

# Learning Consumer Tastes Through Dynamic Assortments

**Citation for published version (APA):**

Ulu, C., Honhon, D. B. L. P., & Alptekinoglu, A. (2012). Learning Consumer Tastes Through Dynamic Assortments. *Operations Research*, 60(4), 833-849. <https://doi.org/10.1287/opre.1120.1067>

**DOI:**

[10.1287/opre.1120.1067](https://doi.org/10.1287/opre.1120.1067)

**Document status and date:**

Published: 01/01/2012

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

# Learning Consumer Tastes Through Dynamic Assortments

Canan Ulu

McCombs School of Business, The University of Texas at Austin, Austin, Texas 78712,  
canan.ulu@mcombs.utexas.edu

Dorothee Honhon

School of Industrial Engineering, Eindhoven University of Technology, 5612 AZ Eindhoven, The Netherlands,  
d.b.l.p.honhon@tue.nl

Aydin Alptekinoglu

Cox School of Business, Southern Methodist University, Dallas, Texas 75275,  
aalp@cox.smu.edu

How should a firm modify its product assortment over time when learning about consumer tastes? In this paper, we study dynamic assortment decisions in a horizontally differentiated product category for which consumers' diverse tastes can be represented as locations on a Hotelling line. We presume that the firm knows all possible consumer locations, comprising a finite set, but does not know their probability distribution. We model this problem as a discrete-time dynamic program; each period, the firm chooses an assortment and sets prices to maximize the total expected profit over a finite horizon, given its subjective beliefs over consumer tastes. The consumers then choose a product from the assortment that maximizes their own utility. The firm observes sales, which provide censored information on consumer tastes, and it updates beliefs in a Bayesian fashion. There is a recurring trade-off between the immediate profits from sales in the current period (exploitation) and the informational gains to be exploited in all future periods (exploration). We show that one can (partially) order assortments based on their information content and that in any given period the optimal assortment cannot be less informative than the myopically optimal assortment. This result is akin to the well-known "stock more" result in censored newsvendor problems with the newsvendor learning about demand through sales when lost sales are not observable. We demonstrate that it can be optimal for the firm to alternate between exploration and exploitation, and even offer assortments that lead to losses in the current period in order to gain information on consumer tastes. We also develop a Bayesian conjugate model that reduces the state space of the dynamic program and study value of learning using this conjugate model.

*Subject classifications:* product assortment; product variety; dynamic programming; Bayesian learning.

*Area of review:* Manufacturing, Service, and Supply Chain Operations.

*History:* Received August 2010; revisions received April 2011, October 2011; accepted February 2012. Published online in *Articles in Advance* August 2, 2012.

## 1. Introduction

Consumers have different tastes, and retailers offering a selection of products to satisfy demand can benefit from information on what consumers prefer. Which colors of a shirt should a retailer carry? Should a grocery store carry only skim and whole milk or should they also carry 1% and 2% milk? Information on consumer preferences helps with anticipating consumers' substitution behavior, forecasting item-level demand, and ultimately making better product assortment decisions.

There are multiple sources of information a firm can utilize to understand consumer preferences better, such as market surveys and (for established products) past sales data. However, this information may not be readily available or may be too costly for new product categories; for breakthrough or innovative products the firm may know little about consumer preferences. In this paper, we study assortment-planning decisions of a retailer who learns

about consumer preferences by experimenting with different assortments and observing sales. Although a retailer carrying established products may not need to experiment with assortments, for products with little information on consumer preferences, the firm may want to change its assortment dynamically to gather better information.

We consider a horizontally differentiated product category; products differ with respect to only one attribute for which consumers have different first choices provided that they cost the same. Products that differ in their fat content, sweetness, spice level, or color are examples. We use a locational choice model à la Hotelling with a discrete set of consumer locations to represent consumer tastes, and a Bayesian framework to model the information-gathering process and the updating of the firm's beliefs on consumer locations. The firm knows the set of all possible consumer locations on the attribute space, starts with a prior distribution over these locations, and after observing sales each period, updates this distribution using Bayes' rule.

This problem exhibits the classic trade-off between exploration and exploitation: offering more products in the assortment allows the firm to gather better information about consumer tastes because fewer consumers would need to substitute away from their ideal products, however, this may result in lower profits in the short term. On the other hand, offering an assortment that provides the maximum profit in the short term may hinder the information-gathering process and result in less profits in the long run. Beyond this classic trade-off, our demand model allows us to study the information content of a given assortment. In a Hotelling model, consumers make choices to maximize their own utilities, and substitution behavior emerges from these optimal choices. Thus, we can study how a given assortment induces substitution and highlight how assortment choice impacts the quality of information a firm can obtain through sales. Because a given assortment may force consumers to substitute, there may be some censoring of consumer taste information if the firm observes sales. Previous models of learning via dynamic assortments used other consumer choice models (e.g., multinomial logit) that do not allow one to analyze the link between product positioning and substitution behavior explicitly.

First, we study the single-period problem where no learning takes place. The firm determines the optimal assortment and prices based on a known, discrete distribution of consumer tastes. We show that the firm only has to solve for the optimal market segments, each one corresponding to a product, because optimal product locations and prices can be easily obtained as a function of these market segments.

Second, we consider the finite-horizon problem with learning under two scenarios to study how censored information on consumer tastes affect assortment planning: (i) the *uncensored* information case, in which the firm is able to observe consumer tastes postpurchase, that is, the firm comes to know the most preferred product for each consumer; and (ii) the *censored* information case, in which the firm is only able to observe sales of each product. The firm is also uncertain about *market size*, i.e., how many consumers visit the store each period, and we initially assume that the firm observes the realization of market size at the end of each period. We then extend our analysis to unobservable market size.

In the uncensored information case, the optimal assortment in a given period is the one that maximizes the immediate profit; in other words, the firm should adopt a myopic policy. In contrast, the optimal assortment in the censored case may be such that it is optimal to sacrifice some immediate profit in order to gather better information about consumer preferences.

The main contribution of this paper is a result that orders assortments based on their information content and shows that the optimal assortment under censored information cannot be less informative than the optimal assortment under uncensored information (the myopically optimal assortment). We believe this is a novel result on

the structure of the optimal policy in product assortment decisions and is of both theoretical and practical interest. Theoretically, it is similar in nature to the “stock more” result in censored newsvendor problems (see, for example, Harpaz et al. 1982, Ding et al. 2002, Lu et al. 2005). If the newsvendor observes only sales and not demand, sales observations provide censored information on demand: depleted inventory only signals that demand must have been higher than the order quantity. In these problems, typically it is optimal to stock more than the myopically optimal order quantity because higher order quantities provide better information about demand. Our paper establishes an analogous result in product assortment problems. We discuss the practical relevance of this result in §4.3; it can decrease the computational effort significantly and can be used in algorithms that search for a good dynamic assortment policy. In §6, we show that our key result is robust in that a generalized version of our result holds when we relax our main assumptions.

Research on assortment planning has advanced rapidly in recent years, particularly on modeling substitution between products using consumer choice theory from the economics and marketing literatures. Here, we focus on papers that consider dynamic assortments. Kök et al. (2009) provide an excellent review of the literature on static assortment planning.

A firm might change assortments over time in response to changes in consumer preferences or in response to fluctuating inventory levels. Caldentey and Caro (2010) model changing consumer preferences as a stochastic process that a retailer tries to follow and use a stylized representation of the retailer’s assortment as a combination of risky assets (fashionable items) and risk-free assets (basic items). Their objective is to maximize the long-term value of dynamically adjusting the menu of products on display. Caro and Martinez-de-Albeniz (2012) propose a satiation-based model of variety-seeking consumers buying from multiple competing retailers; the model implies that consumers spend a higher share of their budget in retailers that refresh assortments at a faster pace. Using this novel choice model, they derive insights on how often retailers should change assortments in a competitive equilibrium. Bernstein et al. (2011) study how assortments should be customized depending on the profile of remaining inventories.

Some dynamic assortment papers explore how a firm might learn consumer preferences. Unlike our work, these papers either take an aggregate view of consumer choice or do not explicitly model the attribute space. Caro and Gallien (2007) use a finite-horizon multiarmed bandit model with several plays per stage. Applying a number of approximations, they obtain a closed-form dynamic index policy that captures the key exploration versus exploitation trade-off, and develop suboptimality bounds. Saure and Zeevi (2011) study a family of assortment-planning problems in which a general random utility model drives consumer choice. They develop dynamic policies that

balance the trade-off between exploration and exploitation and prove that these policies satisfy some performance bounds. Rusmevichientong et al. (2010) construct an algorithm for optimizing dynamic assortments under a multinomial logit choice framework, and explore demand learning issues. The last two papers take an adaptive learning approach, whereas Caro and Gallien (2007) take a Bayesian learning approach, as our paper does. Our results apply to any prior distribution and are not limited to a specific conjugate model as in Caro and Gallien (2007).

Our paper contributes to the literature on assortment planning in three ways. First, we demonstrate how the information on consumer taste that a firm can gather through sales can be censored due to substitution behavior and how the firm can control the quality of information it can obtain through dynamic assortments. This insight is novel mainly because previous models of learning via dynamic assortments do not model the connection between assortments and quality of information obtained from sales directly. Second, we generate new insights on the structure of optimal assortments and on the exploration-exploitation trade-off. Whereas previous papers develop good policies, we obtain the *optimal* policy; our key result compares the information content of the *optimal* assortment and the myopically optimal assortment. Furthermore, although one intuitively expects the optimal policy to explore first and exploit later, we show that the optimal policy is usually more complicated in that it may be optimal to alternate between periods of exploration (offering an assortment that is more informative than the myopically optimal assortment) and periods of exploitation (using the myopically optimal assortment). We also study the *value of learning* by comparing the expected profit obtained under various assumptions regarding the learning process (i.e., uncensored, censored, myopic, and no-learning). We distinguish between *passive*, *active*, and *proactive* learning, study their relative importance, and show that all three values are high when the time horizon is long, and when the firm is uncertain and/or has little information about consumer tastes. Third, our paper also contributes to the literature by highlighting similarities between Bayesian learning models in newsvendor and product assortment problems. In censored newsvendor models sales observations provide censored information on demand. We analyze a different form of censoring, i.e., the censoring of consumer tastes, which occurs when the assortment offered by the firm induces consumers to purchase a product other than their ideal choice. Using a micromodel of consumer choice behavior, we pose a novel censoring problem, which could not be captured by existing dynamic assortment models in the literature.

The rest of the paper is organized as follows: we present in §2 the locational consumer choice model, and in §3 our results in a single-period setting with no learning. We introduce our multiperiod learning model and its dynamic programming (DP) formulation in §4, and

develop our main results. In §5, we present a numerical study, and in §6 we relax some assumptions of our base model. We discuss directions for further research in §7. Throughout the paper, we use simple examples to explain our definitions and results. The electronic companion to this paper is available as part of the online version at <http://dx.doi.org/10.1287/opre.1120.1067>.

## 2. Consumer Choice Model

We consider consumers making a choice among a set of horizontally differentiated products. Let  $\Omega \subseteq \mathbb{R}$ , a convex set of real numbers, represent the attribute space. Each product is characterized by a single taste attribute, an element of  $\Omega$ , which we refer to as product's *location*. We assume that this attribute can be measured on a cardinal scale such as fat content for dairy products.

Consumers choose from  $n$  products with locations  $x_i \in \Omega$  and prices  $p_i$ ,  $i = 1, \dots, n$ . The consumer population is heterogeneous in the sense that each consumer has an ideal product they would purchase if price was not an issue. The location of a consumer's ideal product, which we refer to as consumer's *location* or *taste*, belongs to a discrete set  $Y = \{y_1, y_2, \dots, y_N\}$ , where  $y_j \in \Omega$ , and  $N$  is the number of consumer types ( $N < \infty$ ). We label consumer locations such that  $y_1 < y_2 < \dots < y_N$  and at times refer to consumer location  $y_j$  simply as location  $j$ .

A consumer located at  $y_j$  receives utility  $U(y_j, x_i, p_i) = \bar{p} - p_i - d|y_j - x_i|$  from a product located at  $x_i$  and priced at  $p_i$ . Here,  $\bar{p}$  is the consumer's willingness-to-pay for his ideal product, which we refer to as the *reservation price*, and  $d$  is a disutility that the consumer incurs per unit of distance between his location and the product's location, which we refer to as the *transportation cost*. We assume that the reservation price and transportation cost are the same for all consumers (we relax the former assumption in §6).

Consumers have an outside option with a fixed utility that we set to zero without loss of generality. Each consumer buys the product that gives him the highest utility, or opts out and obtains zero utility from not buying any of the  $n$  products. That is, a consumer located at  $y_j$  buys product  $i$  if  $i = \arg \max_{k \in \{1, 2, \dots, n\}} \{\bar{p} - p_k - d|y_j - x_k|\}$ , and  $\bar{p} \geq p_i + d|y_j - x_i|$ . If such a product does not exist (all  $n$  products result in negative utility), the consumer chooses the outside option.

This locational consumer choice model dates back to Lancaster (1966), who has extended the work of Hotelling (1929) on spatial competition. Lancaster's model involves a uniform density of consumers in a continuous attribute space. Gaur and Honhon (2006) and Alptekinoglu and Corbett (2010) use the model with a continuous and nonuniform distribution of consumer tastes in operations management settings. In contrast, we take a finitely countable set of discrete points in the attribute space as consumer locations, which are known to the firm, and assume a general discrete probability distribution (possibly nonuniform) over it.

### 3. Single-Period Problem

In this section, we study a firm that makes product location and pricing decisions to maximize expected profit in a single period, knowing the probability distribution of consumer tastes.

Suppose the firm offers  $n$  products with locations  $\mathbf{x} \equiv (x_1, x_2, \dots, x_n)$  with  $x_i \in \Omega$ , and prices  $\mathbf{p} \equiv (p_1, p_2, \dots, p_n)$ . Let the products be numbered such that  $x_1 < x_2 < \dots < x_n$ . We define the *market segment* captured by product  $i \in \{1, 2, \dots, n\}$ , denoted by  $a_i \subseteq \Omega$ , as a subset of the attribute space such that all consumers located within  $a_i$  choose to purchase product  $i$ , over all the other products and the outside option, in accordance with the choice behavior described in §2. Clearly,  $a_i$  depend on  $\mathbf{x}$  and  $\mathbf{p}$  through consumer choice; we study this dependence in Lemma 1.

Let  $\theta_j$  be the probability that a randomly chosen consumer is at location  $y_j$ . A random number of consumers,  $\tilde{m}$ , visit the firm's store during the period (we use tilde to distinguish random variables from their realized values). We refer to  $\tilde{m}$  as the *market size*, a discrete random variable independent of consumer locations, and assume that the firm knows its expected value,  $\mu \equiv \mathbb{E}[\tilde{m}]$ . One can also interpret  $\theta_j$  as the proportion of consumers in the population who are located at  $y_j$ , and  $\tilde{m}$  as the size of a randomly selected sample from this population. Thus, the vector of probabilities  $\boldsymbol{\theta} \equiv (\theta_1, \theta_2, \dots, \theta_N)$  with  $\sum_{j=1}^N \theta_j = 1$  specifies the distribution of consumer locations, and the firm knows these probabilities.

The timing of events within the period is as follows. First, the firm determines the locations  $\mathbf{x}$  and prices  $\mathbf{p}$  of products to offer, based on the expected market size  $\mu$  and the vector of probabilities  $\boldsymbol{\theta}$ . Subsequently,  $\tilde{m}$  consumers visit the store and make purchase decisions based on  $\mathbf{x}$  and  $\mathbf{p}$  as described in §2. The firm is not able to price-discriminate based on consumer tastes because it is not able to observe them at the time of product location and pricing decisions.

The demand for product  $i$ ,  $\tilde{D}_i$  depends on its market segment, and it is a random variable with expected value  $\mathbb{E}_{\boldsymbol{\theta}}[\tilde{D}_i(a_i(\mathbf{x}, \mathbf{p}))] = \mu \sum_{j: y_j \in a_i} \theta_j$ . A consumer purchases product  $i$  only if his location falls in  $a_i$ ; this happens with probability  $\sum_{j: y_j \in a_i} \theta_j$ . We emphasize that the expectation above depends on  $\boldsymbol{\theta}$  by writing  $\mathbb{E}_{\boldsymbol{\theta}}[\cdot]$ .

The firm incurs a fixed cost,  $f > 0$ , for each product offered, and a variable cost of production,  $c > 0$ , per unit of product sold. Sales equal demand because we assume away supply-demand mismatches. We assume throughout that the market is profitable,  $c < \bar{p}$ , otherwise the firm would prefer to offer no products. We express the firm's expected profit as follows:

$$r(\mathbf{x}, \mathbf{p}, \boldsymbol{\theta}) = -fn + \sum_{i=1}^n (p_i - c) \mathbb{E}_{\boldsymbol{\theta}}[\tilde{D}_i(a_i(\mathbf{x}, \mathbf{p}))]. \quad (1)$$

We first study how product assortment decisions and market segments that result from consumer choice behavior are related.

LEMMA 1. *Optimal product locations and prices,  $\mathbf{x}$  and  $\mathbf{p}$ , must satisfy the following properties:*

(a) *Market segments that result from  $\mathbf{x}$  and  $\mathbf{p}$  are mutually disjoint and closed intervals in  $\Omega$ , and their boundary points belong to the set of consumer locations, i.e.,  $a_i = [\underline{x}_i, \bar{x}_i] \subseteq \Omega$  with  $\underline{x}_i, \bar{x}_i \in Y$  such that  $\underline{x}_i \leq \bar{x}_i$  for  $i = 1, \dots, n$  and  $\bar{x}_i < \underline{x}_{i+1}$  for  $i = 1, \dots, n - 1$ .*

(b) *Products are located and priced such that consumers at the boundary points of market segments obtain zero net utility from purchase, i.e.,  $x_i = (\underline{x}_i + \bar{x}_i)/2$  and  $p_i = \bar{p} - d(\bar{x}_i - \underline{x}_i)/2$ , for  $i = 1, \dots, n$ .*

The most important implication of Lemma 1 is that the single-period problem reduces to that of setting boundary points for market segments. That is, optimizing product locations  $\mathbf{x}$  and prices  $\mathbf{p}$  is equivalent to optimizing  $\mathbf{a} \equiv (a_1, \dots, a_n) = ([\underline{x}_1, \bar{x}_1], \dots, [\underline{x}_n, \bar{x}_n])$ , which we refer to as the *assortment*. Moreover, finding the optimal assortment is a matter of choosing boundary points for  $n$  disjoint intervals from the set of consumer locations  $Y$ . Once the firm chooses boundary points of market segments optimally, optimal product locations and optimal prices are also effectively set.

The number of products to offer,  $n$ , is itself a decision, and it is possible that the firm may not want to cover the entire market. We say that consumer location  $y_j$  is *covered* if there exists a product  $i$  such that  $y_j \in a_i$ ; otherwise, we say it is *not covered*. Because the optimal market segments are disjoint, each consumer location is covered by at most one product.

Letting  $\mathcal{A}$  be the set of all possible assortments that satisfy the properties shown in Lemma 1, we rewrite the firm's single-period assortment-planning problem as follows:

$$\max_{\mathbf{a} \in \mathcal{A}} r(\mathbf{a}, \boldsymbol{\theta}) = -fn + \sum_{i=1}^n (p_i(a_i) - c) \mathbb{E}_{\boldsymbol{\theta}}[\tilde{D}_i(a_i)],$$

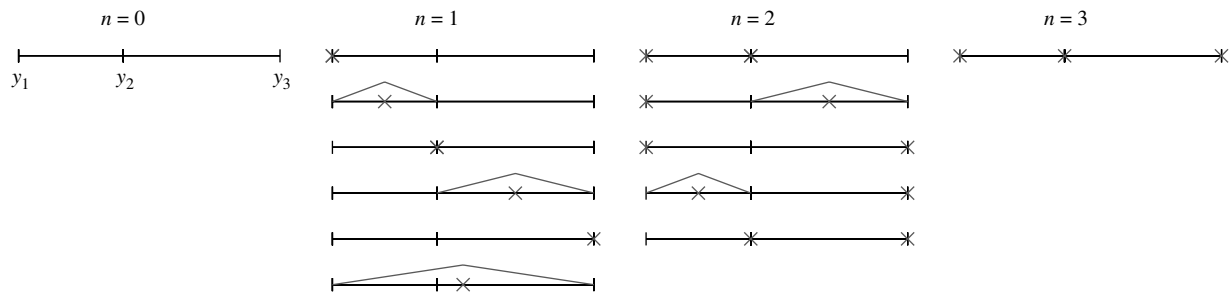
where  $p_i(a_i) = \bar{p} - d((\bar{x}_i - \underline{x}_i)/2)$  and

$$\mathbb{E}_{\boldsymbol{\theta}}[\tilde{D}_i(a_i)] = \mu \sum_{j: y_j \in a_i} \theta_j.$$

The number of possible assortments grows rapidly: for  $N = 2, 3$ , and 4 consumer locations, there are 5, 13, and 34 possible assortments (including the empty one), respectively. Figure 1 shows all possible assortments when  $N = 3$ . Example 1 elaborates on the second assortment in  $n = 2$  column.

EXAMPLE 1. Suppose there are three consumer locations and the firm chooses the market segments  $a_1 = [y_1, y_1]$  and  $a_2 = [y_2, y_3]$ . This is a two-product assortment: product 1 is located at  $x_1 = y_1$  and priced at  $p_1 = \bar{p}$ , whereas product 2 is located at  $x_2 = (y_2 + y_3)/2$  and priced at  $p_2 = \bar{p} - d(y_3 - y_2)/2$ . All consumer locations are covered, and the firm gains an expected profit of  $-2f + (p_1 - c)\mu\theta_1 + (p_2 - c)\mu(\theta_2 + \theta_3)$ .

**Figure 1.** All possible assortments for  $N = 3$  ( $\times$ 's denote product locations; triangles denote market segments).



Solving the above problem for the optimal single-period assortment is equivalent to solving a shortest-path problem of complexity  $O(N^2)$ ; see Alptekinoglu et al. (2012) for further details. Chen et al. (1998) and Alptekinoglu and Corbett (2010) use a similar solution method. Unlike our model, both papers use a continuous attribute space for products and consumers, and neither paper accounts for the possibility of not covering some portions of the market.

#### 4. Multiperiod Model with Learning

Firms do not necessarily know the distribution of consumer tastes for many products, especially for new ones. A firm that is highly uncertain about the distribution of consumer tastes can use product assortments as information-gathering instruments. Such an undertaking must optimally trade off information gathered on consumer tastes by observing sales of an assortment (which can be exploited in the long run) with the profits earned in the short term. We study this trade-off, commonly known as the exploration-exploitation trade-off.

In particular, we consider a firm that makes product location and pricing decisions every period to maximize total discounted expected profits over a finite planning horizon. The timing of events is as follows. At the start of each period, the firm determines the product locations  $\mathbf{x}$  and prices  $\mathbf{p}$ , paying a fixed cost  $f$  for each product offered. In practice,  $f$  may correspond to shelf space and advertising related costs recurring every period for each product. Subsequently, consumers visit the firm's store and make purchase decisions in accordance with §2. We assume that the market size is uncertain, but that the firm knows its distribution; we let  $P(\tilde{m} = m)$  be the probability that the market size in any given period is  $m$  for  $m = 0, 1, \dots, M$  and  $M \leq \infty$ . The firm is also uncertain about consumer tastes, in the sense that although the firm knows the (finite) set of possible consumer locations, it does not know the relative proportions of these locations in the consumer population (we discuss the possibility of unknown consumer locations in §6). We represent this uncertainty by a continuous random variable  $\tilde{\theta} \equiv (\tilde{\theta}_1, \tilde{\theta}_2, \dots, \tilde{\theta}_N)$  with  $\theta_j$  denoting the probability that a randomly chosen consumer is at location  $y_j \in Y$ , or, equivalently, the unknown proportion of consumers from location  $y_j \in Y$  in the consumer population

(in the latter interpretation, each consumer from the population is equally likely to visit the store in a given period). Hence, there are two independent sources of uncertainty: the consumer taste distribution  $\tilde{\theta}$  and the market size  $\tilde{m}$ . In each period the firm observes how many consumers visit the store (including those who choose not to make a purchase) and the purchase decision of each consumer. At the end of the period, the firm uses this information to update its beliefs about the distribution of consumer tastes using Bayes' rule. The updated beliefs are then used in the next period to decide on the product assortment. As in the single-period model, the firm cannot price-discriminate based on consumer tastes because it is not able to observe them at the time of product assortment decisions. Moreover, we assume that the market size is independent of the chosen assortment, which allows us to focus on the quality of information obtained through different assortments.

The assumption of observable market size is realistic for online sales because most companies now have the technology to count the total number of consumers who visit their sites. Even brick-and-mortar stores have a very good understanding, if not a reasonably accurate estimate, of their store traffic (at store entrances one frequently sees greeters with counters in their hands). Nowadays, some retailers even make use of digital imaging to count and track the behavior of customers in various aisles of the store (see Olivares et al. 2011). In §6 we relax the assumption that the market size is observable. We present the case of observable market size first in order to study the learning of consumer tastes in isolation, free of market-size effects. In the unobservable case, the market-size uncertainty also affects the inference on consumer tastes, further complicating the learning process.

We first specify how the firm updates its beliefs on consumer tastes in a generic period.

##### 4.1. Updating Beliefs on Consumer Tastes

Suppose the firm initially has a prior distribution  $\pi(\theta)$  on  $\theta$ . We distinguish two cases: (i) uncensored information and (ii) censored information. The uncensored information case is a theoretical benchmark in which the firm is able to observe all consumer locations. The censored information case is a realistic setting in which the firm

observes sales data only (no consumer locations). Our primary purpose in introducing the uncensored information case is to understand the effect of censoring, which results from consumers' substitution behavior, on dynamic assortment decisions.

*Uncensored information.* Suppose the firm observes market size and consumer tastes at the end of each period, that is, the firm is able to observe the location of each consumer who comes to the store, including those who do not purchase a product. Let  $T_j$  be the number of consumers at location  $j$  and let  $\mathbf{T} = (T_1, \dots, T_N)$  be the consumer tastes vector. We use superscript  $u$  to denote the uncensored information case. Given  $\theta$  and a market-size realization  $m$ , the conditional distribution of consumer tastes (or the likelihood function),  $L^u(\mathbf{T} | \theta, m)$ , is a multinomial distribution

$$L^u(\mathbf{T} | \theta, m) = \frac{m!}{T_1! T_2! \dots T_N!} \theta_1^{T_1} \dots \theta_N^{T_N} \quad (2)$$

defined for all  $\mathbf{T}$  such that  $\sum_{j=1}^N T_j = m$ . If the firm starts with a prior  $\pi$  and observes tastes  $\mathbf{T} = (T_1, \dots, T_N)$ , the *uncensored* posterior distribution is  $\Pi^u(\theta; \pi, \mathbf{T}, m) = L^u(\mathbf{T} | \theta, m) \pi(\theta) / f^u(\mathbf{T}; \pi, m)$ . Here,  $\Pi^u$  is the Bayesian updating operator with *uncensored* taste information and  $f^u(\mathbf{T}; \pi, m) = \int_{\theta} L^u(\mathbf{T} | \theta, m) \pi(\theta) d\theta$  is the probability mass function for consumer tastes conditional on  $m$ .

*Censored information.* In this case, the firm observes sales information for the assortment it offers in addition to the market size at the end of each period. Because the assortment can be such that consumers from multiple locations buy the same product, the taste information may be censored. Let  $\mathbf{D} = (D_1, \dots, D_n)$  denote the observed sales vector for assortment  $\mathbf{a}$  with  $n \leq N$  products. We use the superscript  $c$  to denote the censored information case. The conditional distribution of sales,  $L^c(\mathbf{D} | \theta, m, \mathbf{a})$ , depends on the assortment  $\mathbf{a}$ , and is given by

$$L^c(\mathbf{D} | \theta, m, \mathbf{a}) = \frac{m!}{D_1! \dots D_n! (m - D_0)!} \cdot \prod_{i=1}^n \left( \sum_{j: y_j \in a_i} \theta_j \right)^{D_i} \left( \sum_{j: y_j \notin \cup_{i=1}^n a_i} \theta_j \right)^{m - D_0}, \quad (3)$$

where  $D_0 = \sum_{i=1}^n D_i \leq m$ . Each of the  $n$  terms in the multiplication corresponds to a group of locations that are covered by one product in assortment  $\mathbf{a}$ , and the last term corresponds to all the remaining locations (if any) that are not covered.

**EXAMPLE 2.** Suppose there are three consumer locations. If the firm offers a product that covers location 1 only,  $\mathbf{a} = ([y_1, y_1])$ , the conditional distribution of sales  $D_1$  is  $L^c(D_1 | \theta, m, ([y_1, y_1])) = (m! / (D_1! (m - D_1)!)) \cdot \theta_1^{D_1} (\theta_2 + \theta_3)^{m - D_1}$ . On the other hand, if the firm chooses to offer one product that covers both locations 1 and 2,  $\mathbf{a} = ([y_1, y_2])$ , the conditional distribution of sales is  $L^c(D_1 | \theta, m, ([y_1, y_2])) = (m! / (D_1! (m - D_1)!)) (\theta_1 + \theta_2)^{D_1} \theta_3^{m - D_1}$ .

If the firm starts with prior  $\pi$ , chooses assortment  $\mathbf{a}$ , then observes sales  $\mathbf{D}$  and market size  $m$ , the *censored* posterior distribution on  $\theta$  is  $\Pi^c(\theta; \pi, \mathbf{D}, m, \mathbf{a}) = L^c(\mathbf{D} | \theta, m, \mathbf{a}) \pi(\theta) / f^c(\mathbf{D}; \pi, m, \mathbf{a})$ . Here,  $\Pi^c$  is the Bayesian updating operator with *censored* taste information and  $f^c(\mathbf{D}; \pi, m, \mathbf{a}) = \int_{\theta} L^c(\mathbf{D} | \theta, m, \mathbf{a}) \pi(\theta) d\theta$  is the probability mass function for  $\mathbf{D}$  under  $\mathbf{a}$  conditional on  $m$ .

The analysis in the following sections is general as it holds for any continuous prior  $\pi$ . We discuss conjugate models for both the uncensored and censored information cases in §4.4.

## 4.2. Dynamic Programming Formulation

As in the single-period problem, the firm can decide on the assortment  $\mathbf{a} \in \mathcal{A}$  in each period as opposed to  $\mathbf{x}$  and  $\mathbf{p}$ . Given the assortment,  $\mathbf{x}$  and  $\mathbf{p}$  are as in Lemma 1 in §3.

In the uncensored information case, we write the optimal value function with  $t$  periods remaining,  $v_t^u(\pi)$ , as a DP recursion. With zero periods remaining, the optimal value function is  $v_0^u(\pi) = 0$ . For earlier periods, we take the optimal value function to be

$$v_t^u(\pi) = \max_{\mathbf{a} \in \mathcal{A}} r(\mathbf{a}, \pi) + \delta \mathbb{E}_{\pi} [v_{t-1}^u(\Pi^u(\pi, \tilde{\mathbf{T}}, \tilde{m}))], \quad (4)$$

where  $\delta$  is the discount factor ( $0 \leq \delta \leq 1$ ). The total profit the firm receives from assortment  $\mathbf{a}$  is the sum of immediate profit,  $r(\mathbf{a}, \pi)$ , and the expected profit-to-go,  $\mathbb{E}_{\pi} [v_{t-1}^u(\Pi^u(\pi, \tilde{\mathbf{T}}, \tilde{m}))]$ . Here, and in the remaining sections, we emphasize that expectations depend on our state variable  $\pi$  by writing  $\mathbb{E}_{\pi}[\cdot]$ . As in the single-period problem, we write the immediate profit from assortment  $\mathbf{a}$  as  $r(\mathbf{a}, \pi) = -fn + \sum_{i=1}^n (p_i(a_i) - c) \mathbb{E}_{\pi} [\tilde{D}_i(a_i)]$ , where  $\mathbb{E}_{\pi} [\tilde{D}_i(a_i)] = \mu \sum_{j: y_j \in a_i} \mathbb{E}_{\pi} [\tilde{\theta}_j]$  is the expected demand for product  $i$  and  $\mathbb{E}_{\pi} [\tilde{\theta}_j] = \int \theta_j \pi(\theta) d\theta$  is the expected value of  $\tilde{\theta}_j$  under prior  $\pi$ .

The expected profit-to-go in (4) takes into account the uncertainty about consumer tastes and the market size to be observed next period and can be written more explicitly as

$$\mathbb{E}_{\pi} [v_{t-1}^u(\Pi^u(\pi, \tilde{\mathbf{T}}, \tilde{m}))] = \sum_m \sum_{\mathbf{T}} v_{t-1}^u(\Pi^u(\pi, \mathbf{T}, m)) f^u(\mathbf{T}; \pi, m) P(\tilde{m} = m).$$

Because our DP state variable is a probability distribution, we frequently suppress the domain of the distribution and write the posterior as  $\Pi^u(\pi, \mathbf{T}, m)$  when we want to consider this distribution as a function of the prior  $\pi$ , observed taste vector  $\mathbf{T}$ , and market size  $m$ .

We write an analogous DP recursion for the optimal value function in the censored information case when the firm does not observe consumer tastes directly but has to infer them through sales in each period. With zero periods remaining, we take  $v_0^c(\pi) = 0$ . For earlier periods, we take

$$v_t^c(\pi) = \max_{\mathbf{a} \in \mathcal{A}} \{ r(\mathbf{a}, \pi) + \delta \mathbb{E}_{\pi} [v_{t-1}^c(\Pi^c(\pi, \tilde{\mathbf{D}}, \tilde{m}, \mathbf{a}))] \}, \quad (5)$$

where  $\mathbb{E}_\pi[v_{t-1}^c(\Pi^c(\pi, \tilde{\mathbf{D}}, \tilde{m}, \mathbf{a}))] = \sum_m \sum_{\mathbf{D}} v_{t-1}^c(\Pi^c(\pi, \mathbf{D}, m, \mathbf{a}))f^c(\mathbf{D}; \pi, m, \mathbf{a})P(\tilde{m} = m)$ .

Note that in (4), the assortment  $\mathbf{a}$  affects only the immediate profit,  $r(\mathbf{a}, \pi)$ . Thus, the optimal assortment in the uncensored information case can be solved for myopically using the shortest-path formulation in Alptekinoglu et al. (2012). This is not the case when the consumer taste information is censored; the assortments chosen over multiple periods affect both immediate profits and expected profits-to-go. In the censored information case, the myopically optimal assortment does not necessarily result in the highest profit-to-go; the firm needs to trade off immediate profits with informational gains (learning benefits) when choosing the optimal assortment.

### 4.3. Comparing Optimal Assortments in Censored and Uncensored Information Cases

We first show that the value functions in (4) and (5) are convex in the firm’s beliefs,  $\pi$ , each period. This property helps us study how the optimal assortments compare in censored and uncensored information cases.

**PROPOSITION 1.** *Given any two priors  $\pi_1$  and  $\pi_2$ , let  $\pi_\alpha(\boldsymbol{\theta}) = \alpha\pi_1(\boldsymbol{\theta}) + (1 - \alpha)\pi_2(\boldsymbol{\theta})$  for some  $\alpha \in [0, 1]$ . Then,  $v_t^c(\pi_\alpha) \leq \alpha v_t^c(\pi_1) + (1 - \alpha)v_t^c(\pi_2)$  and  $v_t^u(\pi_\alpha) \leq \alpha v_t^u(\pi_1) + (1 - \alpha)v_t^u(\pi_2)$ , that is, the value functions are convex in  $\pi$ .*

To compare optimal assortments in censored and uncensored information cases, we need to order assortments based on their information content. Consider an assortment  $\mathbf{a}$  with  $n$  products. Let  $P(\mathbf{a})$  be a partition on  $Y$ , which contains  $n$  subsets of  $Y$ , each of which corresponds to a distinct group of consumer locations covered by one product in  $\mathbf{a}$ , plus possibly one more subset that contains all the remaining consumer locations (if any) that are not covered. For example, if there are six consumer locations and  $\mathbf{a} = ([y_1, y_3], [y_5, y_5])$ , then  $P(\mathbf{a}) = \{\{y_1, y_2, y_3\}, \{y_5\}, \{y_4, y_6\}\}$ .

Let  $\mathcal{P}$  be the set of all feasible partitions.<sup>1</sup> We define the notion of a *finer* partition, which gives us a partial order over all assortments in  $\mathcal{A}$  that reflects their information content.

**DEFINITION 1.** A partition  $P = \{S_1, \dots, S_k\}$  is “finer” than partition  $P' = \{S'_1, \dots, S'_{k'}\}$  with  $\bigcup_{i=1}^{k'} S'_i = \bigcup_{i=1}^k S_i$ , if for

each  $S_i \in P$  there exists a set  $S'_i \in P'$  such that  $S_i \subseteq S'_i$ . We say partition  $P$  is “strictly finer” than partition  $P'$  if the subset relationship is strict for at least one  $S_i \in P$ .

**DEFINITION 2.** Assortment  $\mathbf{a}$  is (strictly) “more informative” than assortment  $\mathbf{a}'$ , or  $(\mathbf{a} \succ_I \mathbf{a}') \mathbf{a} \succeq_I \mathbf{a}'$ , if  $P(\mathbf{a})$  is a (strictly) finer partition than  $P(\mathbf{a}')$ . Assortment  $\mathbf{a}$  is “as informative as” assortment  $\mathbf{a}'$ , or  $\mathbf{a} \approx_I \mathbf{a}'$ , if  $P(\mathbf{a}) = P(\mathbf{a}')$ .

Figure 2 shows all possible assortments and corresponding partitions when  $N = 3$ . The information content decreases as one moves from the group of assortments on the left to the group of assortments on the right; any assortment in the first box is more informative than the assortments in the second or third boxes; and any assortment in the second box is more informative than the assortments in the third box. Assortments that imply different partitions in the second box cannot be ordered based on their information content because they censor different portions of the attribute space. For example,  $([y_1, y_1])$  provides censored information on locations 2 and 3, whereas  $([y_1, y_2])$  provides censored information on locations 1 and 2.

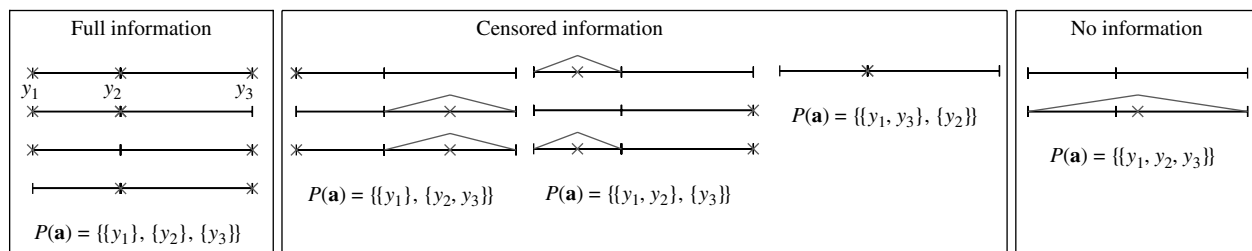
The following lemma shows that a more informative assortment implies higher product variety.

**LEMMA 2.** *If  $\mathbf{a} \succ_I \mathbf{a}'$ , then  $n \geq n'$ , where  $n$  is the number of products in assortment  $\mathbf{a}$  and  $n'$  is the number of products in assortment  $\mathbf{a}'$ .*

However, we note that increasing the number of products in an assortment does not necessarily improve its informativeness;  $([y_1, y_1], [y_2, y_3])$  has more products than  $([y_1, y_1])$  and  $([y_2, y_3])$  but has the same information content. This is because we assume that  $\tilde{m}$  is observable; if the firm cannot observe market-size realizations, then increasing the number of products in an assortment while keeping the implied partition the same improves informativeness. We discuss this and other implications of unobservable market size in §6 and in the e-companion.

Observing sales from any assortment is always (weakly) less informative than obtaining uncensored consumer taste information and therefore results in less profit-to-go. Thus, the optimal profits in the uncensored information case are always higher than the optimal profits in the censored information case. Also, assortments that are less informative

**Figure 2.** All possible assortments and corresponding partitions for  $N = 3$ .



INFORMS holds copyright to this article and distributed this copy as a courtesy to the author(s). Additional information, including rights and permission policies, is available at http://journals.informs.org/.



result in lower profits-to-go. To prove these results, we first prove a lemma on optimal profit-to-go functions.

LEMMA 3. *For any convex function  $u(\pi)$  and assortments  $\mathbf{a} \succeq_I \mathbf{a}'$ , we have*

$$\begin{aligned} \mathbb{E}_\pi[u(\Pi^u(\pi, \tilde{\mathbf{T}}, \tilde{m}))] &\geq \mathbb{E}_\pi[u(\Pi^c(\pi, \tilde{\mathbf{D}}, \tilde{m}, \mathbf{a}))] \\ &\geq \mathbb{E}_\pi[u(\Pi^c(\pi, \tilde{\mathbf{D}}, \tilde{m}, \mathbf{a}'))]. \end{aligned}$$

Using this lemma, we show that the optimal profits without censoring are higher than optimal profits with censoring and that it cannot be optimal to select a less informative assortment in the censored information case than the optimal assortment in the uncensored information case.

THEOREM 1. *For all  $t$ , we have  $v_t^u(\pi) \geq v_t^c(\pi)$ . Moreover, if the optimal assortment in the uncensored information case when there are  $t$  periods to go is  $\mathbf{a}^*$ , then for any assortment  $\mathbf{a}$  such that  $\mathbf{a}^* \succeq_I \mathbf{a}$ , we have  $v_t^c(\pi; \mathbf{a}^*) \geq v_t^c(\pi; \mathbf{a})$ , where  $v_t^c(\pi; \mathbf{a}) = r(\mathbf{a}, \pi) + \delta \mathbb{E}_\pi[v_{t-1}^c(\Pi^c(\pi, \tilde{\mathbf{D}}, \tilde{m}, \mathbf{a}))]$ . Therefore, the value function for the uncensored information case is always higher, and the optimal assortment in the censored information case cannot be less informative than the optimal assortment in the uncensored information case (i.e., the myopically optimal assortment).*

PROOF. We first show that  $v_t^u(\pi) \geq v_t^c(\pi)$ . The proof follows by induction. When there is one period to go,  $v_1^u(\pi) = v_1^c(\pi)$ . Now, assume  $v_{t-1}^u(\pi) \geq v_{t-1}^c(\pi)$ . For any given prior  $\pi$  and assortment  $\mathbf{a}$ , Proposition 1 and Lemma 3 imply the first inequality below, and the induction hypothesis implies the second:  $\mathbb{E}_\pi[v_{t-1}^u(\Pi^u(\pi, \tilde{\mathbf{T}}, \tilde{m}))] \geq \mathbb{E}_\pi[v_{t-1}^u(\Pi^c(\pi, \tilde{\mathbf{D}}, \tilde{m}, \mathbf{a}))] \geq \mathbb{E}_\pi[v_{t-1}^c(\Pi^c(\pi, \tilde{\mathbf{D}}, \tilde{m}, \mathbf{a}))]$ . This then implies that  $v_t^u(\pi) \geq v_t^c(\pi)$ . Now, assume that  $\mathbf{a}^*$  is the optimal assortment in the uncensored information model:  $\mathbf{a}^* = \arg \max r(\mathbf{a}, \pi) + \delta \mathbb{E}_\pi[v_{t-1}^u(\Pi^u(\pi, \tilde{\mathbf{T}}, \tilde{m}))]$ . For any assortment  $\mathbf{a}$  such that  $\mathbf{a}^* \succeq_I \mathbf{a}$ , we have  $r(\mathbf{a}^*, \pi) \geq r(\mathbf{a}, \pi)$ , because the optimal policy in the uncensored information case is myopic. Using Proposition 1 and Lemma 3, we also have  $\mathbb{E}_\pi[v_{t-1}^c(\Pi^c(\pi, \tilde{\mathbf{D}}, \tilde{m}, \mathbf{a}^*))] \geq \mathbb{E}_\pi[v_{t-1}^c(\Pi^c(\pi, \tilde{\mathbf{D}}, \tilde{m}, \mathbf{a}))]$ . Therefore, for any  $\mathbf{a}$  such that  $\mathbf{a}^* \succeq_I \mathbf{a}$ , we have  $v_t^c(\pi; \mathbf{a}^*) \geq v_t^c(\pi; \mathbf{a})$ .  $\square$

This result is analogous to the “stock more” result derived from censored newsvendor problems, in which the

newsvendor is uncertain about demand and gathers information through sales (see, for example, Harpaz et al. 1982, Ding et al. 2002, Lu et al. 2005). Sales observations provide censored information on demand: depleted inventory signals that the demand must have been higher than the order quantity; however, the newsvendor cannot observe the exact realization of demand. One can show that the newsvendor who uses censored demand information (sales) to update its beliefs has to stock more than the newsvendor who can observe demand. In these problems, higher order quantities are more informative because they provide less demand censoring.<sup>2</sup>

Informativeness order in censored newsvendor problems is a complete order, whereas it is a partial order in the product assortment problem (not every assortment pair can be ordered).<sup>3</sup> In the censored information case, it can be optimal to offer an assortment that is neither more nor less informative than the myopically optimal assortment (see Example Problem 2 in §5).

The result of Theorem 1 has some parallels in the dynamic assortment literature. For example, the policy developed by Caro and Gallien (2007) may involve products that do not necessarily have the highest expected demand, but have high variance in order to collect more information.

The lemma below states that two assortments that have the same information content have the same profit-to-go functions defined in (5).

LEMMA 4. *If  $\mathbf{a}, \mathbf{a}' \in \mathcal{A}$  and  $\mathbf{a} \approx_I \mathbf{a}'$ , then*

$$\mathbb{E}_\pi[v_{t-1}^c \cdot (\Pi^c(\pi, \tilde{\mathbf{D}}, \tilde{m}, \mathbf{a}))] = \mathbb{E}_\pi[v_{t-1}^c(\Pi^c(\pi, \tilde{\mathbf{D}}, \tilde{m}, \mathbf{a}'))].$$

The results in this section are not only of theoretical interest, but also provide some computational savings in finding the optimal assortment. Due to Lemma 4, one needs to consider only the assortments that yield highest immediate profit among all the assortments that imply the same partition. This is because all assortments that imply the same partition have the same profit-to-go. Furthermore, due to Theorem 1, one need not consider assortments that are less informative than the myopically optimal assortment. Table 1 illustrates computational savings from Lemma 4 and Theorem 1 with several numerical examples. We compare

Table 1. Computational savings from Lemma 4 and Theorem 1.

	Value function calculations (time in seconds)		
	Method (i)	Method (ii)	Method (iii)
3 periods			
$m = 2$	31,863 (91)	4,291 (15)	111 (0)
$m = 3$	90,301 (312)	11,131 (50)	421 (1)
4 periods			
$m = 2$	5,671,615 (23,962)	278,916 (1,555)	1,111 (1)
$m = 3$	27,090,301 (149,505)	1,168,756 (9,253)	8,421 (9)

the number of value function calculations and computation time when the value functions are computed in three different ways: (i) by comparing the sum of immediate reward and profit-to-go function for every possible assortment (i.e., complete enumeration), (ii) by comparing only the assortments that yield the highest immediate profits for each partition (i.e., using Lemma 4), (iii) by comparing only the assortments that yield the highest immediate profits for each partition and that are not less informative than the myopically optimal assortment (i.e., using Lemma 4 and Theorem 1). We use the following parameters:  $\pi$  has a Dirichlet distribution with  $\alpha = (1, 1, 1, 1)$ ,  $Y = \{0.2, 0.4, 0.6, 0.8\}$ ,  $\delta = 1$ ,  $\bar{p} = 20$ ,  $d = 20$ ,  $c = 3$ , and  $f = 1$  (see §4.4 for the details of how Bayesian updating works under a Dirichlet prior). We vary  $m \in \{2, 3\}$  and the number of periods in  $\{3, 4\}$ .

We see that although Lemma 4 and Theorem 1 do not change the theoretical complexity of the search for the optimal solution, in practice they may lead to significant time savings. This is especially true when  $m$  is large and there are many time periods, which is likely to be the case in practice (we found that the other parameters had little impact on the magnitude of the computational savings).

#### 4.4. Conjugate Models

The analysis so far holds for any prior distribution,  $\pi$ . Here, we study conjugate learning models where posteriors are in the same family of distributions as priors. This property makes Bayesian updating easier and allows us to reduce the state space of our DP model. Also, we use these conjugate models in our numerical study in §5.

*Uncensored Information.* The likelihood function in the uncensored information case is given in (2) and is a multinomial distribution. The conjugate prior for a multinomial distribution is the Dirichlet distribution. Its probability density function with parameters  $\alpha = (\alpha_1, \dots, \alpha_N)$  is  $\pi_D(\theta; \alpha) = (1/B(\alpha))\theta_1^{\alpha_1-1} \dots \theta_N^{\alpha_N-1}$ , where  $B(\alpha) = \prod_{j=1}^N \Gamma(\alpha_j) / \Gamma(\alpha_0)$ ,  $\alpha_0 = \sum_{j=1}^N \alpha_j$  and  $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$ . One can interpret  $\alpha = (\alpha_1, \dots, \alpha_N)$  as past uncensored observations from locations  $y_1, \dots, y_N$ . Then, after observing consumer tastes  $\mathbf{T} = (T_1, \dots, T_N)$ , the posterior distribution is a Dirichlet distribution with parameters  $\alpha + \mathbf{T}$ .

*Censored Information.* The likelihood of sales for a given assortment  $\mathbf{a}$  is given in (3). The conjugate prior for this likelihood is the Extended Dirichlet distribution (Dickey et al. 1987). Its probability density function with parameters  $\alpha = (\alpha_1, \dots, \alpha_N)$  and  $\mathbf{d} = (d_1, \dots, d_K)$  is:  $\pi_{ED}(\theta; \alpha, \mathbf{d}) = \pi_D(\theta; \alpha) \prod_{k=1}^K (\sum_{j=1}^N Z_{jk} \theta_j)^{d_k} / \mathcal{R}(\alpha, Z, \mathbf{d})$ , where  $K = 2^N - N - 2$ ,  $Z$  is an  $N \times K$  binary matrix whose columns correspond to possible combinations of censored locations,<sup>4</sup> and  $\mathcal{R}(\alpha, Z, \mathbf{d})$  is a normalizing constant,

$$\begin{aligned} \mathcal{R}(\alpha, Z, \mathbf{d}) &= \mathbb{E}_{\pi_D} \left[ \prod_{k=1}^K \left( \sum_{j=1}^N Z_{jk} \tilde{\theta}_j \right)^{d_k} \right] \\ &= \int \prod_{k=1}^K \left( \sum_{j=1}^N Z_{jk} \theta_j \right)^{d_k} \pi_D(\theta; \alpha) d\theta. \end{aligned}$$

A Dirichlet distribution is an Extended Dirichlet distribution with  $\mathbf{d} = \mathbf{0}$ . One can interpret  $\alpha$  as past uncensored observations and  $\mathbf{d}$  as past censored observations.

EXAMPLE 3. Suppose there are three consumer locations. Then,

$$Z = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

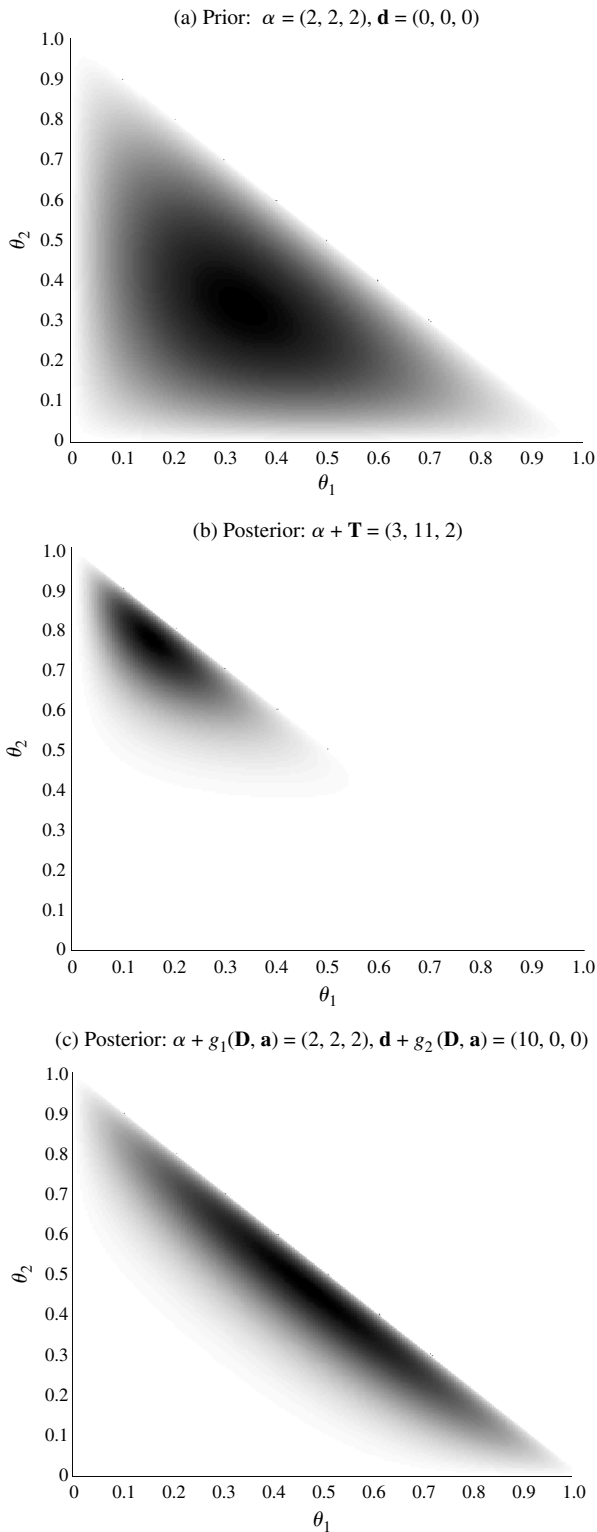
The first column in  $Z$  corresponds to observing censored information on locations 1 and 2; the firm may gather such information by offering either the assortment  $\mathbf{a} = ([y_1, y_2])$  or  $\mathbf{a} = ([y_3, y_3])$ . The second and third columns correspond to censored information on locations  $y_1$  and  $y_3$ , and  $y_2$  and  $y_3$ , respectively. Then, we have  $\pi_{ED}(\theta; \alpha, \mathbf{d}) = \pi_D(\theta; \alpha) (\theta_1 + \theta_2)^{d_1} (\theta_1 + \theta_3)^{d_2} (\theta_2 + \theta_3)^{d_3} / \mathcal{R}(\alpha, Z, \mathbf{d})$ , where  $\mathcal{R}(\alpha, Z, \mathbf{d}) = \mathbb{E}_{\pi_D} [(\tilde{\theta}_1 + \theta_2)^{d_1} (\tilde{\theta}_1 + \theta_3)^{d_2} \cdot (\tilde{\theta}_2 + \tilde{\theta}_3)^{d_3}] = \int (\theta_1 + \theta_2)^{d_1} (\theta_1 + \theta_3)^{d_2} (\theta_2 + \theta_3)^{d_3} \cdot \pi_D(\theta; \alpha) d\theta$ . We show how to calculate  $\mathcal{R}(\alpha, Z, \mathbf{d})$  in the e-companion.

Because the Extended Dirichlet distribution is a conjugate prior for the censored likelihood, formulas for updating beliefs are relatively easy. Sales from an assortment may provide uncensored information on some consumer locations (e.g., if the market segment of a product is a single consumer location) and censored information on others (e.g., if the market segment of a product contains multiple consumer locations); we need to distinguish them. Suppose the firm observes sales  $\mathbf{D} = (D_1, D_2, \dots, D_n)$  from assortment  $\mathbf{a}$ . Let  $\mathbf{g}_1(\mathbf{D}, \mathbf{a})$  be a  $1 \times N$  vector where the  $j$ th component is the number of consumers in location  $j$  observed in  $\mathbf{D}$ ; this is uncensored information. Also, let  $\mathbf{g}_2(\mathbf{D}, \mathbf{a})$  be a  $1 \times K$  vector where the  $j$ th component is the total number of consumers from locations in the  $j$ th column of the  $Z$  matrix; this is censored information. Then, the posterior is an Extended Dirichlet distribution with parameters  $\alpha + \mathbf{g}_1(\mathbf{D}, \mathbf{a})$  and  $\mathbf{d} + \mathbf{g}_2(\mathbf{D}, \mathbf{a})$ .

EXAMPLE 4. Suppose there are three consumer locations, and let  $\pi$  be a Dirichlet distribution with parameters  $\alpha = (2, 2, 2)$ . In the uncensored information case, if the firm receives consumer taste information  $\mathbf{T} = (1, 9, 0)$ , then the posterior is a Dirichlet distribution with parameters  $\alpha + \mathbf{T} = (3, 11, 2)$ . The prior and the posterior distributions are shown in Figures 3(a) and 3(b) on a unit simplex where darker regions correspond to points with higher probability density. Observing many consumers from location  $y_2$  leads the firm to move probability mass towards higher values of  $\theta_2$ .

Take the censored information case now, and suppose that the firm does not receive information on consumer tastes but rather infers them from the sales of a single product that covers the first two locations:  $\mathbf{a} = ([y_1, y_2])$ . If the sales of that product is 10 and the firm observes  $m = 10$ , then the firm's posterior is an Extended Dirichlet with parameters  $\alpha + \mathbf{g}_1(\mathbf{D}) = (2, 2, 2)$  and  $\mathbf{d} + \mathbf{g}_2(\mathbf{D}) =$

**Figure 3.** Prior and posterior distributions in Example 4.



(10, 0, 0). Here, the firm observes high demand from locations  $y_1$  and  $y_2$ ; however, the sales information cannot distinguish between these locations. The posterior moves density away from  $\theta_3$  and distributes it close to  $\theta_1 + \theta_2 = 1$  line. See Figure 3(c).

We next present a DP formulation for the conjugate models. In the uncensored information case, we write the DP recursion with  $\alpha$  as the state variable. With zero periods remaining, the optimal value function is  $v_0^u(\alpha) = 0$ . For earlier periods, we take the optimal value function to be

$$v_t^u(\alpha) = \max_{\mathbf{a} \in \mathcal{A}} r(\mathbf{a}, \alpha) + \delta \mathbb{E}[v_{t-1}^u(\alpha + \tilde{\mathbf{T}})].$$

Here,

$$r(\mathbf{a}, \alpha) = -fn + \sum_{i=1}^n \left( (p_i(a_i) - c)\mu \sum_{j: y_j \in a_i} \mathbb{E}[\tilde{\theta}_j; \alpha] \right)$$

is the immediate profit, and

$$\mathbb{E}[v_{t-1}^u(\alpha + \tilde{\mathbf{T}})] = \sum_m \left( \sum_{\mathbf{T}} v_{t-1}^u(\alpha + \mathbf{T}) f(\mathbf{T}; \alpha; m) \right) P(\tilde{m} = m)$$

is the profit-to-go, where  $\mathbb{E}[\tilde{\theta}_j; \alpha] = \alpha_j / \alpha_0$  and  $f(\mathbf{T}; \alpha, m)$  is a Dirichlet-Multinomial distribution:

$$\begin{aligned} f(\mathbf{T}; \alpha, m) &= \int_{\boldsymbol{\theta}} L^u(\mathbf{T} | \boldsymbol{\theta}, m) \pi_D(\boldsymbol{\theta}; \alpha) d\boldsymbol{\theta} \\ &= \frac{m!}{T_1! \cdots T_N!} \frac{B(\alpha + \mathbf{T})}{B(\alpha)}. \end{aligned} \tag{6}$$

In the censored information case, we write the DP recursion with  $\alpha = (\alpha_1, \dots, \alpha_N)$  and  $\mathbf{d} = (d_1, \dots, d_K)$  as the state variables. With zero periods remaining, the optimal value function is  $v_0^c(\alpha, \mathbf{d}) = 0$ . For earlier periods, we take the optimal value function to be

$$\begin{aligned} v_t^c(\alpha, \mathbf{d}) &= \max_{\mathbf{a} \in \mathcal{A}} \{ r(\mathbf{a}, \alpha, \mathbf{d}) + \delta \mathbb{E}[v_{t-1}^c(\alpha + g_1(\tilde{\mathbf{D}}, \mathbf{a}), \mathbf{d} + g_2(\tilde{\mathbf{D}}, \mathbf{a}))] \}. \end{aligned}$$

Here, the immediate profits are

$$r(\mathbf{a}, \alpha, \mathbf{d}) = -fn + \sum_{i=1}^n \left( (p_i(a_i) - c)\mu \sum_{j: y_j \in a_i} \mathbb{E}[\tilde{\theta}_j; \alpha, \mathbf{d}] \right),$$

where  $\mathbb{E}[\tilde{\theta}_j; \alpha, \mathbf{d}]$  is the expected value of  $\tilde{\theta}_j$  if  $\tilde{\boldsymbol{\theta}}$  has an Extended Dirichlet distribution with parameters  $\alpha$  and  $\mathbf{d}$ . We show how to compute expected values under the Extended Dirichlet distribution in the e-companion.

Profit-to-go function,

$$\mathbb{E}[v_{t-1}^c(\alpha + g_1(\tilde{\mathbf{D}}, \mathbf{a}), \mathbf{d} + g_2(\tilde{\mathbf{D}}, \mathbf{a}))],$$

for the censored information case is

$$\begin{aligned} \sum_m \left( \sum_{\mathbf{D}: D_0 \leq m} v_{t-1}^c(\alpha + g_1(\mathbf{D}, \mathbf{a}), \mathbf{d} + g_2(\mathbf{D}, \mathbf{a})) f(\mathbf{D}; \alpha, \mathbf{d}, m, \mathbf{a}) \right) P(\tilde{m} = m), \end{aligned}$$

where

$$\begin{aligned} f(\mathbf{D}; \alpha, \mathbf{d}, m, \mathbf{a}) &= \int_{\boldsymbol{\theta}} L^c(\mathbf{D} | \boldsymbol{\theta}, m, \mathbf{a}) \pi_D(\boldsymbol{\theta}; \alpha, \mathbf{d}) d\boldsymbol{\theta} \\ &= \frac{m!}{D_1! \cdots D_n! (m - D_0)!} \frac{B(\alpha + g_1(\mathbf{D}, \mathbf{a}))}{B(\alpha)} \\ &\quad \cdot \frac{\mathcal{R}(\alpha + g_1(\mathbf{D}, \mathbf{a}), Z, \mathbf{d} + g_2(\mathbf{D}, \mathbf{a}))}{\mathcal{R}(\alpha, Z, \mathbf{d})}. \end{aligned}$$

This distribution is similar to the Dirichlet-Multinomial distribution in (6) with the exception of the ratio of normalizing constants that addresses the censoring.

In summary, using the conjugate models allows us to change the state variable of the DP from a probability distribution,  $\pi$ , to a finite-sized vector of size  $N$  in the uncensored case and of size  $N \times (2^N - N - 2)$  in the censored case. Furthermore, it allows us to reduce the infinite state space of the original problem, that is, the set of all possible prior distributions for each period, to a countably infinite state space when the market size has infinite support or a finite state space when it has finite support.

## 5. Numerical Study

The purpose of this section is twofold. First, we give some numerical examples that compare optimal assortments under various multiperiod scenarios and highlight interesting aspects of dynamic assortments that relate to the exploration-exploitation trade-off. Second, we study the value of learning.

### 5.1. Exploration vs. Exploitation

To demonstrate the trade-off between learning and maximizing immediate profits, we study five numerical examples. In all cases, we set  $\delta = 1$  and assume a Dirichlet or extended Dirichlet prior and a deterministic market size. We use short time horizons ( $t \leq 3$ ), few consumer locations ( $N = 3$ ), and a small market size ( $m = 2$ ) so that we can list all possible sales vectors and present the optimal dynamic assortment in compact tables. Consistent with the way our DP formulation is laid out, we number periods backward: period 3 precedes period 2, and period 2 precedes period 1. As a shorthand, we write  $y_i$  as opposed to  $[y_i, y_i]$  when a product covers only one location.

**EXAMPLE PROBLEM 1.** In this example, we show that the optimal assortments under the censored and uncensored information cases may be such that they cannot be ordered using the informativeness order in Definition 2. We use two periods and the following parameters:  $Y = \{0.2, 0.4, 0.8\}$ ,  $\alpha = (8, 3, 4)$ ,  $\bar{p} = 20$ ,  $d = 40$ ,  $m = 2$ ,  $c = 5$ , and  $f = 13$ .

The results are shown in Table 2. We list all the partitions in  $\mathcal{P}$  in column 1. Among all the assortments that imply a given partition, we put the assortment that maximizes the immediate profit function in period 2 together with the immediate profit,  $r$ , in column 2 (see Lemma 4). For each assortment in column 2, we list all possible values of the sales vector  $\mathbf{D}$  in column 3 and the probability of observing that sales vector,  $f(\mathbf{D})$ , in column 4. Given a period 2 assortment and a sales vector realization, we present the optimal assortment in period 1 (column 5) and the immediate profit obtained with this assortment (column 6). Then, we compute the expected profit-to-go (column 7) using the profit values in column 6 and the probabilities from column 4. Finally, we present total expected profits over the last two periods (column 8), which is the sum of immediate profits from column 2 and expected profit-to-go from column 7.

The optimal assortment in period 2 offers one product at the first location for a total expected profit of 6.38, whereas the myopically optimal assortment offers one product to cover the first two locations. In this case, the optimal assortment implies the partition  $\{\{1\}, \{2, 3\}\}$ , whereas the myopically optimal assortment implies the partition  $\{\{1, 2\}, \{3\}\}$ . That is, the optimal assortment is neither more nor less informative than the myopically optimal assortment.

**EXAMPLE PROBLEM 2.** In this example, we show that it may be optimal to offer an assortment that leads to a loss in the

**Table 2.** Possible assortments and expected profits for Example Problem 2.

Period 2 partition	Period 2 best assort. (a) and profit	Period 2 sales (D)	Prob. (f(D))	Optimal assort. in period 1	Profit in period 1	Expected profit in period 1	Total expected profit for two periods
{1}, {2}, {3}	(0.2, 0.8) $r = -2.00$	(2, 0)	0.30	(0.2)	4.65	3.47	1.47
		(1, 1)	0.27	(0.2)	2.88		
		(0, 2)	0.08	([0.2, 0.4])	1.24		
		(1, 0)	0.20	([0.2, 0.4])	3.82		
		(0, 1)	0.10	([0.2, 0.4])	2.53		
		(0, 0)	0.05	([0.2, 0.4])	3.82		
{1, 2}, {3}	([0.2, 0.4]) $r = 3.13$	(2)	0.55	([0.2, 0.4])	3.82	3.13	6.26
		(1)	0.37	([0.2, 0.4])	2.53		
		(0)	0.08	([0.2, 0.4])	1.24		
{1}, {2, 3}	(0.2) $r = 3.00$	(2)	0.30	(0.2)	4.65	3.38	6.38
		(1)	0.47	([0.2, 0.4])	3.08		
		(0)	0.23	([0.2, 0.4])	2.34		
{1, 3}, {2}	(0.4) $r = -7.00$	(2)	0.05	([0.2, 0.4])	3.82	3.46	-3.54
		(1)	0.30	([0.2, 0.4])	3.39		
		(0)	0.65	(0.2)	3.47		
{1, 2, 3}	$\emptyset$ $r = 0$	—	1	([0.2, 0.4])	3.13	3.13	3.13

**Table 3.** Possible assortments and expected profits for Example Problem 3.

Period 2 partition	Period 2 best assort. (a) and profit	Period 2 sales (D)	Prob. (f(D))	Optimal assort. in period 1	Profit in period 1	Expected profit in period 1	Total expected profit for two periods
{1}, {2}, {3}	(0.4, 0.8) $r = -11.75$	(2, 0)	0.07	$\emptyset$	0	0.40	-11.35
		(1, 1)	0.29	$\emptyset$	0		
		(0, 2)	0.40	(0.8)	1.00		
		(1, 0)	0.06	$\emptyset$	0		
		(0, 1)	0.15	$\emptyset$	0		
		(0, 0)	0.02	$\emptyset$	0		
{1, 2}, {3}	(0.8) $r = -0.25$	(2)	0.40	(0.8)	1.00	0.40	0.15
		(1)	0.44	$\emptyset$	0		
		(0)	0.15	$\emptyset$	0		
{1}, {2, 3}	([0.4, 0.8]) $r = -6.75$	(2)	0.77	(0.8)	0.05	0.04	-6.71
		(1)	0.21	$\emptyset$	0		
		(0)	0.02	$\emptyset$	0		
{1, 3}, {2}	(0.4) $r = -11.50$	(2)	0.07	$\emptyset$	0	0.25	-11.25
		(1)	0.35	$\emptyset$	0		
		(0)	0.57	(0.8)	0.44		
{1, 2, 3}	$\emptyset$ $r = 0$	—	1	$\emptyset$	0	0	0

current period to learn about consumer tastes. We use two periods and the following parameters:  $Y = \{0.2, 0.4, 0.8\}$ ,  $\alpha = (2, 4, 10)$ ,  $\bar{p} = 20$ ,  $d = 40$ ,  $m = 2$ ,  $c = 5$ , and  $f = 19$ . The results are shown in Table 3. We see that it is optimal to offer a product at the third location for a total profit of 0.15 in period 2 even though this assortment yields a negative immediate profit ( $r = -0.25$ ). By offering a product at this location in the first period, the firm hopes to observe sales that will allow the firm to conclude that there are enough consumers to justify offering a product there in the last period and generate positive profits.

If the firm was to optimize myopically in each period, it would conclude that it is optimal not to offer any products in periods 2 and 1. The expected total profit for the myopic firm is 0, whereas the firm that changes the assortment dynamically stands to make a strictly positive expected profit.

**EXAMPLE PROBLEM 3.** In this example, we show that the firm may alternate between exploitation and exploration and that variety and market coverage of the optimal assortment are nonmonotone with respect to time. We use three periods, four consumer locations, and the following parameters:  $Y = \{0.2, 0.4, 0.6, 0.8\}$ ,  $\alpha = (5, 1, 9, 2)$ ,  $\bar{p} = 20$ ,  $d = 40$ ,  $m = 2$ ,  $c = 5$ , and  $f = 10$ .

The optimal assortment in period 3 is (0.6), i.e., it is optimal to cover only the third consumer location. This is also the assortment that maximizes immediate profits in period 3. If no consumer buys in period 3, i.e.,  $\mathbf{D} = (0)$ , which happens with probability 0.24, then the myopically optimal assortment in period 2 is again (0.6) and the optimal assortment in period 2 is (0.2, 0.6), which is strictly more informative than the myopically optimal assortment. Finally, if in period 2 the firm sells two units of the product located at 0.6, i.e.,  $\mathbf{D} = (0, 2)$ , then the optimal assortment

in period 1 is (0.6), which is of course myopically optimal because it is the last period.

In this example, we see that the firm alternates between exploitation, i.e., choosing the myopically optimal assortment (in periods 3 and 1), and exploration (in period 2), i.e., not choosing the myopically optimal assortment. Therefore, a policy that has a number of periods of exploration followed by periods of exploitation such as the “adaptive assortment” of Rusmevichientong et al. (2010) and the “separation-based policy” of Saure and Zeevi (2011) is not necessarily optimal in our setting. (Saure and Zeevi 2011 also study the performance of “refined dynamic assortment policy,” which may alternate between exploration and exploitation.) Here, observing no sales for the product offered in period 3 comes as a surprise; therefore, the firm wants to learn about tastes and, in particular, investigate the relative proportion of consumers in location 1, which the firm believes to be the next most likely location. We also see that the information content of the optimal assortment is not monotone: the assortment in period 2 is more informative than the assortment in periods 1 and 3. Finally, we see that the number of products in the optimal assortment and its market coverage are not monotone.

The following examples demonstrate that some very intuitive properties regarding the exploration-exploitation trade-off need not hold.

**EXAMPLE PROBLEM 4.** In this example, we show that it is not necessarily optimal to drop a product that receives no sales. This result is analogous to the suboptimality of “switch-from-a-loser” strategy in bandit problems (see, for example, Berry and Fristedt 1985). We use two periods and the following parameters:  $Y = \{0.25, 0.75\}$ ,  $\alpha = (5, 1)$ ,  $\bar{p} = 20$ ,  $d = 40$ ,  $m = 1$ ,  $c = 5$ , and  $f = 10$ . With two periods to go, it is optimal to offer assortment (0.25). If the firm

sees zero sales, which happens with probability 0.17, then it is still optimal to offer assortment (0.25) with one period to go. This is because the new observation is not enough to counterbalance the prior distribution, which indicates a strong likelihood that most customers are in location 1.

**EXAMPLE PROBLEM 5.** In this example, we show that it may be optimal to stop covering a consumer location even after observing the highest sales possible for the product that covers this location. This result is at odds with the “stay-with-a-winner” rule shown to be optimal in some bandit problems (Berry and Fristedt 1985). We use two periods and the following parameters:  $Y = \{0.1, 0.5, 0.75, 0.95\}$ ,  $\alpha = (1, 2, 2, 4)$ ,  $\mathbf{d} = (0, \dots, 0, 3, 0, \dots, 0)$  where 3 is the value of parameter  $d$  that corresponds to censoring of locations 3 and 4,  $\bar{p} = 20$ ,  $d = 40$ ,  $m = 3$ ,  $c = 5$ , and  $f = 10$ . With two periods to go, the optimal assortment is  $([0.5, 0.75], 0.95)$ . If the firm sells three units of the product that covers locations 2 and 3, which happens with probability 0.1034, the optimal assortment with one period to go becomes  $([0.75, 0.95])$ ; that is, location 2 is no longer covered.

### 5.2. Value of Learning

In this section, we study the *value of learning* by comparing the performance of (i) a firm that receives uncensored information and uses the optimal policy, (ii) a firm that receives censored information and uses the optimal policy, (iii) a myopic firm that receives censored information and updates its beliefs, and (iv) a no-learning firm that does not update its beliefs at all. We break down the *value of learning* into three components (*proactive*, *active*, and *passive learning*) and study how they change with the prior distribution and the length of the planning horizon.

Regarding the myopic firm, note that there may be more than one assortment that maximizes immediate reward in a given period. These assortments do not necessarily have the same expected profit-to-go. Hence, expected profits of a myopic firm need to be defined carefully. We assume that the myopic firm uses a randomized policy to choose one of the myopically optimal assortments (if multiple). Let  $\mathcal{A}^{ml}(\pi) = \{\mathbf{a}: r(\mathbf{a}, \pi) = \max_{\mathbf{a}' \in \mathcal{A}} r(\mathbf{a}', \pi)\}$  be the set of assortments that maximize the immediate reward given  $\pi$

and  $v_t^{ml}$  be the expected profit made by the myopic firm. We have

$$v_t^{ml}(\pi) = \sum_{\mathbf{a} \in \mathcal{A}^{ml}(\pi)} \frac{1}{|\mathcal{A}^{ml}(\pi)|} (r(\mathbf{a}, \pi) + \delta \mathbb{E}_{\pi} [v_{t-1}^{ml}(\Pi^c(\pi, \tilde{\mathbf{D}}, \tilde{m}, \mathbf{a}))]),$$

where  $|A|$  denotes the size of set  $A$ . Note that the myopic firm does learn about consumers' taste distribution as it updates its prior distribution based on sales; previous literature has referred to this form of learning as *passive learning* (see Aviv and Pazgal 2005). However, it does not *actively* learn because it chooses not to explore; that is, it never sacrifices immediate profits in order to gather information on consumer tastes, which could improve future profits.

The no-learning firm does not update its distribution on customer preferences, and hence, it offers the assortment that is optimal at the beginning of the time horizon in every period. Let  $v_t^u$  be the expected profit made by the no-learning firm. Discounting the same immediate profit over  $t$  periods, we obtain  $v_t^{nl}(\pi) = ((1 - \delta^{t+1}) / (1 - \delta)) \max_{\mathbf{a} \in \mathcal{A}} r(\mathbf{a}, \pi)$ .

Obviously, we have  $v_t^u(\pi) \geq v_t^c(\pi) \geq v_t^{ml}(\pi) \geq v_t^{nl}(\pi)$ . We refer to  $v_t^u(\pi) - v_t^c(\pi)$  as the *value of proactive learning*, which measures the expected gains from collecting information about customer tastes at the time of their visit to the store (e.g., via a survey). We refer to  $v_t^c(\pi) - v_t^{ml}(\pi)$  as the *value of active learning*, which measures the expected gains from choosing exploration over exploitation in some periods in order to improve the information gathering. Finally, we refer to  $v_t^{ml}(\pi) - v_t^{nl}(\pi)$  as the *value of passive learning*, which measures the expected gains from updating the information about customer preferences in each period but taking advantage of it only myopically. We report all three quantities as a percentage of  $v_t^u$ , i.e.,  $(v_t^u - v_t^c)/v_t^u$ ,  $(v_t^c - v_t^{ml})/v_t^u$ , and  $(v_t^{ml} - v_t^{nl})/v_t^u$ . Table 4 shows how these values vary with the number of periods and the parameter  $\alpha$  of the prior distribution. We use  $N = 4$ ,  $Y = \{0.2, 0.4, 0.6, 1\}$ ,  $\delta = 1$ ,  $\bar{p} = 15$ ,  $d = 40$ ,  $m = 14$ ,  $c = 3$ , and  $f = 3$ .

In Table 4, we see that the biggest part of the value of learning stems from passive learning. Hence, it is very important for firms to update their beliefs over the distribution of customer preferences over time. The value of proactive learning is the second-biggest part, but in practice

**Table 4.** Value of learning (in percentages) as a function of the time horizon and prior distribution.

$\alpha$	Two periods			Three periods			Four periods			Five periods			Six periods		
	Pro.	Act.	Pass.	Pro.	Act.	Pass.	Pro.	Act.	Pass.	Pro.	Act.	Pass.	Pro.	Act.	Pass.
(1, 1, 1, 1)	1.69	1.69	28.81	7.38	2.38	34.85	8.79	3.27	38.70	9.21	3.84	41.12	9.55	4.36	42.96
(2, 2, 2, 2)	1.10	1.10	18.68	4.28	1.72	24.06	4.86	3.27	38.70	5.15	2.82	30.72	5.41	3.13	32.85
(3, 3, 3, 3)	0.81	0.81	13.82	3.38	1.39	18.39	3.70	1.95	22.03	3.74	2.38	24.40	3.91	2.64	26.45
(5, 1, 1, 1)	0.00	0.00	0.00	0.41	0.00	0.00	0.79	0.00	0.00	1.00	0.04	0.02	1.25	0.05	0.09
(4, 1, 1, 2)	0.00	0.67	3.33	1.50	0.78	4.41	2.04	0.90	5.45	2.26	0.96	6.26	2.50	1.09	6.97
(3, 1, 2, 2)	0.00	3.57	12.39	2.89	3.62	16.75	4.12	3.73	19.39	4.02	3.80	21.82	4.03	4.16	23.66

it requires an important investment in resources because one would need to question every customer who exits the store in order to know which product they truly wanted to buy (which may differ from the one they bought). The value of active learning is the smallest of all three, but it remains significant, especially when the number of periods is large. We see that as the number of periods increases, all three measures of the value of learning increase. This shows that as the time horizon expands, there might be a greater incentive to learn about consumer preferences. From looking at the first three rows, we also see that as the sum of the  $\alpha$  vector increases while keeping all  $\alpha_i$ s equal, all three measures of the value of learning decrease. This is because  $\alpha_0 = \sum_{i=1}^N \alpha_i$  can be interpreted as the number of past uncensored data points. The larger this number is, the more confident the firm is about the distribution of consumer tastes, therefore, the less it needs to learn. In the last three rows, we vary the prior distribution while keeping  $\alpha_0$  constant, which means that we change firm's beliefs on the relative likelihood of the four consumer locations. We find that the relationship is not always monotone, but we generally find that the value of learning tends to be higher when the firm's beliefs on the consumers' taste distribution is more diffuse, i.e., when the  $\alpha$  vector is more evenly distributed. This is because, in that case, there is a greater potential for learning compared to the case when the firm is already convinced that most of the customers are in one end of the attribute space.

## 6. Extensions

In this section we discuss three extensions to our model: (i) unknown consumer locations, (ii) heterogenous reservation price, and (iii) unobservable market size.

*Unknown Consumer Locations.* Our model assumes that the firm knows the discrete set of possible consumer locations,  $Y$ , and is learning about the relative likelihood of consumers in the population being at each of these locations. If there were consumer locations that the firm does not know about (outside of  $Y$ ), then it would not be able to learn about their relative likelihood. We suggest two ways of handling this issue.

First, one could pick a very large set  $Y$ . For most practical purposes, the domain of consumer locations is discrete. Take the example of milk. Consumers can only discern differences in fat content beyond a certain value, which can be used to discretize the attribute space (e.g., 0.5%, 1%, 1.5%, 2%, 2.5%). There are many other limits to human perception, such as ability to differentiate colors, that set ultimate limits on the level of granularity needed on product attributes. By assigning a very small (though positive) mass on each of these potential locations, the firm would be able to learn about their relative probabilities. Although this approach is challenging computationally when  $Y$  is large, the model and insights developed in this paper would still apply.

Alternatively, one can treat the distribution of consumer tastes as continuous. This way, Bayesian updating can determine which locations have higher posterior density. The problem with this approach is that there does not exist a procedure to solve for the optimal assortment explicitly. As an approximation, one could divide the attribute space into a fixed number of disjoint intervals and define  $\theta_j$  as the probability mass over the  $j$ th interval rather than at one point. The assortment problem becomes one of covering intervals with products, i.e., a product can cover one or more adjacent intervals and its price can be set appropriately. Our entire analysis extends to this setting. Alptekinoglu et al. (2012) show that this approximation method performs very well even with a small number of intervals: in most cases, it is possible to achieve 95% of the optimal expected profit in single-period settings by dividing the attribute space into only 10 intervals.

We also test our model for robustness in case of a misspecification of the set of consumer locations. More specifically, we assume that  $Z = \{z_1, \dots, z_S\}$  is the set of consumer locations, but that the firm thinks consumers are located in  $Y = \{y_1, \dots, y_N\}$  where  $Y \subseteq Z$ , that is, the firm is missing some consumer locations. First, note that the firm may capture sales from customers located in  $Z \setminus Y$ , but incorrectly attribute those sales to customers in  $Y$ . For example, suppose that  $Z = \{0.25, 0.5, 0.75\}$  and  $Y = \{0.25, 0.75\}$ . If the firm offers assortment  $([0.25, 0.75])$  and one customer from the central location comes to the store, the firm will sell a unit of the product but will wrongly interpret this sales as coming from a customer located at 0.25 or 0.75.

We numerically evaluate the impact of such misspecification on a firm's expected profit. In period  $t$ , we (i) calculate the optimal assortment  $\mathbf{a} = (a_1, \dots, a_n)$  by solving for  $v_t^c$ , (ii) simulate the demand  $D = (D_1, \dots, D_n)$  using the true distribution of customers, (iii) calculate the realized profit equal to  $-fn + \sum_{i=1}^n (p_i(a_i) - c)D_i$ , and (iv) update the prior distribution as described in §4.4. We do so for two periods, then add up the realized profits to obtain the two-period realized profit value (we found that the insights did not change as we increased the number of periods). We repeat this process over 10,000 instances and calculate the average two-period profit value.

We use  $S = 3$ ,  $N = 2$ ,  $Z = \{0.25, 0.5, 0.75\}$ ,  $\delta = 1$ ,  $\bar{p} = 25$ ,  $d = 40$ ,  $f = 10$ , and  $c = 3$ . The market size is unobservable<sup>5</sup> and can be equal to 1 or 2 with equal probability. For the true distribution of customer tastes we use  $\theta = (\theta_1, \theta_2, \theta_3) = ((36 - p)/72, p/36, (36 - p)/72)$  where  $p \in \{2, 4, 6, 8, 10, 12\}$ . We consider two cases:  $Y = Z$  and  $Y = \{0.25, 0.75\}$ . When  $Y = Z$ , we use  $\alpha = 36\theta$  and when  $Y \subset Z$ , we use  $\alpha = ((36 - p)/2, (36 - p)/2)$ , so that in both cases, the  $\alpha$  vector of the prior distribution is proportional to the  $\theta$  vector from the true distribution. The percentage loss in average two-period profit with  $Y = \{0.25, 0.75\}$  compared to  $Y = Z$  is 0, 0.31, 1.94, 9.32, 15.69, 21.42

when  $p = 2, 4, 6, 8, 10, 12$ , respectively. As  $p$  increases, it becomes more likely that customers are located in the center location. We see that, as expected, the profit loss due to misspecification increases as the missing location becomes more likely. In practice, the firm should have a good idea of where most of the customers are, implying that missing locations should have a low probability. In such instances, we would not expect the impact of misspecification to be significant.

*Heterogenous Reservation Price.* Next, we discuss the assumption of a fixed and known reservation price  $\bar{p}$ . Suppose, instead, that consumers belong to two segments: those with a high ( $H$ ) reservation price,  $\bar{p}^H$ , and those with a low ( $L$ ) reservation price,  $\bar{p}^L$ , with (known) respective probabilities  $\beta$  and  $1 - \beta$ . In this case, one can define an assortment as  $\mathbf{a} = (\mathbf{a}^H, \mathbf{a}^L)$  where  $\mathbf{a}^H = (a_1^H, \dots, a_n^H) = ([x_1^H, \bar{x}_1^H], \dots, [x_n^H, \bar{x}_n^H])$  are the high-type market segments and  $\mathbf{a}^L = (a_1^L, \dots, a_n^L) = ([x_1^L, \bar{x}_1^L], \dots, [x_n^L, \bar{x}_n^L])$  are the low-type market segments. The introduction of heterogeneous reservation price makes this problem considerably more complex; however, we show in the e-companion to this paper that our results—in particular Theorem 1—extend to this case, provided we redefine the informativeness order over assortments as follows.

Let  $P(\mathbf{a})$  be a partition on  $Y \times \{H, L\}$ , which contains  $n$  subsets of  $Y \times \{H, L\}$ , each of which corresponds to a distinct group of consumer locations and types covered by one product in  $\mathbf{a}$ , plus possibly one more subset that contains all the remaining consumer locations and types (if any) that are not covered. For example, if there are two consumer locations and  $\mathbf{a}^H = ([y_1, y_2])$  and  $\mathbf{a}^L = ([y_1, y_1])$ , then  $P(\mathbf{a}) = \{(y_1, L), (y_1, H), (y_2, H)\}, \{(y_2, L)\}$ .

**DEFINITION 3.** Assortment  $\mathbf{a} = (\mathbf{a}^H, \mathbf{a}^L)$  is “more informative” than assortment  $\mathbf{a}' = (\mathbf{a}'^H, \mathbf{a}'^L)$ , or  $\mathbf{a} \succeq_I \mathbf{a}'$ , if  $P(\mathbf{a})$  is a finer partition than  $P(\mathbf{a}')$ .

This definition generalizes our previous definition of informativeness. For an assortment to be more informative, implied partitions on locations and types need to be finer.

*Unobservable Market Size.* Our model assumes that the firm is able to observe the market size together with the consumer taste vector (in the uncensored information case) or the sales vector (in the censored information case). In practice, the firm may not be able to count the number of consumers who come to the store but leave empty-handed (this is especially true for brick-and-mortar stores). Having to infer consumer tastes from sales only, the firm needs to take into account the additional uncertainty created by the market size. For example, with three consumer locations, if the market size is observable,  $([y_1, y_1], [y_2, y_2])$  is as informative an assortment as  $([y_1, y_1], [y_2, y_2], [y_3, y_3])$ , because the firm can infer the number of consumers at location  $y_3$  even if it is not covered. However, when the market size is unobservable, the former is strictly less informa-

tive, which makes it necessary to redefine the notion of informativeness.

**DEFINITION 4.** Assortment  $\mathbf{a} = (a_1, \dots, a_n)$  is (strictly) “strongly more informative” than assortment  $\mathbf{a}' = (a'_1, \dots, a'_n)$ , or  $(\mathbf{a} \succ_{SI} \mathbf{a}') \mathbf{a} \succeq_{SI} \mathbf{a}'$ , if  $P(\mathbf{a})$  is a (strictly) finer partition than  $P(\mathbf{a}')$ , and assortment  $\mathbf{a}$  covers more locations than assortment  $\mathbf{a}'$ , or  $\bigcup_{i=1}^n a_i \supseteq \bigcup_{i=1}^n a'_i$ .

We show in the e-companion to this paper that our results—in particular Theorem 1—extend to the case of unobservable market size, provided we redefine the informativeness order over assortments as in Definition 4. Clearly, this definition implies our previous definition (Definition 2) and imposes a stronger partial order. An assortment  $\mathbf{a}$  that is more informative than  $\mathbf{a}'$  under the previous definition is not necessarily more informative than  $\mathbf{a}'$  under this definition. Increasing the number of products within a group in Figure 2 increases the informativeness when the market size is unobservable, which can be shown to hold in general. Moreover, a strongly more informative assortment contains higher variety (see Lemma EC.3 in the e-companion).

## 7. Conclusion

In this paper, we develop a model of dynamic assortment planning for horizontally differentiated products when a retailer is learning about consumer tastes. We study how the firm should adjust assortments over time so as to learn consumer tastes better. The classic exploration-exploitation trade-off is present in our model: the firm may choose to forgo immediate profits to learn about consumer tastes, which would help the firm in the coming periods.

We model horizontal differentiation on a Hotelling line where there is a finite number of locations (discrete points) that represent consumers’ ideal products or tastes. A firm offers an assortment of products on the Hotelling line, setting their locations and prices, and paying a fixed cost per product. Each consumer has a choice to buy one unit of a product that the firm offers or opt for an outside option. We study the firm’s assortment decisions in single- and multiperiod settings. We highlight especially the learning aspect of assortments: over multiple periods, with dynamically changing assortments, the firm is able to infer consumer locations (perfectly or imperfectly) from past sales data, learn their tastes better, and use this information to fine-tune its offering.

First, we show that the firm’s product location and pricing problem in the single-period case simplifies to a search for the optimal product market segments and that there is only a finite number of possible assortments to consider.

Next, we introduce a multiperiod model with learning, and study the optimal assortment and pricing decisions in this dynamic setting. We emphasize how sales of a given assortment may provide the firm with censored information on consumer tastes. Intuitively, if the firm caters



to many types of consumers with a single product, consumers substitute away from their ideal product and the firm receives censored information about their tastes. On the other hand, if a firm offers many dedicated products, fewer consumers need to substitute away from ideal products and the firm may gather better information about consumer tastes. We formalize the notion of how informative an assortment is and construct a partial order based on it. Depending on what the firm can infer about consumer tastes from past sales data, some assortments are more informative than others (e.g., at the extreme, offering a product targeted for each possible consumer location is the most informative assortment). Using our partial order, we establish the following result: it cannot be optimal for the firm to select a less informative assortment than the optimal assortment under uncensored information (i.e., myopically optimal assortment). This is reminiscent of the “stock more” result shown in censored newsvendor settings where sales are observed, but not demand (see for example Ding et al. 2002). This result is robust as it extends to the case of unobservable market size and heterogeneous reservation price provided that one redefines the informativeness order of assortments carefully.

As an important special case of our multiperiod model, we also develop a family of conjugate learning models, using the Dirichlet and Extended Dirichlet distributions. We then present a number of numerical examples using these conjugate models and highlight interesting properties of the optimal dynamic assortments and value of learning over multiple periods. For instance, it is possible that the optimal assortment may bring losses to the firm in early periods, a phenomenon entirely driven by the firm’s desire to learn about consumer tastes.

We believe that the concept of dynamic assortments offers rich analytical and empirical challenges to researchers. For instance, exploring demand learning issues with alternative consumer choice models (e.g., preference ranking models) and possibly for products with multiple attributes would expand the applicability of insights gathered in the literature so far. Competition between firms learning through dynamic assortment also remains largely unexplored (in contrast, competition with “static assortment” has been studied, e.g., Alptekinoğlu and Corbett 2008). Building on our model, we think allowing consumer heterogeneity on transportation cost, modeling nonstationary consumer preferences, or treating the case where the firm learns about consumer tastes and the market size simultaneously, are all interesting and challenging problems. Another interesting avenue for future research is to consider the case where the firm holds inventory of the products and is not able to observe lost sales. In this case, there are two types of censoring: the censoring of consumer tastes due to (assortment-based) substitution and the censoring of the market size and product demands due to the stockouts, and it will be interesting to study the interaction between them.

## Electronic Companion

An electronic companion to this paper is available as part of the online version at <http://dx.doi.org/10.1287/opre.1120.1067>.

## Endnotes

1. Not every partition of  $Y$  is feasible, because each market segment is a closed interval on  $\Omega$  (Lemma 1). In a partition induced by an assortment, we can have at most one subset with nonadjacent locations; that subset is the one with consumer locations (if any) that are not covered.
2. Chen and Plambeck (2008) show that in a multiperiod Bayesian inventory management problem with nonperishable inventory, the optimal inventory level may be lower than the myopic level because reducing inventory helps mitigate risk of overstocking in subsequent periods.
3. Because of this partial order “not less informative” is not the same as “at least as informative.”
4. Each column has at least two 1s, and column  $k$  means that the subset of locations  $Y_k \equiv \{y_j: Z_{jk} = 1\}$  belongs to at least one feasible partition  $P \in \mathcal{P}$ . Column  $k$  is relevant only when the assortment contains a product that exactly covers all locations in  $Y_k$  or when the assortment leaves all locations in  $Y_k$  not covered.
5. If the market size was observable, the firm who uses  $Y \subset Z$  would understand that some consumer locations are missing when it believes its current assortment covers the whole market but obtains a total sales lower than  $m$ .

## Acknowledgments

The authors are grateful for helpful comments provided by Jim Smith, Sridhar Seshadri, and the anonymous associate editor and referees.

## References

- Alptekinoğlu A, Corbett CJ (2008) Mass customization vs. mass production: Variety and price competition. *Manufacturing Service Oper. Management* 10(2):204–217.
- Alptekinoğlu A, Corbett CJ (2010) Leadtime-variety tradeoff in product differentiation. *Manufacturing Service Oper. Management* 12(4):569–582.
- Alptekinoğlu A, Honhon D, Ulu C (2012) Positioning and pricing of horizontally differentiated products. Working paper, Southern Methodist University, Dallas.
- Aviv Y, Pazgal A (2005) Dynamic pricing of short life-cycle products through active learning. Working paper, Washington University, St. Louis.
- Bernstein F, Kök AG, Xie L (2011) Dynamic assortment customization with limited inventories. Working paper, Duke University, Durham, NC.
- Berry DA, Fristedt B (1985) *Bandit Problems: Sequential Allocation of Experiments* (Chapman & Hall, London).
- Caldentey R, Caro F (2010) Dynamic assortment planning. Working paper, New York University, New York.
- Caro F, Gallien J (2007) Dynamic assortment with demand learning for seasonal consumer goods. *Management Sci.* 53(2):276–292.
- Caro F, Martínez-de-Albeniz V (2012) Product and price competition with satiation effects. *Management Sci.* 58(7):1357–1373.
- Chen F, Eliashberg J, Zipkin P (1998) Customer preferences, supply-chain costs, and product-line design. Ho T-H, Tang CS, eds. *Product Variety Management: Research Advances* (Kluwer Academic Publishers, Norwell, MA), 123–144.

- Chen L, Plambeck EL (2008) Dynamic inventory management with learning about the demand distribution and substitution probability. *Manufacturing Service Oper. Management* 10(2):126–156.
- Dickey JM, Jiang J, Kadane JB (1987) Bayesian methods for censored categorical data. *J. Amer. Statist. Assoc.* 82(399):773–781.
- Ding X, Puterman ML, Bisi A (2002) The censored newsvendor and the optimal acquisition of information. *Oper. Res.* 50(3):517–527.
- Gaur V, Honhon D (2006) Product variety and inventory decisions under a locational consumer choice model. *Management Sci.* 52(10):1528–1543.
- Harpaz G, Lee WY, Winkler RL (1982) Learning, experimentation, and the optimal output decisions of a competitive firm. *Management Sci.* 28(6):589–603.
- Hotelling H (1929) Stability in competition. *Econom. J.* 39(153):41–57.
- Kök AG, Fisher M, Vaidyanathan R (2009) Assortment planning: Review of literature and industry practice. Agrawal N, Smith SA, eds. *Retail Supply Chain Management* (Springer, New York), 99–153.
- Lancaster K (1966) A new approach to consumer theory. *J. Political Econom.* 74(2):132–157.
- Lu X, Song J, Zhu K (2005) On “the censored newsvendor and the optimal acquisition of information.” *Oper. Res.* 53(6):1024–1027.
- Olivares O, Lu Y, Musalem A, Schilkut A (2011) Measuring the effect of queues on customer purchases. Working paper, Columbia University, New York.
- Rusmevichientong P, Shen M, Shmoys D (2010) Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Oper. Res.* 58(6):1666–1680.
- Saure D, Zeevi A (2011) Optimal dynamic assortment planning with demand learning. Working paper, Columbia University, New York.

---

**Canan Ulu** is an assistant professor in the Information, Risk and Operations Management Department at the McCombs School of Business, University of Texas at Austin. Her research interests include Bayesian learning in sequential decision problems and the impact of behavioral decision theory on decision analysis methods.

**Dorothee Honhon** is an assistant professor in the Department of Industrial Engineering at the Eindhoven University of Technology in the Netherlands. Her research interests include inventory management, assortment planning, retail operations, behavioral operations management, and transportation theory.

**Aydin Alptekinoglu** is an assistant professor of operations management at Cox School of Business, Southern Methodist University. His broad research interest is product strategy. He studies product variety management, assortment planning, customization of goods and services, retail returns, and other aspects of product strategy that lie at the interface of operations and marketing.