International Journal of Research in Marketing 35 (2018) 242–257

Contents lists available at ScienceDirect

IJRM
International Journal of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Full Length Article

A simultaneous model of multiple-discrete choices of variety and quantity☆

Ralf van der Lans

Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

ARTICLE INFO

Article history:
First received on May 8, 2017 and under review for 5 months
Available online 3 January 2018

Senior Editor: Tülin Erdem

Keywords:
Bayesian estimation
Consumer choice
Quantity discounts
Variety

ABSTRACT

In many categories, consumers purchase discrete quantities of multiple varieties. For example, when doing grocery shopping for cereals, consumers may purchase in each category three units of brand A, four of brand B, and one of brand C. These decisions are often influenced by nonlinear pricing strategies such as quantity discounts. Modeling such multiple-discrete choices is challenging, as they violate assumptions of standard choice models. In this research, the author introduces a computationally attractive choice model that simultaneously captures 1) variety, 2) discrete quantity, and 3) nonlinear pricing strategies, such as quantity discounts. The model assumes that consumers maximize variety of the choice outcome, while taking into account constraints on utilities of alternatives. Application of the proposed model to two datasets demonstrates the superior fit compared to several rival models. Counterfactual analyses demonstrate that the model is a valuable tool for assortment and pricing decisions.

1. Introduction

Choices of variety and discrete quantity are common in nearly every product category. For instance, when doing grocery shopping, consumers buy multiple varieties of yogurts, soft drinks and cereals. Likewise, customers of an ice-cream store may buy two scoops of chocolate and one scoop of strawberry. Marketers often stimulate these variety and quantity decisions by nonlinear pricing strategies, such as set menus in restaurants, quantity discounts, bundling promotions and waiving shipping fees for larger orders in online retail stores (Foubert & Gijsbrechts, 2007; Lewis, Singh, & Fay, 2006). Although multiple-discrete choices of variety and quantity are common, and understanding them is vital for optimizing marketing mix decisions, analyzing such data is challenging as the choice model needs to capture the following three important features. First, it needs to incorporate the possibility that consumers choose multiple products at the same time. Second, as in the examples above, demand for quantity is usually discrete. Third, given that the unit price of a product often depends on the quantity and variety chosen, it is important for the choice model to accommodate nonlinear pricing strategies. Standard choice models assume that consumers choose only one unit of one product, and are, thus, not able to analyze variety and quantity choices simultaneously. Although there is an extensive body of research that extended standard choice models, they have become computationally burdensome as the complexity grows

☆ The author would like to thank Xi Chen, Jun Kim, Song Lin, Anirban Mukhopadhyay, Jason Roos, Wenbo Wang, Michel Wedel and Qiang Zhang for their valuable suggestions. He also wishes to thank Ashley To for the data collection of the second empirical application, and Sanghak Lee for providing the data and code of their model for the first application. Special thanks to Jacopo Contiero and Celine Wong for making their ice-cream shop available for this research and the many scoops of delicious ice-cream flavors that provided the calories necessary to run the models. This research was supported by a grant from the Hong Kong Research Grant Council (GRF645511). Finally, the author thanks the editor, senior editor, and three anonymous reviewers for valuable suggestions.

E-mail address: rlans@ust.hk (R. van der Lans).

https://doi.org/10.1016/j.ijresmar.2017.12.007
0167-8116/

exponentially in the number of alternatives. The goal of this research is to develop a new, computationally attractive, simultaneous choice model that is able to capture 1) discrete quantities, 2) variety, and 3) nonlinear pricing strategies, and does not involve a complex optimization strategy.

To capture multiple-discrete choices of variety and quantity, previous research has extended economic demand models of consumer choice (Hanemann, 1984). In these models, consumers are assumed to maximize utility subject to budget constraints. Maximizing utility is a complex task and demands high cognitive effort (Bettman, Luce, & Payne, 1998), which becomes even harder if consumers consider choices for variety and quantity. As a consequence, the computational complexity of extended economic demand models has become nondeterministic polynomial-time hard (NP-hard) (Howell, Lee, & Allenby, 2016; Lee & Allenby, 2014) and can, therefore, only be applied to situations with a limited number of alternatives (Kim, Allenby, & Rossi, 2002). To reduce computational complexity, previous research decomposed the choice task into simpler tasks, such as quantity first and then choice decisions (Dubé, 2004; Foubert & Gijsbrechts, 2007; Gupta, 1988; Harlam & Lodish, 1995). Although these quantity-then-choice models allow for variety and discrete quantity decisions, they assume that consumers solve the different subtasks in isolation, which may lead to suboptimal decisions (Kim, Telang, Vogt, & Krishnan, 2010) and biased parameter estimates (Campo, Gijsbrechts, & Nisol, 2003).

This research introduces an alternative modeling approach for multiple-discrete choices of variety and quantity. The approach is based on theories of consumer choice that show that consumers often prefer variety when choosing more than one product, even if this leads to choosing less preferred alternatives (Ratner, Kahn, & Kahneman, 1999; Zeithammer & Thomadsen, 2013). Hence, instead of the objective to maximize utility, the proposed model assumes that consumers maximize variety of their choice subject to constraints on utility. An important feature of the proposed model is that the variety maximization objective is relatively easy to solve and its computational complexity grows only linearly in the number of alternatives. This makes estimation of the proposed model feasible to large assortments, which is also consistent with the limited computational capabilities of decision makers (Bettman et al., 1998). Two empirical applications demonstrate that the proposed model efficiently captures choices for variety and discrete quantity. The first empirical application focuses on a relatively small choice problem, which allows a comparison with recently developed economic choice models (Lee & Allenby, 2014) as well as quantity-then-choice models. Surprisingly, despite the reduction in computational complexity, these model comparisons (both in-sample and holdout prediction samples) demonstrate that the proposed model predicts choices much more accurately. In a second empirical application, we demonstrate that the proposed model can be applied to situations that are infeasible for existing economic choice models, containing many choice alternatives and a nonlinear pricing strategy. Counterfactual analyses in this empirical application demonstrate that the proposed methodology is a useful tool for strategic assortment and (nonlinear) pricing optimization decisions.

2. Literature review of models for variety and quantity choices

Modeling consumer choice can be challenging, especially when choice data is lumpy, containing many zeros (corner solutions) and discrete purchase quantities (Chintagunta & Nair, 2011). As a consequence, researchers need to make simplifying assumptions about consumers' choice rules. Previous research mostly assumed that consumers assign utilities to all choice alternatives and choose the alternative with maximum utility, as in the multinomial logit/probit model (McFadden, 1974), or all products for which the utility passes a certain threshold, as in the multivariate probit model (Duvvuri, Ansari, & Gupta, 2007; Gentzkow, 2007; Manchanda, Ansari, & Gupta, 1999). Although these models do not allow for purchase quantity decisions, it is possible to define each possible choice bundle as a single outcome of a choice model (Bradlow & Rao, 2000; Niraj, Padmanabhan, & Seetharaman, 2008; Russell & Petersen, 2000). This approach also allows for quantity discounts, but the number of possible choice bundles grows exponentially in the number of choice alternatives and quantities, enabling feasible analysis of choices only for few alternatives and small quantities.

A popular approach to model quantity decisions is due to Hanemann (1984), and is based on economic demand models. In these models, consumers maximize utility subject to a budget constraint. Consumers are assumed to optimally spend their budget on at most one of the choice alternatives and an outside good. This model has been extensively applied in the marketing literature (Arora, Allenby, & Ginter, 1998; Chiang, 1991; Chintagunta, 1993). However, these demand models have certain limitations, as they allow for the choice of at most one product (in addition to the outside good), and thus assuming that products are perfect substitutes. Furthermore, they assume that quantity decisions are continuous (Chintagunta & Nair, 2011, p. 981) and they do not allow for nonlinear pricing strategies.

Several researchers have extended the basic versions of consumer demand models to accommodate these limitations. First, continuous quantity decisions of multiple products have been incorporated. For instance, Song and Chintagunta (2007) assumed that consumers shop in multiple categories and choose at most one alternative in each category, while allowing for substitution and complementarity between categories. Others incorporated satiation effects in quantity decisions, which may result in the choice of multiple products (Bhat, 2008; Kim et al., 2002; Wales & Woodland, 1983). Second, previous research extended the basic demand model to allow for discrete quantity decisions. For instance, Lee and Allenby (2014) developed a direct utility model that allows for discrete quantities of multiple products, where the choice of multiple products is due to satiation effects. Based on this model, Lee, Kim, and Allenby (2013) allowed for multiple category choices, in which alternatives within a category are assumed to be substitutes, but between categories are allowed to be complements. Third, specific forms of nonlinear pricing strategies have been incorporated. Allenby, Shively, Yang, and Garratt (2004) incorporated quantity discounts in discrete quantity decisions, but assumes that consumers choose only one product. Howell et al. (2016) allow for variety and quantity decisions and incorporate quantity discounts for situations in which the price per unit of a specific product decreases after a certain quantity threshold. Both models do not allow for bundled promotions or waiving shipping fees, in which discounts are a function of combination of products. Moreover, these models have become increasingly complex, which not only makes them difficult to estimate on choice sets with many alternatives, but also assumes an excessive computational burden on the consumer's cognitive system (Bettman et al., 1998).

To reduce model complexity, previous research has tried to decompose the choice task into simpler subtasks. One popular approach is to assume that purchase quantity and choice are independent decisions (Gupta, 1988). Follow up research extended this approach and allowed for correlated error terms between the quantity and choice decisions (Bell, Chiang, & Padmanabhan, 1999; Krishnamurthi & Raj, 1988; Zhang & Krishnamurthi, 2004). A second approach assumes that consumers first make quantity decisions, and conditional on quantity make choice decisions. For instance, Harlam and Lodish (1995) assumed that quantity is given, and that consumers make repeated choices that follow a logit model. Dillon and Gupta (1996) extended this quantity-then-choice approach and modeled purchase quantity by a zero-truncated Poisson distribution. Foubert and Gijsbrechts (2007) extended this approach to capture nonlinear pricing strategies for bundle promotions. Similarly, Dubé (2004) modeled the number of future consumption occasions as a Poisson process and developed a demand model with discrete quantity decisions that conditions on these occasions (see also Hendel, 1999). These approaches have simplified the choice process while allowing for multiple-discrete choices of variety and quantity. However, they all decompose the choice problem into different subtasks that are solved separately under different constraints and assumptions, which may lead to suboptimal decisions (Kim et al., 2010).

3. Model development

This section introduces a new approach to capture multiple-discrete choices of quantity and variety, which assumes that consumers maximize the variety of the choice outcome, while taking into account the utility of choice alternatives. This objective is based on theories of consumer choice that show that consumers often seek variety when choosing multiple options from a choice set (Ratner et al., 1999; Ratner & Kahn, 2002; Zeithammer & Thomadsen, 2013). There are various explanations for why consumers prefer variety, such as anticipated satiation (Read & Loewenstein, 1995), the need for stimulation and exploring new options (Raju, 1980; Zeithammer & Thomadsen, 2013), uncertainty about future preferences (Simonson, 1990), and social pressure (Ratner & Kahn, 2002). The proposed model formulates the variety maximization goal as a lexicographic optimization problem (Hooker & Williams, 2012; Stidsen, Andersen, & Dammann, 2014), which allows consumers to select their preferred choice bundle, subject to utility constraints. The constraints assure that utilities exceed prices in addition to costs (or benefits) caused by substitution (complementarity) between alternatives in the choice outcome. The proposed optimization problem can be efficiently solved using a sequential optimization algorithm, which allows it to apply the model to large assortments. Moreover, restricted versions of the model enclose several important choice models, such as the multinomial and ordinal probit models as demonstrated in Web Appendix A.

Next, we first introduce an example of a customer in an ice-cream store who decides to choose different scoops of flavors. This example, which is also the context of the second empirical application, involves variety, discrete quantity and a nonlinear pricing strategy. Second, using this example, we introduce the proposed model. Third, an algorithm is presented that efficiently solves the optimization problem of the proposed model. Finally, we discuss identification and estimation of the model.

3.1. Example of a choice problem

Consider a consumer c on shopping trip i visiting an ice-cream store that contains J different flavors. For simplicity, assume that $J = 3$ (Chocolate, Strawberry, and Vanilla), and that the price $p_{cij}(q_{ci})$ of a scoop of flavor j depends on the number of chosen scoops $q_{ci} = \{q_{ci1}, q_{ci2}, q_{ci3}\}$, with $q_{cij} \in \{0, 1, 2, \ldots\}$ representing the number of scoops of flavor j. For example, the price of the first scoop may equal \$2, and the second scoop onwards \$1. In this situation, a cone containing two scoops of chocolate and vanilla, and one scoop of strawberry (i.e., $q_{ci} = \{2,1,2\}$) is priced at \$6, with the corresponding average price of one scoop of flavor j equal to $p_{cij}(q_{ci}) = \frac{6}{\sum_{k=1}^{J} q_{cik}} = \frac{6}{5} = 1.20$. The task of this consumer is to buy a combination of scoops $q_{ci} = \{q_{ci1}, q_{ci2}, q_{ci3}\}$. Although this seems to be a simple example, note that previously discussed economic demand models are not able to capture this situation, as it contains 1) choice of variety, 2) discrete quantities, and 3) nonlinear pricing for a combination of flavors.

3.2. The model

3.2.1. Utility of alternatives

Similar to previous choice models, we assume that consumer c at shopping trip i assigns utilities u_{cij} to each of the J flavors that are defined as follows:

$$u_{cij} = x_{cij}'\beta_c + \varepsilon_{cij}, \forall j \in \{1, \ldots, J\}. \tag{1}$$

In Eq. (1), x_{cij} and β_c are $(K \times 1)$-vectors containing, respectively, observed consumer, shopping trip, product characteristics, and interactions between these factors (x_{cij}) and their corresponding importance (β_c), and ε_{cij} is an error term. In the ice-cream example, x_{cij} could represent dummy variables for each flavor (i.e., $x_{cijk} = 1$, if $k = j$, zero otherwise) and β_{ck} represents consumer c's preference for flavor $k = j$. Because of its flexibility compared to the extreme value distribution, the vector of disturbance terms ε_{ci} is assumed to follow a multivariate normal distribution with mean zero and covariance matrix Σ.

3.2.2. The variety optimization objective

Previous choice models mostly assume that consumers aim to maximize the sum of utilities of chosen alternatives. Under this assumption, the utility structure in Eq. (1) implies that consumers select only one alternative j (Chiang, 1991; Kim et al., 2002). To accommodate variety choices under this assumption, previous research adjusted the utility specification in Eq. (1) to reflect a specific process that leads to variety choices, such as satiation (Bhat, 2008; Kim et al., 2002; Lee & Allenby, 2014) or different future consumption preferences (Dubé, 2004). While these specific utility specifications lead to variety choices, they also strongly increase the computational complexity (Kim et al., 2002). Moreover, there are many other reasons for variety choices. Examples are the need for stimulation and exploring new options (Raju, 1980), social pressure (Ratner & Kahn, 2002), buying for others (Choi, Kim, Choi, & Yi, 2006), a desire to balance attributes across alternatives (Farquhar & Rao, 1976), and higher retrospective evaluations of variety choices (Ratner et al., 1999). Assuming one specific underlying process for variety decisions, may therefore result in a suboptimal model if the assumed underlying process is incorrect. In this research, instead of modifying the utility function (Eq. (1)) to capture a specific process leading to variety choices, we assume that consumers aim to maximize variety prior to utility. This assumption corresponds to empirical observations that show that consumers prefer variety when their purchase volume increases (Simonson & Winer, 1992).

There are several ways to measure variety $V(q_{ci},u_{ci})$ of a choice outcome q_{ci}. Examples are entropy, dispersion and association between attributes of products (Swait & Marley, 2013; van Herpen & Pieters, 2002). These measures assume that each product can be decomposed into attributes, which is difficult in many relevant choice situations including the ice-cream example (Kim et al., 2002, p. 231). Moreover, these measures do not take into account replicates of choice combinations, i.e., consumers perceive less variety when facing a choice bundle containing one scoop of three flavors each, compared to a choice bundle with two scoops of three flavors (Kahn & Wansink, 2004). Therefore, we use the actual variety measure proposed by Kahn and Wansink (2004), which consist of two components: 1) the number of different flavors j, and 2) the number of replicated flavor combinations. This variety measure can be expressed as follows:

$$V(q_{ci}, u_{ci}) = \text{Lex}\quad \max \left(q_{ci(J)}, q_{ci(J-1)}, \ldots, q_{ci(1)}, \sum_{j:q_{cij}>0} \left(u_{cij} - p_{cij}(q_{ci}) \cdot q_{cij} \right) \right). \tag{2}$$

In Eq. (2), $q_{ci(j)}$ is the j-th order statistic of selected quantities, such that $q_{ci(j+1)} \le q_{ci(j)}$, and 'Lex max' represents the lexicographic maximization function (Hooker & Williams, 2012; Stidsen et al., 2014). The lexicographic maximization function entails that the $J + 1$ expressions $(q_{ci(J)}, q_{ci(J-1)}, \ldots, q_{ci(1)}, \sum_{j:q_{cij}>0} (u_{cij} - p_{cij}(q_{ci}) \cdot q_{cij}))$ are maximized in the order in which they appear and that expression j has priority over expression $j + 1$ (Stidsen et al., 2014). In other words, the variety $V(q_{ci},u_{ci})$ of a choice bundle increases if the quantity of the least chosen alternative $q_{ci(J)}$ increases. Subsequently, if the least chosen alternative is maximized, variety increases if the second least chosen alternative $q_{ci(J-1)}$ increases, etc. Finally, given the lexicographically maximized choice quantities q_{ci}, consumers aim to choose those alternatives that maximize utility relative to the price $\sum_{j:q_{cij}>0} (u_{cij} - p_{cij}(q_{ci}) \cdot q_{cij})$. In contrast to previous demand models that multiply quantity and utilities in the utility maximization objective, the utility of the choice outcome u_{cij} only depends on whether an alternative j is selected (i.e., $q_{cij} > 0$), but does not depend on the specific quantity level itself. Thus, in our model, quantity decisions are driven by the variety maximization objective as well as utility constraints, which we will discuss below.

In the example above, assume that the utilities of chocolate, strawberry, and vanilla are, respectively, $u_{ci} = (3.3; 1.75; 3.1)$. A consumer choosing $q_{ci} = \{3,2,3\}$ obtains a variety that equals $V(q_{ci},u_{ci}) = (2,3,3,-0.85)$. This choice outcome has more variety than a bundle containing three scoops of chocolate, one scoop of strawberry and three scoops of vanilla $q_{ci} = \{3,1,3\}$ (i.e., $V(q_{ci},u_{ci}) = (1,3,3,0.15)$). Although the latter choice outcome has a larger utility-price difference, the quantity of the least chosen alternative is lower ($q_{ci(J)}$ equals 1, vs. 2 for the first choice outcome), resulting in a lexicographically lower variety measure (i.e., $(1,3,3,0.15) <_{\text{lex}} (2,3,3,-0.85)$, with $<_{\text{lex}}$ representing lexicographic inequality).

3.2.3. Utility constraints

Without any constraints, the variety maximization objective (Eq. (2)) does not have a solution, because increasing choice quantities always increases variety. Obviously, such solutions are infeasible as they exceed the budget constraints of consumers. Therefore, following standard assumptions of choice models, the utilities of chosen alternatives need to exceed thresholds that are a function of the costs for the consumer. While previous research mostly considered prices to reflect these costs, consumers also face other costs, such as inventory and search costs (Satomura, Kim, & Allenby, 2011). When choosing varieties, consumers incur costs by forgoing a preferred alternative for a least preferred alternative (Ratner et al., 1999). In the context of coupon redemption choices, Bawa and Shoemaker (1987) termed this "substitution costs", and we follow their terminology. In our model, consumers will only increase variety if adding an alternative exceeds the additional substitution costs. Hence, substitution costs may prevent customers from selecting all alternatives, even if their utilities exceed the price constraint. However, in some situations adding variety could lead to benefits if the alternatives are complements (Manchanda et al., 1999). In such situations, substitution costs become negative.

To capture substitution costs between m different flavors, substitution parameter γ_{cm} is introduced, $m \in \{2,\ldots,M\}$, with $M > 1$ and $M \le J$ the maximum order of substitution. This parameter represents consumer c's perceived substitution costs between combinations of m different flavors. In this research, we assume that perceived substitution costs depend on the number of flavors in the choice bundle, but not on the specific flavors in the choice bundle. While this is potentially a strong assumption for categories that contain both substitutes and complements, in many categories substitution effects do not vary strongly if alternatives are similarly priced and branded (Horváth & Fok, 2013; Sethuraman, Srinivasan, & Kim, 1999).[1] Importantly, this assumption significantly reduces computational complexity, as the number of substitution parameters γ_{cm} would otherwise grow exponentially with the number of alternatives. Under these assumptions, the m-th order substitution costs $S_j^{(m)}(q_{cij}, q_{ci,-j}; \gamma_c)$ between q_{cij} scoops of flavor j and $q_{ci,-j} = \{q_{ci1}, \ldots, q_{ci,j-1}, q_{ci,j+1}, \ldots, q_{ciJ}\}$ scoops of all other flavors, depend on 1) the number of different flavors in the choice bundle, and 2) the number of replicated flavor combinations, and is defined as follows:

$$S_j^{(m)}\left(q_{cij}, q_{ci,-j}; \gamma_c\right) = \gamma_{cm} \sum_{j_1 \neq j} \cdots \sum_{j_m \notin \{j, j_1, \ldots, j_{m-1}\}} \min\left(q_{cij}, q_{cij_1}, \ldots, q_{cij_m}\right). \tag{3}$$

Subsequently, the total substitution costs $S_j^{tot}(q_{cij}, q_{ci,-j}; \gamma_c)$ between q_{cij} scoops of flavor j and $q_{ci,-j} = \{q_{ci1}, \ldots, q_{ci,j-1}, q_{ci,j+1}, \ldots, q_{ciJ}\}$ scoops of all other flavors are the sum over all orders m of substitution costs:

$$S_j^{\text{tot}}\left(q_{cij}, q_{ci,-j}; \gamma_c\right) = \sum_{m=2}^{M} S_j^{(m)}\left(q_{cij}, q_{ci,-j}; \gamma_c\right). \tag{4}$$

To illustrate the substitution costs $S_j^{(m)}(q_{cij}, q_{ci,-j}; \gamma_c)$ and $S_j^{\text{tot}}(q_{cij}, q_{ci,-j}; \gamma_c)$ for flavor j in Eqs. (3) and (4), consider an example in which a consumer chooses $q_{ci} = \{2,1,2\}$ and assume that $\gamma_{c2} = 0.25$ and $\gamma_{c3} = -0.10$. Applying Eq. (3) to chocolate ($j = 1$), the second-order substitution costs are $\gamma_{c2} \cdot (\min(q_{ci1}, q_{ci2}) + \min(q_{ci1}, q_{ci3})) = 0.25 \cdot (1 + 2) = 0.75$, and the third-order substitution costs are $\gamma_{c3} \cdot \min(q_{ci1}, q_{ci2}, q_{ci3}) = -0.10 \cdot 1 = -0.10$. Hence, the total substitution costs $S_j^{\text{tot}}(q_{cij}, q_{ci,-j}; \gamma_c)$ between chocolate and the other two flavors are 0.75–0.10 = 0.65 (see Eq. (4)). Similarly, applying Eq. (3) to strawberry ($j = 2$) results in second-order substitution costs of $\gamma_{c2} \cdot (\min(q_{ci2}, q_{ci1}) + \min(q_{ci2}, q_{ci3})) = 0.25 \cdot (1 + 1) = 0.50$, and third-order substitution costs of $\gamma_{c3} \cdot \min(q_{ci1}, q_{ci2}, q_{ci3}) = -0.10 \cdot 1 = -0.10$. Hence, applying Eq. (4) results in total substitution costs between strawberry and the other two flavors of $0.50 - 0.10 = 0.40$. Finally, the substitution costs between vanilla and the other flavors equal that of chocolate.

Taking into account price and substitution costs, the utilities u_{cij} of individual flavors and the sum of all utilities in the choice bundle $\sum_{j: q_{cij} > 0} u_{cij}$ need to satisfy the following constraints:

$$u_{cij} \geq p_{cij}(q_{ci}) \cdot q_{cij} + S_j^{\text{tot}}\left(q_{cij}, q_{ci,-j}; \gamma_c\right), \forall j : q_{cij} > 0, \tag{5}$$

$$\sum_{j: q_{cij} > 0} u_{cij} \geq \sum_{j=1}^{J} p_{cij}(q_{ci}) \cdot q_{cij} + \sum_{j=1}^{J} \sum_{m=2}^{M} \frac{S_j^{(m)}\left(q_{cij}, q_{ci,-j}; \gamma_c\right)}{m}. \tag{6}$$

Eq. (5) implies that, if flavor j is chosen ($q_{cij} > 0$), the utility u_{cij} of flavor j needs to exceed the price $p_{cij}(q_{ci}) \cdot q_{cij}$ of q_{cij} scoops of flavor j in addition to the perceived substitution costs between flavor j and the other flavors in the choice bundle. Similarly, Eq. (6) implies that the sum of utilities of selected flavors needs to exceed the total price $\sum_{j=1}^{J} p_{cij}(q_{ci}) q_{cij}$ in addition to the total substitution costs between all possible combinations of flavors. To avoid double counting substitution costs, all m-th order substitution costs are divided by m.

Inequalities (5) and (6) capture two important features of multiple discrete choices: 1) substitution and complementarity between alternatives, and 2) a rationale for providing quantity discounts. First, to capture substitution and complementarity, an important property of inequality (6) is that is always satisfied if inequality (5) is satisfied and $\gamma \geq 0$. However, if alternatives are complements (i.e., $\gamma < 0$), satisfying Eq. (5) does not imply that Eq. (6) is satisfied. This captures potential indifferences between no-choice and choice of two (or more) alternatives, an important property of complementarity (Gentzkow, 2007). To demonstrate[2] this, imagine that the utilities of two alternatives j and j' satisfy inequalities (5) but not inequality (6). If the utility (price) of alternative j increases (decreases) such that inequality (6) is satisfied, this consumer will purchase both alternatives j and j'. Hence, the choice of alternative j' depends on the utility and price of alternative j, which reflects complementarity (Gentzkow, 2007; Manchanda et al., 1999). Second, while previous research assumed satiation effects to rationalize quantity discounts, our model rationalizes quantity discounts through the utility thresholds (5) and (6). These thresholds capture the change in marginal costs caused by quantity discounts (Lewis et al., 2006). Moreover, while choosing higher quantities results in higher utility thresholds, utilities u_{cij} do not change as a function of purchase quantity. This assumption is in line with independent models of quantity and choice, which modeled quantity decisions using ordinal regression models (Gupta, 1988).

To illustrate the constraints on utilities (5) and (6) numerically, consider again the example with $\gamma_{c2} = 0.25$ and $\gamma_{c3} = -0.10$, and $q_{ci} = \{2, 1, 2\}$. According to Eq. (5), the utilities of chocolate and vanilla need to exceed $p_{cij}(q_{ci}) \cdot q_{cij} + S_j^{\text{tot}}(q_{cij}, q_{ci,-j}; \gamma_c) = (1.20 \cdot 2) + 0.65 = 3.05$, and that of strawberry $(1.20 \cdot 1) + 0.40 = 1.60$. Subsequently, Eq. (6) indicates that the sum of utilities

[1] In both empirical applications, alternatives are similarly priced and from the same brand.
[2] See Web Appendix A for a detailed graphical illustration of inequalities (5) and (6) and how they capture substitution and complementarity.

needs to exceed $\sum_{j=1}^{J} p_{cij}(q_{ci})q_{cij} + \sum_{j=1}^{J}\sum_{m=2}^{M}\frac{S_j^{(m)}(q_{cij},q_{ci,-j};\gamma_c)}{m}$. The price of the choice outcome equals $\sum_{j=1}^{J} p_{cij}(q_{ci})q_{cij} = 1.20\cdot(2+1+2) = 6$. The sum of all second-order substitution costs equals $\sum_{j=1}^{J}\frac{S_j^{(2)}(q_{cij},q_{ci,-j};\gamma_c)}{2} = \frac{0.75+0.50+0.75}{2} = 1$, and the sum of all third-order substitution costs equals $\sum_{j=1}^{J}\frac{S_j^{(3)}(q_{cij},q_{ci,-j};\gamma_c)}{3} = \frac{-0.10-0.10-0.10}{3} = -0.10$. Hence, the sum of utilities needs to exceed 6.90.

3.2.4. The maximization problem

Combining Eqs. (1) to (6), customer c solves the following lexicographic optimization problem to select q_{ci} at shopping trip i:

$$\begin{aligned} &\text{Lex max}\left(q_{ci(J)}, q_{ci(J-1)}, \ldots, q_{ci(1)}, \sum_{j:q_{cij}>0}\left(u_{cij} - p_{cij}(q_{ci})\cdot q_{cij}\right)\right)\\ &\text{subject to } u_{cij} \geq p_{cij}(q_{ci})\cdot q_{cij} + S_j^{\text{tot}}\left(q_{cij}, q_{ci,-j};\gamma_c\right), \forall j: q_{cij}>0,\\ &\qquad \sum_{j:q_{cij}>0} u_{cij} \geq \sum_{j=1}^{J} p_{cij}(q_{ci})\cdot q_{cij} + \sum_{j=1}^{J}\sum_{m=2}^{M}\frac{S_j^{(m)}\left(q_{cij}, q_{ci,-j};\gamma_c\right)}{m},\\ &\qquad q_{cij}\in\{0,1,2,\ldots\}, \forall j = 1,\ldots,J. \end{aligned} \tag{7}$$

The optimization problem in Eq. (7) requires customers to simultaneously maximize variety $V(q_{ci},u_{ci})$ while taking into account constraints on utilities of alternatives. This objective follows quantity-then-choice models that have been widely adopted in the literature to model multiple-item choices (e.g., Bucklin & Gupta, 1992; Dillon & Gupta, 1996; Dubé, 2004; Foubert & Gijsbrechts, 2007; Harlam & Lodish, 1995; Hendel, 1999). An important difference between the choice principle in Eq. (7) and quantity-then-choice models is that the variety decision is not independent from the utility maximization decision. The variety decision that determines quantities, directly affects utility through the restrictions of the optimization approach, while quantity-then-choice models assume that these are independent. Compared with direct utility models that only focus on utility maximization (Kim et al., 2002; Lee & Allenby, 2014), the proposed model keeps a simple additive linear utility structure as in Hanemann (1984) and Chiang (1991). This keeps the number of parameters manageable and significantly reduces computational complexity, making it applicable to situations with many alternatives (Dubé, 2004). Moreover, the reduction in complexity allows us to estimate a full covariance matrix Σ of the error terms, while many demand models assume independent and identically distributed (i.i.d.) errors (Kim et al., 2002; Lee & Allenby, 2014). Covariances of the error term are important in variety decisions and may capture co-incidence effects, which is "the set of all reasons except purchase complementarity and consumer heterogeneity that could induce joint purchase of items across categories" (Manchanda et al., 1999, p. 98). Examples of co-incidence are economic (spreading costs of a shopping trip across many items), habit, mood, time pressure, physical store environment, or any other unobserved reasons that increases joint purchases.

Finally, Web Appendix A illustrates that restricted versions of the proposed model are equivalent to popular choice models, such as the multinomial probit model (Chintagunta, 1992), specific versions of the multivariate probit model (Gentzkow, 2007; Manchanda et al., 1999), and the multivariate ordinal probit model (Lee & Allenby, 2014).

3.3. An efficient sequential optimization algorithm

Solving multi-objective optimization problems, such as the lexicographic optimization problem in Eq. (7), can be challenging (Ehrgott & Gandibleux, 2000; Stidsen et al., 2014). Interestingly, model (7) allows for a simple and efficient sequential optimization approach that is summarized in Fig. 1. If the r.h.s of Eqs. (5) and (6) are increasing functions of quantity, a condition that is met when total the total price is an increasing function of quantity,[3] this algorithm provides the optimal solution.

The algorithm is initialized by starting with an empty choice bundle $q_{ci} = 0$. After the initialization, the algorithm searches for multivariate binary choice bundles $y_{ci} = \{y_{ci1}, y_{ci2},\ldots,y_{ciK}\}$, $y_{cij} \in \{0,1\}$, with maximum variety that satisfy restrictions (5) and (6). If this binary choice bundle is non-empty (i.e., $\sum_{j=1}^{J} y_{ci}>0$), the algorithm updates choice quantities q_{ci} by adding the binary choice bundle to the current choice quantities (i.e., $q_{ci} + y_{ci}$). Given the updated q_{ci}, the algorithm returns to the multivariate binary choice step and aims to find a new choice bundle y_{ci} with maximum variety that satisfies restrictions (5) and (6). This process continues until y_{ci} is empty, in which case the optimal choice outcome equals q_{ci}. Note that it is relatively straightforward to solve the lexicographic maximization problem in the multivariate binary choice step. After ranking choice alternatives in decreasing order according to their difference between utility and price (i.e., $u_{cij} - p_{cij}(q_{ci}+1)$), it involves checking restrictions (5) and (6) for up to J possible binary choice outcomes. Moreover, under the assumptions that the r.h.s. of inequalities (5) and (6) are increasing functions of quantity, if a specific choice bundle y_{ci} does not satisfy inequalities (5) or (6) in a specific iteration, it won't satisfy this inequality in any subsequent iteration.

[3] The algorithm also does not converge if γ is too negative (i.e., complementary benefits are much higher than the price of adding two alternatives). In the estimation algorithm, γ is drawn from a truncated distribution to avoid such cases (see Web Appendix B).

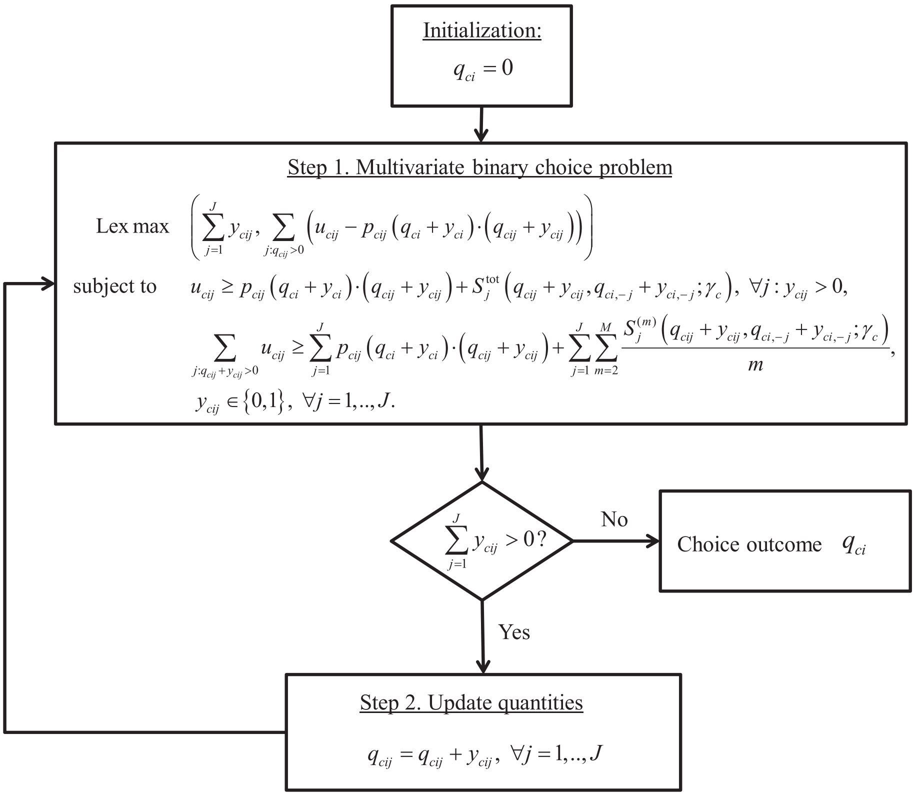

Fig. 1. Sequential optimization procedure.

3.3.1. Illustration of the optimization algorithm

Consider again the ice-cream example with three flavors ($J=3$), and a consumer c at shopping trip i with utility vector $u_{ci}=(3.3;1.75;3.1)$ and substitution parameters $\gamma_{c2}=0.25$ and $\gamma_{c3}=-0.10$. The ranking of the J alternatives in decreasing order in the difference between utilities and prices results in the following order: {1,3,2}. In the multivariate binary choice step of the first iteration, both inequalities (5) and (6) are satisfied after initializing $q_{ci}=0$ and when $y_{ci}=\{1,1,1\}$. In this case, all utilities u_{ci} exceed $p_{cij}(q_{ci}+y_{ci})\cdot(q_{cij}+y_{cij})+S_j^{\text{tot}}(q_{cij}+y_{cij},q_{ci,-j}+y_{ci,-j};\gamma_c)=\frac{4}{3}\cdot 1+0.25(1+1)-0.10\cdot 1=1.73$, and the sum of utilities (8.15) exceeds $\sum_{j=1}^{3} p_{cij}(q_{ci}+y_{ci})\cdot(q_{cij}+y_{cij})+\sum_{j=1}^{3}\sum_{m=2}^{3}\frac{S_j^{(m)}(q_{cij}+y_{cij},q_{ci,-j}+y_{ci,-j};\gamma_c)}{m}=\frac{4}{3}(1+1+1)+\frac{0.5\cdot 3}{2}-\frac{0.1\cdot 3}{3}=4.65$. Using the multivariate binary choice $y_{ci}=\{1,1,1\}$, results in $q_{ci}=\{1,1,1\}$ after updating the choice quantities. In the next multivariate binary choice step of the second iteration, inequality (5) is not satisfied for u_{ci2} and u_{ci3} when $y_{ci}=\{1,1,1\}$ (i.e., the rhs of inequality (5) equals $\frac{7}{6}\cdot 2+0.25(2+2)-0.10\cdot 2=3.13$). However, when $y_{ci}=\{1,0,1\}$, this equality is satisfied by both u_{ci1} and u_{ci3} (i.e., the rhs of inequality (5) equals $\frac{6}{5}\cdot 2+0.25(2+1)-0.10\cdot 1=3.05$). Similarly, inequality (6) is satisfied when $y_{ci}=\{1,0,1\}$ as the sum of utilities (8.15) exceeds the rhs of inequality (6) (i.e., $\frac{6}{5}\cdot 5+\frac{0.75+0.5+0.75}{2}-\frac{0.3}{3}=6.9$). After updating quantities in the second iteration, $q_{ci}=\{2,1,2\}$. In the subsequent multivariate binary choice step of the third iteration, inequality (5) is not satisfied for both u_{ci1} and u_{ci3} when $y_{ci}=\{1,0,1\}$ (i.e., the rhs of inequality (5) equals $\frac{8}{7}\cdot 3+0.25(3+1)-0.10\cdot 1=4.33$). Similarly, for $y_{ci}=\{1,0,0\}$ inequality (5) is not satisfied (i.e., $u_{ci1}=3.3<\frac{7}{6}\cdot 3+0.25(2+1)-0.10\cdot 1=4.15$). Hence, the binary choice step in the third iteration leads to an empty choice set, $y_{ci}=\{0,0,0\}$, which terminates the algorithm and leads to final choice outcome $q_{ci}=\{2,1,2\}$.

3.4. Identification

Similar to probit models, identification of the proposed model needs to consider invariance to location and scale transformations of the parameters and utilities u, common to latent variable models (Albert & Chib, 1993; Gentzkow, 2007). In contrast to the probit model in which the thresholds are parameters, the thresholds in the proposed model are determined by the prices of the alternatives as specified in Eqs. (5) and (6). Consequently, the location of the thresholds cannot be moved or rescaled, and thus the proposed model is identified for location and scale transformations without the need to put any restrictions on the parameters.[4] To illustrate this, consider a market with two products A and B with prices p_A and p_B, respectively. For simplicity,

[4] Identification of the scale is similar to surplus models that set the price coefficient to unity (Jedidi, Jagpal, & Manchanda, 2003).

assume that products A and B are neither complements nor substitutes, such that $\gamma_2 = 0$. In this case, Eqs. (5) and (6) imply that a customer will choose q_A units of product A if its utility u_A satisfies the following conditions $p_A q_A \le u_A < p_A(q_A + 1)$. Similarly, this customer chooses q_B units of product B if the utility u_B of product B satisfies $p_B q_B \le u_B < p_B(q_B + 1)$. Note that identification of the scale makes estimation of the parameters relatively straightforward compared to the probit model, as there are no restrictions on the covariance matrix Σ (Albert & Chib, 1993; McCulloch, Polson, & Rossi, 2000).

The remaining challenge is to identify both the substitution parameters γ and the covariance of the error terms Σ, because both parameters may affect the correlations between quantity decisions q. Interestingly, whereas both parameters affect correlations between quantity decisions, their effects on the sum of quantities $\sum_{j=1}^{J} q_j$ and the probability of no choice vastly differ, which identifies γ and Σ. To illustrate this intuitively, consider again a market with two products A and B. In this market, higher values between the covariance of the error terms σ_{AB} result in higher correlations $r(q_A, q_B)$ between choice quantities q_A and q_B. Similarly, $r(q_A, q_B)$ also increases if γ_2 decreases, because products A and B become stronger complements (or weaker substitutes). In contrast, the sum of choice quantities $q_A + q_B$ is negatively affected by γ_2, whereas the effect of σ_{AB} on $q_A + q_B$ depends on γ_2. If $\gamma_2 = 0$, $q_A + q_B$ is unaffected by σ_{AB}, whereas $q_A + q_B$ is positively (negatively) affected by σ_{AB} if γ_2 is negative (positive), as substitution and complementarity effects are amplified by positively correlated preferences. Another difference is the effect of σ_{AB} and γ_2 on the probability of no choice $p(q_A + q_B = 0)$. Whereas, $p(q_A + q_B = 0)$ increases in σ_{AB}, this probability does not change for different levels of substitution (i.e., positive values of γ_2). However, for negative values of γ_2, $p(q_A + q_B = 0)$ decreases when γ_2 becomes more negative. In this case, products A and B become stronger complements and are therefore more likely to be chosen (i.e., the thresholds to choose both products A and B in Eqs. (5) and (6) become weaker). Table 1 illustrates this in a simulation analysis that compares these three statistics for different combinations of σ_{AB} and γ_2.

In sum, the proposed model (7) is identified if there is sufficient variation of choice outcomes in the data, involving choices of variety and quantity. As discussed by Gentzkow (2007, p. 721), choice models are identified if the number of moments is greater than or equal to the number of free parameters. Consider the market situation of two flavors in which consumers do not choose more than two scoops of each flavor (see also Web Appendix A). In this market, there are nine possible choice outcomes, which correspond to eight moments. In total, this model contains six free parameters (i.e., β_1, β_2, σ_{11}, σ_{22}, σ_{12}, and γ_2), and is identified. Adding more flavors and/or quantities to the model causes the number of moments to increase exponentially. For instance, adding one flavor to this market increases the number of choice outcomes from 3^2 to 3^3. At the same time, the number of parameters grows much slower to eleven (i.e., β_3, γ_3, σ_{33}, σ_{13}, and σ_{23} are added to the model). Thus, increasing the number of flavors J, as well as the number of possible choice outcomes through different choice quantities, increases the degrees of freedom for model estimation.

3.5. Estimation

Due to the advancement of Bayesian estimation techniques, estimating latent variable models, such as the multivariate probit model and the proposed model, has become relatively straightforward (Albert & Chib, 1993; Manchanda et al., 1999). Moreover, because the proposed model does not contain identification restrictions on the covariance matrix Σ, all posterior distributions are standard. This allows for standard Bayesian approaches, which makes the model computationally attractive. Following previous research, a hierarchical structure is formulated to capture heterogeneity, such that $\beta_c \sim N(\mu_\beta, \Omega_\beta)$ and $\gamma_c \sim N(\mu_\gamma, \Omega_\gamma)$. We use standard diffuse conjugate priors for μ_β, μ_γ, Ω_β, Ω_γ and Σ. Substitution parameters γ and utilities u are drawn using the procedure proposed by Neal (2003), which avoids the need to compute truncation points. Web Appendix B presents the MCMC algorithm that was used for model estimation, and Web Appendix C provides the results of a simulation study, which demonstrates that the algorithm is able to recover true parameters accurately. All estimations used 70,000 iterations, thinned 1 in 10 for inference, after a burn-in period of 30,000 iterations. Convergence was assessed using diagnostics presented in Geweke (1992).

An important advantage of the Bayesian estimation procedure is that latent variables are explicitly generated in the MCMC algorithm. These latent variables represent the utilities of consumers in each shopping trip, and can also be generated in situations when flavors are missing (Zeithammer & Lenk, 2006). Such situations are common due to stockouts or retailer decisions (Bruno & Vilcassim, 2008; Campo et al., 2003), which complicates estimation of the covariance matrix Σ. Simulated utilities for missing

Table 1
Comparative statistics[a].

Error covariance σ_{AB}	Statistic	Substitution parameter γ_2		
		−0.5	0	0.5
−0.5	$r(q_A, q_B)$	0.00	−0.24	−0.29
	$q_A + q_B$	0.79	0.67	0.64
	$p(q_A + q_B = 0)$	0.46	0.48	0.48
0	$r(q_A, q_B)$	0.42	0.00	−0.20
	$q_A + q_B$	0.90	0.67	0.60
	$p(q_A + q_B = 0)$	0.51	0.53	0.53
0.5	$r(q_A, q_B)$	0.76	0.36	−0.01
	$q_A + q_B$	1.01	0.67	0.54
	$p(q_A + q_B = 0)$	0.57	0.59	0.59

[a] The comparative statistics are based on 1 million simulations, with preferences $\beta_A = \beta_B = 0.4$ and $\sigma_A = \sigma_B = 1$.

Table 2
Descriptive statistics yogurt purchases: empirical application 1.

Flavor	Blueberry	Strawberry	Vanilla	Raspberry	Key Lime	Peach
Average price	0.72	0.72	0.72	0.72	0.72	0.72
Purchase incidence	541	455	420	474	383	393
Purchase quantity	1074	774	1025	840	813	739
Stockout/unavailable	337	193	370	244	457	351
Zero	5046	5132	5167	5113	5204	5194
One	243	240	157	274	170	183
Two	192	151	133	127	122	137
Three	41	33	34	25	36	27
Four and above	65	21	96	48	55	46

flavors are not used in the choice decision (i.e., consumers are not allowed to choose unavailable flavors, even if these are preferred), but can be used in computing the covariance matrix Σ, which makes its posterior distribution standard. This procedure is applied in both empirical applications to account for stockouts and assortment decisions.

4. Empirical application 1: household panel data

In the first application, the proposed model is validated against advanced economic demand models and quantity-then-choice models. The model validation is performed on the 6-ounce Yoplait yogurt purchases from the IRI household panel dataset (Bronnenberg, Kruger, & Mela, 2008). This dataset involves linear pricing and contains variety and quantity choices of households among six different yogurt flavors, making it feasible to estimate advanced demand benchmark models, as well as quantity-then-choice models. This dataset was also used by Lee and Allenby (2014) to validate their discrete demand model.

4.1. Data description

In total, the dataset contains shopping trips of 109 households, each with at least 28 trips of which at least 5 contained yogurt purchases. Table 2 provides descriptive statistics of the dataset. On average, the price of one unit of each flavor was $0.72 and a household bought on average 3.58 units, conditional on a purchase. However, on most shopping trips (92%), a household did not purchase yogurt. Out of the 5587 shopping trips, 5.8% of the yogurt flavors were unavailable due to stockouts or other reasons. For a more detailed description of the dataset, please refer to Lee and Allenby (2014).

4.2. Model comparisons

Two versions of the proposed model were estimated. In the first version, we followed Lee and Allenby (2014) and assumed that there were no stockouts. Similar to Lee and Allenby, we imputed the mean price for flavors that were unavailable during specific shopping trips. In the second version, we took into account stockouts and used the procedure proposed by Zeithammer and Lenk (2006) to deal with different assortments across shopping trips. In both versions, we only included second-order substitution effects so that the number of parameters equals the number of parameters in Lee and Allenby (2014).

To validate the proposed model, we estimated for both versions four nested models: 1) without substitution costs ($\gamma_2 = 0$), 2) independent error terms, 3) independent and identically distributed (i.i.d.) error terms, and 4) i.i.d. error terms without substitution costs. In addition, we also compared it with three rival models for the first version that ignored stockouts, and one rival model for the second version that takes into account stockouts. First, for both versions, we estimated a quantity-then-choice model. To estimate choice quantity, we used a zero-inflated Poisson model based on the Bayesian procedure proposed by Ghosh, Mukhopadhyay, and Lu (2006) and Damien, Wakefield, and Walker (1999). The parameters of the Poisson model were a function of the average price of flavors on a shopping trip. Second, for the version that ignores stockouts, we compared the proposed model to discrete and continuous demand models (see Lee & Allenby, 2014). Parameters in all models were allowed to be heterogeneous across households. Moreover, we estimated full covariance matrices for the error terms.[5]

To compute model fit, we computed “true” log likelihoods (True LLs, Lee & Allenby, 2014) and hit rates (Manchanda et al., 1999). The “true” log likelihood is a disaggregate fit measure and represents the joint log probabilities of the observed data that can be compared across model specifications. The hit rate is an aggregate level fit measure that compares the number of predicted choice incidences (i.e., presence vs. absence of yogurt flavors in choice outcomes) with observed incidences.[6] Hit rates are between

[5] We also estimated a model with correlated preferences. The fit in the estimation sample remained the same, while the True LL of the holdout sample slightly decreased from −1124 to −1131, and from −1109 to −1116, respectively in the models without and with stockouts.

[6] We focused on choice incidence, because the number of choice outcomes is unlimited. For example, choice outcome $q_{ci} = (2,1,0,5,0,0)$ is transformed to choice incidence (1,1,0,1,0,0), resulting in a total of $2^K = 2^6 = 64$ different choice incidence outcomes. Finally, to assure that hit rates are bounded between zero and one, we divided the sum of absolute differences between predicted and actual choice incidences by two times the number of shopping trips (see Manchanda et al., 1999, Equation 18 on p. 110).

Table 3
Model comparisons: empirical application 1.

Model	True Log-likelihood		Hit rates[b]	
	Estimation sample	Holdout sample	Estimation sample	Holdout sample
Not taking into account stockouts				
Proposed model	−8175	−1124	0.939	0.919
– No substitution costs: $\gamma_2 = 0$	−8247	−1127	0.942	0.921
– Uncorrelated errors	−8444	−1167	0.879	0.859
– i.i.d. errors[a]	−8498	−1173	0.875	0.856
– i.i.d. errors, $\gamma_2 = 0$	−8875	−1224	0.840	0.825
Quantity-then-choice-model	−9381	−1276	0.939	0.915
Discrete demand model	−8691	−1237	–	–
Continuous demand model	−9192	−1278	–	–
Taking into account stockouts				
Proposed model	−8060	−1109	0.939	0.917
– No substitution costs: $\gamma_2 = 0$	−8133	−1112	0.942	0.918
– Uncorrelated errors	−8321	−1152	0.881	0.860
– i.i.d. errors	−8371	−1159	0.878	0.856
– i.i.d. errors, $\gamma_2 = 0$	−8744	−1210	0.843	0.825
Quantity-then-choice-model	−9351	−1308	0.932	0.905

[a] i.i.d. indicates independent and identically distributed.
[b] Lee and Allenby (2014) did not report hit rates for demand models.

zero and one, with higher numbers corresponding to a better model fit. Table 3 presents the model fit statistics of the estimation and holdout samples. The true log likelihood fit measures indicate that taking into account stockouts significantly improves model fit (True LL estimation sample: −8060 vs. −8175; True LL holdout sample: −1109 vs. −1124, respectively for the model with and without taking into account stockouts). Most importantly, the proposed model strongly outperforms all benchmark models both in the estimation sample as well as the holdout sample. Compared to the discrete demand model by Lee and Allenby (2014), model fit strongly improves both in the estimation sample (True LL: −8175 vs. −8691) as well as the holdout sample (True LL: −1124 vs. −1237). As reported by Lee and Allenby, accounting for discreteness improves model fit compared to continuous demand models (estimation sample True LL: −9192; holdout sample True LL: −1278). Quantity-then-choice rival models do not do well, especially compared to the proposed model (estimation sample True LL: −9351 and −9381; holdout sample True LL: −1308 and −1276, respectively for the version with and without stockouts). Finally, the hit rates indicate that both the proposed and the quantity-then-choice rival model predict aggregate choice incidences well (all hit rates are between 0.90 and 0.94, both in the estimation as well as the holdout samples).

Comparisons across the nested model versions, illustrate how different model components contribute to the improved model fit. Clearly, ignoring correlated error terms significantly decreases model fit both at disaggregate and aggregate levels (True LLs: −8321 and −1152; hit rates: 0.88 and 0.86, respectively for the estimation and holdout samples). Finally, a model that assumes i.i.d. error terms without substitution costs has a similar true log likelihood as the discrete demand model. These fit statistics illustrate that the proposed choice model is not only computationally attractive, it also predicts choices much more accurately compared to demand and quantity-then-choice models. Moreover, it also illustrates the importance of capturing both substitution costs as well as co-incidence effects through a full covariance specification of the error terms.

4.3. Estimation results

Table 4 presents median estimated parameters of preferences (μ), heterogeneity of preferences (σ), standard deviations of the error terms (σ_ε) for different flavors, as well as the mean and heterogeneity of the substitution parameter. The estimates show that preferences for strawberry are on average highest ($\mu_\beta = -1.47$), while vanilla is the least preferred ($\mu_\beta = -3.25$). Moreover, there is strong heterogeneity in preferences across households, especially for vanilla ($\sigma = 2.16$), which explains why vanilla is the most frequently chosen yogurt flavor (see Table 2). Interestingly, the substitution parameter μ_γ is negative ($\mu_\gamma = -.07$, 99.9% of posterior draws are negative), suggesting that yogurt flavors complement each other. However, there is heterogeneity across consumers, implying that some consumers perceive yogurt flavors as substitutes, while others perceive them as complements. Moreover, the posterior draws of all error terms covariances are positive, and their medians vary between 0.86 (Blueberry; Raspberry) and 1.64 (Raspberry; Key Lime), illustrating strong co-incidence effects for yogurt purchases.

5. Empirical application 2: Italian ice-cream consumption

An important advantage of the proposed model is that it is computationally attractive, and that it can be applied to situations with many alternatives, involving nonlinear pricing. To illustrate this, the proposed model is applied to sales data form an ice-cream shop that sold 45 different flavors and used a nonlinear pricing strategy as a function of the number of scoops ordered.

Table 4
Estimation results[a]: empirical application 1.

Parameter	Intercept μ	Heterogeneity σ	Error σ_ε
Blueberry	−1.79 (−2.16; −1.46)	1.37 (1.13; 1.66)	1.43 (1.32; 1.55)
Strawberry	−1.47 (−1.78; −1.19)	0.98 (0.80; 1.21)	1.27 (1.15; 1.40)
Vanilla	−3.25 (−3.90; −2.69)	2.16 (1.75; 2.73)	1.86 (1.69; 2.04)
Raspberry	−1.77 (−2.11; −1.47)	1.03 (0.84; 1.27)	1.47 (1.34; 1.62)
Key Lime	−2.80 (−3.39; −2.32)	1.78 (1.44; 2.25)	1.70 (1.55; 1.89)
Peach	−1.93 (−2.33; −1.60)	1.14 (0.93; 1.42)	1.41 (1.28; 1.58)
Substitution/complementarity			
γ	−0.07 (−0.12; −0.02)	0.11 (0.08; 0.14)	–

[a] 95 percent posterior interval between brackets.

This application is also used to illustrate that the proposed model is a powerful tool to perform counterfactual analysis for assortment and pricing decisions.

5.1. Data description

Sales of ice creams have increased substantially in recent years, with annual US sales of 5.8 billion dollar in 2016, a 2.8 percent increase compared to the previous year.[7] Moreover, brands such as Ben and Jerry's, Cold Stone Creamery, and Häagen Dazs increasingly open ice-cream shops around the world, contributing to 80% of Häagen Dazs' revenues in China.[8] The data for this study were collected in an independent Italian ice-cream shop, located in a small harbor village in Southeast Asia. This village attracts tourists and visitors from a nearby city, who dine on seafood, go to the beach, rent a boat, or to take a hike in a nearby country park. Hence, there are few loyal customers, and most customers visit this ice-cream shop only once, which makes identification of heterogeneity across individuals not possible.

During nine weeks on weekdays, for five hours each day, a research assistant recorded the choices for each customer. During this period, 1408 customers visited the shop and they could choose different scoops of ice-cream flavors from an assortment of 20 different flavors that varied across days.[9] In total, the shop sold 45 different flavors divided across 36 different assortments during the period the data were collected. A customer could choose one scoop, two scoops or three scoops of any combination of ice-cream flavors available in the assortment. In addition to this, a customer could also buy other products, such as a large take-home ice-cream box or a cup of Italian espresso. These outside goods were recorded as no choice. During the first eight weeks, the prices of all flavors equaled $3.48, $5.03, and $5.80 for respectively, one scoop, two scoops, and three scoops. However, in week 9, there was an additional quantity discount on all fruit flavors, except durian (a popular local fruit), such that the price for any combination of fruit flavors was $4.51 and $5.29, respectively for two and three scoops.

In addition to quantity discounts, the shop changed two additional marketing mix variables during the observation period. First, the display with ice-cream flavors was each day randomly rearranged in a two by ten matrix-type display layout. To capture the effect of display location, we coded for each available flavor whether it was located in the front (i.e., the ten flavors located closer to the consumer), the right (i.e., the ten flavors located towards the right of the consumer), and center (i.e., four flavors located in the center of the display). Previous research suggests that flavors located in the center may attract more attention and increase the probability of choice (Atalay, Bodur, & Rasolofoarison, 2012). Second, recommendations for specific flavors by the salesperson were coded. The explanatory variables x_{cij} are coded in a (49×1) vector. The first 45 variables are zero, except for element j, which equals one, and measure the preferences for flavors. Variables 46 to 49 are dummy variables that indicate the location of the flavor on the display, i.e., front, right, and center, respectively and whether the flavor was recommended by the salesperson.

Table 5 presents descriptive statistics of choice outcomes and the availability of different flavors. Most customers (918) choose only one scoop, followed by two scoops of different flavors (374). Only few customers (25 in total) choose an ice cream that consisted of two scoops of the same flavor, and there were no customers that choose three scoops of the same flavor. From these choice outcomes it is unclear whether ice-cream flavors are substitutes or complements. On the one hand, many customers choose only one flavor, suggesting substitution. On the other hand, if customers choose two or more flavors, they prefer variety over quantity, indicating complementarity. Table 5 also presents the top ten flavors in terms of sales. Chocolate scoops were sold most frequently (196 scoops of chocolate were chosen during the observation period). Durian was a close second, with 192 scoops sold. However, it is difficult to tell from this data which flavor is most preferred, as the availability of flavors varies. For instance, while chocolate was available to 1352 customers in the sample, durian was only available to 1327 customers. Moreover, the assortments varied across customers. Therefore, important managerial questions are: (1) which are the most preferred flavors, (2) what is the optimal assortment, and (3) what is the optimal pricing strategy? Next, we first discuss the performance

[7] http://www.nielsen.com/us/en/insights/news/2016/whats-the-scoop-digging-in-to-americans-love-of-ice-cream.html.
[8] http://seekingalpha.com/article/1903081-general-mills-moving-ahead-with-continuous-innovation-in-china.
[9] Due to stock replenishments, during four short time periods, the assortment contained 19 flavors, and in another short time period the assortment contained only 18 flavors.

Table 5
Descriptive statistics for Italian ice-cream store: frequency of choice outcomes and availability of top 10 flavors.

Choice outcome	Frequency	Availability
$q_{(1:3)} = \{0,0,0\}$	32	
$q_{(1:3)} = \{1,0,0\}$	918	
$q_{(1:3)} = \{1,1,0\}$	374	
$q_{(1:3)} = \{1,1,1\}$	59	
$q_{(1:3)} = \{2,0,0\}$	20	
$q_{(1:3)} = \{2,1,0\}$	5	
$q_{(1:3)} = \{3,0,0\}$	0	
Total	1408	
1. Chocolate	196 (0.103)[a]	1352
2. Durian	192 (0.101)	1327
3. Pistachio	136 (0.072)	967
4. Mint and chocolate	106 (0.056)	1354
5. Mango	103 (0.054)	1326
6. Coffee	102 (0.054)	1215
7. Yogurt	83 (0.044)	1380
8. Coconut	79 (0.042)	1297
9. Yuzu	77 (0.041)	1134
10. Strawberry	71 (0.037)	1408
Others	753 (0.40)	15,016
Total	1898	27,776

[a] Choice percentages between brackets.

Table 6
Model comparisons: empirical application 2.

Model	True log likelihood		Hit rates	
	Estimation sample	Holdout sample	Estimation sample	Holdout sample
Proposed model	−4126	−1978	0.828	0.686
– No substitution costs: $\gamma_2 = \gamma_3 = 0$	−5682	−2709	0.266	0.212
Quantity-then-choice-model	−4447	−2040	0.619	0.562

of the model compared to several rival models, after that we discuss the estimation results (addressing managerial question 1), and use counterfactual analysis to address managerial questions 2 and 3.

5.2. Model comparisons

As discussed earlier, it is not straightforward to benchmark the proposed model against economic demand models, as the data contain many flavors with a nonlinear pricing strategy in a combination of flavors.[10] Therefore, the proposed model is only compared with a quantity-then-choice model. Quantity-then-choice models (Harlam & Lodish, 1995) are frequently used to deal with bundled promotions that involve nonlinear pricing strategies (Foubert & Gijsbrechts, 2007). Based on Gupta (1988), we used an ordinal probit model for quantity decisions, which takes into account that customers can only choose a maximum of three scoops of ice cream and includes the no choice possibility (quantity equals zero). Moreover, we followed Lewis et al. (2006) and incorporated quantity discounts through the thresholds of the ordinal probit model. For the choice model, we used a probit model with full covariance matrix in which the variance of the first product was fixed to one for identification, and we followed the Bayesian estimation procedure by McCulloch et al. (2000). In addition to the quantity-then-choice rival model, we also estimated one nested version of the proposed model without substitution costs (i.e., $\gamma_2 = \gamma_3 = 0$). For all models, we applied the procedure proposed by Zeithammer and Lenk (2006) to account for different assortments and stockouts.

We randomly divided the data in estimation (965 customers) and holdout samples (443 customers) and computed true log likelihoods and hit rates, as in application 1. Table 6 presents the fit statistics of the model comparisons. The proposed model significantly fits the choice data better, both in the estimation (True LL: −4126 vs. −4447; hit rate: 0.828 vs. 0.619) as well as the holdout sample (True LL: −1978 vs. −2040; hit rate: 0.686 vs. 0.562). Moreover, the nested model fit statistics (True LLs: −5682 and −2709; hit rates: 0.266 and 0.212, respectively for the estimation and holdout samples) demonstrate the importance of taking into account substitution costs, as ignoring this strongly decreased model fit. The results of the proposed model, estimated on 1408 customers, are presented next.

[10] The demand model by Howell et al. (2016) only allows quantity discounts as a function of one alternative, but not as a function of a combination of alternatives, making this demand model not applicable to this situation. Moreover, the complexity of this model grows exponentially in the number of alternatives and discounts.

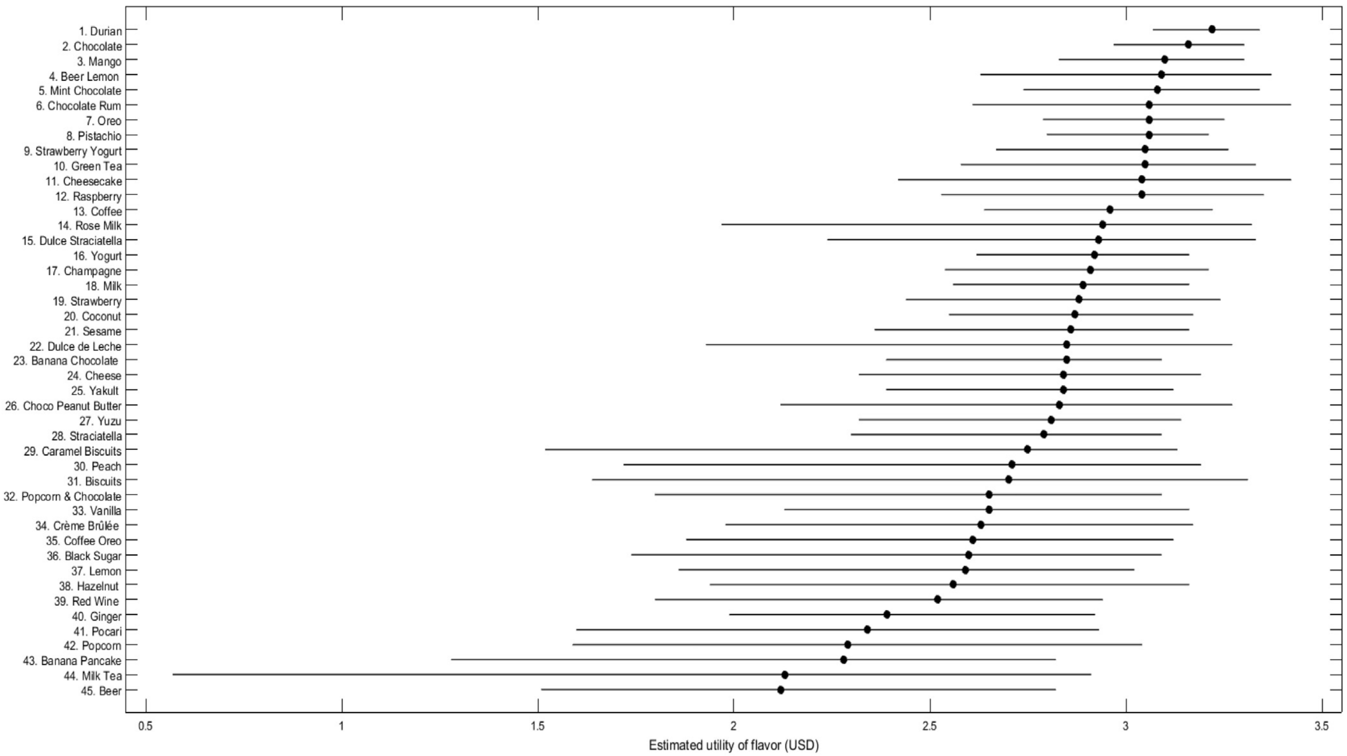

Fig. 2. Estimation results for utilities of Italian ice-cream: median intercepts and 95 percent posterior intervals.

Table 7
Estimation results for Italian ice-cream choices: substitution and marketing mix variables.[a]

Variable	Parameter
Substitution/complementarity	
Between two flavors γ_2	1.52 (1.48; 1.56)
Between three flavors γ_3	−0.74 (−0.80; −0.69)
Influence on direct utility	
Shelf position	
Front	0.01 (−0.02; 0.04)
Right	−0.00 (−0.04; 0.03)
Center	0.02 (−0.01; 0.05)
Recommendation	0.57 (0.45; 0.70)

[a] 95 percent posterior interval between brackets.

5.3. Estimation results

Fig. 2 presents the estimated preferences (β) of all 45 flavors, with the flavors ordered according to their median posterior preferences. Note that the median estimates can be interpreted as the average perceived monetary value in US$ that customers assign to each flavor. Corroborating the descriptive statistics in Table 5, durian and chocolate are the flavors with highest preferences ($\beta_{\text{durian}} = 3.22$ and $\beta_{\text{chocolate}} = 3.16$), although the rank order is reversed, but this difference is not significant (74% of the posterior draws of durian are higher than chocolate). Similar rank order reversals occur between other flavors, such as pistachio that is ranked third according to choice frequency (Table 5) and only eight in terms of estimated preferences. An explanation for these reversals is that the standard deviation of the error terms of preference for chocolate is higher than that of durian ($\sigma_{\text{chocolate}} = 0.74$ vs. $\sigma_{\text{durian}} = 0.64$, 87% of posterior draws of chocolate are higher), similarly for pistachio and mango ($\sigma_{\text{pistachio}} = 0.82$ vs. $\sigma_{\text{mango}} = 0.58$, 96% of posterior draws of pistachio are higher). This suggests that preferences of the pistachio and chocolate flavors vary more across customers compared to durian and mango. Although the difference between standard deviations between pistachio and mango is relatively large, there is relatively little variation in this parameter across all flavors (median estimates between 0.55 and 0.92 for all flavors). Furthermore, most error terms, except for 26 combinations out of 990, are uncorrelated based on 95 percent posterior intervals, suggesting that utilities of flavors do not co-vary systematically. Although differences in standard deviations do not vary much across flavors, median preferences vary substantially from 3.22 for durian to 2.12 for beer flavor. While Fig. 2 provides insights into the most preferred flavors, Table 7 presents the effects of the marketing mix variables (shelf position and recommendation) on the utility of flavors. In contrast to previous findings (Atalay et al., 2012), the effect of shelf position on choice probability is insignificant. However, as expected, recommendations have strong effects on the utility of a flavor ($\beta_{\text{recommendation}} = 0.57$, all posterior draws are positive).

Table 7 also presents the posterior estimates of the substitution parameters. As indicated in this table, the second-order substitution parameter γ_2 is positive (1.52, all posterior draws positive), suggesting that combinations of two ice-cream flavors are substitutes. Moreover, the third-order substitution parameter is negative (median estimate: −0.74, all posterior draws negative). Combining the third and second-order substitution parameters indicates that substitution costs of three ice-cream flavors are $S_j^{tot}(q_{cij}, q_{ci,-j}; \gamma_c) = 2.3$ (i.e., $1.52(1 + 1) - 0.74$, see Eqs. (3) and (4)). These results suggest diminishing substitution costs of increasing levels of variety, as the substitution costs for two flavors (1.52) are relatively high compared to the substitution costs of three flavors (2.3).

5.4. Counterfactual analysis: assortment and pricing strategies

Although it is useful to understand the preference for flavors, the owner of the ice-cream shop ultimately wants to determine a pricing and assortment strategy that increases profit. In order to test this, we obtained variable cost data of producing a scoop of each flavor. The variable production costs are relatively low and stable across flavors, with pistachio the most expensive ($4.72/kg or $0.60 for 1 scoop) and champagne the cheapest ($2.14/kg or $0.27 for 1 scoop).[11] Based on this information, the contribution margin in the observation period equaled $4800. To find out whether it is possible to increase this margin, we did a counterfactual analysis using the proposed model as well as the rival model discussed above. In this counterfactual analysis, we used the estimation results to simulate choice behavior of 100,000 customers under different pricing and assortment conditions. Because there were 45 flavors, the number of possible assortments consisting of 20 flavors is extremely large (i.e., $3.17 \cdot 10^{12}$). Therefore, for both models,[12] we first determined the optimal pricing strategy for the top 20 flavors. To determine the optimal price for scoops of one, two and three flavors, we only considered price increments of $0.25. Given the optimal price, we then determined the optimal assortment by considering the top 23 flavors as predicted by each model, which reduced the number of possible assortments of sizes of 20 flavors to 1771.

[11] The average weight per scoop was 127 g for an ice cream with 1 scoop, 105 g for an ice cream with 2 scoops, and 86 g for an ice cream with 3 scoops.

[12] Note that it is not possible for the quantity-then-choice rival model to determine an optimal pricing strategy, because the thresholds are modeled as dummy variables and not as a function of specific price levels. I therefore used the existing pricing strategy for this counterfactual analysis.After going through my changes, I would like to see the updated version before it gets published.After going through my changes. I would like to view the updated version, before it gets published.

Table 8
Counterfactual analysis: expected profits (US$).[a]

Model	Expected profits (US$)
Full model	5164 (7.6)
Quantity-then-choice-model	4986 (3.9)

[a] Percentage improvement compared to expected profits (US$ 4,800) of current pricing and assortment strategy in brackets.

The counterfactual analysis revealed that the proposed model is able to increase the contribution margin to $5164 (SD = $22.9), corresponding to a 7.6 percent increase compared to the observed assortments and pricing strategy. The optimal assortment consists of all top 20 flavors in Fig. 2, except strawberry yogurt, which is replaced by banana chocolate. The optimal price for one scoop is similar to the current pricing strategy and equals $3.50, while the optimal price for two and three scoops is about $0.50 lower, respectively $4.50 and $5.25. This illustrates that a nonlinear pricing strategy is profitable and that it is important for choice models to capture these strategies. Because the recommended strategy involves popular flavors and a price reduction for two and three scoops, the increase in profit is due to an increase in consumption. On average, in the recommended strategy, customers choose 1.88 scoops, compared to 1.35 scoops in the current strategy.

Table 8 reports the expected change in contribution margin for the strategy recommendations of the rival model, with the expected contribution margin computed using simulations of the proposed model. The quantity-then-choice model leads to a higher expected contribution margin ($4986, SD = $29.3) than the current strategy, but it is significantly outperformed by the proposed model ($5164, SD = $22.9).

In sum, the second empirical application demonstrates that the proposed model is able to analyze choice scenarios with many alternatives and nonlinear bundling pricing strategies. Moreover, the model uncovers whether choice alternatives are substitutes or complements and it allows for counterfactual analyses to improve pricing and assortment strategies.

6. Discussion

Choice models play a central role in the understanding of consumer decision making, but the analysis of choice data is complex. Choice outcomes are qualitative in nature, and the underlying behavioral process leading to these outcomes is unobserved. Hence, previous research assumed that consumers assign utilities to alternatives in the choice set and that they select alternatives that maximize utility. This research introduces a choice model which assumes that consumers maximize the variety of the choice outcome instead of utility. This approach significantly reduces computational complexity and allows application of the model to situations that involve 1) choice of multiple products, 2) discrete quantity, and 3) nonlinear pricing strategies that are common in practice. Estimation of the proposed model involves standard Bayesian MCMC techniques as posterior distributions are closed form.

Application of the model to two datasets demonstrates that the proposed model outperforms alternative choice models and is applicable to large assortments. The first empirical application illustrates that the proposed model predicts choices more accurately compared with advanced economic choice models as well as quantity-then-choice models. The second empirical application involves choices of ice-cream flavors from a large assortment that contains a nonlinear bundling pricing strategy. This application illustrates that the model is flexible and can be applied to challenging situations that cannot be captured by previously developed economic choice models. Moreover, follow-up counterfactual analysis showed that the model is a powerful tool for developing assortment and pricing strategies. By selecting the most preferred flavors and allowing for quantity discounts, the store is able to increase contribution margins by 7.6%.

Although the proposed model is flexible and can be applied to many different choice settings, it has some limitations that can be extended in future research. First, in both empirical applications, preference for flavors was modeled separately. In some situations, it is possible to reduce model complexity even further by the incorporation of product attributes to describe choice outcomes (Kim, Allenby, & Rossi, 2007). Second, although the proposed model extended previous literature by allowing combinations of alternatives to be either substitutes or complements, it did not allow the strength of these effects to vary across different combinations of products. It is possible, however, that combinations of some alternatives are substitutes, while others are complements. Future research could extend the modeling approach to allow for different patterns of substitution and complementarity across flavors. To reduce the curse of dimensionality, because the number of substitution parameters grows exponentially in the number of choice alternatives, it is possible to specify a hierarchical model for this parameter following Wedel and Zhang (2004). Third, it is interesting to extend the model with a store choice decision that depends on assortments size. In the counterfactual analysis of the second empirical application, we kept assortment size constant as different assortment sizes may result in a different probability of choice incidence (Zhang & Krishna, 2007). However, smaller assortments may significantly reduce costs and therefore increase potential profits. Finally, researchers could apply the model to answer important managerial questions, such as the cross-category effects of the marketing mix on choice, quantity and bundle promotions.

Web Appendix

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijresmar.2017.12.007.

References

Albert, J. H., & Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, *88*(422), 669–679.

Allenby, G. M., Shively, T. S., Yang, S., & Garratt, M. J. (2004). A choice model for packaged goods: Dealing with discrete quantities and quantity discounts. *Marketing Science*, *23*(1), 95–1008.

Arora, N., Allenby, G. M., & Ginter, J. L. (1998). A hierarchical Bayes model of primary and secondary demand. *Marketing Science*, *17*(1), 29–44.

Atalay, A. S., Bodur, H. O., & Rasolofoarison, D. (2012). Shining in the center: Central gaze cascade effect on product choice. *Journal of Consumer Research*, *39*(4), 848–866.

Bawa, K., & Shoemaker, R. W. (1987). The effects of direct mail coupon on brand choice behavior. *Journal of Marketing Research*, *24*(4), 370–376.

Bell, D. R., Chiang, J., & Padmanabhan, V. (1999). The decomposition of promotional response: An empirical generalization. *Marketing Science*, *18*(4), 504–526.

Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. *Journal of Consumer Research*, *25*, 187–217.

Bhat, C. R. (2008). The multiple discrete-continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions. *Transportation Research Part B*, *42*, 274–303.

Bradlow, E. T., & Rao, V. R. (2000). A hierarchical Bayes model for assortment choice. *Journal of Marketing Research*, *37*(May), 259–268.

Bronnenberg, B. J., Kruger, M. W., & Mela, C. F. (2008). Database paper: The IRI marketing data set. *Marketing Science*, *27*(4), 745–748.

Bruno, H. A., & Vilcassim, N. J. (2008). Structural demand estimation with varying product availability. *Marketing Science*, *27*(6), 1126–1131.

Bucklin, R. E., & Gupta, S. (1992). Brand choice, purchase incidence, and segmentation: An integrated modeling approach. *Journal of Marketing Research*, *39*, 210–215.

Campo, K., Gijsbrechts, E., & Nisol, P. (2003). The impact of retailer stockouts on whether, how much, and what to buy. *International Journal of Research in Marketing*, *20*, 273–286.

Chiang, J. (1991). A simultaneous approach to whether, what and how much to buy questions. *Marketing Science*, *10*(4), 297–315.

Chintagunta, P. K. (1992). Estimating a multinomial probit model of brand choice using the method of simulated moments. *Marketing Science*, *11*(4), 386–407.

Chintagunta, P. K. (1993). Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Science*, *12*(2), 184–208.

Chintagunta, P. K., & Nair, H. S. (2011). Discrete-choice models of consumer demand in marketing. *Marketing Science*, *30*(6), 977–996.

Choi, J., Kim, B. K., Choi, I., & Yi, Y. (2006). Variety-seeking tendency in choice for others: Interpersonal and intrapersonal causes. *Journal of Consumer Research*, *32*(4), 590–595.

Damien, P., Wakefield, J., & Walker, S. (1999). Gibbs sampling for Bayesian non-conjugate and hierarchical models by using auxiliary variables. *Journal of the Royal Statistical Society: Series B*, *61*, 331–344.

Dillon, W. R., & Gupta, S. (1996). A segment-level model of category volume and brand choice. *Marketing Science*, *15*(1), 38–59.

Dubé, J. -P. (2004). Multiple discreteness and product differentiation: Demand for carbonated soft drinks. *Marketing Science*, *23*(1), 66–81.

Duvvuri, S. D., Ansari, A., & Gupta, S. (2007). Consumers' price sensitivities across complementary categories. *Management Science*, *53*(12), 1933–1945.

Ehrgott, M., & Gandibleux, X. (2000). A survey and annotated bibliography of multiobjective combinatorial optimization. *OR-Spektrum*, *22*, 425–460.

Farquhar, P. H., & Rao, V. R. (1976). A balance model for evaluating subsets of multiattributed items. *Management Science*, *22*(5), 528–539.

Foubert, B., & Gijsbrechts, E. (2007). Shopper response to bundle promotions for packaged goods. *Journal of Marketing Research*, *44*(Nov), 647–662.

Gentzkow, M. (2007). Valuing new goods in a model with complementarity: Online newspapers. *The American Economic Review*, *97*(3), 713–744.

Geweke, J. (1992). *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments.* Oxford: Oxford University Press.

Ghosh, S. K., Mukhopadhyay, P., & Lu, J. -C. J. (2006). Bayesian analysis of zero-inflated regression models. *Journal of Statistical Planning and Inference*, *136*, 1360–1375.

Gupta, S. (1988). Impact of sales promotions on when, what, and how much to buy. *Journal of Marketing Research*, *25*(Nov), 342–355.

Hanemann, W. M. (1984). Discrete/continuous models of consumer demand. *Econometrica*, *52*(3), 541–561.

Harlam, B. A., & Lodish, L. M. (1995). Modeling consumers' choices of multiple items. *Journal of Marketing Research*, *32*(4), 404–418.

Hendel, I. (1999). Estimating multiple-discrete choice models: An application to computerization returns. *The Review of Economic Studies*, *66*(2), 423–446.

van Herpen, E., & Pieters, R. (2002). The variety of an assortment: An extension to the attribute-based approach. *Marketing Science*, *21*(3), 331–341.

Hooker, J. N., & Williams, H. P. (2012). Combining equity and utilitarianism in a mathematical programming model. *Management Science*, *58*(9), 1682–1693.

Horváth, C., & Fok, D. (2013). Moderating factors of immediate, gross, and net cross-brand effects of price promotions. *Marketing Science*, *32*(1), 127–152.

Howell, J. R., Lee, S., & Allenby, G. M. (2016). Price promotions in choice models. *Marketing Science*, *35*(2), 319–334.

Jedidi, K., Jagpal, S., & Manchanda, P. (2003). Measuring heterogeneous reservation prices for product bundles. *Marketing Science*, *22*(1), 107–130.

Kahn, B. E., & Wansink, B. (2004). The influence of assortment structure on perceived variety and consumption quantities. *Journal of Consumer Research*, *30*(March), 519–533.

Kim, J., Allenby, G. M., & Rossi, P. E. (2002). Modeling consumer demand for variety. *Marketing Science*, *21*(3), 229–250.

Kim, J., Allenby, G. M., & Rossi, P. E. (2007). Product attributes and models of multiple discreteness. *Journal of Econometrics*, *138*, 208–230.

Kim, Y., Telang, R., Vogt, W. B., & Krishnan, R. (2010). An empirical analysis of mobile voice service and SMS: A structural model. *Management Science*, *56*(2), 234–252.

Krishnamurthi, L., & Raj, S. P. (1988). A model of brand choice and purchase quantity price sensitivities. *Marketing Science*, *7*(1), 1–20.

Lee, S., & Allenby, G. M. (2014). Modeling indivisible demand. *Marketing Science*, *33*(3), 364–381.

Lee, S., Kim, J., & Allenby, G. M. (2013). A direct utility model for asymmetric complements. *Marketing Science*, *32*(3), 454–470.

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

Manchanda, P., Ansari, A., & Gupta, S. (1999). The "shopping basket": A model for multicategory purchase incidence decisions. *Marketing Science*, *18*(2), 95–114.

McCulloch, R. E., Polson, N. G., & Rossi, P. E. (2000). A Bayesian analysis of the multinomial probit model with fully identified parameters. *Journal of Econometrics*, *99*, 173–193.

McFadden, D. (1974). *Conditional logit analysis of qualitative choice behavior.* New York: Academic Press, Inc.

Neal, R. M. (2003). Slice sampling. *Ann. Statist.*, *31*(3), 705–741.

Niraj, R., Padmanabhan, V., & Seetharaman, P. B. (2008). A cross-category model of households' incidence and quantity decisions. *Marketing Science*, *27*(2), 225–235.

Raju, P. S. (1980). Optimum stimulation level: Its relationship to personality, demographics, and exploratory behavior. *Journal of Consumer Research*, *7*(3), 272–282.

Ratner, R. K., & Kahn, B. E. (2002). The impact of private versus public consumption on variety-seeking behavior. *Journal of Consumer Research*, *29*(2), 246–257.

Ratner, R. K., Kahn, B. E., & Kahneman, D. (1999). Choosing less-preferred experiences for the sake of variety. *Journal of Consumer Research*, *26*(1), 1–15.

Read, D., & Loewenstein, G. (1995). Diversification bias: Explaining the discrepancy in variety seeking between combined and separated choices. *Journal of Experimental Psychology: Applied*, *1*(1), 34–49.

Russell, G. J., & Petersen, A. (2000). Analysis of cross category dependence in market basket selection. *Journal of Retailing*, *76*(3), 367–392.

Satomura, T., Kim, J., & Allenby, G. M. (2011). Multiple-constraint choice models with corner and interior solutions. *Marketing Science*, *30*(3), 481–490.

Sethuraman, R., Srinivasan, V., & Kim, D. (1999). Asymmetric and neighborhood cross-price effects: Some empirical generalizations. *Marketing Science*, *18*(1), 23–41.

Simonson, I. (1990). The effect of purchase quantity and timing on variety-seeking behavior. *Journal of Marketing Research*, *27*(May), 150–162.

Simonson, I., & Winer, R. S. (1992). The influence of purchase quantity and display format on consumer preference for variety. *Journal of Consumer Research*, *19*(1), 133–138.

Song, I., & Chintagunta, P. K. (2007). A discrete-continuous model for multicategory purchase behavior of households. *Journal of Marketing Research*, *34*(Nov), 595–612.

Stidsen, T., Andersen, K. A., & Dammann, B. (2014). A branch and bound algorithm for a class of Biobjective mixed integer programs. *Management Science*, *60*(4), 1009–1032.

Swait, J., & Marley, A. A. J. (2013). Probabilistic choice (models) as a result of balancing multiple goals. *Journal of Mathematical Psychology*, *57*(1–2), 1–14.

Wales, T. J., & Woodland, A. D. (1983). Estimation of consumer demand systems with binding non-negativity constraints. *Journal of Econometrics*, *21*, 263–285.

Wedel, M., & Zhang, J. (2004). Analyzing brand competition across subcategories. *Journal of Marketing Research*, *41*(Nov), 448–456.

Zeithammer, R., & Lenk, P. (2006). Bayesian estimation of multivariate-normal models when dimensions are absent. *Quantitative Marketing and Economics*, *4*(3), 241–265.

Zeithammer, R., & Thomadsen, R. (2013). Vertical differentiation with variety-seeking consumers. *Management Science*, *59*(2), 390–401.

Zhang, J., & Krishna, A. (2007). Brand-level effects of stockkeeping unit reductions. *Journal of Marketing Research*, *44*(Nov), 545–559.

Zhang, J., & Krishnamurthi, L. (2004). Customizing promotions in online stores. *Marketing Science*, *23*(4), 561–578.