European Journal of Operational Research 294 (2021) 1120–1138

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Production, Manufacturing, Transportation and Logistics

# Capacitated strategic assortment planning under explicit demand substitution

Nagihan Çömez-Dolgan[a], Nilgun Fescioglu-Unver[b,*], Ecem Cephe[c], Alper Şen[d]

[a] Faculty of Management, Ozyegin University, Cekmekoy, Istanbul, Turkey
[b] Department of Industrial Engineering, TOBB University of Economics and Technology, Ankara, Turkey
[c] Turkish Aerospace Industries, Inc. (TAI), Ankara, Turkey
[d] Department of Industrial Engineering, Bilkent University, Ankara, Turkey

ARTICLE INFO

*Article history:*
Received 4 April 2020
Accepted 5 February 2021
Available online 11 February 2021

*Keywords:*
Production
Production variety
Substitution
Assortment capacity
Exogenous demand

ABSTRACT

Buyers have easier access to a variety of products with the rise of multi-channel distribution strategies and the increase in new product introductions. On the other hand, firms experience greater pressure in offering the correct product variety given that the manufacturing infrastructure often imposes physical and financial constraints in attaining variety. This study examines a firm's optimal assortment planning problem under an exogenous demand model, where each customer has a predetermined preference for each product from a potential set. Proportional demand substitutions are allowed from out-of-assortment products to those available. We show that the problem is NP-complete. We also show that an optimal assortment is composed of some number of the highest margin products, if one product having a higher margin than another implies that the former product has a lower demand rate than the latter. The firm's assortment capacity is fully utilized at the optimum if the customers' substitution ratio does not exceed a particular threshold. We also introduce several approximate assortment policies that can be easily implemented, and test these policies through extensive numerical analyses. The results reveal that some of the policies can provide less than a 1% profit gap with an optimal solution for a 20-product set. The policy's performance highly depends on the firm's assortment capacity-to-product set size ratio. Moreover, we provide performance bounds for two of these well-performing approximate policies.

© 2021 Elsevier B.V. All rights reserved.

## 1. Introduction

When the Ford Motor Company revealed a newly revised Ford Fusion sedan in March 2018, they also announced that they standardized many features of the vehicle—for example, Ford's Co-Pilot360 driver-assist technology—while leaving only a few additional options (Ford, 2018). The company states that this strategy substantially decreased the number of orderable configurations for the Ford Fusion, from approximately 2000 to 36. By decreasing these configurations, they decreased their manufacturing complexity with the aim to reduce costs.

In parallel to the reduced number of configurations per model, the number of car models produced in each plant also significantly decreased with a similar incentive of obtaining a leaner production plant operating at lower costs. Choudhary, Hasija, and Netessine (2018) report that between 1998 and 2006, the number of models produced in assembly plants by three major U.S. automobile manufacturers—Chrysler, Ford, and General Motors—is decreasing. For example, while Ford manufactured up to nine models per plant in 1998, the maximum number of models per plant decreased to three in 2006. The authors note that each plant should reach a particular level of variety to balance the demand satisfaction benefits and excessive set-up costs.

One critical decision for the automotive industry involves assigning models and their subsequent configurations to production plants. These assignments are determined by tooling and capacity investment decisions that must be made between one to three years before production begins. Further, this may require substantial investment, such as funds for a new assembly line, additional tooling, or employee training (Jordan & Graves, 1995). The problem of assigning products to automobile production plants is a strategic assortment-planning problem that does not often consider inventory.

Generally, the firm's limited assortment, or set of products offered at any time, should be carefully set consistent with the firm's

* Corresponding author.
*E-mail addresses:* nagihan.comez@ozyegin.edu.tr (N. Çömez-Dolgan), nfunver@etu.edu.tr (N. Fescioglu-Unver), ecem.cephe@gmail.com (E. Cephe), alpersen@bilkent.edu.tr (A. Şen).

https://doi.org/10.1016/j.ejor.2021.02.026
0377-2217/© 2021 Elsevier B.V. All rights reserved.

strategy. Assortment planning is the process of deciding (i) the number of categories, called the *breadth*; (ii) the number of products in each category, called the *depth*; and (iii) the corresponding inventory levels for each product to be offered at any time. An assortment of a certain size involves its relevant operational complexities and costs as well as customer sales potential. Further, assortment planning aims to offer an optimal variety of products to customers to maximize the total profit from sales relative to the given costs and limitations of this variety.

Briesch, Chintagunta, and Fox (2009) report that customers' brand choice decisions can be more sensitive to assortments than to prices. Further, firms must attempt to implement periodical assortment planning to consider customers' changing preferences over time, seasons, and the launch of new products on the market (Kök, Fisher, & Vaidyanathan, 2015). The assortment-planning decision is complex due to several trade-offs between having a rich versus limited assortment. From the customer's perspective, on the one hand, a rich and variable assortment draws higher customer traffic (Timonina-Farkas, Katsifou, & Seifert, 2020). For example, expanding an assortment in a retail setting can decrease consumers' search even with unprofitable products, consequently increasing profit (Cachon, Terwiesch, & Xu, 2005). On the other hand, a narrow assortment can make customers' decision easier thus increasing the probability of purchase (de Vries-van Ketel, 2006; Mantrala et al., 2009). Moreover, Boatwright and Nunes (2001) reveal that reducing the assortment by up to 54% increases average sales by 11%.

From an operational perspective, firms may also experience space and budget constraints regarding the variety offered. Each product requires substantial investment, such as funds for new assembly lines, additional tooling, or employee training. For example, Toyota Europe announced that they will invest 300 million euros in its plant in France to build a platform to enable the production of new Toyota models (Toyota-Europe, 2018). After a product is included in the assortment, operational costs per product are incurred because of material handling and warehousing, as well as merchandize presentation (Smith & Agrawal, 2000), recordkeeping, and reordering. While it might be more challenging to compute the fixed assortment costs per product, a constrained assortment size often inevitably arises in practice. Subsequently, manufacturers are limited in the number of assembly lines they can place into service, and each line can only produce a few models. Each offered product involves handling, replenishment, and inventory costs, as a certain quantity of inventory should be maintained. All products in an assortment are intertwined through the total available budget. If all products have symmetrical or similar space and/or financial needs, then assortment constraints can be reduced to a cardinality constraint.

Given that the assortment size will be limited because of both demand and cost perspectives, firms should also consider the behaviors of consumers faced with a limited variety. When customers visit a firm, they typically demand a specific product, and its unavailability may lead them to consider either leaving the firm without purchasing, or switching to another product. The act of switching to an alternative product when the favored product is unavailable is known as *substitution* (Shin, Park, Lee, & Benton, 2015). This can occur when either a shortage exists in the product's inventory, called *stockout-based substitution*, or the product is not offered within the assortment, called *assortment-based substitution*. Corsten and Gruen (2004) report that almost half of customers may tend to switch to different products when their favorite is unavailable. Customers' substitution-behavior effects can also further complicate firms' assortment choices. A survey of U.S. vehicle dealers indicated that 15–30% of customers switched from the car they originally sought to one available on the lot (Stalk, Stephenson, & King, 1997). Further, Mahajan and van Ryzin (2001b) indicate that firms can stock relatively more quantities of popular products and relatively fewer unpopular ones under substitution than in the event in which substitutions are not allowed; thus, inventory will be more evenly spread across variants.

This paper examines the optimal assortment of a manufacturing firm whose assortment size constrained by manufacturing infrastructure requirements (Hart & Rafiq, 2006) and that explicitly considers customers' substitution behavior. It is studied as a strategic problem, thus tactical level inventory and/or production capacity decisions are not incorporated at assortment planning stage. The existence of product substitution complicates the assortment planning problem. We actually show that the problem is NP-complete, which invalidates the use of simple "greedy" algorithms for optimal solution. So, we aim to illustrate the properties of optimal assortments to understand the effects of substitution and capacity constraints on the choice of assortment.

We demonstrate that the optimal assortment contains most preferred products, and that its capacity is always fully utilized if products have different customer preference rates but equal profit margins. When products also vary in their profit margins, the optimal assortment cannot be obtained with a "greedy" algorithm. If all products can be sorted monotonically in increasing order of their profit margins and decreasing order of their demand probabilities, the optimal assortment is composed of some number of most dominant (profitable) products, which can be lower than the assortment capacity. It is shown that by keeping a low profit margin, but high demand product out of the assortment, its demand can be directed to higher margin substitutes.

If all products do not posit monotonic ordering of profit margins and demand probabilities, some number of highly dominant products can omitted from the assortment under a high substitution ratio, which increases the probability of retaining high-margin substitutes. It is proven that when the substitution ratio is smaller, it is more likely that the assortment capacity is fully utilized and an optimal assortment will be composed of the most profitable products. We introduce an upper limit on the substitution ratio below which the assortment capacity of the firm is always fully utilized at optimality. Alternatively, when the assortment capacity is high, capacity utilization may decrease and the optimal assortment may include some less profitable products while excluding some that are more profitable to direct customers to high-margin products in the assortment. We prove that the firm may take the risk of a shallower assortment and expect customers to substitute their demands with those that are high-margin under a high substitution ratio.

Next, our work benefits from obtained optimality properties to introduce seven heuristic assortment-planning algorithms, the complexities of which vary according to their optimality properties and differ in a range from simple sorting to complete profit computation. Our numerical analyses demonstrate that a firm should consider the product set's size, the assortment capacity, and computational capability when deciding not only whether to use a heuristic policy, but also which to utilize.

We contribute to strategic assortment-planning literature by analyzing a generalized exogenous demand model with product-specific demand rates and profit margins under assortment capacity. We analytically show how substitution ratio affects the product choice as well as the optimal assortment capacity utilization. Moreover, the approximate algorithms we introduce are efficient in computation and effective in obtaining nearly optimal assortments proved by their performance bounds. The assortment optimization problem can also be formulated as a mixed-integer model, but the resulting problem can still be computationally very challenging to solve (Chung, Ahn, & Jasin, 2019) and would not provide the insights on optimal assortments that we obtain in our current study.

We consider manufacturing firms' strategic assortment decisions, thus excluding inventory decisions in assortment planning. The proposed methodologies are also applicable to other assortment problems without significant inventory concerns during the assortment planning stage, such as when either the inventory management is relatively easier as with slow-moving goods or all inventory is not carried on shelves, but in a depot with limited shelf facings (Kök et al., 2015). For example, Fisher and Vaidyanathan (2014) study the assortment-planning problem for slow-moving products with no inventory concerns motivated by retailers who carry a fixed, often small inventory for each SKU. Feldman and Topaloglu (2017) provide a detailed list of similar assortment problems independent of inventory decisions.

While we primarily express our motivation using automobile manufacturers, the problem setting and results are largely generalizable to other manufacturing industries where the inclusion of each product in a production assortment requires substantial investment. For example, Akçay and Tan (2008) state that small to medium-sized enterprises (SMEs), and textile manufacturers in particular, collect orders from large buyers and procure accordingly (i.e., make-to-order firms). However, each SME specializes in the production of certain types of fabric and subsequently constructs their assortments accordingly. Thus, each SME commits to producing a limited number of varieties due to the significant assortment costs per product. Tan and Akçay (2014) provide the furniture industry as an example, in which a manufacturer has a production catalog with a limited number of models, and each model may require a specific expertise and tooling set-up. Thus, our model and insights are also valid for manufacturers operating in these industries.

The remainder of this paper is organized as follows. In Section 2, we review related studies in literature. In Section 3, we present the problem and show that it is NP-Complete. We reveal the analytical properties of an optimal assortment in Section 4. We provide numerical analysis results to further explain and illustrate the optimal assortment's properties in Section 5. Next, we introduce several approximate assortment policies that are easy to use and compare their performances with system parameters in Section 6. Finally, Section 7 presents our final remarks by summarizing our findings and obtained insights from the perspective of subsequent research.

## 2. Literature review

Assortment planning has two main inputs. One of them is customer-related, as assortment affects customer traffic and sales. The other is operations-related, as the assortment size determines many cost terms, such as handling, shelving, and replenishment, as discussed in Section 1. Thus, both operations management and marketing researchers work on assortment planning. This section presents a brief survey of related literature. Extensive reviews of assortment planning literature are provided by Pentico (2008), Mantrala et al. (2009), Misra (2010), Chernev (2012), and Kök et al. (2015). It is also worth to note that studies incorporating strategic level supply chain concerns into assortment planning problems are still very scarce, where one of the contributions of this study stands in (Umpfenbach, Dalkiran, Chinnam, & Murat, 2018a).

Assortment planning literature dates back to the 1950s, when Sadowski (1959) was likely the first to label the "assortment problem" (Pentico, 2008). Past studies differ in terms of the model characteristics considered, such as the consumer demand model, demand substitution pattern, the inclusion of inventory-level decisions, and the consideration of assortment capacity. Mantrala et al. (2009) differentiate assortment studies according to resulting trade-offs in optimization by considering these characteristics separate from the consumer, retailer, and environmental perspectives. Table 1 lists the most related past work to the current study by categorizing according to demand model, whether inventory decision is included, whether assortment capacity is used, substitution model, and solution methodology utilized. Besides, main contribution of each study to the literature is noted in the table.

Bernstein, Kök, and Xie (2015) categorize earlier work according to the customer choice model used—such as the multinomial logit (MNL) and exogenous demand models—which greatly affect the other problem characteristics that can be modeled, such as the substitution type. Chung et al. (2019) state that estimating consumer behavior precisely is one of the most challenging issues in assortment problems and more realistic and detailed models sacrifice from the tractability of the problem solution. The MNL model is one of the most commonly used customer behavior models in assortment problems given its robust practical estimations, as noted by van Ryzin and Mahajan (1999), Mahajan and van Ryzin (2001a), and Cachon and Kök (2007). This utility-based model easily incorporates pricing into a customer demand model. Further, the MNL model assumes that each customer visiting a store associates a utility with each product that can be decomposed into two parts: deterministic and random components. This assumes that customers deciding from a discrete set of products are utility-maximizing individuals. However, two primary criticisms exist regarding the MNL model. First, the independence from irrelevant alternatives (IIA) property states that a customer's ratio of choice probabilities for two products is independent of other available choices in the overall set. Therefore, omitting a product from the model will change the parameter estimates of all the remaining items at the same relative rate, which cannot always be correct. Nonetheless, this is commonly applied in the market research, economics, and logistics fields, among others. A second criticism of the MNL model involves its restricted modeling of substitutions.

In an exogenous demand model, the demand for each product is specified ex-ante for all possible products, and thus, does not depend on a selected assortment (Smith & Agrawal, 2000). Moreover, customers' substitution behaviors are predefined independent of the choice of assortment set. When a customer's most favored product is unavailable—either due to stock-out conditions or it has been omitted from the assortment—this demand is substituted by the customer's second-favorite product, which is not necessarily within the available assortment, with a predetermined probability. The number of substitutions can be fixed at a certain number, or it can continue until an available product can be reached. Exogenous demand models are heavily used in the literature, especially in studies that involve real life applications. Important examples include Kök and Fisher (2007), Fisher and Vaidyanathan (2014) and Bernales, Guan, Natarajan, Gimenez, and Tajes (2017). The exogenous demand model has more degrees of freedom and provides more flexibility in modeling substitution behavior (Kök et al., 2015). A recent work by Chung et al. (2019) show that using exogenous demand models in assortment planning do not lead to significant shortfall in revenue when compared to using the exact choice models such as mixtures of MNLs and are shown to approximate any random utility model with a desired level of accuracy (McFadden & Train, 2000). The exogenous demand model's primary shortcoming is its lack of an underlying consumer behavior model defining demand rates, which consequently requires significant data collected for an application.

Gallego, Ratliff, and Shebalov (2014) introduced the generalized attraction model (GAM) as a customer demand model, which is a generalized model that may be reduced to the MNL or exogenous demand models in special cases. The MNL model ignores the consumer search option when the first choice is unavailable, and this yields to the overestimation of recaptured demand. In contrast, the exogenous demand model ignores the switching option from direct demand, subsequently yielding an underestimation of overall

**Table 1**
Overview of the most related past studies.

| Authors | Demand model[(1)] | Capacity | Substitution[(2)] | Solution method[(3)] | Main contribution |
|---|---|---|---|---|---|
| **Our work** | Exogenous | ✓ | Assortment | Optimal & Approximate | Explicit relationship between substitution and optimal assortment |
| Aouad et al. (2018) | Choice | | Assortment | Approximate | Approximability bounds under a general choice model |
| Baloch and Gzara (2020) | Empirical | ✓ | Both | MIO | Assortment, inventory and substitution policies for kiosks with real data |
| Bernstein et al. (2015) | MNL | | Stock-out | Optimal & Approximate | Customized assortments based on available inventory |
| Besbes and Saure (2016) | MNL | ✓ | Assortment | Optimal | Joint assortment and price competitions under the MNL model |
| Blanchet et al. (2016) | Choice | | Assortment | Approximate | Computationally tractable approach to choice model |
| Cachon et al. (2005) | MNL | | Assortment | Optimal | Accounting consumer search in assortment planning |
| Cachon and Kök (2007) | MNL | | Assortment | Optimal& Approximate | Competing retailers for basket shopping consumers |
| Chung et al. (2019) | Exogenous | ✓ | Assortment | MIO | Approximation to any random-utility choice model |
| Dèsir et al. (2020) | Choice | ✓ | Assortment | Approximate | Approximation for constrained problems under choice model |
| Fadıloglu et al. (2010) | Exogenous | | Assortment | Optimal | Optimization model with minimal data requirement |
| Feldman and Topaloglu (2015a) | MNL | ✓ | Assortment | Optimal & Approximate | Tractable solutions under constrained nested logit model |
| Feldman and Topaloglu (2015b) | Mix MNL | | Assortment | Approximate | Upper bound on the optimal solution for mixed MNL model |
| Feldman and Topaloglu (2017) | Choice | | Assortment | Optimal & LP | LP to obtain the optimal solution with Markov chain choice model |
| Feldman et al. (2019) | Choice | | Assortment | Approximate | Customer choice model with a limit on customer substitutions |
| Fisher and Vaidyanathan (2014) | Exogenous | ✓ | Assortment | Approximate | Real example for parameter estimation and heuristic application |
| Gallego and Topaloglu (2014) | Nested | ✓ | Assortment | LP & Approximate | Cardinality and space constraints with the nested logit model |
| Golrezaei et al. (2014) | Choice | ✓ | - | Approximate | Algorithms for real-time personalized assortments |
| Goyal et al. (2016) | Choice | ✓ | Both | Approximate | Algorithm with provable performance under dynamic substitution |
| Honhon et al. (2012) | Choice | | Assortment | Optimal | Practically motivated special cases of assortment |
| Jagabathula (2014) | Choice | ✓ | Assortment | Approximate | Local search heuristic |
| Jagabathula and Rusmevichientong (2017) | Choice | | Assortment | Optimal& Approximate | Nonparametric approach for joint assortment and price optimization |
| Kök and Fisher (2007) | Exogenous | ✓ | Both | Iterative Approximate | Practical assortment planning approach |
| Mahajan and van Ryzin (2001b) | Utility | | Both | Optimal& Gradient | Efficient computational approach using gradients |
| Nip et al. (2017) | Choice | | Assortment | Optimal & MIO | Seller can recommend products for substitution |
| Rusmevichientong et al. (2009) | Nested | ✓ | Assortment | Approximate | Approximation for capacitated nested logit choice model |
| Rusmevichientong et al. (2010) | MNL | ✓ | Assortment | Optimal | Capacitated problem both in static and dynamic settings |
| Smith and Agrawal (2000) | Exogenous | ✓ | Both | Lagrange relaxation | Demand substitution and customer service level for all items |
| Şen et al. (2018) | Mix MNL | ✓ | Assortment | MIO | Conic quadratic MIO to solve large size capacitated problems |
| van Ryzin and Mahajan (1999) | MNL | | Assortment | Optimal | Theoretical insights on joint assortment and inventory problem |
| Wang (2012) | MNL | ✓ | Assortment | Optimal | Capacitated assortment and price optimization under the MNL model |
| Wang (2013) | GAM | ✓ | Assortment | Optimal & Approximate | Optimal assortment under GAM |
| Yücel et al. (2009) | Exogenous | ✓ | Both | MIO | Practical and flexible model |

**Acronyms:** (1) Choice: Consumer choice model, Nested: Nested logit choice model (2) Assortment: Assortment-based, Stock-out: Stock-out-based, Both: Both assortment and stock-out based (2) MIO: Mixed Integer Optimization, LP: Linear Program, Optimal: Optimality Characterization

demand. In addition to the direct attraction values, this demand model also considers switching attraction values.

Some recent work has used a ranking-based consumer choice model to represent consumer preferences such that each customer has a ranking of the potential product. Golrezaei, Nazerzadeh, and Rusmevichientong (2014) propose approximate policies that use real-time inventory information to offer personalized assortments for arriving customers. Jagabathula and Rusmevichientong (2017) incorporate price thresholds for each customer in addition to general preference lists, and propose an approximation algorithm to jointly determine an assortment set and product prices. Honhon, Jonnalagedda, and Pan (2012) use a ranking-based consumer choice model. For the special case of one-way substitutions, they obtain optimality properties to offer an efficient optimization algorithm. Aouad, Farias, Levi, and Segev (2018) prove the complexity of an assortment optimization when customers' choices are modeled through arbitrary ranking-based preference lists, but also demonstrate that the widely studied revenue-ordered assortments achieve the best possible approximation performance. Feldman, Paul, and Topaloglu (2019) note that even when customers' preference lists are incredibly limited, the assortment problem is NP-hard for which they develop an approximation algorithm.

To model customers' multiple substitution attempts, Blanchet, Gallego, and Goyal (2016) propose an iterative Markov search model where the substitution probability is indicated as a transition probability in a Markov chain. In this model, a customer continues searching until she finds a preferred product or decides to leave the system with no-purchase. The proposed model approximates the true choice model well and is a good approximation of existing choice models, including logit and exogenous demand models. They introduce a polynomial-time algorithm to solve the corresponding assortment problem. Feldman and Topaloglu (2017) introduce a linear-programming-based solution for an assortment problem with a similar Markov chain choice model. Nip, Wang, and Wang (2017) study a Markov chain choice model with single transition, where the seller controls the set of products to recommend for this transition. They show that the problem is generally NP-Hard, so they provide polynomial time algorithms for special cases, such as each product can only transit to one other product. Dèsir, Goyal, Segev, and Chun (2020) study an constrained assortment problem under Markov chain choice model and introduce an approximation algorithm with a provable worst-case gap.

The current studies that use general demand choice or Markov chain choice either derive efficient algorithms under specific parametric structures, or propose approximate solutions with certain performance bounds (Jagabathula, 2014). On the other hand, we aim to obtain the structural properties of optimal assortments to understand their dynamics and also use these results to introduce easily implementable assortment-planning policies. Moreover, we explicitly study the substitution effect on assortment planning, which to our knowledge would be extremely challenging with a more generalized demand choice model (Chung et al., 2019).

This study uses an exogenous demand model to explicitly capture customer demand substitutions. We also consider assortment capacities, which are relatively limited compared to non-capacity models. Assortment planning studies that do not consider assortment capacity solve for the trade-off between the extra revenue brought by including each product in the assortment and the product's additional operational cost. In this category, van Ryzin and Mahajan's (1999) seminal paper considers all product variants as having an identical retail price and unit cost, and assortment-based substitutions are allowed. They reveal that the optimal assortment consists of a certain number of the most popular products, or those with the highest demand rates.

Literature less often studies assortment planning problems that consider constraints on the assortment capacity, such as the number of products, space consumption, and substitution ratio. Smith and Agrawal (2000) use an exogenous demand model to observe the effects of substitution in deciding inventory levels subject to resource constraints; they consider either unequal product profit margins or demand rates, but not both. Through illustrative examples, they reveal that substitution reduces the optimal number of items to stock under fixed costs. Even if fixed costs are neglected, it is not always optimal to offer all items in the assortment when items have different profit margins.

Kök and Fisher (2007) study an assortment-planning problem for which they first present a procedure to estimate the substitution and demand parameters using sales summary data, then develop an iterative heuristic to discover the assortment's structural properties, such as deciding the priority order of the products in the assortment. For each product subcategory, each product's number of facings is determined subject to shelf-space constraints. Fadıloglu, Karasan, and Pınar (2010) provide an optimization model to eliminate some SKUs from the shelf to prevent product pollution. The demand substitution from out-of-assortment products is defined by exogenously establishing a substitution ratio, as in our model, as well as the relative weight of sales for each SKU. Rusmevichientong, Shen, and Shmoys (2010) study MNL demand for an assortment problem with a cardinality constraint; they solve for the optimal assortment using a static model in which customer preferences are known. Goyal, Levi, and Segev (2016) study an assortment problem under dynamic substitution, stochastic demand, and a total inventory capacity. After illustrating that the problem is NP-hard, they provide an efficient algorithm with near-optimal performance guarantees.

Wang (2012) studies a joint-assortment and price-optimization problem under a cardinality constraint and MNL demand. This study proves that an optimal assortment's size equals the capacity when prices are jointly determined with the assortment. When prices are set ex-ante, some low-margin products can be set out of the assortment to direct customers to those high-margin products. Wang (2013) studies a cardinality-constrained assortment problem with a GAM and fixed prices. The study provides an efficient algorithm to discover the optimal assortment in polynomial time for a static problem, and to establish a time threshold structure for a dynamic problem. Baloch and Gzara (2020) investigate the assortment and stocking decisions of medications at pharmacy kiosks, which are limited by stocking capacity. They develop several different mixed integer optimization models that use sales data. Besbes and Saure (2016) analyze retailers' joint assortment and price competitions under the MNL model. They reveal that the optimal assortment has a nested structure to allow the products to be simply ranked by quality values and costs; further, the optimal assortment includes the top products in the ranking to fully utilize the capacity.

Rusmevichientong, Shen, and Shmoys (2009) address an assortment problem under the nested logit choice model, formulated as an integer-programming problem involving a sum of ratios that reduces to a knapsack problem. The authors provide a polynomial time approximation for the optimization problem with a budget constraint. Gallego and Topaloglu (2014) also use the nested logit model with cardinality and space constraints for each nest. They indicate that the optimal assortment under cardinality constraints can be obtained through a linear program, but under space constraints the problem is NP-hard. Feldman and Topaloglu (2015a) also study a similar problem, but they impose a common capacity constraint on all nests. Another variant of the MNL model is the mixed MNL model, in which customers are grouped into multiple segments, each of which has a separate MNL model. Feldman and Topaloglu (2015b) and Şen, Atamturk, and Kaminsky (2018) consider constrained assortment optimization under this model. Şen et al. (2018) introduce a conic quadratic mixed-integer formulation to optimally solve relatively large size problems. Feldman and Topaloglu (2015a) present an efficient algorithm to compute the optimal assortment for the cardinality constrained case and approximation under the general capacity constraint.

## 3. Assortment planning model

We model the assortment-planning problem of a manufacturing firm that needs to determine its limited product portfolio due to significant investment limitations, which is ultimately treated as a cardinality constraint. Each product in the possible set has a predetermined customer preference and a profit margin. The goal is to select the right products to be offered in the assortment to maximize the firm's total profit from sales relative to the capacity limitation.

The set of all potential products is $N$. Without loss of generality, the total demand of the firm is reduced to a unit. The probability of demand coming for a specific product in $N$ is known as customer preference and is independent of the offered assortment. Literature notes this demand pattern as the exogenous demand model (Smith & Agrawal, 2000; Yücel, Karaesmen, Salman, & Türkay, 2009). Each product $i$ has a certain probability of being the first choice of a visiting customer, denoted as $\alpha_i$, such that $0 < \alpha_i \le 1$ and $\sum_{i \in N} \alpha_i = 1$.

If an arriving customer's favorite product $i$ is in the firm's assortment, he or she will pay $r_i$ to purchase it. If the firm does not offer product $i$, then the customer may substitute another product with probability $\theta \le 1$. The substitution from an unavailable product to another product, regardless of whether it is in the assortment, can occur in different ways: randomly, or to any other product with the same probability; adjacent, or to neighboring products according to some attributes; or proportionally according to the demand rates of other products in the potential set $N$ (Kök et al., 2015). Here, we define proportional or "market share-based" substitutions (Karabati, Tan, & Ozturk, 2009; Smith & Agrawal, 2000). If the customer decides to substitute his first choice with a second one, the probability that the customer substitutes product $i$ with product $j$ is

$$\delta_{ij} = \left(\frac{\alpha_j}{1-\alpha_i}\right)$$

where $\delta_{xx}$= 0. If the second-favorite product $j$ is in the assortment, then the customer's demand is satisfied; otherwise, the customer leaves the system without any purchase.

Kök and Fisher (2007) provide a detailed discussion on how to estimate initial customer demand for each product $\alpha_i$ and the assortment-based substitution ratio $\theta$ in an exogenous demand model using sales data. Briefly, at a store that operates with almost hundred percent service level and offers an assortment that is smaller than the potential set, demand rates are estimated from the past sales, which may include substitutions as well. The demand rates without substitutions can be estimated from sales data of a similar store that carries a full assortment. If the demand rates at the store with less than full assortment are higher than those in full assortment store, it can be concluded that the substitution ratio is positive and substitution ratio can be estimated by regression model.

As this paper does not consider inventory decisions, stock-out-based substitutions—also called dynamic substitutions—are outside of this paper's scope. Consequently, the term "substitution" refers to assortment-based substitution. Other assortment studies in literature have also considered this form of demand substitution, also known as static substitution, such as works by Besbes and Saure (2016), Kök and Fisher (2007), Yücel et al. (2009), and Umpfenbach, Dalkiran, Chinnam, and Murat (2018b). We implicitly assume that any demand for the products existing in the assortment can be satisfied. We primarily allow for at most one substitution attempt due to analytical tractability, which does not significantly contribute to the model because it is possible to approximate multiple substitution behaviors with a single substitution by increasing the substitution ratio (Karabati et al., 2009; Kök, 2003; Kök & Fisher, 2007; Smith & Agrawal, 2000).

The firm's objective is to maximize its total expected system profit $\Pi(.)$ by selecting the best assortment set $S$ subject to the capacity constraint $C$. The quantity of products in the assortment is denoted by $|S|$, where $|X|$ is the cardinality of a set $X$. Subsequently, the assortment-planning problem is noted as follows:

$$\max \Pi(S) = \sum_{i\in S}\left(\alpha_i r_i + \theta\sum_{j\notin S}\alpha_j\delta_{ji}r_i\right) \quad (1)$$

$$s.t., \ |S| \le C$$

The profit function is composed of the expected profit from the firm's direct sales of customers' first-choice demands and the expected profit from substituting the first-choice products for the products in the firm's assortment. Table 2 summarizes the notation used for modeling.

We first show that the profit of a given assortment problem can be expressed as follows.

$$\Pi(S) = \sum_{i\in S}\left(\alpha_i r_i + \theta\sum_{j\notin S}\alpha_j\frac{\alpha_i}{1-\alpha_j}r_i\right) = \sum_{i\in S}\alpha_i r_i\left(1+\theta\sum_{j\notin S}\gamma_j\right), \quad (2)$$

where $\gamma_j = \alpha_j/(1-\alpha_j)$.

**Table 2**
Notation Used in Modeling.

| Parameters | |
|---|---|
| $N$ | The set of all possible products |
| $C$ | The firm's assortment capacity |
| $\alpha_i$ | Probability that product $i$ is a customer's first choice, $i \in N$ |
| $r_i$ | Profit margin per unit product $i$, $i \in N$ |
| $\theta$ | Substitution ratio if a customer cannot find her first choice |
| $\delta_{ij}$ | Substitution probability for product $i$ with product $j$, $\delta_{ii} = 0$ and $i \in N$ |
| Variables | |
| $S$ | The firm's assortment set |

We now show that the problem is NP-Complete even when there is no cardinality constraint. We first state the decision version of the problem. Given a set of products $N$, product profit margins $r_i$, $i \in N$, purchase probabilities $\alpha_i$, $i \in N$ and substitution ratio $\theta$, is there an assortment whose profit is larger than or equal to $H$? We state this problem formally as CAPACITATED STRATEGIC ASSORTMENT PLANNING UNDER EXPLICIT DEMAND SUBSTITUTION.

**Theorem 1.** *CAPACITATED STRATEGIC ASSORTMENT PLANNING UNDER EXPLICIT DEMAND SUBSTITUTION is NP-Complete.*

Using (2), the problem given in (1) can be formulated as a mixed integer program (MIP) as follows.

$$\max_{x_i\in\{0,1\}} \Pi = \left(M - \theta\sum_{j=1}^{|N|}\gamma_j x_j\right)\sum_{i=1}^{|N|}\alpha_i r_i x_i \quad (3)$$

$$s.t., \sum_{i=1}^{|N|} x_i \le C \quad (4)$$

where

$$M = 1 + \theta\sum_{i=1}^{|N|}\gamma_i.$$

.

Note that the MIP given above is a mixed integer quadratic program (MIQP) as the objective function in (3) is in quadratic form. Since the problem is NP-complete, it is not possible reformulate the problem as a linear program as is done by Davis, Gallego, and Topaloglu (2013) for the assortment problem under an MNL model and a cardinality constraint where they linearize a fractional objective function and relax the integrality constraints as the matrix for the cardinality constraint is totally unimodular.

## 4. Properties of optimal assortments

In Section 3, we proved that assortment problem we study is NP-complete. So, this section aims to obtain structural properties optimal assortment. We first analyze the assortment-planning problem under symmetric product profit margins ($r_i = r$), in which products only differ in their customer demand rates $\alpha_i$ as noted by Cachon et al. (2005) and Alptekinoglu and Grasas (2014). Symmetrical profit margins result that the firm's optimal assortment includes a set of its most popular products. Thus, the optimal assortment is denoted as the *popular (assortment) set P*. Further, the popular assortment is the set of products in descending order according to their purchase probabilities (Kök & Xu, 2011). The most popular product $i$, which has the largest $\alpha_i$, is indexed by $i$=1; accordingly, the remaining products are sorted in descending order of $\alpha_i$. The popular set also includes a null set, or $P = \{\{\}, \{1\}, \{1, 2\}, ., \{1, 2, ., N\}\}$. The finding is stated as Theorem 2.

**Theorem 2.** *When all products have equal profit margins $r_i = r$, for $i \in N$, the optimal assortment of the firm (i) is in the popular set and (ii) fully utilizes the capacity C.*

According to Theorem 2, the firm will fully utilize its assortment capacity at the optimality and include only its most popular products. Wang (2012) demonstrates that the assortment capacity will be fully used when prices are endogenous. We can illustrate that when profit margins are exogenously set at the same value, the assortment capacity is also fully utilized, as no incentive exists to direct customers from one product to another by excluding some products from the assortment. As a result of Theorem 2, the optimal assortment can be easily obtained by including $C$ number of products with the highest demand probability.

In a more general setting, in which products may differ in both their demand probabilities $\alpha_i$ and profit margins $r_i$, it is crucial to consider the expected profitability of each product, defined by $\alpha_i r_i$, which is also called the *profit rate*. Let the product indices be set such that $\alpha_x r_x \geq \alpha_y r_y$, for all $x \leq y$, and $x, y \in N$. Product $x$ is called *dominant* over product $y$. Let $a(i)$ denote the index of the product with the $i^{th}$-largest demand rate, such that $\alpha_{a(1)} \geq \alpha_{a(2)} \ldots \geq \alpha_{a(N)}$. Lemma 1 explains the priority order of products as a function of demand rates and profit margins.

**Lemma 1.** *Product x has always a higher priority than a less dominant product y ($x \leq y$) to be in the optimal assortment*

(i) *if $\alpha_x \leq \alpha_y$ (where $r_x \geq r_y$ will also hold trivially);*
(ii) *or when $\alpha_x > \alpha_y$, if $\theta$ does not exceed the threshold level $\hat{\theta}_{xy}$, such that*

$$\hat{\theta}_{xy} = \frac{(\alpha_x r_x - \alpha_y r_y)(1-\alpha_x)(1-\alpha_y)}{(\alpha_x - \alpha_y)\sum_{i=1}^{C-1} \alpha_i r_i - \alpha_x \alpha_y (r_x - r_y) - (\alpha_x r_x - \alpha_y r_y)\left((1-\alpha_x)(1-\alpha_y)\sum_{i=C+1}^{|N|} \frac{\alpha_{a(i)}}{1-\alpha_{a(i)}} - \alpha_x \alpha_y\right)}. \quad (5)$$

If product $x$ is dominant over $y$—or $\alpha_x r_x \geq \alpha_y r_y$—three cases are possible regarding their profit margins and demand rates. Case (i): $\alpha_x \leq \alpha_y$ and $r_x \geq r_y$. As the profit margin $r_x$ and the profit rate $\alpha_x r_x$ of product $x$ are both higher, $x$ is more preferable for the assortment. It is advantageous for the firm to keep product $x$ with a lower demand rate in the assortment, because product $y$ has a higher demand rate, but a lower margin can be set out of the assortment and its demand can be directed to product $x$ through substitution. Here, this aims to satisfy the demand for a low-margin item by offering a higher margin substitute although customers express less preference for the high-margin product.

Case (ii): $\alpha_x > \alpha_y$ and $r_x \leq r_y$. When the demand rate of product $x$ is higher than that of product $y$, $x$ is always the priority if the substitution ratio $\theta$ does not exceed a threshold limit. If the $\theta$ surpasses this threshold, product $y$ might be preferred as it has a higher margin despite its lower demand rate. The firm may benefit more from directing customers to higher-margin products when the $\theta$ is higher.

Case (iii): $\alpha_x > \alpha_y$ and $r_x \geq r_y$. This is the most compelling case out of the three. Although product $x$ has a higher margin as well as higher popularity—which directly leads to the dominance of $x$ over $y$, it is still not guaranteed that $x$ has a higher priority than $y$ to be in the optimal assortment. This result can be illustrated by the following example: Let $|N| = 4$ with $\alpha_1 = 0.4$, $\alpha_2 = 0.3$, $\alpha_3 = 0.2$, $\alpha_4 = 0.1$, $r_1 = 5.1$, $r_2 = 6$, $r_3 = 5$, $r_4 = 9$, $\theta = 0.9$, and $C = 3$. For this problem setting, the firm's optimal assortment is $S = \{2, 3, 4\}$, with a total expected profit of $\Pi(S) = 5.92$. Thus, product 1, which has a higher margin and demand rate than product 3, is out of the optimal assortment and product 3 is in. When the problem is solved with $\theta = 0.8$ by keeping everything else the same, the firm's optimal assortment is $S = \{1, 2, 4\}$, where product 1 is added into the assortment by removing product 3. This result can be explained by the higher demand rate of product 1 than product 3, which also results in a higher substitution potential to other high-margin available products, such as products 2 and 4, when product 1 is out of assortment. When the $\theta$ is high enough, a highly popular product that also has a high margin can be set out of the assortment to benefit from the product's popularity to direct customers to other products with even higher margins. Li (2007) notes a similar behavior, but explains that the situation occurs due to high overage costs and the high demand variability among the high-margin and high-demand rate products, where substitution is not allowed. However, our model strategically encourages substitutions under customers' high substitution probabilities.

Note that the limit on the substitution ratio (5) is a sufficient condition, but not necessary; regarding the above example, $\theta_{13} = 0.63$. Thus, for any $\theta \leq \theta_{13} = 0.63$, product 1 is always preferred over product 3 in the optimal assortment. However, and as previously reported, the optimal assortment is $S = \{1, 2, 4\}$ when $\theta = 0.8$. Thus, product 1 is preferable to product 3, even if the $\theta$ surpasses the threshold $\theta_{13}$. Benefiting from Lemma 1, the optimal assortment is characterized in Theorem 3.

**Theorem 3.**

(i) *If all products satisfy the relationships $r_x \geqslant r_y$ and $\alpha_x \leqslant \alpha_y$ for $x \leqslant y$, then the optimal assortment is composed of some number of the most dominant products.*
(ii) *Otherwise, it is more probable for the optimal assortment to include only some number of the most dominant products, when the $\theta$ and/or capacity C are smaller.*

Theorem 3i states that if all products can be sorted—such that for each product pair in the possible set $N$, a dominant product has a lower demand rate, and thus, a higher margin due to the definition of dominance—then the optimal assortment can be obtained by selecting a certain number of the most dominant products. Consequently, an optimal assortment can be obtained by a "greedy" algorithm, by adding one product to the assortment at every iteration until $|S| = C$ and searching for a positive improvement in the total profit at each iteration.

Theorem 3ii illuminates the substitution ratio and assortment capacity's effects on the optimal assortment's structure. When $\theta$ is high, an incentive exists to exclude some highly dominant products from the assortment. When the statement is evaluated together with Lemma 1, it can be concluded that a dominant product with a higher demand rate can be omitted from the assortment to benefit from its demand rate to direct the customer to other higher-profit margin products. Similarly, the assortment can contain more products under a higher assortment capacity, which increases the probability of keeping high-margin substitutes for out-of-assortment products in the assortment. Thus, the optimal assortment may no longer be in the dominant set.

The proof of Theorem 3ii regarding the effect of assortment capacity $C$ also provides an additional insight; the proof is based on the result that as the capacity $C$ increases, the marginal benefit of adding product $x$ over product $y$ decreases, which may pass through zero only once from positive to negative. This implies that once it is more profitable to include product $y$ over $x$, for $y \geq x$ in the dominance order, it is never optimal to replace product $y$ with $x$ in the optimal assortment as the assortment has increased. This result is important in developing successful "greedy" heuristic algorithms.

**Table 3**
Example of an Optimal Assortment for $\alpha_i = \{0.27, 0.21, 0.19, 0.16, 0.13\}$, and $r = (20, 15, 10, 10, 9)$.

| C/Θ | 0,1 | 0,2 | 0,3 | 0,4 | 0,5 | 0,6 | 0,7 | 0,8 | 0,9 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | {1} | {1} | {1} | {1} | {1} | {1} | {1} | {1} | {1} | (3) {1} |
| 2 | {1,2} | {1,2} | {1,2} | {1,2} | {1,2} | {1,2} | {1,2} | {1,2} | {1,2} | {1,2} |
| 3 | {1,2,3} | {1,2,3} | {1,2,3} | {1,2,3} | {1,2,3} | {1,2,3} | (2) {1,2,3} | {1,2,3} | {1,2,4} | {1,2,4} |
| 4 | {1,2,3,4} | {1,2,3,4} | {1,2,3,4} | {1,2,3,4} | {1,2,3,4} | {1,2,3,4} | {1,2,3,4} | {1,2,3,4} | (4) {1,2,4} | {1,2,4} |
| 5 | {1,2,3,4,5} | {1,2,3,4,5} | {1,2,3,4,5} | {1,2,3,4,5} | {1,2,3,4,5} | {1,2,3,4,5} | {1,2,3,4} | {1,2,3,4} | {1,2,4} | {1,2,4} |

(1)

While it is revealed that the assortment capacity is fully utilized when all product margins are identical, this might not be the case when products also differ in their profit margins. This is because customers are directed to high-margin products by excluding low-margin products from the assortment. Theorem 4 proves that the firm may take a risk in making the assortment more shallow and expecting customers to substitute their demands with high-margin products under a high substitution ratio. It also introduces an upper limit on the substitution ratio, below which the assortment capacity will always be fully utilized at optimality.

**Theorem 4.**

(i) *The firm's assortment capacity is always fully utilized in an optimal solution if $\theta$ does not exceed the threshold level $\hat{\theta}_c$, such that*

$$\hat{\theta}_c = \min_{x \in N} \left\{ \hat{\theta}_x = \frac{-r_x}{r_x \sum_{i=C+1}^{|N|} \frac{\alpha_{a(i)}}{1-\alpha_{a(i)}} - \frac{1}{1-\alpha_x} \sum_{i=1}^{C-1} \alpha_i r_i} \mid \hat{\theta}_x > 0 \right\}.$$

(ii) *Capacity utilization is non-increasing in $\theta$ and $C$.*

Note that $\hat{\theta}_c$ is a sufficiency condition, but not necessary; thus, if the substitution ratio is small enough, the number of iterations to obtain the optimal solution is the capacity-combination of the size of the set of all products, which is $\binom{|N|}{C}$. So, under Theorem 4i, the number of iterations is unimodal in $C$. When the capacity limit $C$ is equal to $\lceil |N|/2 \rceil$, the number of iterations required reaches the highest level. If the conditions defined by Theorems 3i and 4 i simultaneously hold, the result is straightforward: the optimal assortment is composed first $C$ of the most dominant products.

To given an idea about how obtained optimality properties decrease the computational complexity, for a product set with $|N| = 20$ and a capacity limit of 15, it is possible to decrease the number of candidate assortments by 99.99%—from 1,042,380 to 16 if all products can be monotonically sorted according to profit margins and demand probabilities as stated by Theorem 3i. The knowledge of full utilization of capacity at optimality leads to a 98.5% reduction in the number of candidate assortments, to 15,504. If all products satisfy the dominance rule and $\theta$ does not exceed the threshold level $\hat{\theta}_c$, then the optimal set is exactly composed of the 15 most dominant products.

We further expand Theorems 3 and 4 with a numerical example. Table 3 reports the optimal assortment sets for a problem run with different values of the assortment capacity $C$ and substitution ratio $\theta$. The result indicates how the $\theta$ and $C$ affect the assortment set's dominance and capacity utilization.

Table 3 shows that with an increase in $\theta$, the capacity utilization rate decreases and the inclusion of less-dominant products increases, indicated by (1) and (4), respectively. Moreover, as $C$ increases, the capacity utilization rate decreases and the inclusion of less-dominant products increases, indicated by (2) and (3), respectively. Note that all positive and negative changes are stated in loose terms.

## 5. Computational insights on optimal assortments

This section investigates the optimal assortments' sensitivity to the changes in three parameters: the substitution ratio $\theta$, the variance of product demand rates VAR($\alpha$), and the variance of product profit margins VAR($r$). The product set's size is $|N| = 10$ for the numerical tests in this section. For each sensitivity analysis, we use 8 different levels of the parameter under test and observe the changes in the optimal solutions' properties.

We generated the problem instances in three steps. Following are the generation steps of the substitution rate sensitivity analysis problem instances. Generation of product demand rate variance and profit margin problems follow the same steps.

**Step 1:** Generate 10 different problem instances. Set the substitution rate of all problems equal to the first level (the lowest level tested for $\theta$=0.07). Set all other parameters randomly according to the distribution specified in Table 4. Scale all $\alpha_i$ values such that $\sum_{i \in N} \alpha_i = 1$. These problems constitute problem set 1.
**Step 2:** Take problem set 1 and increase the substitution rate of all problems to the next level. Do not change any other parameter value.
**Step 3:** Repeat step 2 until the substitution level reaches its 8th level (the highest level tested for $\theta$=0.97).

We solve the same problem sets for five different values of the assortment capacity $C = \{3, 5, 7, 9, 10\}$. Therefore, for each parameter's sensitivity analysis we solve a total of 10*8*5 = 400 individual problem instances. All problem instances are solved with complete enumeration.

Tables 5, 6, and 7 report the optimal assortment solution's respective sensitivities to the substitution rate $\theta$; the variance in profit margins, denoted by VAR($r$); and the variance in demand rates of all potential products in the set $N$, denoted by VAR($\alpha$). These tables report for each problem set, the assortment capacity $C$, the value of the parameter tested, and the average performance measures in the optimal solution, or specifically: the average capacity utilization $\overline{C}$, the average percentage of profit from direct sales to the total profit $\overline{DS\%}$, the average percentage of profit from the substituted demand sales in the total profit $\overline{SS\%}$, and the average percentage of change in the total profit relative to the first problem set $\overline{\Delta\Pi^1\%}$. For a problem set $s$, $\overline{\Delta\Pi^1\%}$ is computed using (6), where $p$ is the problem instance number, $s$ is the problem set number, and $\Pi_{sp}$ is the total expected system profit for the problem instance $p$ of the problem set $s$.

$$\overline{\Delta\Pi^1\%} = \frac{\sum_{p=1\ldots10} \frac{\Pi_{sp}-\Pi_{1p}}{\Pi_{1p}} \times 100}{10} \quad \forall s \in 1\ldots8. \tag{6}$$

Table 5 indicates that as the $\theta$ increases, the capacity utilization decreases, which confirms Theorem 4. Moreover, Theorem 3 states that when the $\theta$ is larger, the optimal assortment is more likely to include a non-dominant set of products. Consequently, as $\theta$ increases for a fixed $C$, there might be fewer number of products

**Table 4**
Distributions of Randomly Generated Parameters $0 < \alpha_i \leq 1$ and $\sum_{i\in N} \alpha_i = 1$ for $i =\in N$.

| $r_i$ | U(1,10) | $\theta$ | U(0, 1) | $\alpha_i$ | U(0, 1) |
|---|---|---|---|---|---|

**Table 5**
The Optimal Assortments' Sensitivity to $\theta$. $|N| = 10$.

| $C$ | $\theta$ | $\bar{C}$ | $\overline{DS\%}$ | $\overline{SS\%}$ | $\overline{\Delta\Pi^1\%}$ | $\theta$ | $\bar{C}$ | $\overline{DS\%}$ | $\overline{SS\%}$ | $\overline{\Delta\Pi^1\%}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 0.07 | 3.0 | 95.98 | 4.02 | 0.00 | 0.60 | 3.0 | 72.46 | 27.54 | 32.42 |
| 5 | | 5.0 | 98.03 | 1.97 | 0.00 | | 5.0 | 81.37 | 18.63 | 18.51 |
| 7 | | 7.0 | 99.22 | 0.78 | 0.00 | | 6.7 | 85.92 | 14.08 | 10.65 |
| 9 | | 9.0 | 99.77 | 0.23 | 0.00 | | 7.5 | 87.26 | 12.74 | 6.19 |
| 10 | | 10.0 | 100.00 | 0.00 | 0.00 | | 7.5 | 87.26 | 12.74 | 5.41 |
| 3 | 0. 20 | 3.0 | 88.83 | 11.17 | 8.04 | 0.73 | 3.0 | 67.65 | 32.35 | 40.60 |
| 5 | | 5.0 | 93.73 | 6.27 | 4.11 | | 4.9 | 76.20 | 23.80 | 23.74 |
| 7 | | 7.0 | 97.38 | 2.62 | 1.58 | | 6.6 | 80.82 | 19.18 | 14.57 |
| 9 | | 9.0 | 99.07 | 0.93 | 0.49 | | 7.2 | 82.10 | 17.90 | 9.60 |
| 10 | | 9.8 | 99.66 | 0.34 | 0.04 | | 7.2 | 82.10 | 17.90 | 8.79 |
| 3 | 0. 33 | 3.0 | 82.53 | 17.47 | 16.15 | 0.87 | 3.0 | 63.91 | 36.09 | 48.90 |
| 5 | | 5.0 | 89.18 | 10.82 | 8.68 | | 4.9 | 71.92 | 28.08 | 29.27 |
| 7 | | 7.0 | 93.03 | 6.97 | 3.99 | | 6.6 | 77.28 | 22.72 | 18.63 |
| 9 | | 8.6 | 96.50 | 3.50 | 1.67 | | 7.1 | 78.95 | 21.05 | 13.23 |
| 10 | | 8.8 | 96.84 | 3.16 | 0.95 | | 7.1 | 78.95 | 21.05 | 12.40 |
| 3 | 0. 47 | 3.0 | 77.16 | 22.84 | 24.29 | 0.97 | 3.0 | 61.37 | 38.63 | 55.12 |
| 5 | | 5.0 | 84.87 | 15.13 | 13.59 | | 4.9 | 69.68 | 30.32 | 33.47 |
| 7 | | 6.7 | 88.66 | 11.34 | 7.16 | | 6.5 | 73.90 | 26.10 | 21.87 |
| 9 | | 8.0 | 92.24 | 7.76 | 3.51 | | 6.9 | 74.94 | 25.06 | 16.22 |
| 10 | | 8.0 | 92.24 | 7.76 | 2.75 | | 6.9 | 74.94 | 25.06 | 15.37 |

**Table 6**
The Optimal Assortments' Sensitivity to Product Margin Variances $VAR(r)$, where $\mu(r)$ =6.5. $|N| = 10$.

| $C$ | $VAR(r)$ | $\bar{C}$ | $\overline{DS\%}$ | $\overline{SS\%}$ | $\overline{\Delta\Pi^1\%}$ | $VAR(r)$ | $\bar{C}$ | $\overline{DS\%}$ | $\overline{SS\%}$ | $\overline{\Delta\Pi^1\%}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 0.00 | 3.0 | 81.12 | 18.88 | 0.00 | 2.99 | 3.0 | 79.94 | 20.06 | 5.08 |
| 5 | | 5.0 | 89.55 | 10.45 | 0.00 | | 5.0 | 87.02 | 12.98 | 4.45 |
| 7 | | 7.0 | 95.76 | 4.24 | 0.00 | | 7.0 | 91.81 | 8.19 | 3.57 |
| 9 | | 9.0 | 99.42 | 0.58 | 0.00 | | 8.6 | 93.45 | 6.55 | 2.40 |
| 10 | | 10.0 | 100.00 | 0.00 | 0.00 | | 9.1 | 94.04 | 5.96 | 2.02 |
| 3 | 0.19 | 3.0 | 80.96 | 19.04 | 0.72 | 4.68 | 3.0 | 78.61 | 21.39 | 7.73 |
| 5 | | 5.0 | 89.44 | 10.56 | 0.46 | | 5.0 | 86.35 | 13.65 | 6.56 |
| 7 | | 7.0 | 95.52 | 4.48 | 0.40 | | 6.9 | 89.47 | 10.53 | 5.28 |
| 9 | | 9.0 | 98.48 | 1.52 | 0.30 | | 8.4 | 92.99 | 7.01 | 3.61 |
| 10 | | 9.8 | 98.96 | 1.04 | 0.24 | | 8.8 | 93.16 | 6.84 | 3.14 |
| 3 | 0.75 | 3.0 | 80.96 | 19.04 | 1.83 | 6.73 | 3.0 | 78.42 | 21.58 | 10.86 |
| 5 | | 5.0 | 89.10 | 10.90 | 1.46 | | 5.0 | 85.18 | 14.82 | 9.03 |
| 7 | | 7.0 | 94.96 | 5.04 | 1.10 | | 6.8 | 88.50 | 11.50 | 7.38 |
| 9 | | 8.8 | 97.56 | 2.44 | 0.74 | | 8.1 | 89.84 | 10.16 | 5.08 |
| 10 | | 9.6 | 98.09 | 1.91 | 0.62 | | 8.4 | 89.98 | 10.02 | 4.47 |
| 3 | 1.68 | 3.0 | 80.90 | 19.10 | 3.22 | 9.17 | 3.0 | 78.14 | 21.86 | 14.47 |
| 5 | | 5.0 | 88.22 | 11.78 | 2.70 | | 5.0 | 83.73 | 16.27 | 11.86 |
| 7 | | 7.0 | 93.12 | 6.88 | 2.10 | | 6.8 | 88.50 | 11.50 | 9.55 |
| 9 | | 8.8 | 95.18 | 4.82 | 1.43 | | 8.0 | 89.84 | 10.16 | 6.84 |
| 10 | | 9.4 | 95.36 | 4.64 | 1.15 | | 8.3 | 89,97 | 10,03 | 6,11 |

in the optimal assortment where some less dominant products are kept, while more dominant products will be set outside. Thus, the ratio of substitute sales to total profit would increase, as can be observed in Table 5. The increase in substitute sales leads to increased profit, as denoted by $\overline{\Delta\Pi^1\%}$. On the one hand, the increase in $\overline{\Delta\Pi^1\%}$ is higher when $C$ is smaller; a firm with a smaller capacity carries a smaller variety of products, although the average capacity utilization $\bar{C}$ rate is higher, as stated in Theorem 3. Therefore, the higher $\theta$ would result in a higher benefit when the $C$ is small and the firm relies more on substitution sales.

Table 6 illustrates the optimal solution's sensitivity to all potential products' variance in profit margins, denoted by VAR($r$). Initially VAR($r$) equals zero which means all products have equal profit margins. VAR($r$) gradually increases and the variance of product margins of all products in the set is equal to 9.17 in the last set. In order to increase the variance, for a product set of $|N| = 10$, the profit margins of the first 5 products increased whereas the profit margins of the last 5 products decreased gradually in a symmetrical manner while keeping the mean value constant at $\mu(r) = 6.5$ level.

The results indicate that as VAR($r$) increases, the capacity utilization may decrease, and the rate of decrease is higher when the $C$ is larger. This is because low-margin items are omitted from the assortment as the variance in products' margins increases, which will direct customers to higher-margin products and increase the substitute sales percentage, regardless of whether the substitution rate changes. The total profit also increases, as now it is easier to differentiate the higher-margin products and keep them in the assortment. Intuitively, the effect of VAR($r$) on the increase in profit is smaller when the capacity is larger, as less of a need exists to differentiate products according to their margins when more products can be kept in the assortment. It is noteworthy that in some instances the average profit increases with the increase in VAR($r$), although the assortment set remains exactly the same. If the

**Table 7**
The Optimal Assortments' Sensitivity to Variances in Product Demand Rates $VAR(\alpha)$, where $\mu(\alpha)$ =0.1. $|N| = 10$.

| $C$ | $VAR(\alpha)$ | $\overline{C}$ | $\overline{DS\%}$ | $\overline{SS\%}$ | $\overline{\Delta\Pi^1\%}$ | $VAR(\alpha)$ | $\overline{C}$ | $\overline{DS\%}$ | $\overline{SS\%}$ | $\overline{\Delta\Pi^1\%}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 0.00000 | 3.0 | 73.94 | 26.06 | 0.00 | 0.0014 | 3.0 | 75.90 | 24.10 | 10.64 |
| 5 | | 5.0 | 79.38 | 20.62 | 0.00 | | 5.0 | 80.98 | 19.02 | 4.77 |
| 7 | | 6.8 | 85.00 | 15.00 | 0.00 | | 6.8 | 85.93 | 14.07 | 2.41 |
| 9 | | 8.0 | 88.56 | 11.44 | 0.00 | | 8.0 | 88.37 | 11.63 | 0.99 |
| 10 | | 8.3 | 88.83 | 11.17 | 0.00 | | 8.3 | 88.55 | 11.45 | 0.91 |
| 3 | 0.00009 | 3.0 | 74.03 | 25.97 | 0.62 | 0.0023 | 3.0 | 76.43 | 23.57 | 15.00 |
| 5 | | 5.0 | 79.63 | 20.37 | 0.39 | | 5.0 | 81.91 | 18.09 | 6.73 |
| 7 | | 6.8 | 85.20 | 14.80 | 0.35 | | 6.8 | 86.23 | 13.77 | 3.38 |
| 9 | | 8.0 | 88.48 | 11.52 | 0.15 | | 8.0 | 88.29 | 11.71 | 1.40 |
| 10 | | 8.3 | 88.78 | 11.22 | 0.21 | | 8.3 | 88.45 | 11.55 | 1.17 |
| 3 | 0.00036 | 3.0 | 74.49 | 25.51 | 3.29 | 0.0033 | 3.0 | 77.59 | 22.41 | 19.83 |
| 5 | | 5.0 | 80.12 | 19.88 | 1.38 | | 5.0 | 82.81 | 17.19 | 8.95 |
| 7 | | 6.8 | 85.41 | 14.59 | 0.83 | | 6.7 | 86.14 | 13.86 | 4.52 |
| 9 | | 8.0 | 88.40 | 11.60 | 0.34 | | 7.9 | 87.92 | 12.08 | 1.83 |
| 10 | | 8.3 | 88.72 | 11.28 | 0.44 | | 8.2 | 88.06 | 11.94 | 1.44 |
| 3 | 0.00081 | 3.0 | 74.91 | 25.09 | 6.63 | 0.0044 | 3.0 | 78.13 | 21.87 | 24.90 |
| 5 | | 5.0 | 80.61 | 19.39 | 3.02 | | 5.0 | 83.51 | 16.49 | 11.27 |
| 7 | | 6.8 | 85.72 | 14.28 | 1.57 | | 6.7 | 86.63 | 13.37 | 5.71 |
| 9 | | 8.0 | 88.29 | 11.71 | 0.61 | | 8.0 | 88.49 | 11.51 | 2.34 |
| 10 | | 8.3 | 88.64 | 11.36 | 0.67 | | 8.3 | 88.49 | 11,51 | 1,72 |

optimal solution already includes the products with the highest margins, increasing the variance in margins increases both direct and substitute sales profits.

Table 7 illustrates the optimal solution's sensitivity to the variance in demand rates of all potential products, denoted by VAR($\alpha$). The VAR($\alpha$) is increased by dividing the product set into two groups and increasing the demand rates of the products in one group while decreasing the demand rates in the other group. Note that as all products' total demand rate always equals one by definition, then $\mu(\alpha)$ =0.1. The results indicate a negligible change in capacity utilization. Consistently, the average percentage distribution of profit from direct and substituted demand sales is almost monotone in VAR($\alpha$). Alternatively, the average profit increases with the increase in variance in demand rates. The effect on profit is even more significant in small-capacity settings ($C$ = 3, 5, and 7). In fact, the results are noteworthy compared to those sensitivity effects observed in Tables 5 and 6, in which an increase in total profit occurred due to the higher substitution sales with narrower assortments. With the increase in VAR($\alpha$), we do not observe any shrinkage in assortment size or increase in percentage of profit from substitute sales. The total profits increase because the demand rates of products in the assortment increase with an increase in VAR($\alpha$), and thus, the outside products' demand rates decrease, awarding higher direct sales. In contrast, if the demand rates of products in the assortment decrease, then the demand rates increase for the outside products, bringing higher profits from substitute sales. Thus, even if the profit distribution is almost constant, the system benefits from higher sales through both higher direct sales and substitutions. As small-capacity problems have a larger product set left out of their assortment list, they have more options to consider and can benefit more from variance changes than large capacity problems. Additionally, small-capacity problems' profits are less than that of large capacity problems; consequently, the same amount of profit increase leads to a larger percentage change in small-capacity problems.

## 6. Approximate assortment policies

Firms should make essentially two decisions during assortment planning under capacity constraints: how much of the capacity to use and which products to include in the assortment set. The decision as to whether to use the complete capacity depends on the product substitution expectations. Section 5 demonstrates that as the substitution rate $\theta$ decreases, the number of products kept in the optimal solution also decreases. The decision regarding the products to include in the assortment set depends on several factors, such as the substitution rate $\theta$ and individual demand probabilities and profit margins of the products in the set of all potential products $N$. It is noted that the optimal assortment can be determined by including all of these parameters, and a closed-form solution to the optimal assortment does not seem to exist. Nonetheless, a firm might prefer to use a quicker, rational solution in which all the potential products in $N$ are sorted in a specific order according to a predefined policy and include some of the highest-order products in the assortment set $S$, with the objective of increasing the total profit $\Pi$.

Therefore, we introduce seven different policies to sort the potential products and select some of them according to this order to be included in the assortment. These policies are based on two decisions: how to arrange the products and how many of these products can be included in the assortment set. These policies can be classified in two sets: policies that use all available capacity $C$, denoted by "Full-Cap," and policies that use the capacity selectively, denoted by "Select-Cap." Three policies belong to the first group and four policies belong to the second; the following explains each policy and its rationale.

Policies that use the firm's full capacity provide valuable information when the firm's capacity is less than the size of the set of potential products. These policies sort the products using a given rule and individually insert the products in the given order until the capacity is full. The first two policies Max $\boldsymbol{\alpha_i r_i}$|Full-Capand Max$(\mathbf{1-\alpha_i})\boldsymbol{r_i}$|Full-Cap are relatively easier to implement, and specifically, they can be useful when it is hard to predict the substitution rate $\theta$, which primarily determines how much of the capacity should be utilized.

**Max $\boldsymbol{\alpha_i r_i}$| Full-Cap** Sort the products in descending order of $\boldsymbol{\alpha_i r_i}$, or from most dominant to least. Multiplying the profit margin and demand rate denotes a product's expected profitability. This policy sorts the products by the descending order of their expected profitability, and includes the most profitable until the capacity is full.

**Max $(\mathbf{1-\alpha_i})\boldsymbol{r_i}$|Full-Cap** Sort the products in descending order of $(\mathbf{1-\alpha_i})\boldsymbol{r_i}$ and include the highest $C$ of them in the assortment set. When product substitution is non-negligible, it can be profitable to leave a product with a high demand rate and low margin out of the firm's assortment. Thus, the customer demand for this product will be partially distributed among the other products in the assortment set that can bring the firm higher profits.

**Max** *Priority*|**Full-Cap** Sort the products by descending priority, where a product's priority is defined as the total number of products it has priority over. A product has priority over another product if one of the conditions in Lemma 1 is satisfied. This policy counts the number of products each product has priority over, and individually includes the products in the assortment set by checking if this inclusion violates any priority order. A product is ineligible for inclusion if another product with priority over this product has not yet been included. In each step, policy checks if the product with the highest priority in the assortment set is eligible for inclusion. If the eligibility condition is satisfied, the product is added to the assortment until the assortment capacity is full.

As the optimal assortment may include fewer number of products than the assortment capacity, we next introduce policies that do not enforce full capacity usage. These policies individually add products to the assortment set sequentially while checking the total profit. The total expected profit for the current solution and the expected profit after adding the candidate product are computed before adding the product into the assortment set, and the product is added if the total profit improves. Otherwise, the next candidate product is considered for inclusion in the assortment set. For ease of implementation, no reversing iteration occurs; specifically, a product added into the assortment set is not taken out in further iterations.

The first three policies, Max $\boldsymbol{\alpha_i r_i}$|Select-Cap, Max$\mathbf{(1-\alpha_i)r_i}$|Select-Cap, and Max *Priority*|Select-Cap use the same sorting rules as the policies in the first group as previously defined. They include products in the assortment set added individually from the sorted list if an addition improves the total profit and available capacity exists. Thus, the policy stops including more products in the assortment if either none of the out-of-assortment products positively improves profits or the assortment capacity is full. We then introduce a fourth selective policy, Max $\Pi$|Select-Cap, which facilitates dynamic ordering and selection.

**Max** $\Pi$|**Select-Cap** Sort the products in descending order of the expected profits at each iteration. This policy at each iteration evaluates all products outside the assortment set as a candidate. It computes the expected profit resulting from including each of these candidate products to the current assortment individually, and selects the one that results in the highest positive expected profit improvement. Product inclusion continues until either no product is left, which positively increases profit, or there is no available capacity.

Firms can execute these policies as follows: If the firm has the computation capability to compute the total profit in (1) and the knowledge or a reliable prediction about the substitution rate $\theta$, then the policies can be executed as previously described. However, if the firm cannot compute total profit, then the policies Max$\boldsymbol{\alpha_i r_i}$|Select-Cap and Max$\mathbf{(1-\alpha_i)r_i}$|Select-Cap can still be executed through in-time monitoring of the total profit. The firm can order the products in the set as described in one of these policies, include a product in its assortment list, and wait for one period; at the end of the period, the firm can then check if this inclusion improved total profit. If total profit increased, the firm can add the next product from the ordered set and continue including one more product in each period. If the total profit decreases at any point in time, the firm can remove the last added product and include the following product in the out-of-assortment set in its assortment set. The firm can continue adding products as long as capacity is available and the total profit increases. In-time monitoring of total profit and period-by-period product selection is applicable for systems that keep no inventory and produce in a make-to-order setting.

We next test these policies' performances and compare them to each other and with the optimal solution under various problem settings. Section 6.1 tests these policies' performances with different product set sizes and assortment capacities and also provide performance bounds for the two best performing policies. Section 6.2 investigates these policies' sensitivities to changes in the substitution rate $\theta$, variance in customer demand rate $\alpha_i$, and variance in profit margins $r_i$. The optimal solution in all problem sets is obtained through total enumeration for precision.

### 6.1. Policy performances

This section compares the approximate policies' performances under different assortment capacities $C$ and product set sizes $|N|$. For this purpose, we generated 100 problems with randomly generated parameters for $|N| \in \{6, 10, 15, 18, 20\}$. All parameters are generated using the distributions given in Table 4. Further, each set of 100 problems generated with the product set size $|N|$ is solved separately for $C = \{3, 4, \ldots, |N|\}$. We obtain the optimal assortment and assortment solutions for a problem instance through each of the approximate policies as previously introduced. The percentage of decrease in profit from using an approximate policy over the optimal solution $\Delta\Pi\%$ is obtained, and the average percentage of decrease in profit for an $|N| - C$ pair is reported by calculating over 100 problem instances, denoted as $\overline{\Delta\Pi\%}$.

Figs. 1 and 2 illustrate $\overline{\Delta\Pi\%}$ by using approximate policies for product set sizes 6, 10, 15, and 20, respectively. These figures plot both selective- and full-capacity policies using solid and dashed lines, respectively. These figures indicate that selective capacity policies perform better than their full-capacity counterparts for all $|N|$, and the performance gap increases with the increase in capacity. This is consistent with Theorem 4, which states that capacity utilization tends to decrease with the increase in $C$ in the optimal solution; thus, selective policies closely follow the optimal solution. All full-capacity heuristics result in the same percentage profit gap when $|N| = C$, as all these policies include all products in their final assortment solution. Max $\Pi$|Select-Cappolicy performs the best for almost all $|N| - C$ values. Max$\mathbf{(1-\alpha_i)r_i}$|Select-Capis the poorest-performing heuristic policy, but its performance substantially improves when $|N|$ approaches $C$. The Max *Priority*|Select-Capheuristic does not perform well when capacity is small relative to $|N|$.

Table 8 reports more results on the performances of approximate policies with selective capacity by displaying the average percentage of decrease in profit from the optimal solution $\overline{\Delta\Pi\%}$, maximum percentage profit decrease Max $\Delta\Pi(\%)$, and the number of instances for which the approximate and optimal solutions coincide, called # hits, within the set of 100 problem instances tested for each $|N| - C$ pair.

Table 8 demonstrates that Max $\Pi$|Select-Cap policy outperforms the other policies in most cases. Moreover, the Max $\alpha_i r_i$|Select-Cap policy performs much better, and especially as the product set $N$ is broader in terms of both the lower average and maximum profit gaps. It is also noteworthy that when $|N|$ is large, the average profit gap is still less than %1 even if the number of hits is low, or almost %50. As the Max $\alpha_i r_i$|Select-Cap policy includes products in descending order of their expected product profitability and the products omitted from the assortment have relatively lower profitability, a difference in the final assortment occurs compared to the optimal solution due to these lower-profitability products, which results in a quite low total expected profit difference. Alternatively, the two other policies Max$(1-\alpha_i)r_i$|Select-Cap and Max *Priority*|Select-Cap do not

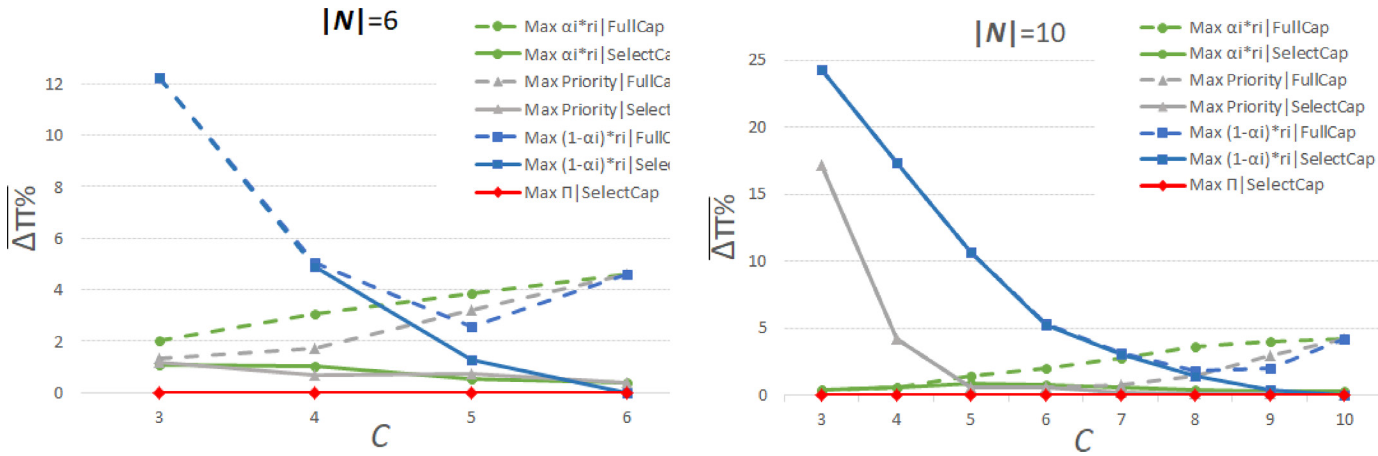


**Fig. 1.** Average profit gap $\overline{\Delta\Pi\%}$ versus the assortment capacity for all heuristics when $|N| = 6$ and $|N| = 10$.

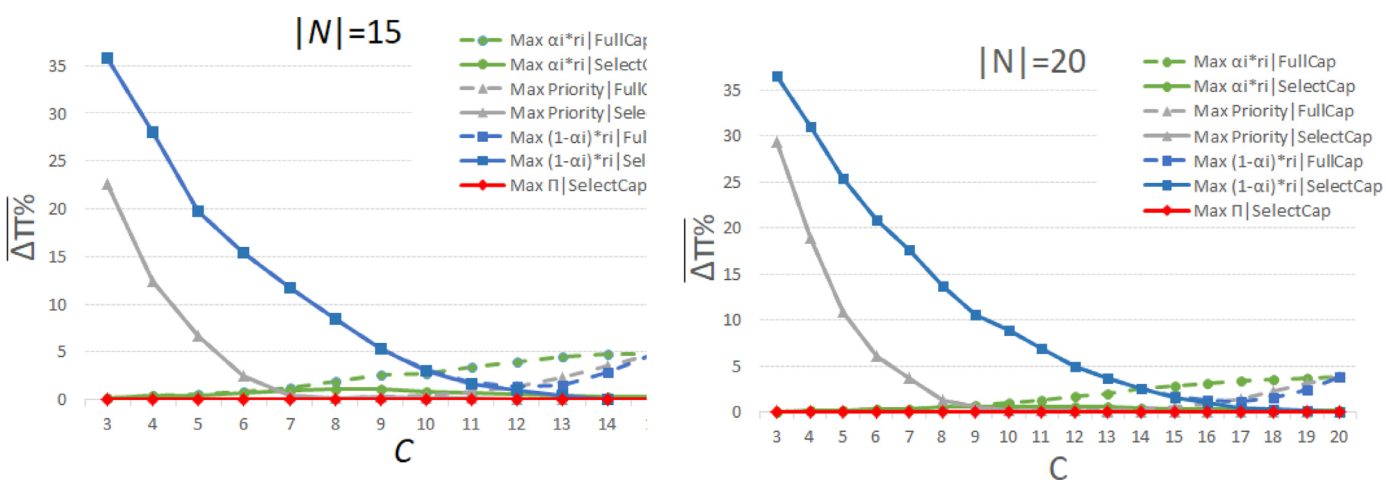


**Fig. 2.** Average profit gap $\overline{\Delta\Pi\%}$ versus the assortment capacity for all heuristics when $|N| = 15$ and $|N| = 20$.

**Table 8**
Selective Capacity Policies' Performance for Different $|N| - C$ Pairs.

| N | C | Max Π | | | Max $\alpha_i r_i$ | | | Max *Priority* | | | Max $(1-\alpha_i)r_i$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\overline{\Delta\Pi\%}$ | *Max* $\Delta\Pi\%$ | # hits | $\overline{\Delta\Pi\%}$ | *Max* $\Delta\Pi\%$ | # hits | $\overline{\Delta\Pi\%}$ | *Max* $\Delta\Pi\%$ | # hits | $\overline{\Delta\Pi\%}$ | *Max* $\Delta\Pi\%$ | # hits |
| 6 | 3 | 0.00 | 0.00 | 100 | 1.07 | 19.60 | 76 | 1.19 | 27.58 | 81 | 12.23 | 71.23 | 36 |
| | 5 | 0.00 | 0.47 | 99 | 0.54 | 12.36 | 80 | 0.70 | 48.21 | 88 | 1.27 | 24.43 | 77 |
| | 6 | 0.00 | 0.00 | 100 | 0.39 | 12.36 | 85 | 0.36 | 26.58 | 93 | 0.00 | 0.00 | 100 |
| 10 | 3 | 0.00 | 0.00 | 100 | 0.37 | 10.26 | 87 | 17.16 | 77.44 | 38 | 24.30 | 82.15 | 7 |
| | 5 | 0.00 | 0.02 | 99 | 0.84 | 7.80 | 62 | 0.54 | 12.57 | 76 | 10.65 | 41.41 | 13 |
| | 7 | 0.00 | 0.00 | 100 | 0.55 | 5.69 | 63 | 0.19 | 3.15 | 79 | 3.09 | 21.55 | 33 |
| | 9 | 0.02 | 2.17 | 98 | 0.30 | 4.35 | 76 | 0.06 | 2.27 | 84 | 0.43 | 5.88 | 75 |
| | 10 | 0.02 | 2.17 | 98 | 0.28 | 4.35 | 82 | 0.06 | 2.27 | 88 | 0.00 | 0.24 | 99 |
| 15 | 3 | 0.00 | 0.00 | 100 | 0.16 | 3.78 | 86 | 22.55 | 78.54 | 25 | 35.79 | 88.54 | 3 |
| | 5 | 0.00 | 0.04 | 99 | 0.45 | 6.62 | 65 | 6.69 | 45.62 | 45 | 19.70 | 63.12 | 5 |
| | 7 | 0.00 | 0.00 | 100 | 0.93 | 8.37 | 61 | 0.39 | 4.84 | 77 | 11.65 | 50.60 | 6 |
| | 9 | 0.00 | 0.00 | 100 | 1.03 | 6.48 | 36 | 0.25 | 2.78 | 65 | 5.28 | 23.61 | 15 |
| | 13 | 0.00 | 0.00 | 99 | 0.42 | 5.45 | 54 | 0.06 | 1.00 | 81 | 0.37 | 4.23 | 72 |
| | 15 | 0.00 | 0.00 | 99 | 0.28 | 2.99 | 68 | 0.02 | 0.58 | 85 | 0.00 | 0.00 | 100 |
| 18 | 3 | 0.00 | 0.00 | 100 | 0.02 | 0.88 | 96 | 26.80 | 72.96 | 14 | 33.82 | 85.87 | 2 |
| | 7 | 0.00 | 0.00 | 100 | 0.58 | 4.54 | 56 | 2.76 | 38.87 | 56 | 15.71 | 42.52 | 1 |
| | 11 | 0.00 | 0.01 | 99 | 0.73 | 6.31 | 42 | 0.14 | 2.13 | 72 | 6.19 | 31.84 | 5 |
| | 15 | 0.00 | 0.00 | 100 | 0.32 | 3.23 | 49 | 0.06 | 1.37 | 77 | 1.07 | 9.63 | 52 |
| | 17 | 0.00 | 0.00 | 100 | 0.21 | 3.16 | 61 | 0.03 | 0.62 | 83 | 0.22 | 3.15 | 79 |
| | 18 | 0.00 | 0.00 | 100 | 0.18 | 3.16 | 72 | 0.02 | 0.62 | 91 | 0.00 | 0.00 | 100 |
| 20 | 3 | 0.00 | 0.00 | 100 | 0.08 | 1.88 | 89 | 29.32 | 69.50 | 12 | 36.48 | 85.34 | 1 |
| | 7 | 0.00 | 0.00 | 100 | 0.38 | 2.68 | 59 | 3.64 | 29.72 | 44 | 17.60 | 41.08 | 1 |
| | 11 | 0.00 | 0.02 | 98 | 0.62 | 4.04 | 45 | 0.17 | 1.73 | 73 | 6.92 | 24.21 | 1 |
| | 15 | 0.00 | 0.41 | 99 | 0.37 | 3.12 | 46 | 0.08 | 1.13 | 69 | 1.62 | 11.89 | 34 |
| | 17 | 0.00 | 0.41 | 98 | 0.25 | 2.87 | 58 | 0.05 | 1.02 | 78 | 0.44 | 5.75 | 65 |
| | 20 | 0.00 | 0.41 | 98 | 0.16 | 2.87 | 76 | 0.03 | 1.45 | 91 | 0.00 | 0.00 | 100 |

**Table 9**
Selective Capacity Policies' Sensitivity to $\theta$. $|N| = 10$.

| $\theta$ | C | Max $\Pi$ | | | Max $\alpha_i \times r_i$ | | | Max *Priority* | | | Max $(1-\alpha_i) \times r_i$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\overline{\Delta\Pi\%}$ | *Max* $\Delta\Pi\%$ | # hits | $\overline{\Delta\Pi\%}$ | *Max* $\Delta\Pi\%$ | # hits | $\overline{\Delta\Pi\%}$ | *Max* $\Delta\Pi\%$ | # hits | $\overline{\Delta\Pi\%}$ | *Max* $\Delta\Pi\%$ | # hits |
| 0.07 | 3 | 0.00 | 0.00 | 10 | 0.05 | 0.54 | 9 | 7.65 | 21.92 | 5 | 31.45 | 77.80 | 1 |
| | 5 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 17.59 | 44.04 | 1 |
| | 7 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 9.69 | 19.93 | 0 |
| | 9 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.75 | 4.24 | 6 |
| | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 |
| 0.33 | 3 | 0.00 | 0.00 | 10 | 0.48 | 4.20 | 7 | 7.52 | 19.22 | 5 | 28.69 | 74.93 | 1 |
| | 5 | 0.00 | 0.00 | 10 | 0.29 | 1.58 | 7 | 0.14 | 0.91 | 7 | 12.91 | 38.24 | 2 |
| | 7 | 0.00 | 0.00 | 10 | 0.43 | 1.28 | 5 | 0.31 | 1.28 | 6 | 4.84 | 13.10 | 3 |
| | 9 | 0.00 | 0.00 | 10 | 0.10 | 1.05 | 9 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 |
| | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 |
| 0.60 | 3 | 0.00 | 0.00 | 10 | 0.87 | 6.79 | 7 | 11.62 | 51.02 | 4 | 26.63 | 72.65 | 1 |
| | 5 | 0.00 | 0.00 | 10 | 0.55 | 2.72 | 6 | 0.22 | 2.03 | 8 | 9.81 | 33.32 | 2 |
| | 7 | 0.00 | 0.00 | 10 | 1.07 | 5.05 | 7 | 0.00 | 0.00 | 10 | 2.54 | 6.87 | 4 |
| | 9 | 0.00 | 0.00 | 10 | 0.31 | 1.59 | 7 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 |
| | 10 | 0.00 | 0.00 | 10 | 0.31 | 1.59 | 7 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 |
| 0.87 | 3 | 0.00 | 0.00 | 10 | 1.30 | 8.71 | 5 | 14.45 | 50.44 | 4 | 25.12 | 70.79 | 1 |
| | 5 | 0.05 | 0.48 | 9 | 1.29 | 5.08 | 4 | 0.66 | 2.99 | 6 | 7.51 | 29.09 | 3 |
| | 7 | 0.00 | 0.00 | 10 | 0.47 | 3.06 | 7 | 0.03 | 0.32 | 9 | 1.13 | 4.12 | 5 |
| | 9 | 0.00 | 0.00 | 10 | 0.32 | 2.98 | 8 | 0.00 | 0.00 | 10 | 0.02 | 0.24 | 9 |
| | 10 | 0.00 | 0.00 | 10 | 0.32 | 2.98 | 8 | 0.13 | 1.33 | 9 | 0.02 | 0.24 | 9 |
| 0.97 | 3 | 0.00 | 0.00 | 10 | 1.47 | 9.31 | 5 | 14.41 | 50.25 | 4 | 24.67 | 70.18 | 1 |
| | 5 | 0.11 | 1.10 | 9 | 1.63 | 5.88 | 4 | 0.58 | 2.92 | 7 | 6.84 | 27.65 | 3 |
| | 7 | 0.00 | 0.00 | 10 | 0.83 | 4.23 | 4 | 0.18 | 1.29 | 6 | 0.79 | 3.22 | 6 |
| | 9 | 0.00 | 0.00 | 10 | 0.71 | 4.44 | 6 | 0.12 | 0.90 | 8 | 0.04 | 0.42 | 9 |
| | 10 | 0.00 | 0.00 | 10 | 0.71 | 4.44 | 6 | 0.12 | 0.90 | 8 | 0.04 | 0.42 | 9 |

**Table 10**
Selective Capacity Policies' Sensitivity to the Product Margin Variance $VAR(r)$, where $\mu(r)$ =6.5. $|N| = 10$.

| $VAR(r)$ | C | Max $\Pi$ | | | Max $\alpha_i \times r_i$ | | | Max *Priority* | | | Max $(1-\alpha_i) \times r_i$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\overline{\Delta\Pi\%}$ | *Max* $\Delta\Pi\%$ | # hits | $\overline{\Delta\Pi\%}$ | *Max* $\Delta\Pi\%$ | # hits | $\overline{\Delta\Pi\%}$ | *Max* $\Delta\Pi\%$ | # hits | $\overline{\Delta\Pi\%}$ | *Max* $\Delta\Pi\%$ | # hits |
| 0.00 | 3 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 15.90 | 59.74 | 5 | 78.28 | 88.58 | 0 |
| | 5 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 58.90 | 68.28 | 0 |
| | 7 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.31 | 3.08 | 9 | 34.60 | 51.37 | 0 |
| | 9 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.17 | 1.14 | 8 | 8.80 | 21.20 | 0 |
| | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 |
| 0.75 | 3 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 7.62 | 34.87 | 5 | 52.11 | 71.70 | 0 |
| | 5 | 0.00 | 0.00 | 10 | 0.18 | 1.36 | 8 | 0.36 | 2.29 | 8 | 26.71 | 55.08 | 0 |
| | 7 | 0.00 | 0.00 | 10 | 0.16 | 1.04 | 7 | 0.25 | 1.90 | 7 | 14.63 | 34.35 | 0 |
| | 9 | 0.00 | 0.00 | 10 | 0.07 | 0.43 | 7 | 0.51 | 2.17 | 5 | 6.23 | 19.09 | 1 |
| | 10 | 0.00 | 0.00 | 10 | 0.06 | 0.43 | 8 | 0.17 | 1.12 | 8 | 0.09 | 0.93 | 9 |
| 2.99 | 3 | 0.00 | 0.00 | 10 | 0.21 | 1.07 | 8 | 6.44 | 24.46 | 5 | 39.10 | 67.13 | 0 |
| | 5 | 0.05 | 0.48 | 9 | 0.64 | 3.38 | 6 | 0.30 | 1.47 | 7 | 16.61 | 40.36 | 0 |
| | 7 | 0.15 | 1.55 | 9 | 0.75 | 2.41 | 6 | 0.52 | 2.12 | 7 | 5.75 | 17.92 | 2 |
| | 9 | 0.21 | 2.10 | 9 | 0.61 | 3.40 | 5 | 0.79 | 3.40 | 4 | 0.47 | 3.43 | 6 |
| | 10 | 0.21 | 2.10 | 9 | 0.42 | 2.10 | 7 | 0.64 | 2.39 | 6 | 0.00 | 0.02 | 9 |
| 4.68 | 3 | 0.00 | 0.00 | 10 | 1.07 | 3.08 | 6 | 8.04 | 28.12 | 3 | 32.36 | 60.75 | 0 |
| | 5 | 0.00 | 0.00 | 10 | 1.14 | 6.01 | 7 | 0.81 | 2.60 | 6 | 11.88 | 37.33 | 0 |
| | 7 | 0.01 | 0.10 | 8 | 0.96 | 4.58 | 7 | 0.70 | 4.30 | 5 | 4.53 | 16.57 | 4 |
| | 9 | 0.00 | 0.00 | 10 | 0.93 | 5.28 | 7 | 0.34 | 2.52 | 5 | 0.31 | 2.72 | 6 |
| | 10 | 0.00 | 0.00 | 10 | 0.75 | 3.79 | 8 | 0.12 | 0.51 | 6 | 0.02 | 0.20 | 9 |
| 9.17 | 3 | 0.00 | 0.00 | 10 | 1.59 | 8.77 | 7 | 11.79 | 36.89 | 4 | 24.16 | 58.53 | 0 |
| | 5 | 0.00 | 0.00 | 10 | 1.03 | 5.95 | 7 | 0.21 | 1.66 | 8 | 9.02 | 31.08 | 3 |
| | 7 | 0.00 | 0.00 | 10 | 1.30 | 7.34 | 8 | 0.12 | 1.08 | 8 | 2.54 | 13.86 | 4 |
| | 9 | 0.00 | 0.00 | 10 | 0.81 | 5.46 | 8 | 0.01 | 0.13 | 9 | 0.14 | 1.29 | 8 |
| | 10 | 0.00 | 0.00 | 10 | 0.81 | 5.46 | 8 | 0.01 | 0.13 | 9 | 0.01 | 0.13 | 9 |

exhibit a monotonous performance in $|N|$, but can be said to perform better as $C/|N|$ is higher, which is clearly observed in Fig. 3; this figure illustrates selective capacity heuristics' performances for different capacity-to-product set size ratios $C/|N|$.

One exception is that the Max$(1-\alpha_i)r_i$|Select-Cap policy performs extremely well when $|N| = C$, and even in some cases better than Max $\Pi$|Select-Cap. As explained at the beginning of Section 6, a firm must compute its total expected profit from adding each candidate product at each iteration to use Max $\Pi$|Select-Cap policy, but the other approximate policies can be used by period-by-period profit monitoring. Therefore, when $|N| = C$, it is more efficient to use Max$(1-\alpha_i)r_i$|Select-Cap policy; otherwise, the firm can select either the Max *Priority*|Select-Cap or Max $\alpha_i r_i$|Select-Cap policies according to its capacity. Fig. 3 shows that apart from the extreme performance of Max $\Pi$|Select-Cap, if the capacity-to-product set size ratio is less than the 0.5 level, the Max $\alpha_i r_i$|Select-Cap policy performs well. When $C/|N|$ surpasses 0.5, the firm should prefer the Max *Priority*|Select-Cap policy.

Relying on superior performance of Max $\Pi$|Select-Cap and relatively well performance of Max $\alpha_i r_i$|Select-Cap heuristics in

**Table 11**
Selective Capacity Policies' Sensitivity to the Product Demand Rate Variance $VAR(\alpha)$, where $\mu(\alpha) = 0.1$, $|N| = 10$.

| $VAR(\alpha)$ | C | Max Π | | | Max $\alpha_i \times r_i$ | | | Max Priority | | | Max $(1-\alpha_i) \times r_i$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\overline{\Delta\Pi\%}$ | Max ΔΠ% | # hits | $\overline{\Delta\Pi\%}$ | Max ΔΠ% | # hits | $\overline{\Delta\Pi\%}$ | Max ΔΠ% | # hits | $\overline{\Delta\Pi\%}$ | Max ΔΠ% | # hits |
| 0 | 3 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 |
| | 5 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 |
| | 7 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 |
| | 9 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 |
| | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 |
| 0.0014 | 3 | 0.00 | 0.00 | 10 | 0.55 | 3.75 | 8 | 18.73 | 54.09 | 4 | 11.50 | 34.80 | 1 |
| | 5 | 0.00 | 0.00 | 10 | 0.38 | 2.01 | 7 | 0.55 | 3.71 | 7 | 3.56 | 12.68 | 4 |
| | 7 | 0.00 | 0.00 | 10 | 0.27 | 1.70 | 6 | 0.01 | 0.09 | 9 | 2.74 | 9.25 | 4 |
| | 9 | 0.00 | 0.00 | 10 | 0.17 | 1.70 | 9 | 0.00 | 0.00 | 10 | 0.36 | 2.40 | 8 |
| | 10 | 0.00 | 0.00 | 10 | 0.17 | 1.70 | 9 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 |
| 0.0023 | 3 | 0.00 | 0.00 | 10 | 0.19 | 1.29 | 8 | 19.97 | 59.25 | 4 | 19.18 | 43.34 | 1 |
| | 5 | 0.00 | 0.00 | 10 | 0.82 | 6.14 | 7 | 0.64 | 4.32 | 7 | 4.97 | 16.67 | 3 |
| | 7 | 0.00 | 0.00 | 10 | 0.68 | 4.18 | 6 | 0.38 | 3.72 | 7 | 3.90 | 13.00 | 4 |
| | 9 | 0.00 | 0.00 | 10 | 0.26 | 2.03 | 8 | 0.05 | 0.54 | 9 | 0.58 | 3.27 | 8 |
| | 10 | 0.00 | 0.00 | 10 | 0.26 | 2.03 | 8 | 0.05 | 0.54 | 9 | 0.00 | 0.00 | 10 |
| 0.0033 | 3 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 | 19.60 | 65.43 | 5 | 23.58 | 51.48 | 1 |
| | 5 | 0.00 | 0.00 | 10 | 0.73 | 4.67 | 7 | 0.41 | 2.49 | 8 | 6.55 | 20.35 | 3 |
| | 7 | 0.00 | 0.00 | 10 | 0.55 | 2.40 | 6 | 0.23 | 2.23 | 8 | 4.95 | 16.74 | 4 |
| | 9 | 0.00 | 0.00 | 10 | 0.36 | 2.40 | 7 | 0.03 | 0.25 | 8 | 0.79 | 4.10 | 8 |
| | 10 | 0.00 | 0.00 | 10 | 0.27 | 2.40 | 8 | 0.03 | 0.25 | 8 | 0.00 | 0.00 | 10 |
| 0.0044 | 3 | 0.00 | 0.00 | 10 | 0.22 | 2.21 | 9 | 17.31 | 63.41 | 4 | 27.81 | 59.29 | 1 |
| | 5 | 0.00 | 0.00 | 10 | 0.40 | 3.23 | 7 | 0.06 | 0.59 | 9 | 10.33 | 26.53 | 2 |
| | 7 | 0.00 | 0.00 | 10 | 0.62 | 3.34 | 6 | 0.01 | 0.09 | 9 | 5.98 | 20.33 | 4 |
| | 9 | 0.00 | 0.00 | 10 | 0.28 | 2.79 | 9 | 0.00 | 0.00 | 10 | 1.07 | 5.07 | 7 |
| | 10 | 0.00 | 0.00 | 10 | 0.28 | 2.79 | 9 | 0.00 | 0.00 | 10 | 0.00 | 0.00 | 10 |

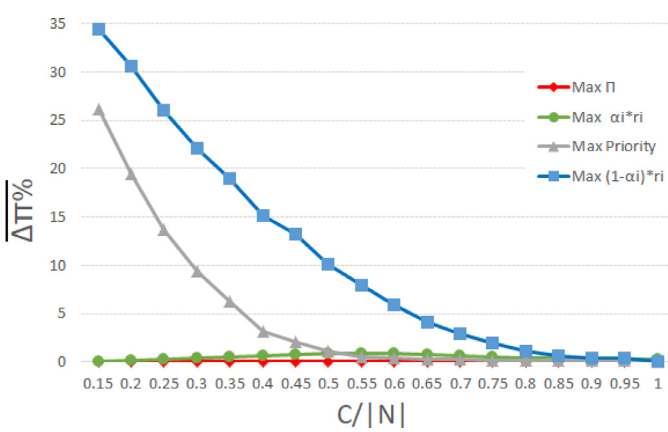


**Fig. 3.** The average profit gap $\overline{\Delta\Pi\%}$ versus the capacity rate $C/|N|$ for selective capacity heuristics.

numerical results, next we theoretically prove their worst case performances. In order to develop a performance guarantee for the greedy heuristic Max Π|Select-Cap , we first show that the set function Π is submodular. Remember that for any set Ω, a set function $f$ is submodular if $f(S \cup \{j\}) - f(S) \geq f(T \cup \{j\}) - f(T)$ for any $S \subset \Omega$ and $T \subset \Omega$ with $S \subset T$ and $j \in \Omega \setminus T$.

**Lemma 2.** *The set function Π is submodular.*

We next show the conditions under which the set function Π is non-decreasing.

**Lemma 3.** *The set function Π is non-decreasing if*

$$\alpha_j r_j + \theta\left(\alpha_j r_j \sum_{i \notin S \cup \{j\}} \gamma_i - \gamma_j \sum_{i \in S} \alpha_i r_i\right) \geq 0, \quad \text{for all } S \subset N \text{ and } j \in N \setminus S. \tag{7}$$

Note that the first term in the parenthesis in (7) is large whereas the second term is small for small $S$ and therefore the entire expression in the parenthesis is likely to be positive. When the set $S$ is large, it is possible that the expression in the parenthesis is negative. Lemma 3 shows that the set function Π (profit function) is monotone non-decreasing in the assortment as long as $\theta$ is sufficiently small to satisfy

$$\theta \leq \frac{\alpha_j r_j}{\gamma_j \sum_{i \in S} \alpha_i r_i - \alpha_j r_j \sum_{i \notin S \cup \{j\}} \gamma_i}. \tag{8}$$

Using the results in Lemmas 2 and 3, we are now ready show the performance bound on the Max Π|Select-Cap , i.e., the greedy heuristic. Theorem 5 is simply using the result in Nemhauser, Wolsey, and Fisher (1978), which shows that under a cardinality constraint, a greedy heuristic produces a solution that is at least $1 - ((C-1)/C)^C$ times the optimal value where $C$ is the maximum cardinality allowed.

**Theorem 5.** *If Π is non-decreasing, Max Π|Select-Cap heuristic has performance guarantee of*

$$1 - \left(\frac{C-1}{C}\right)^C.$$

Note that a lower bound for the performance guarantee in Theorem 5 is $1 - 1/e \approx 0.632$ which is the limit of the guarantee as $C$ goes to infinity.

Next, we prove the performance bound for Max $\alpha_i r_i$|Select-Cap , which is shown to perform quite well in numerical analyses.

**Theorem 6.** *The Max $\alpha_i r_i$|Select-Cap heuristic has performance guarantee of*

$$\frac{r_{\min}(\theta\, r_{\min} + r_{\max})}{r_{\max}(\theta\, r_{\max} + r_{\min})}$$

*where $r_{\min} = \min_{j \in N} r_j$ and $r_{\max} = \max_{j \in N} r_j$.*

### 6.2. Sensitivity of approximate policies

As heuristic selective capacity policies dominate their full-capacity counterparts in terms of performance, this section investigates the selective capacity policies' sensitivities to the changes in the substitution rate $\theta$, and variances in the demand rates VAR($\alpha$) and profit margins VAR($r$) of all potential products in the set $N$. The policies' sensitivities are tested using the same problem sets as described in Section 5, but the results are only reported for $|N| = 10$.

Table 9 demonstrates that Max $\Pi$|Select-Cap policy continues to outperform for various $\theta$ values. The Max $\alpha_i r_i$|Select-Cap and Max *Priority*|Select-Cap perform better when the $\theta$ is smaller, while the Max $\alpha_i r_i$|Select-Cap policy performs relatively better with higher $\theta$ values. An improvement in performance for Max $(1-\alpha_i)r_i$|Select-Cap is expected because it becomes more beneficial to include high-margin and low-demand items in the assortment set when the substitution rate increases. As the substitution rate is high, the omitted products' high demand rates contribute to substitute products' sales. As Max $(1-\alpha_i)r_i$|Select-Cap policy selects the high-margin, low-demand items, it starts to perform well, even at mid-capacity levels.

Table 10 exhibits the policy performances relative to the changes in product margin variances. The results reveal that an increase in product margin variances improves the performance of Max $(1-\alpha_i)r_i$|Select-Cap . As this variance increases, Max *Priority*|Select-Cap performs more poorly than Max $\alpha_i r_i$|Select-Cap in most cases, even with a relatively large capacity.

Table 11 indicates that with an increase in demand rate variances, the Max *Priority*|Select-Cap again outperforms Max $\alpha_i r_i$|Select-Cap in large-capacity cases.

## 7. Concluding remarks

This paper examines the strategic assortment optimization problem of a firm. It is a strategic level decision, because manufacturing infrastructure investment is based on the assortment selected. Our proposed methodology is also applicable to other problem settings without significant inventory concerns during the assortment optimization stage. We consider the cardinality constraint on the assortment and customer demand is defined with an exogenous demand model, where each customer has a predetermined preference for each product from the potential set. Proportional demand substitutions are also considered to explain customer behavior for the out of assortment products.

The analytical study of the proposed model shows that the firm's optimal assortment is composed of the most popular products and fully utilizes the assortment capacity when all products have symmetric profit margins. So, the optimal assortment can be easily obtained by including the products with the highest demand probability to fill the capacity. When products have asymmetric profit margins, it is necessary to examine each product's expected profitability. If all products can be sorted monotonically in increasing order of their profit margins and decreasing order of their demand probabilities, the optimal assortment is composed of some number of most dominant (profitable) products. The rationale behind this optimality property is to keep a low profit margin, but high demand products out of the assortment, so that their demands can be directed to higher margin substitutes. Knowing that the optimal assortment is composed of some number of the most dominant products, it is possible to significantly decrease the number of possible assortments compared to full enumeration. If all products do not posit monotonic ordering of profit margins and demand probabilities, some number of highly dominant products can omitted from the assortment under a high substitution ratio, which increases the probability of retaining high-margin substitutes.

Capacity is not always fully utilized when profit margins are asymmetrical; it is better to exclude some high-demand and low-margin products from the assortment to direct customers to higher margin products. We prove a threshold value on the substitution ratio, below which the capacity is always fully utilized at optimality. This can lead to an additional reduction in the number of possible assortments to find the optimal solution, despite a closed-form solution to the optimal assortment does not seem to exist. To further simplify assortment planning in practice to obtain a quick and rational solution, seven different heuristic algorithms are defined by relying on the optimality properties obtained in this study and compared to optimal solution. Among these heuristics, the greedy policy, which adds the product among all remaining candidate products with the highest positive expected profit improvement to the current assortment, performs the best. This policy results in an assortment solution with a less than 1% average profit gap of the optimal solution for a possible product set size of 20. It is also demonstrated that policies that add products to the assortment selectively perform better than those that enforce assortment capacity usage to the limit.

This study contributes to assortment optimization literature by demonstrating properties of optimal assortment in a generalized model where both demand rates and profit margins can be product specific and demand substitutions among products are explicitly allowed. The current study can also be further detailed and extended. Specifically, one compelling but challenging extension would involve endogenously deciding production capacities for the products to be included in the optimal assortment. This is particularly difficult because both capacity-and assortment-based substitutions should be considered; however, it would be valuable to observe how the resulting assortments' characteristics could change under a more comprehensive future approach.

## Acknowledgment

This research has been partially supported by the Scientific and Technological Research Council of Turkey (TUBITAK) Grant 110M488.

## Appendix A. Proofs of Lemmas/Theorems

**Proof of Theorem 1.** Clearly, the problem is in NP, since one can compute the profit of a given assortment and verify whether it is larger than or equal to $H$ or not in polynomial time. We will show that an instance of the Subset Sum problem can be transformed to an equivalent strategic assortment planning problem under explicit demand substitution. The Subset Sum is a known NP-complete problem (Garey & Johnson, 1979) and can be stated as follows. Given a finite set $N$, size $s_j \in Z^+$ for each $j \in N$, and a positive integer $B$, is there a subset $S \subseteq N$ such that $\sum_{j\in S} s_j = B$?

We first show that the profit of a given assortment problem can be expressed as follows.

$$\sum_{i\in S}\left(\alpha_i r_i + \theta\sum_{j\notin S}\alpha_j\frac{\alpha_i}{1-\alpha_j}r_i\right) = \sum_{i\in S}\alpha_i r_i\left(1+\theta\sum_{j\notin S}\gamma_j\right),$$

where $\gamma_j = \alpha_j/(1-\alpha_j)$. The profit can be then expressed as

$$\left(\sum_{i\in S}\alpha_i r_i\right)\left(1+\theta\sum_{i\in N}\gamma_i - \theta\sum_{i\in S}\gamma_i\right). \tag{9}$$

Assume that $\sum_{i\in N} s_i$ is even. Now consider the assortment optimization problem with the following parameters

$$\theta = 1,\quad \alpha_i = \frac{s_i}{s_i+K},\quad r_i = \frac{s_i+K}{K},$$

where $K$ is a sufficiently large even integer which ensures $\sum_{i\in N}\alpha_i \le 1$ (One can always find $K$ that is strictly smaller than

$\sum_{i\in N} s_i$). The decision problem is whether there is any assortment $S \subseteq N$ with a profit that is larger than or equal to

$$H = \frac{1}{4K^2}\left(K + \sum_{i\in N} s_i\right)^2.$$

The profit in (9) can be written as

$$f\left(\sum_{i\in S} s_i\right) = \frac{1}{K^2}\left(\sum_{i\in S} s_i\right)\left(K + \sum_{i\in N} s_i - \sum_{i\in S} s_i\right)$$

The function $f$ is concave and obtains its maximum at

$$\frac{1}{2}\left(K + \sum_{i\in N} s_i\right)$$

leading to optimum value

$$\frac{1}{K^2}\left(K + \sum_{i\in N} s_i\right)^2.$$

But this is only achievable if there is a subset $S$ where

$$\sum_{i\in S} s_i = \frac{1}{2}\left(K + \sum_{i\in N} s_i\right).$$

Hence the assortment problem has a solution with profit that is larger than or equal to $H = \frac{1}{4K^2}\left(K + \sum_{i\in N} s_i\right)^2$ if and only if the Subset Sum problem with

$$B = \frac{1}{2}\left(K + \sum_{i\in N} s_i\right)$$

has an affirmative answer. Since Subset Sum is NP-Complete so is Capacitated Strategic Assortment Planning Under Explicit Demand Substitution. □

**Proof of Theorem 2. (i)** Consider adding either product $x$ or $y$ to an existing assortment $S$, such that $\alpha_x \geqslant \alpha_y$. Let $\vartheta_x(S)$ denote the marginal benefit of adding product $x$ to an existing assortment $S$, assuming that $|S| \le C-1$, or specifically, the assortment capacity is not exceeded by adding product $x$. The increase in profit is denoted by adding $x$ as follows:

$$\vartheta_x(S) = \Pi(S\cup\{x\}) - \Pi(S) = r\alpha_x + r\theta \sum_{i\notin S\cup\{x\}} \alpha_i\delta_{ix} - r\theta\sum_{i\in S}\alpha_x\delta_{xi}.$$

Next, we consider adding product $y$ to assortment $S$.

$$\vartheta_y(S) = \Pi(S\cup\{y\}) - \Pi(S) = r\alpha_y + r\theta \sum_{i\notin S\cup\{y\}} \alpha_i\delta_{iy} - r\theta\sum_{i\in S}\alpha_y\delta_{yi}.$$

When we replace $\delta_{ik}$ with its open form $\frac{\alpha_k}{1-\alpha_i}$, the difference between the two alternative assortments' marginal profits is obtained as follows:

$$\begin{aligned}
&\vartheta_x(S) - \vartheta_y(S)\\
&= r(\alpha_x-\alpha_y) + r\theta\left[\alpha_x\sum_{i\notin S\cup\{x\}}\frac{\alpha_i}{1-\alpha_i} - \alpha_y\sum_{i\notin S\cup\{y\}}\frac{\alpha_i}{1-\alpha_i}\right.\\
&\quad\left. - \sum_{i\in S}\alpha_i\left(\frac{\alpha_x}{1-\alpha_x} - \frac{\alpha_y}{1-\alpha_y}\right)\right]\\
&= r(\alpha_x-\alpha_y) + r\theta\left[\sum_{i\notin S\cup\{x,y\}}\frac{\alpha_i(\alpha_x-\alpha_y)}{1-\alpha_i} - \frac{\alpha_x\alpha_y(\alpha_x-\alpha_y)}{(1-\alpha_x)(1-\alpha_y)}\right.\\
&\quad\left. - \sum_{i\in S}\frac{\alpha_i(\alpha_x-\alpha_y)}{(1-\alpha_x)(1-\alpha_y)}\right]\\
&= r(\alpha_x-\alpha_y)\left[1 - \frac{\theta(\alpha_x\alpha_y + \sum_{i\in S}\alpha_i)}{(1-\alpha_x)(1-\alpha_y)} + \theta\sum_{i\notin S\cup\{x,y\}}\frac{\alpha_i}{1-\alpha_i}\right]. \quad (10)
\end{aligned}$$

Consider the first and second terms inside the square brackets in (10). We can demonstrate that

$$\begin{aligned}
&1 \ge \frac{\theta(\alpha_x\alpha_y + \sum_{i\in S}\alpha_i)}{(1-\alpha_x)(1-\alpha_y)}\\
&1 - \alpha_x - \alpha_y + \alpha_x\alpha_y \ge \theta(\alpha_x\alpha_y + \sum_{i\in S}\alpha_i)\\
&1 - \alpha_x - \alpha_y - \theta\sum_{i\in S}\alpha_i \ge (\theta-1)\alpha_x\alpha_y \quad (11)
\end{aligned}$$

As $1-\alpha_x-\alpha_y \ge \sum_{i\in S}\alpha_i$ holds by definition, and given that $\theta \le 1$, the left-hand side of (11) is non-negative, while the right-hand side is non-positive. Thus, (11) is always satisfied, and the sum of the terms inside the square brackets in (10) is non-negative. Consequently, it is never better to include $y$ rather than $x$ for $\alpha_x \geqslant \alpha_y$. Hence, a single firm's optimal assortment exists in the popular set when all products have the same profit margin.

**(ii)** Consider the marginal benefit of adding product $x$ to an existing assortment $S$. After replacing $\delta_{ik}$ with its open form $\frac{\alpha_k}{1-\alpha_i}$, $\delta_x(S)$ is rearranged as follows:

$$\begin{aligned}
\vartheta_x(S) &= r\alpha_x + r\alpha_x\theta\sum_{i\notin S\cup\{x\}}\frac{\alpha_i}{1-\alpha_i} - r\theta\frac{\alpha_x}{1-\alpha_x}\sum_{i\in S}\alpha_i\\
&= r\alpha_x\left(1 + \theta\sum_{i\notin S\cup\{x\}}\frac{\alpha_i}{1-\alpha_i} - \theta\frac{1}{1-\alpha_x}\sum_{i\in S}\alpha_i\right). \quad (12)
\end{aligned}$$

Consider the first and third terms inside the parentheses in (12). We can demonstrate that

$$\begin{aligned}
&1 \ge \theta\frac{1}{1-\alpha_x}\sum_{i\in S}\alpha_i\\
&1 \ge \alpha_x + \theta\sum_{i\in S}\alpha_i.
\end{aligned}$$

The result holds given that $\theta \le 1$, $\sum_{i\in N}\alpha_i$=1, and $S\cup\{x\} \subseteq N$. Thus, it is always profitable to add a product to the assortment as long as capacity is available. Hence, capacity is fully utilized at the optimal assortment when all products have the same profit margin. □

**Proof of Lemma 1. (i)** Consider adding product $x$ to an existing assortment $S$. The marginal profit of adding $x$ is as follows; recall that $\vartheta_x(S)$ denotes the marginal benefit of adding product $x$ to an existing assortment $S$ assuming that $|S| \le C-1$, or that the assortment capacity is not exceeded by adding product $x$:

$$\vartheta_x(S) = \alpha_x r_x + r_x\theta\sum_{i\notin S\cup\{x\}}\alpha_i\delta_{ix} - \theta\sum_{i\in S}\alpha_x\delta_{xi}r_i.$$

Now we consider adding product $y$ to the existing assortment $S$.

$$\vartheta_y(S) = \alpha_y r_y + r_y\theta\sum_{i\notin S\cup\{y\}}\alpha_i\delta_{iy} - \theta\sum_{i\in S}\alpha_y\delta_{yi}r_i.$$

When we replace $\delta_{ik}$ with its open form—and after some rearrangement—the difference in the two alternative assortments' marginal profits is noted as follows:

$$\begin{aligned}
&\vartheta_x(S) - \vartheta_y(S)\\
&= (\alpha_x r_x - \alpha_y r_y) + \theta(\alpha_x r_x - \alpha_y r_y)\sum_{i\notin S\cup\{x,y\}}\frac{\alpha_i}{1-\alpha_i}\\
&+ \theta\left(\frac{\alpha_y}{(1-\alpha_y)}\sum_{i\in S\cup\{x\}}\alpha_i r_i - \frac{\alpha_x}{(1-\alpha_x)}\sum_{i\in S\cup\{y\}}\alpha_i r_i\right)\\
&= (\alpha_x r_x - \alpha_y r_y) + \theta(\alpha_x r_x - \alpha_y r_y)\sum_{i\notin S\cup\{x,y\}}\frac{\alpha_i}{1-\alpha_i}\\
&+ \theta\left(\alpha_x\alpha_y\left(\frac{r_x}{(1-\alpha_y)} - \frac{r_y}{(1-\alpha_x)}\right) + \sum_{i\in S}\alpha_i r_i\left(\frac{\alpha_y}{(1-\alpha_y)} - \frac{\alpha_x}{(1-\alpha_x)}\right)\right) \quad (13)
\end{aligned}$$

In (13), the first term is the profit difference from direct customer demands, the second term occurs due to the substitution from out-of-assortment products, and the last term is the profit difference between the substitution from product $y$ to other existing products and the substitution from $x$ to others. The first and second terms in (13) are non-negative, as indicated by the dominance of $x$ over $y$, or $\alpha_x r_x \geq \alpha_y r_y$. Within the third term, the second part is non-negative, because $\alpha_y \geq \alpha_x$ and $(1-\alpha_y) \leq (1-\alpha_x)$. Regarding the first part within the third term, as $r_x\alpha_x \geq r_y\alpha_y$ and $\alpha_x \leq \alpha_y$ by definition, it follows that $(1-\alpha_y) \leq (1-\alpha_x)$ and $r_x \geq r_y$. Then it is easy to see that $r_x/(1-\alpha_y) \geq r_y/(1-\alpha_x)$. This proves that the third term in (13) is also non-negative. Thus, the expected marginal profit from adding product $x$ to an existing assortment rather than $y$ is non-negative under the given conditions.

**(ii)** Reconsider the difference in marginal benefits of adding product $x$ and $y$ to an existing assortment $S$ stated by (13). (13) can be rearranged as follows:

$$\vartheta_x(S) - \vartheta_y(S) = (\alpha_x r_x - \alpha_y r_y) + \theta(\alpha_x r_x - \alpha_y r_y) \sum_{i \notin S\cup\{x,y\}} \frac{\alpha_i}{1-\alpha_i} - \theta \frac{(\alpha_x-\alpha_y)\sum_{i\in S}\alpha_i r_i}{(1-\alpha_x)(1-\alpha_y)} + \theta \frac{\alpha_x\alpha_y(r_x(1-\alpha_x) - r_y(1-\alpha_y))}{(1-\alpha_x)(1-\alpha_y)}. \quad (14)$$

Under the condition that $r_x\alpha_x \geq r_y\alpha_y$ and $\alpha_x > \alpha_y$, (14) decreases with a decrease in the second term and an increase in the third term. In the second term, the summation is over $i \notin S\cup\{x,y\}$ and in the third term, the summation is over $i \in S$. So, when the set $i \notin S\cup\{x,y\}$ gets smaller, the set $i \in S$ is expanded. For a given $x$ and $y$, $\vartheta_x(S) - \vartheta_y(S)$ obtains its minimum value when the current set $S$ contains maximum number of products, so $i \notin S\cup\{x,y\}$ gets it smallest size. By definition, $\vartheta_i(S)$ denotes the marginal benefit of adding product $i$ to an existing assortment $S$, so it should hold that $|S| \leq C-1$. Then, when $|S| = C-1$, set $i \notin S\cup\{x,y\}$ is minimized and the set $i \in S$ is maximized. Recall that $a(i)$ denotes the index of the product with the $i^{th}$ largest product preference, such that $\alpha_{a(1)} \geq \alpha_{a(2)} \ldots \geq \alpha_{a(N)}$. Subsequently, $\sum_{i\notin S\cup\{x,y\}} \frac{\alpha_i}{1-\alpha_i}$ has a minimum value of $\sum_{i=C+1}^{|N|} \frac{\alpha_{a(i)}}{1-\alpha_{a(i)}}$. Moreover, $\sum_{i\in S}\alpha_i r_i$ obtains its maximum value when the most dominant products are in $S$, such that $\sum_{i\in S}\alpha_i r_i = \sum_{i=1}^{|C-1|}\alpha_i r_i$. Thus, we can state that

$$\vartheta_x(S) - \vartheta_y(S) \geq (\alpha_x r_x - \alpha_y r_y)\left(1 + \theta \sum_{i=C+1}^{|N|} \frac{\alpha_{a(i)}}{1-\alpha_{a(i)}}\right) - \theta \frac{(\alpha_x - \alpha_y)\sum_{i=1}^{|C-1|}\alpha_i r_i - \alpha_x\alpha_y(r_x(1-\alpha_x) - r_y(1-\alpha_y))}{(1-\alpha_x)(1-\alpha_y)}.$$

For this minimum marginal profit difference to be non-negative, it should hold that

$$\theta \leq \frac{(\alpha_x r_x - \alpha_y r_y)(1-\alpha_x)(1-\alpha_y)}{(\alpha_x-\alpha_y)\sum_{i=1}^{|C-1|}\alpha_i r_i - \alpha_x\alpha_y(r_x - r_y) - (\alpha_x r_x - \alpha_y r_y)((1-\alpha_x)(1-\alpha_y)\sum_{i=C+1}^{|N|}\frac{\alpha_{a(i)}}{1-\alpha_{a(i)}} - \alpha_x\alpha_y)},$$

which is called $\hat{\theta}_{xy}$. □

**Proof of Theorem 3.**

**(i)** This directly follows from Lemma 1i.

**(ii)** Consider the equation $\vartheta_x(S) - \vartheta_y(S)$ in (14). We know that the sum of first two terms is always non-negative by definition, and the last term is a linear function of $\theta$. Thus, if the last term with a minus sign in front is positive, and the total is positive, (14) linearly decreases in $\theta$, which can pass through zero from positive to negative a maximum of once. Consequently, there is no chance that (14) can switch from negative to positive as the $\theta$ increases. Next, we first note that if $r_x \geq r_y$ and $\alpha_x \leq \alpha_y$ do not simultaneously hold, then either $r_x \geq r_y$ and $\alpha_x \geq \alpha_y$ or $r_x \leq r_y$ and $\alpha_x \geq \alpha_y$ from the definition of the dominance relationship between $x$ and $y$ for $x \leq y$. Thus, let $\alpha_x \geq \alpha_y$; as $|S|$ increases while the second (always non-negative) term decreases, the third term with a minus sign in front increases. Thus, (14) decreases overall. As the capacity $C$ increases, $|S|$ can increase, and thus, the marginal benefit of adding product $x$ over product $y$ decreases, which may pass through zero only once, from positive to negative. □

**Proof of Theorem 4. (i)** Consider the change in profit by adding one more product, such as $x$, to an existing assortment set $S$:

$$\begin{aligned}\vartheta_x(S) &= \alpha_x r_x + r_x\theta \sum_{i\notin S\cup\{x\}} \alpha_i\delta_{ix} - \theta\sum_{i\in S}\alpha_x\delta_{xi}r_i \\ &= \alpha_x r_x + \alpha_x r_x\theta \sum_{i\notin S\cup\{x\}} \frac{\alpha_i}{1-\alpha_i} - \theta\frac{\alpha_x}{1-\alpha_x}\sum_{i\in S}\alpha_i r_i \\ &= \alpha_x\left[r_x + \theta\left(r_x \sum_{i\notin S\cup\{x\}}\frac{\alpha_i}{1-\alpha_i} - \frac{1}{1-\alpha_x}\sum_{i\in S}\alpha_i r_i\right)\right]. \quad (15)\end{aligned}$$

Let the value of $\theta$, which makes $\vartheta_x(S) = 0$, be called $\theta_{Sx}$, such that

$$\theta_{Sx} = \left(\frac{-r_x}{r_x\sum_{i\notin S\cup\{x\}}\frac{\alpha_i}{1-\alpha_i} - \frac{1}{1-\alpha_x}\sum_{i\in S}\alpha_i r_i}\right).$$

For $\theta \leq \theta_{Sx}$, adding product $x$ to the current assortment $S$ is not harmful, and thus, product $x$ can be added. If $\theta_{Sx} \leq 0$, then product $x$ is profitable if included in the assortment independent of the $\theta$. If $\theta_{Sx} > 0$, over all possible $S$ sets, $\theta_{Sx}$ obtains its smallest value for the minimum value of $\sum_{i\notin S\cup\{x\}}\frac{\alpha_i}{1-\alpha_i}$ and maximum value of $\sum_{i\in S}\alpha_i r_i$. Given a capacity of $C$, the minimum set of $i \notin S\cup\{x\}$ is obtained when $|S| = C-1$. Recall that $a(i)$ denotes the index of the product with the $i^{th}$-largest product preference, such that $\alpha_{a(1)} \geq \alpha_{a(2)} \ldots \geq \alpha_{a(N)}$. Subsequently, $\sum_{i\notin S\cup\{x\}}\frac{\alpha_i}{1-\alpha_i}$ has the minimum value of $\sum_{i=C+1}^{|N|}\frac{\alpha_{a(i)}}{1-\alpha_{a(i)}}$. Moreover, $\sum_{i\in S}\alpha_i r_i$ has a maximum value when $|S| = C-1$ of $\sum_{i=1}^{C-1}\alpha_i r_i$, such that $\alpha_1 r_1 \geq \alpha_2 r_2 \ldots \geq \alpha_N r_N$ by definition. As a result, $\theta_{Sx}$ achieves its minimum non-negative value as denoted by $\hat{\theta}_x$, such that

$$\hat{\theta}_x = \max\left\{0, \frac{-r_x}{r_x\sum_{i=C+1}^{|N|}\frac{\alpha_{a(i)}}{1-\alpha_{a(i)}} - \frac{1}{1-\alpha_x}\sum_{i=1}^{C}\alpha_i r_i}\right\}.$$

Among all non-zero $\hat{\theta}_x$, the minimum value of $\hat{\theta}_x$ over all $x$ is denoted by $\hat{\theta}_c$, which is the critical substitution level to add any product $i$, or $i \in N$. If $\theta \leq \hat{\theta}_c$, it is always profitable to add a product to any given assortment as long as the capacity limit is not exceeded. Further, $\hat{\theta}_c$ can be formally defined as

$$\begin{aligned}\hat{\theta}_c &= \min_{x\in N}\left\{\hat{\theta}_x \mid \hat{\theta}_x > 0\right\} \\ &= \min_{x\in N}\left\{\hat{\theta}_x > 0 \mid \hat{\theta}_x = \frac{-r_x}{r_x\sum_{i=C+1}^{|N|}\frac{\alpha_{a(i)}}{1-\alpha_{a(i)}} - \frac{1}{1-\alpha_x}\sum_{i=1}^{C-1}\alpha_i r_i}\right\}.\end{aligned}$$

On the one hand, **(ii)** if $\theta > \hat{\theta}_c$, it is possible that (15) can be negative if the equation in parentheses is negative. If this is the case, as $\theta$ increases the $\delta_x(S)$ decreases, which may lead to the capacity's underutilization. On the other hand, for a fixed $\theta$, if the term in parentheses in (15) decreases, $\delta_x(S)$ also decreases. This may occur as the set of $i \notin S \cup \{x\}$ decreases, which may be the result of an increase in capacity $C$. □

**Proof of Lemma 2.** First note that

$$\Pi(S) = \left(\sum_{i\in S}\alpha_i r_i\right)\left(1+\theta\sum_{i\in S}\gamma_i\right)$$

$$\Pi(S\cup\{j\}) = \left(\sum_{i\in S}\alpha_i r_i + \alpha_j r_j\right)\left(1+\theta\sum_{i\in S}\gamma_i - \theta\gamma_j\right)$$

leading to

$$\Pi(S\cup\{j\}) - \Pi(S) = -\left(\sum_{i\in S}\alpha_i r_i\right)\theta\,\gamma_j + \alpha_j r_j\left(1+\theta\sum_{i\notin S}\gamma_i\right) - \alpha_j r_j\theta\,\gamma_j.$$

Similarly, we have

$$\Pi(T\cup\{j\}) - \Pi(T) = -\left(\sum_{i\in T}\alpha_i r_i\right)\theta\,\gamma_j + \alpha_j r_j\left(1+\theta\sum_{i\notin T}\gamma_i\right) - \alpha_j r_j\theta\,\gamma_j.$$

Then,

$$\Pi(S\cup\{j\}) - \Pi(S) - (\Pi(T\cup\{j\}) - \Pi(T)) = \left(\sum_{i\in T}\alpha_i r_i - \sum_{i\in S}\alpha_i r_i\right)\theta\gamma_j + \alpha_j r_j\theta\left(\sum_{i\notin S}\gamma_i - \sum_{i\notin T}\gamma_i\right).$$

Since $S \subset T$, and $\gamma_j > 0$, $\alpha_j > 0$, $r_j > 0$ for all $j$, the expression above is strictly larger than 0, which leads to the desired result. □

**Proof of Lemma 3.** Note that

$$\Pi(S\cup\{j\}) - \Pi(S) = -\left(\sum_{i\in S}\alpha_i r_i\right)\theta\,\gamma_j + \alpha_j r_j\left(1+\theta\sum_{i\notin S}\gamma_i\right) - \alpha_j r_j\theta\,\gamma_j.$$

Simplifying the expression and ensuring that it is non-negative gives the desired result. □

**Proof of Theorem 6.** Consider the case of two products with $C = 1$. Assume $r_1\alpha_1 > r_2\alpha_2$, but $r_2 > r_1$. The heuristic Max $\alpha_i r_i$|Select-Cap will select product 1 as its assortment. Resulting profit is

$$r_1\alpha_1 + \theta r_1\alpha_2.$$

On the other hand an assortment with product 2 has the following profit

$$r_2\alpha_2 + \theta r_2\alpha_1.$$

The assortment with product 2 has a larger profit and is optimal if and only if

$$\theta \geq \frac{r_1\alpha_1 - r_2\alpha_2}{r_2\alpha_1 - r_1\alpha_2}.$$

The performance of the Max $\alpha_i r_i$|Select-Cap heuristic is then

$$\frac{r_1\alpha_1 + \theta r_1\alpha_2}{r_2\alpha_2 + \theta r_2\alpha_1} = \left(\frac{r_1}{r_2}\right)\left(\frac{\alpha_1 + \theta\alpha_2}{\alpha_2 + \theta\alpha_1}\right).$$

The second term in the right hand side is increasing in $\alpha_1$ and decreasing in $\alpha_2$. However we have the conditions $r_1\alpha_1 > r_2\alpha_2$ and $\alpha_1 + \alpha_2 \leq 1$. Under these restrictions the second term obtains its minimum at

$$\frac{r_1\,\theta + r_2}{r_2\,\theta + r_1}.$$

Using this we obtain the performance bound

$$\frac{r_1(\theta\, r_1 + r_2)}{r_2(\theta\, r_2 + r_1)}.$$

Generalizing this for more than two products leads to the desired result. □

## References


Akçay, Y., & Tan, B. (2008). On the benefits of assortment-based cooperation among independent producers. *Production and Operations Management, 17*(6), 620–640.

Alptekinoglu, A., & Grasas, A. (2014). When to carry eccentric products? optimal retail assortment under consumer returns. *Production and Operations Management, 23*(5), 877–892.

Aouad, A., Farias, V., Levi, R., & Segev, D. (2018). The approximability of assortment optimization under ranking preferences. *Operations Research, 66*(6), 1457–1759.

Baloch, G., & Gzara, F. (2020). Capacity and assortment planning under one-way supplier-driven substitution for pharmacy kiosks with low drug demand. *European Journal of Operational Research, 282*, 108–128.

Bernales, P. J., Guan, Y., Natarajan, H. P., Gimenez, P. S., & Tajes, M. X. A. (2017). Less is more: Harnessing product substitution information to rationalize SKUs at intcomex. *INFORMS Journal on Applied Analytics, 47*(3), 230–243.

Bernstein, F., Kök, A. G., & Xie, L. (2015). Dynamic assortment customization with limited inventories. *Manufacturing & Service Operations Management, 17*(4), 538–553.

Besbes, O., & Saure, D. (2016). Product assortment and price competition under multinomial logit demand. *Production and Operations Management, 25*(1), 114–127.

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

Boatwright, P., & Nunes, J. C. (2001). Reducing assortment: An attribute-based approach. *Journal of Marketing, 65*(3), 50–63.

Briesch, R. A., Chintagunta, P. K., & Fox, E. J. (2009). How does assortment affect grocery store choice? *Journal of Marketing Research, 46*(2), 176–189.

Cachon, G. P., & Kök, A. G. (2007). Category management and coordination of categories in retail assortment planning in the presence of basket shoppers. *Management Science, 53*(6), 934–951.

Cachon, G. P., Terwiesch, C., & Xu, Y. (2005). Retail assortment planning in the presence of consumer search. *Manufacturing Service Operations Management, 7*(4), 330–346.

Chernev, A. (2012). Product assortment and consumer choice: An interdisciplinary review. *Foundations and Trends in Marketing, 6*(1), 1–61.

Choudhary, V., Hasija, S., & Netessine, S. (2018). Do flexibility and chaining really help? an empirical analysis of automotive plant networks. In *INSEAD working paper no. 2018/58/TOM*. Available at SSRN: https://ssrn.com/abstract=3301302

Chung, H., Ahn, H. S., & Jasin, S. (2019). (rescaled) multi-attempt approximation of choice model and its application to assortment optimization. *Production and Operations Management, 28*(2), 341–353.

Corsten, D., & Gruen, T. (2004). Stock-outs cause walkouts. *Harvard Business Review, 82*(5), 26–28.

Davis, J., Gallego, G., & Topaloglu, H. (2013). Assortment planning under the multinomial logit model with totally unimodular constraint structures. In *Working paper*.

Dèsir, A., Goyal, V., Segev, D., & Chun, Y. (2020). Constrained assortment optimization under the Markov chain-based choice model. *Management Science, 66*(2), 698–721.

Fadılоglu, M. M., Karasan, O. E., & Pınar, M. c. (2010). A model and case study for efficient shelf usage and assortment analysis. *Annals of Operations Research, 180*, 105–124.

Feldman, J. B., Paul, A., & Topaloglu, H. (2019). Assortment optimization with small consideration sets. *Operations Research*. Forthcoming. Available at SSRN 3013357

Feldman, J. B., & Topaloglu, H. (2015a). Capacity constraints across nests in assortment optimization under the nested logit model. *Operations Research, 63*(4), 812–822.

Feldman, J. B., & Topaloglu, H. (2015b). Bounding optimal expected revenues for assortment optimization under mixtures of multinomial logits. *Production and Operations Management, 24*, 1598–1620.

Feldman, J. B., & Topaloglu, H. (2017). Revenue management under the Markov chain choice model. *Operations Research, 65*(5), 1322–1342.

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

Ford (2018). Available at http://fordauthority.com/2018/03/ford-fusion-updated-for-2019-just-enough-to-keep-it-fresh. Retrieved on May 5, 2019.

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.

Gallego, G., & Topaloglu, H. (2014). Constrained assortment optimization for the nested logit model. *Management Science, 60*(10), 2583–2601.

Garey, M. R., & Johnson, D. S. (1979). *Computers and intractability: A guide to the theory of NP-Completeness*. New York, NY, USA: W. H. Freeman & Co..

Golrezaei, N., Nazerzadeh, H., & Rusmevichientong, P. (2014). Real-time optimization of personalized assortments. *Management Science, 60*(6), 1532–1551.

Goyal, V., Levi, R., & Segev, D. (2016). Near-optimal algorithms for the assortment planning problem under dynamic substitution and stochastic demand. *Operations Research, 64*(1), 219–235.

Hart, C., & Rafiq, M. (2006). The dimensions of assortment: A proposed hierarchy of assortment decision making. *The International Review of Retail, Distribution and Consumer Research, 16*(3), 333–351.

Honhon, D., Jonnalagedda, S., & Pan, X. A. (2012). Optimal algorithms for assortment selection under ranking-based consumer choice model. *Manufacturing& Service Operations Management, 14*(2), 279–289.

Jagabathula, S., & Rusmevichientong, P. (2017). A nonparametric joint assortment and price choice model. *Management Science, 63*(9), 2773–3145.

Jagabathula, S., & (2014). Assortment optimization under general choice (october 21, 2014). Available at SSRN:https://ssrn.com/abstract=2512831 or https://doi.org/10.2139/ssrn.2512831

Jordan, W. C., & Graves, S. C. (1995). Principles on the benefits of manufacturing process flexibility. *Management Science, 41*(4), 577–594.

Karabati, S., Tan, B., & Ozturk, O. C. (2009). A method for estimating stock-out-based substitution rates by using point-of-sale data. *IIE Transactions, 41*(5), 408–420.

de Vries-van Ketel, E. (2006). How assortment variety affects assortment attractiveness: A consumer perspective. In *ERIM ph.d. series research in management. erasmus research institute of management*. Available from: http://hdl.handle.net/1765/7193

Kök, A. G. (2003). *Management of product variety in retail operations*. The Wharton School, University of Pennsylvania Ph.d. dissertation.

Kök, A. G., & Fisher, M. L. (2007). Demand estimation and assortment optimization under substitution: Methodology and application. *Operations Research, 55*(6), 1001–1021.

Kök, A. G., Fisher, M. L., & Vaidyanathan, R. (2015). Assortment planning: Review of literature and industry practice. In N. Agrawal, & S. Smith (Eds.), *Retail supply chain management* (2nd ed.). New York, NY: Springer Science+Business Media.

Kök, A. G., & Xu, Y. (2011). Optimal and competitive assortments with endogenous pricing under hierarchical consumer choice models. *Management Science, 57*(9), 1546–1563.

Li, Z. (2007). A single-period assortment optimization model. *Production and Operations Management, 16*(3), 369–380.

Mahajan, S., & van Ryzin, G. (2001a). Inventory competition under dynamic consumer choice. *Operations Research, 49*(5), 646–657.

Mahajan, S., & van Ryzin, G. (2001b). Stocking retail assortments under dynamic consumer substitution. *Operations Research, 49*(3), 334–351.

Mantrala, M. K., Levy, M., Barbara, E. K., Fox, E. J., Gaidarev, P., Dankworth, B., & Shah, D. (2009). Why is assortment planning so difficult for retailers? a framework and research agenda. *Journal of Retailing, 85*(1), 71–83.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics, 15*(5), 447–470.

Misra, K. (2010). Understanding retail assortments in competitive markets. *Working paper*. London: London Business School. http://faculty.london.edu/kmisra/assets/documents/AssortmentKM.pdf

Nemhauser, G. L., Wolsey, L. A., & Fisher, M. L. (1978). An analysis of approximations for maximizing submodular set functions i. *Mathematical Programming, 14*(1), 265–294.

Nip, K., Wang, Z., & Wang, Z. (2017). Assortment optimization under a single transition model. Available at SSRN 2916110.

Pentico, D. W. (2008). The assortment problem: A survey. *European Journal of Operational Research, 190*, 295–309.

Rusmevichientong, P., Shen, Z.-J. M., & Shmoys, D. B. (2009). A PTAS for capacitated sum of ratios optimization. *Operations Research Letters, 37*(4), 230–238.

Rusmevichientong, P., Shen, Z.-J. M., & Shmoys, D. B. (2010). Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Operations Research, 58*(6), 1666–1680.

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

Sadowski, W. (1959). A few remarks on the assortment problem. *Management Science, 6*(1), 13–24.

Şen, A., Atamturk, A., & Kaminsky, P. (2018). A conic integer programming approach to constrained assortment optimization under the mixed multinomial logit model. *Operations Research, 66*(4), 994–1003.

Shin, H., Park, S., Lee, E., & Benton, W. C. (2015). A classification of the literature on the planning of substitutable products. *European Journal of Operational Research, 246*(3), 686–699.

Smith, S. A., & Agrawal, N. (2000). Management of multi-item retail inventory systems with demand substitution. *Operations Research, 48*(1), 50–64.

Stalk, G., Stephenson, S. G., & King, T. (1997). Searching for fulfillment: Order–to-delivery opportunities in the auto industry. *Supply Chain Management Review*. September

Tan, B., & Akçay, Y. (2014). Assortment-based cooperation between two make-to-stock firms. *IIE Transactions, 46*(3), 213–229.

Timonina-Farkas, A., Katsifou, A., & Seifert, R. W. (2020). Product assortment and space allocation strategies to attract loyal and non-loyal customers. *European Journal of Operational Research, 285*(3), 1058–1076.

Toyota-Europe (2018). Available at https://newsroom.toyota.eu/toyota-announces-new-investment-in-tmmf. Retrieved on April 24, 2019.

Umpfenbach, E. L., Dalkiran, E., Chinnam, R. B., & Murat, A. E. (2018a). Promoting sustainability of automotive products through strategic assortment planning. *European Journal of Operational Research, 269*, 272–285.

Umpfenbach, E. L., Dalkiran, E., Chinnam, R. B., & Murat, A. E. (2018b). Optimization of strategic planning processes for configurable products. *Journal of Operational Research Society, 69*(11), 1834–1853.

Wang, R. (2012). Capacitated assortment and price optimization under the multinomial logit model. *Operations Research Letters, 40*(6), 492–497.

Wang, R. (2013). Assortment management under the generalized attraction model with a capacity constraint. *Journal of Revenue and Pricing Management, 12*(3), 254–270.

Yücel, E., Karaesmen, F., Salman, F. S., & Türkay, M. (2009). Optimizing product assortment under customer-driven demand substitution. *European Journal of Operational Research, 199*, 759–768.