smith and Agrawal (2003) study the effect of dynamic substitution on assortment planning, inventory competition, assortment and inventory optimization for complementary products and category management with basket-shopping customers.

This paper contributes to the retail operations literature in multiple ways. First, while most of the papers in the literature of retail fulfillment focus on brick-and-mortar centers, this paper concentrates on pure online retailers. Second, many papers in the literature deal with solely one of the inventory, assortment, or fulfilment planning problems. This paper, however, focuses on fulfillment planning models with the capability to address the inventory management problem simultaneously. Finally, this paper enriches the literature by providing a more detailed framework that includes many realistic situations such as multiple fulfillment options, stock-out and substitution effects, limited storage capacity, and stochastic market demand.

III. PROBLEM DESCRIPTION AND ASSUMPTIONS

We seek the optimal fulfillment plan for a profit-maximizing online retailer that has demand for a set of finite products from different locations. Demand at each point can be fulfilled by a combination of three alternatives: fulfillment by the local FC (closest to the demand point potentially with less fulfillment costs), neighbor FCs (any center other than the local), or directly drop-shipped from the suppliers. The following graph illustrates the network of possible fulfillment scenarios for a simple case of only two fulfillment centers (FC-1 and FC-2) each covering a corresponding local demand region. For simplicity, we assumed that there is only one supplier for the products (S).

Total costs in the profit function includes purchasing costs, carrying costs of products at FCs, and shipping costs from the origins (FCs or suppliers) to the customers. Throughput this paper, we assume that the unit purchasing cost of a product would be cheaper if the retailer ordered it beforehand and stored it in his warehouse rather than fulfilling by direct shipments from the suppliers. i.e., $w_i^D \leq w_i^S$ where w_i^D denote the wholesale unit cost of product i if ordered and stored in FCs, while w_i^S would be unit cost if directly shipped from suppliers. This is consistent with the current practice and also with the Newsvendor problem with the reorder in which buying prices increase after the retailer observes the demand (Cachon and Netessine, 2006).

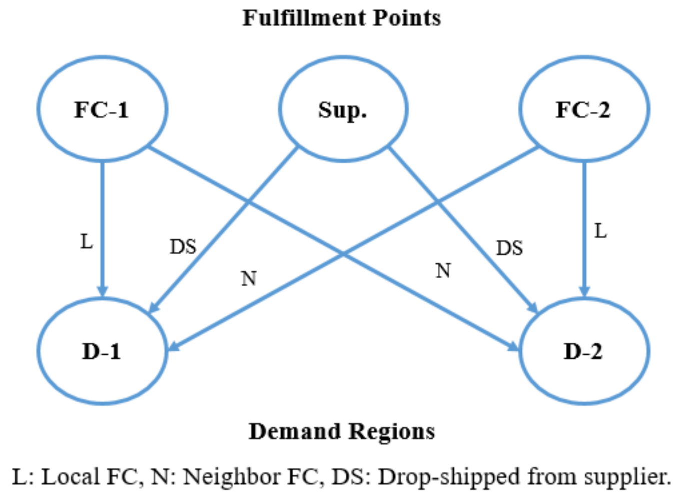

FIGURE 1. FULFILLMENT OPTIONS FOR A RETAILER WITH TWO FCS

We consider products to be substitutable, in the sense that if the demand for a product is not met by the retailer, a portion of its demand can be absorbed by other products that are provided. In our models, there might be two possible causes for substitution: 1) when a product is not offered by the retailer, neither carried out in FC nor drop-shipped from the supplier, 2) when the demand exceeds the total supply, the extra unmet demand will be substituted with other substitute products (if any). The second usually happens when demand is stochastic.

If a product is not profitable, retailer can choose not to offer it. Assuming products i and l are substitutable, θ_{li} denotes the percentage of demand for product l that in case of shortage will be substituted with product i. The total proportion of demand for product l that is substituted by other products would be $\Sigma_i \theta_{li}$, where $\Sigma_i \theta_{li} \leq 1$.

We also assume an assortment costs for products carried at FCs. This will make sure that the retailer does not carry infinitesimal amounts of a product in the storage and only store products that have reasonable demand. We let γ_{ik} denote the assortment cost that the retailer incurs if product i is stored at FC k.

Online retailer's main decisions are type and quantity of products stored in different FCs and the fulfillment plan. In addition to these, amount of sales, shortage, and surplus are other decision variables in the model. Following is the list of parameters and decision variables that are used in the formulation of the problem.

Parameters:

i= index of products (i= 1, 2 … I); I: total number of products

l = index of products with shortage that may substitute (l= 1, 2 … I); I: total number of products

j= index of demand locations (j= 1, 2 …J); J: total number of FCs

k= index of fulfillment locations (k=1, 2…J); J: total number of FCs

P_i = retail price of product i

w_i^D = wholesale unit cost of product i if ordered and stored at FCs

w_i^S = wholesale unit cost of product i if supplied by suppliers

C_{ijk} = shipping cost per unit of product i from FC k to the region covered by FC j
F_{ij}= shipping cost per unit of product i shipped directly from supplier to the region covered by FC j
γ_{ik}: assortment cost of product i at FC k
h_i: cost of excess capacity for unit of product i [holding cost of unused products]
d_{ij}= potential demand for product i in demand location j (STOCHASTIC)
θ_{li} = percentage of demand for product l that substitute to product i in case of shortage
t_i = shelf space needed per unit for product i (ft^3)
β_k= space capacity of FC k (ft^3)

Decision variables:
Z_{ijk} = amount of product i fulfilled from FC k for location j
Y_{ij} = amount of product i fulfilled from suppliers for location j
X_{ik} = amount of product i stored at FC k
Γ_{ik} = binary variable indicating if product i is stored in FC k
S_{ij} = Sale of product i at location j
H_{ij} = Slack (Unsatisfied demand for product i at location j)
O_{ij} = Overage (excess supply for product i at location j)

3.1 Deterministic Model

We first focus on a static fulfillment model in which the demand is assumed deterministic and certain. The results in this section are relevant to situations where demand is highly predictable, which can be the case for foods and necessary consumable items such as shampoos, toothpastes, napkins, etc.

$$(P1) = \text{Max } \{\sum_{i=1}^{I}\sum_{j}^{J} P_i S_{ij} - \sum_{i=1}^{I}\sum_{j}^{J} w_i^D X_{ij} - \sum_{i=1}^{I}\sum_{j}^{J} w_i^S Y_{ij} - \sum_{i=1}^{I}\sum_{j}^{J} F_{ij} Y_{ij} - \sum_{k=1}^{J}\sum_{i=1}^{I}\sum_{j=1}^{J} C_{ijk} Z_{ijk} - \sum_{j=1}^{J}\sum_{i=1}^{I} h_i O_{ij} - \sum_{k=1}^{J}\sum_{i=1}^{I} \gamma_{ik}\Gamma_{ik}\}$$

Subject to:

$$\sum_{k=1}^{J} Z_{ijk} + Y_{ij} - (d_{ij} + \sum_{\substack{l=1 \\ l\neq i}}^{I} H_{lj}\,\theta_{li}) = O_{ij} - H_{ij} \quad (1)$$

$$\sum_{k=1}^{J} Z_{ijk} + Y_{ij} - O_{ij} = S_{ij} \quad (2)$$

$$X_{ik} \leq \beta_k \;\Gamma_{ik} \quad (3)$$

$$\sum_{j=1}^{J} Z_{ijk} \leq X_{ik} \quad (4)$$

$$\sum_{i=1}^{I} X_{ik}\; t_i \leq \beta_k \quad (5)$$

$$H_{ij}, S_{ij}, X_{ij}, Z_{ijk}, Y_{ij} \geq 0$$

The objective is to maximize the total profit function that is defined as the total revenue from products minus total purchasing costs of products, total transportation costs, holding costs of unsold products, and assortment costs. Constraint (1) measures excess or shortage of each product at each demand location by subtracting total demand from total supply. Constraint (2) measures the total sale of each product at each demand location by subtracting the excess supply from the total amount of supply. Constraint (3) is defined to ensure that product i is considered "carried" in FC k (i.e., $\Gamma_{ik} = 1$) only if some positive amount of that product is stored at FC k (i.e., $X_{ik} > 0$). Constraint (4) ensures that the total amount of product i shipped from FC k does not exceed the stored amount of that product in FC k. Constraint (5) applies the storage capacity limitation.

3.2 Stochastic Model

In this section, we allow the products' demand to be stochastic. Having uncertain parameters in the right-hand-side vector of the formulation can affect the whole model as now the constraints will directly engage with probabilistic parameters while the objective function will indirectly be affected by corresponding changes on the value of decision variables. For the case of stochastic demand, we use Chance Constrained Programming as an appropriate method for linear programing problems with uncertain parameters.

Chance Constrained Programming (CCP) as the second type of stochastic programming, developed by Charnes and Cooper (1959), attempts to reconcile optimization over uncertain constraints. The constraints, which contain random variables, are guaranteed to be satisfied with a certain probability. In CCP, it is required that the objectives should be achieved with the stochastic constraints held at least α percent of time, where α is provided as an appropriate safety margin by the decision maker (aka service level).

Assume that x is a decision vector, ξ is a stochastic vector and $g_j(x, \xi)$ are stochastic constraint functions, j= 1, 2… p. Since the stochastic constraints $g_j(x, \xi) \leq 0$, j= 1, 2… p do not define a deterministic feasible set, they need to be hold with a confidence level α. Thus chance constraint is represented as follows (Liu, 2009):

$$\Pr\{g_j(x, \xi) \leq 0, j = 1, 2, \ldots, p\} \geq \alpha \quad (6)$$

Which is called a joint chance constraint, and when we want to consider them separately it is shown as follows:

$$\Pr\{g_j(x, \xi) \leq 0\} \geq \alpha_j, \quad j = 1, 2, \ldots, p \quad (7)$$

In general, for a linear-form function $g(x,\xi) = a_1x_1 + a_2x_2 + \cdots + a_nx_n - b$, if the coefficients $\zeta = (a_1, a_2, \ldots, a_n, b)$ are assumed to be independently-distributed normal random variables, then $\Pr\{g(x, \xi) \leq 0\} \geq \alpha$ if and only if

$$\sum_{i=1}^{n} E[a_i]x_i + \Phi^{-1}(\alpha)\sqrt{\sum_{i=1}^{n} Var[a_i]x_i^2 + V[b]} \leq E[b] \quad (8)$$

where Φ is the standard normal distribution function[1]. For this problem, we assume demand ($\widetilde{d_{lj}}$) is stochastic and normally distributed with the mean of $E[\widetilde{d_{lj}}]$ and variance of $Var(\widetilde{d_{lj}})$. Therefore, we can rewrite the first constraint as:

$$\sum_{k=1}^{J} Z_{ijk} + Y_{ij} - \sum_{\substack{l=1 \\ l\neq i}}^{I} \widetilde{H_{lj}}\,\theta_{li} + \phi^{-1}(1-\alpha)\sqrt{Var(\widetilde{d_{lj}})} \geq E[\widetilde{d_{lj}}] \quad (9)$$

where $\widetilde{H_{lj}}$ is the shortage of product l at demand location j, which itself is random depending on the realization of demand $\widetilde{d_{lj}}$. Setting the same α at the product level will make the retailer incur enormous amount of carrying and fulfillment costs. To fix this, we allow the retailer to set different values of α for different products. Therefore, if the retailer finds a product uncritical or unworthy of the carrying cost, he may set a low α for that product. On the other hand, for the high-demand and critical products, he may set a high value of α. This gives the retailer the flexibility to adjust its risk of shortage according to the profitability of a product.

[1] Refer to Liu (2009) for the proof.

This formulation of the stochastic model also allows taking the same service level α for a whole category of products if necessary.

It is important to note that with the assumption of random demand, the three variables of sales, excess supply, and shortage ($\widetilde{S_{ij}}$, $\widetilde{O_{ij}}$, $\widetilde{H_{ij}}$) become random variables, which depend on the realization of product demands. Our approach in this model would be to approximate these three random variables with their realized values when demand is equal to the average demand, i.e., $E[\widetilde{d_{ij}}]$. As a result, we use $S_{ij}^{\mu} = \widetilde{S_{ij}}\left(E[\widetilde{d_{ij}}]\right)$, $O_{ij}^{\mu} = \widetilde{O_{ij}}\left(E[\widetilde{d_{ij}}]\right)$, $H_{ij}^{\mu} = \widetilde{H_{ij}}\left(E[\widetilde{d_{ij}}]\right)$ as approximate values for those random variables and by doing so, we will end up with a deterministic-approximate model for when demand is assumed stochastic.

$$(P2) = \text{Max } \{\textstyle\sum_{i=1}^{I}\sum_{j}^{J} P_i S_{ij}^{\mu} - \sum_{i=1}^{I}\sum_{j}^{J} w_i^D X_{ij} - \sum_{i=1}^{I}\sum_{j}^{J} w_i^S Y_{ij} - \sum_{i=1}^{I}\sum_{j}^{J} F_{ij}Y_{ij} - \sum_{k=1}^{J}\sum_{i=1}^{I}\sum_{j=1}^{J} C_{ijk}Z_{ijk} - \sum_{j=1}^{J}\sum_{i=1}^{I} h_i O_{ij}^{\mu} - \sum_{k=1}^{J}\sum_{i=1}^{I} \gamma_{ik}\Gamma_{ik}\}$$

Subject to:

$$\sum_{k=1}^{J} Z_{ijk} + Y_{ij} - \sum_{\substack{l=1\\l\neq i}}^{I} H_{lj}^{\mu}\,\theta_{li} + \phi^{-1}(1-\alpha_i)\sqrt{Var(\widetilde{d_{ij}})} \geq E[\widetilde{d_{ij}}] \quad (10)$$

$$\sum_{k=1}^{J} Z_{ijk} + Y_{ij} - (\mu_{ij}^{d} + \sum_{\substack{l=1\\l\neq i}}^{I} H_{lj}^{\mu}\,\theta_{li}) = O_{ij}^{\mu} - H_{ij}^{\mu} \quad (11)$$

$$\sum_{k=1}^{J} Z_{ijk} + Y_{ij} - O_{ij}^{\mu} = S_{ij}^{\mu} \quad (12)$$

$$X_{ik} \leq \beta_k\ \Gamma_{ik} \quad (13)$$

$$\sum_{j=1}^{J} Z_{ijk} \leq X_{ik} \quad (14)$$

$$\sum_{i=1}^{I} X_{ik}\ t_i \leq \beta_k \quad (15)$$

$$H_{ij}, S_{ij}, X_{ij}, Z_{ijk}, Y_{ij} \geq 0$$

The objective is to maximize an approximated value for the total expected profit function that is defined as the total sales revenue at average demand minus total purchasing costs of products, transportation costs, holding costs of unsold products assuming average demand, and assortment costs. Constraint (10) defines the chanced constraint applying the required service level. Constraint (11) measures the excess or shortage of each product for when the demand is at the average level as a proxy for the expected excess and shortage. Constraint (12) measures the total sales revenue of each product if demand is equal to the mean values as a proxy for the total expected revenue. Constraint (13) is defined to ensure that the binary variable Γ_{ik} takes one only when product i is stored at FC k (i.e., $X_{ik} > 0$). Constraint (14) ensures that the total amount of product i shipped from FC k does not exceed the stored amount of that product in FC k. Constraint (15) does not allow extra storage above the normal capacity of each FC.

IV. NUMERICAL ANALYSIS

4.1 Sample category of products

To check the effectiveness and robustness of our models, we performed multiple numerical analyses. For illustration, we included a hypothetical small-scale fulfillment problem, but the results are general enough to be extended to larger problems. We consider an online retailer with three FCs (J=3) that offers a category of ten products (I=10). Table 1 summarizes for each product the price, unit costs, shipping costs per unit, unit volume, and holding costs per unit. Table 2 presents the demand (mean and variance) and the substitution rates of each product. Storage capacities in FCs 1 to 3 are respectively 805, 713, and 854.

TABLE 1. HYPOTHETICAL NUMERICAL EXAMPLE (PRICE, UNIT COSTS, SHIPMENT COSTS, UNIT VOLUME AND HOLDING COST)

i	P_i	W_D	W_S	c_{ij1}			c_{ij2}			c_{ij3}			F_{ij}			t_i	h_i
				j=1	j=2	j=3	j=1	j=2	j=3	j=1	j=2	j=3	j=1	j=2	j=3		
1	15	1	1.6	0.2	1.2	0.8	1.6	1	2.1	0.7	1.1	0.3	0.5	1.2	0.9	1.6	2.2
2	12	1	2.9	1.3	1.6	1.4	2	0.5	0.8	1.4	2.4	0.7	1.4	0.8	1	2.8	2.3
3	15	1.7	3	1.2	2	1.3	0.9	0.6	0.7	1.1	0.8	0.5	1.4	0.9	0.8	1.3	0.4
4	10	2.1	2.2	1.4	2.5	2.4	3.2	1.2	1.4	2.9	3.2	1.3	3.4	2.5	2.6	3.8	3.2
5	13	0.8	0.9	1.1	2.2	2.2	1.8	0.4	0.6	1.5	1.9	1.1	2.3	1.2	2	1.7	2.1
6	16	2	3	0.1	0.2	0.6	0.9	0.8	1.3	0.9	1.1	0.85	0.6	0.7	0.9	1.8	1.5
7	16	1	1.8	3.6	4.1	3.7	1.7	0.55	0.65	1.4	2.2	1.2	3.5	3.4	2.7	3.9	2.4
8	20	4	6	3.2	3.3	3.5	3.1	1.9	2.8	2.9	3.1	2	3.4	2.6	3	2.5	2.9
9	12	1.2	1.6	0.1	0.2	0.8	1.3	0.75	0.85	1.7	1.6	1.2	1.2	0.8	1.4	0.9	1.1
10	19	7	8	0.4	1	0.8	1.7	1.6	3	1.1	0.4	0.3	0.6	1.1	1.3	0.2	0.9

TABLE 2. HYPOTHETICAL NUMERICAL EXAMPLE (DEMAND AND SUBSTITUTION PARAMETERS)

i	$E(d_{ij})$			$Var(d_{ij})$			θ_{li}									
	j=1	j=2	j=3	j=1	j=2	j=3	l=1	l=2	l=3	l=4	l=5	l=6	l=7	l=8	l=9	l=10
1	105	111	52	856	404	70	0	40%	0	0	0	0	0	0	0	0
2	141	101	63	1699	439	500	40%	0	0	0	0	0	0	0	0	0
3	53	67	147	11	26	2542	0	0	0	30%	30%	0	0	0	0	0
4	131	100	114	2315	1480	2020	0	0	30%	0	30%	0	0	0	0	0
5	139	109	97	1059	1925	634	0	0	30%	30%	0	0	0	0	0	0
6	138	103	140	2583	1925	2915	0	0	0	0	0	0	10%	50%	0	0
7	130	68	145	381	52	2319	0	0	0	0	0	0	0	40%	0	0
8	102	65	90	1496	238	267	0	0	0	0	0	0	0	0	0	0
9	136	72	71	175	816	65	0	0	0	0	0	0	0	0	0	0
10	93	109	111	476	1728	205	0	0	0	0	0	0	0	0	0	0

4.2 Result: Deterministic Model

We first run the deterministic version of the model for the hypothetical example. The demand for each product is assumed to be equal to the mean demand provided at Table 2. The proposed Mixed Integer Programming model (P1) is solved using the MATLAB R2019a. For our hypothetical example, the maximum profit level is $33,061 under the deterministic model. The details of the optimal decision variables are reported in Tables 3 and 4.

TABLE 3. OPTIMAL INVENTORY AND FULFILLMENT FOR DETERMINISTIC MODEL

i	X_{ij}			Y_{ij}			Z_{ij1}			Z_{ij2}			Z_{ij3}		
	j=1	j=2	j=3	j=1	j=2	j=3	j=1	j=2	j=3	j=1	j=2	j=3	j=1	j=2	j=3
1	0	0	0	105	111	52	0	0	0	0	0	0	0	0	0
2	0	0	0	141	101	63	0	0	0	0	0	0	0	0	0
3	0	0	371	0	0	0	0	0	0	0	0	0	92	97	181
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	178	139	131	0	0	0	0	0	0	0	0	0
6	273	0	0	0	0	108	138	103	32	0	0	0	0	0	0
7	0	183	0	130	0	30	0	0	0	0	68	115	0	0	0
8	0	0	149	43	65	0	0	0	0	0	0	0	59	0	90
9	279	0	0	0	0	0	136	72	71	0	0	0	0	0	0
10	313	0	0	0	0	0	93	109	111	0	0	0	0	0	0

TABLE 4. OPTIMAL SALE, SHORTAGE, EXCESS, AND STORAGE DECISIONS FOR DETERMINISTIC MODEL

i	S_{ij}			H_{ij}			O_{ij}			Γ_{ik}		
	j=1	j=2	j=3	j=1	j=2	j=3	j=1	j=2	j=3	k=1	k=2	k=3
1	105	111	52	0	0	0	0	0	0	0	0	0
2	141	101	63	0	0	0	0	0	0	0	0	0
3	92	97	181	0	0	0	0	0	0	0	0	1
4	0	0	0	131	100	114	0	0	0	0	0	0
5	178	139	131	0	0	0	0	0	0	0	0	0
6	138	103	140	0	0	0	0	0	0	1	0	0
7	130	68	145	0	0	0	0	0	0	0	1	0
8	102	65	90	0	0	0	0	0	0	0	0	1
9	136	72	71	0	0	0	0	0	0	1	0	0
10	93	109	111	0	0	0	0	0	0	1	0	0

There are three noteworthy points in Tables 3 and 4:

1) For all products, the retailer stores them at most in one location and chooses to cross-ship from neighbor FCs. Obviously, as the size of the problem increases, the retailer will probably store each product in a greater number of FCs, but taking advantage of neighbor FCs is a viable option especially when the assortment cost is significant. In such cases, the retailer prefers to pay the extra shipping cost of Neighbor FCs to avoid the assortment and carrying costs in multiple FCs.
2) The retailer has chosen not to offer product 4 to the market (permanent stockout), neither through its FCs nor directly through suppliers. This is because by not offering it, only 40% of its demand will be lost as the remaining 60% will be substituted by products 3 and 5 (see θ_{li} in Table 2). Moreover, since the profit contribution of its

substitutes are significantly higher, the retailer will have enough incentive to eliminate it from the list of offered products.

3) The retailer has chosen not to carry products 1, 2, and 5 in its FCs, but it offers them through drop-shipping. This case happens when the supplier of a product does not charge significantly higher than the normal price for drop-shipping. In those cases, the retailer would avoid the high assortment cost by not carrying it locally at its FCs. Indeed, considering the profit margin of the product and affordable cost of shipment from suppliers, drop-shipping becomes the optimal alternative. This is especially true for products that have high demands but are not highly substitutable.

TABLE 5. OPTIMAL INVENTORY AND FULFILLMENT FOR STOCHASTIC MODEL

i	X_{ij}			Y_{ij}			Z_{ij1}			Z_{ij2}			Z_{ij3}		
	$j=1$	$j=2$	$j=3$	$j=1$	$j=2$	$j=3$	$j=1$	$j=2$	$j=3$	$j=1$	$j=2$	$j=3$	$j=1$	$j=2$	$j=3$
1	0	0	0	142	137	63	0	0	0	0	0	0	0	0	0
2	0	0	0	194	128	92	0	0	0	0	0	0	0	0	0
3	0	245	0	0	0	98	0	0	0	70	76	99	0	0	0
4	0	0	0	193	149	172	0	0	0	0	0	0	0	0	0
5	0	0	0	181	165	129	0	0	0	0	0	0	0	0	0
6	572	0	0	0	0	0	203	159	209	0	0	0	0	0	0
7	0	0	439	0	0	0	0	0	0	0	0	0	155	77	207
8	0	347	0	0	0	0	0	0	0	152	85	111	0	0	0
9	343	0	0	0	0	0	153	109	81	0	0	0	0	0	0
10	121	0	0	0	162	129	121	0	0	0	0	0	0	0	0

TABLE 6. OPTIMAL SALE, SHORTAGE, EXCESS, AND STORAGE DECISIONS FOR STOCHASTIC MODEL

i	S_{ij}			H_{ij}			O_{ij}			Γ_{ik}		
	$j=1$	$j=2$	$j=3$	$j=1$	$j=2$	$j=3$	$j=1$	$j=2$	$j=3$	$k=1$	$k=2$	$k=3$
1	105	111	52	0	0	0	37	26	11	0	0	0
2	141	101	63	0	0	0	53	27	29	0	0	0
3	53	67	147	0	0	0	4	7	65	0	1	0
4	131	100	114	0	0	0	62	49	58	0	0	0
5	139	109	97	0	0	0	42	56	32	0	0	0
6	138	103	140	0	0	0	65	56	69	1	0	0
7	130	68	145	0	0	0	25	9	62	0	0	1
8	102	65	90	0	0	0	50	20	21	0	1	0
9	136	72	71	0	0	0	17	37	10	1	0	0
10	93	109	111	0	0	0	28	53	18	1	0	0

4.3 Result: Stochastic Model

In this subsection, we consider stochastic demands for products. We assume product demands follow normal distribution with the means and variances provided in Table 2. To address the uncertainty embedded in our model, we use the Chance Constrained Programming method described in Subsection 3.2. The model (P2) is solved for the same hypothetical example described in Subsection 4.1 using the MILP solver of MATLAB R2019a. We consider the same service level coefficient of $\alpha_i = 0.9$ for all products, $\forall i \in \{1, \ldots, 10\}$. The maximum profit level would be \$23,608 under the stochastic model with a 90% service level for all the products. The optimal values of decision variables are provided in Tables 5 and 6.

There are three important points in Tables 5 and 6:

1) Similar to the deterministic model, the retailer stores all the items at most in one location and benefits from cross shipments from neighbor FCs, which reinforces the observation that the retailer makes the tradeoff between the extra shipping cost of Neighbor FCs and the assortment and carrying costs in multiple FCs.
2) Previously, in deterministic model, the retailer opted for a permanent stockout of product 4 by not offering it to the market whereas in this stochastic solution, the retailer supplies all products including product 4 through either FCs or supplier. Note that this is only because the solution in Table 5 and 6 is solved for service levels at the individual product levels (i.e., α_i). Permanent stockout of products should become possible in stochastic models if a composite service level α is defined at the category or store level.
3) The retailer has chosen not to carry in its FCs products 1, 2, 4, and 5 and instead, drop-ship them directly from the supplier. Note that due to the uncertain demand, the retailer needs to store more than the average demand if it chooses to carry a product. Given a limited capacity, this is only possible by decreasing the number of carried products.

4.4 Sensitivity Analysis

To assess the robustness of the stochastic model and to extract some insights from the results, we conduct sensitivity analyses on two of the key parameters: space capacity of FCs (β_k) and service level of products (α_i). In the first part of the sensitivity analysis, we change the storage capacity of FCs and trace its effect on the retailer's fulfillment decisions. We are specifically interested in percentage of the total fulfillment from the local and neighbor FCs as well as shipments from suppliers. For this sensitivity analysis, we assume FCs have equal storage capacities. We let the total storage capacity of the retailer ($\sum_k \beta_k$) change from 10% to 100% of the sum of average demands, $\sum_i \sum_j E(d_{ij})$. The total fulfillment for our hypothetical example is 4,199 units of products assuming the same service level for all products, $\alpha_i = 0.9, \ \forall i \in \{1, \ldots, 10\}$. Table 7 represents the percentage of total fulfillment from each fulfillment channel.

According to Table 7, as the storage capacity increases, the model uses the full storage for fulfillment through its FCs rather than sending directly from suppliers. As such, the percentage of orders fulfilled directly by suppliers decreases as the storage capacity increases. This would allow the retailer to take advantage of the reduced purchasing costs notwithstanding the holding and assortment costs. Between fulfillments from local and neighbor FCs, the share of local fulfillment increases in general as storage capacity increases. The key observation, however, is that after some specific level of storage capacity, increasing the FCs' capacity does not change the optimal fulfillment plan as the model does not use the entire storage capacity and instead, always meets a proportion of the total fulfillment from the suppliers.

TABLE 7. RESULTS SENSITIVITY TO FC'S CAPACITY

	$\sum_k \beta_k / \sum_i \sum_j E(d_{ij})$									
	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
$\sum_{i, j\neq k} Z_{ijk}$ (Neighbor)	2.1%	4.6%	7.1%	9.1%	12.0%	15.4%	23.5%	33.7%	31.7%	31.7%
$\sum_{i, j=k} Z_{ijk}$ (Local)	1.1%	2.2%	4.4%	8.6%	13.2%	16.9%	21.6%	25.7%	31.2%	31.1%
$\sum_{i,j} Y_{ij}$ (suppliers)	96.9%	93.2%	88.5%	82.3%	74.8%	67.7%	54.9%	40.6%	37.1%	37.1%

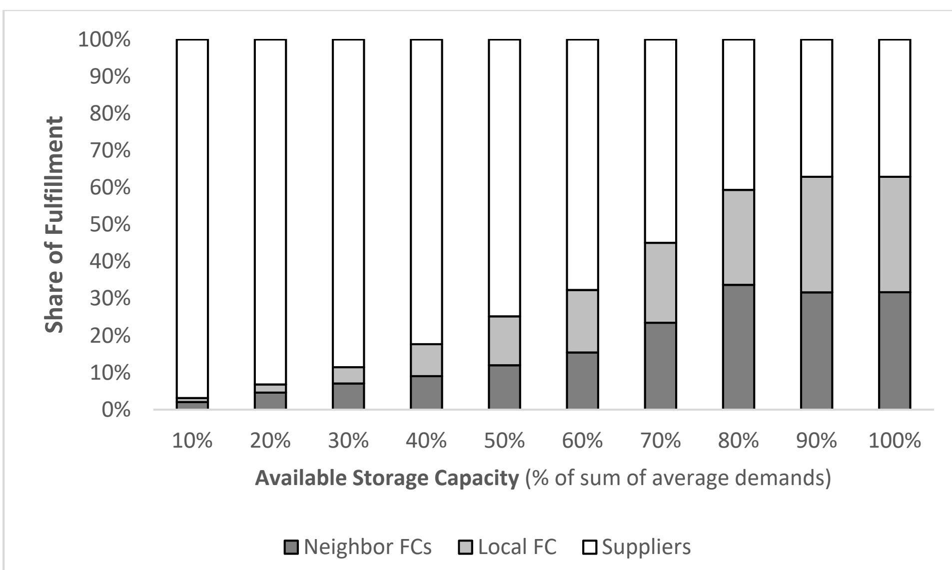

FIGURE 2. PERCENTAGE OF FULFILLMENT OPTIONS FOR DIFFERENT STORAGE CAPACITIES

In the second part of this analysis, the sensitivity of the results of the stochastic model to changes in the service level (α_i) is analyzed. We use the same hypothetical example described in Subsection 4.1 and use the same service level for all products, i.e., $\alpha_i = \alpha;\ \ \forall i \in \{1, \ldots, 10\}$. We let the value of α change from 0.8 to 0.99 with increments of 0.05. Tables 8 presents the estimated value of profit function, total amount of products fulfilled from suppliers, and the estimated excess of products (unsold) in the end of period for five different values of α. As can be seen from this table, as service level α increases, the overall estimated profit of the retailer decreases. This is because as α increases, the retailer must carry more products which ultimately leads to more holding and shipping costs as well as more expected number of unsold products. These factors together lead to decrease in the profit. Also, assuming a constant storage capacity, as the service level increases, direct shipment from suppliers takes a larger share among different fulfillment options.

TABLE 8. RESULTS' SENSITIVITY TO DIFFERENT SERVICE LEVEL (A)

α	**0.8**	**0.85**	**0.9**	**0.95**	**0.99**
Total Excess $\sum_{j=1}^{J}\sum_{i=1}^{I} O_{ij}^{\mu}$	720	887	1,097	1,408	1,991
Shipped from suppliers ($\sum_{j=1}^{J}\sum_{i=1}^{I} Y_{ij}$) % out of total expected demand	1,614 (44%)	1,943 (48%)	2,133 (51%)	2,396 (55%)	3,025 (63%)
Profit ($)	26,578	24,813	23,608	22,267	18,003

V. CONCLUSION

The problem of fulfillment optimization is a prominent issue for online retailers. In this paper, we proposed Mixed Integer Linear Programming models for the fulfillment optimization problem faced by an online retailer. We formulated the models for deterministic and stochastic demand assumptions and solved them using a small-scale fulfillment problem. When the demand of products is deterministic, the solution of the optimization problem is certainly optimized. However, under stochastic demand, our proposed model provides an estimate for the optimal solution.

The tractability of the proposed MILP models makes it possible for retailers to find a near optimal solution for their fulfillment problem. MATLAB R2019a was used for solving these models. Robustness of our models was tested using sensitivity analysis with respect to two important parameters, total capacity of fulfillment centers and the service level. Our analysis shows that in some cases it is beneficial for the retailer not to store products at their local stores but instead fulfill a portion of demand using direct shipments from suppliers. This is especially true for the products where the difference between the purchasing cost before and after realizing the demand is negligible.

In our models, product prices were considered as given parameters. Including pricing to the list of endogenous decision variables can be an extension of this research. We developed single-stage fulfillment decision models. Future extensions could also consider multiple stages of decision-making.

REFERENCES

Agatz N., Moritz Fleischmann and Jo van Nunen, "E-Fulfillment and Multi-Channel Distribution – A Review Niels", *European Journal of Operational Research*, 187 (2), 2008, 339-356.

Amazon, "Amazon Fulfillment Centers", 2019, https://bit.ly/2sRN2MI (accessed December 20, 2019).

Bendoly E., Blocher, D., Bretthauer, K.M. and Venkataramanan, M.A., "Service and cost benefits through clicks-and-mortar integration: implications for the centralization/decentralization debate", *European Journal of Operational Research*, 180 (1), 2007, 426-442.

Bretthauer K.M., Mahar, S. and Venakataramanan, M.A., "Inventory and distribution strategies for retail/e-tail organizations", *Computers and Industrial Engineering*, 58 (1), 2010, 119-132.

Cachon, G. P., & Netessine, S., "Game theory in supply chain analysis. In Models, methods, and applications for innovative decision making", *INFORMS*, 2006, 200-233.

Charnes, A. and Cooper, W.W., "Chance-constrained Programming", *Management Science*, 6, 1959, 73–79.

Dinlersoz EM and Li H, "The shipping strategies of Internet retailers: Evidence from Internet book retailing", *Quant. Marketing Economy,* 4(4), 2006. 407–438.

Howland D., "Online operating costs crushing brick-and-mortar retailers", *Retail Dive* (May 4, 2016), http://bit.ly/23pFWoF.

Hübner A., Holzapfel, A. and Kuhn, H., "Operations management in multi-channel retailing: an exploratory study", *Operations Management Research*, 8 (3), 2015, 84-100.

Kök, A. G., M. Fisher, and R. Vaidyanathan, *Assortment planning: Review of literature and industry practice. In Retail Supply Chain Management* (2nd ed.) N. Agrawal and S. Smith, Springer, 2014.

Liu, B., *Theory and Practice of Uncertain Programming* (3rd ed), UTLAB, 2009.

Lei Y., J. Stefanus, A. Sinha, "Joint Dynamic Pricing and Order Fulfillment for E-Commerce Retailers", *Manufacturing and Service Operations Management,* 2019, 269-284.

Liu, B., Zhang, R. Xiao, M., "Joint decision on production and pricing for online dual Channel Supply Chain system", *Applied Mathematical Modeling*, 34, 2010, 4208-4218.

Macrotrends, "Amazon Operating Margin 2006-2019", 2019, https://bit.ly/34BoULK (accessed December 20, 2019).

Mahajan, S., G. van Ryzin, "Stocking retail assortments under dynamic consumer substitution", *Operations Research*. 49(3), 2001a, 334-351.

Mahajan, S., G. van Ryzin, "Inventory competition under dynamic consumer choice", *Operations Research*, 49(5), 2001b, 646-657.

Newsgram, A., "USA AMGRAM Standing Orders: Drop Shipping To Other Than Direct Distributor". *Amway Corporation*, 6 (1), 1963.

Nitin Chaturvedi, Mirko Martich, B. R. and Ulker, N., "The future of retail grocery online". *McKinsey & Company report*, 2013.

Randall T., S. Netessine, N. Rudi, "An empirical examination of the decision to invest in fulfillment capabilities: A study of Internet retailers", *Management Science*, 52(4), 2005, 567–580.

Rigby, D., *Online shopping isn't as profitable as you think*, 2014, http://bit.ly/1RR7XPy.

Rodríguez, B. and Aydın, G., "Pricing and assortment decisions for a manufacturer selling through dual channels", *European Journal of Operational Research*, 242(3), 2015, 901-909.

Rusmevichientong P., H. Topaloglu, "Robust assortment optimization in revenue management under the multinomial logit choice model", *Operations Research*, 60(4), 2012, 865-882.

Smith, S.A., N. Agrawal, "multi-item retail inventory systems with demand substitution", *Operations Research*, 48, 2000, 50-64.

U.S. Census Bureau, *Latest quarterly e-commerce report*, 2019, http://bit.ly/1cDTnLk.

Van Ryzin G., S. Mahajan, "On the relationship between inventory costs and variety benefits in retails assortments", *Management Science*, Online Publication, Nov. 1999.