10-16-2013

Online retailers' promotional pricing, free-shipping threshold, and inventory decisions : a simulation-based analysis

Rafael BECERRIL-ARREOLA
University of South Carolina

Mingming LENG
Department of Computing and Decision Sciences, Lingnan University

Mahmut PARLAR
McMaster University

Online Retailers' Promotional Pricing, Free-Shipping Threshold, and Inventory Decisions: A Simulation-Based Analysis[1]

Rafael Becerril-Arreola[2],   Mingming Leng[3],   Mahmut Parlar[4]

We consider a two-stage decision problem, in which an online retailer first makes optimal decisions on his profit margin and free-shipping threshold, and then determines his inventory level. We start by developing the retailer's expected profit function. Then, we use publicly-available statistics to find the best-fitting distribution for consumers' purchase amounts and the best-fitting function for conversion rate (i.e., probability that an arriving visitor places an online order with the retailer). We show that: (i) a reduction of the profit margin does not significantly affect the standard deviation of consumers' order sizes (purchase amounts) but increases the average order size; whereas, (ii) variations in a positive finite free-shipping threshold affect both the average value and the standard deviation of the order sizes. We then use Arena to simulate the online retailing system and OptQuest to find the retailer's optimal decisions and maximum profit. Next, we perform a sensitivity analysis to examine the impact of the ratio of the unit holding and salvage cost to the unit shipping cost on the retailer's optimal decisions. We also draw some important managerial insights.

Key words: Profit margin, contingent free-shipping, inventory, simulation with Arena, optimization with OptQuest, sensitivity analysis.

---

1   Introduction

Online retailing is an important distribution channel and it is continuing to expand quickly. As information technology (IT) advances rapidly, consumers can access the Internet more conveniently, safely and cheaply, and they are thus more interested in online shopping. To successfully compete, online retailers frequently implement promotions. Promotions are of special importance to business-to-consumer (B2C) firms because search costs are lower in online markets than in traditional ones, and the competition in online markets is thus very stiff. Moreover, Advertising.com's survey [1] indicated that around 56% of consumers consider price discounts as the most enticing promotion offered by online retailers. Price promotions are thus a critical marketing strategy for online retailers.

For online retailing operations, shipping costs are also an important determinant of a consumer's purchase decision. As Dinlersoz and Li [12] discussed, shipping costs are of capital importance to online markets, even though they are irrelevant to consumers' purchase decisions in offline distribution channels for most product categories. To reduce the negative impact of the shipping fees, online retailers have been implementing a variety of shipping-related promotions, which can be divided into the following three categories: (i) *unconditional* free shipping (UFS) policy, under which an online retailer absorbs the shipping costs for all orders; (ii) *contingent* free shipping (CFS) policy, under which a retailer pays for the shipping costs but *only* for those orders equal to or larger than a predefined cutoff level; and (iii) consumers absorb the shipping costs by paying the fees that are increasing in their order sizes. In [26] Lewis reported that the CFS policy is the most effective promotion in increasing the revenues of online retailers. However, a critical question for the CFS policy remains: What CFS cutoff level should be set by the online retailer?

As the above discussion illustrates, pricing and CFS strategies are two most effective tactics that influence consumers' online purchases. In order to satisfy consumers' demand during the promotion period, the retailer also needs to make a decision on how many products should be stocked for the sales. We learn from Anderson, Fitzsimons, and Simester [3] and Breugelmans et al. [10] that the inventory decision for a promotion period is also of great significance to the success of an online retailer.

We conclude from our above discussion that decisions on (i) pricing, (ii) CFS threshold and, (iii) inventory play important roles in the quickly-growing online retailing industry. We note from Advertising.com's survey [1] and Yao [31] that, in practice, many online retailers have been implementing a two-stage decision process. In the first stage of this process, they first determine their pricing and CFS thresholds and announce their decisions to the market before a promotion takes place. Then, at the second stage, they make the stocking decision to satisfy consumers' online orders during the promotion period. For example, Amazon.com announced its 2007 Christmas pricing decisions for most toy gifts and its free-shipping promotions for all orders over $25 several months before the Christmas promotion period.

In this paper we develop a two-stage sequential decision model for an online retailer who first makes optimal decisions on the retail price and the CFS cutoff level, and then makes the inventory decision. We assume that the online retailer sells multiple products in a single category; this is a common case in practice. For example, Barnesandnoble.com mainly sells books and CDs online, and Ashford.com mainly sells watches, sunglasses and fragrances. From Anderson, Fitzsimons, and Simester [4] and Blattberg and Neslin [8] who have shown that in practice the *profit margins* of different products in the same category are *identical*, we learn that making the pricing decision is equivalent to determining a single profit margin for all products in a single category. Our goal is to find the optimal profit margin, CFS threshold and inventory that maximize the retailer's expected profit during a promotion period.

We note that the online retailer's expected profit function for the two-stage decision problem is too complicated to be optimized analytically. The complexity of the problem stems

from the fact that it involves optimizing three decisions that jointly determine profits both at the individual and aggregate level. Therefore, we use simulation to find the optimal pricing, CFS and inventory decisions for the retailer. In order to obtain realistic results based on real data, we simulate the online retailing system using the empirical statistics reported by Lewis, Singh, and Fay [27]. These include percentages of small, medium and large orders and conversion rates (i.e., probability that an arriving visitor places an online order and thus becomes a customer of the retailer) for different free-shipping and pricing policies. For our simulation, we first find the decision-dependent distributions that best fit the purchase distributions given in [27]. We then determine the best-fitting random purchase amount distribution and conversion rate function, and use these specifications to simulate the operations of the online retailer and to compute optimal two-stage decisions.

We use Arena, a primary simulation software, to efficiently simulate the retailing system, and then use OptQuest, an optimization add-on for Arena, to search for the optimal decisions that maximize the retailer's profit. Arena and OptQuest have been widely used in practice to simulate a variety of real systems; but, very few academic publications in the business and economics fields used this simulation technique. Academic publications concerning simulation with Arena and optimization with OptQuest include Aras et al. [5] and Askin and Chen [6].

Additionally, we perform several sensitivity analyses to investigate the impact of the ratio of the inventory cost parameter to the unit shipping cost on the retailer's optimal decisions and maximum profit. The sensitivity analyses are conducted because: (i) these parameters were estimated for the particular firm in the empirical study [27], so they may not be general; and (ii) from the sensitivity analysis we are able to draw some important managerial insights.

The remainder of this paper is organized as follows. In Section 2 we review major relevant publications to show the originality of our paper. In Section 3 we develop our two-stage decision model for an online retailer. In Section 4 we find the best-fitting distribution function for the purchase amount and the best-fitting conversion rate function, and use these best-fitting functions to simulate the retailing system. In Section 5 we find the optimal decisions and conduct a sensitivity analysis. This paper ends with a summary of our contributions in Section 6.

2   Literature Review

We classify the relevant literature into five categories: Category 1 includes papers concerning the effects of shipping costs and promotions on purchasing behavior of consumers in eBusiness; Category 2 includes papers on the joint pricing and CFS decisions; Category 3 focuses on the joint pricing and inventory/purchasing decisions; Category 4 is concerned with the inventory decisions of online retailers; and Category 5 is about the optimal inventory policies in the presence of order quantity requirements.

We start with Category 1 in which, as a representative publication, Lewis et al. [27] performed an empirical study of the impact of nonlinear shipping and handling fees on consumers'

purchasing behaviors. Using a database from an online retailer, the authors showed that consumers are more sensitive to shipping charges than to prices. In Category 2, representative publications include Leng and Becerril-Arreola [24] and Gümüş et al. [15]. In [24] Leng and Becerril-Arreola examined the impact of an online retailer's joint pricing and CFS free-shipping decisions on the purchase behavior of consumers, and computed the optimal price and CFS cutoff level for the retailer. Gümüş et al. [15] discussed two price partitioning formats—i.e., (i) a price is partitioned into a product price and a separate shipping and handling surcharge and (ii) free shipping service is offered to customers with a non-partitioned price which includes, in part or in whole, the shipping cost. The above two papers did not consider the inventory decision but only the pricing and CFS threshold decisions.

There are several publications in Category 3. For example, van der Heijden [30] investigated a two-echelon network consisting of a central depot and multiple local warehouses, and developed a periodic-review inventory model to make the optimal decisions on the initial inventory allocation to the local warehouses, the maximum stock in the central depot, and the shipping frequency. Jang and Kim [20] analyzed a single-period inventory problem to find a supplier's joint optimal decisions on the production quantity, the inventory allocation (among multiple customers), and the shipping sequence of his products for the customers. Jang, Kim, and Park [21] investigated the joint inventory allocation and shipping frequency decisions for a manufacturer with a limited capacity, who serves multiple types of retailers with a single product. Hua, Wang, and Cheng [17] obtained the price and lead time decisions for a dual-channel supply chain where a manufacturer sells its products through both a physical retail channel and an online direct channel. In [18] and [19], Hua, Wang, and Cheng determined an online retailer's optimal order lot size and retail price under a given CFS policy in the long-run and the newsvendor settings, respectively. Although Category 3 contains some relevant publications, we still find that very few papers in this category considered the CFS decisions for online retailers. Leng and Parlar [25] investigated a free-shipping decision problem in B2B transactions, where an online seller announces his cutoff level decision and a buyer chooses her purchase amount. The authors developed a leader-follower game and find the Stackelberg equilibrium for the game.

The recently published and relevant publications in Category 4 are reviewed as follows. Ayanso, Diaby, and Nair [7] developed a Monte Carlo simulation model to examine a continuous-review inventory problem for an online retailer selling multiple non-perishable and made-to-stock items. Jing and Lewis [22], who used empirical data from an online grocery store to discuss the impact of stockouts and fulfillment rates on the purchase behaviors of consumers in different segments. This paper did not consider pricing and CFS decisions for online retailers. For a comprehensive review of other relevant publications in Category 4, see Agatz, Fleischmann, and van Nunen [2].

As a representative publication in Category 5, Zhao and Katehakis [32] characterized inventory policies that minimize ordering, holding, and backorder costs when the order size is either zero or at least a minimum number. For an inventory system similar to that in [32], Bradford

and Katehakis [9] addressed an inventory allocation problem, and Zhou, Zhao, and Katehakis [34] proposed and evaluated a simple heuristic ordering policy. Zhou, Katehakis, and Zhao [33] examined a retailer's CFS policy, under which the shipping cost is a positive constant if the order quantity is smaller than a threshold and is zero otherwise. The authors derived optimal inventory control policies for both a single-period problem and a multi-period problem. The above publications considered order size constraints and optimal inventory policies; but, they did not involve pricing or CFS decisions.

The above brief review of representative literature reveals that our paper differs from extant publications in that we simultaneously consider the optimal pricing, CFS threshold, and inventory decisions for an online retailer, rather than analyzing only one or two of these three decisions as in the publications reviewed above. Since the discussion in Section 1 **indicated** that the joint decisions on pricing, CFS threshold, and inventory are important to online retailers, our research problem is worth investigating, and our results should contribute to the literature in the operations management area and help practitioners make their optimal decisions.

3 Two-Stage Decision Model

We now consider a two-stage decision model for an online retailer, who maximizes his expected profit during a promotion period by first determining its profit margins (prices) and CFS thresholds and then making his stocking decision. Note that in practice, the products promoted by the retailer in a promotional period such as Christmas and Valentines' day are usually seasonal or perishable products, e.g., Christmas trees and roses. Accordingly, we assume seasonable or perishable products in this paper, and consider a single-period decision problem for the retailer. The retailer serves consumers with multiple products of a single seasonal or perishable category. Since the profit margins of different products in the same category are often identical (see Anderson, Fitzsimons, and Simester [4] and Blattberg and Neslin [8]), we use the concept of profit margin to make the pricing decision for the online retailer.

Assumption 1 The retailer sets the price of each product as $p = (1 + m)c$, where m and c denote the profit margin and the unit purchase cost of the product, respectively. ◁

For recent applications of the above assumption to marketing-operations problems, see Cachon and Kok [11], Dong et al. [13], etc. The retailer offers a contingent free-shipping (CFS) policy to his customers by setting the CFS cutoff level as u dollars; for example, Amazon.com's CFS cutoff level is currently $u = \$25$. We let Y denote a consumer's random purchase amount (in dollars). If a consumer's purchase amount $Y = y$ is equal to or greater than u, then the consumer qualifies for free shipping and the retailer pays for the shipping cost. Otherwise, if $y < u$, then the consumer absorbs the shipping cost, even when the sum of the purchase amount $\$y$ and the shipping cost exceeds the threshold u. This occurs because, in practice, whether or not a consumer qualifies for free shipping only depends on the comparison between

his or her purchase amount y and the cutoff level u. The shipping cost is only an additional cost to the consumer when $y < u$.

3.1 Description of the Decision Problem

We learn from Lewis, Singh, and Fay [27] that the profit margin (pricing) and free-shipping decisions greatly impact consumers' order amounts (i.e., dollar purchase amounts) and the conversion rate, which is defined as the ratio of the number of consumers who place online orders to the number of visitors. Note that the number of visitors includes the consumers who place online orders and the persons who visit the retailer's website but do not buy any product online. Thus, given a number of visitors, we can calculate the number of online purchases as the conversion rate times the number of visitors.

Assumption 2 The number of online purchases is expressed as $N(m, u) = R(m, u)B$, where $R(m, u)$ and B denote the conversion rate and the number of visitors. ◁

Next, we consider a consumer's purchase amount Y, which also depends on the profit margin m and the CFS cutoff level u. Since in practice consumers are likely to spend different purchase amounts, we make the following assumption.

Assumption 3 A consumer's purchase amount Y is a random variable with the probability density function (p.d.f.) $f_{m,u}(y) \equiv f(y \mid m, u)$, the cumulative distribution function (c.d.f.) $F_{m,u}(y) \equiv F(y \mid m, u)$, and the expected value $\mu_{m,u} \equiv E(Y \mid m, u)$. We can then calculate the cumulative (aggregate) dollar sales for the promotion period as $Z = \sum_{i=1}^{N(m,u)} Y_i$, which is a random variable with a distribution defined by the p.d.f. $g_{m,u}(z) \equiv g(z \mid m, u)$, the c.d.f. $G_{m,u}(z) \equiv G(z \mid m, u)$, and the expected value $\mu_{m,u}^Z \equiv E(Z \mid m, u) = N(m, u)\mu_{m,u}$. ◁

Note that in the above assumption, Y_i, $i = 1, \ldots, N(m, u)$, are i.i.d. random variables that denote the stochastic purchase amounts of consumers $i = 1, \ldots, N(m, u)$. Note that Y_i, $i = 1, \ldots, N(m, u)$, are i.i.d. and $N(m, u)$ is a *stopping time* for the sequence of the purchase amounts of all consumers arriving during the single period (i.e., promotion period). We can thus apply Wald's Equation (see, e.g., Ross [28, p. 105]) to compute the retailer's expected total profit as given in Assumption 3.

In addition to determining the optimal values of the profit margin m and the CFS cutoff level u, we should compute the retailer's optimal stocking decision. We determine the optimal dollar value of inventory (denoted by v) rather than product quantity in stock; for extant applications of the dollar value of inventory, see, e.g., Ettl et al. [14] and Thompson [29]. Hence, a natural question arises: how does one determine the stocking level (in quantity) for each product from the optimal dollar value of inventory? One possibility is to allocate proportionally the optimal dollar value among all products, and then determine the stocking level of each product for the promotion period. Thus, we assume, without loss of generality, that the online retailer sells n products each with purchase cost c_i, $i = 1, 2, \ldots, n$. Given the stocking level q_i for product $i = 1, 2, \ldots, n$, the dollar value of product i is $c_i q_i$ and the total

dollar value of inventory is $\sum_{i=1}^{n} c_i q_i$. When the optimal dollar value of all products in stock is found as v^*, we need to determine the optimal stocking levels q_i^* such that $\sum_{i=1}^{n} c_i q_i^* = v^*$. Here, v^* may be regarded as the optimal budget for the online retailer, who needs to decide how to allocate v^* among n products. As an easy-to-implement and practical approach, the retailer can use *past* promotions' sale data to compute the ratio of dollar value of each product's sales to total dollar value of all sold products. Accordingly, we make the following assumption regarding the inventory allocation.

Assumption 4 Denote the allocation ratio for product i by w_i, $i = 1, 2, \ldots, n$ such that $\sum_{i=1}^{n} w_i = 1$. Since the optimal dollar value (budget) is $v^* = \sum_{i=1}^{n} c_i q_i^*$, we can compute the allocation of v^* to product $i = 1, 2, \ldots, n$ as $w_i v^*$, and then find the "optimal" stocking level of product i as $q_i^* = w_i v^*/c_i$. $\lhd$

In addition to the method suggested by the above assumption, we may consider the optimal allocation solution that solves the following bounded knapsack problem: $\max_{q_i, i=1,2,\ldots,n} \sum_{i=1}^{n} m c_i q_i$, subject to $\sum_{i=1}^{n} c_i q_i = v^*$ and $q_i \in \{0, 1, 2, \ldots, (1+\eta_i)\bar{q}_i\}$, for $i = 1, 2, \ldots, n$. Note that $\bar{q}_i$ denotes the average sales of product i—which can be estimated from historical data—and η_i represents a percentage increment—which can be obtained by using historical data to estimate the mean value $\bar{q}_i$ and the standard deviation σ_{q_i} and then calculating the safety stock $\eta_i \hat{q}_i$ according to the retailer's service level. Note that our subsequent analysis is independent of the allocation solution, and therefore we do not need to provide the optimal solution for the above knapsack problem (which can be actually solved by using dynamic programing).

Next, we describe the two-stage decision problem of an online retailer who *sequentially* makes optimal decisions to maximize his expected profit for a promotion period. In particular, at the first stage, the retailer does not consider the inventory-related costs, and only determines the optimal profit margin m^* and CFS cutoff level u^* that maximize his expected profit $\pi_1(m, u)$, which is computed as expected sales revenue minus the sum of expected purchase cost and expected shipping cost. At the second stage, given the optimal profit margin m^* and CFS cutoff level u^*, the retailer determines the optimal stocking decision v^* that minimizes his expected inventory-related cost (sum of the expected holding and salvage cost and the expected shortage cost), which is denoted by $\pi_2(v \mid m^*, u^*)$. Hence, the retailer's maximum system-wide expected profit is found as

$$\pi(m^*, u^*, v^*) = \pi_1(m^*, u^*) - \pi_2(v^* \mid m^*, u^*).$$

To solve the two-stage problem, we use the following approach. We first minimize the second-stage function $\pi_2(v \mid m, u)$, for fixed values of m and u, to find the stocking decision $v^*(m, u)$ (as a function of m and u). We then maximize the expected profit $\pi(m, u, v^*(m, u)) = \pi_1(m, u) - \pi_2(v^*(m, u) \mid m, u)$ to obtain the retailer's optimal profit margin m^* and CFS cutoff level u^*. Then, substituting m^* and u^* into $v^*(m, u)$ gives the optimal stocking decision $v^* = v^*(m^*, u^*)$. Next, we develop the functions $\pi_1(m, u)$ and $\pi_2(v \mid m, u)$.

3.2 Profit Function

When a consumer places an online order with purchase amount $Y = y$, the retailer's sale revenue is y. Under the CFS policy, the retailer incurs acquisition costs and may absorb the shipping costs. We now compute the retailer's acquisition cost of products that the consumer buys by spending y. This acquisition cost can be computed as $c = y/(1+m)$, since the profit margin is m and the total sales price is $y = (1+m)c$. Moreover, the profit that the retailer gains from the purchase amount y is computed as $y - y/(1+m) = my/(1+m)$.

As Lewis, Singh, and Fay [27] discussed, the shipping-related cost includes the shipping fee and other fees (e.g., insurance fee, taxes) associated with shipping the orders. Thus, the shipping-related cost for a consumer depends on the dollar value of the products that the consumer buys. For example, as estimated by the manager of the firm in the empirical study in [27], the shipping-related cost is $6.5 for a small order (with order size smaller than or equal to $50), $7.5 for a medium order (with order size greater than $50 but smaller than or equal to $75), and $10 for a large order (with order size greater than $75). In our paper, we denote the shipping-related cost by $S(y)$ and assume that it is increasing in y. Note that who pays for the shipping-related cost depends on the comparison of y and u. Specifically, if $y \geq u$, then the retailer absorbs $S(y)$; otherwise, the consumer pays for the shipping-related cost.

As shown in [27], a high free-shipping cutoff level may trigger some customers to increase their purchase amounts to qualify for FS, and it may thus lead to a larger average purchase amount. On the other hand, a low cutoff level allows more customers to qualify for FS and convinces fewer customers to increase their purchase amounts. Each customer's reaction to a given FS cutoff level u can therefore be modeled as a probabilistic process dependent on u. For our model, we let $\rho(y, u)$ be the probability that, given the FS cutoff level u, a consumer increases her purchase amount from y to $u + X$, where X is a *positive* random variable with the mean value $E(X) = \beta$. We consider the term X because, given that a consumer increases her purchase amount to qualify for free shipping, the consumer needs to purchase more products and spend $(u - y)$ or more.

Since a consumer may be more willing to increase the purchase amount when the difference between u and y is smaller, the function $\rho(y, u)$ is assumed to be decreasing in u but increasing in y; that is, $\partial \rho(y, u)/\partial u < 0$ and $\partial \rho(y, u)/\partial y > 0$. If $y > u$, a consumer with an (initial) amount y qualifies for free shipping and does not need to increase the order size. Hence, we assume that $\rho(y, u) = 0$ for $y > u$. In addition, when $y = u$, the consumer can obtain the free shipping service with the initial purchase amount y (equal to u); thus, for this case, $\rho(y, u) = \rho(u, u) = 1$. Note that, when u is a significantly large number, no consumer would increase the purchase amount to obtain this costly free-shipping service. Accordingly, we can assume that $\rho(y, u) = 0$ when u approaches infinity.

We now construct the retailer's expected profit (excluding the inventory-related cost) function $\pi_1(m, u)$ for the first stage. If $y \geq u$, then the retailer's profit is calculated as $my/(1+m) - S(y)$. If $y < u$, then the consumer may keep the original purchase amount y with the probability $1 - \rho(y, u)$. In this case, the retailer does not need to pay the shipping

cost, and his profit is $my/(1+m)$. Alternatively, the consumer increases her purchase to $\$(u+X)$—where X is a r.v. with the mean value $E(X) = \beta$—with the probability $\rho(y,u)$ and qualifies for free shipping. As a result, the retailer receives the expected sale revenue $m(u+\beta)/(1+m)$ and pays the expected shipping cost $S(u+\beta)$. Thus, in this case, the retailer's expected profit is $m(u+\beta)/(1+m) - S(u+\beta)$. The retailer's expected profit function $\pi_1(m,u)$ is then

$$
\begin{aligned}
\pi_1(m,u) \;=\; N(m,u) & \left\{ \int_u^\infty \left[\frac{my}{1+m} - S(y) \right] f_{m,u}(y) \; dy + \int_0^u [1 - \rho(y,u)] \frac{my}{1+m} f_{m,u}(y) dy \right. \\
& \left. + \int_0^u \rho(y,u) \left[\frac{m(u+\beta)}{1+m} - S(u+\beta) \right] f_{m,u}(y) \; dy \right\}.
\end{aligned}
$$

Next, we compute the inventory-related cost for the second stage at which the retailer determines the value of v (i.e., the dollar value of inventory), given the values of m and u. When the aggregate (cumulative) sales of the promotion period are $Z = z$, we find that, if $z \le v$, then the products worth $\$(v-z)$ remain unsold at the end of the period. Note that after the promotional period, the retailer salvages the unsold seasonal or perishable products (by, for example, largely reducing its retail price) and absorbs a disposal cost in addition to a cost of holding the unsold products for the period. Therefore, for the unsold products worth $\$(v-z)$, the retailer incurs the holding and salvage cost $I(v-z)$, where I (% per dollar) denotes the cost of holding and salvaging products worth $1 for a promotion period; that is, such a cost is proportional to the dollar value of inventory. This proportionality method has been widely used in the inventory management. As Hadley and Whitin [16, p. 422] discussed, in practice I ranges from 15% to 35% with the typical value being 20%.

If $z > v$, then products worth $\$(z-v)$ are out of stock during the the promotion period, and the retailer incurs a shortage cost. For the single-period problem, we can simply use the lost profit to estimate the shortage cost. We do not consider the goodwill cost for the online retailer, because of the following fact: As the empirical study in Jing and Lewis [22] indicates, the stock-out of a retailer's products is very likely to result from customers' strong preference for the retailer, and thus have no significant long-term impact on the sales. That is, the stock-out for the current period (e.g., Christmas season in 2011) does not significantly affect the sales during the next period (e.g., Christmas season in 2012). Moreover, since the products in the single category are likely to be substitutable—which is accordingly assumed in this paper—the stockout of a product may lead consumers to buy another product in the category. Because of the substitutability of the products, the retailer focuses on the aggregate demand for all products. That is, when shortage occurs, all products rather than one or a few products are sold out, which means that we do not need to consider the short-term impact of stock-out. According to the above, the retailer's shortage cost is calculated as the lost profit $\$(z-v)$ if $z > v$ and zero otherwise.

Since the retailer gains the profit $\$m/(1+m)$ from selling the products at price $1, the lost profit of the retailer is $\$m(z-v)/(1+m)$ when products worth $\$(z-v)$ are out of stock. Thus,

given m and u, the retailer chooses the optimal dollar value v^* that minimizes the following expected inventory-related cost function:

$$\pi_2(v \mid m, u) = I \int_0^v (v - z)g_{m,u}(z) \, dz + \frac{m}{1 + m} \int_v^\infty (z - v)g_{m,u}(z) \, dz. \tag{1}$$

Theorem 1 Given the values of m and u, for the second stage, the retailer's optimal dollar value $v^*(m, u)$ can be *uniquely* obtained by solving the following equation for $v(m, u)$:

$$G_{m,u}(v(m, u)) = \frac{m}{I(1 + m) + m}. \tag{2}$$

Proof. The first- and second-order derivatives of $\pi_2(v \mid m, u)$ in (1) w.r.t. v are computed as $d\pi_2(v \mid m, u)/dv = m/(1 + m) - [I + m/(1 + m)] \int_0^v g_{m,u}(z) \, dz$ and $d^2\pi_2(v \mid m, u)/dv^2 = [I + m/(1 + m)]g_{m,u}(v) > 0$, which implies that $\pi_2(v \mid m, u)$ in (1) is strictly concave in v. Solving $d\pi_2(v \mid m, u)/dv = 0$ for v gives the result as in this theorem. ∎

Substituting $v^*(m, u)$ in (2) into (1) gives,

$$\pi_2(v^*(m, u) \mid m, u) = \frac{m}{1 + m}\mu_{m,u}^Z - \left(I + \frac{m}{1 + m}\right) \int_0^{v^*(m,u)} z g_{m,u}(z) \, dz.$$

Then, the retailer's two-stage expected profit function $\pi(m, u, v^*(m, u))$—which is computed as $\pi_1(m, u) - \pi_2(v^*(m, u) \mid m, u)$, as discussed in Section 3.1—is found as,

$$
\begin{aligned}
\pi(m, u, v^*(m, u)) &= N(m, u)\left\{\int_0^u \rho(y, u)\left[\frac{m(u + \beta - y)}{1 + m} - S(u + \beta)\right] f_{m,u}(y) \, dy \right. \\
&\quad \left. - \int_u^\infty S(y)f_{m,u}(y) \, dy\right\} + \left(I + \frac{m}{1 + m}\right) \int_0^{v^*(m,u)} z g_{m,u}(z) \, dz. \tag{3}
\end{aligned}
$$

The retailer should maximize $\pi(m, u, v^*(m, u))$ to find the optimal profit margin m^* and the optimal CFS cutoff level u^*. Since $\pi(m, u, v^*(m, u))$ is a very complex function, except in some very special cases, it would be impossible to find m^* and u^* analytically. In fact, even if the p.d.f. $f_{m,u}(y)$ were given, it would still be intractable to find the p.d.f. $g_{m,u}(z)$ of the random variable $Z = \sum_{i=1}^{N(m,u)} Y_i$ (except for the special case where Y_i are normal, or exponential). These observations lead us to consider a numerical study for our two-stage decision problem combined with simulation to determine the optimal solution. To assure the uniqueness of the optimal solution in our numerical study, we need to show the convergence of the above two-stage procedure. However, because of the intractable complexity of the function in (3), we cannot prove the convergence analytically but will later perform numerical experiments to demonstrate that our two-stage procedure converges to a unique optimal solution.

4   Empirical Study to Estimate the Model Components

In [27] Lewis, Singh, and Fay describe a distribution of consumers' purchase amounts, i.e., the percentage of large orders (each with an order size larger than or equal to $75), the percentage of medium orders (order size between $50 and $75) and the percentage of small orders (order size smaller than $50) along with average purchase amounts and conversion rates. For a fixed profit margin of $m = 25\%$, they provide the conversion rates and the order-size distributions corresponding to the following three free-shipping policies, as summarized in Table B in online Appendix B.

1. Base policy ($m = 25\%, u = \infty$): No free shipping service is offered to any order. That is, for this policy, the retailer sets the CFS cutoff level at infinity, i.e., $u = \infty$.

2. Free shipping to all orders ($m = 25\%, u = 0$): Free shipping service is offered to all orders; thus, for this policy, $u = 0$.

3. Free shipping only to large orders ($m = 25\%, u = 75$): Free shipping service is offered to an order only when the size of the order is greater than or equal to $75, that is, $u = 75$.

When free-shipping service is not offered, i.e., $u = \infty$, Lewis, Singh and Fay [27] provide the conversion rates and the order-size distributions for the following two pricing (profit margin) policies as summarized in online Table C.

1. Base policy ($m = 25\%, u = \infty$). This is of course the same base policy as above.

2. 10% price discount ($m = 12.5\%, u = \infty$): Here, the price of each product is reduced by 10% which results in $m = 12.5\%$.

In the remainder of this section we use the data in Tables B and C in online Appendix B to estimate the components of the optimization model (3), by adopting the following three-stage procedure.

Stage 1: Best-fitting distributions of purchase amount for all policies above. For each policy, we estimate the probability distribution of a consumer's random purchase amount that best fits the limited empirical data provided in Lewis, Singh, and Fay [27] and summarized in online Tables B and C. Then, observing the best-fitting distributions for all policies, we draw conclusions regarding the impact of m and u on the distribution of the purchase amount.

Stage 2: Purchase amount random variable for an arbitrary policy. We use our conclusions in Stage 1 to construct (i) each consumer's stochastic initial purchase amount function Y and (ii) the probability $\rho(y, u)$ that a consumer with the initial purchase amount $0 < y < u$ increases y to u or $u + X$ in order to qualify for free shipping. Note that X is a random variable [with the p.d.f. $\phi(x)$] denoting the amount that the consumer spends in excess of u when she qualifies for free shipping.

Next, we use empirical data in [27]—which is summarized in Tables B and C in online Appendix B—to estimate the parameter values of Y, $\rho(y, u)$, and the p.d.f. $\phi(x)$. Then, we develop each consumer's final purchase amount function $\bar{Y}(m, u)$ for any arbitrary policy.

Stage 3: Estimated parameter values of $R(m, u)$. We build a logit model of the conver-

sion rate $R(m, u)$, and use empirical data in [27] to estimate the parameters in $R(m, u)$.

4.1 Best-Fitting Probability Distributions of Purchase Amount for All Policies

For given values of m and u corresponding to each policy in online Tables B and C we can find the estimated p.d.f. $\hat{f}_{m,u}(y)$ that best fits the order-size data while assuring that the expected value of the purchase amount Y equals the empirical average order size. We do this by using the percentage values in the tables and solving a nonlinear programming problem where the decision variables are the parameters of a tested density function. We proceed as follows:

Step 1: We consider five commonly-used distributions (normal, lognormal, Erlang, gamma, Weibull) and for each distribution solve the following minimization problem:

$$\min \ \Pi = \left[\int_0^{50} \hat{f}_{m,u}(y) \ dy - P_S \right]^2 + \left[\int_{50}^{75} \hat{f}_{m,u}(y) \ dy - P_M \right]^2, \text{ s.t. } \mu_{m,u}(y) = A, \quad (4)$$

where P_S, P_M, and A denote % of small orders, % of medium orders, and the average order size, and are all given in online Tables B and C. Here, the decision variables are the parameters of the fitting distribution. Note that we do not consider the squared deviation for the percentage of *large* orders because this percentage is equal to one minus the sum of percentages of small and medium orders.

Step 2: We then compare the minimum squared deviations obtained in Step 1 for all distributions, and find the best-fitting distribution as the one with the smallest value of Π.

To illustrate our approach in Step 1, we provide an example in which we consider the Weibull distribution and solve the minimization problem (4) for the base policy (i.e., $m = 25\%$ and $u = \infty$). For the base policy the purchase amount random variable is denoted by Y_0.

Example 1 For the base policy in online Tables B and C, we consider the Weibull distribution with the probability density function,

$$\hat{f}_0(y_0) \equiv f_{0.25,\infty}(y_0) = \begin{cases} \kappa_1 \kappa_2^{-\kappa_1} y_0^{\kappa_1 - 1} \exp[-(y_0/\kappa_2)^{\kappa_1}], & \text{for } y_0 > 0, \\ 0, & \text{otherwise,} \end{cases} \quad (5)$$

where $\kappa_1 > 0$ and $\kappa_2 > 0$. We solve the minimization problem (4) to find the optimal values of the parameters κ_1 and κ_2 as $\kappa_1^* = 1.243$ and $\kappa_2^* = 62.844$. The minimum squared deviation for the Weibull distribution is thus obtained as $\Pi^* = 0.0000595$, which shows that the best Weibull distribution with κ_1^* and κ_2^* fits the empirical data well. ◁

As in Example 1, we find the optimal values of parameters and the minimum squared deviations for the other common distributions. Table 1 shows that Weibull distribution is the one that best fits the empirical data for the base policy because it has the smallest minimum squared deviation.

Distributions	Weibull	Normal	Lognormal	Erlang	Gamma
Min. Sq. Dev.	0.0000595	0.0146	0.00517	0.00381	0.0873

Table 1: Comparison among all common distributions for the base policy ($m = 25\%$ and $u = \infty$).

Remark 1 Using the above two-step procedure, we also find the best-fitting distributions for other policies, i.e., "free shipping to all orders" ($u = 0$), "free shipping to large orders" ($u = 75$) and "10% discount" ($m = 12.5\%$) and observe the following:

1. For all policies, the Weibull distribution is always the one that best fits the empirical data.

2. For all policies, we use the best-fitting Weibull distributions to compute the mean values and standard deviations of the random purchase amount Y. The results are given in Table 2, where we find that the price discount (or, the reduction of profit margin) can increase the average purchase amount but doesn't significantly impact the standard deviation.

3. For the base policy and the policy of free shipping to all orders, the mean value changes but the standard deviation doesn't significantly change. However, when the free-shipping cutoff level is $75, we find that both mean and standard deviation change significantly. This occurs mainly because, as shown by Lewis, Singh and Fay [27], when the free shipping threshold is nonzero and finite (e.g., $75), a consumer with an (initial) order size smaller than the cutoff level $75 may increase his or her order size to $75 or more, in order to qualify for free shipping. ◁

	Mean	Standard Deviation
Base policy ($m = 25\%$, $u = \infty$)	58.61	5.82
Free shipping to all orders ($m = 25\%$, $u = 0$)	60.91	5.81
Free shipping to large orders ($m = 25\%$, $u = 75$)	63.13	6.98
10% discount ($m = 12.5\%$, $u = \infty$)	59.51	5.75

Table 2: Mean values and standard deviations of the random purchase amount for all policies.

4.2 Purchase Amount Random Variable for an Arbitrary Policy

It is important to note that the density-estimation procedure we have described above gives results for only the four policies for which Lewis, Singh and Fay [27] have provided empirical data. Naturally, in order to optimize the objective function $\pi(m, u, v^*(m, u))$ of the two-stage problem in (3), we need a more general procedure to estimate the p.d.f. of the purchase amount random variable for any combination of (m, u).

We learn from Remark 1 that, for the policy of "free shipping to large orders" ($m = 25\%$, $u = 75$), each consumer may increase his or her initial purchase amount to quality for free shipping, whereas for the other three policies, the free shipping policy is independent of each consumer's initial purchase amount, and thus all consumers spend their initial purchase amount. As a result, the mean and the standard deviation for the policy with $m = 25\%$ and

$u = 75$ significantly differ from those for other policies. Next, we first temporarily ignore the policy of "free shipping to large orders," and construct a stochastic function Y to describe each consumer's initial purchase amount. Then, building a function $\rho(y, u)$ to model the probability that a consumer increases his or her initial purchase amount y to u or above, we develop the consumer's random final purchase amount $\bar{Y}(m, u)$, which can apply to all policies in Table 2.

4.2.1 Initial Purchase Amount Random Variable

For now, temporarily ignoring the policy with $m = 25\%$ and $u = 75$, we note from Remark 1 that the standard deviation of the random variable Y when $u = 0$ is almost the same as that when $u = \infty$, and the reduction of profit margin (from the base value of 25%) does not significantly affect the standard deviation. We thus construct each consumer's stochastic initial purchase amount function in *additive* form as,

$$Y = K(m, u) + Y_0, \tag{6}$$

where $K(m, u)$ is a deterministic function dependent of the decision variables m and u; Y_0 represents the random purchase amount for the base policy ($m = 25\%$, $u = \infty$). As noted in Example 1, we find that Y_0 satisfies the Weibull distribution with the p.d.f. (5) in which the values of parameters are $\kappa_1 = 1.243$ and $\kappa_2 = 62.844$.

Remark 2 We now determine the $K(m, u)$ function considering the following three conditions:

1. For the base policy, we must have $K(25\%, \infty) = 0$, and thus, $Y = Y_0$.
2. The figures in Table 2 suggest that the policy of "free shipping to all orders" and the base policy have different average order sizes but almost identical standard deviations. Thus, for the policy of "free shipping to all orders", we will have $K(25\%, 0) = 60.91 - 58.61 = 2.3$.
3. Table 2 indicates that the policy of "10% discount" and the base policy have different average order sizes but almost identical standard deviations. Thus, for the "10% discount" policy, we have $K(12.5\%, \infty) = 59.51 - 58.61 = 0.9$. ◁

To satisfy the above three conditions, we set

$$K(m, u) \equiv \alpha_1 (0.25 - m) + \alpha_2 \exp(-\gamma u) \tag{7}$$

where $\alpha_1 > 0$ is the weight of the profit margin m, which measures the sensitivity of the purchase amount Y to m; $\alpha_2 > 0$ and $\gamma > 0$ are two parameters for the free-shipping cutoff u, which measure the sensitivity of the initial purchase amount Y to u. For details regarding the function $K(m, u)$ in (7), see online Appendix A.1.

4.2.2 Qualifying for Free Shipping

As we discussed previously, a consumer with an order size smaller than u may increase her order size to a level equal to or larger than u, in order to qualify for free shipping. In order to reflect this, we consider the following two possibilities:

Initial Purchase Amount is Increased to Qualify for Free Shipping When u is finite, a consumer with the initial purchase amount of $y < u$ may increase her purchase amount with the probability $\rho(y, u)$. More specifically, if $u - y$ is small, then the consumer would increase y to at least u with a high probability. Otherwise, the consumer increases y to u or more with a small probability. Thus, we write $\rho(y, u)$ as,

$$\rho(y, u) = \begin{cases} e^{-r(u-y)}, & \text{if } y \leq u, \\ 0, & \text{otherwise,} \end{cases} \tag{8}$$

where r denotes the sensitivity of the probability to the difference between u and y.

The function $\rho(y, u)$ measures the probability that a consumer will increase her initial purchase amount y to $u+x$, where the p.d.f. $\phi(x)$ of X is right-skewed because most consumers spends just enough to qualify for free shipping. We thus specify $\phi(x)$ as follows:

$$\phi(x) = \begin{cases} \dfrac{1}{\beta} e^{-x/\beta}, & \text{for } x > 0, \\ 0, & \text{otherwise,} \end{cases} \tag{9}$$

which has mean $E(X) = \beta$ and variance $\sigma^2 = \beta^2$.

Initial Purchase Amount is not Increased For finite u, a consumer with $y < u$ will not increase her purchase amount to u or more with probability $1 - e^{-r(u-y)}$, and thus will not qualify for free shipping.

4.2.3 Estimates of Parameter Values

Because of the possibility that a consumer (with an order size $y < u$) may increase y to $u+x$, we write the consumer's random final purchase amount $\bar{Y}(m, u)$ as,

$$\bar{Y}(m, u) = \begin{cases} u + X, & \text{with probability } \rho(y, u); \\ Y = K(m, u) + Y_0, & \text{with probability } 1 - \rho(y, u). \end{cases} \tag{10}$$

where $K(m, u)$ is given in (7). To estimate the parameter values in (10), we perform the following two steps. In the first step, we use Remark 2 to calculate the parameters α_1 and α_2 in $K(m, u)$ as $\alpha_1 = 7.2$ and $\alpha_2 = 2.3$. In the second step, we consider the policy of "free shipping to large orders" ($m = 25\%$, $u = \$75$) for the estimation of other parameter values. Specifically, using the policy's empirical data in online Table B, we solve a minimization problem similar to that in (4), and find that $\gamma = 10.55$ in $K(m, u)$, $r = 0.07$ in $\rho(y, u)$ [which

is given in (8)], and $\beta = 26.64$ in the p.d.f. $\phi(x)$ of X [which is given in (9)]. For details regarding this procedure, see online Appendix A.2.

Remark 3 As indicated in Section 4.2.1, the probability that Y assumes negative values can be shown to be negligibly small. We use (10) to find that $E(Y) > 7.2(0.25 - m) + E(Y_0)$ since $2.3 \exp(-10.55u) > 0$, and the standard deviation of Y is the same as that of Y_0. Recall from Table 2 that $E(Y_0) = 58.61$ and $\sqrt{\mathrm{Var}(Y)} = \sqrt{\mathrm{Var}(Y_0)} = 5.82$. Even if we assume that $m = 200\%$, we find that $E(Y) > 46.01$, which is at least $7.9 \ (= 46.01/5.82)$ times the standard deviation of 5.82. Hence, with a standard deviation that is significantly smaller than its expected value, Y is unlikely to assume negative values. ◁

Now we can calculate the total sales as $Z = \sum_{i=1}^{N(m,u)} \bar{Y}_i(m, u)$, and find the expected value of Z as,

$$E(Z) = E[\bar{Y}_i(m, u)]N(m, u) = E[\bar{Y}_i(m, u)]R(m, u)B, \tag{11}$$

where we compute $E[\bar{Y}_i(m, u)]$ using (18) in online Appendix A.2; $R(m, u)$ is the conversion rate, which is analyzed below; and B is the customer base size (for our numerical study, we set $B = 10,000$, as in the empirical study by Lewis, Singh and Fay [27].)

4.3 Conversion Rate

We define the conversion rate as the probability that an arriving visitor places an order with the online retailer and thus becomes a customer. The empirical data in Tables B and C (in online Appendix B) suggest the following (which we later use to construct our conversion rate model).

1. Table B shows that, as u increases, the conversion rate decreases.
2. Table C shows that, as m increases, the conversion rate decreases.
3. Table B indicates that, when $u \to \infty$, the conversion rate is 18.37%. We also note that, as $m \to \infty$, no consumer buys any product and the conversion rate should thus be zero.
4. Table B indicates that, when $u = 0$, the conversion rate is 21.65%.

Next, we construct a conversion rate function that reflects the above observations. To do this, we divide all arriving visitors into two groups: "buy" (i.e., consumers who place online orders) and "no buy" (i.e., visitors who leave without any purchase). Thus, the conversion rate function can be constructed as a binary choice (response) model. Jing and Lewis [22] used a logistic specification to describe online consumers' purchase incidence. Similarly, we model the conversion rate as the following logistic function:

$$R(m, u) = \frac{\exp[\theta_1 m + (\omega + \theta_2 u)^{-1} + \tau]}{\exp[\theta_1 m + (\omega + \theta_2 u)^{-1} + \tau] + 1}, \tag{12}$$

where $\theta_1 < 0$ is the parameter of profit margin m and reflects the sensitivity of the conversion rate to the profit margin m. The parameters $\omega > 0$ and $\theta_2 > 0$ are related to the CFS cutoff level u and reflect the sensitivity of the conversion rate to the free-shipping cutoff level u. The

constant parameter τ can be positive or negative. It is easy to see that our logistic model (12) satisfies the first three conditions above.

We include the parameter $\omega > 0$ in (12) because without it the model $R(m, u)$ in (12) reduces to $R(m, u) = \exp[\theta_1 m + (\theta_2 u)^{-1} + \tau]/\{\exp[\theta_1 m + (\theta_2 u)^{-1} + \tau] + 1\} = 1/\{1 + \exp[-\theta_1 m - (\theta_2 u)^{-1} - \tau]\}$. But in this case when $u = 0$, we have $R(m, 0) = 1$, which implies that, if $u = 0$, then all arriving visitors will place online orders. This is obviously contrary to the fourth condition above.

Using the empirical data in Tables B and C in online Appendix B, we obtain the parameter values in the conversion rate model $R(m, u)$ as $\theta_1 = -1.17$, $\theta_2 = 0.13$, $\omega = 4.85$, and $\tau = 1$, as shown in online Appendix A.3.

5   Optimal Decisions

We now use the random purchase amount function $\bar{Y}(m, u)$ in Section 4.2.3 and the conversion rate $R(m, u)$ in Section 4.3 to find the optimal values of the profit margin m, the free-shipping threshold u, and the inventory value v for the online retailer. To solve the two-stage model in Section 3.2, we need to first solve (2) for v^*, given the values of m and u. This calculation requires the p.d.f. $g_{m,u}(z)$ of the random variable $Z = \sum_{i=1}^{N(m,u)} \bar{Y}_i(m, u)$, where $\bar{Y}_i(m, u)$ is the random final purchase amount as computed in (10). Unfortunately, it is quite difficult to obtain the p.d.f. $g_{m,u}(z)$ analytically. For this reason, we follow an alternate route and use the *simulation* software Arena and its optimization add-on OptQuest to search for the optimal v^*. We begin by listing three main issues that we will address in this section.

1. We first use (11) to compute the *exact* values of the expected total purchase amount $E(Z)$ for a number of pairs of m and u. The exact values of $E(Z)$ that are computed without simulation will be later compared with our corresponding simulated results to show the accuracy of our simulation.

2. We use Arena with OptQuest to search for the optimal solution (m^*, u^*, v^*) that maximizes the retailer's expected profit $\pi(m, u, v^*(m, u))$ in (3).

3. We perform sensitivity analysis to investigate the impact of the ratio of the inventory cost parameter I to the unit shipping costs for three order categories (i.e., small, medium, and large orders) on the retailer's optimal decisions and maximum expected profit.

To test the accuracy of the simulation, we consider the value of $E(Z)$ only. We do so because, for any numerical study without simulation, it would be impossible to evaluate the integrals in (3) that involve the density $g_{m,u}(z)$, even numerically. Thus, we cannot numerically compute the retailer's two-stage expected profit $\pi(m, u, v^*(m, u))$, given the values of m and u. This is a major reason why we use simulation to search for the optimal solutions.

5.1   Computation of the Expected Total Purchase Amount $E(Z)$

We now use (11) to compute the expected total purchase amount $E(Z)$, given the values of m and u. For example, we consider the policy of "free shipping to large orders" ($m = 25\%, u =$

75), and evaluate (11) to find that $E(Z) = \$122,575$. Next, we perform a sensitivity analysis to examine the impacts of m and u on $E(Z)$, as shown by Figure 1. The computational results are presented in Table A in online Appendix A.4.
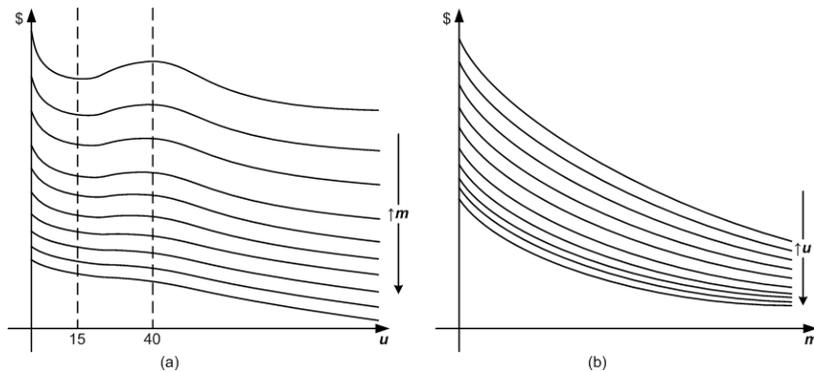

Figure 1: The impacts of m and u on $E(Z)$. Graph (a) indicates the impacts of u for values of m that range from 12.5% (corresponding to the top curve) to 125% (corresponding to the bottom curve) in increments of 12.5%. Graph (b) indicates the impacts of m for values of u that range from 0 (corresponding to the top curve) to 135 (corresponding to the bottom curve) in increments of 15.

From Figure 1(a), we find that, given the value of m, the expected total purchase amount is the highest when the retailer provides the free-shipping service to all online orders, i.e., $u = 0$. When the value of u is increased to a low nonzero level (e.g., a level smaller than \$15), we find that the expected total purchase amount decreases. When the retailer raises the CFS threshold to a level greater than \$15 but smaller than \$40, we find that the expected purchase amount increases. However, if the retailer increases the value of u to a high level (e.g., a level greater than \$40), then many visitors may leave without any purchase, and the retailer would thus have fewer sales even though some consumers still increase their purchase amounts.

We find from Figure 1(b) that, when the value of u is fixed, the expected total purchase amount $E(Z)$ is decreasing in the profit margin m. Thus, we can draw the conclusion that, if the retailer hopes to increase the total sales (purchase amount), then the retailer should reduce the profit margin to a small but acceptable level, and set the CFS threshold to a proper level. However, if the retailer pursues a high profit margin, then the retailer may need to provide the free-shipping service to all orders.

5.2 Simulation with Arena to Search for the Optimal Solution (m^*, u^*, v^*)

We now use simulation to compute the optimal solution. We plot a flowchart in Figure A (which is given in online Appendix A.5) to depict the online retailing process in our simulation model. We use Arena—a primary simulation software used in industrial applications—to build our simulation model. Arena provides a variety of modules for simulation. For more information regarding how to use Arena for simulation, see, e.g., Kelton et al. [23].

$E(Z)$ ($\times 100$) [%]		m									
		0.125	0.25	0.375	0.5	0.625	0.75	0.875	1	1.125	1.25
u	0	1500 [0.16]	1320 [0.11]	1160 [0.31]	1010 [−0.05]	881 [0.09]	766 [0.21]	667 [0.75]	579 [1.22]	501 [1.61]	432 [1.84]
	15	1450 [0.05]	1280 [0.17]	1120 [−0.22]	984 [0.15]	858 [0.08]	746 [0.01]	648 [0.11]	561 [0.12]	484 [−0.01]	417 [−0.06]
	30	1460 [0.14]	1290 [0.46]	1130 [0.38]	985 [0.15]	857 [0.05]	744 [0.01]	645 [0.10]	557 [0.02]	480 [−0.04]	412 [−0.31]
	45	1450 [0.11]	1270 [−0.20]	1110 [−0.36]	972 [0.01]	844 [−0.16]	733 [−0.03]	634 [−0.05]	547 [−0.10]	471 [−0.13]	404 [−0.36]
	60	1430 [0.35]	1250 [−0.06]	1100 [0.56]	955 [0.16]	828 [−0.07]	719 [0.13]	622 [0.20]	535 [−0.07]	460 [−0.18]	396 [0.03]
	75	1400 [0.19]	1230 [0.35]	1070 [−0.12]	935 [0.19]	810 [−0.07]	702 [−0.01]	606 [−0.10]	522 [−0.18]	448 [−0.42]	385 [−0.34]
	90	1370 [−0.02]	1200 [−0.12]	1050 [0.03]	915 [0.09]	794 [0.04]	687 [−0.03]	593 [−0.10]	511 [−0.11]	439 [−0.22]	377 [−0.18]
	105	1350 [0.28]	1180 [−0.01]	1030 [−0.07]	899 [0.18]	780 [0.13]	675 [0.10]	582 [−0.06]	501 [−0.15]	431 [−0.10]	370 [−0.08]
	120	1330 [0.31]	1160 [−0.17]	1020 [0.51]	886 [0.29]	769 [0.30]	665 [0.20]	573 [−0.01]	493 [−0.13]	424 [−0.10]	364 [−0.07]
	135	1310 [0.06]	1150 [0.23]	1010 [0.81]	875 [0.33]	759 [0.29]	657 0.31	566 0.09	487 −0.03	418 −0.19	358 [−0.39]

Table 3: Total sales obtained through simulation and the difference (in %) between our simulation and computational results (in brackets).

Comparison of the Simulation Results to the Exact Results for $E(Z)$ We simulate the online system—when the values of m and v are given as in Section 5.1—to find the total purchase amounts, and compare the simulation results with those computed by evaluating $E(Z)$ in (11). For our simulation, we perform 100 replications each terminating when 10,000 visitors have arrived. For example, when $m = 25\%$ and $u = 75$, we find that the average value of the total purchase amount from our simulation is \$122,997. In order to examine the accuracy of this simulation result, we use Output Analyzer, an add-on software of Arena, to find that the half width for the 95% confidence level is 651, which is significantly small compared with the simulated purchase amount \$122,997. In fact, this simulation result is very close to the result $E(Z) = \$122,571$ that is computed using (11). The difference between our simulation and the computational results is only 0.35%.

In online Table 3, we present the simulation values of $E(Z)$ for each pair of m and u, when m is increased from 0.125 to 1.25 in steps of 0.125 and u is increased from 0 to 135 in steps of 15. Moreover, we compare these results with those computational results in online Table A, and compute the differences in percentage. We find from Table 3 that the maximum difference is only 1.84% (when $m = 125\%$ and $u = 0$); this demonstrates that the results generated by our simulation model are accurate and thus reliable for our subsequent search for optimal solutions.

Finding the Optimal Solution with OptQuest   We use OptQuest—which is an optimization add-on for Arena—to search for the optimal solution (m^*, u^*, v^*). Note that, given the values of m, u and v, the service level $G_{m,u}(v)$ (generated by our simulation model) is the ratio of the total realized sales to the sum of the total realized sales and the lost sales. We find from (2) that the optimal service level $G_{m,u}(v^*)$, when m and u are given, is computed as $m/[I(1 + m) + m]$. Thus, we search for the optimal solution (m^*, u^*, v^*) that maximize the profit (which is computed as, sale revenue − the acquisition cost − total shipping cost − inventory cost) subject to the service level (generated by the simulation) equal to the optimal level $m/[I(1 + m) + m]$.

In agreement with [16], we set $I = 0.385\%$. In addition, we find from [27] that, when a consumer's *final* purchase amount is $\bar{Y}(m, u) = \bar{y}$, the shipping cost $S(\bar{y})$ is given as

$$S(\bar{y}) = \begin{cases} 6.5, & \text{if } 0 \leq \bar{y} \leq 50; \\ 7.5, & \text{if } 50 < \bar{y} \leq 75; \\ 10, & \text{if } \bar{y} > 75. \end{cases} \tag{13}$$

Note that, as Lewis et al. [27] stated, the shipping cost $S(\bar{y})$ in (13) was estimated by the firm in their empirical study.

Setting the initial point $(m^0, u^0, v^0) = (25\%, 100, 85000)$, we use Arena and OptQuest to search for the optimal solution, which is found as $(m^*, u^*, v^*) = (52.13\%, 156.07, 88783.64)$. The resulting expected (average) profit is $\$26,876.5$, and the expected service level is 98.89%, which equals what we find by computing $m/[I(1+m)+m]$. The statistical summary generated by Arena's Output Analyzer is provided as follows:

	Average	Minimum	Maximum	Half Width for the 95% Conf. Level
Profit	$\$26,876.5$	$\$26,611.5$	$\$27,141.5$	265
Service level	98.89%	98.33%	99.45%	0.56%

From the above statistical summary, we find that, for both profit and service level, the half width for the 95% confidence level is significantly smaller than the average value. This demonstrates that the number of replications (i.e., 100) is enough to assure the accuracy of our results.

The above results indicate that, given an initial point, we can use Arena and OptQuest to find an optimal solution for our two-stage problem. To show the uniqueness of the above optimal solution [i.e., $(m^*, u^*, v^*) = (52.13\%, 156.07, 88783.64)$], we need to examine whether or not the search process with different initial points can converge to this solution. For the convergence test, we consider 20 different initial points by, *ceteris paribus*, (i) increasing the value of m^0 from 5% to 50% in increments of 5%, and (ii) increasing the value of u^0 from 70 to 160 in steps of 10. Note that we do not set different initial points of v for the convergence test, because, as Theorem 1 shows, the retailer's optimal inventory value v^* is uniquely defined by the values of m^* and u^*.

For each initial point, we repeat the above search process to find a corresponding optimal solution, as given in Table 4, where we find that the optimal solutions for all initial points are very close to that found above when the initial point is $(m^0, u^0) = (25\%, 100)$. Such a result demonstrates the convergence of the two-stage decision model that is specified in Section 3.

From our simulation result we find that, if the online retailer adopts the two-stage decision process, then the retailer should set a high CFS cutoff level so that most consumers absorb the shipping costs. However, setting a high free-shipping threshold may deter some consumers from making purchases. To assure profitability, this retailer should also increase the profit

m^0	m^*	u^*	u^0	m^*	u^*
5%	52.11% (↓ 0.04%)	156.063 (↓ 0.004%)	70	52.12% (↓ 0.02%)	156.074 (↑ 0.003%)
10%	52.13% (↓ 0%)	156.074 (↑ 0.003%)	80	52.12% (↓ 0.02%)	156.066 (↓ 0.003%)
15%	52.13% (↓ 0%)	156.067 (↓ 0.002%)	90	52.13% (↓ 0%)	156.073 (↑ 0.002%)
20%	52.12% (↓ 0.02%)	156.071 (↑ 0.001%)	**100**	**52.13%**	**156.07**
25%	**52.13%**	**156.07**	110	52.13% (↓ 0%)	156.068 (↓ 0.001%)
30%	52.13% (↓ 0%)	156.066 (↓ 0.003%)	120	52.13% (↓ 0%)	156.071 (↑ 0.001%)
35%	52.13% (↓ 0%)	156.073 (↑ 0.002%)	130	52.13% (↓ 0%)	156.062 (↓ 0.005%)
40%	52.12% (↓ 0.02%)	156.061 (↓ 0.006%)	140	52.11% (↓ 0.04%)	156.070 (↓ 0%)
45%	52.13% (↓ 0%)	156.072 (↑ 0.001%)	150	52.12% (↓ 0.02%)	156.069 (↓ 0.001%)
50%	52.14% (↑ 0.02%)	156.061 (↓ 0.006%)	160	52.13% (↓ 0%)	156.075 (↑ 0.003%)

Table 4: The convergence of our two-stage decision model. Note that the mark "↑" or "↓" represents the percentage increase or decrease compared with the optimal solution when the initial point is $(m^0, u^0) = (25\%, 100)$.

	$I/S(\bar{y})$			**Optimal Decisions**			
Case	Small Orders	Medium Orders	Large Orders	m^*	u^*	v^*	$\pi(m^*, u^*, v^*)$
1	0.13%	0.12%	0.11%	72.11%	0	91,341.52	30,512.33
2	0.12%	0.11%	0.10%	64.34%	0	92,197.31	28,442.13
3	0.11%	0.10%	0.09%	59.27%	39.16	89,072.61	27,532.36
4	0.10%	0.09%	0.08%	55.93%	91.56	88,399.14	27,219.42
5	0.09%	0.08%	0.07%	54.38%	128.73	88,451.16	27,130.38
6	0.08%	0.07%	0.06%	53.21%	141.11	88,523.82	27,002.52
7	0.07%	0.06%	0.05%	52.16%	150.29	88,672.44	26,962.65
8	0.06%	0.05%	0.04%	52.13%	156.07	88,783.64	26,876.50
9	0.05%	0.04%	0.03%	52.12%	157.16	89,491.52	26,641.21
10	0.04%	0.03%	0.02%	52.11%	159.68	89,221.14	26,522.23

Table 5: The online retailer's optimal decisions and maximum profit obtained from simulation with different values of the ratios of I to the unit shipping costs for three order categories.

margin to a high level.

5.3 Sensitivity Analysis

For our simulation in Section 5.2, we need the values of I (% of inventory value in dollar) and the unit shipping cost $S(\bar{y})$ (for the purchase amount $\bar{y}$), which, however, may be inaccurate. Hence, we next perform a sensitivity analysis to address the following critical question: How will the online retailer's optimal decisions and maximum profit change if the values of I and $S(\bar{y})$ are different from the above estimated values?

To simplify our analysis, we consider the impact of the ratio of I to the unit shipping costs for small, medium, and large orders on the optimal decisions (m^*, u^*, v^*) and the maximum expected profit $\pi(m^*, u^*, v^*)$. Noting from our numerical experiment in Section 5.2 that the ratios of I to the unit shipping costs for small, medium, and large orders are 0.06%, 0.05%, and 0.04%, we increase the ratio of I to the unit shipping cost for small orders from 0.04% to 0.13% in increments of 0.01%. We also increase the ratio of I to the unit shipping cost for medium orders from 0.03% to 0.12% in steps of 0.01%, and the ratio of I to the unit shipping cost for large orders from 0.01% to 0.11% in steps of 0.01%. At each step, we simulate the operation of the online retailer for the corresponding ratios for three categories of orders. Thus, as shown in Table 5, we consider ten cases for the sensitivity analysis.

To observe the impact of the ratio on m^* and u^*, we plot Figure 2(a), which shows that, as the ratio decreases, the retailer should decrease his optimal profit margin but increase his optimal CFS cutoff level at a decreasing rate. Another important finding is drawn as follows: when the ratio is larger than a specific level (e.g., the ratios for the small, medium and large orders are 0.09%, 0.08%, and 0.07%, respectively), the optimal profit margin and CFS threshold change largely; but, when the ratio is smaller than the specific level, the optimal profit margin and free-shipping threshold change at a small rate.
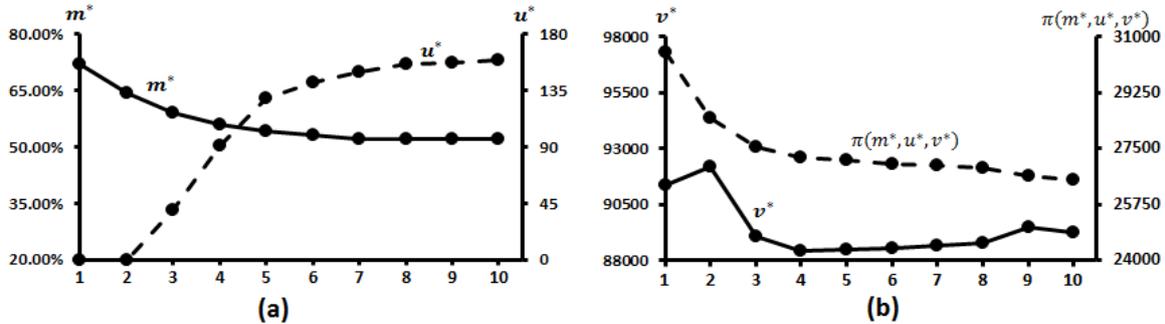

Figure 2: The impact of the ratios of I to the unit shipping costs for three order categories on the online retailer's optimal decisions (i.e., m^*, u^*, and v^*) and maximum profit $\pi(m^*, u^*, v^*)$. Note that the horizontal axis represents ten cases each corresponding to a set of three ratios for the small, medium and large orders, as shown in Table 5.

As Figure 2(b) indicates, the plot of the optimal inventory value includes four segments, which are specified as follows:

1. The first segment corresponds to the situation in which the shipping cost relative to the inventory cost is sufficiently small such that the ratios for the small, medium and large orders are larger than 0.12%, 0.11%, and 0.10%, respectively. In this segment, when the ratio decreases, the optimal inventory value should increase.

2. The second segment corresponds to the situation in which the ratios for the small, medium and large orders are respectively smaller than 0.12%, 0.11%, and 0.10% but greater than 0.10%, 0.09%, and 0.08%. For this segment, we find that, as the ratio decreases (i.e., the relative shipping cost increases), the optimal inventory value decreases.

3. The third segment corresponds to the situation in which the ratios for the small, medium and large orders are respectively smaller than 0.10%, 0.09%, and 0.08% but greater than 0.05%, 0.04%, and 0.03%. We find that, in this segment, the retailer should increase his inventory as the ratio decreases.

4. The fourth segment corresponds to the situation in which the ratio for the small, medium and large orders are respectively smaller than 0.05%, 0.04%, and 0.03%. For this segment, the retailer should decrease his inventory as the ratio is reduced.

We find that the curve for the total purchase amount in Figure 1(a) and that for the inventory value in Figure 2(b) are very similar. The reason is as follows. From (2) we find that the optimal service level $G_{m,u}(v^*)$ depends on I and m. Thus, when we increase the shipping cost $S(\bar{y})$, m decreases [as shown in 2(a)] and thus $G_{m,u}(v^*)$ also decreases. However,

according to our simulation, we find that the service level decreases at a very small rate. Note that the service level in our simulation model can be actually approximated as: min(inventory value,total purchase amount)/total purchase amount. As a result, the optimal inventory should be roughly proportional to the total purchase amount.

We also find from Figure 2(b) that, as the ratio decreases (i.e., the shipping cost relative to the inventory cost increases), the retailer's maximum profit decreases. This reveals the important fact that the shipping cost significantly affects the profitability of the online retailer. Even though the retailer uses the pricing (profit margin) and free-shipping strategies, the retailer still experiences loss of profits if the shipping cost is increased.

6   Conclusion

We consider a two-stage decision process in which the retailer first makes optimal decisions on the profit margin and the contingent free shipping (CFS) threshold, and then determines the optimal inventory value that maximizes the expected profit for a promotion period. Since the two-stage decision model is very complicated, it is impossible to find the optimal solution analytically and numerically without simulation. Hence, with empirical data provided by Lewis, Singh and Fay [27], we use Arena to simulate the operation of the online retailer, and then use OptQuest to find the optimal solutions. In addition, we conduct sensitivity analyses to examine how the retailer's optimal decisions and maximum profit change when the ratio of the unit holding and salvage cost to the unit shipping cost is different from that in the empirical study in [27]. The sensitivity analysis also delivers some important managerial insights.

The main contributions of our paper are summarized as follows.

1. We simulate an online retailing operation using the statistics of the retailer's sales (e.g., percentages of small, medium and large orders) provided by Lewis, Singh, and Fay [27]. Even with such limited data, we find the best-fitting purchase amount function and conversion rate. Thus, we provide a framework for numerical studies with limited empirical data.

2. We find that the reduction of profit margins does not significantly change the standard deviation of consumers' order sizes (purchase amounts) but increases the average order size. However, if the retailer applies a nonzero, finite CFS threshold, then both average value and standard deviation of consumers' order sizes are changed.

3. We use Arena to simulate the online retailing operation and use OptQuest to find the optimal decisions and maximum profit. The simulation with Arena and the optimization with OptQuest are extensively used in industry but, to our knowledge, have not been widely applied in academia, as discussed in Section 1. In our paper, we provide a particular discussion about how to use Arena for simulation and how to use OptQuest for optimization. This is expected to help other scholars use this technique for their research.

4. We conduct a sensitivity analysis, and draw some important managerial insights. For

example, the ratio of inventory to shipping cost significantly affects the retailer's optimal decision and maximum profit. As the ratio decreases, the retailer should decrease his optimal profit margin but increase his optimal CFS threshold at a decreasing rate. Moreover, the expected profit is increasing in the ratio.

In this paper, we have made four major model assumptions, which may limit the applicability of our model. In the future, it would be useful to develop a new (and feasible) method to address our research problems without these assumptions. For example, we may relax Assumption 1 to consider multiple products with different profit margins. The relaxation of one or more assumptions can generate more interesting results even though it would result in very complicated optimization problems.

References

[1] Advertising.Com. Holiday shopping study: Consumer shopping patterns and advertising performance, 2007. Optigence Research report.

[2] N. A. H. Agatz, M. Fleischmann, and J. A. E. E. van Nunen. E-fulfillment and multi-channel distribution—a review. *European Journal of Operational Research*, 187(2):339–356, 2008.

[3] E. T. Anderson, G. J. Fitzsimons, and D. Simester. Measuring and mitigating the costs of stockouts. *Management Science*, 52(11):1751–1763, 2006.

[4] S. P. Anderson, A. de Palma, and J. F. Thisse. *Discrete Choice Theory of Product Differentiation*. MIT Press, Cambridge, MA, 1992.

[5] N. Aras, V. Verter, and T. Boyaci. Coordination and Priority Decisions in Hybrid Manufacturing/Remanufacturing Systems. *Production and Operations Management*, 15(4):528–543, 2006.

[6] R.G. Askin and J. Chen. Dynamic task assignment for throughput maximization with worksharing. *European Journal of Operational Research*, 168(3):853–869, 2006.

[7] A. Ayanso, M. Diaby, and S. K. Nair. Inventory rationing via drop-shipping in Internet retailing: A sensitivity analysis. *European Journal of Operational Research*, 171(1):135–152, 2006.

[8] R. C. Blattberg and S. A. Neslin. *Sales Promotion: Concepts, Methods and Strategies*. Prentice Hall, Englewood Cliffs, NJ, 1990.

[9] P. Bradford and M. Katehakis. Constrained inventory allocation and its applications. *WSEAS Transactions on Mathematics*, 6(2):263–270, 2007.

[10] E. Breugelmans, K. Campo, and E. Gijsbrechts. Opportunities for active stock-out management in online stores: The impact of the stock-out policy on online stock-out reactions. *Journal of Retailing*, 82(3):215–228, 2006.

[11] G. P. Cachon and A. G. Kok. Category management and coordination in retail assortment planning in the presence of basket shopping consumers. *Management Science*, 53(6):934–951, June 2007.

[12] E. M. Dinlersoz and H. Li. The shipping strategies of Internet retailers: Evidence from internet book retailing. *Quantitative Marketing and Economics*, 4:407–438, 2006.

[13] L. Dong, P. Kouvelis, and Z. Tian. Dynamic pricing and inventory control of substitute products. *Manufacturing & Service Operations Management*, 11(2):317–339, Spring 2009.

[14] M. Ettl, G. E. Feigin, G. Y. Lin, and D. D. Yao. A supply network model with base-stock control and service requirements. *Operations Research*, 48(2):216–232, March–April 2000.

[15] M. Gümüş, S. Li, W. Oh, and S. Ray. Shipping fees or shipping free? a tale of two price partitioning strategies in online retailing, 2012. To appear in *Production and Operations Management*.

[16] G. Hadley and T. M. Whitin. *Analysis of Inventory Systems*. Prentice-Hall, Englewood Cliffs, N.J., 1963.

[17] G. Hua, S. Wang, and T. C. E. Cheng. Price and lead time decisions in dual-channel supply chains. *European Journal of Operational Research*, 205(1):113–126, 2010.

[18] G. Hua, S. Wang, and T. C. E. Cheng. Optimal order lot sizing and pricing with free shipping. *European Journal of Operational Research*, 218(2):435–441, 2012.

[19] G. Hua, S. Wang, and T. C. E. Cheng. Optimal pricing and order quantity for the newsvendor problemwith free shipping. *International Journal of Production Economics*, 135(1):162–169, 2012.

[20] W. Jang and D. Kim. Integrated production and allocation policies with one direct shipping option. *European Journal of Operational Research*, 181(2):716–732, 2007.

[21] W. Jang, D. Kim, and K. Park. Inventory allocation and shipping when demand temporarily exceeds production capacity. *European Journal of Operational Research*, 227(3):464–470, 2013.

[22] X. Jing and M. Lewis. Stockouts in online retailing. *Journal of Marketing Research*, XLVIII:342–354, April 2011.

[23] W. D. Kelton, R. P. Sadowski, and D. T. Sturrock. *Simulation with Arena*. McGraw-Hill, New York, fourth edition, 2007.

[24] M. Leng and R. Becerril-Arreola. Joint pricing and contingent free-shipping decisions in B2C transactions. *Production and Operations Management*, 19(4):390–405, July–August 2010.

[25] M. Leng and M. Parlar. Free shipping and purchasing decisions in B2B transactions: A game-theoretic analysis. *IIE Transactions*, 37:1119–1128, 2005.

[26] M. Lewis. The effect of shipping fees on customer acquisition, customer retention, and purchase quantities. *Journal of Retailing*, 82(1):13–23, 2006.

[27] M. Lewis, V. Singh, and S. Fay. An empirical study of the impact of nonlinear shipping and handling fees on purchase incidence and expenditure decisions. *Marketing Science*, 25(1):51–64, January–February 2006.

[28] S. Ross. *Stochastic Processes*. John Wiley, New York, 1983.

[29] H. E. Thompson. Inventory management and capital budgeting: A pedagogical note. *Decision Sciences*, 6(2):383–398, April 1975.

[30] M. C. van der Heijden. Multi-echelon inventory control in divergent systems with shipping frequencies. *European Journal of Operational Research*, 116(2):331–351, 1999.

[31] D. Yao. Christmas creeps into stores, 2006. `http://www.signonsandiego.com/uniontrib/20061025/news_1n25creep.html` (URL last accessed in April 8, 2013.).

[32] Y. Zhao and M. Katehakis. On the structure of optimal ordering policies for stochastic inventory systems with minimum order quantity. *Probability in the Engineering and Informational Sciences*, 20(2):257–270, 2006.

[33] B. Zhou, M. Katehakis, and Y. Zhao. Managing stochastic inventory systems with free shipping option. *European Journal of Operational Research*, 196(1):186–197, 2009.

[34] B. Zhou, Y. Zhao, and M. Katehakis. Effective control policies for stochastic inventory systems with a minimum order quantity and linear costs. *International Journal of Production Economics*, 106(2):523–531, 2007.

Online Appendix A   Estimation and Computation Results

A.1   Explanation of the Function $K(m, u)$ in (7)

Note that in (7) we use a linear function for the first term in order to reflect the fact that a sufficiently large profit margin (m) results in very high retail prices (for all products) and thus no consumer would purchase from the online retailer. However, a large value of m in (7) could drive $K(m, u)$ negative which may result in a situation where the initial purchase amount r.v. $Y = K(m, u) + X_0$ could assume negative values; this is an unrealistic situation since purchase amounts are nonnegative quantities. Fortunately, for the problem we are studying with the empirical data in online Tables B and C, and Table 2, we later show in Remark 3 that the probability of negative purchase amounts is negligibly small.

We use an exponential function (rather than a linear function) for the second term of $K(m, u)$ in (7), because as u (the CFS cutoff level) approaches infinity (i.e., the retailer doesn't offer free shipping service), some consumers would still buy from the retailer. If we use a linear function (as in the first term), then our model will imply that no consumer buys the products online when u is sufficiently large; this is inconsistent with the empirical data (for example, 18.37% of all visitors place online orders when $u = \infty$ and $m = 25\%$) in online Tables B and C.

A.2   Estimate of $\bar{Y}(m, u)$ in (10)

We now estimate the values of the parameters that determine $\bar{Y}$ in (10). It is easy to see that as indicated in Section 4.2, for the base policy ($m = 25\%$ and $u = \infty$), we have $K(25\%, \infty) = 0$ and $Y = Y_0$; this means that the first condition above is satisfied by the function (7). For the second condition (i.e., $K(25\%, 0) = 2.3$), we find that $\alpha_2 = 2.3$, and the function in (7) can be re-written as

$$K(m, u) = \alpha_1(0.25 - m) + 2.3 \exp(-\gamma u). \tag{14}$$

For the third condition above (i.e., $K(0.125, \infty) = 0.9$), we solve $0.125\alpha_1 = 0.9$ and find $\alpha_1 = 7.2$; thus, the function in (14) can be re-written as

$$K(m, u) = 7.2(0.25 - m) + 2.3 \exp(-\gamma u). \tag{15}$$

Substituting (15) into (6) gives

$$Y = K(m, u) + Y_0 = 7.2(0.25 - m) + 2.3 \exp(-\gamma u) + Y_0. \tag{16}$$

We now return to the policy of "free shipping to large orders" ($m = 25\%$, $u = \$75$) and use the empirical data corresponding to this policy to estimate the value of γ. For this policy, the random purchase amount function (16) is

$$Y = K(0.25, 75) + Y_0 = 2.3 \exp(-75\gamma) + Y_0. \tag{17}$$

We learn from Section 4.2 that, with the probability $\rho(y, u)$ in (8), a consumer with an order size y that is smaller than u may increase his or her order size from y to $u + x$, where X is a r.v. with the p.d.f $\phi(t)$ in (9). Note that, when a consumer's initial purchase amount y is greater than or equal to u, the consumer doesn't need to increase y so that $\rho(y, u) = 0$ and $\bar{Y}(m, u) = Y$. Thus, for finite u, we can use (16) to calculate the expected final purchase amount of each customer as follows:

$$
\begin{aligned}
E[\bar{Y}(m, u)] \;=\; & \int_{u-K(m,u)}^{\infty} [y_0 + K(m, u)]\hat{f}_0(y_0)\ dy_0 \\
& + \int_0^{u-K(m,u)} \{\rho(y, u)(u + \beta) + [1 - \rho(y, u)][y_0 + K(m, u)]\}\hat{f}_0(y_0)\ dy_0,
\end{aligned}
$$

or more simply,

$$
\begin{aligned}
E[\bar{Y}(m, u)] \;=\; & E(Y_0) + K(m, u) \\
& + \int_0^{u-K(m,u)} \rho(y_0 + K(m, u), u)\{u + \beta - [y_0 + K(m, u)]\}\hat{f}_0(y_0)\ dy_0, \tag{18}
\end{aligned}
$$

where $E(Y_0) = 58.61$, because Y_0 represents the random purchase amount for the base policy. When $u = \$75$ and $m = 25\%$, $Y = K(0.25, 75) + Y_0 = 2.3 \times \exp(-75\gamma) + Y_0$, as shown in (17); and we thus have the expected purchase amount as

$$E[\bar{Y}(0.25, 75)] = 58.61 + K(0.25, 75) + \int_0^{75-K(0.25,75)} \rho(y, 75)[K(0.25, 75) + \beta - y]\hat{f}_0(y_0)\ dy_0, \tag{19}$$

where $y = y_0 + K(0.25, 75)$, and

$$
\rho(y, 75) = \begin{cases} \exp[-r(75 - y_0 - K(0.25, 75))], & \text{if } y_0 \le 75 - K(0.25, 75), \\ 0, & \text{otherwise.} \end{cases} \tag{20}
$$

We need to estimate the values of parameters γ [in the random purchase amount function (17)], β [in the p.d.f. (9)] and ρ [in the function (20)]. From online Table B, we can find the empirical data (i.e., average order size and order-size distribution) for the policy of "free

shipping to large orders." Next, we find the values of the parameters γ, β and ρ that best fit the empirical data. The average order size $E(\bar{Y} \mid 0.25, 75)$ is computed as in (19). We calculate the order-size distribution below.

If a consumer with an initial order size $Y = y$—which is computed as $K(0.25, 75) + y_0$, as discussed above—is smaller than \$50, then the consumer increases his or her initial purchase amount to the free-shipping threshold \$75 or more with probability $\rho(y, 75)$, and keeps the initial purchase amount with probability $[1 - \rho(y, 75)]$. Therefore, the percentage of small orders (each with an order size smaller than or equal to \$50) is computed as $\Pr(\bar{Y} \leq 50) = \int_0^{50-K(0.25,75)} [1 - \rho(y, 75)] \hat{f}_0(y_0) \, dy_0$, where $y = K(0.25, 75) + y_0$. Similarly, we can calculate the percentage of medium orders (each with an order size between \$50 and \$75) as $\Pr(50 < \bar{Y} \leq 75) = \int_{50-K(0.25,75)}^{75-K(0.25,75)} [1 - \rho(y, 75)] \hat{f}_0(y_0) \, dy_0$, and calculate the percentage of large orders (each with an order size greater than \$75) as $\Pr(\bar{Y} > 75) = 1 - \Pr(\bar{Y} \leq 50) - \Pr(50 \leq \bar{Y} \leq 75)$.

From online Table B, we find that, for the policy of "free shipping to large orders" ($u = \$75$), the average order size is \$63.13; and the percentages of small, medium and large orders are 52.12%, 6.94% and 40.94%. Like in Example 1, we solve the following minimization problem to find the optimal values of (γ^*, r^*, β^*) that best fit the above empirical data.

$$\min_{\gamma, r, \beta} \ \Pi = [\Pr(\bar{Y} \leq 50) - 0.5212]^2 + [\Pr(50 \leq \bar{Y} \leq 75) - 0.0694]^2, \text{ s.t. } E(\bar{Y}) = 63.13.$$

We find that the optimal solution is $(\gamma^*, r^*, \beta^*) = (10.55, 0.07, 26.64)$, which results in the minimum squared deviation 1.32/1000. Thus, we can obtain the parameter values in $\bar{Y}(m, u)$ as given in Section 4.2.3.

A.3   Estimate of the Conversion Rate Model

We next use the data (on the conversion rates for all policies) in online Tables B and C to estimate the parameters of the model for $R(m, u)$ in (12). For both base policy and the policy of "10% discount", no free-shipping service is offered. That is, for these two policies, the CFS cutoff level u approaches infinity, i.e., $u \to \infty$ and as a result, we find that $R(m, \infty) = \exp(\theta_1 m + \tau)/[\exp(\theta_1 m + \tau) + 1]$. From online Table B, we find that, for the base policy ($m = 25\%, u = \infty$), the conversion rate is 18.37%, and we thus have

$$R(0.25, \infty) = \frac{\exp(\theta_1 \times 0.25 + \tau)}{\exp(\theta_1 \times 0.25 + \tau) + 1} = 0.1837. \tag{21}$$

In addition, we find from online Table C that, for the policy of "10% discount" ($m = 12.5\%, u = \infty$), the conversion rate is 20.67% and thus,

$$R(0.125, \infty) = \frac{\exp(\theta_1 \times 0.125 + \tau)}{\exp(\theta_1 \times 0.125 + \tau) + 1} = 0.2067. \tag{22}$$

We note that (21) and (22) are simultaneous equations in the parameters θ_1 and τ. Solving (21) and (22) for θ_1 and τ gives $\theta_1 = -1.17$ and $\tau = -1.2$. We substitute the values of θ_1 and

τ into the incidence function $R(m, u)$ in (12), and find that

$$R(m, u) = \frac{\exp[-1.17m + (\omega + \theta_2 u)^{-1} - 1.2]}{\exp[-1.17m + (\omega + \theta_2 u)^{-1} - 1.2] + 1}. \tag{23}$$

We then find from Table B that, for the policy of "free shipping to all orders" ($m = 25\%, u = 0$), the conversion rate is 21.65%, and thus,

$$R(0.25, 0) = \frac{\exp(-1.17 \times 0.25 + \omega^{-1} - 1.2)}{\exp(-1.17 \times 0.25 + \omega^{-1} - 1.2) + 1} = 0.2165,$$

and solving it for ω gives $\omega = 4.85$. We substitute the value of ω into the incidence function (23), and find

$$R(m, u) = \frac{\exp[-1.17m + (4.85 + \theta_2 u)^{-1} - 1.2]}{\exp[-1.17m + (4.85 + \theta_2 u)^{-1} - 1.2] + 1}, \tag{24}$$

where only the parameter θ_2 needs to be estimated. As online Table B indicates, for the policy of "free shipping to large orders" ($m = 25\%, u = 75$), the conversion rate is 19.40%. Thus, we have

$$R(0.25, 75) = \frac{\exp[-1.17 \times 0.25 + (4.85 + \theta_2 \times 75)^{-1} - 1.2]}{\exp[-1.17 \times 0.25 + (4.85 + \theta_2 \times 75)^{-1} - 1.2] + 1} = 0.194$$

Solving this equation for θ_2 gives $\theta_2 = 0.13$. Hence, we find the conversion rate model as given in Section 4.3.

A.4   Computed Total Purchase Amount

As discussed in Section 5.1, we can solve (11) to find the expected total purchase amount $E(Z)$, for the given values of m and u. Now, we increase the value of m from 0.125 to 1.25 in steps of 0.125, and increase the value of u from 0 to 135 in steps of 15; and then, for each pair of m and u, we calculate the corresponding value of $E(Z)$. For our computational results, see Table A. Note that we use the data Table A to plot Figure 1 (a) and (b).

					m					
u	0.125	0.25	0.375	0.5	0.625	0.75	0.875	1	1.125	1.25
0	149,764.9	131,850.9	115,634.1	101,052.3	88,021.4	76,442.3	66,206.1	57,199.1	49,306.7	42,417.1
15	144,930.3	127,782.6	112,242.4	98,251.9	85,732.8	74,592.2	64,727.9	56,033.4	48,401.1	41,725.9
30	145,944.4	128,413.7	112,572.3	98,349.9	85,657.1	74,390.4	64,438.6	55,687.1	48,021.7	41,331.7
45	144,847.3	127,258.3	111,398.9	97,189.9	84,533.7	73,320.2	63,432.7	54,752.0	47,160.7	40,545.0
60	142,505.5	125,078.2	109,387.4	95,348.6	82,860.1	71,808.5	62,074.9	53,538.5	46,080.9	39,587.9
75	139,736.9	122,571.1	107,130.2	93,327.2	81,058.5	70,209.9	60,662.0	52,294.3	44,988.6	38,631.9
90	137,030.7	120,148.8	104,972.6	91,413.9	79,368.8	68,723.3	59,358.4	51,154.7	43,995.2	37,768.0
105	134,620.7	118,004.7	103,073.6	89,739.0	77,897.1	67,434.6	58,233.4	50,175.4	43,144.9	37,031.4
120	132,586.6	116,201.5	101,481.9	88,339.6	76,671.2	66,364.1	57,301.6	49,366.4	42,444.2	36,425.9
135	130,926.9	114,733.3	100,188.5	87,204.6	75,678.8	65,499.2	56,550.0	48,715.0	41,881.0	35,940.0

Table A: The expected total sales $E(Z)$ obtained by solving the function (11).
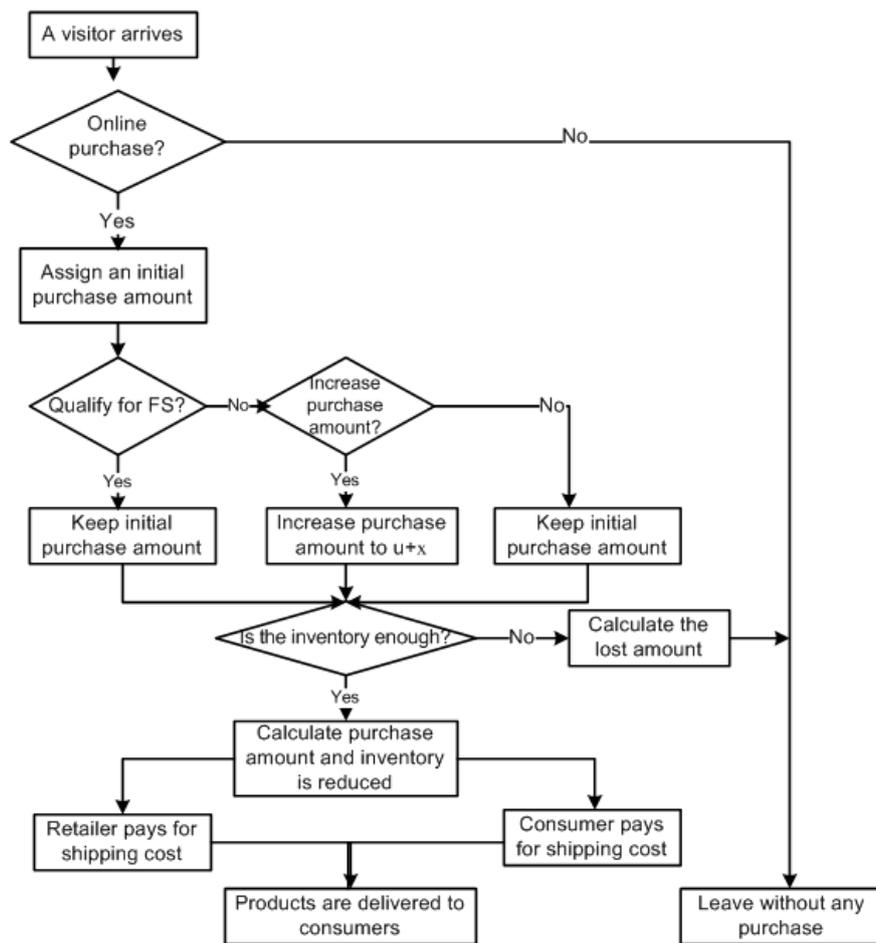
A.5 The Flow Chart of the Online Retailing Process

Figure A: The flow chart of the online retailing process for the simulation model.

Online Appendix B   Empirical Data from Lewis, Singh and Fay [27]

In Sections 4 and 5, we use the empirical data in [27] to estimate the model components and conduct simulation to search for optimal solutions. The empirical data is summarized as in online Tables B and C.

($m = 25\%$) Policy	Conversion Rate	Average Order Size (A)	Order-Size Distribution		
			% of Small Orders (P_S)	% of Medium Orders (P_M)	% of Large Orders (P_L)
Base policy ($u = \infty$)	18.37%	$58.61	52.38%	17.76%	29.86%
FS to all ($u = 0$)	21.65%	$60.91	50.21%	16.11%	33.68%
FS to large ($u = \$75$)	19.40%	$63.13	52.12%	6.94%	40.94%

Table B: The conversion rates and order-size distributions for three free-shipping policies—i.e., base policy, free shipping to all orders, and free shipping only to large orders.

($u = \infty$) Policy	Conversion Rate	A	Order-Size Distribution		
			P_S	P_M	P_L
Base policy ($m = 25\%$)	18.37%	$58.61	52.38%	17.76%	29.86%
10% discount ($m = 12.5\%$)	20.67%	$59.51	51.28%	17.59%	31.14%

Table C: The conversion rates and order-size distributions for two pricing (profit margin) policies—i.e., base policy and 10% discount. Note that the data for the base policy is the same as that in Table B.