Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Stockout-Based Substitution and Inventory Planning in Textbook Retailing

Joonkyum Lee, Vishal Gaur, Suresh Muthulingam, Gary F. Swisher

MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 18, No. 1, Winter 2016, pp. 104–121
ISSN 1523-4614 (print) | ISSN 1526-5498 (online)
http://dx.doi.org/10.1287/msom.2015.0551

Stockout-Based Substitution and Inventory Planning in Textbook Retailing

Joonkyum Lee
Sogang Business School, Sogang University, Seoul 121-742, Korea, jklee@sogang.ac.kr

Vishal Gaur
Johnson Graduate School of Management, Cornell University, Ithaca, New York 14853, vg77@cornell.edu

Suresh Muthulingam
Smeal College of Business, The Pennsylvania State University, University Park, Pennsylvania 16802, suresh@psu.edu

Gary F. Swisher
The Cornell Store, Cornell University, Ithaca, New York 14853, garyswisher@cornell.edu

We demonstrate the value of utility-based choice models to estimate demand and plan inventory for new and used textbooks in the presence of consumer choice and stockout-based substitution at a university textbook retailer. Demand information is censored, the exact time of stockout is not observed, and the short selling season often does not allow for replenishment. Using data for 26,749 book titles from 2007 to 2011 and a simulation experiment calibrated on real data, we show that an attribute-based choice model generates accurate demand estimates (mean absolute percentage error less than 1%) even when nearly 90% of the textbooks in the fit sample experience stockout. This performance is driven by the heterogeneity of product attributes and is robust to the occurrence of product returns. We implement this model at the bookstore in a controlled field experiment and obtain over 10% increase in profit. The results show that accounting for asymmetric and stockout-based substitution in demand estimation and inventory planning enables us to make systematic corrections in inventory mix and inventory level compared to the existing process.

Keywords: OM practice; retailing; inventory theory and control; demand estimation; stockout-based substitution
History: Received: November 13, 2012; accepted: May 16, 2015. Published online in *Articles in Advance* October 12, 2015.

1. Introduction

The estimations of consumer choice and substitution behavior are important problems in retailing. In most such contexts, the demand rate for one product depends on the set of products in the assortment, their attributes, and their availability in stock. Moreover, the demand and time of stockout are often not observed, and only sales data are available. Thus, without modeling choice and substitution behavior, demand will not be estimated accurately, and inventory decisions will be suboptimal. The recent literature in operations management has made significant advances in developing methods for choice-based estimation. However, there is limited evidence on the extent of benefit in practice, especially when the set of products changes each season, data are incomplete, stockouts are common, and optimization must be combined with the estimation of parameters. Our paper provides such evidence in the context of the textbook business of a college bookstore.

There are over 4,500 college stores in the United States catering to the needs of more than 20 million students with annual sales of over $10 billion. According to the National Association of College Stores, textbook sales constituted nearly 61% of the sales from college bookstores in 2011. The Cornell Store, a nonprofit college store owned and managed by Cornell University, has the objective of providing all textbooks prescribed by instructors for their courses at the university. However, the cost of meeting this objective has been increasing over time on account of competition from online retailers. Hence, the bookstore is interested in improving the profitability of its textbook business by better matching supply with demand. The store sells approximately 3,500 book titles or international standard book numbers (ISBNs), a unique numeric commercial identifier given to each book title, for various courses and instructors each semester. Each title can be stocked in two substitutable SKUs (product types), new and used books. Students may have varying preferences to buy new books, to buy used books, and to substitute between new and used books in the event of stockouts. From 2007 to 2011, at least one type of book stocked out in 44.2% of instances, indicating that stockouts and potential substitution can have a significant impact on business.

The buyers at the Cornell Store face several challenges in demand estimation and inventory planning. These challenges are common to university bookstores, which have similar business environments and managerial processes, and are also generalizable to other retail settings. First, only censored demand information is available. Although sales of each book type are observed, demand is not observed as a result of stockouts. Further, the bookstore does not track the times of purchase transactions, so the time of stockout occurrences is not known. Therefore, historical sales data do not serve as a good proxy for future demand. Second, the supply of used books is limited and uncertain because they are purchased from students and third-party vendors. Thus, it may not be possible to stock the desired mix of book types for each book title. This complicates inventory planning because the number of new books to be stocked is a function of the number of used books available due to substitution. Third, the textbook business has a short selling season of approximately two weeks at the start of a semester, which provides limited opportunity for replenishment and puts a premium on accurate demand estimation.

The objective of our paper is to investigate the value of using an attribute-based choice model to estimate demand and plan inventory in a real world context with the above challenges. We address this question in three steps. First, we set up and estimate a choice model with stockout-based substitution using data on 26,749 book titles spanning nine semesters from 2007 to 2011 for the Cornell Store. An attribute-based choice model is used because our data consist of observations on dissimilar short-life-cycle products, and we need to apply the model to estimate demand for new products introduced each semester. The attributes include the retail price of a new book, the number of courses a book title is used for, the average number of books per course, whether a textbook is required or optional, and others. Most discrete choice models using aggregate data have to deal with unknown market size and unobserved no-purchases. We are able to effectively handle the no-purchase option by using the number of students enrolled in all courses that have the book listed in their syllabus as an estimate of market size.

Second, we assess the accuracy of demand estimation in the presence of stockouts and substitution. In a simulation calibrated on real data, we find that when an attribute-based choice model is estimated on a cross-section of heterogeneous products, then accurate demand estimates can be obtained even under high stockout situations. For example, when 89% of SKUs get stocked out, the mean percentage error (MPE) in demand of our model is 0.5%; the benchmark model that ignores stockout information and substitution by assuming that demand equals sales significantly underestimates demand with an MPE of −26.7%; and the benchmark model that ignores substitution but accounts for stockout information by not censoring data overestimates demand with an MPE of 8.5%. In the online supplement (available at http://dx.doi.org/10.1287/msom.2015.0551), we test the robustness of our results to the heterogeneity of product attributes in the data set, the specification of the choice model, the size of the calibration data set, and the occurrence of unobserved product returns.

In the third step, we conduct a controlled field experiment at the Cornell Store to determine the profitability impact of optimal inventory planning based on our model in practice. The experiment involves a matched sample of 3 groups of 24 book titles each: group 1 is a control group with the Cornell Store managers deciding on the stocking levels using the existing process; for group 2, we provide managers with demand estimates from our model; and for group 3, we provide managers with suggested stocking levels from our inventory optimization model. Group 2 realizes 8.4% higher profit than group 1, and group 3 realizes 10.2% higher profit than group 1, both statistics being significant according to several tests. Most of the gain occurs because of improvement in demand estimation. In other words, managers are able to dramatically improve their stocking decisions when given model-based demand estimates. Optimization of the stocking quantity provides additional benefit beyond that from demand estimation.

This paper contributes to the literature by presenting practical evidence of the benefits from modeling substitution behavior in demand estimation and inventory planning using a large data set and a controlled experiment. It shows that high incidence of stockouts or unknown stockout times need not be deterrents in the accurate estimation of demand. The existing literature has investigated the value of stocking more inventory to learn about demand under censoring. Our analysis shows another approach; i.e., in situations when there are many products of heterogeneous attributes to be stocked, one can achieve similar learning without increasing inventory. Finally, our paper presents insights into why optimization-based inventory decisions yield higher profit than the stocking decisions made by buyers for substitutable products. These insights can be useful to improve managerial decision making even in situations where optimization is costly to implement.

2. Literature Review

This paper builds on the literature on choice-based demand estimation, assortment planning, and practical applications of these models. We discuss relevant papers and describe the contribution of our paper to this literature.

Multinomial logit (MNL) choice models are widely used in the operations literature (e.g., Vulcano et al. 2012, Musalem et al. 2010) and in economics and marketing (e.g., Train 2009, Guadagni and Little 1983). Researchers (e.g., Guadagni and Little 1983) show that the MNL model works well from a practical point of view in estimating demand for substitutable products. The MNL framework involves random utility maximization (RUM) with products characterized as combinations of attributes. It enables us to use a small number of variables to describe the choice model and facilitates application of the model fitted on historical data to new products. However, it suffers from the restrictive property of independence from irrelevant alternatives (IIA), which can be alleviated through more general choice models, such as the nested logit model (Ben-Akiva and Lerman 1985).

The seminal paper examining demand estimation for substitutable products in the presence of stockouts is by Anupindi et al. (1998), who propose an expectation-maximization (EM) algorithm to estimate demand rates and substitution probabilities while treating the time of stockout as missing data. Applying this model to a beverages vending machine example, they show that estimated demand rates differ from observed sales rates because of substitution. An implication of their model is that the number of parameters increases rapidly with the number of products. Conlon and Mortimer (2013) and Musalem et al. (2010) build on the work of Anupindi et al. (1998) by using an attribute-based choice model to estimate demand. This model has a parsimonious representation with only a few parameters, and it can be applied to nonstationary demand data for dissimilar or short-life-cycle products. Conlon and Mortimer (2013) develop an EM method to estimate choice parameters of a RUM model in this setting. Their method can become computationally complex as the number of stocked-out products increases, and they propose a heuristic for multiple stockouts to apply their EM method. Musalem et al. (2010) address this shortcoming by presenting a different method, a Bayesian approach that employs a Markov chain Monte Carlo simulation to estimate the posterior distribution of parameters for a random coefficients MNL choice model. They use their method to estimate lost sales, substitution patterns, and the effect of promotions.

Both Conlon and Mortimer (2013) and Musalem et al. (2010) are applicable to our setting. In particular, like those papers, we have nonstationary demand with dissimilar products, and the total number of consumers is known whereas information on stockouts is missing. The main difference is that, whereas those papers focus on estimating demand and lost sales, we build on their work by assessing the effect of stockouts on the accuracy of demand estimation and by demonstrating the profit impact in practice through a controlled field experiment. Moreover, since we have substitution between only two product types, new and used books, we can simplify demand estimation by writing the likelihood function of the choice model in closed form.

Vulcano et al. (2010) study demand estimation with substitution in a different setting where the time of stockout is observed. They demonstrate the impact of incorporating a choice model on demand estimation and report revenue improvement using simulation on airline industry data. Similar to Vulcano et al. (2010), Vulcano et al. (2012) propose an efficient EM method to estimate MNL choice parameters, demand rate, and lost sales. They show the effectiveness of their method using examples from the airline and the retail industries. van Ryzin and Vulcano (2015) generalize the work of Vulcano et al. (2012) by addressing the problem of jointly estimating the arrival rate and the nonparametric choice patterns using an EM approach. Newman et al. (2014) present an efficient method based on a marginal log likelihood function to estimate MNL parameters and arrival rate without using an EM algorithm. They show superior computational performance compared to the EM algorithm using data from the hotel industry. The data structures of these papers are different from ours because they have information on product availability at all time instances.

Our paper also builds on previous practice-based research that has demonstrated the benefits of optimization-based assortment planning. Kök and Fisher (2007) develop methods to estimate choice parameters and substitution rate from sales data under assortment-based substitution, and to optimize the assortment and number of facings of each item in a product category under an MNL choice model and a shelf space constraint. They implement their methods in a supermarket chain and report a significant increase in profit. Fisher and Vaidyanathan (2014) construct an attribute-based choice model with preference ordering and develop heuristics to find the optimal assortment for a convenience-store chain and for a tire retailer. Their method can estimate the demand for a product that was included in earlier assortments in only a subset of stores, and they consider the impact of substitution on assortment planning. Our paper contributes to this practice-based research by focusing on the impact of stockout-based substitution, rather than assortment-based substitution, on demand estimation, inventories, and profit.

Multi-item assortment/inventory planning problems with substitution constitute an important research area in which many theoretical advances have been made. van Ryzin and Mahajan (1999) solve this problem using an MNL choice model; Talluri and van Ryzin (2004) study an assortment problem under a general discrete choice model in the context of revenue management; Smith and Agrawal (2000) consider an exogenous choice model; and Nagarajan and Rajagopalan (2008) consider an aggregate-level demand specification for substitutable products. Mahajan and van Ryzin (2001) and Honhon et al. (2010) focus on stockout-based substitution taking sample-path-based approaches that explicitly account for the sequence of stochastic choices made by consumers and the resulting evolution of inventory levels. Kök et al. (2015) provide an extensive review of this literature. These studies focus on the optimization problem when the parameters of the choice models are known. We add to this literature by addressing whether the demand parameters can be estimated accurately under stockouts and substitution.

3. Estimation of Choice Model

This section presents our approach for estimating the parameters of the choice model from censored sales data. We set up the choice model for two products, used and new textbooks. The model is developed from attribute-based choice models used in the literature and can easily be generalized to other settings.

Let $\mathcal{I} := \{1, 2, \ldots, I\}$ denote the set of books or ISBNs stocked at the Cornell Store in a given semester. If a book is sold in more than one semester, we assign it a different index i in each semester. We follow this cross-sectional estimation approach because a large number of new books are introduced each semester. To estimate demand for such books, we need to assume that the form of the utility function is similar across books, thus leading to a cross-sectional model. Each ISBN is sold in two substitutable types, new and used, thus defining the consideration set $\mathcal{K}_i := \{n, u\}$. Each book type has M attributes. Let $x_{ik} := (x_{ik1}, \ldots, x_{ikM})'$ denote the vector of attributes of type k, with $X_i := (x_{in}, x_{iu})$ and $X := (X_1, \ldots, X_I)$. Some attributes, e.g., whether a book is required or optional for a course, are common across both book types. We also consider the no-purchase option as a book type indicated by subscript 0. The utility of consumer c buying book type k for ISBN i is given by

$$Utility_{cik} = \beta_k' x_{ik} + \varepsilon_{cik}.$$

Here, $\beta_k := (\beta_{k1}, \ldots, \beta_{kM})'$ are the coefficients to be estimated, and ε_{ik} represents the unknown random component. The utility of the no-purchase option is defined as $Utility_{ci0} = \varepsilon_{ci0}$. We index the coefficients β_k by k to capture differences in the coefficients of attributes across new and used books. Consumers may have a higher price elasticity for new books than used books. Consumers may also derive a higher utility from used books when a course prescribes a larger number of books.

We assume that substitution occurs between the types of a given ISBN i, but there is no substitution across ISBNs. In other words, we treat each ISBN as a separate product category. (In practice, different ISBNs for a course may form a single consideration set. We partially account for this by including the number of required and optional books listed in a course as an explanatory variable in the utility model.) We also assume that the random components ε_{cik} are independently and identically distributed random variables drawn from a Gumbel distribution. This assumption leads to an MNL choice model (Anderson et al. 1992). Thus, the probability that consumer c chooses product type k is given by

$$P_{cik}(\beta \mid \mathcal{K}_{i1}, X_i) = \frac{e^{\beta_k' x_{ik}}}{1 + \sum_{l \in \mathcal{K}_{i1}} e^{\beta_l' x_{il}}}$$

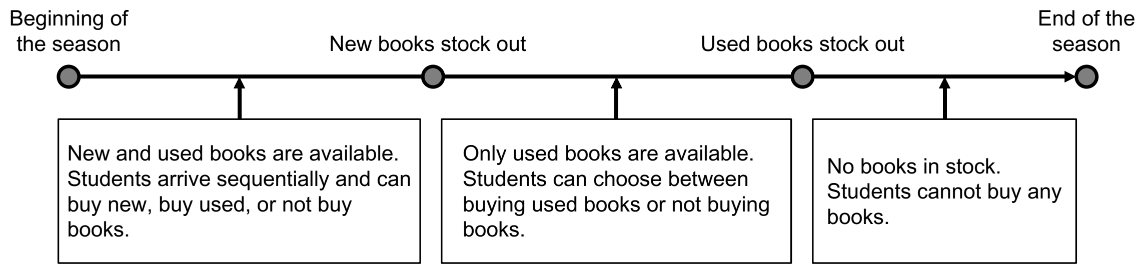
if k is included in the assortment $\mathcal{K}_{i1} \subseteq \mathcal{K}_i$, and 0 otherwise. The probability of no-purchase is $1 - \sum_{l \in \mathcal{K}_{i1}} P_{il}(\beta \mid \mathcal{K}_{i1}, X_i)$. We suppress the consumer subscript c throughout the rest of this paper.

This choice model is subject to the restrictions imposed by the IIA property. Although we use the MNL model at the Cornell Store for its simplicity, we evaluate the robustness of our results using a random coefficients MNL model in Appendix A.1 in the online supplement.

Our objective is to estimate the parameters β using observed data. The data set includes sales of each type of each textbook, whether a book type stocked out, several attributes of each book, a mapping of books to course–instructor combinations, and the estimated enrollment of each course–instructor combination available at the time of planning. Our setup is similar to those in Conlon and Mortimer (2013) and Musalem et al. (2010). Note that we have enrollment data, which is analogous to their assumption that the total number of consumers entering a retail store is known.

Consumers arrive sequentially and make buying decisions based on the availability of different book types at the time of the visit. A student may choose to buy a new book, buy a used book, or not buy any book. As the inventory gets depleted, if one type of book stocks out, then subsequent students can buy the other type or choose not to buy. If both book types stock out before the selling period ends, then subsequent students cannot buy any books from the store.

Figure 1 **Time Line Illustrating a Possible Sequence of Events in a Selling Period**



We do not distinguish between students who did not visit the store and those who visited but did not buy any books. Figure 1 illustrates a possible sequence of events wherein both book types are stocked initially, and new books stock out first followed by used books.

If the times of customer arrivals and purchases were observed, then we could apply methods such as in Vulcano et al. (2010) to compute the likelihood function and estimate the choice parameters. However, because we do not observe when each book type stocks out, we must account for all possible sequences of stockouts that yield observed sales. Musalem et al. (2010) solve this problem by using Gibbs sampling for an arbitrary number of product types. Since we have only two book types, it is feasible and efficient to compute the likelihood function directly. The likelihood function is constructed as follows: there are eight possible combinations of the initial assortment and final stockout situation. Table 1 lists all these scenarios and identifies them using the variable STI. For example, $STI = 3$ denotes the scenario in which both new and used books are included in the assortment and both types stock out. The likelihood function for this scenario is the most complex, and hence, we illustrate it in the following paragraph. Because all other scenarios are special cases of this scenario, we do not detail their likelihood functions here.

Consider ISBN i for which both new and used books stock out. Let s_{ik} denote the total sales of book type k for this ISBN, with $S_i := (s_{in}, s_{iu})$ and $S := (S_1, \ldots, S_I)$, and E_i denote its estimated enrollment, with $E := (E_1, \ldots, E_I)$. Let $L_i(\beta \mid STI_i = 3, S_i, E_i, X_i)$ denote the total likelihood for this observation as a function of the input sales S_i, enrollment E_i, and book attributes X_i. First, consider the set of sample paths along which new books stock out before used books. The total likelihood of the occurrence of all such sample paths is given by

$$\sum_{h_u=0}^{s_{iu}-1}\sum_{h_0=0}^{E_i-s_{in}-s_{iu}}\left[\frac{(s_{in}-1+h_u+h_0)!}{(s_{in}-1)!\,h_u!\,h_0!}(P_{in,nu})^{s_{in}}(P_{iu,nu})^{h_u}(P_{i0,nu})^{h_0}\cdot\left\{\sum_{h_{00}=0}^{E_i-s_{in}-s_{iu}-h_0}\frac{(s_{iu}-1-h_u+h_{00})!}{(s_{iu}-1-h_u)!\,h_{00}!}\cdot(P_{iu,u})^{(s_{iu}-h_u)}(P_{i0,u})^{h_{00}}\right\}\right]. \quad (1)$$

Here, h_u enumerates the number of used books sold before the last new book is sold, h_0 enumerates the number of customers who choose not to purchase anything when both book types are in stock, and h_{00} enumerates the number of customers who choose not to purchase anything when only used books are in stock. $P_{in,nu}$ denotes the probability of purchase of a new book when both book types are in stock, and so on. Thus, to compute the likelihood function, we enumerate and sum over combinatorial probability terms in which s_{in} new book purchases, h_u used book purchases, and h_0 no-purchase decisions occur before any stockout; then $s_{iu} - h_u$ used book purchases and $E_i - h_0 - s_{in} - s_{iu}$ no-purchase decisions occur after new books stock out.

Similarly, consider the set of sample paths along which used books stock out before new books. The total likelihood of the occurrence of all such sample paths is analogously given by

$$\sum_{h_n=0}^{s_{in}-1}\sum_{h_0=0}^{E_i-s_{in}-s_{iu}}\left[\frac{(s_{iu}-1+h_n+h_0)!}{(s_{iu}-1)!\,h_n!\,h_0!}(P_{in,nu})^{h_n}(P_{iu,nu})^{s_{iu}}(P_{i0,nu})^{h_0}\cdot\left\{\sum_{h_{00}=0}^{E_i-s_{in}-s_{iu}-h_0}\frac{(s_{in}-1-h_n+h_{00})!}{(s_{in}-1-h_n)!\,h_{00}!}\cdot(P_{in,n})^{(s_{in}-h_n)}(P_{i0,n})^{h_{00}}\right\}\right]. \quad (2)$$

Table 1 **Definition of Stockout Indicator *STI* Based on Potential Outcomes**

Stockout indicator (STI)	Initial assortment at Cornell Store	Stockout occurrence: New book stocks out	Stockout occurrence: Used book stocks out	Frequency in Cornell Store data (%)
0	(n, u)	No	No	21.7
1	(n, u)	Yes	No	8.0
2	(n, u)	No	Yes	22.9
3	(n, u)	Yes	Yes	6.9
4	(n)	No	Not applicable	29.5
5	(n)	Yes	Not applicable	5.5
6	(u)	Not applicable	No	4.5
7	(u)	Not applicable	Yes	0.9

Here, s_{iu} used book purchases, h_n new book purchases, and h_0 no-purchase decisions occur before any stockout; then $s_{in} - h_n$ used book purchases and $E_i - h_0 - s_{in} - s_{iu}$ no-purchase decisions occur after new books stock out. Thus, $L_i(\beta \mid STI_i = 3, S_i, E_i, X_i)$ is given by the sum of (1) and (2), and the total likelihood of observing the aggregate sales data is given by

$$L(\beta \mid STI, S, E, X) = \prod_{i \in \mathcal{J}} L_i(\beta \mid STI_i, S_i, E_i, X_i). \qquad (3)$$

The maximum likelihood estimator (MLE), $\hat{\beta}_{ML}$, for the parameters of the choice model is thus obtained by solving a nonlinear optimization problem in β:

$$\hat{\beta}_{ML} = \arg\max_{\beta} \log L(\beta \mid STI, S, E, X). \qquad (4)$$

We use a quasi-Newton method with the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm to solve this problem. The objective function can have multiple local optima. Therefore, we test a wide range of starting points and verify that they result in the same solution.

4. Simulation-Based Evaluation

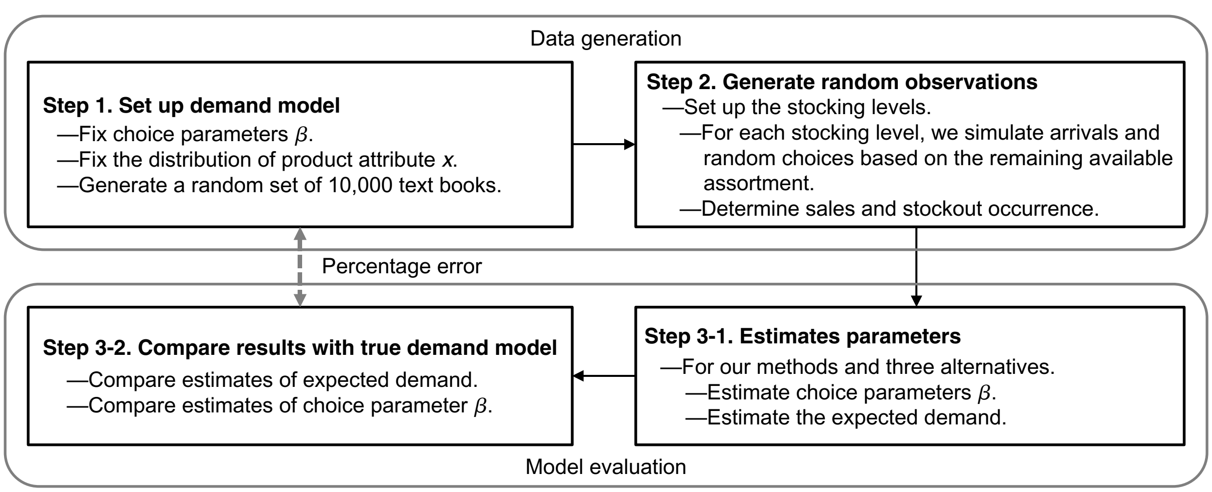
A key question for the success of the application at the Cornell Store is whether demand can be estimated accurately from available data. A simulation-based approach is necessary to address this question because true demand is unobserved in the real sales data. Thus, in this section, we present a simulation study to evaluate the ability of our model to provide accurate demand estimates under various stockout situations against benchmark models. The main insight from the results is that heterogeneity of product attributes in the data set helps to mitigate the effect of demand censoring and enables us to obtain accurate demand estimates even under high stockout rates. In Appendix A of the online supplement to the paper, we extend the simulation study to examine the robustness of its results to the number of book titles in the data set used for estimation and the occurrence of random product returns.

4.1. Experiment Design

The simulation experiment involves three steps as illustrated in Figure 2. In step 1, we extract a random set of 10,000 textbooks by sampling from the Cornell Store data. The random sample is representative of our data set and contains books with varying prices, other attributes, and course enrollments. Then we use two attributes of each book i, the price of a new book NP_i and whether the course is for freshmen $CL1_i$, to formulate the consumer's utility function. (The complete set of attributes used in the implementation is presented in §5.1. We use two of the most salient attributes here to focus on the main insights.) Using these attributes, we represent consumer choice using an MNL model. Thus, there are six parameters in the choice model: the intercept, the coefficient of NP, and the coefficient of $CL1$, each for new and used books. We set the parameters β_n and β_u as $(-1.2, -0.2, 0.7)$ and $(-1, -0.1, 0.4)$, respectively, to simulate demand data for the simulation. These numeric values are obtained from the Cornell Store data; we also test other parameter values for robustness and find that the insights are essentially the same.

In step 2, we simulate sales and stockout occurrences for each book i. To examine the effect of stocking level on the performance of the demand models, we select seven stocking levels corresponding to 0.5, 0.75, 1, 1.25, 1.5, 1.75, and 2 times the expected demand for each SKU. Thus, we have $7 \times 10{,}000$ combinations of book titles and stocking levels. We stock both new and used books for each book title to evaluate the effect of stockout-based substitution.

Figure 2 Simulation Process for Evaluating Our Demand Model Against Benchmark Models



For each stocking level and each textbook, we simulate sales by assuming that enrolled students arrive sequentially and make random choices from the remaining assortment based on the utility model.

In step 3, we utilize information on observed sales, stockouts, and product attributes to estimate the choice parameters according to our demand model and the four benchmark models, and we compare the results with the true expected demand, given by the values of β_n and β_u that were used to generate the data.

4.2. Benchmark Models

We compare the performance of our model against those of the four benchmark models of increasing sophistication to assess the effect of incorporating stockout information and substitution on demand estimation.

In model 1, we use only those observations that are not censored, i.e., those in which no stockout occurred. This avoids the problem of having to estimate stockout-based substitution. Although this model is simple, it suffers from selection bias. It can underestimate demand because it captures only small realizations of demand. In model 2, we use all observations, but assume that demand equals sales. Such a method is commonly used in practice. It ignores the effect of stockouts on own demand as well as the demand for substitute products. It also underestimates demand because of censoring but does not suffer from selection bias. In model 3, we "uncensor" demand for each book type independently. This model captures the effect of stockouts on own demand but ignores substitution. Note that model 3 has the same form as our model for those observations in our data set in which there is no stockout or only one book type is stocked initially. Finally, model 4 is a full-information benchmark in which the times of stockout occurrences are known. We use model 4 to assess the benefit of knowing the exact time of stockout occurrences for the accuracy of demand estimates.

Comparing our results with models 1 and 2 enables us to assess the effect of stockout information on demand estimation, a comparison with model 3 helps to evaluate the implications of modeling substitution, and a comparison with model 4 helps to assess the value of knowing the time of stockout. Some of these models have been used as benchmarks in the literature: Anupindi et al. (1998), Conlon and Mortimer (2013), and Musalem et al. (2010) compare their results with model 2 to evaluate the effect of stockouts and substitution; in addition, Conlon and Mortimer (2013) use another model similar to model 1. Appendix B in the online supplement presents the likelihood functions for the benchmark models.

4.3. Results

Table 2 shows the mean absolute percentage error (MAPE) between estimated and true expected demand for each model at the different stocking levels. MAPE is calculated as

$$\frac{1}{10{,}000}\sum_{i=1}^{10{,}000}\Big|(\text{estimated expected demand of } i \\ -\text{true expected demand of } i) \\ \cdot(\text{true expected demand of } i)^{-1}\Big|\times 100\%.$$

In addition to Table 2, Figure 3 shows a plot of the mean percentage error (MPE) to compare all the demand models with respect to bias. MPE can, in general, yield different results than MAPE because a model with a high MAPE can give either low or high MPE depending on the extent to which positive and negative errors offset each other. In our setting, we observe from Figure 3 that the absolute values of MPE of all models are close to their corresponding MAPE values. The relative performance of the models is also the same on both performance metrics. Thus, MAPE and MPE yield similar inferences in our setting.

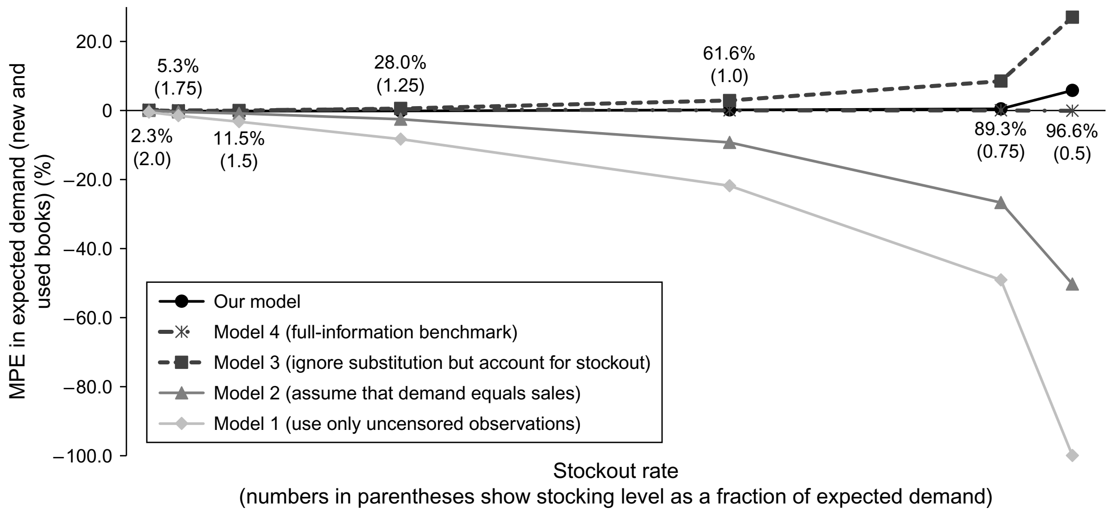
We observe from Table 2 that benchmark model 4 yields the best performance, followed by our model, and models 3, 2, and 1 in that order. Our model yields statistically better performance than models 3, 2, and 1 in every case, except the stocking level of 1.5, for which the MAPE of model 3 is 0.2% whereas that of our model is 0.3%. The relative performance ranking of these models is not surprising

Table 2 Results of the Simulation Experiment: Mean Absolute Percentage Error (MAPE) Between Estimated and True Expected Demand for Each Model at Different Stocking Levels

	0.5 (96.6%)	0.75 (89.3%)	1 (61.6%)	1.25 (28.0%)	1.5 (11.5%)	1.75 (5.3%)	2 (2.3%)
Model 1: Use only uncensored observations	99.9	49.1	21.8	8.3	3.3	1.5	0.6
Model 2: Assume that demand equals sales	50.3	26.7	9.3	2.5	0.9	0.4	0.5
Model 3: Ignore substitution but account for stockout	27.1	8.5	2.9	0.6	0.2	0.4	0.5
Model 4: Full-information benchmark	0.4	0.4	0.4	0.6	0.4	0.4	0.3
Our model	7.4	0.8	0.5	0.3	0.3	0.4	0.3

Note. Stocking level as a fraction of expected demand (figures in parentheses show the average stockout rate at each stocking level) (%).

Figure 3 Mean Percentage Error (MPE) Between Estimated and True Expected Demand for Each Model at Different Stocking Levels



per se, because the models are of decreasing levels of sophistication in the above order and also because the true demand data are generated with stockouts and substitution according to our choice model. Indeed, from Figure 3, we see that model 1, which uses only uncensored observations, underestimates demand. Model 2, which uses all demand observations but equates demand with sales, underestimates demand less severely than model 1. Interestingly, model 3, which ignores substitution but accounts for stockouts, overestimates expected demand. Such a model is often suggested as suitable for single-product models, but our result shows that it is no longer accurate in multiproduct models with substitution. After a product type stocks out, the demand for the other product type consists of two sources: demand from consumers who choose it as their first choice and spillover (substitute) demand from the other product type. By ignoring this spillover effect, model 3 overestimates the expected demand when two product types are stocked initially and at least one product type stocks out.

Although the performance ranking is intuitive, what is surprising and useful is that our model estimates demand accurately even under high stockout rates. For instance, when nearly 90% of the textbooks experience stockout, the MAPE for our model is less than 1%, whereas the MAPE values for models 1, 2, and 3 are 49.1%, 26.7%, and 8.5%, respectively. Thus, incorporating stockouts and substitution has substantial benefits in estimating demand. Moreover, the performance of our model is similar to that of the full-information benchmark model 4 for stocking levels from 0.75 to 2. Only for the stocking level of 0.5 times the mean demand does model 4 outperform our model. This indicates that knowing the time of stockouts only provides a marginal improvement in estimation accuracy. This occurs in our setting because we can utilize data on multiple products and incorporate stockout-based substitution. The knowledge of the time of stockout can be valuable in other settings. Jain et al. (2015) show that knowing the time of stockouts is valuable in a single-product multiperiod Bayesian inventory model with censored demand.

Besides evaluating the five models with respect to their abilities to predict demand, we also evaluate their abilities to estimate the underlying choice parameters. The true value, the estimates, and the MPE for each model for stocking levels 0.75 and 1 are shown in Table 3. We see that, in most instances, the estimates of our model are closer to the true parameter values than those of models 1, 2, and 3. The results for other stocking levels yield similar insights and are omitted for brevity.

In summary, the simulation results show that incorporating stockouts and substitution leads to accurate demand estimates even for a high incidence of stockouts. The magnitude of this benefit is substantial compared to benchmark models. Moreover, the value of knowing the time of stockout is marginal. Next, we analyze the reasons for the superior performance of our model.

4.4. Drivers of Estimation Error

The low estimation error achieved by our model in the presence of high stockouts seemingly contrasts with the body of literature that highlights the negative impact of stockouts on demand estimation. Whereas intuition suggests that a firm should overstock to learn about its demand, our results suggest that overstocking is not necessary. We hypothesize

Table 3 Results of the Simulation Experiment: Estimates of Choice Parameters from Each Method for Stocking Levels of 0.75 and 1

Stocking level as a fraction of expected demand = 0.75 (overall stockout rate = 89.3%)

	New books						Used books					
	Intercept		New-book price		Course level		Intercept		New-book price		Course level	
	Estimate	PE (%)	Estimate	PE (%)	Estimate	PE (%)	Estimate	PE (%)	Estimate	PE (%)	Estimate	PE (%)
Model 1	−2.00	−66.5	−0.43	−116.1	0.67	−4.1	−1.77	−77.3	−0.20	−100.9	0.32	−20.1
Model 2	−1.68	−40.0	−0.18	7.6	0.64	−8.9	−1.47	−47.4	−0.08	19.8	0.33	−18.4
Model 3	−1.04	13.6	−0.24	−18.2	0.81	15.4	−0.89	11.3	−0.11	−13.0	0.52	29.2
Model 4	−1.22	−1.7	−0.18	−11.8	0.69	−1.2	−0.98	−1.9	−0.12	20.2	0.40	1.2
Our model	−1.17	2.4	−0.22	−9.8	0.72	2.6	−1.02	−1.5	−0.08	17.2	0.40	0.3

Stocking level as a fraction of expected demand = 1.0 (over all stockout rate = 61.6%)

	New books						Used books					
	Intercept		New-book price		Course level		Intercept		New-book price		Course level	
	Estimate	PE (%)	Estimate	PE (%)	Estimate	PE (%)	Estimate	PE (%)	Estimate	PE (%)	Estimate	PE (%)
Model 1	−1.59	−32.1	−0.24	−21.2	0.75	7.7	−1.37	−36.7	−0.09	11.0	0.40	0.5
Model 2	−1.34	−12.0	−0.22	−12.2	0.71	2.0	−1.13	−12.9	−0.12	−21.2	0.40	−1.2
Model 3	−1.12	6.3	−0.23	−13.1	0.71	1.6	−0.91	8.9	−0.15	−50.1	0.42	4.8
Model 4	−1.18	1.5	−0.21	7.4	0.71	0.9	−1.00	0.2	−0.10	0.7	0.40	−0.1
Our model	−1.18	1.9	−0.21	−7.3	0.69	−1.2	−0.96	4.4	−0.14	−43.4	0.40	−0.9

Note. PE denotes percentage error.

that our result arises because of using an attribute-based choice model and estimating it on a large data set of dissimilar products with heterogeneous product attributes. An attribute-based model enables us to represent demand parsimoniously as a function of only a few parameters, which can then be estimated accurately by exploiting the heterogeneity of the data. We examine this hypothesis in what follows.

When the stockout rate is high, demand information is more censored, and it is more difficult to restore the demand distribution. For instance, in the statistics literature on estimating the rate parameter of an exponential demand distribution from censored data of a single product type, it is shown that the asymptotic variance increases as the stockout probability increases (Deemer and Votaw 1955, Bartholomew 1957). On this basis, the Bayesian inventory management literature shows in many models (e.g., Harpaz et al. 1982, Lariviere and Porteus 1999, Ding et al. 2002) that it is optimal to stock more, especially in the initial selling periods, in order to learn about demand. Besbes and Muharremoglu (2013) state that this "stock more" result is absent under nonparametric methods for inventory planning. They show that the exploration–exploitation tradeoff depends on the granularity of the demand distribution and that the need for stocking more is considerably lessened if demand has a continuous distribution or if partial censoring information is available. The "stock more" result is also reversed when customers can substitute and both the demand distribution and substitution rate are unknown (Chen and Plambeck 2008).

The above papers are based on a single-product multiperiod setting in which a single observation of sales occurs in each period. In contrast, our setting involves multiple short-life-cycle products stocked in each time period. Thus, we have a large number of observations with varying product attributes. This gives us a different estimation problem. The econometrics literature highlights that a higher degree of heterogeneity of explanatory variables in a data set improves the accuracy of parameter estimates. For instance, in linear regression, the standard error of coefficients' estimates decreases as the variance of explanatory variables increases (Greene 2007). Similarly, we seek to learn about the parameters of an attribute-based choice model. Is it possible to estimate parameters such as price elasticity accurately from data for many products with varying prices even under demand censoring? This is the argument we test. Thus, we investigate whether an increase in the variance of product attributes leads to a decrease in the standard errors of parameters' estimates.

The standard error of parameters' estimates can be computed in closed form for an MNL choice model in the absence of stockouts. Let $\mathcal{K}_i$ denote the initial assortment of ISBN i and P_{ik} denote the probability that a consumer chooses book type k for ISBN i. The log likelihood function, LL, of observing the sales data is given by

$$LL = \sum_i \sum_{k\in \mathbf{K}_i} s_{ik}\boldsymbol{\beta}_k'\mathbf{x}_i - \sum_i \sum_{k\in\mathcal{K}_i} s_{ik}\ln\left(\sum_{l\in\mathcal{K}_i} e^{\boldsymbol{\beta}_l'\mathbf{x}_i}\right) + \sum_i \ln\binom{E_i}{s_{i1}s_{i2}\cdots s_{iK}}.$$

The first and the second derivatives of the log likelihood function are as follows:

$$\partial LL/\partial\boldsymbol{\beta}_{k_1} = \sum_{i:k_1\in\mathcal{K}_i} s_{ik_1}\mathbf{x}_i - \sum_{i:k_1\in\mathcal{K}_i}\sum_{k_1\in\mathcal{K}_i} s_{ik}\left(e^{\boldsymbol{\beta}_{k_1}'\mathbf{x}_i}\mathbf{x}_i \Big/ \sum_{l\in\mathcal{K}_i} e^{\boldsymbol{\beta}_l'\mathbf{x}_i}\right) \quad \text{for } k_1\in\{0,\dots,K\}.$$

$$\begin{aligned}\frac{\partial^2 LL}{\partial\boldsymbol{\beta}_{k_1}\partial\boldsymbol{\beta}_{k_2}'} &= -\sum_{i:k_1,k_2\in\mathcal{K}_i}\sum_{k\in\mathcal{K}_i} s_{ik} \cdot \frac{(\sum_{l\in\mathcal{K}_i} e^{\boldsymbol{\beta}_l'\mathbf{x}_i})\mathbf{x}_i(\partial e^{\boldsymbol{\beta}_{k_1}'\mathbf{x}_i}/\partial\boldsymbol{\beta}_{k_2}') - e^{\boldsymbol{\beta}_{k_1}'\mathbf{x}_i}\mathbf{x}_i e^{\boldsymbol{\beta}_{k_2}'\mathbf{x}_i}\mathbf{x}_i'}{(\sum_{l\in\mathcal{K}_i} e^{\boldsymbol{\beta}_l'\mathbf{x}_i})^2}\\
&= -\sum_{i:k_1,k_2\in\mathcal{K}_i}\left(\sum_{k\in\mathcal{K}_i} s_{ik}\right)\left(\frac{\mathbf{x}_i(\partial e^{\boldsymbol{\beta}_{k_1}'\mathbf{x}_i}/\partial\boldsymbol{\beta}_{k_2}')}{\sum_{l\in\mathcal{K}_i} e^{\boldsymbol{\beta}_l'\mathbf{x}_i}} - \frac{e^{\boldsymbol{\beta}_{k_1}'\mathbf{x}_i}}{\sum_{l\in\mathcal{K}_i} e^{\boldsymbol{\beta}_l'\mathbf{x}_i}}\frac{e^{\boldsymbol{\beta}_{k_2}'\mathbf{x}_i}}{\sum_{l\in\mathcal{K}_i} e^{\boldsymbol{\beta}_l'\mathbf{x}_i}}\mathbf{x}_i\mathbf{x}_i'\right)\\
&= -\sum_{i:k_1,k_2\in\mathcal{K}_i} E_i(\mathbf{1}(k_1=k_2) - P_{ik_1})P_{ik_2}\mathbf{x}_i\mathbf{x}_i' \quad \text{for } k_1,k_2\in\{0,\dots,K\}.\end{aligned}$$

The resulting Hessian matrix, $\mathbf{H}$, has $K\times K$ submatrices, each of dimension $M\times M$. Note that the Hessian is not a function of the sales realizations; i.e., there are no s_{ik} terms in the Hessian. Thus, we have $\mathrm{E}[\mathbf{H}]=\mathbf{H}$. Therefore, the asymptotic variance–covariance matrix of $\hat{\boldsymbol{\beta}}$ can be computed as

$$\operatorname{Var}[\hat{\boldsymbol{\beta}}] = (-E[\mathbf{H}])^{-1} = (-\mathbf{H})^{-1} = \left(-\frac{\partial^2 LL}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}\right)^{-1}.$$

The structure of $\mathbf{H}$ in the presence of stockouts is more complex. Therefore, we conduct a numeric experiment to investigate the impact of heterogeneity in product attributes on the standard errors of the estimated parameters. We use the same setting as in §4.1 but vary the heterogeneity of new-book prices (*NP*) by sampling from a normal distribution with the same mean as the original data set and with five different values of the coefficients of variation: 0.25, 0.5, 0.75, 1, and 1.25. We also use a smaller sample size for computational speed. Thus, we construct five data sets of 1,000 book titles each, varying in the heterogeneity of prices. The coefficient of variation of *NP* in the original data set is 1.02.

When the true value of β is not known, the expected Hessian in the sample, $E[\hat{\mathbf{H}}]$, is used as an estimator of $E[\mathbf{H}]$, where $E[\hat{\mathbf{H}}]$ is evaluated at $\hat{\beta}$ (Train 2009). We follow this approach. Table 4 shows the results of the experiment. The standard error of $\hat{\beta}_{NP}$ decreases in two factors: the coefficient of variation of *NP* and the inventory level.

Table 4 The Impact of Heterogeneity in New-Book Price *NP* on the Estimation Error at Different Stocking Levels

			Estimates						Standard error in estimates					
			New books			Used books			New books			Used books		
Stocking level	Stockout rate (%)	Coefficient of variation of new-book price	Intercept	New-book price	Course level	Intercept	New-book price	Course level	Intercept	New-book price	Course level	Intercept	New-book price	Course level
0.5	97.5	0.25	−3.16	1.78	0.58	−2.20	1.13	0.26	1.098	1.097	0.142	1.057	1.054	0.145
	97.6	0.50	−1.22	−0.19	0.70	−2.18	1.08	0.45	0.509	0.503	0.139	0.553	0.542	0.143
	97.4	0.75	−1.36	−0.05	0.74	−1.46	0.34	0.43	0.331	0.325	0.140	0.346	0.345	0.142
	97.5	1.00	−1.55	0.14	0.73	−1.26	0.16	0.42	0.265	0.259	0.139	0.271	0.269	0.142
	97.7	1.25	−1.49	0.09	0.73	−1.24	0.14	0.42	0.208	0.198	0.148	0.213	0.205	0.141
1	61.9	0.25	−1.14	−0.27	0.73	−0.96	−0.14	0.41	0.301	0.300	0.030	0.273	0.272	0.029
	62.3	0.50	−1.12	−0.28	0.73	−0.92	−0.17	0.42	0.145	0.144	0.030	0.130	0.129	0.029
	62.7	0.75	−1.12	−0.28	0.73	−0.93	−0.17	0.42	0.095	0.093	0.030	0.086	0.084	0.029
	61.5	1.00	−1.12	−0.28	0.73	−0.92	−0.18	0.42	0.074	0.072	0.030	0.068	0.066	0.029
	62.2	1.25	−1.14	−0.27	0.74	−0.99	−0.19	0.40	0.056	0.053	0.030	0.052	0.049	0.029
1.5	13.0	0.25	−1.14	−0.26	0.73	−0.95	−0.15	0.43	0.261	0.261	0.026	0.240	0.239	0.025
	13.4	0.50	−1.21	−0.19	0.56	−1.06	−0.02	0.23	0.123	0.122	0.026	0.114	0.113	0.025
	13.3	0.75	−1.12	−0.28	0.67	−1.02	−0.10	0.45	0.082	0.080	0.026	0.075	0.074	0.025
	13.2	1.00	−1.21	−0.20	0.75	−0.99	−0.12	0.41	0.064	0.063	0.026	0.059	0.058	0.025
	13.1	1.25	−1.18	−0.23	0.74	−1.01	−0.08	0.41	0.050	0.047	0.026	0.046	0.043	0.025
2	3.7	0.25	−1.27	−0.14	0.74	−1.05	−0.09	0.37	0.260	0.259	0.026	0.239	0.238	0.025
	3.4	0.50	−1.20	−0.20	0.72	−1.07	−0.11	0.43	0.122	0.121	0.026	0.113	0.112	0.025
	3.5	0.75	−1.20	−0.20	0.72	−1.06	−0.10	0.41	0.081	0.080	0.026	0.075	0.073	0.025
	3.7	1.00	−1.21	−0.19	0.73	−1.06	−0.09	0.40	0.064	0.062	0.026	0.059	0.057	0.025
	3.4	1.25	−1.21	−0.19	0.72	−1.07	−0.10	0.39	0.049	0.047	0.026	0.046	0.043	0.025

First, the more the heterogeneity in product attributes, the less are the standard errors of the estimates. This relationship is monotone and is observed for each stocking level. For example, in Table 4, at the stocking level of 0.5, the standard error of new books drops from 1.097 to 0.198 as the coefficient of variation of *NP* increases from 0.25 to 1.25. There are also diminishing returns to the increase in variation of *NP*. Increasing the variation of *NP* brings a larger benefit when the coefficient of variation is small. For completeness, note that the stockout rate is affected only marginally when the coefficient of variation of *NP* is changed. In other words, the result is not attributed to change in the stockout rate.

Second, the higher the stocking level, the less are the standard errors of $\hat{\beta}_{NP}$ of new and used books. For instance, Table 4 shows that when the coefficient of variation of *NP* is 1.0, as the stocking level increases from 0.5 and 2, the standard error of $\hat{\beta}_{NP}$ for new books drops from 0.259 to 0.062 and the stockout rate goes down from 97.5% to 3.4%, respectively. However, the magnitude of decrease is small after the stocking level of 1.

These results show that there are two ways to obtain more accurate parameters' estimates: increasing the stocking level and increasing the heterogeneity of product attributes in the data set. For example, we observe the estimated standard error of $\hat{\beta}_{NP}$ to be 1.097 when the stocking level is 0.5 and the underlying coefficient of variation of *NP* is 0.25. We can reduce the estimated standard error of $\hat{\beta}_{NP}$ to approximately 0.26 by either increasing the stocking level to 2 or increasing the coefficient of variation of *NP* to 1.0. Therefore, the impact of increasing the variation in the data set is similar to that of increasing the stocking level. In the results shown in §4.3, we could obtain accurate estimates in spite of the stockout rate of 95% because the coefficient of variation of *NP* was 1.0; i.e., there was a large heterogeneity in our data.

We examine the sensitivity of the above results to the number of titles in the data set by conducting additional numeric experiments with fewer titles. Appendix A.2 of the online supplement shows the results obtained for 100 and 300 titles in the data set as compared to 1,000 titles used in Table 4. We observe that the standard errors of the estimates are higher than those in Table 4. The standard errors of price coefficient obtained from data sets of 100 and 300 titles, respectively, are four times as large as and twice as large as those from 1,000 titles. However, the effects of stocking level and heterogeneity in product attributes are supported by these smaller data sets as well.

Thus, heterogeneity in product attributes is an additional factor, besides stocking more inventory, to improve the ability to learn about demand in the presence of censoring.

5. Application to the Cornell Store

We estimate the demand model using data from the Spring 2007 to Spring 2011 semesters from the Cornell Store and then test it by conducting a controlled pilot experiment in the Spring semester of 2012. Next, we describe the data set, the results of demand estimation, and the details of the field experiment.

Our interaction with the Cornell Store started in October 2009. We met with the manager and the buyers over the next 22 months to collect historical data and evaluate the accuracy of demand estimates. Then we got in-principle agreement for the pilot experiment. The details of the experiment were finalized over the next five months. The experiment was conducted in the spring semester of 2012. After the pilot experiment, we continued to meet with them to explain the results and transfer the findings to the Cornell Store.

5.1. Data Description

We obtained data from the Cornell Store on 31,424 book titles spanning 9 semesters from 2007 to 2011, an average of 3,492 book titles per semester. One table in the data set contains information for each used and new textbook sold at the Cornell Store, such as its stocking level, sales, selling price, purchase cost, salvage value, and whether it stocked out. A second table provides the mapping of book titles to course–instructor combinations, including information on each course such as its course level, department, and estimated enrollment at the time when books are ordered. We merge these tables using the ISBN level as the unit of analysis. For instance, if two instructors teach one course using different books, then their courses will have separate observations in our data set. We exclude 4,675 records because of missing information regarding the stocking level of new or used books. Our final data set has information on 26,749 textbooks. If an ISBN is sold in more than one semester, we treat each semester as a separate textbook since its attributes may have changed. Table 5 provides summary statistics of the variables used in our analysis.

We expect consumer utility to be a function of the following product attributes. These attributes and their effects on demand were identified through discussions with Cornell Store managers.

—*New-book price* (NP_i): The retail price of new book i varies from \$1 to \$360 with a mean (median) of \$39.4 (\$23). The bookstore sets new book prices based on its assessment of competition and norms in the college bookstore industry. According to these norms, the procurement cost and the salvage value of a new book are typically 0.6 and 0.48 times the retail price. The retail price, cost, and salvage value of a used book are also constant fractions of the price of a new book:

Table 5 Summary Statistics of Variables in the Cornell Store Data Set for the Period 2007–2011 (26,749 ISBNs)

Variables		Mean	Median	Standard deviation	Minimum	Maximum
New-book price	*NP*	39.4	23	40.4	1	360
Number of courses	*NC*	1.5	1	1.1	1	22
Average number of books per course	*NB*	8.1	6	8.8	1	86
Proportion required	*PR*	0.9	1	0.3	0	1
New ISBN	*NI*	0.5	1	0.5	0	1
Estimated enrollment	*E*	65.3	34	90.3	2	990
Stocking level of new books	q_n	20.1	10	37.0	0	821
Stocking level of used books	q_u	10.3	3	21.7	0	547
Sales of new books	S_n	13.6	5	30.1	0	617
Sales of used books	S_u	8.2	2	18.6	0	507

0.75, 0.375, and 0.3, respectively. Hence, we do not need to define an additional variable for used book price. We expect consumer utility from both new and used books to be decreasing in new-book price, but consumer utility from a used book to be decreasing at a lower rate. The bookstore does not vary prices during the selling season because it does not want to sell a textbook to different students at different prices within the same semester.

—*Number of courses* (NC_i): On average a book title is used for 1.5 courses. Students may derive a higher utility from books that are used for multiple courses. On the other hand, students may already have books that are used for multiple courses. Therefore, demand may increase or decrease in NC_i.

—*Average number of books per course* (NB_i): This represents the average number of book titles used in the courses that use book title i. If a course uses multiple books, the utility of a specific book among those books might be lower than the utility of a book that is the only book used by a course. Therefore, we expect consumer utility to be decreasing in NB_i. The mean (median) of this variable is 8.1 (6).

—*Proportion required* (PR_i): This represents the fraction of courses for which book title i is a required textbook. The utility of a required book should be higher than that of an optional book. Because a book can be used for multiple courses, it can be required for one course while being optional for another. Therefore, we calculate PR_i as the fraction of courses for which a book is required, weighted by the enrollment of those courses. The mean of PR_i in our data is 0.9.

—*New ISBN* (NI_i): This variable identifies whether book title i will be sold for the first time at the Cornell Store. We expect that the utility of a book will be higher if it is a new ISBN because the book will be a relatively new edition and there will be limited availability of used books in the marketplace (i.e., at competing online stores or in the peer-to-peer market). The value of NI_i is 1 if a book title is sold for the first time and 0 otherwise. The mean of NI_i is 0.5 in our data set.

—*Course level* (CL_i): This variable seeks to identify whether book title i is used for freshmen, sophomore, junior, and senior/graduate courses. It captures the effect of course levels on the utility of a book title. Students in lower-level courses may be more inclined to buy books than those in higher-level courses. A book can be used for multiple courses with different levels. Therefore, we define CL_i as the mode of the course levels for which book i is used. Course levels 1, 2, 3, and 4 account for 21.1%, 18.7%, 21.3%, and 38.8% of the observations, respectively, in our data set. We use three binary variables, $CL1_i$, $CL2_i$, and $CL3_i$, for estimation.

—*Department* (DE_i): This variable captures the impact of the department for which book title i is used. DE_i accounts for heterogeneity across departments. For instance, students from professional schools such as law and management may be more inclined to buy books compared to students from other schools. The value of DE_i is defined as the mode of the departments for which book i is used, weighted by enrollment. Agriculture, architecture, art and science, engineering, hotel administration, human ecology, industrial and labor relations, management, and law departments respectively account for 7.6%, 2.0%, 72.7%, 5.3%, 1.9%, 3.3%, 3.7%, 1.1%, and 2.5% of the ISBNs in our data. In our estimation, we use eight binary variables, omitting law.

Altogether, we have 34 variables, 16 each for new and used textbooks and two constant terms, to capture the attributes that the utility function is based on. We do not include the level of inventory at the time when a customer visits the store as an attribute of the utility function. The recent literature has investigated endogeneity between demand and product availability; see, e.g., Balakrishnan et al. (2004), Cachon et al. (2013), or Stock and Balachander (2005). Musalem et al. (2010) describe a different issue: the inventory level stocked by buyers may be correlated with demand shocks, and since demand is censored by inventory, the parameters' estimates may be biased.

This problem can be solved by constructing an instrument for the correlation of inventory with demand shocks or by introducing additional demand forecasting variables into the model such as the attractiveness of consignment purchase contracts with different book publishers or competitive prices at online retailers. The Cornell Store does not have such data available at present, and our application can be developed further by incorporating these features.

The following additional data are used to model the choice and substitution process for each ISBN.

—*Enrollment* (E_i): This is the sum of estimated enrollment of the courses that use book title i. It varies from 2 to 990 with a mean (median) of 65.3 (34). It represents the number of potential consumers of a book title. We use estimated enrollment because the actual enrollment is known only after the selling season.

—*Stocking level* (q_{in}, q_{iu}): The inventory of new books (used books) stocked by the Cornell Store for a book title has a mean of 20.1 (10.3) with a standard deviation of 37.0 (21.7). The managers at the Cornell Store make stocking decisions using historical data, the availability of new and used books, and their experience regarding students' purchase behavior.

—*Sales data* (s_{in}, s_{iu}): Sales of new books (used books) has a mean of 13.6 (8.2) and a standard deviation of 30.1 (18.6). The sales data are compiled after each selling season.

—*Stockout indicator* (STI_i): This variable identifies the stockout situation based on initial assortment, as defined in Table 1. Overall, there is no stockout in 55.7% (= 21.7 + 29.5 + 4.5) of the cases, whereas at least one product type stocks out in the remaining 44.3% of observations.

5.2. Demand Estimation

In this section, we present the parameters' estimates obtained for the entire pooled data set from 2007 to 2011. It may be noted that the purchase probability of books has changed marginally over this period as a result of online competition and changes in the regulatory environment. We discuss the implications of these changes toward the end of the section.

Table 6 shows the 34 estimated choice parameters obtained from solving the MLE problem (4) for the above data from 2007 to 2011. We find that 29 parameters are statistically significant at $p < 0.01$; 2 parameters, coefficients of *NB* for used books and *Management* for new books, are significant at $p < 0.05$; and 3 parameters for used books are not significant.

The utility of a book decreases as *New-book price* (NP_i) increases as expected. Moreover, the utility of new books is more sensitive to NP_i than that of used books. *Number of courses* (NC_i) has a negative effect on utility. Though books listed for multiple courses may provide higher utility, it is likely that students may already have those books, thus depressing demand. The utility of both new and used textbooks increases in *Proportion required* (PR_i). That is, the utility from a book is higher when it is required for a course than when it is optional. Moreover, this effect is higher for new books than for used books. *New ISBN* (NI_i) also has a positive effect on utility. As *Course level* (CL_i) increases, utility decreases. In particular, freshmen are more likely to purchase books from the Cornell Store than other students. Finally, utility varies across various departments (DE_i).

Table 6 Estimated Parameters of the Choice Model for the Cornell Store Data Set for the Period 2007–2011

		New books		Used books	
Attributes		Estimates	Standard errors	Estimates	Standard errors
Intercept	*Intercept*	-1.903^{***}	0.022	-0.083	0.054
New-book price	*NP*	-0.276^{***}	0.006	-0.157^{***}	0.009
Number of courses	*NC*	-0.137^{***}	0.005	-0.137^{***}	0.005
Average number of books per course	*NB*	-0.017^{***}	0.002	0.005^{**}	0.002
Proportion required	*PR*	1.220^{***}	0.012	0.064^{***}	0.014
New ISBN	*NI*	0.376^{***}	0.012	0.122^{***}	0.018
Course level 1	*CL1*	0.622^{***}	0.016	0.706^{***}	0.020
Course level 2	*CL2*	0.543^{***}	0.013	0.405^{***}	0.018
Course level 3	*CL3*	0.227^{***}	0.015	0.408^{***}	0.021
Course level 4		–		–	
Agriculture	*AGR*	-0.926^{***}	0.022	-0.696^{***}	0.068
Architecture	*ARC*	-1.231^{***}	0.046	-0.900^{***}	0.071
Art and science	*AAS*	-1.220^{***}	0.018	-0.739^{***}	0.061
Engineering	*ENG*	-0.805^{***}	0.031	-0.843^{***}	0.070
Hotel administration	*HAD*	-1.026^{***}	0.035	-0.679^{***}	0.069
Human ecology	*HEC*	-0.673^{***}	0.032	-0.559^{***}	0.064
Industrial and labor relations	*ILR*	-1.134^{***}	0.030	-0.078	0.071
Management	*MGT*	-0.128^{**}	0.056	1.252	5.020
Law		—		—	

$^{**}p < 0.05$; $^{***}p < 0.01$.

An intuitive way to interpret these results is to compare the purchasing probabilities of books in different situations. We calculate average purchasing probabilities over the data set from 2007 to 2011 using the estimated model parameters. The mean purchasing probabilities of new and used books averaged over all observations are 8.5% and 31.9% with standard deviations of 3.3% and 7.3%, respectively, when both types are available. The mean purchasing probability of new books increases to 12.8% when only new books are available, and that of used books increases to 35.0% when only used books are available. The purchasing probability of new books is most sensitive to *Proportion required* (PR_i), and that of used books is less sensitive to this variable. For example, when both types are available, the mean purchasing probabilities of required and optional new books are 9.1% and 2.9%, respectively, and those of required and optional used books are 32.1% and 31.5%, respectively.

Table 7 **Change in Purchase Probability of New and Used Textbooks Over the Period Before and After the Introduction of the Higher Education Opportunity Act**

	2008 spring semester	2009 spring semester	Difference
Average purchase probability of new books when new and used books are available	0.0941	0.0843	-0.0098^{***}
Average purchase probability of used books when new and used books are available	0.3328	0.3042	-0.0286^{***}
Average purchase probability of new textbooks when only new books are available	0.1447	0.1242	-0.0205^{***}
Average purchase probability of used textbooks when only used books are available	0.3678	0.334	-0.0338^{***}

$^{***}p < 0.001$.

Whereas the above estimates are for our entire data set, the purchase probability of books has evolved over this period because of intensifying online competition and because the Higher Education Opportunity Act (HEOA) enacted in 2008 requires colleges to disclose the list of textbooks used for various courses through the Internet before the semester start date. Although the year-to-year impact of these changes is marginal, we observe a significant cumulative impact over time. Table 7 shows the effect of the introduction of the HEOA by comparing the purchase probabilities of new and used books for the semesters before and after the introduction of the act. Observe that all purchase probabilities declined significantly over this period. To address such changes over time, we calibrate our model on data from the previous year, instead of the entire pool of historical data, to plan inventory for a given semester. We find that, although there is a trend over time, the changes in coefficients of the utility function are small in magnitude and statistically insignificant from one year to the next. Moreover, the size of the data set in each year is sufficiently large to get accurate estimates of parameters.

5.3. Field Experiment

We conduct a controlled field experiment to assess the benefit of the incorporating stockout-based substitution in inventory planning and to assess whether stocking decisions based on our MNL choice model perform better than the decisions made by store managers. Since our comparisons are necessarily across products, we use a matched sample of control and test groups similar to the approach adopted in other studies (e.g., Caro and Gallien 2010).

5.3.1. Experiment Design. We create three groups of ISBNs for the experimental study: one control group and two test groups. For the control group (group 1), the Cornell Store managers decide stocking levels using their existing process. For the first test group (group 2), the Cornell store managers are provided with a demand estimation tool. They use demand estimates from our model but make stocking decision themselves. For the last test group (group 3), we provide optimal stocking levels to the managers.

We select 72 book titles matched by product attributes into 24 sets of three each, representing a wide range of values of product attributes in our entire data set. The attributes of these titles are shown in Table 8. The books within a set have the same values of *PR* and *CL*, and they have marginal differences in the values of other attributes. The last two rows of Table 8 highlight how closely the books were matched based on two metrics. The first metric is the proportion of the sets where all three corresponding books have identical attributes. On all attributes other than new-book price (*NP*), over 91% of the sets had identical attributes. The second metric shows the mean and median of the maximum percentage difference in attribute values. The maximum percentage difference for an attribute is calculated as

$$\frac{\max\{value_{l1}, value_{l2}, value_{l3}\} - \min\{value_{l1}, value_{l2}, value_{l3}\}}{\max\{value_{l1}, value_{l2}, value_{l3}\}} \times 100\%,$$

where $value_{lg}$ is the value of the attribute of the book title in group g of set l. The mean of the maximum difference is 7.1% for *NP* and less than 3% for all other attributes. Overall, the matching results show that the three book titles in a set have similar attributes. Each book title in a set is assigned to one of three experimental groups. The 72 book titles comprise 3.4% of the total book titles for Spring 2012 for which the Cornell Store had information in November 2011.

For group 1, the Cornell Store managers used the following process to make stocking decisions. They first set a target total stocking level based on their assessment of demand. Then they purchase used books from students, third-party sellers, and wholesalers based on their assessment of demand for used books. Finally, they make up the difference in the target total stocking level by purchasing new books from publishers. Note that this method does not involve tradeoffs between individual choice probabilities of new and used books and the substitution rate between them. In particular, it does not address the dependence of demand for new books on the number of used books stocked and vice versa.

For each book title in group 2, we provide the bookstore with the mean and the standard deviation of demand for three situations: first, if only new books are stocked; second, if only used books are stocked; and third, if both new and used books are stocked in sufficient quantities so that there is no stockout-based

Table 8 Product Attributes of the Three Matching Groups for the Field Experiment and the Result of the Matching Process

	New-book price (in U.S. dollars)			Number of courses			Average number of books per course			Enrollment			New ISBN					
Group/Set	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	Proportion required	Course level	Department
1	9	10	11	1	1	1	8	8	8	150	150	150	0	0	0	1	1	AAS
2	14.99	15	15	1	1	1	19	19	13	18	18	18	0	0	0	1	1	AAS
3	13.95	14	14	1	1	1	6	6	6	45	45	45	0	0	0	1	2	AAS
4	7.99	8	8	1	1	1	8	8	8	96	96	96	0	0	0	1	2	AAS
5	25	24.95	24.99	1	1	1	16	16	16	50	50	50	0	0	0	1	2	AAS
6	14.95	14.95	15	2	2	2	7	7	7	110	110	110	0	0	0	1	2	AAS
7	14.95	15	16	2	2	2	9	9	9	50	50	50	0	0	0	1	2	AAS
8	13.95	14	15	3	3	3	7	7	7	25	25	25	0	0	0	1	2	AAS
9	15.75	15.95	16	1	1	1	16	16	16	50	50	50	1	0	0	1	2	AAS
10	10.95	10.95	11.95	2	2	2	9	9	9	35	35	35	0	0	0	1	3	AAS
11	5.5	5.5	5.95	2	2	2	11	11	11	25	25	25	0	0	0	1	3	AAS
12	16.95	17.95	18	2	2	2	11	11	11	25	25	25	0	0	0	1	3	AAS
13	7	8	8	2	2	2	12	12	12	30	30	30	0	0	0	1	3	AAS
14	12.95	13	13	2	2	2	12	12	12	30	30	30	0	0	0	1	3	AAS
15	7.95	8	7.95	1	1	1	10	10	10	25	25	25	1	1	1	1	3	AAS
16	6.95	6.95	6.95	3	3	3	9.3	9.3	9.3	25	25	25	1	1	1	1	3	AAS
17	15	15	15	2	2	2	10	10	10	16	16	16	0	0	0	1	4	AAS
18	15	15.99	16.95	6	6	6	10	10	10	17	17	17	0	0	0	1	4	AAS
19	6.95	7.95	7.95	2	2	2	9	9	9	25	25	25	1	1	1	1	4	AAS
20	90	99.95	103	1	1	3	1	1	1	25	25	25	0	0	0	1	4	ENG
21	29.95	33.25	29.99	1	1	1	2	2	2	40	40	45	0	0	0	1	4	HAD
22	15	16	16	1	1	1	7	7	7	80	80	80	1	0	0	0	4	HEC
23	14.95	14.99	15	1	1	1	5	5	5	30	30	30	0	0	0	1	4	ILR
24	24.95	25	26.95	2	2	2	3	3	3	35	35	35	0	0	0	1	4	ILR
Proportion of the sets with identical attributes (%)																		
		8.3			95.8			95.8			95.8			91.7		100.0	100.0	100.0
Mean (median) maximum difference																		
		7.1% (6.8%)			2.8% (0.0%)			1.3% (0.0%)			0.5% (0.0%)			—			No difference	

Note. AAS, art and science; ENG, engineering; HAD, hotel administration; HEC, human ecology; ILR, industrial and labor relations.

substitution. Note that, in a situation with stockout-based substitution, demand becomes a function of inventory. Our method of presenting data allows us to decouple demand estimation from the inventory planning decision and provides only demand estimates to the managers. The managers take this information into account when they make stocking decisions.

For group 3, we use an optimization procedure to recommend stocking levels for each book title. For example, for the first title in group 3, our procedure recommends that 16 new books and 64 used books should be stocked. Altogether, both new and used books are recommended for 14 titles, only new books are recommended for 3 titles, and only used books are recommended for 7 titles. The total stocking levels are shown in Table 9.

The optimization procedure works as follows. For a given book title, we first construct an initial solution using the independent newsvendor heuristic (Mahajan and van Ryzin 2001). This heuristic has a simple structure and is computationally efficient when the number of product types is small. Since we have two types of books, we have three possible initial assortments: new books only, used books only, and new and used books. This gives us three initial candidate solutions for a book title. Searching over the neighborhoods of those solutions, we find the stocking level that yields the highest profit under

Table 9 Results of the Pilot Experiment: Enrollment, Stocking Levels, Sales, and Profits of All Test SKUs

	Group 1	Group 2	Group 3
Estimated enrollment	1,057	1,057	1,062
Stocking level			
New books	289	262	112
Used books	228	280	305
Total	517	542	417
Sales			
New books	161	161	96
Used books	184	201	229
Total	345	362	325
Stockout rate			
New books (%)	8.7	10.5	77.8
Used books (%)	26.3	26.3	38.1
Total profit ($)	1,611.59	1,747.03	1,775.48

stockout-based substitution. Let $q = (q_n, q_u) =$ (initial inventory level of new books, initial inventory level of used books) denote a solution proposed by the heuristic. Then, we conduct a search over the range of points $(q_n \pm C_n, q_u \pm C_u) \cap \Re^{2+}$ using a simulation-based procedure to incorporate stockout-based substitution. The value of C_n is set to be the maximum of $0.2\times$ mean demand and 5 units, where mean demand is the demand for new books if only new books are stocked. The value of C_u is set analogously for used books. These limits are adjusted if the resulting solution is close to a boundary of the range. An important consideration here is that used books are in limited supply. They are procured before new books, so the quantity of new books to be procured depends on the quantity of used books obtained. Thus, if the initial recommended stocking level is not feasible, we undertake the optimization again with a constraint on used book availability.

5.3.2. Results and Implications. In this section, we compare the profit performance of the three groups in our experimental study. We trace their differences in profit to differences in stocking decisions and product mix, and we obtain generalizable insights on the implications of consumer choice and substitution on inventory planning in practice.

Profit Comparison: Table 9 presents the total enrollment, stocking levels, sales, and profits for each group. The total realized profits of groups 1, 2, and 3 are \$1,622, \$1,747, and \$1,776, respectively. Group 3 achieves 10.2% higher profit than group 1, and group 2 obtains 8.4% higher profit than group 1. We test the statistical significance of the difference in profit across the three groups using a sign test, a paired t-test, and the Wilcoxon signed-rank test. The results of these tests are shown in Table 10.

We first compare groups 2 and 3 against group 1. Group 3 yields higher profit than group 1 in 17 of 24 instances (71%), which is significant at $p < 0.05$ in the sign test. Group 2 yields higher profit than group 1 in 16 of 24 instances (67%), which is significant at $p < 0.1$. The paired t-tests indicate that the mean paired difference in profit of groups 2 and 3 over group 1 is significant at $p < 0.1$. The Wilcoxon signed-rank test, which examines the median difference in profits, also provides similar inferences. Thus, we obtain robust evidence that our test groups (groups 2 and 3) yield higher profit than the control group (group 1).

We now compare group 2 with group 3 to determine the incremental value of optimization-based inventory planning over improved demand forecasting. We find that, although group 3 achieves the highest profit, the difference in profit between groups 2 and 3 is not statistically significant in any test. This suggests that improved demand forecasting brings most of the benefit.

Reasons for Differences in Profit: To identify why our model-based inventory planning achieves higher profit, we examine the variation in stocking levels across the three groups. From Table 9, we observe that the total inventory of new and used books across the 24 titles in group 1 is 289 and 228, respectively; the total number of books stocked is 517. The corresponding figures for group 2 are 262 (new), 280 (used), and 542 (total); for group 3, they are 112 (new), 305 (used), and 417 (total). Group 1 stocks the most new books and the fewest used books. This product mix contrasts with the choice model, which predicts a higher purchase probability for used books than new books. Group 2 partially rectifies this discrepancy in product mix. The number of used books increases significantly and the number of new books drops

Table 10 Pairwise Comparison of Groups 1, 2, and 3 with Respect to Differences in Profits, Stocking Levels, Sales, and Leftover Inventory

	Group 3 vs. Group 1	Group 2 vs. Group 1	Group 3 vs. Group 2
Paired difference in profits			
Proportion of positive difference (sign test)	0.71**	0.67*	0.50
Mean paired difference (paired t-test)	6.83*	5.64*	1.19
Median paired difference (Wilcoxon signed-rank test)	3.55**	7.21*	−0.26
Paired difference in stocking levels			
Total: Mean paired difference (paired t-test)	−4.17***	1.04*	−5.21***
New books: Mean paired difference (paired t-test)	−7.37***	−1.13	−6.25***
Used books: Mean paired difference (paired t-test)	3.21*	2.17	1.04
Paired difference in sales			
New books: Mean paired difference (paired t-test)	−2.71*	0.00	−2.71***
Used books: Mean paired difference (paired t-test)	1.88	0.71	1.17
Paired difference in leftover inventory			
New books: Mean paired difference (paired t-test)	1.33***	1.46*	−3.54***
Used books: Mean paired difference (paired t-test)	1.33*	1.46*	−0.13

*, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively, for one-tailed tests.

marginally. The total inventory increases. Group 3 stocks the fewest new books and the most used books, with the net effect that total inventory decreases from 517 to 417. Thus, the improvement in performance from model-based inventory planning can be traced to two drivers: a correction in product mix, and a correction in safety stock.

In discussions with the Cornell Store buyers, we find that the discrepancies in Groups 1 and 2 can be attributed to the processes used to decide inventory levels. Two alternative heuristics are used in practice. In one heuristic, the total demand forecast and order quantity are first determined at the category level, and then the order quantity is allocated across substitutable products within the category. This method is convenient because forecasting demand at the category level is easier than at the SKU level. Decisions in group 1 are due to this method: buyers decide a total quantity, then purchase used books, and finally make up the difference in new books. This method overlooks the dependence of the amount of inventory of one book type on the amount of inventory procured of the other book type. It also ignores asymmetry in substitution rates across book types. Consequently, group 1 differs from the optimal solution in product mix and stocking levels.

A second heuristic treats substitutable products as independent; i.e., it forecasts demand and sets inventory levels for them separately. This heuristic, known as the independent newsvendor heuristic, takes initial choice probabilities into account but ignores stockout-based subsitution (Mahajan and van Ryzin 2001). Group 2, by providing demand forecasts to buyers for new and used books, anchors them on this heuristic. Thus, it sets inventory levels that are too high. In contrast, the stocking decisions for group 3 are based on the optimal solution, which typically yields lower total inventory than the solution from the independent newsvendor approach (Honhon et al. 2010). It also results in lower new-book stocking levels because the substitution rate of used books for new books (13% on average across all book titles) is lower than the substitution rate of new books for used books (36% on average across all book titles).

Having shown the reasons for the differences in profit, we now present supporting statistical evidence. Table 10 shows that the above differences in stocking quantities are statistically significant. The total stocking level of group 3 is lower than those of groups 1 and 2 (significant at $p < 0.01$), group 3 stocks fewer new books than both groups 1 and 2 (significant at $p < 0.01$), and group 3 stocks more used books than group 1 (significant at $p < 0.1$).

The differences in stocking levels affect the stockout rates, sales, and leftover inventory in the three groups. Some of these effects are predictable, but some are not because of substitution. The lower new-book stocking levels of group 3 naturally lead to higher new-book stockout rates for group 3 (77.8%) than those for groups 1 and 2 (8.7% and 10.5%, respectively), lower new book sales, and lower new book leftover inventory. In contrast, the higher used-book stocking levels of group 3 do not result in lower used-book stockout rates. Group 3 has a higher used-book stockout rate (38.1%) than either group 1 or group 2 (both are 26.3%). This is because group 3 experiences more stockout-based substitution of new books for used books compared to groups 1 and 2. Thus, group 3 has higher sales of used books as well as a higher stockout rate of used books than groups 1 and 2. In summary, group 3 achieves a higher total profit than group 1 despite stocking fewer new books, which have higher profit per unit than used books. The volume of sales of used books and substitution of demand of new for used books more than compensate for the decrease in sales of more profitable new books.

6. Conclusions

We show that in a multiproduct demand estimation problem, incorporating stockout-based substitution significantly improves the accuracy of demand estimates even when stockout rates are high and the times of stockout are not observed. The improvement in accuracy in our model is driven by heterogeneity in product attributes and the use of an attribute-based choice model. In contrast, ignoring stockout and substitution leads to underestimation of demand, whereas accounting for stockouts and ignoring substitution leads to overestimation of demand. We use our model in a field experiment and realize over 10% increase in average profits compared to the previous practice. Improvement in demand estimation accounts for most of the increase in profitability because the store buyers are able to achieve a better inventory mix based on demand estimates from our model.

Our project culminated in a spreadsheet-based planning tool for the Cornell Store. For given attributes of books, the tool forecasts demand for new and used books and recommends inventory levels. Its usage has reduced inventory planning errors at the store and enabled buyers to focus on exceptions, i.e., books whose attributes change during the planning period, and on higher value-added problems such as pricing. The insights from the paper are also applicable to over 4,500 college stores in the United States that cater to the needs of more than 20 million students with annual text book sales of over $10 billion in 2010–2011.

Our application has some restrictions that can be examined in future research. For instance, the findings of our paper can be examined in other retail settings with larger product assortments. The modeling of product returns requires further research in the specification of choice models as well as estimating demand from data that are affected by returns. Our model could be extended to capture the potential effect of inventory on demand as well as the dependence of inventory levels on demand shocks. Finally, our model assumes that the prices of new and used books are exogenous and it also does not include competitive prices in the demand estimation model. The reason for this is that competitive prices for the selling season are not known at the time when the bookstore makes its stocking decision. It would be useful to build predictive models of competitive prices in retailing and use them in demand estimation, inventory planning, and pricing decisions.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/msom.2015.0551.

Acknowledgments

The authors are grateful to the Cornell Store for collaborating with them in this research, to Amr Farahat for his suggestions in the early stages of this project, and to Stephen C. Graves (former editor-in-chief of *M&SOM*) and anonymous associate editor and reviewers for their valuable feedback, which led to significant improvements in this paper.

References

Anderson SP, de Palma A, Thisse JF (1992) *Discrete Choice Theory of Product Differentiation* (MIT Press, Cambridge, MA).

Anupindi R, Dada M, Gupta S (1998) Estimation of consumer demand with stockout based substitution: An application to vending machine products. *Marketing Sci.* 17(4):406–423.

Balakrishnan A, Pangburn MS, Stavrulaki E (2004) Stack them high, Let 'em fly: Lot-sizing policies when inventories stimulate demand. *Management Sci.* 50(5):630–644.

Bartholomew DJ (1957) A problem in life testing. *J. Amer. Statist. Assoc.* 52(279):350–355.

Ben-Akiva M, Lerman SR (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand* (MIT Press, Cambridge, MA).

Besbes O, Muharremoglu A (2013) On implications of demand censoring in the newsvendor problem. *Management Sci.* 59(6): 1407–1424.

Cachon G, Gallino S, Olivares M (2013) Does adding inventory increase sales? Evidence of a scarcity effect in U.S. automobile dealerships. Columbia Business School Research Paper 13-60, Columbia University, New York. http://ssrn.com/abstract=2286800.

Caro F, Gallien J (2010) Inventory management of a fast-fashion retail network. *Oper. Res.* 58(2):257–273.

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

Conlon CT, Mortimer JH (2013) Demand estimation under incomplete product availability. *Amer. Econom. J.: Microeconomics* 5(4): 1–30.

Deemer WL, Votaw DF (1955) Estimation of parameters of truncated or censored exponential distributions. *Ann. Math. Statist.* 26(3):498–504.

Ding X, Puterman ML, Bisi A (2002) The censored newsvendor and the optimal acquisition of information. *Oper. Res.* 50(3):517–527.

Fisher ML, Vaidyanathan R (2014) A demand estimation procedure for retail assortment optimization with results from implementations. *Management Sci.* 60(10):2401–2415.

Greene WH (2007) *Econometric Analysis* (Prentice Hall, Upper Saddle River, NJ).

Guadagni P, Little J (1983) A logit model of brand choice calibrated on scanner data. *Marketing Sci.* 2(3):203–238.

Harpaz G, Lee WY, Winkler RL (1982) Learning, experimentation, and the optimal output decisions of a competitive firm. *Management Sci.* 28(6):589–603.

Honhon D, Gaur V, Seshadri S (2010) Assortment planning and inventory decisions under stockout-based substitution. *Oper. Res.* 58(5):1364–1379.

Jain A, Rudi N, Wang T (2015) Demand estimation and ordering under censoring: Stock-out timing is (almost) all you need. *Oper. Res.* 63(1):134–150.

Kök AG, Fisher ML (2007) Demand estimation and assortment optimization under substitution: Methodology and application. *Oper. Res.* 55(6):1001–1021.

Kök AG, Fisher ML, Vaidyanathan R (2015) Retail assortment planning: Review of literature and industry practice. Agrawal N, Smith SA, eds. *Retail Supply Chain Management* (Springer, New York), 175–236.

Lariviere MA, Porteus EL (1999) Stalking information: Bayesian inventory management with unobserved lost sales. *Management Sci.* 45(3):346–363.

Mahajan S, van Ryzin G (2001) Stocking retail assortments under dynamic consumer substitution. *Oper. Res.* 49(3):334–351.

Musalem A, Olivares M, Bradlow ET, Terwiesch C, Corsten D (2010) Structural estimation of the effect of out-of-stocks. *Management Sci.* 56(7):1180–1197.

Nagarajan M, Rajagopalan S (2008) Inventory models for substitutable products: Optimal policies and heuristics. *Management Sci.* 54(8):1453–1466.

Newman JP, Garrow L, Ferguson M, Jacobs TL, Purnomo H (2014) Estimation of choice-based models using sales data from a single firm. *Manufacturing Service Oper. Management* 16(2):184–197.

Smith SA, Agrawal N (2000) Management of multi-item retail inventory system with demand substitution. *Oper. Res.* 48(1): 50–64.

Stock A, Balachander S (2005) The making of a "hot product": A signaling explanation of marketers' scarcity strategy. *Management Sci.* 51(8):1181–1192.

Talluri K, van Ryzin G (2004) Revenue management under a general discrete choice model of consumer behavior. *Management Sci.* 50(1):15–33.

Train KE (2009) *Discrete Choice Methods with Simulation* (Cambridge University Press, New York).

van Ryzin G, Mahajan S (1999) On the relationship between inventory costs and variety benefits in retail assortments. *Management Sci.* 45(11):1496–1509.

van Ryzin G, Vulcano G (2015) A market discovery algorithm to estimate a general class of nonparametric choice models. *Management Sci.* 61(2):281–300.

Vulcano G, van Ryzin G, Chaar W (2010) Choice-based revenue management: An empirical study of estimation and optimization. *Manufacturing Service Oper. Management* 12(3):371–392.

Vulcano G, van Ryzin G, Ratliff R (2012) Estimating primary demand for substitutable products from sales transaction data. *Oper. Res.* 60(2):313–334.