ARTICLE IN PRESS
JID: EOR [m5G;September 7, 2022;19:26]

European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Stochastics and Statistics

Assortment optimization using an attraction model in an omnichannel environment

Andrey Vasilyev[a,*], Sebastian Maier[b,a], Ralf W. Seifert[a,c]

[a] College of Management of Technology, Ècole Polytechnique Fèdèrale de Lausanne (EPFL), Switzerland
[b] Department of Statistical Science, University College London (UCL), United Kingdom
[c] International Institute for Management Development (IMD), Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 21 October 2021
Accepted 3 August 2022
Available online xxx

Keywords:
Retailing
Omnichannel
Assortment optimization
Discrete choice modeling

ABSTRACT

Making assortment decisions is becoming an increasingly difficult task for many retailers worldwide as they implement omnichannel initiatives. Discrete choice modeling lies at the core of this challenge, yet existing models do not sufficiently account for the complex shopping behavior of customers in an omnichannel environment. In this paper, we introduce a discrete choice model called the multichannel attraction model (MAM). A key feature of the MAM is that it specifically accounts for both the product substitution behavior of customers within each channel and the switching behavior between channels. We formulate the corresponding assortment optimization problem as a mixed integer linear program and provide a computationally efficient heuristic method that can be readily used for obtaining high-quality solutions in large-scale omnichannel environments. We also present three different methods to estimate the MAM parameters based on aggregate sales transaction data. Finally, we describe general effects of the implementation of widely-used omnichannel initiatives on the MAM parameters, and carry out numerical experiments to explore the structure of optimal assortments, thereby gaining new insights into omnichannel assortment optimization. Our work provides the analytical framework for future studies to assess the impact of different omnichannel initiatives.

1. Introduction

Omnichannel retailing is a major trend in modern commerce. Its aim is to create a seamless customer shopping experience by integrating multiple retail channels with each other. One of the most common omnichannel initiatives is buy-online-and-pick-up-in-store (BOPS), also called click-and-collect. It allows customers to use online services to reserve a product for collection in a retail store. Other examples of omnichannel initiatives include providing online customers with in-store inventory availability information, and installing digital help desks in brick-and-mortar stores so that customers can readily access information about the retailer's online store, such as its assortment, prices and delivery options. With regard to supply chain management, one of the most prominent examples of channel integration is fulfilling a customer's online order from a local brick-and-mortar store in order to best leverage available inventories or respond in a timely fashion.

Modern consumers demand a variety of purchasing and delivery options, and retailers have to adapt their services to changes in customer expectations in order not to lose market share. For example, the rivalry between Amazon and Walmart induced the latter to push the BOPS functionality in order to provide customers with an option comparable in convenience to Amazon's same-day delivery but without charging for shipping (Petro, 2020). According to the recent Global Shopper Trends Report by iVend Retail (2019), 81.4% of consumers reported using BOPS, which represents a growth of nearly 30% from last year's survey. The study of Sopadjieva, Dholakia, & Benjamin (2017), which is based on a survey of 46,000 customers who made a purchase between June 2015 and August 2016, found that 73% of participants used multiple channels during their shopping journey compared to 20% of store-only shoppers and 7% of online-only shoppers. It also revealed that omnichannel retailers are more likely to retain customers. In fact, customers who had an omnichannel shopping experience took 23% more repeat shopping trips to the retailer's stores within 6 months after the purchase than those who shopped

* Corresponding author.
E-mail addresses: andrey.vasilyev@epfl.ch (A. Vasilyev), s.maier@ucl.ac.uk (S. Maier), ralf.seifert@epfl.ch (R.W. Seifert).

https://doi.org/10.1016/j.ejor.2022.08.002
0377-2217/

Please cite this article as: A. Vasilyev, S. Maier and R.W. Seifert, Assortment optimization using an attraction model in an omnichannel environment, European Journal of Operational Research, https://doi.org/10.1016/j.ejor.2022.08.002

through a single channel. It is therefore not surprising that the State of Omnichannel Retail report (Brightpearl, 2017) found that 87% of retailers agree that omnichannel is crucial to their business success.

However, implementing omnichannel strategies remains a difficult task for retailers, with estimating the effect of such strategies on demand being one of the key problems. Moreover, since an omnichannel environment is characterized by a high level of integration between retail channels, a change in assortment in one channel affects the demand across all channels, which makes assortment optimization extremely challenging. At the same time, omnichannel retailing is a rather recent research area with a relatively small number of analytical research papers published. In particular, there is a lack of works which consider the problem of demand and choice modeling in an omnichannel environment. Addressing this problem is an integral part of estimating the expected profit of a retailer, yet existing models do not account sufficiently and adequately for the complex nature of omnichannel shopping behavior. Our paper aims to fill this gap in the extant literature.

This paper makes the following contributions. We introduce a discrete choice model called the multichannel attraction model (MAM) that captures the complex shopping behavior of customers in an omnichannel environment. Importantly, the choice probabilities under the MAM are expressed through simple functional forms of the model parameters, making them easily interpretable. We prove that the assortment optimization problem under the MAM can be reformulated as a mixed integer linear program. We propose a heuristic method to approximate its value in case of large-scale problems, and show numerically that our method is extremely efficient in terms of both computational performance and solution quality. We also present three different methods to estimate the MAM parameters based on sales history data, focusing on the case where only limited data are available. Next, we perform a sensitivity analysis of the model parameters in the two-channel case (with online and physical channels), which leads to insights into omnichannel assortment optimization. For example, we show that a product with a relatively high unit profit in one of the channels may not be included in the corresponding optimal assortment, and vice versa – a product with a relatively low unit profit may be offered. We also analyze how the sizes of optimal assortments depend on the ratio of customers whose primary choice is to shop online to those whose primary choice is to go to a retail store, and on the proportion of customers willing to switch from one channel to another in case of absence of a certain product. We find that implementing the BOPS initiative can be unprofitable if the proportion of online customers using BOPS is too large compared to the additional traffic attracted to the offline channel. Finally, we demonstrate the benefits of omnichannel assortment optimization as opposed to optimizing siloed assortments in a multichannel environment.

The remainder of this paper is organized as follows: In Section 2, we review the two main streams of literature related to our research. In Section 3, we present a discrete choice model for omnichannel retailing, referred to as the MAM, and provide the intuition behind it. In Section 4, we formulate the corresponding assortment optimization problem as a mixed integer linear program. An efficient heuristic method for solving the assortment optimization problem for very large numbers of products is provided in Section 5. The subsequent section is devoted to estimating the parameters of the MAM. In Section 7, we describe the impact of implementing widely-used omnichannel initiatives on the MAM parameters and present a numerical study which investigates the structure of optimal assortments. We summarize our contributions and discuss future research directions in Section 8.

2. Theoretical background and related literature

2.1. Related discrete choice models

Since discrete choice modeling is a vast and complex area of research, here we review only the works most relevant to our paper. A good introduction to discrete choice modeling can be found, for example, in Ben-Akiva & Lerman (1985) and Train (2002).

The multinomial logit model (MNL) formulated by McFadden (1973) is one of the most prominent discrete choice models. For clarity and introduction of notation, we provide a short formal description of the MNL. The choice probabilities under this model are derived as follows. Let $\mathcal{N} = \{1, 2, \ldots, n\}$ denote a set of products, $S \subseteq \mathcal{N}$ denote an offered set, and 0 denote the no-purchase alternative. A customer selects either one product from the offered set, or the no-purchase alternative. Each alternative $j \in \mathcal{N} \cup \{0\}$ has utility U_j, which is given by the sum of $\hat{U}_j$, a constant representing the known part of the utility, and ξ_j, which is a Gumbel-distributed random variable representing the unobserved part of the utility. Random variables ξ_j are assumed to be independent and identically distributed (i.i.d.) for all the alternatives, and are commonly considered to be normalized so that their mean is zero and the variance is $\pi^2/6$. Furthermore, it is assumed that each customer selects the alternative with the highest utility among the available choices (the no-purchase alternative is available by default). Then, it can be shown that the probability of a customer selecting product j from the offered set S is

$$\pi_j(S) = \frac{e^{\hat{U}_j}}{e^{\hat{U}_0} + \sum_{k \in S} e^{\hat{U}_k}}. \tag{1}$$

The MNL can be viewed as a special case of the basic attraction model (BAM) developed by Luce (1959). Let v_j represent the “attractiveness” value of product j, and v_0 represent the attractiveness value of the no-purchase alternative. Under the BAM, the probability of a customer selecting product $j \in S$ is the ratio of the attractiveness value of product j to the sum of attractiveness values of all available alternatives, that is

$$\pi_j(S) = \frac{v_j}{v_0 + \sum_{k \in S} v_k}. \tag{2}$$

Clearly, the MNL choice probabilities (1) take the form (2) if we set the attractiveness value of each product j to $v_j = e^{\hat{U}_j}$, and the attractiveness value of the no-purchase alternative to $v_0 = e^{\hat{U}_0}$.

Gallego, Ratliff, & Shebalov (2014) proposed a generalization of the BAM called the general attraction model (GAM). As noted by the authors, the BAM may be too optimistic in estimating recapture probabilities as it ignores the possibility that a customer can choose to buy product $j \in \mathcal{N} \backslash S$ from another vendor or at a later time. For that reason, they modified formula (2) in the following way:

$$\pi_j(S) = \frac{v_j}{v_0 + \sum_{k \in S} v_k + \sum_{i \in \mathcal{N} \backslash S} w_i}, \tag{3}$$

where $w_i \in [0, v_i]$ represents the “shadow attractiveness” value of getting product $i \in \mathcal{N} \backslash S$ from another source. The meaning of the shadow attractiveness is that a customer does not consider the opportunity of buying a product somewhere else as long as it is present in the assortment. If the product is not available, however, a customer can decide to purchase it from another source. Note that under the GAM, the no-purchase probability is $\pi_0(S) = \dfrac{v_0 + \sum_{i \in \mathcal{N} \backslash S} w_i}{v_0 + \sum_{k \in S} v_k + \sum_{i \in \mathcal{N} \backslash S} w_i}$. Further note that the case $w_i = 0 \ \forall i \in \mathcal{N}$ results in the BAM, and the case $w_i = v_i \ \forall i \in \mathcal{N}$ leads to the independent demand model as the choice probability of any product $j \in S$ does not depend on S.

Gallego et al. (2014) also formulated the sales-based linear program (SBLP) for network revenue management under the GAM. The decision variables in the SBLP are sales quantities rather than the offered set. In the case of infinite capacity and a single market segment, the SBLP takes the form of the assortment optimization problem under the GAM. Below we provide the formulation of this problem since we refer to it further in the text. Let r_j be the gross profit per unit of product j and Λ be the total number of customers. Then, the following linear program can be used to find the optimal assortment if the choice probabilities are given by (3):

$$\max_{x} \sum_{j\in\mathcal{N}} r_j x_j \tag{4a}$$

$$\text{s.t. } \frac{\tilde{v}_0}{v_0}x_0 + \sum_{j\in\mathcal{N}} \frac{\tilde{v}_j}{v_j}x_j = \Lambda, \tag{4b}$$

$$\frac{x_j}{v_j} - \frac{x_0}{v_0} \le 0 \qquad \forall j \in \mathcal{N}, \tag{4c}$$

$$x_0, x_j \in \mathbb{R}_{\ge 0} \qquad \forall j \in \mathcal{N}, \tag{4d}$$

where x_j is the sales quantity of product j, $\tilde{v}_j = v_j - w_j$ and $\tilde{v}_0 = v_0 + \sum_{j\in\mathcal{N}} w_j$. Constraint (4b) is the balance constraint, and constraints (4c) are the scale constraints.

Furthermore, the authors noted that the GAM is the limit of the nested logit model in which customers first select a nest constructed from offerings of the same product by different vendors and then select a vendor which offers this product, while assuming that the dissimilarity parameter of products in each nest tends to zero. The nested logit model, introduced by Domencich & McFadden (1975), is another well-known discrete choice model where each choice probability can be decomposed into the product of two standard logit probabilities: the probability that a certain nest is chosen and the probability that a certain alternative is chosen given the nest.

Finally, it is worth mentioning the Markov chain choice model in which the product substitution behavior of customers is represented by transitions in a Markov chain. Blanchet, Gallego, & Goyal (2016) showed that such a model provides a simultaneous approximation to all random utility choice models including the MNL, the nested logit model and mixed MNLs. They also proved that this approximation becomes exact in the case of GAM choice probabilities. However, the number of parameters of the general Markov chain choice model is $(n+1)^2$, where n is the total number of products, and, moreover, there are no interpretable functional forms for the choice probabilities as computing them for a certain assortment requires a matrix inversion where the matrix entries depend on the assortment.

2.2. Choice modeling and assortment optimization in omnichannel retailing

As mentioned earlier, the problem of choice modeling in an omnichannel environment is underrepresented in the literature. Since most of the research conducted on omnichannel is either purely empirical or qualitative, there is a considerable lack of works presenting analytical models. Most of the analytical papers focus on supply chain and inventory management questions rather than on discrete choice modeling or assortment optimization in an omnichannel environment. For example, Schneider & Klabjan (2013) analyzed when known inventory control policies are optimal in the presence of two sales channels. He, Xu, & Wu (2020) developed a newsvendor model of a two-channel retailer that accounts for cross-channel product returns. Several papers studied the effect of ship-from-store operations on the optimal inventory policy (Seifert, Thonemann, & Sieke, 2006), fulfillment policy (Bayram & Cesaret, 2021), or both policies combined (Govindarajan, Sinha, & Uichanco, 2018). However, to the best of our knowledge, very few papers address at least one of the following questions, which we consider jointly in this paper: discrete choice modeling in the presence of *multiple* retail channels; the impact of omnichannel initiatives on demand allocation; and, ultimately, omnichannel assortment optimization.

Cao, So, & Yin (2016) proposed a theoretical framework to analyze the effect of adding the BOPS channel (called “online-to-store” channel) to existing sales channels on demand allocation. A major limitation of their study is that it considers a single-product setup. The authors use utility functions associated with different channels to model customers’ channel choices, where each utility function is a linear function of the following parameters: the product value, the price per unit of the product in the corresponding channel, the delivery cost (in the case of the online channel), as well as certain inconvenience costs and factors. The paper shows that it may not be profitable to implement the BOPS functionality for some products depending on their characteristics.

Gao & Su (2016) presented a somewhat more complex approach to analyze the impact of the BOPS channel on demand allocation. Similar to the work of Cao et al. (2016), the authors model customers’ channel choices using utility functions associated with different channels, but they also account for the cross-selling effect and inventory management considerations. One of the key findings of their paper confirms that not all products are well-suited for the BOPS functionality. However, their model does not account for the product substitution effect in the case when a product is removed from the assortment.

More recently, Harsha, Subramanian, & Ettl (2019a) studied the pricing problem in an omnichannel environment with a chain of brick-and-mortar stores in the presence of other (online) channels. In their paper, the authors first consider a single-product setup where a customer only selects a source to buy the product from. It is assumed that the brick-and-mortar stores are located in geographically distributed zones, meaning that customers’ choices in a zone do not depend on the parameters of other zones. At the zone level, each customer obtains a utility for choosing a channel depending on the product’s price in this channel. The choices of customers are defined by a BAM where all attractiveness values are expressed through positive and strictly increasing functions of price. The price optimization problem is then formulated as a mixed integer linear program. The authors also extend their analysis to a multiproduct setup using a nested attraction model where nests correspond to channels and each nest is comprised of products included in the channel assortment. For their extension, they provide a mixed integer linear program to solve the price optimization problem approximately. In a related paper, Harsha, Subramanian, & Uichanco (2019b) studied the omnichannel price optimization problem in a single-product setup, whilst accounting for both exogenous cross-channel fulfillment flows and inventory constraints.

With regard to multichannel assortment optimization for multiple products, Bhatnagar & Syam (2014) presented an integer program to determine the optimal item allocation for a hybrid retailer that manages both a chain of physical stores and an online store. They found that the retailer’s profitability can be increased by removing products with high carrying costs from the physical stores and making them available exclusively online, thereby reducing the inventory carrying costs. However, their model relies on a number of strong assumptions, including that the demand for each product is a fixed parameter, meaning that the product demands do not depend on the assortment.

A different angle on multichannel assortment optimization was provided by Dzyabura & Jagabathula (2018). They studied the problem of determining the subset of products from the retailer's online channel to offer in the offline channel in order to maximize the aggregate revenue. It is assumed that each product is defined by a set of attributes, and there is a utility associated with each attribute which depends on whether the product is offered in the offline channel. The intuition behind this approach is that the utilities of the attributes change when customers learn about products by inspecting them in a brick-and-mortar store. Their choice model is the MNL based on the utilities of the attributes. The paper shows that accounting for the impact of the retailer's offline assortment on the online sales can lead to substantial gains in expected revenue.

Lo & Topaloglu (2022) addressed the same problem as Dzyabura & Jagabathula (2018) but in a different setup. The key difference between these is that the work of Dzyabura & Jagabathula (2018) assumes that there exists a product for every potential combination of feature values, while in the model of Lo & Topaloglu (2022) it is assumed that the product portfolio can be characterized by a features tree, where each leaf corresponds to a product, and its ancestors are the features. The authors consider a mixture of customers: offline customers who shop in the physical store, and online customers who first visit the physical store to inspect the products offered there and then choose a product from the full assortment offered online. They showed that the assortment optimization problem in this setup is NP-hard, and leveraged the features tree structure to provide a fully polynomial time approximation scheme (FPTAS) based on dynamic programming that allows to determine approximately optimal assortments.

Finally, Hense & Hübner (2021) studied omnichannel assortment optimization while taking into account both in-channel and cross-channel demand substitution. In contrast to our work, the authors consider the exogenous demand (ED) model instead of leveraging discrete choice modeling techniques. In their approach, the base demands are assumed to be pairwise independent, and if one product is not available, then the proportion of the demand that is substituted by another product is given by a parameter. Such an approach has certain advantages because, in addition to finding optimal assortments, it also allows the authors to determine optimal shelf space and inventory levels across channels. However, the ED model does not allow to capture some of the complexity of customers' product substitution behavior. For instance, according to this model, it is assumed that if a product is not available then its demand would be shifted to another product in the same channel, but the associated demand is lost entirely if this other product is also not available. Moreover, the relatively large number of parameters in their proposed approach makes the problem of estimating the parameters from sales history data particularly challenging.

3. Multichannel attraction model

3.1. Model formulation

In this subsection we present a discrete choice model that captures the complex customer shopping behavior in an omnichannel environment. Our proposed model is a generalization of the GAM to a setup where a retailer can sell products across several channels. Importantly, by generalizing the GAM our model inherits a number of desirable features. First of all, the GAM itself generalizes the basic attraction model (BAM) and the multinomial logit model (MNL), which is arguably the most widely used discrete choice model. At the same time, choice probabilities under the GAM are formulated using simple closed-form expressions. We can also leverage the concept of shadow attractiveness since it has an interpretation that is easily adaptable to the case of an omnichannel retailer. We refer to the proposed model as the multichannel attraction model (MAM).

The MAM is most suitable for the case of a multichannel retailer that offers a range of substitutable products of which customers select at most one product. For instance, a sneakers subdivision of a large apparel retailer is a useful illustration to keep in mind for further reading. For tractability reasons, we consider a retailer with two channels: an online store and a physical store (or a chain of physical stores). However, our model can easily be generalized to the case of a larger number of channels (see Appendix A). The main idea behind the MAM is to develop a framework which allows to manage assortments in both channels jointly, taking into account customers who switch from one channel to another if certain products are unavailable. We therefore separate customers into two groups: the first group comprises customers whose primary choice is to purchase a product in the first channel if all products are available in both channels, whereas customers from the second group shop through the second channel under the same condition. For both groups of customers, we model their choices using our proposed generalization of the GAM, where each shadow attractiveness value is divided into two parts which determine how likely the customers are to switch to another channel to buy the corresponding product.

Let $\mathcal{N} = \{1, 2, \ldots, n\}$ denote a set of products which can be offered in both channels, and let 0 denote the no-purchase alternative. Furthermore, let $c \in \mathcal{C} = \{1, 2\}$ denote a channel index, $\bar{c} = \mathcal{C} \setminus \{c\}$ denote the other channel index, and $S_c \subseteq \mathcal{N}$ denote the set of products offered in channel c. By type-c customers we mean customers whose primary choice would be to shop in channel c if all products were available in both channels. For type-c customers, we use the following notation:

- $v_j^{(c)}$: attractiveness value of purchasing product $j \in S_c$ in channel c;
- $v_0^{(c)}$: attractiveness value of the no-purchase alternative;
- $u_i^{(c)} + w_i^{(c)}$: shadow attractiveness value of purchasing product $i \in \mathcal{N} \setminus S_c$ from another source (that is, either from channel $\bar{c}$ or from another retailer);
- $u_i^{(c)} / (u_i^{(c)} + w_i^{(c)})$: proportion of customers switching to channel $\bar{c}$ out of those willing to purchase product i outside of channel c (if $i \notin S_c$).

Superscript (c) indicates type-c customers' characteristics (who might shop in both channels), whereas subscript c refers to channel-specific features. Note that $u_i^{(c)}$ and $w_i^{(c)}$ are not in themselves attractiveness values. They provide an idea of how likely customers are to switch to another channel or to go to another retailer, but they do not represent utilities of different alternatives as such (see Subsection 3.2 for more details).

Also, similarly to the GAM, we assume that $u_j^{(c)} + w_j^{(c)} \in [0, v_j^{(c)}]$ for all products $j \in \mathcal{N}$, meaning that the shadow attractiveness value of purchasing product j from another source (including channel $\bar{c}$) does not exceed the attractiveness value of purchasing this product in channel c. This also implies that $\mathcal{N}$ is a set of substitutable products, and discarding a product from the assortment in channel c increases the demand for the remaining products generated by type-c customers. It is important to note that the demand substitution assumption is conventional for a setup where each customer purchases at most one product. The opposite effect – when discarding a product leads to a lower demand for some other products – is also possible in such a setup, e.g. if one product highlights the advantages of another product, thus creating a synergy. A recent example of a modeling framework devoted to assortment optimization in the presence of the product synergy effect can be found in Lo & Topaloglu (2019). However, accounting for this effect is outside the scope of our research.

Considering customers of two types associated with two channels is rational for two main reasons. Firstly, these two types of customers are likely to have noticeably distinct shopping preferences, which can be captured by different sets of parameters associated with different customer types. Secondly, the two flows of customers can differ considerably in volume (e.g., the number of customers associated with the online channel can be several times higher or lower than the one associated with the offline channel), and this has to be taken into account when making assortment decisions. Since the MAM is inherently a mixture of models where each model is associated with a customer type, it would be straightforward to extend the model by considering more types of customers. Nevertheless, in this work we focus on two types of customers in order to keep the model tractable.

Let us define the choice probabilities under the MAM given assortments in both channels in the following way. The probability that a type-c customer buys product j in channel c is

$$\pi_{cj}^{(c)}(S_c) = \begin{cases} \dfrac{v_j^{(c)}}{v_0^{(c)} + \sum_{k\in S_c} v_k^{(c)} + \sum_{i\in\mathcal{N}\setminus S_c}(u_i^{(c)} + w_i^{(c)})} & \text{if } j \in S_c, \\ 0 & \text{otherwise;} \end{cases} \tag{5}$$

and the probability that a type-c customer buys product j in channel $\bar{c}$ is

$$\pi_{\bar{c}j}^{(c)}(S_c, S_{\bar{c}}) = \begin{cases} \dfrac{u_j^{(c)}}{v_0^{(c)} + \sum_{k\in S_c} v_k^{(c)} + \sum_{i\in\mathcal{N}\setminus S_c}(u_i^{(c)} + w_i^{(c)})} & \text{if } j \in S_{\bar{c}}\setminus S_c, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

It is straightforward to check that for each customer type, the sum of all choice probabilities (including the no-purchase probability) equals one.

Now, suppose that the total expected number of type-c customers is $\Lambda^{(c)} = \int_0^T \lambda^{(c)}(t)dt$, where $\lambda^{(c)}(t)$ is the arrival rate of type-c customers at time t, and T is the time horizon. Let $x_{cj}^{(c)}(S_c) = \pi_{cj}^{(c)}(S_c)\Lambda^{(c)}$ denote the expected number of type-c customers purchasing product j in channel c, and let $x_{\bar{c}j}^{(c)}(S_c, S_{\bar{c}}) = \pi_{\bar{c}j}^{(c)}(S_c, S_{\bar{c}})\Lambda^{(c)}$ denote the expected number of type-c customers purchasing product j in channel $\bar{c}$. Hereafter, we slightly abuse the notation by writing just $x_{cj}^{(c)}$ and $x_{\bar{c}j}^{(c)}$. Note that the overall probability that a customer (of any type) buys product j in channel c is as follows:

$$\pi_{cj}(S_c, S_{\bar{c}}) = \frac{x_{cj}^{(c)} + x_{cj}^{(\bar{c})}}{\Lambda^{(c)} + \Lambda^{(\bar{c})}}.$$

Importantly, we focus on the modeling setup with unlimited inventories, i.e., we do not consider the possibility of stockouts. As a result, the shopping behavior of customers is fully determined by assortments offered at the beginning of the sales period. Moreover, it also means that terms "demand" and "sales" can be used interchangeably. The assumption of unlimited inventories – which is common in the assortment planning literature – does not prevent our modeling framework from being applicable in a number of relevant and important practical situations. For example, this assumption is valid in the make-to-order setting, in which a company produces a product only after receiving an order and thus avoids carrying a lot of stock. It is also a valid assumption for companies which rarely have stockouts due to a high level of inventory.

3.2. Model discussion

In this subsection, we provide the intuition behind the formulation of the MAM. Generally speaking, when it comes to omnichannel retailing, the product substitution behavior is not trivial, so our MAM requires a detailed description. At a high level, we assume that under the MAM, customers are subject to the following product substitution behavior. Suppose that if all products were available in all channels, a certain type-c customer would purchase product j in channel c. If this product is not offered in channel c, then the customer may either be determined to purchase product j anyway (potentially in channel $\bar{c}$), or decide to purchase another product $k \neq j$ instead, or leave without purchasing anything from this retailer. The outcomes of these three alternatives are summarized in Fig. 1. In essence, if product j is not offered in channel c (i.e. $j \notin S_c$) and it is the first choice of a type-c customer, then this customer will either purchase another product $k \in S_c$, or a product $l \in S_{\bar{c}}\setminus S_c$, where $l \neq k$ but possibly $l = j$, or nothing at all.

Let us consider the first two cases in more detail. First, suppose that the customer decides to stick with product j. In this case, the customer may be willing to search for product j in channel $\bar{c}$, or to go looking for this product somewhere else. To keep the model formulation tractable, we assume that if a customer decides to stick with a certain product and thus switches to another channel searching for it, then the product is always purchased if it is available in that channel, or else the customer leaves without purchasing anything from this retailer (e.g., goes to a competitor). Second, if the customer decides to purchase another product k, then there are two possibilities: If product k is offered in channel c, then the customer purchases it there; otherwise, the customer either continues the search in channel c, or looks for product k in channel $\bar{c}$, or looks for this product somewhere else. Similar to the previous case, we assume that if the customer is willing to switch to channel $\bar{c}$ in order to look for product k, then the product is always purchased if it is available there, otherwise the customer just leaves without purchasing anything.

Let us also provide additional insights into the structure of this product substitution behavior by showing the link between the MAM and Markov chain choice model (MCCM).

Proposition 1. *The MAM can be represented as a mixture of MCCMs where choices of type-c customers are characterized by the MCCM with the following parameters:*

$$\begin{aligned}
&\lambda_{j_c}^{(c)} = v_j^{(c)}, \quad \lambda_{j_{\bar{c}}}^{(c)} = 0,\\
&\rho_{j_c i_c}^{(c)} = \frac{v_i^{(c)}(v_j^{(c)} - u_j^{(c)} - w_j^{(c)})}{v_j^{(c)} - v_j^{(c)}(v_j^{(c)} - u_j^{(c)} - w_j^{(c)})},\\
&\rho_{j_c j_{\bar{c}}}^{(c)} = \frac{u_j^{(c)}}{v_j^{(c)}}, \quad \rho_{j_c i_{\bar{c}}}^{(c)} = 0,\\
&\rho_{j_c 0}^{(c)} = \frac{v_0^{(c)}(v_j^{(c)} - u_j^{(c)} - w_j^{(c)}) + u_j^{(c)} + w_j^{(c)}}{v_j^{(c)} - v_j^{(c)}(v_j^{(c)} + u_j^{(c)} + w_j^{(c)})} - \frac{u_j^{(c)}}{v_j^{(c)}},\\
&\rho_{j_{\bar{c}} i_{\bar{c}}}^{(c)} = 0, \quad \rho_{j_{\bar{c}} j_c}^{(c)} = 0, \quad \rho_{j_{\bar{c}} i_c}^{(c)} = 0, \quad \rho_{j_{\bar{c}} 0}^{(c)} = 1,
\end{aligned} \tag{7}$$

where j_c denotes product j in channel c, $\lambda^{(c)}$ is the vector of arrival probabilities, and $\rho^{(c)}$ is the matrix of transition probabilities.

The formal proof can be found in Appendix B. The intuition behind these expressions is as follows. We leverage the fact that the MAM restricted to type-c customers and products in channel c is equivalent to the GAM. Therefore, for such customers, the arrival probabilities of products in channel c as well as the transition probabilities between such products are defined by analogy to the probabilities that result in the GAM (see Blanchet et al., 2016).

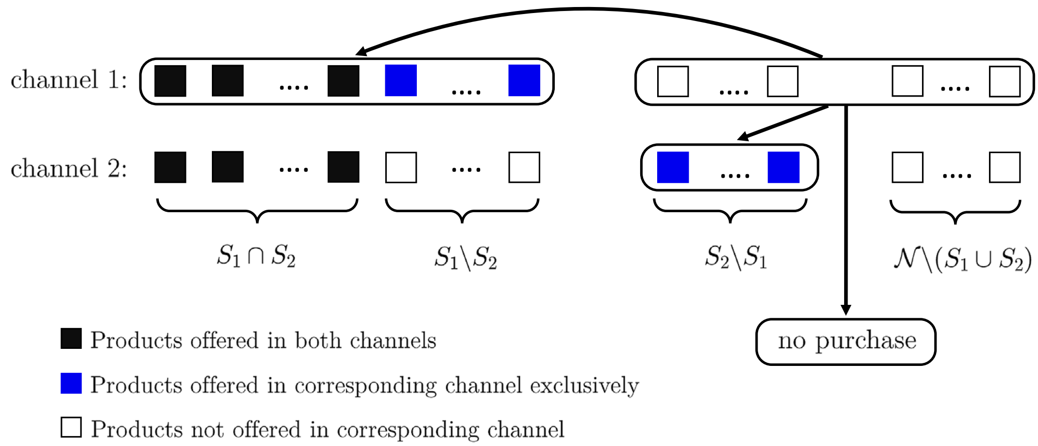


Fig. 1. Product substitution pattern.

However, in contrast to the GAM, we split the transition probability from product j_c to the no-purchase alternative into two parts: one corresponds to choosing the no-purchase alternative directly, and the other one corresponds to first choosing product $j_{\bar{c}}$, and then – in case product j is not available in channel $\bar{c}$ – choosing the no-purchase alternative with probability 1. Note that we impose the constraint that type-c customers cannot purchase product $j_{\bar{c}}$ directly as they attempt to purchase product j_c first. This means, in turn, that type-c customers will never buy product j in channel $\bar{c}$ if this product is available in channel c. We believe that this assumption is not only reasonable, but also essential as it enables us to obtain simple analytical formulas (and thus achieve tractability and ensure interpretability of the model) for the choice probabilities, which would not be the case if we used a mixture of MCCMs in a general setting.

Let us also highlight the connection between the MAM and the random utility theory. First, note that the GAM is a random utility model (RUM), i.e., there exists a joint distribution of random utilities over a certain set of alternatives such that if each customer chooses an alternative with a maximum realization of utility, then the product choice probabilities are consistent with the GAM choice probabilities. Indeed, the GAM can be viewed as the nested logit model in the limit (see Gallego et al., 2014), and the nested logit model is a special case of the generalized extreme value (GEV) model, which is a RUM. Now since the GAM is a RUM, it is straightforward to show that for each customer type, there exists a joint distribution of random utilities over a certain set of alternatives that is consistent with the MAM choice probabilities. This can be done by using the fact that the condition of existence of a joint probability distribution of random utilities is equivalent to the condition of existence of a probability distribution over rankings of alternatives consistent with a given set of choice probabilities (see Block & Marschak, 1959). The MAM thus belongs to the class of RUMs and we can formulate the following proposition:

Proposition 2. *The MAM is a mixture of RUMs (one model per customer type), and as such it is also a RUM.*

The formal proof of this proposition can be found in Appendix C. However, note that our modelling approach is motivated by incorporating probabilistic cross-channel transitions into the GAM to account for the complex shopping behavior of customers in an omnichannel environment, rather than by proposing a new discrete choice model through specifying random utility of alternatives within a random utility maximization framework.

Finally, let us demonstrate the benefit of our formulation approach as opposed to a more traditional, utility-based way of generalizing the GAM to the omnichannel setting. As mentioned above, Gallego et al. (2014) showed that the GAM can be represented as a nested logit model in a limit. Following the same logic, one could have formulated the MAM so that it would emerge as a mixture of nested logit models (one model per customer type) in a limit, where each nest corresponds to a product and the dissimilarity parameter of each nest tends to zero. In this case, each nest would comprise three alternatives: purchasing the product in channel c, in channel $\bar{c}$, and from another source. Then, the choice probabilities would take the following form:

$$\pi_{cj}^{(c)}(S_c, S_{\bar{c}}) = \begin{cases} \dfrac{v_j^{(c)}}{v_0^{(c)} + \sum_{k\in S_c} v_k^{(c)} + \sum_{i\in(\mathcal{N}\backslash S_c)\cap S_{\bar{c}}} \max\{u_i^{(c)}, w_i^{(c)}\} + \sum_{l\in\mathcal{N}\backslash(S_c\cup S_{\bar{c}})} w_l^{(c)}} & \text{if } j \in S_c, \\ 0 & \text{otherwise;} \end{cases}$$

and the probability that a type-c customer buys product j in channel $\bar{c}$ is

$$\pi_{\bar{c}j}^{(c)}(S_c, S_{\bar{c}}) = \begin{cases} \dfrac{u_j^{(c)}}{v_0^{(c)} + \sum_{k\in S_c} v_k^{(c)} + \sum_{i\in(\mathcal{N}\backslash S_c)\cap S_{\bar{c}}} \max\{u_i^{(c)}, w_i^{(c)}\} + \sum_{l\in\mathcal{N}\backslash(S_c\cup S_{\bar{c}})} w_l^{(c)}} & \text{if } j \in S_{\bar{c}}\backslash S_c \text{ and } u_j^{(c)} \geq w_j^{(c)}, \\ 0 & \text{otherwise.} \end{cases}$$

Importantly, in the above formulas the parameters $u_j^{(c)}$ and $w_j^{(c)}$ have a different interpretation compared to our model: for type-c customers, $u_j^{(c)}$ and $w_j^{(c)}$ here represent the shadow attractiveness of purchasing product j in channel $\bar{c}$ and elsewhere, respectively, whereas in our case $u_j^{(c)}$ is defined through the proportion of customers willing to purchase product j in channel $\bar{c}$ if it is not available in channel c (see Subsection 3.1).

However, this alternative formulation based on a mixture of nested logit models has important limitations. Indeed, it means that if $j \in S_{\bar{c}}\backslash S_c$, then either all type-$c$ customers that choose nest j purchase product $j_{\bar{c}}$ (if $u_j^{(c)} > w_j^{(c)}$), or all of them leave the retailer (if $u_j^{(c)} < w_j^{(c)}$). Both these cases are rather extreme and hence not sufficiently realistic. There is also the special case when $u_j^{(c)} = w_j^{(c)}$, in which exactly half of the considered customers purchase product $j_{\bar{c}}$ and the other half leave the retailer, but this is also too restrictive. In contrast, our approach to formulating the MAM probabilities does not suffer from these limitations as it allows any partitioning of customers' choices between purchasing product $j_{\bar{c}}$ and leaving the retailer. We therefore believe that our approach is appealing because it is not only capable of more realistically representing omnichannel customer behaviour but it also provides for more flexibility without overcomplicating the choice model.

4. Assortment optimization problem

The assortment optimization problem under the MAM requires finding offer sets such that the total expected profit is maximized, that is determining the optimizers of the following problem:

$$\max_{S_c, S_{\bar{c}}} \sum_{c\in\mathcal{C}} \sum_{j\in\mathcal{N}} r_{cj}\pi_{cj}(S_c, S_{\bar{c}}), \tag{8}$$

where r_{cj} denotes the gross profit per unit of product j sold through channel c. This problem can be reformulated in terms of binary variables. Let z_{cj}, $c \in \mathcal{C}$, $j \in \mathcal{N}$ be a binary variable such that $z_{cj} = 1$ if $j \in S_c$, and 0 otherwise. Then, problem (8) can be written as follows:

$$\max_{S_c, S_{\bar{c}}} \sum_{c\in\mathcal{C}} \sum_{j\in\mathcal{N}} r_{cj}\pi_{cj}(S_c, S_{\bar{c}}) =$$

$$\max_{z_{cj}, z_{\bar{c}j}} \sum_{c\in\mathcal{C}} \sum_{j\in\mathcal{N}} \frac{r_{cj}}{\Lambda^{(c)} + \Lambda^{(\bar{c})}} \left(\frac{v_j^{(c)}\Lambda^{(c)} z_{cj}}{v_0^{(c)} + \sum_{k\in\mathcal{N}} v_k^{(c)} z_{ck} + \sum_{i\in\mathcal{N}} (u_i^{(c)} + w_i^{(c)})(1 - z_{ci})} + \frac{u_j^{(\bar{c})}\Lambda^{(\bar{c})} z_{cj}(1 - z_{\bar{c}j})}{v_0^{(\bar{c})} + \sum_{k\in\mathcal{N}} v_k^{(\bar{c})} z_{\bar{c}k} + \sum_{i\in\mathcal{N}} (u_i^{(\bar{c})} + w_i^{(\bar{c})})(1 - z_{\bar{c}i})} \right).$$

This optimization problem is extremely difficult to solve as it includes binary decision variables and a nonlinear objective function. However, we can formulate an equivalent mixed integer linear program (MILP) that can be solved using standard, off-the-shelf optimization software. Consider the following problem:

$$\max_{x} \sum_{c\in\mathcal{C}} \sum_{j\in\mathcal{N}} r_{cj}\left(x_{cj}^{(c)} + x_{cj}^{(\bar{c})}\right) \tag{9a}$$

$$\text{s.t.} \quad \frac{\tilde{v}_0^{(c)}}{v_0^{(c)}} x_{c0}^{(c)} + \sum_{j\in\mathcal{N}} \frac{\tilde{v}_j^{(c)}}{v_j^{(c)}} x_{cj}^{(c)} = \Lambda^{(c)} \qquad \forall c \in \mathcal{C}, \tag{9b}$$

$$\frac{x_{cj}^{(c)}}{v_j^{(c)}} + \frac{x_{\bar{c}j}^{(c)}}{u_j^{(c)}} \le \frac{x_{c0}^{(c)}}{v_0^{(c)}} \qquad \forall c \in \mathcal{C},\ j \in \mathcal{N}, \tag{9c}$$

$$\frac{x_{c0}^{(c)}}{v_0^{(c)}} - \frac{x_{cj}^{(c)}}{v_j^{(c)}} \le \frac{\Lambda^{(c)}}{\tilde{v}_0^{(c)}}(1 - z_{cj}) \qquad \forall c \in \mathcal{C},\ j \in \mathcal{N}, \tag{9d}$$

$$x_{cj}^{(c)} \le H_j^{(c)} z_{cj} \qquad \forall c \in \mathcal{C},\ j \in \mathcal{N}, \tag{9e}$$

$$x_{\bar{c}j}^{(c)} \le K_j^{(c)} x_{\bar{c}j}^{(\bar{c})} \qquad \forall c \in \mathcal{C},\ j \in \mathcal{N}, \tag{9f}$$

$$x_{c0}^{(c)}, x_{cj}^{(c)}, x_{\bar{c}j}^{(c)} \in \mathbb{R}_{\ge 0} \qquad \forall c \in \mathcal{C},\ j \in \mathcal{N}, \tag{9g}$$

$$z_{cj} \in \{0, 1\} \qquad \forall c \in \mathcal{C},\ j \in \mathcal{N}, \tag{9h}$$

where $x = \left\{x_{c0}^{(c)},\ x_{cj}^{(c)},\ x_{\bar{c}j}^{(c)}, z_{cj}\right\}_{c\in\mathcal{C}, j\in\mathcal{N}}$, $\tilde{v}_0^{(c)} = v_0^{(c)} + \sum_{i\in\mathcal{N}}\left(u_i^{(c)} + w_i^{(c)}\right)$, $\tilde{v}_j^{(c)} = v_j^{(c)} - \left(u_j^{(c)} + w_j^{(c)}\right)$, and constants $H_j^{(c)}$ and $K_j^{(c)}$ are given by

$$H_j^{(c)} = \frac{v_j^{(c)}\Lambda^{(c)}}{v_0^{(c)} + v_j^{(c)} + \sum_{i\in\mathcal{N}\setminus\{j\}}\left(u_i^{(c)} + w_i^{(c)}\right)},$$

$$K_j^{(c)} = \frac{u_j^{(c)}\Lambda^{(c)}}{\tilde{v}_0^{(c)}} \Bigg/ \frac{v_j^{(\bar{c})}\Lambda^{(\bar{c})}}{v_0^{(\bar{c})} + \sum_{k\in\mathcal{N}} v_k^{(\bar{c})}}.$$

We call this problem formulation the sales-based mixed integer linear program (SBMILP) in analogy to the sales-based linear program presented by Gallego et al. (2014) for the GAM. Constraints (9b) are similar to the balance constraint in the SBLP, and constraints (9c) are modified scale constraints. However, due to the multichannel structure of the MAM and hence a more complex product substitution behavior of customers, we need additional constraints with binary variables. The meaning of each constraint as well as the equivalence of problems (8) and (9) becomes evident from the proof of the following Theorem (see Appendix D):

Theorem 1. *The SBMILP is a valid formulation of the assortment optimization problem under the MAM, that is, the optimal value of problem (9) is equal to the optimal value of problem (8) multiplied by the constant* $(\Lambda^{(c)} + \Lambda^{(\bar{c})})$.

Remark 1. It is straightforward to verify that constraints (9d), (9e) and (9f) cannot be tightened, that is, the constant coefficients on the right-hand side of these constraints cannot be reduced.

Remark 2. The SBMILP can easily be modified to incorporate additional constraints. For example, if there is a cost a_{cj} associated with product j offered in channel c, and the total cost induced by products offered in this channel is limited by the upper bound L_c, then the following constraint has to be added to the SBMILP:

$$\sum_{j\in\mathcal{N}} a_{cj} z_{cj} \le L_c. \tag{10}$$

If channel c is the physical channel, then constraint (10) can be viewed as a shelf-space constraint where a_{cj} represents the shelf space required for product j to be offered in channel c, and L_c is the total shelf space in that channel.

Adding the shelf space constraint described in Remark 2 to the SBMILP is of particular practical importance (typically omnichannel retailers cannot offer all products in the physical channel due to limited shelf space), but has implications in terms of computational complexity:

Proposition 3. *The assortment optimization problem represented by the shelf-space-constrained SBMILP, which is given by adding constraint (10) to the SBMILP formulation (9), is NP-hard.*

The proof can be found in Appendix E. It is important to note that there is another modification of the assortment optimization problem under the MAM that not only makes the modified problem NP-hard, but may also suggest that the original SBMILP is NP-hard too. In particular, if assortments in both channels have to be the same (i.e. $S_1 = S_2$), then this problem is equivalent to the assortment optimization problem under a mixture of two GAMs in a single channel. Such a problem is NP-hard as a generalization of the assortment optimization problem under a mixture of two MNL models, which has been shown by Rusmevichientong, Shmoys, Tong, & Topaloglu (2014) to be NP-hard. This well-known result, formulated for the simplest illustrative case of the assortment optimization problem under a mixture of discrete choice models, strongly indicates that the original SBMILP formulation (9) is also NP-hard since the MAM itself is essentially a mixture of discrete choice models (one model per customer type). While an interesting problem, formally establishing NP-hardness of the original SBMILP is left for future research.

If the assortment in one of the channels is fixed and equals $\mathcal{N}$ (i.e., all products are offered), then we can build upon some of the results obtained for the SBLP in Gallego et al. (2014) and establish certain analytical properties of the optimal assortment in the other channel. Let $R^{(c)}(S_c, S_{\bar{c}})$ be the total profit generated by type-c customers given assortments S_c and $S_{\bar{c}}$. Suppose that the assortment in channel $\bar{c}$ is fixed so that $S_{\bar{c}} = \mathcal{N}$, thus no customers of type

$\bar{c}$ will switch to channel c. We can then formulate the following proposition:

Proposition 4. *Let $S_{\bar{c}} = \mathcal{N}$ and, without loss of generality, suppose that all products are sorted in descending order of the ratio $(r_{cj}v_j^{(c)} - r_{\bar{c}j}u_j^{(c)})\Lambda^{(c)}/\tilde{v}_j^{(c)}$. Then, the optimal assortment in channel c is given by*

$$z_{cj} = \begin{cases} 1 & \text{if } j \le m, \\ 0 & \text{otherwise,} \end{cases} \tag{11}$$

where $m = \max\left\{ j \in \mathcal{N} : R^{(c)}\left(\{1, \ldots, j\}, \mathcal{N}\right) < (r_{c,j+1}v_{j+1}^{(c)} - r_{\bar{c},j+1}u_{j+1}^{(c)})\Lambda^{(c)}/\tilde{v}_{j+1}^{(c)} \right\}$.

The proof can be found in Appendix F. The intuition behind this finding is clear. Even if the gross profit per unit of a certain product j in channel c is high and the product has a high attractiveness value (i.e., the demand for product j in channel c is high compared to other products), it may not be profitable to include this product in the channel c assortment.

Further developing the idea behind the proof of Proposition 4, we can formulate the following property of the optimal assortment in both channels:

Proposition 5. *Let $(S_c, S_{\bar{c}})$ be the optimal combination of assortments, and suppose that $k \in S_c$. Then $j \in S_c$ as well if $F_c(j, S_c, S_{\bar{c}}) \ge F_c(k, S_c, S_{\bar{c}})$, where*

$$F_c(j, S_c, S_{\bar{c}}) = \frac{r_{cj}v_j^{(c)} - r_{\bar{c}j}u_j^{(c)}\mathbb{1}_{j\in S_{\bar{c}}}}{\tilde{v}_j^{(c)}} + \frac{\left(v_0^{(c)} + \sum_{k\in S_c\cup\{j\}} v_k^{(c)} + \sum_{i\in\mathcal{N}\setminus(S_c\cup\{j\})} (u_i^{(c)} + w_i^{(c)})\right) r_{cj}u_j^{(\bar{c})}\Lambda^{(\bar{c})}/\Lambda^{(c)}\mathbb{1}_{j\in\mathcal{N}\setminus S_{\bar{c}}}}{\tilde{v}_j^{(c)}\left(v_0^{(\bar{c})} + \sum_{k\in S_{\bar{c}}} v_k^{(\bar{c})} + \sum_{i\in\mathcal{N}\setminus S_{\bar{c}}} (u_i^{(\bar{c})} + w_i^{(\bar{c})})\right)}. \tag{12}$$

This property, whose proof is provided in Appendix G, could serve as a base for a heuristic algorithm that allows to determine a high-quality solution if solving the SBMILP is computationally infeasible. In particular, $F_c(j, S_c, S_{\bar{c}})$ could be approximated with an expression that depends only on product j and channel-$\bar{c}$ assortment, which would allow for the approximate characterization of the optimal assortment in channel c given the assortment in channel $\bar{c}$. Then, the heuristic algorithm could take form of an iterative procedure, where in each iteration the assortment in one of the channels is determined given the assortment in the other channel. In this work, however, we focus on the development of a heuristic algorithm based on a relaxation approach for situations in which directly solving the SBMILP is computationally challenging.

5. Heuristic method

The SBMILP formulation (9) has $2n$ binary variables and $4n$ constraints containing binary variables. For large values of n, this problem may become more challenging to solve. Let us therefore consider the LP relaxation of the SBMILP derived from formulation (9) by removing binary variables z_{cj} together with the corresponding constraints (9d) and (9e). Importantly, preliminary numerical experiments showed that a solution to the relaxed problem satisfies the removed constraints for almost all $c \in \mathcal{C}$, $j \in \mathcal{N}$, and this is a key observation underlying our heuristic. We propose the following two-step algorithm:

1. Solve the relaxed problem. Let $\left\{\hat{x}_{c0}^{(c)},\ \hat{x}_{cj}^{(c)},\ \hat{x}_{\bar{c}j}^{(c)}\right\}_{c\in\mathcal{C}, j\in\mathcal{N}}$ be its optimal solution, and $\mathcal{J}^{(c)}$ be the set of indexes $j \in \mathcal{N}$ such that either $\frac{\hat{x}_{cj}^{(c)}}{v_j^{(c)}} = \frac{\hat{x}_{c0}^{(c)}}{v_0^{(c)}}$ or $\hat{x}_{cj}^{(c)} = 0$.
2. Solve problem (9) with the following additional constraints:
$$z_{cj} = \mathbb{1}_{\hat{x}_{cj}^{(c)} > 0} \quad \forall c \in \mathcal{C},\ j \in \mathcal{J}^{(c)}, \tag{13}$$
where $\mathbb{1}$ represents the indicator function. The obtained solution is the heuristic output.

To gain insights into the computational performance of the proposed heuristic method, we generate the parameters of the omnichannel assortment optimization problem in the following way:

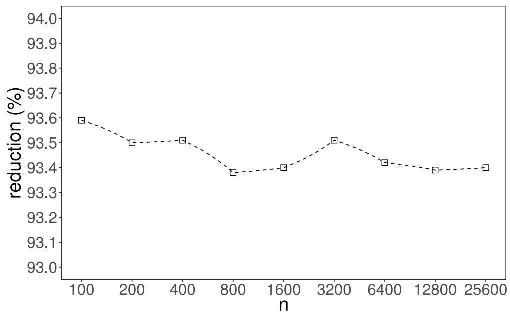
(i) $r_{1j} = u(0,1) + \varepsilon$, $r_{2j} = r_{1j}(1 + u(0, 0.5))\ \forall j \in \mathcal{N}$;
(ii) $v_j^{(1)} = u(0,1) + \varepsilon$, $v_j^{(2)} = u(0,1) + \varepsilon\ \forall j \in \mathcal{N} \cup \{0\}$;
(iii) $u_j^{(1)} = u(0, 0.5)v_j^{(1)}$, $u_j^{(2)} = u(0, 0.5)v_j^{(2)}\ \forall j \in \mathcal{N}$;
(iv) $w_j^{(1)} = u(0, 0.5)v_j^{(1)}$, $w_j^{(2)} = u(0, 0.5)v_j^{(2)}\ \forall j \in \mathcal{N}$,

where $u(a, b)$ denotes a value sampled from the uniform distribution $\mathcal{U}(a, b)$, and $\varepsilon = 0.01$. Also, we normalize the attractiveness values so that $v_0^{(c)} + \sum_{j\in\mathcal{N}} v_j^{(c)} = 1\ \forall c \in \mathcal{C}$. While the generated values of the parameters may not necessarily be representative of a real-world example, they are well-suited for the purpose of evaluating the computational effort required to solve the problem instance. The only meaningful restriction we impose is that the gross profit per unit of a product is higher for the online channel than for the offline channel, which can be justified by the difference in holding costs. Lastly, we fix the values of parameters $\Lambda^{(1)}$ and $\Lambda^{(2)}$ at 10^4 and $3 \cdot 10^4$, respectively.

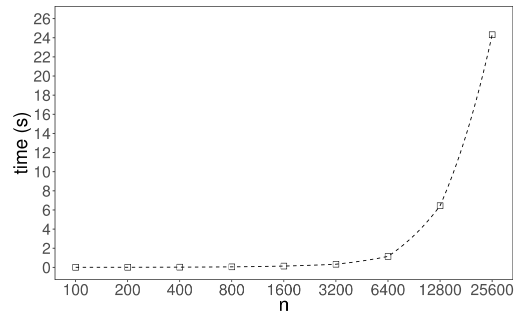
We use the SBMILP as a benchmark to evaluate the comparative performance of our heuristic method. The computational study was carried out on a laptop with Intel Core i7-8650U CPU (1.90 GHz), 8 GB RAM and 64-bit Windows 10 OS. To solve our MILPs we used Gurobi (version 8.0.1). We ran 100 experiments for each value of $n \in \{100, 200, 300\}$. The results showed that, on average, the SBMILP can be solved to optimality in 0.29 seconds if $n = 100$, in 1.48 seconds if $n = 200$, and in 10.72 seconds if $n = 300$. However, running 100 experiments for $n = 400$ turned out to be not particularly feasible, as some instances took up to 500 seconds to solve (even though some other instances were solved in less than 7 seconds). This highlights the strong need for a computationally efficient heuristic method to solve the assortment optimization problem.

Our proposed heuristic method allows to drastically reduce both the number of binary variables and the number of constraints containing binary variables in the SBMILP formulation. This is because each binary variable z_{cj} such that $j \in \mathcal{J}^{(c)}$ turns into a parameter. The numerical results in terms of reduction in the number of binary variables and solving time for different values of n are shown in Fig. 2(a) and (b), respectively (note that the x-axis is log-scaled with base 2). The results are averaged over 100 generated instances. It can be seen that the reduction in the number of binary variables is more than 93% and this is independent of the actual value of n. As a consequence, the optimization problem (9) with additional constraints (13) is computationally much easier to solve than the original one. The heuristic still requires solving a MILP whose size grows linearly in n, so the solving time grows exponentially in n. However, even if n is of order 10^4, the heuristic solution can be found in a matter of seconds, making this method attractive for most practical applications.

For moderate values of n – in the range from 50 to 300 – we also compared the output and the solving time of the heuristic to those of the SBMILP. The results, which again represent averages over 100 generated instances, are given in Table 1. As can be seen, on average, the profit yielded by the heuristic is around 0.99997 of the optimal profit, and the proportion of mismatched assortment decisions (i.e., the proportion of z_{cj} values which are different for the heuristic output and for the optimal output) is almost



(a) Reduction in the number of binary variables.



(b) Solving time of the heuristic method.

Fig. 2. Computational performance of the heuristic method as a function of the product set size (n).

Table 1
Comparative performance of the heuristic method.

n	ratio of solving times	ratio of optimal values	proportion of mismatched decisions
50	8.95	0.999984	0.003400
100	39.76	0.999975	0.003950
150	86.17	0.999979	0.003400
200	138.61	0.999960	0.003925
250	218.32	0.999974	0.003540
300	595.31	0.999966	0.004050

always less than 0.004. At the same time, the ratio of the solving time of the original formulation to the solving time of the heuristic is around 9 if $n = 50$, and around 595 if $n = 300$. These results demonstrate that the developed heuristic method yields close-to-optimal assortments whilst generally being substantially superior in terms of solving time, especially as n grows.

Finally, for large values of n, we verified numerically that our heuristic method yields close-to-optimal solutions. For each $n \in \{5000, 10000, 15000, 20000, 25000\}$, we generated 100 random MAM instances. Then, we compared the heuristic outputs with the solutions to the linear relaxations of the corresponding SBMILPs, which provide upper bounds on the true optimal values. For each problem instance, we computed the value of $(obj_{rel} - obj_{heur})/obj_{heur}$, where obj_{rel} is the value of the linear relaxation of the SBMILP and obj_{heur} is the value of the heuristic. As can be seen in Fig. 3, the value of these gaps never exceeded 0.2%, with the average gap between 0.15% and 0.16%, confirming the high solution accuracy of the proposed heuristic.

Note that since the MAM is a regular choice model – i.e., adding a product to an assortment in one of the channels cannot lead to an increase in the probability of customers choosing any other product – the revenue-ordered heuristic algorithm (see, e.g., Berbeglia & Joret, 2020) can be used to solve the assortment optimization problem under the MAM. In particular, if all products (in both channels simultaneously) are ranked in ascending order of the price, and the heuristic solution is determined by finding the best cutoff in this ranking, then this solution approximates the optimum revenue within a factor of $\frac{1}{1+\ln(r_{\max}/r_{\min})}$, where $r_{\max}$ and $r_{\min}$ are the maximum and the minimum price, respectively. This method can prove to be very useful if prices are sufficiently close to each other. By contrast, the quality of our proposed heuristic method, which is very promising in terms of computational time and solution accuracy, does not rely on the closeness of prices of products, meaning it is readily applicable in a general setting.

6. Parameter estimation

We start with describing a basic method to estimate the parameters of the MAM. Recall that $\pi_{cj}(S_c, S_{\bar{c}})$ is the probability that a customer buys product j in channel c given assortments S_c and $S_{\bar{c}}$ in channels c and $\bar{c}$, respectively. Suppose that for all $i, j \in \mathcal{N}$, $c \in \mathcal{C}$ the probabilities $\pi_{cj}(\mathcal{N}, \mathcal{N})$, $\pi_{cj}(\mathcal{N}\backslash\{i\}, \mathcal{N})$ and $\pi_{cj}(\mathcal{N}, \mathcal{N}\backslash\{i\})$ are known along with the respective values of $\Lambda^{(c)}$ and $\Lambda^{(\bar{c})}$. Such information can be obtained from data comprising the aggregate demand values (or, equivalently, the aggregate sales, assuming that there are no shortages) for the corresponding assortments together with the shares of customers of each type who chose the no-purchase alternative if all the products from $\mathcal{N}$ had been offered in both channels. The latter means that the ratios $v_0^{(c)}/\sum_{j\in\mathcal{N}} v_j^{(c)}$ $\forall c \in \mathcal{C}$ are estimated exogenously. While the assumption of having these exogenous estimates is fairly restrictive, it is in line with the existing literature (see Vulcano, van Ryzin, & Ratliff, 2012).

Without loss of generality, we can assume $v_0^{(c)} + \sum_{j\in\mathcal{N}} v_j^{(c)} = 1$ $\forall c \in \mathcal{C}$. For all $c \in \mathcal{C}$, $j \in \mathcal{N}$, the MAM parameters can then be determined using the following expressions:

$$v_j^{(c)} = \pi_{cj}(\mathcal{N}, \mathcal{N})\frac{\Lambda^{(c)} + \Lambda^{(\bar{c})}}{\Lambda^{(c)}}, \quad v_0^{(c)} = 1 - \sum_{j\in\mathcal{N}} v_j^{(c)},$$

$$u_j^{(c)} = \frac{\pi_{\bar{c}j}(\mathcal{N}\backslash\{j\}, \mathcal{N}) - \pi_{\bar{c}j}(\mathcal{N}, \mathcal{N})}{\pi_{ck}(\mathcal{N}\backslash\{j\}, \mathcal{N})} v_k^{(c)},$$

$$w_j^{(c)} = \frac{\pi_{ck}(\mathcal{N}, \mathcal{N})}{\pi_{ck}(\mathcal{N}\backslash\{j\}, \mathcal{N})} + v_j^{(c)} - u_j^{(c)} - 1, \quad (14)$$

where $k \in \mathcal{N}$ is any product different from j. Expressions (14) can be verified by straightforward calculations. This method, however, requires specific information on the product demands for a certain set of assortments, which may be difficult to obtain in practice. Therefore, there is a need for more general parameter estimation methods for situations in which only limited data about product demands is available. The two most common parameter estimation techniques for discrete choice models are maximum likelihood estimation (MLE) and least squares estimation. In their recent empirical study, Berbeglia, Garassino, & Vulcano (2021) compared the estimation results produced by these two standard estimation techniques and, considering a range of prominent discrete choice models, found that the quality of estimates of these two methods is very similar. Importantly, least squares estimation requires only aggregate sales data (i.e., how many units of each product were sold during each period with a fixed assortment), whereas MLE is typically used when the information about all individual sales

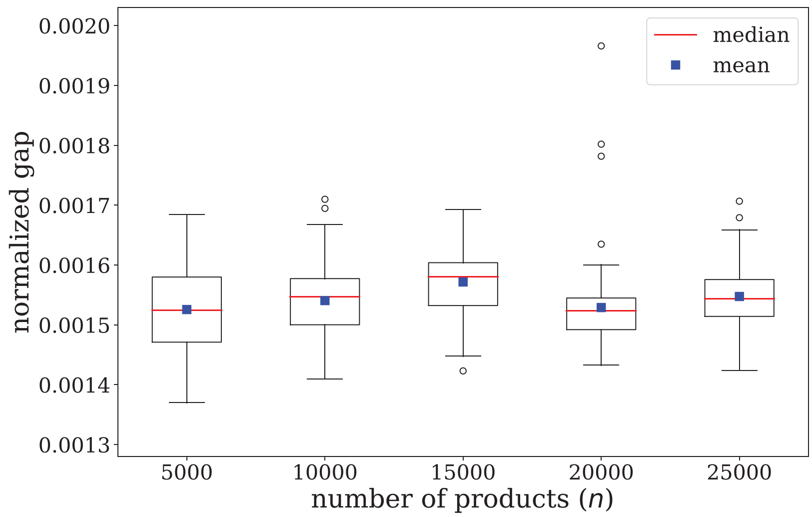


Fig. 3. Heuristic performance bounds for large numbers of products.

transactions is available. In this work, we focus on the limited setting of aggregate sales transaction data (i.e. only sales quantities and product availability in each period are observed) in which direct maximization of the (log-)likelihood function is computationally unappealing (see Vulcano et al. 2012), so we resort to an estimation based on least squares which is readily applicable in incomplete data situations.

Suppose that product demands which arise from the MAM are observed in both channels for T periods. In other words, we implicitly assume customers make choices according to an MAM in a homogeneous market (i.e. preferences of customer are homogeneous across the selling horizon, meaning their choice behaviour can be modelled by a single MAM). Note that although this assumption is standard in the estimation of discrete choice models, it can be straightforwardly relaxed (see, for example, the discussion in Vulcano et al. 2012). For each period t, we denote the assortment in channel c by S_{ct} and the observed demand for product j in channel c by d_{cjt}. We also assume that the demand rate is constant for both channels, i.e. $\Lambda_t^{(c)} = \Lambda^{(c)}\ \forall t \in \{1,\dots,T\}$ and that the market size is sufficiently large (i.e. observed historical sales are representative of expected sales). If the available data is composed solely of the demand values and the corresponding assortments, we can obtain estimates of the MAM parameters by minimizing the sum of squared residuals, i.e., by solving the following optimization problem:

$$\begin{aligned}
\min_{\substack{v^{(c)},\Lambda^{(c)}\\ u^{(c)},w^{(c)}}} \quad & \sum_{t=1}^{T}\sum_{c\in\mathcal{C}}\sum_{j\in\mathcal{N}}\Bigg(\frac{v_j^{(c)}\Lambda^{(c)}\mathbb{1}_{j\in S_{ct}}}{v_0^{(c)}+\sum_{k\in S_{ct}} v_k^{(c)}+\sum_{i\in\mathcal{N}\setminus S_{ct}}(u_i^{(c)}+w_i^{(c)})}+\\
& \frac{u_j^{(\bar c)}\Lambda^{(\bar c)}\mathbb{1}_{j\in S_{ct}\setminus S_{\bar c t}}}{v_0^{(\bar c)}+\sum_{k\in S_{\bar c t}} v_k^{(\bar c)}+\sum_{i\in\mathcal{N}\setminus S_{\bar c t}}(u_i^{(\bar c)}+w_i^{(\bar c)})}-d_{cjt}\Bigg)^2\\
\text{s.t.} \quad & u_j^{(c)}+w_j^{(c)} \le v_j^{(c)} && \forall c\in\mathcal{C},\ j\in\mathcal{N},\\
& v_0^{(c)}+\sum_{j\in\mathcal{N}} v_j^{(c)} = 1 && \forall c\in\mathcal{C},\\
& \Lambda^{(c)}, v_0^{(c)}, v_j^{(c)}, u_j^{(c)}, w_j^{(c)} \in \mathbb{R}_{\ge 0} && \forall c\in\mathcal{C},\ j\in\mathcal{N}.
\end{aligned} \tag{15}$$

Importantly, despite the fact that problem (15) is nonconvex, some off-the-shelf solvers are able to cope with this problem quite well. In particular, the IPOPT package developed by Wächter & Biegler (2006) for nonlinear optimization shows a surprisingly good performance. Following Vulcano et al. (2012) and Gallego et al. (2014), we illustrate the performance of the above parameter estimation method by considering an exemplary setup with $n = 5$ and $T = 15$. We simulate 100 instances of demand arising from the MAM with fixed parameters (which are randomly generated in the way described in Section 5), with each instance corresponding to a set of randomly simulated assortments S_{ct}, $c \in \mathcal{C}$, $t \in \{1,\dots,T\}$. We consider two cases: in the first case, we assume that the values of the ratios $v_0^{(c)} / \sum_{j\in N} v_j^{(c)}\ \forall c \in \mathcal{C}$ (or, equivalently, the values of $v_0^{(c)}$ if $\sum_{j\in\mathcal{N}} v_j^{(c)} = 1$) are given exogenously, whereas no such information is available in the second case. The least squares estimates obtained by solving problem (15) and averaged over 100 instances are presented in Table 2. It can be seen that these estimates are particularly close to the true values of the parameters if the values of $v_0^{(c)}$ are known, which highlights the importance of information availability. Indeed, if a firm has access to accurate exogenous estimates of the attractiveness of the no-purchase option (e.g., by keeping track of the no-purchase outcomes), then the accuracy of parameter estimation is shown to improve dramatically.

An alternative way to estimate the MAM parameters is to build upon the Expectation Maximization (EM) algorithm which was developed by Vulcano et al. (2012) for estimating the parameters of the BAM when only aggregate sales data are available, which makes the standard MLE approach extremely computationally challenging. Their algorithm was later adapted by Gallego et al. (2014) to estimate the parameters of the GAM. The idea behind these algorithms is to estimate the model parameters iteratively using estimates of the first-choice demand. For large-scale problems, such an approach can be more effective than solving the least squares problem. However, similar to the algorithm presented by Gallego et al. (2014), the method based on an adaptation of the EM algorithm to the MAM suffers from an important limitation: its convergence is not theoretically guaranteed. We provide a detailed description and performance examples of this method in Appendix H of the supplementary appendix material.

7. Impact of omnichannel initiatives and sensitivity analysis

7.1. General effects of omnichannel initiatives

In this subsection, we consider widely used omnichannel initiatives and discuss general effects of their implementation on the MAM parameters. Importantly, we do not account for implementation and maintenance costs associated with these initiatives – our goal is to estimate and explore the demand evolution. In the following, we assume that channel 1 is the retailer's offline channel (i.e., a physical store or a chain of physical stores) and channel 2 is the online channel.

Table 2
Least squares estimates of the MAM parameters.

	j	$v_j^{(c)}$ True	$v_j^{(c)}$ Estimated ($v_0^{(c)}$ known)	$v_j^{(c)}$ Estimated ($v_0^{(c)}$ unknown)	$u_j^{(c)}$ True	$u_j^{(c)}$ Estimated ($v_0^{(c)}$ known)	$u_j^{(c)}$ Estimated ($v_0^{(c)}$ unknown)	$w_j^{(c)}$ True	$w_j^{(c)}$ Estimated ($v_0^{(c)}$ known)	$w_j^{(c)}$ Estimated ($v_0^{(c)}$ unknown)	$\Lambda^{(c)}$ True	$\Lambda^{(c)}$ Estimated ($v_0^{(c)}$ known)	$\Lambda^{(c)}$ Estimated ($v_0^{(c)}$ unknown)
channel 1	1	0.068	0.068	0.083	0.013	0.018	0.021	0.027	0.028	0.033	10000	10124	11991
	2	0.145	0.146	0.173	0.045	0.045	0.058	0.031	0.042	0.055			
	3	0.233	0.227	0.304	0.106	0.106	0.131	0.017	0.044	0.054			
	4	0.096	0.102	0.117	0.047	0.050	0.057	0.039	0.031	0.038			
	5	0.221	0.219	0.255	0.035	0.029	0.036	0.095	0.101	0.133			
channel 2	1	0.105	0.106	0.122	0.021	0.022	0.025	0.046	0.046	0.059	30000	29884	23254
	2	0.199	0.194	0.249	0.068	0.067	0.081	0.093	0.087	0.121			
	3	0.195	0.197	0.248	0.081	0.081	0.092	0.083	0.084	0.098			
	4	0.236	0.234	0.313	0.061	0.062	0.070	0.088	0.086	0.129			
	5	0.008	0.013	0.013	0.000	0.004	0.003	0.001	0.004	0.005			

The most straightforward effect takes place when customers have in-store access to information about the availability of online inventory, e.g., through in-store digital help desks. In this case, values of $u_j^{(1)}$, $j \in \mathcal{N}$ have to increase as ratios $u_j^{(1)}/(u_j^{(1)} + w_j^{(1)})$ determine the probability that customers are willing to switch from the offline to online channel when looking for the desired product, and offline customers have an additional incentive to check the online assortment. In other words, if product j is not available in the offline channel but it is available in the online channel, then the demand for product j in the online channel is expected to increase due to the fact that more customers switch from the offline to online channel when looking for this product. At the same time, one can expect that the shadow attractiveness values of purchasing products from another source, $u_j^{(1)} + w_j^{(1)}$, should remain constant for all $j \in \mathcal{N}$ since there should be no impact on choices of customers who select products which are offered in the offline channel. That is, the increase in demand for products in the online channel caused by this omnichannel initiative takes place on account of type-1 customers who would otherwise go to a different retailer. However, note that parameters $u_j^{(1)}$ may not increase considerably because nowadays customers have an option to access the retailer's online store using their smartphones. It is also reasonable to assume that the effects on all other parameters is negligible as no additional customers are drawn to the store, and customers willing to buy something from the in-store assortment are not likely to switch to the online channel looking for other products. Therefore, the benefits of this initiative may be outweighed by its implementation costs. It would thus be interesting for future studies to compare the additional revenue generated by increased cross-channel demand with the implementation costs of this initiative through real-world case studies.

We now investigate the effect of BOPS, which is arguably the most prevalent omnichannel initiative. We assume that BOPS orders are fulfilled from the physical store inventory. Otherwise, if a customer is only allowed to pick up a product after it is delivered from a warehouse to the store, there is no substantial difference between regular online transactions and BOPS transactions. The main distinction is in the delivery cost – if the delivery is carried out by the retailer and it is not paid separately for by customers, then the delivery cost in the latter case should at least not be higher than in the former because deliveries to the store can be organized in batches. Thus, the major issue for the retailer is to compare the benefits generated by attracting new online customers through introducing the BOPS functionality with the associated implementation costs (including those for adjusting the supply chain). This is, however, outside the scope of our research.

Within the framework of our research, it is more interesting to study the effect of BOPS if orders have to be fulfilled from the physical store inventory. This is also a common practice for omnichannel retailers, usually due to the need to have items ready for collection shortly after the order was placed (see Gallino & Moreno, 2014). In general, the overall traffic of customers should increase because customers who do not want to wait for a delivery have an additional convenient way to receive a product. However, even if we do not consider implementation costs, introducing BOPS can be unprofitable. Despite counting purchases made through BOPS as online transactions, the empirical analysis carried out by Gallino & Moreno (2014) revealed that the introduction of the BOPS functionality generally leads to a reduction in online sales and an increase in offline sales. They explained this phenomenon by the impact of sharing reliable inventory availability information on customers' decisions. If it is guaranteed that a certain product is available in-store, customers may choose to go to the store (even without reserving the product using BOPS) rather than order it online. A similar effect can take place if the retailer provides online information about the current stock level of each product in each store. As a consequence, the retailer may lose part of its online customers while attracting more in-store customers instead. This can lead to losses under the assumption that the gross profits per unit of each product in the online are higher than those in the offline channel.

Based on these assertions, we can describe the general effects of implementing the BOPS functionality on the MAM parameters as follows: Firstly, one can expect an increase in the expected number of customers visiting at least one of the retail channels, $\Lambda^{(1)} + \Lambda^{(2)}$, together with a decrease in the number of online customers $\Lambda^{(2)}$. Secondly, since online customers now have information about the in-store assortment, parameters $u_j^{(2)}$ should go up, while the shadow attractiveness values of purchasing products from another source, $u_j^{(2)} + w_j^{(2)}$, should remain constant (similar to the impact of in-store information about online inventory availability). Note that in this case, we count purchases made through BOPS as in-store purchases.

A more rigorous way to study the effect of BOPS on customer choices would be to introduce a separate channel for BOPS transactions, and suitably adjust the parameters related to other chan-

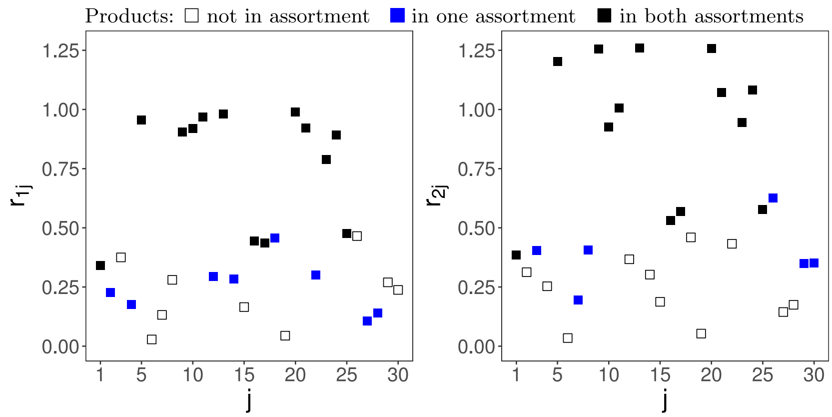


Fig. 4. Relation between the optimal assortments and r_{cj} for channels 1 (left) and 2 (right).

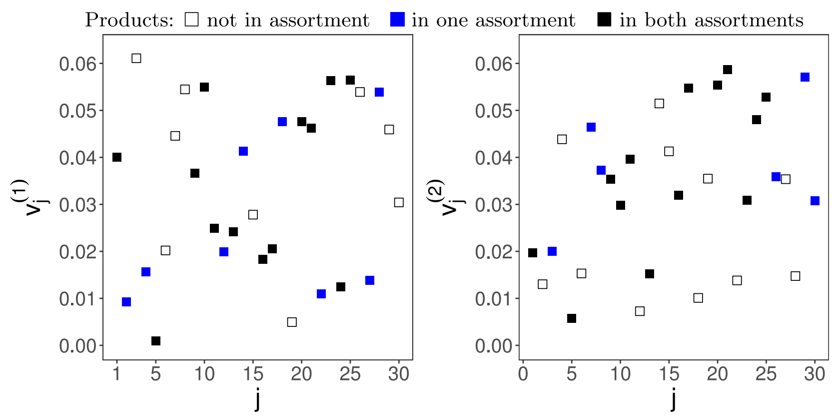


Fig. 5. Relation between the optimal assortments and $v_j^{(c)}$ for channels 1 (left) and 2 (right).

nels. This approach would also allow us to formulate an optimization problem for finding optimal assortments in all three channels, thereby determining the products for which it is profitable to implement the BOPS functionality. Future research should focus on studying the impact of this and other omnichannel initiatives in order to gain a better understanding of, and novel insights into the profitability of adopting such initiatives.

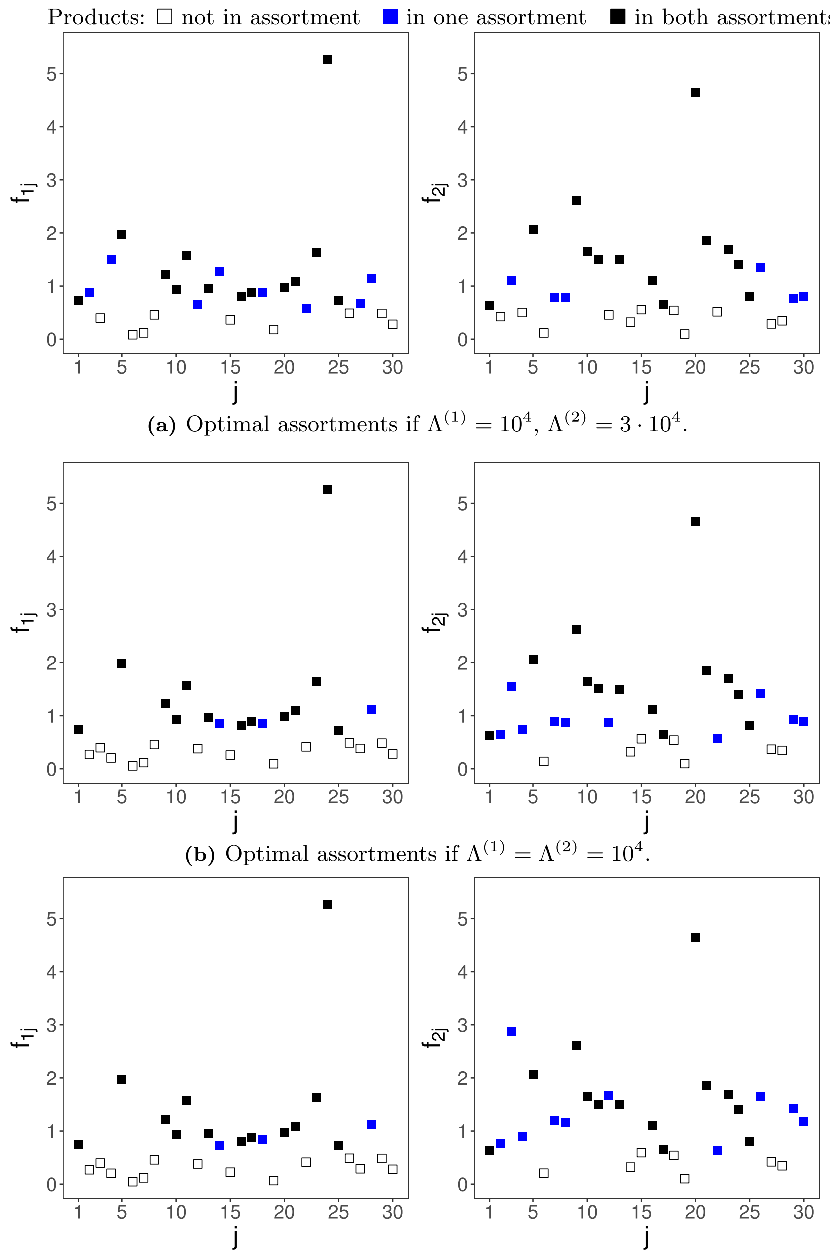
7.2. Numerical analyses and managerial insights

For the first part of our numerical analyses, we generate the MAM parameters as described in Section 5 and keep them fixed throughout the numerical experiments. More specifically, we consider the case $n = 30$ in order to be able to clearly visualize the optimal assortments. In the below figures, we indicate products that belong to the optimal assortment of both channels by black squares, products that belong to one optimal channel assortment exclusively by blue squares, and the remaining products – i.e. those not offered at all – by white squares (similar to Fig. 1). Each figure consists of two plots, with the plot(s) on the left- and right-hand side corresponding to channel 1 (offline) and channel 2 (online), respectively.

It is important to recall that the generated parameter values are convenient for carrying out this analysis, but they are not necessarily realistic. For example, one can expect that if the gross profit per unit of a certain product j is high compared to other products, then the proportion of the demand for product j to the total demand in the corresponding channel is probably small, i.e., there is a negative correlation between the values of r_{cj} and $v_j^{(c)}$. Moreover, we assume that the values of r_{cj} and $v_j^{(c)}$ are uniformly distributed, which is unlikely to be the case in practice. However, our goal is to study the relationship between the parameters values and the optimal solution to the SBMILP, rather than to investigate a real-world case study, and using such generated data ideally fits this purpose.

First, let us explore the relation between the gross profit per unit of each product and the optimal assortments given by the optimal solution of the SBMILP (see Fig. 4). We observe that the higher the gross profit per unit of a product, the more likely it is that this product belongs to the optimal assortment. However, such a relation is not always evident, i.e., a product with a relatively high unit profit in one of the channels may not be included in the corresponding assortment, whereas a product with a relatively low unit profit may be offered. Moreover, the relationship between unit profits and optimal assortments seems to be more pronounced in channel 2 than in channel 1. It can possibly be explained by the fact that we set the total number of type-2 customers to be considerably higher than that of type-1 customers ($\Lambda^{(1)} = 10^4$ and $\Lambda^{(2)} = 3 \cdot 10^4$), which could reflect a situation where channel 2 represents an online channel. Also, note that if a product is not included in the assortment in one channel, and the profit per unit of the product is low in both channels, it can still be profitable to offer this product in the other channel due to demand generated by customers who switch to that channel and those who shop there in the first place.

Next, we create a similar visualization but for the attractiveness values $v_j^{(c)}$ instead of the gross profit values, with the results shown by Fig. 5. Unlike Fig. 4, it can be observed that there is no



(a) Optimal assortments if $\Lambda^{(1)} = 10^4$, $\Lambda^{(2)} = 3 \cdot 10^4$.
(b) Optimal assortments if $\Lambda^{(1)} = \Lambda^{(2)} = 10^4$.
(c) Optimal assortments if $\Lambda^{(1)} = 3 \cdot 10^4$, $\Lambda^{(2)} = 10^4$.

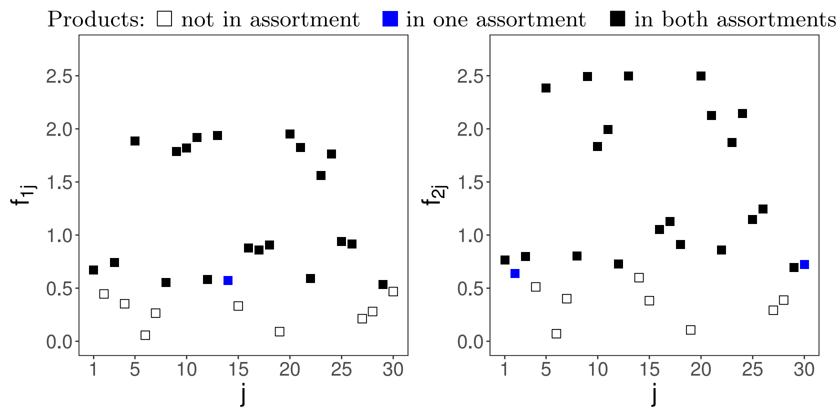
Fig. 6. Optimal assortments for different values of $\Lambda^{(1)}/\Lambda^{(2)}$ for channels 1 (left) and 2 (right).

apparent relation. To further analyze the structure of optimal assortments, recall the finding formulated in Proposition 5. Based on $F_c(j, S_c, S_{\bar{c}})$, we can derive a more tractable expression that approximately characterizes the optimal assortment in one channel given the assortment in the other channel. Indeed, consider the following expression:

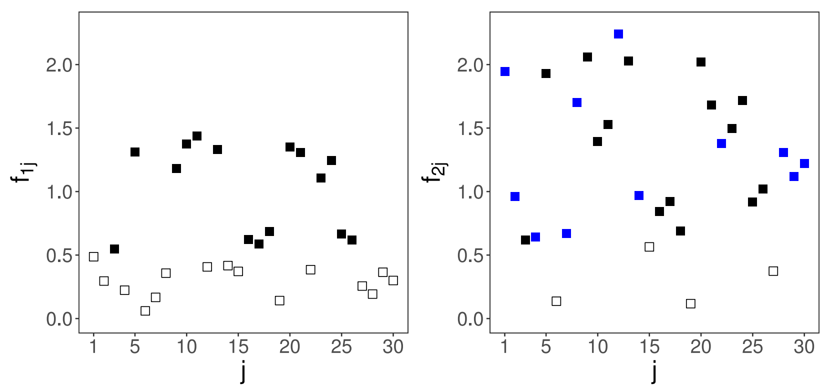
$$f_{cj}(S_{\bar{c}}) = \begin{cases} \dfrac{r_{cj}v_j^{(c)} - r_{\bar{c}j}u_j^{(c)}}{\tilde{v}_j^{(c)}} & \text{if } j \in S_{\bar{c}}, \\ \dfrac{r_{cj}v_j^{(c)} + r_{cj}u_j^{(\bar{c})}\Lambda^{(\bar{c})}/\Lambda^{(c)}}{\tilde{v}_j^{(c)}} & \text{otherwise.} \end{cases} \quad (16)$$

The values of $f_{cj}(S_{\bar{c}})$ $\forall c \in \mathcal{C},\ j \in \mathcal{N}$ for the optimal assortments are illustrated by Fig. 6a. We observe that unlike the attractiveness values, these expressions do provide an approximate characterization of the optimal assortments. Indeed, as can be seen, if $f_{cj}(S_{\bar{c}}) > f_{ck}(S_{\bar{c}})$ and product k belongs to the optimal assortment in channel c, then product j tends to belong to the optimal assortment as well. In the following, we use the values of f_{cj} to produce more illustrative plots.

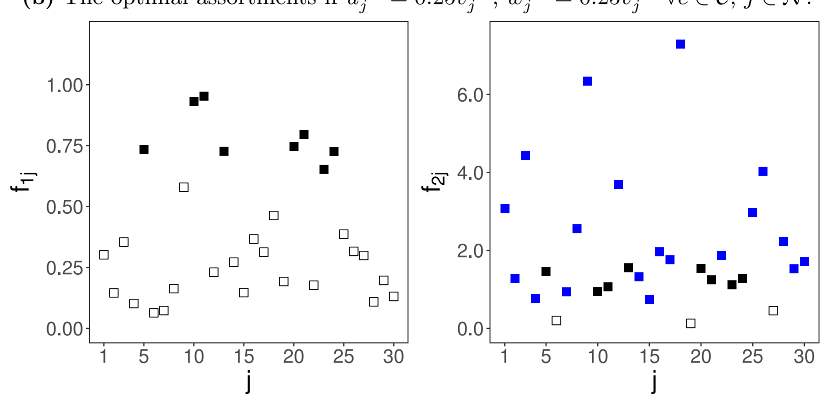
To study how the optimal assortments are being affected by a change in the ratio $\Lambda^{(1)}/\Lambda^{(2)}$, we consider two additional cases: $\Lambda^{(1)} = \Lambda^{(2)} = 10^4$ and $\Lambda^{(1)} = 3 \cdot 10^4$, $\Lambda^{(2)} = 10^4$. The results are summarized in Fig. 6b and 6 c. Interestingly, it can be seen that with an increase in the ratio $\Lambda^{(1)}/\Lambda^{(2)}$ from 1/3 to 1, the optimal assortment in channel 1 becomes smaller in size, whereas the optimal assortment in channel 2 becomes larger. However, with a further increase in the ratio $\Lambda^{(1)}/\Lambda^{(2)}$ from 1 to 3, the optimal assortments in both channels do not change. To sum up, the smaller the ratio of customers whose primary choice is to shop online to those whose primary choice is to go to a retail store, the smaller the optimal assortment in the physical channel and the larger the assortment in the online channel (up to a certain limit). This is intuitively clear given that an online purchase is generally more



(a) The optimal assortments if $u_j^{(c)} = 0.01v_j^{(c)}$, $w_j^{(c)} = 0.49v_j^{(c)}$ $\forall c \in \mathcal{C}$, $j \in \mathcal{N}$.



(b) The optimal assortments if $u_j^{(c)} = 0.25v_j^{(c)}$, $w_j^{(c)} = 0.25v_j^{(c)}$ $\forall c \in \mathcal{C}$, $j \in \mathcal{N}$.



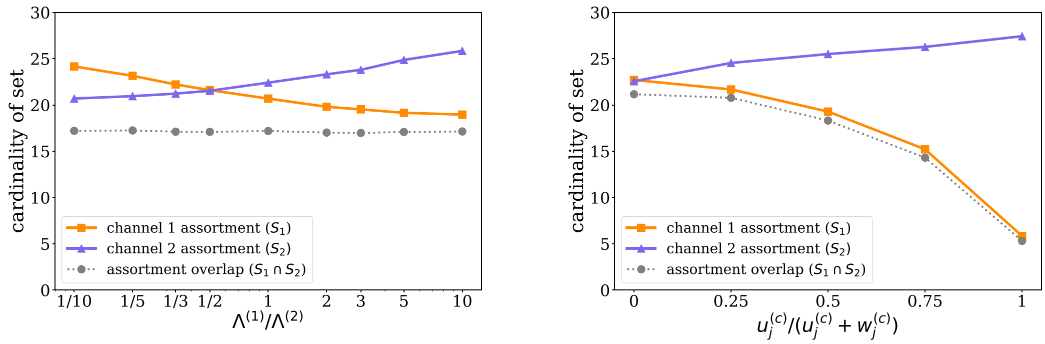
(c) The optimal assortments if $u_j^{(c)} = 0.49v_j^{(c)}$, $w_j^{(c)} = 0.01v_j^{(c)}$ $\forall c \in \mathcal{C}$, $j \in \mathcal{N}$.

Fig. 7. Optimal assortments for different values of $u_j^{(c)}$ and $w_j^{(c)}$ for channels 1 (left) and 2 (right).

profitable for the retailer than the in-store purchase of the same product, so the retailer is interested in a high online traffic and therefore increases the online assortment.

We now investigate the effect of the values of $u_i^{(c)}/(u_i^{(c)} + w_i^{(c)})$ on the optimal assortments. These ratios determine how willing customers are to switch to another channel when looking for the desired product. We therefore fix the values of all parameters except for the values of $u_j^{(c)}$ and $w_j^{(c)}$, and also set $u_j^{(c)} + w_j^{(c)} = 0.5v_j^{(c)}$. We then consider the following three cases: $u_j^{(c)} = 0.01v_j^{(c)}$, $u_j^{(c)} = 0.25v_j^{(c)}$ and $u_j^{(c)} = 0.49v_j^{(c)}$ $\forall c \in \mathcal{C}$, $j \in \mathcal{N}$. The optimal assortments are displayed in Fig. 7. It is interesting to note that a similar effect to the one described above for different values of $\Lambda^{(1)}/\Lambda^{(2)}$ can be observed when increasing the ratio $u_j^{(c)}/(u_j^{(c)} + w_j^{(c)})$: given a fixed sum $u_j^{(c)} + w_j^{(c)}$ $\forall c \in \mathcal{C}$, $j \in \mathcal{N}$, the optimal assortment in channel 1 becomes smaller in size, whereas the optimal assortment in channel 2 becomes larger. In other words, the larger the proportion of customers willing to switch from one channel to another in case of absence of a certain product, the larger the optimal assortment in the online channel and the smaller the optimal assortment in the physical channel. Trivially, in the case when 100% of in-store customers are willing to switch to the online channel if their primary-choice product is not available, the optimal assortment in the physical channel is the empty set.

Next, we investigate the extent to which the implementation of the BOPS functionality affects the total profit of the retailer. To this end, we use the initially generated values of parameters $u_j^{(c)}$ and $w_j^{(c)}$ $\forall c \in \mathcal{C}$, $j \in \mathcal{N}$. Recall that the impact of introducing the BOPS functionality on the MAM parameters can be described approxi-



(a) Impact of ratio of numbers of different customer types.

(b) Impact of proportions of customers switching to another channel.

Fig. 8. Effect of ratio of physical to online traffic (left) and of proportions of customers switching channels (right) on both the average size of assortments ($|S_1|$, $|S_2|$) and the degree of assortment overlap ($|S_1 \cap S_2|$).

Table 3
Different impacts of the BOPS implementation on total profit.

change in $\Lambda^{(1)}$	change in $\Lambda^{(2)}$	change in $u_j^{(2)}$	objective value
–	–	–	24025.16
7500	−6000	–	23489.87
9000	−6000	–	24133.95
10500	−9000	–	22362.52
12000	−9000	–	22931.03
13500	−9000	–	23499.54
13500	−9000	$0.2w_j^{(2)}$	23514.09
13500	−9000	$0.5w_j^{(2)}$	23582.26
13500	−9000	$0.8w_j^{(2)}$	23723.54
15000	−9000	–	24068.05

mately as follows: the sum $\Lambda^{(1)} + \Lambda^{(2)}$ increases, $\Lambda^{(2)}$ decreases, and $u_j^{(2)}$ increases under the condition that the sum $u_j^{(2)} + w_j^{(2)}$ does not change $\forall j \in \mathcal{N}$. We carried out several numerical experiments considering different degrees of such impact and report results in Table 3. Based on these results, the following key observations can be made. The implementation of the BOPS functionality can be unprofitable if the proportion of online customers using this option is too large compared to the additional traffic attracted to the offline channel. Furthermore, the primary effect on total profit is due to the change in the parameters $\Lambda^{(1)}$ and $\Lambda^{(2)}$, whereas the effect of an increase of $u_j^{(2)}$-values is less pronounced. This is not surprising since $u_j^{(2)}$-values only reflect the cross-channel demand volume generated by customers switching from channel 2 to channel 1, while $\Lambda^{(1)} + \Lambda^{(2)}$ represents the total number of customers visiting the retailer. Ultimately, our results indicate that the profitability of adopting the BOPS functionality needs to be evaluated on a case-by-case basis, which is in line with the empirical findings of Gallino & Moreno (2014).

Unlike the above analyses, which were carried out for illustrative purposes for one problem instance, we now analyze the effect of the considered parameter changes on both the size of assortments and the degree of assortment overlap for a large number of problem instances (simulated in the same way). First, we simulate 1000 problem instances, and for each problem instance, we fix all the parameters except for $\Lambda^{(1)}$ and $\Lambda^{(2)}$. We then consider several cases with different values of the $\Lambda^{(1)}/\Lambda^{(2)}$-ratio. Note that the absolute values of $\Lambda^{(1)}$ and $\Lambda^{(2)}$ are irrelevant since the objective function in the SBMILP can be divided by the constant $\Lambda^{(2)}$ (leaving the optimal solution unchanged), which makes it a function of $\Lambda^{(1)}/\Lambda^{(2)}$. Subsequently, we solve the SBMILP in each considered case and record the size of assortments in both channels and the degree of assortment overlap. The aggregated results are shown in Fig. 8(a) (note that the x-axis is log-scaled). We observe that the average assortment sizes follow the same trends as the ones described earlier for one problem instance: with an increase in the ratio $\Lambda^{(1)}/\Lambda^{(2)}$, on average the size of assortment in channel 1 (physical channel) decreases while the size of assortment in channel 2 (online channel) increases. Interestingly, we also observe that on average, the degree of assortment overlap remains virtually constant. This can be explained by fact that the trend lines are almost symmetric, that is the absolute values of their slopes are very close to each other.

Next, we study the effect of the values of $u_i^{(c)}/(u_i^{(c)} + w_i^{(c)})$ on both the size of assortments and the degree of assortment overlap. As previously, we simulate 1000 problem instances. For each problem instance, we fix all parameters except for u and w. Similar to the analysis of one problem instance, we set $u_i^{(c)} + w_i^{(c)} = 0.5v_i^{(c)}$ and consider several cases with the ratio $u_i^{(c)}/(u_i^{(c)} + w_i^{(c)})$ taking values from 0 to 1. The results are summarized by Fig. 8(b). We can see that if $u_i^{(c)}/(u_i^{(c)} + w_i^{(c)})$ tends zero – that is if there is no cross-channel demand – then the average sizes of assortments in both channels are the same. This is not surprising because if there is no cross-channel demand, then each assortment can be optimized independently. Since the parameters related to each channel are simulated randomly, on average the assortment sizes are virtually identical, i.e. $|S_1| \approx |S_2|$. Also, we can observe that on average, an increase in $u_i^{(c)}/(u_i^{(c)} + w_i^{(c)})$ leads to an increase in the size of channel 1 assortment ($|S_1|$) and a decrease in the size of channel 2 assortment ($|S_2|$). This is consistent with our previous observations for one problem instance. Lastly, we observe that the line representing the average overlap degree closely follows the line representing the average size of channel 1 assortment – in other words, there are hardly any products that belong to channel 1 assortment but not to channel 2 assortment. This might be explained by the fact that since channel 2 assortment comprises almost all products, the few products that are not part of it are characterized by very low unit gross profits. In this case, their unit gross profits are also low in channel 1 (due to the parameter simulation procedure), making them also less likely to be part of the optimal assortment in that channel.

Finally, let us evaluate the revenue benefits of solving the omnichannel assortment problem as opposed to optimizing the two assortments in siloed channels, which is often referred to as multichannel assortment optimization. To this end, we consider several problem sizes with the number of products, n, ranging from 30 to 300. For each n, we simulate 1000 problem instances in the way described at the beginning of Section 5. Then, for each problem instance, we compute two revenues: one that corresponds to omnichannel assortment optimization; and one that corresponds to

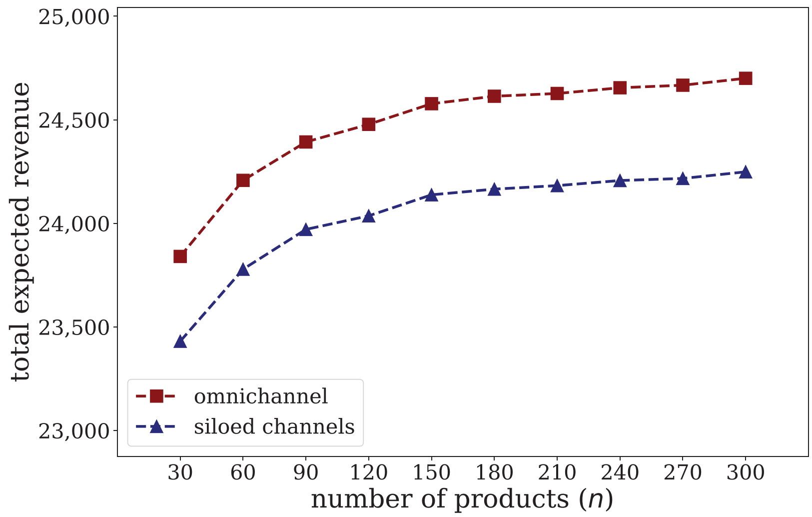


Fig. 9. Profitability of omnichannel assortment optimization and assortment optimization in siloed channels.

assortment optimization in siloed channels. The former is obtained by solving the SBMILP, whereas the latter is determined in the following way: First, we solve the assortment optimization problem under the GAM for each channel separately, using the fact that the MAM restricted to choices of type-c customers in channel c is equivalent to the GAM. Once assortments in the two channels are identified, we compute the corresponding total expected revenue assuming that the customers make their choices according to the MAM. Importantly, by doing this we do not neglect the cross-channel demand – even though the assortments are optimized in siloed channels, the underlying demand model is the same as in the omnichannel case. Fig. 9 shows that there is a clear benefit of solving the omnichannel assortment optimization problem as opposed to optimizing assortments in siloed channels. In fact, for the n-values under consideration, the omnichannel solution turned out to between 1.8% and 1.9% more profitable on average than the solution obtained for siloed channels, with maximum improvements of up to 6.7%, highlighting the considerable revenue gains that may be achieved by omnichannel assortment planning.

8. Conclusions and future work

In this paper, we have developed an analytical framework for both modeling product demand and making assortment decisions in an omnichannel environment. In particular, we have introduced a discrete choice model referred to as the multichannel attraction model (MAM) that specifically accounts for the complex nature of omnichannel shopping behavior. Compared to the single-channel setup – which corresponds to the general attraction model (GAM) – the multichannel structure of our choice model substantially increases the complexity of associated problems. For example, the assortment optimization problem under the GAM can be formulated as a linear program, and the optimal assortment can actually be found analytically. For the MAM, on the other hand, we have formulated the sales-based mixed integer linear program (SBMILP) – a tight MILP formulation of the corresponding assortment optimization problem – and proved that the optimal assortment in one channel can be found analytically if all products are available in the remaining channel. We have proposed a computationally efficient heuristic method to approximately solve the SBMILP, and showed numerically that its output is extremely close to the optimal solution. We have also presented three different methods to estimate the parameters of the MAM, and demonstrated that if product demands are only known for a limited number of assortments in each channel, the MAM parameters can be estimated fairly accurately by simply solving a least squares problem.

We have analyzed general effects of the implementation of widely-used omnichannel initiatives on the MAM parameters, and have carried out numerical experiments to investigate the structure of optimal assortments. We demonstrated that in an omnichannel environment, optimal assortments cannot be characterized by a single factor alone as they are affected by a combination of several factors. In fact, in our experiments we identified a relation between the sizes of optimal assortments and the following two factors: the ratio of customers whose primary choice is to shop online to those whose primary choice is to go to a retail store, and the proportion of customers willing to switch from one channel to another in case of absence of a certain product. We also showed numerically that implementing the buy-online-and-pick-up-in-store (BOPS) initiative is not always profitable, which supports previous findings in the omnichannel literature. Finally, we evaluated the benefits of omnichannel assortment optimization as opposed to optimizing siloed assortments (also known as multichannel assortment optimization) and showed that the former can result in substantial revenue gains for omnichannel retailers. In fact, our numerical analysis showed that the omnichannel solution is 1.8%-1.9% more profitable on average than the solution obtained for siloed channels, with maximum gains of up to 6.7%. These findings are encouraging and demonstrate that our framework can be beneficially used for omnichannel assortment planning as well as for exploring the profitability of implementing different omnichannel initiatives, which should be of great interest to decision-makers in the retailing industry.

The proposed framework can be the basis for a range of important further developments. Firstly, a better understanding of the structure of optimal assortments in each channel could be obtained through an extensive numerical analysis based on values of the MAM parameters estimated using a real-world dataset. Secondly, additional theoretical results related to the assortment optimization problem under the MAM could be derived, with the question of NP-hardness of the SBMILP being of particular interest. In a similar vein, the possibility of developing an FPTAS (e.g. along the lines of the recent work of Désir, Goyal, & Zhang (2022)) may be explored. Also, the MAM could be subject to multiple extensions. For example, the MAM could be calibrated on the product features level. In this case, it would be very interesting to compare the results obtained using such a modified version of the MAM to those obtained by Dzyabura & Jagabathula (2018) and Lo & Topaloglu

(2022). Another promising research direction would be to formulate and explore a stochastic version of the SBMILP for the case of uncertain demand in each channel. It would also be beneficial to study the effects of BOPS on customer choices by introducing a separate channel for BOPS transactions. Such an approach would allow to investigate not only whether but also for which specific products it is profitable to implement the BOPS initiative. Finally, deriving analytical properties of solutions to the SBMILP should lead to additional managerial insights and result in a more efficient SBMILP formulation.

Acknowledgments

The authors thank the three anonymous reviewers for many insightful comments and valuable suggestions. The work of the first and the last author was supported by the Swiss National Science Foundation (SNSF) under Grant 192545.

Appendix A. The MAM Formulation in the General Setup

Suppose that $\mathcal{C} = \{1, \dots, K\}$ is the set of channels with $K > 2$. We adjust and extend the notation presented in Section 3 in the following way. Let $u_{di}^{(c)}/v_i^{(c)}$ be the proportion of type-c customers willing to purchase product $i \in \mathcal{N}\backslash S_c$ in channel $d \in \mathcal{C}\backslash\{c\}$ if it is not available in channel c, and let $u_i^{(c)}$ denote the sum $\sum_{d\in\mathcal{C}\backslash\{c\}} u_{di}^{(c)}$. In other words, $u_i^{(c)}/v_i^{(c)}$ is the proportion of type-c customers willing to purchase product i in any of the retailer's channels if it is not available in channel c.

Then, the choice probabilities under the MAM given assortments in all channels can be defined by analogy to the two-channel case. The probability that a type-c customer buys product j in channel c is

$$\pi_{cj}^{(c)}(S_c) = \begin{cases} \dfrac{v_j^{(c)}}{v_0^{(c)} + \sum\limits_{k\in S_c} v_k^{(c)} + \sum\limits_{i\in\mathcal{N}\backslash S_c}(u_i^{(c)} + w_i^{(c)})} & \text{if } j \in S_c, \\ 0 & \text{otherwise;} \end{cases}$$

and the probability that a type-c customer buys product j in channel d is

$$\pi_{dj}^{(c)}(S_c, S_d) = \begin{cases} \dfrac{u_{dj}^{(c)}}{v_0^{(c)} + \sum\limits_{k\in S_c} v_k^{(c)} + \sum\limits_{i\in\mathcal{N}\backslash S_c}(u_i^{(c)} + w_i^{(c)})} & \text{if } j \in S_d\backslash S_c, \\ 0 & \text{otherwise.} \end{cases}$$

Appendix B. Proof of Proposition 1

First, note that the MAM restricted to type-c customers and products in channel c is equivalent to the GAM for products in this channel. At the same time, Blanchet et al. (2016) provided the set of MCCM parameters under which the choice probabilities can be expressed as the GAM probabilities. Suppose that we only consider choices of type-c customers in channel c. Then, the MCCM parameters that result in the GAM probabilities are as follows (see Blanchet et al., 2016):

$$\begin{aligned} &\lambda_{j_c}^{(c)} = v_j^{(c)}, \quad \rho_{j_c i_c}^{(c)} = \frac{v_i^{(c)}(v_j^{(c)} - u_j^{(c)} - w_j^{(c)})}{v_j^{(c)} - v_j^{(c)}(v_j^{(c)} - u_j^{(c)} - w_j^{(c)})}, \\ &\rho_{j_c 0}^{(c)} = \frac{v_0^{(c)}(v_j^{(c)} - u_j^{(c)} - w_j^{(c)}) + u_j^{(c)} + w_j^{(c)}}{v_j^{(c)} - v_j^{(c)}(v_j^{(c)} + u_j^{(c)} + w_j^{(c)})}. \end{aligned} \tag{B.1}$$

Now, let us also take into consideration products in channel $S_{\bar{c}}$. We therefore split the transition probability from product j_c to the no-purchase alternative into two parts: one corresponds to switching to the no-purchase alternative directly, and the other one corresponds to first purchasing product $j_{\bar{c}}$, and then – in case this product is not available – selecting the no-purchase option with probability 1. Formally, let us define

$$\rho_{j_c 0}^{(c)} = \frac{v_0^{(c)}(v_j^{(c)} - u_j^{(c)} - w_j^{(c)}) + u_j^{(c)} + w_j^{(c)}}{v_j^{(c)} - v_j^{(c)}(v_j^{(c)} + u_j^{(c)} + w_j^{(c)})} - \frac{u_j^{(c)}}{v_j^{(c)}},$$

$$\rho_{j_c j_{\bar{c}}}^{(c)} = \frac{u_j^{(c)}}{v_j^{(c)}}, \quad \rho_{j_{\bar{c}} 0}^{(c)} = 1,$$

and leave $\lambda_{j_c}^{(c)} = v_j^{(c)}$ and $\rho_{j_c i_c}^{(c)}$ unchanged as in (B.1).

Finally, let us set the remaining parameters $\lambda_{j_{\bar{c}}}^{(c)}$, $\rho_{j_c i_{\bar{c}}}^{(c)}$, $\rho_{j_{\bar{c}} i_{\bar{c}}}^{(c)}$, $\rho_{j_{\bar{c}} j_c}^{(c)}$ and $\rho_{j_{\bar{c}} i_c}^{(c)}$ to be zero, which gives us exactly the set of parameters shown in (7). By construction, if type-c customers make their choices according to the MCCM with this set of parameters, then the probability that such a customer buys product j in channel c is exactly the MAM probability (5). What remains to be proven is that the probability of such a customer purchasing product j in channel $\bar{c}$ is the MAM probability (6).

Suppose that product j is offered in channel $\bar{c}$. Note that type-c customers interested in this product consider alternatives only in channel c before switching to channel $\bar{c}$. Therefore, in order to obtain probability $\pi_{\bar{c}j}^{(c)}(S_c, S_{\bar{c}})$, we can consider a subchain of our constructed Markov chain comprising all products in channel c, product j in channel $\bar{c}$, and the no-purchase alternative. Following Blanchet et al. (2016), let B denote the transition probability submatrix from alternatives that are not in the assortment to alternatives from the assortment (or to the no-purchase alternative), and let C denote the transition probability submatrix among the alternatives that are not in the assortment. Furthermore, let $\hat{\lambda}$ be the vector of arrival probabilities to alternatives that are not in the assortments. Finally, let $e_{j_{\bar{c}}}$ be the standard unit vector such that the product $Be_{j_{\bar{c}}}$ corresponds to the vector of transition probabilities from alternatives that are not in the assortment to product j in channel $\bar{c}$. Then, using the formula for Markov chain choice probabilities (see Blanchet et al., 2016), we obtain the following:

$$\begin{aligned} \pi_{\bar{c}j}^{(c)}(S_c, S_{\bar{c}}) &= 0 + \hat{\lambda}^T(I - C)^{-1}Be_{j_{\bar{c}}} = \hat{\lambda}^T\left(\sum_{q=0}^{\infty} C^q\right)Be_{j_{\bar{c}}} \\ &= \left(\sum_{q=0}^{\infty}\left(\sum_{i\in\mathcal{N}\backslash S_c} v_i^{(c)} - \sum_{i\in\mathcal{N}\backslash S_c}(u_i^{(c)} + w_i^{(c)})\right)^q\right)\hat{\lambda}^T Be_{j_{\bar{c}}} \\ &= \frac{\hat{\lambda}^T Be_{j_{\bar{c}}}}{1 - \left(\sum_{i\in\mathcal{N}\backslash S_c} v_i^{(c)} - \sum_{i\in\mathcal{N}\backslash S_c}(u_i^{(c)} + w_i^{(c)})\right)} \\ &= \frac{v_j^{(c)}\rho_{j_c j_{\bar{c}}}^{(c)}}{v_0^{(c)} + \sum_{k\in S_c} v_k^{(c)} + \sum_{i\in\mathcal{N}\backslash S_c}(u_i^{(c)} + w_i^{(c)})} \\ &= \frac{u_j^{(c)}}{v_0^{(c)} + \sum_{k\in S_c} v_k^{(c)} + \sum_{i\in\mathcal{N}\backslash S_c}(u_i^{(c)} + w_i^{(c)})}, \end{aligned}$$

which concludes the proof.

Appendix C. Proof of Proposition 2

Recall that the GAM can be viewed as a limited case of the nested logit model where the dissimilarity parameter of each nest goes to zero (see Gallego et al., 2014). In this model, each nest corresponds to a product, and alternatives within the nest correspond to different sources where this product can be bought. Since the dissimilarity parameter of each nest tends to zero, it can be assumed that each nest comprises only two sources: The retailer itself (with the highest utility overall) and the outside source (with the highest utility among other available sources). Since the nested logit model is a RUM, the GAM is a RUM as well and it can be represented through a distribution over rankings (permutations) of alternatives. Readers are referred to Block & Marschak (1959), who showed how to construct the distribution over rankings from the joint distribution of random utilities, and vice versa. Importantly, if the number of products is n, then each ranking comprises $2n+1$ alternatives since each product can be bought either from the retailer or from the outside source (and there is also the no-purchase alternative).

We can use the distribution over rankings that is consistent with the GAM to construct the distribution over rankings that is consistent with the MAM choice probabilities for one customer type. First, note that the choice probabilities of type-c customers selecting alternatives from channel c are exactly the same as the GAM choice probabilities defined over this set of products. Let us consider the corresponding distribution over rankings. One can insert the alternative of buying a certain product j from channel $\bar{c}$ into the rankings and modify the distribution in such a way that the probabilities of choosing products from channel c do not change while the probability of choosing product j from channel $\bar{c}$ is exactly the same as the MAM choice probability. Indeed, let us consider a ranking $r=(r_1, r_2, \dots r_{2n+1})$ with $\mathbb{P}(r)=p$. Suppose that r_k corresponds to the option of buying product j from the outside source (which is always available). Let s denote the alternative of buying a certain product j from channel $\bar{c}$. Then, let us replace ranking r with two rankings $r'=(r_1, r_2, \dots, r_{k-1}, s, r_k, r_{k+1} \dots, r_{2n+1})$ and $r''=(r_1, r_2, \dots, r_{k-1}, r_k, s, r_{k+1} \dots, r_{2n+1})$, such that $\mathbb{P}(r')=pu_j^{(c)}/(u_j^{(c)}+w_j^{(c)})$ and $\mathbb{P}(r'')=pw_j^{(c)}/(u_j^{(c)}+w_j^{(c)})$. Note that $u_j^{(c)}/(u_j^{(c)}+w_j^{(c)})$ is exactly the probability of type-c customers switching to channel $\bar{c}$ given that they are willing to purchase product j outside of channel c. If we do this for each ranking, then the probabilities of choosing products from channel c will not change by construction of rankings r' and r'' and since $\mathbb{P}(r')+\mathbb{P}(r'')=\mathbb{P}(r)$ for all rankings r. At the same time, the probability of choosing product $j \in S_{\bar{c}}$ will be defined as the probability of choosing the outside option according to the GAM multiplied by the coefficient $u_j^{(c)}/(u_j^{(c)}+w_j^{(c)})$, that is

$$\frac{(u_j^{(c)}+w_j^{(c)})\mathbb{1}_{j\in\mathcal{N}\setminus S_c}}{v_0^{(c)}+\sum_{k\in S_c} v_k^{(c)}+\sum_{i\in\mathcal{N}\setminus S_c}(u_i^{(c)}+w_i^{(c)})}\cdot\frac{u_j^{(c)}\mathbb{1}_{j\in S_{\bar{c}}}}{u_j^{(c)}+w_j^{(c)}}$$
$$=\frac{u_j^{(c)}\mathbb{1}_{j\in S_{\bar{c}}\setminus S_c}}{v_0^{(c)}+\sum_{k\in S_c} v_k^{(c)}+\sum_{i\in\mathcal{N}\setminus S_c}(u_i^{(c)}+w_i^{(c)})},$$

which is exactly the MAM choice probability $\pi_{\bar{c}j}^{(c)}(S_c, S_{\bar{c}})$. If we repeat this procedure n times for all products in channel $\bar{c}$, then we obtain the distribution over rankings that is consistent with the MAM choice probabilities. The MAM is thus a mixture of RUMs.

Finally, note that a mixture of RUMs is also a RUM. Indeed, from the distributions over rankings of alternatives consistent with each individual model in the mixture, one can straightforwardly construct a distribution over rankings corresponding to the mixture model. In the case of the MAM, suppose that the distribution over rankings that corresponds to choice probabilities of type-c customers is as follows:

$$D^{(c)}=\{(r^1, p_1^{(c)}), (r^2, p_2^{(c)}), \dots, (r^h, p_h^{(c)})\},$$

where each r^i is a ranking and $p_i^{(c)}$ is the associated probability. Similarly, let

$$D^{(\bar{c})}=\{(r^1, p_1^{(\bar{c})}), (r^2, p_2^{(\bar{c})}), \dots, (r^h, p_h^{(\bar{c})})\}$$

be the distribution over rankings of alternatives that corresponds to choice probabilities of type-$\bar{c}$ customers. Note that both distributions $D^{(c)}$ and $D^{(\bar{c})}$ are defined over the same set of rankings of alternatives which is composed of all permutations of all possible purchase outcomes (including the no-purchase option). In particular, each ranking r^i contains $3n+1$ alternatives (the no-purchase option together with products in channel c, channel $\bar{c}$, and the outside source) and h equals the total number of permutations of alternatives, i.e., $h=(3n+1)!$. Then, it is easy to see that the following distribution corresponds to the MAM choice probabilities:

$$D=\{(r^1, \omega p_1^{(c)}+(1-\omega)p_1^{(\bar{c})}), (r^2, \omega p_2^{(c)}+(1-\omega)p_2^{(\bar{c})}), \dots,$$
$$(r^h, \omega p_h^{(c)}+(1-\omega)p_h^{(\bar{c})})\},$$

where $\omega=\Lambda^{(c)}/(\Lambda^{(c)}+\Lambda^{(\bar{c})})$. Thus, the MAM is also a RUM.

Appendix D. Proof of Theorem 1

In order to prove the theorem, we need to show the following:

1) Let $S_c, S_{\bar{c}}$ be arbitrary subsets of $\mathcal{N}$. If $x_{cj}^{(c)}, x_{\bar{c}j}^{(c)}$ correspond to probabilities (5), (6) multiplied by $\Lambda^{(c)}$, z_{cj} take the corresponding binary values, and $x_{c0}^{(c)}$ satisfy

$$x_{c0}^{(c)}=\frac{v_0^{(c)}\Lambda^{(c)}}{v_0^{(c)}+\sum_{k\in S_c} v_k^{(c)}+\sum_{i\in\mathcal{N}\setminus S_c}(u_i^{(c)}+w_i^{(c)})},$$

then $\{x_{c0}^{(c)}, x_{cj}^{(c)}, x_{\bar{c}j}^{(c)}, z_{cj}\}_{c\in\mathcal{C}, j\in\mathcal{N}}$ is a feasible solution to the SBMILP.

2) If $\{x_{c0}^{(c)}, x_{cj}^{(c)}, x_{\bar{c}j}^{(c)}, z_{cj}\}_{c\in\mathcal{C}, j\in\mathcal{N}}$ is an optimal solution to the SBMILP, then there exist offer sets $S_c, S_{\bar{c}} \subseteq \mathcal{N}$ such that $x_{cj}^{(c)}, x_{\bar{c}j}^{(c)}$ correspond to probabilities (5), (6) multiplied by $\Lambda^{(c)}$, and z_{cj} take the corresponding binary values.

The first part can be shown straightforwardly by substituting $\{x_{c0}^{(c)}, x_{cj}^{(c)}, x_{\bar{c}j}^{(c)}, z_{cj}\}_{c\in\mathcal{C}, j\in\mathcal{N}}$ into the SBMILP constraints. For the second part, note that from constraints (9e) and the form of the objective function it follows that $z_{cj}=1$ if and only if $x_{cj}^{(c)}>0$ (since z_{cj} are binary and $x_{cj}^{(c)}$ are nonnegative). Furthermore, from constraints (9c) and (9d) we can see that

$$\frac{x_{cj}^{(c)}}{v_j^{(c)}}=\begin{cases}\frac{x_{c0}^{(c)}}{v_0^{(c)}} & \text{if } z_{cj}=1,\\ 0 & \text{otherwise.}\end{cases} \tag{D.1}$$

Let us set $S_c=\{j\in\mathcal{N}: \ z_{cj}=1\}$. Substituting (D.1) into constraints (9b), we obtain the following:

$$\frac{\tilde{v}_0^{(c)}}{v_0^{(c)}}x_{c0}^{(c)}+\sum_{j\in\mathcal{N}}\frac{\tilde{v}_j^{(c)}}{v_j^{(c)}}x_{cj}^{(c)}=\frac{\tilde{v}_0^{(c)}}{v_0^{(c)}}x_{c0}^{(c)}+\sum_{k\in S_c}\frac{\tilde{v}_k^{(c)}}{v_0^{(c)}}x_{c0}^{(c)}=\Lambda^{(c)}.$$

Hence,

$$x_{c0}^{(c)}=\frac{v_0^{(c)}\Lambda^{(c)}}{\tilde{v}_0^{(c)}+\sum_{k\in S_c}\tilde{v}_k^{(c)}}=\frac{v_0^{(c)}\Lambda^{(c)}}{v_0^{(c)}+\sum_{k\in S_c} v_k^{(c)}+\sum_{i\in\mathcal{N}\setminus S_c}(u_i^{(c)}+w_i^{(c)})}. \tag{D.2}$$

Subsequently, by combining (D.1) and (D.2), we obtain the desired values of $x_{cj}^{(c)}$. Finally, note that constraints (9c) and (9f) imply that $x_{\bar{c}j}^{(c)}\neq 0$ if and only if $x_{cj}^{(c)}=0$ and $x_{\bar{c}j}^{(\bar{c})}\neq 0$. In that case, $x_{\bar{c}j}^{(c)}$ attains its maximum value at

$$x_{\bar{c}j}^{(c)}=\frac{u_j^{(c)}\Lambda^{(c)}}{v_0^{(c)}+\sum_{k\in S_c} v_k^{(c)}+\sum_{i\in\mathcal{N}\setminus S_c}(u_i^{(c)}+w_i^{(c)})},$$

which concludes the proof.

Appendix E. Proof of Proposition 3

Let all shadow attractiveness values together with all parameters related to channel $\bar{c}$ be zero. Then, the assortment optimization problem under the MAM with shelf space constraint (10) takes the following form:

$$\max_{z_{cj}\in\{0,1\}^n} \frac{\sum_{j\in\mathcal{N}} r_{cj} v_j^{(c)} z_{cj}}{v_0^{(c)} + \sum_{k\in\mathcal{N}} v_k^{(c)} z_{ck}} \quad \text{s.t.} \sum_{j\in\mathcal{N}} a_{cj} z_{cj} \le L_c. \tag{E.1}$$

Problem (E.1) is essentially the shelf-space-constrained assortment optimization problem under the MNL, which has been shown by Désir et al. (2022) to be NP-hard. Indeed, suppose that $r_{cj} = 1$ for all $j \in \mathcal{N}$ and, following Désir et al. (2022), note that function $f(x) = \frac{x}{v_0^{(c)} + x}$ is increasing in x, meaning that the objective function of problem (E.1) can be replaced by $\sum_{j\in\mathcal{N}} v_j^{(c)} z_{cj}$. The resulting problem is equivalent to the knapsack problem. Therefore, the shelf-space constrained SBMILP – i.e. the SMBILP formulation (9) with additional constraint (10) – is NP-hard by reduction from the knapsack problem.

Appendix F. Proof of Proposition 4

First, let us show that $R^{(c)}(S_c \cup \{j\}, \mathcal{N})$ can be expressed as a convex combination of $R^{(c)}(S_c, \mathcal{N})$ and $(r_{cj} v_j^{(c)} - r_{\bar{c}j} u_j^{(c)})\Lambda^{(c)}/\tilde{v}_j^{(c)}$. Consider the following:

$$\alpha = \frac{\tilde{v}_j^{(c)}}{v_0^{(c)} + \sum_{k\in S_c\cup\{j\}} v_k^{(c)} + \sum_{i\in\mathcal{N}\setminus(S_c\cup\{j\})} (u_i^{(c)} + w_i^{(c)})}.$$

Note that $\alpha \in (0, 1)$. Then, it is easy to verify that

$$\alpha \frac{(r_{cj} v_j^{(c)} - r_{\bar{c}j} u_j^{(c)})\Lambda^{(c)}}{\tilde{v}_j^{(c)}} + (1-\alpha) R^{(c)}(S_c, \mathcal{N}) = R^{(c)}(S_c \cup \{j\}, \mathcal{N}).$$

This fact is sufficient for showing that the assortment given by (11) is optimal for channel c. Indeed, suppose that there is an optimal assortment S_c such that $q \in S_c$ and $p \notin S_c$ for some $p < q$. Then, since S_c is optimal, $R^{(c)}(S_c\backslash\{q\}, \mathcal{N}) \le R^{(c)}(S_c, \mathcal{N}) \le (r_{cq} v_q^{(c)} - r_{\bar{c}q} u_q^{(c)})\Lambda^{(c)}/\tilde{v}_q^{(c)}$. At the same time, since $p < q$ and hence $(r_{cq} v_q^{(c)} - r_{\bar{c}q} u_q^{(c)})\Lambda^{(c)}/\tilde{v}_q^{(c)} < (r_{cp} v_p^{(c)} - r_{\bar{c}p} u_p^{(c)})\Lambda^{(c)}/\tilde{v}_p^{(c)}$, it follows that $R^{(c)}(S_c, \mathcal{N}) < (r_{cp} v_p^{(c)} - r_{\bar{c}p} u_p^{(c)})\Lambda^{(c)}/\tilde{v}_p^{(c)}$. Therefore, $R^{(c)}(S_c \cup \{p\}, \mathcal{N})$ is greater than $R^{(c)}(S_c, \mathcal{N})$ as a convex combination of $R^{(c)}(S_c, \mathcal{N})$ and $(r_{cp} v_p^{(c)} - r_{\bar{c}p} u_p^{(c)})\Lambda^{(c)}/\tilde{v}_p^{(c)}$, which contradicts the assumption that S_c is optimal. It means that the optimal allocation has to be in descending order of $(r_{cj} v_j^{(c)} - r_{\bar{c}j} u_j^{(c)})\Lambda^{(c)}/\tilde{v}_j^{(c)}$. Finally, the fact that the optimal assortment has to be of the form (11) with index m specified in Proposition 4 follows from the same convex combination observation.

Appendix G. Proof of Proposition 5

Let $(S_c, S_{\bar{c}})$ be the optimal combination of assortments. Consider channel c and suppose that all products in that channel are sorted in descending order of the value of expression (5). Suppose that the proposition does not hold, i.e., for some $p < q$ we have that $q \in S_c$ and $p \notin S_c$. First, note that the total revenue generated by type-c customers is as follows:

$$R^{(c)}(S_c, S_{\bar{c}}) = \frac{\left(\sum_{k\in S_c} r_{ck} v_k^{(c)} + \sum_{i\in S_{\bar{c}}\setminus S_c} r_{\bar{c}i} u_i^{(c)}\right)\Lambda^{(c)}}{v_0^{(c)} + \sum_{k\in S_c} v_k^{(c)} + \sum_{i\in\mathcal{N}\setminus S_c} (u_i^{(c)} + w_i^{(c)})},$$

and the revenue generated by type-$\bar{c}$ customers purchasing product j in channel c is:

$$R_{cj}^{(\bar{c})}(S_c, S_{\bar{c}}) = \frac{r_{cj} u_j^{(\bar{c})} \Lambda^{(\bar{c})} \mathbb{1}_{j\in S_c\setminus S_{\bar{c}}}}{v_0^{(\bar{c})} + \sum_{k\in S_{\bar{c}}} v_k^{(\bar{c})} + \sum_{i\in\mathcal{N}\setminus S_{\bar{c}}} (u_i^{(\bar{c})} + w_i^{(\bar{c})})}.$$

Given that $q \in S_c$ and $p \notin S_c$, one can check that

$$\beta_1 \Lambda^{(c)} F_c(p, S_c, S_{\bar{c}}) + (1-\beta_1) R^{(c)}(S_c, S_{\bar{c}}) = R^{(c)}(S_c \cup \{p\}, S_{\bar{c}}) + R_{cp}^{(\bar{c})}(S_c \cup \{p\}, S_{\bar{c}}) \tag{G.1}$$

and

$$\beta_2 \Lambda^{(c)} F_c(q, S_c, S_{\bar{c}}) + (1-\beta_2) R^{(c)}(S_c\backslash\{q\}, S_{\bar{c}}) = R^{(c)}(S_c, S_{\bar{c}}) + R_{cq}^{(\bar{c})}(S_c, S_{\bar{c}}), \tag{G.2}$$

where

$$\beta_1 = \frac{\tilde{v}_p^{(c)}}{v_0^{(c)} + \sum_{k\in S_c\cup\{p\}} v_k^{(c)} + \sum_{i\in\mathcal{N}\setminus(S_c\cup\{p\})} (u_i^{(c)} + w_i^{(c)})}$$

and

$$\beta_2 = \frac{\tilde{v}_q^{(c)}}{v_0^{(c)} + \sum_{k\in S_c} v_k^{(c)} + \sum_{i\in\mathcal{N}\setminus S_c} (u_i^{(c)} + w_i^{(c)})}.$$

Since $(S_c, S_{\bar{c}})$ is assumed to be the optimal combination of assortments, it holds that $R^{(c)}(S_c\backslash\{q\}, S_{\bar{c}}) < R^{(c)}(S_c, S_{\bar{c}}) + R_{cq}^{(\bar{c})}(S_c, S_{\bar{c}})$. Therefore, from relation (G.2) it follows that $\Lambda^{(c)} F_c(q, S_c, S_{\bar{c}}) > R^{(c)}(S_c, S_{\bar{c}}) + R_{cq}^{(\bar{c})}(S_c, S_{\bar{c}}) \ge R^{(c)}(S_c, S_{\bar{c}})$. Since $p < q$ and products are sorted in descending order of the value of expression (5), we obtain that $\Lambda^{(c)} F_c(p, S_c, S_{\bar{c}}) > R^{(c)}(S_c, S_{\bar{c}})$. Thus, from relation (G.1) it follows that $R^{(c)}(S_c \cup \{p\}, S_{\bar{c}}) + R_{cp}^{(\bar{c})}(S_c \cup \{p\}, S_{\bar{c}}) > R^{(c)}(S_c, S_{\bar{c}})$. This means that the combination of assortments $(S_c \cup \{p\}, S_{\bar{c}})$ is more profitable than $(S_c, S_{\bar{c}})$, which contradicts the initial assumption and thereby concludes the proof.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ejor.2022.08.002.

References

Bayram, A., & Cesaret, B. (2021). Order fulfillment policies for ship-from-store implementation in omni-channel retailing. *European Journal of Operational Research, 294*(3), 987–1002.

Ben-Akiva, M. E., & Lerman, S. R. (1985). *Discrete choice analysis: Theory and application to travel demand*. MIT Press Cambridge (Mass.).

Berbeglia, G., Garassino, A., & Vulcano, G. (2021). A comparative empirical study of discrete choice models in retail operations. *Management Science*.

Berbeglia, G., & Joret, G. (2020). Assortment optimisation under a general discrete choice model: A tight analysis of revenue-ordered assortments. *Algorithmica, 82*(4), 681–720.

Bhatnagar, A., & Syam, S. S. (2014). Allocating a hybrid retailer's assortment across retail stores: Bricks-and-mortar vs online. *Journal of Business Research, 67*(6), 1293–1302.

Blanchet, J., Gallego, G., & Goyal, V. (2016). A markov chain approximation to choice modeling. *Operations Research, 64*(4), 886–905.

Block, H., & Marschak, J. (1959). Random Orderings and Stochastic Theories of Response. *Cowles Foundation Discussion Papers*. Cowles Foundation for Research in Economics, Yale.

Brightpearl (2017). The State of Omnichannel Retail – Survey of Leading Retailers. White paper, Accessed on 01/05/2022, https://www.brightpearl.com/ecommerce-guides/an-introduction-to-omnichannel-retail.

Cao, J., So, K. C., & Yin, S. (2016). Impact of an "online-to-store" channel on demand allocation, pricing and profitability. *European Journal of Operational Research, 248*(1), 234–245.

Désir, A., Goyal, V., & Zhang, J. (2022). Technical note-capacitated assortment optimization: hardness and approximation. *Operations Research, 70*(2), 893–904.

Domencich, T., & McFadden, D. L. (1975). *Urban travel demand: A Behavioral analysis*. North-Holland Pub. Co..

Dzyabura, D., & Jagabathula, S. (2018). Offline assortment optimization in the presence of an online channel. *Management Science, 64*(6), 2767–2786.

Gallego, G., Ratliff, R., & Shebalov, S. (2014). A general attraction model and sales-based linear program for network revenue management under customer choice. *Operations Research, 63*(1), 212–232.

Gallino, S., & Moreno, A. (2014). Integration of online and offline channels in retail: the impact of sharing reliable inventory availability information. *Management Science, 60*(6), 1434–1451.

Gao, F., & Su, X. (2016). Omnichannel retail operations with buy-online-and-pick-up-in-store. *Management Science, 63*(8), 2478–2492.

Govindarajan, A., Sinha, A., & Uichanco, J. (2018). Joint inventory and fulfillment decisions for omnichannel retail networks. Ross School of Business Paper No. 1341. https://onlinelibrary.wiley.com/doi/abs/10.1002/nav.21969. Accessed on 01/05/2022.

Harsha, P., Subramanian, S., & Ettl, M. (2019a). A practical price optimization approach for omnichannel retailing. *INFORMS Journal on Optimization, 1*(3), 241–264.

Harsha, P., Subramanian, S., & Uichanco, J. (2019b). Dynamic pricing of omnichannel inventories. *Manufacturing & Service Operations Management, 21*(1), 47–65.

He, Y., Xu, Q., & Wu, P. (2020). Omnichannel retail operations with refurbished consumer returns. *International Journal of Production Research, 58*(1), 271–290. https://www.tandfonline.com/doi/full/10.1080/00207543.2019.1629672.

Hense, J., & Hübner, A. (2021). Assortment optimization in omni-channel retailing. *European Journal of Operational Research*. (In Press)

Lo, V., & Topaloglu, H. (2019). Assortment optimization under the multinomial logit model with product synergies. *Operations Research Letters, 47*(6), 546–552.

Lo, V., & Topaloglu, H. (2022). Omnichannel assortment optimization under the multinomial logit model with a features tree. *Manufacturing & Service Operations Management, 24*(2), 1220–1240.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. Wiley New York.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behaviour. In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 105–142). NY, USA: Academic Press New York.

Petro, G. (2020). Walmart Challenges Amazon With 'Less Is More' Strategy. Forbes. https://www.forbes.com/sites/gregpetro/2020/02/07/walmart-challenges-amazon-with-less-is-more-strategy/?sh=74ddc2777ff1. Accessed on 01/05/2022.

iVend Retail (2019). Global Shopper Trends Report. White paper. https://ivend.com/survey-report-2019/. Accessed on 01/05/2022.

Rusmevichientong, P., Shmoys, D., Tong, C., & Topaloglu, H. (2014). Assortment optimization under the multinomial logit model with random choice parameters. *Production and Operations Management, 23*(11), 2023–2039.

Schneider, F., & Klabjan, D. (2013). Inventory control in multi-channel retail. *European Journal of Operational Research, 227*(1), 101–111.

Seifert, R. W., Thonemann, U. W., & Sieke, M. A. (2006). Integrating direct and indirect sales channels under decentralized decision-making. *International Journal of Production Economics, 103*(1), 209–229.

Sopadjieva, E., Dholakia, U. M., & Benjamin, B. (2017). A Study of 46,000 Shoppers Shows That Omnichannel Retailing Works. Harvard Business Review. https://hbr.org/2017/01/a-study-of-46000-shoppers-shows-that-omnichannel-retailing-works. Accessed on 01/05/2022.

Train, K. (2002). *Discrete choice methods with simulation*. Cambridge University Press.

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

Wächter, A., & Biegler, L. (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming, 106*, 25–57.