MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
https://pubsonline.informs.org/journal/msom
Vol. 00, No. 0, Xxxxx 0000, pp. 000–000
Submitted to *Manufacturing & Service Operations Management*
ISSN 1523-4614, EISSN 1526-5498

Store Layout Optimization with Endogenous Consumer Search

Victor Martínez-de-Albéniz
IESE Business School, University of Navarra, valbeniz@iese.edu

Laura Wagner
IESE Business School, University of Navarra, lwagner@iese.edu

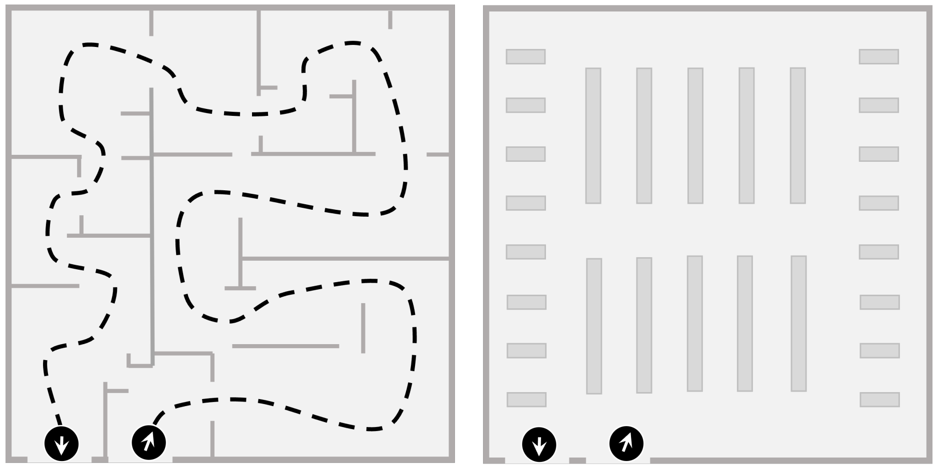
Abstract. Retailers employ various store layouts to drive traffic, conversions, and ultimately sales. At one extreme is the race-track layout, which guides shoppers through a defined route, maximizing exposure to products and encouraging unplanned purchases but at the expense of convenience. At the other extreme, open layouts allow easy movement between aisles, enhancing convenience but reducing category exposure. This paper develops a model where layout and category placement influence consumer search and purchase behavior, explicitly modeling search costs, shopper decisions to visit aisles and the resulting category-specific revenue. The retailer's layout and assignment optimization problem results in a multiple knapsack integer program, which is solvable under certain conditions (either exactly or with an FPTAS). In other cases, the problem is hard to approximate, so there is no hope of obtaining solutions with guaranteed near-optimal performance. Key insights from the model suggest that race-track layouts are optimal when search costs have economies of scale, while open-space configurations work best with diseconomies of scale. The optimal aisle composition strategically pairs high-traffic anchor categories with impulse categories. Finally, we demonstrate the model's application through a real case study, discussing estimation challenges and practical implications.

1. Introduction

Store design is an essential element in successful retailing. State-of-the-art design principles often involve architectural, aesthetic aspects, but seldom explicitly account for consumer behaviors with respect to navigating the layout, and even more rarely they involve optimization. Despite the lack of data-driven tools, practice has distilled two dominant designs worth reviewing.

Ikea, the global furniture and home goods retailer, has pioneered a highly distinctive in-store experience, in which the shopper is exposed to a winding corridor that showrooms the entire Ikea

Figure 1 Store layout choices: race-track (left) vs. open-space (right).



catalog. To some, the shopping experience is a fascinating discovery of numerous, attractive decoration options. To others, it may become a chilling experience, where the endless array of impeccably designed room setups induces a sense of disorientation in the unsettling depths of the Ikea maze. Such design is called a *race-track*. It was in fact pioneered by Piggly Wiggly supermarkets in the US in 1917, which even patented it (Tarazano and Daemmrich 2020), before Ikea popularized it globally. During the COVID pandemic in 2020, the concept was borrowed by Walmart or Kroger, who used it to enforce one-way flows and social distancing.

In contrast to this peculiar store organization, most retailers nowadays employ an *open-space* store design, which maximizes the ease of access to most parts of the store. The consumer can easily transfer from one aisle to the next, without much effort. This approach provides better visibility on the available merchandize, makes navigation of the store more flexible and less linear, and facilitates shortcuts and shorter in-store visits. In contrast, a *race-track* organization forces longer ones. Figure 1 illustrates *race-track* and *open-space* configurations.

Which one of these two approaches is most effective? From a qualitative point of view, race-track will 'trap' consumers in the store, which may increase unplanned spending (Hui et al. 2013), but with ambivalent consequences on shopper convenience and satisfaction. In contrast, open-space may reduce search costs and increase conversion (Underhill 2009, Caro et al. 2021), at the expense of having shoppers discover a low percentage of the store offer – Hui et al. (2013) reports that the average consumer only explores a third of the store. To quantify how a layout affects consumer store exploration, it thus seems essential to incorporate the cost of entering a specific area, accounting for time and physical effort. This needs to be compared to the potential utility achieved, which is determined by the product composition included in that area.

In other words, we cannot study store layouts without also incorporating the assignment of categories to locations within the store. Indeed, locating good products at the right location is likely to shape shopper in-store flows: to encourage broader in-store exploration, retailers frequently employ *anchors* or *magnets*, which are categories or products designed to generate traffic (a concept also seen in shopping mall design, see Pashigian and Gould 1998). For instance, in grocery stores, milk – a staple on most shopping lists – is often placed at the back to encourage shoppers to pass through other sections. This placement strategy channels foot traffic into less accessible aisles, boosting engagement in areas that might otherwise see fewer visitors.

The objective of this paper is precisely to develop a model that connects layout and category-location assignments with consumer behavior, and search and purchase decisions in particular. Our model should not only help us understand how architectural structures will affect store commercial success, but also provide a prescriptive decision support for retailers to optimize their store layouts.

For this purpose, we conceptualize a store as a partition of space into multiple aisles, in which race-track (a partition into a single set containing the entire assortment) and open-space (a partition into single categories) are extreme cases. In this view, the number of possible layouts is enormous and contains as many hybridizations of race-track and open-space as one desires. Furthermore, the composition of each aisle is also a decision variable for the retailer. We then explicitly model the visit decision of shoppers to each aisle, as well as the expected revenue generated at the category, contingent on aisle visit. This is done by introducing a search cost to the aisle, which is compared to the ex-ante expected utility derived from inspecting the aisle, which is only revealed once the consumer pays the search cost. This premise implies that by combining two categories in a single aisle, the retailer can cross-subsidize search costs and specifically use the excess utility from one attractive category to reduce search costs for a second category that is insufficiently attractive to generate exploration on its own.

After characterizing the consumer purchase process, we reformulate the retailer decision problem as an assignment program. When consumers are homogeneous, the problem can be simplified by introducing constraints on aisle composition that ensure that consumers visit it. The resulting formulation is a multiple knapsack integer program, which turns out to be solvable under certain conditions (either exactly or with a FPTAS). In other cases, the problem is hard even to approximate, so we cannot expect algorithms with guaranteed near-optimal performance. When consumers are heterogeneous, the formulation is less tractable but we can still develop integer formulations amenable to numerical optimization.

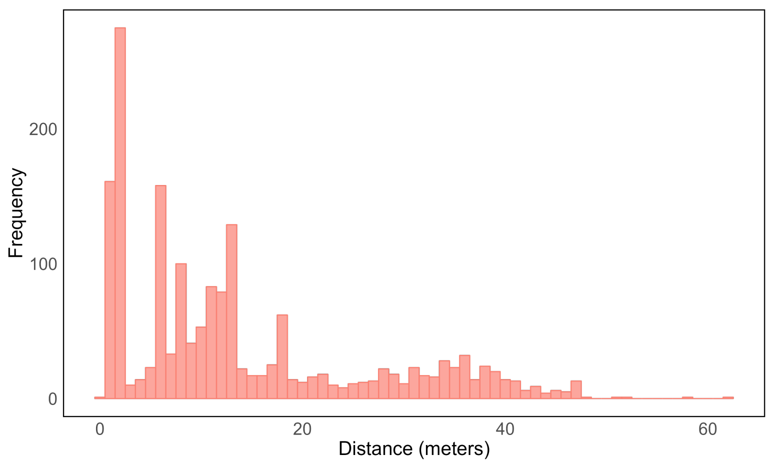
Our model reveals several intuitive and useful insights. First and foremost, the optimal store structure should resemble a race-track when search costs are submodular, i.e., there are economies of scale in search; in contrast, open-space configurations are best when there are diseconomies of scale and shoppers prefer seeing several small aisles to a larger one. Second, the optimal composition of an aisle combines *anchor* categories (such as dairy or produce) that attract shoppers, with *impulse* categories (like snacks or beverages) that are profitable but involve a relative low expected utility compared to the search cost. Anchor categories de facto subsidize the search cost incurred with impulse categories, and this is why combining them increases exposure to impulse categories without putting off consumers. Furthermore, *exploration* categories, which attract shoppers who enjoy discovering new items (such as specialty or gourmet foods), should be kept separately. Third, our numerical study indicates that higher consumer heterogeneity leads to a higher number of aisles, i.e., more open spaces. In a way, the combination of consumers with different search costs makes the race-track concept impractical, because it makes high-search cost consumers prefer not visiting the store.

In addition to our theoretical and numerical findings, we apply our framework to a real case study, (1) to demonstrate how model parameters can be estimated, and (2) to show that different stores may require different layout structures. Interestingly, in practice there is already plenty of variation on aisle structure and category co-location, suggesting that retailers may already be customizing store design to local conditions. Interestingly, such variation also implies that our model primitives should be easy to estimate. For example, Figure 2 compares the layouts of two very similar Target stores in shape and structure. In store A (Abington, MA), the *Kitchen* category is positioned near *Pets*, while in store B (Albuquerque Wyoming, NM), these categories are located much farther apart. Figure 3 quantifies this variability across 1,821 Target store layouts – all those available as of 2024,– revealing that only 26.4% of the time, *Kitchen* and *Pets* are within 4.5 meters of each other, approximately the width of the leading aisles in large U.S. hypermarkets. This variability in co-locating categories underscores the complexity that retailers face in store design, and the need for advanced decision support tools to guide them in this task.

The rest of the paper is organized as follows. Section 2 reviews the relevant literature. We formulate the retailer layout problem in §3, and analyze it for homogeneous and heterogeneous consumers in §§4-5 respectively. Section 6 includes a numerical study and applies the model to a grocery case study, and §7 concludes. Proofs are contained in the Appendix.

Figure 2 Location of categories in two Target stores.

Figure 3 Frequency of distance (in meters) between *Kitchen* and *Pets* category, for 1,821 Target stores.



2. Literature Review

Our research is related to several streams of work: research on multi-stage choice models, including consideration sets; on consumer trajectories in store; and on product bundling.

Perhaps the closest connection to our model is the literature that describes shopping as a multi-stage process. Lemon and Verhoef (2016) review the literature about the customer journey. Our model includes the decision of entering the aisle first, and then the decision to purchase within the categories within the aisle. This resembles the models in which consumers first build a consideration set of the products that they are interested in, and then choose one of them for purchase. Wang and Sahin (2018) determine that the size of the optimal consideration set depends on the search cost that the consumer incurs, and derive recommendations for assortment planning and pricing. Assortment planning under consider-then-choose process has also been studied by Aouad et al. (2021), who propose efficient dynamic programming techniques. Estimation of consideration sets from real data is discussed in Jagabathula et al. (2023) and Ziaei et al. (2023). These approaches can be embedded into multi-level consumer decisions, as in Martínez-de Albéniz et al. (2020). Similarly, our model lets consumers endogenously choose whether to consider an aisle and visit the categories contained in it. Our paper is also related to the engineering of the nests in which products are located. Hence, there is also a connection to nested logit choice models, see e.g., Wen and Koppelman (2001), Gallego and Topaloglu (2014), Feldman and Topaloglu (2015) or Li et al. (2015). Similar to our aisle bundling decision is Kök and Xu (2011), who compare choices under nesting by type of product vs. brand. In our model, we let the retailer decide which categories are placed in the same nest.

Another relevant line of work is the research that models and empirically describes how consumers move within stores. Hui et al. (2009b) provides a general discussion of how to study path data in retail stores, if one can precisely track shopper movements. Hui et al. (2009a) use path data to show that shoppers follow certain behavioral rules, including exploration fatigue and reduced conversion under congestion. Hui et al. (2013) show that unplanned spending increases when shoppers' paths get longer. Aouad et al. (2022) apply similar models to museum visitor trajectories, and highlight visit fatigue and congestion disliking as key drivers of path choices. They are also able to calibrate the impact of distance between two points on the probability of a transition occurring. Our model similarly introduces the notion of search cost to drive shopper decisions to enter an aisle. Finally, Caro et al. (2021) show that conversion rates decrease when products are located further away from the entrance, everything else being equal. These works suggest that the cost of search is a relevant factor that shapes consumer behavior, which is influenced by consumer factors (fatigue, congestion) as well as store factors (distance). In contrast with these works, we use the cost structure to guide optimal aisle composition and specifically identify sub- and super-modularity of the cost function as the key driver of the optimal solution.

Finally, the mechanism by which combining different categories within the same aisle induces more visits is related to the advantages of bundling. From the seminal paper of Adams and Yellen (1976), combining multiple items into a single bundle has been shown to dominate separate selling, with the full bundle being optimal for information goods (Bakos and Brynjolfsson 1999). In our case, we force consumers to either access a bundle of categories, within an aisle, or skip it; hence we do not allow separate access to a given category. This literature highlights production costs as determinant of the advantages of bundling (Fang and Norman 2006, Ma and Simchi-Levi 2021). In our case, we have zero production costs and consumer utilities are nonnegative, so that in principle pure bundling is optimal. However, our model also introduces ex-ante search costs that are paid before learning utilities, which deter some consumers from participation, and hence may result in smaller bundles or individual selling as the optimal structure. Like Honhon and Pan (2017) but with a very different context, we show that, even when there are no complementarities between categories, it may be beneficial to bundle categories together when consumers are pushed to buy more. We note that, as generically suggested in proposition 11 by Stremersch and Tellis (2002), subadditive costs lead to bundling.

3. Model Formulation

3.1. Retailer's layout decision

We consider a retailer (he) that is interested in optimizing his store layout and category arrangement. This involves a decision on the *planogram*, which describes where each category is located in the store. One option is to adopt a so-called *race-track* or forced-path layout, which intentionally directs shoppers from the store entrance to the checkout, forcing them to pass by every category of products along the way. Such a layout guarantees that shoppers will view all available categories, potentially leading them to purchase items they had not planned to buy, with the caveat that shoppers may find it inconvenient and abandon the retailer altogether. On the opposite end of the spectrum, the retailer could opt to display as few products together as possible to allow shoppers to quickly grab an item and go. There is, of course, a middle ground, whereby similar product categories are bundled together and organized in long separate aisles, for example, ensuring that shoppers only have to visit the particular aisle their category is placed on, reducing time while maximizing exposure for all categories in the same aisle. We build a stylized model that captures the main trade-off between shopping search costs – the smaller aisle, the more convenient to access – and category exposure – the larger aisle, the more information passed to the consumer.

To focus on this basic trade-off and abstract away from the geometric constraints of real floor plans, we conceptualize the retailer's decision process as two main decisions: (i) how many aisles $j \in J$ to place in a given store and (ii) which categories $i \in I$ to allocate to which aisle. We denote the set of categories included in an individual aisle j as A_j. Note that decision (i) is directly related to aisle size and, hence, the effort required to explore the aisle. In contrast, decision (ii) involves the interaction of the consumer with individual categories. Interestingly, as shown below, two categories within the same aisle will become correlated in the purchase process, even if they provide utilities to the consumer that are independent from each other, just because they are co-located.

3.2. Shoppers' decision process

Once the retailer has decided on the store layout design and category allocation, a shopper (she) proceeds through a two-stage process. In the first stage, she decides which aisle(s) to visit, possibly all of them, some or none at all. Upon entering aisle j, she is exposed to all the categories in A_j. Subsequently, during the second stage, a shopper having just entered an aisle, decides which categories she purchases. We allow for heterogeneous shopper types $k \in K$ (we specify later the various dimensions on which shoppers can differ). In summary, shopper k visits an aisle j initially, and then decides which categories to purchase from, conditional on visiting the aisle. We next discuss the two-stage process in more detail.

Second stage: purchase process. Let U_i^k be the utility derived from purchasing products in category i, where $U_i^k = u_i^k + \varepsilon_i^k$. The utility is thus composed of a shopper type specific product valuation u_i^k, and an independent random shock ε_i^k, where the latter is realized in the purchase stage.

We assume that categories in an aisle are independent, such that a shopper purchase categories in an aisle as long as her utility is larger than the outside option, i.e., $U_i^k \geq 0$. The specification that the outside option is valued at zero is without loss of generality, as any outside option value can be incorporated in u_i^k. In sum, we let the probability of purchasing category i, contingent on visiting the aisle in which i is located, be

$$\theta_i^k := Pr[U_i^k \geq 0 | u_i^k] = Pr[\varepsilon_i^k \geq -u_i^k] = 1 - F^k(-u_i^k), \tag{1}$$

where F^k is the c.d.f. of ε_i^k, in which we allow the distribution to vary across shoppers. For instance, under the widely-used Multinomial logit model (MNL), uncertainty arising from the product evaluation is drawn from an independent and identically distributed Gumbel random variable, which is then compared to an outside option that is valued as another Gumbel shock. This leads to ε_i^k being the difference of two Gumbel random variables, and results in having a common $F(z) = e^z/(1+e^z)$ across shoppers, who might still differ in their product valuations u_i^k. This results in $\theta_i^k = 1 - e^{-u_i^k}/(1+e^{-u_i^k}) = e^{u_i^k}/(1+e^{u_i^k})$.

In other words, in the second stage, shoppers make independent purchases in all the categories within the aisle. One could easily incorporate substitution patterns as well, but this would unnecessarily complicate our formulation, while not being overly important given that categories are precisely defined as independent product groupings without strong interactions with each other.

First stage: visit process. During the first stage, shoppers determine whether they want to visit an aisle, without a clear knowledge of which (if any) categories they will purchase within the aisle. Visiting an aisle is time consuming and thus entails a cost, which the shopper should compare to the benefit – in expected utility terms – obtained from entering the aisle. We assume that visiting an aisle j incurs a cost $c_j^k = c_0^k + \sum_{i \in A_j} c_i^k$. Hence, the cost of visiting an aisle comprises a fixed cost c_0^k as well as an additional cost for each category included in the aisle. One can interpret c_0^k as the effort or the willingness to walk into the aisle, and c_i^k as the inspection cost of category i. Intuitively, larger categories should have a higher c_i^k since more facings are encountered and more distance needs to be travelled.

Note that we do not make any assumptions on the sign of c_0^k and c_i^k. It may be possible that they are negative. An interpretation of negative search cost could be that shoppers enjoy browsing through the store and/ or that they derive an experiential utility from searching through long aisles. For example, while some people (including the authors) may feel that IKEA's store layout is a hassle, others may find it is an enjoyable treasure hunt. Whichever case applies, shoppers in the first stage weigh the expected utility of choosing to visit aisle j or not, against the cost of entering. In other words, shopper k enters aisle j if and only if

$$\sum_{i \in A_j} E_{\varepsilon_i^k} \max\{U_i^k, 0\} \geq c_0^k + \sum_{i \in A_j} c_i^k. \tag{2}$$

We let $W^k = -c_0^k$ and

$$w_i^k = c_i^k - G^k(u_i^k), \tag{3}$$

where $G^k(u_i^k) := E_{\varepsilon_i^k} \max\{u_i^k + \varepsilon_i^k, 0\}$. Note that we can relate θ_i to w_i as follows: Since $\left(G^k\right)'(u_i^k) = E_{\varepsilon_i^k} 1_{u_i^k + \varepsilon_i^k \geq 0} = 1 - F^k(-u_i^k) = \theta_i^k$, which implies that w_i is related to θ_i, for example, for the MNL, $G^k(u_i^k) = \ln\left(1 + e^{u_i^k}\right) = -\ln(1 - \theta_i^k)$. Since $(G^k)' \geq 0$ and $(F^k)' \geq 0$ (it is equal to the p.d.f. of ε_i^k evaluated at $-u_i^k$), θ_i^k is always decreasing in w_i^k. Table 1 illustrates the relation between θ_i and w_i for various common distributional assumptions.

Table 1 Expressions for θ_i^k and w_i^k for various usual distributions.

Distribution of ε_i^k	θ_i^k	w_i^k
Logistic	$e^{u_i^k}/(1+e^{u_i^k})$	$c_i^k - \ln\left(1+e^{u_i^k}\right)$
Uniform in $[0,1]$ with $-1 \le u_i^k \le 0$	u_i^k	$c_i^k - \left(u_i^k\right)^2/2$
Exponomial with $u_i^k \ge -E$	$1-e^{-(E+u_i^k)}$	$c_i^k - e^{-(E+u_i^k)} + 1 - (u_i^k + E)$
Pareto with $u_i^k \le 0$	$\left(1-u_i^k\right)^{-\alpha}$	$c_i^k - \left(1-u_i^k\right)^{1-\alpha}/(1-\alpha)$

With this notation, Equation (2) is equivalent to $\sum_{i\in A_j} w_i^k \le W^k$. The term w_i^k can be interpreted as the net cost, i.e., the difference between search costs minus expected utility: positive values of w_i^k indicate that the search cost outweighs the expected benefit, consuming the available search budget W^k, whereas negative values suggest that the expected utility exceeds the search cost, effectively increasing the budget.

As we will see below, the sign of W^k has an essential influence on optimal structures. When $W^k \ge 0$, i.e., $c_0^k \le 0$, the search cost structure is supermodular (or superadditive), because joining two aisles j and j' leads to a cost of $c_0^k + \sum_{i\in A_j \cup A_{j'}} c_i^k$, which is higher than $2c_0^k + \sum_{i\in A_j\cup A_{j'}} c_i^k$. This means that by having two separate aisles, the retailer can 'create' an additional budget $W^k \ge 0$ available to the consumer during the search process, and should intuitively lead to using more aisles to induce more exploration. In contrast, when $W^k \le 0$, i.e., $c_0^k \ge 0$, the cost function is submodular (or subadditive) and there fewer aisles should be beneficial, because they mitigate the negative budgets that consumers carry.

During the first stage, the uncertainties around the shopper type, i.e., the realizations of w_i^k and W^k, are realized. (In Section 5, we discuss in detail in which dimensions shopper types vary, and how each of the dimension affect the solution to retailer's problem.) In all its generality, the probability of a typical shopper entering aisle j can be expressed as $P\left(\sum_{i\in A_j} w_i^k \le W^k\right)$.

3.3. Retailer's optimization problem.

Letting r_i be the expected revenue for each category i, which may include average unit prices as well as number of units sold within the category. Then, the profit function for the retailer can be expressed as follows:

$$\Pi = \sum_{j\in J} P\left(\sum_{i\in A_j} w_i^k \le W^k\right) E\left[\sum_{i\in A_j} r_i\theta_i^k \middle| \sum_{i\in A_j} w_i^k \le W^k\right]$$

The retailer's planogram optimization problem can thus be formulated as follows:

$$\max_{x_{ij}=0,1} \sum_{j\in J}\sum_{i\in I} r_i x_{ij} E\left[\theta_i^k 1_{\sum_{i'\in I} w_{i'}^k x_{i'j} \le W^k}\right] \qquad (4)$$
$$\text{s.t.} \sum_{j\in J} x_{ij} \le 1 \text{ for } i\in I,$$

where the constraint ensures that one category can be placed at maximum into one aisle. As one can see, Equation (4) is an assignment problem, with the difficulty that the objective function is non-linear in the decision variables. Specifically, the revenue of category i in aisle j depends on the rest of categories present within the same aisle, through the factor $1_{\sum_{i'\in I} w_{i'}^k x_{i'j} \le W^k}$. In summary, while we assumed that the category purchase decisions are independent, the search process makes them complementary. As we will see next, this is a hard problem to solve. For this reason, we start with the case of homogeneous consumers, for which we obtain complete analytical results. We then consider the general case, with gradually increasing complexity, in Section 5.

4. Homogeneous Consumers

When consumers are homogeneous, they all share the same utility values for each product, i.e., $u_i^k = u_i$, so there is no heterogeneity about product preferences, making $\theta_i^k = \theta_i = 1 - F(-u_i)$ a deterministic quantity. They also share the same cost structure so that $c_i^k = c_i$. This implies that $w_i^k = w_i$ which becomes a deterministic quantity, and so does $W^k = W$. As a result, the random variable $1_{\sum_{i'\in I} w_{i'}^k x_{i'j} \le W^k}$ takes value 0 or 1 depending on the assignment of categories to aisles. Letting y_j denote the decision to *activate* aisle j, we can rewrite problem (4) into

$$\max_{x_{ij}=0,1; y_j=0,1} \sum_{j\in J}\sum_{i\in I} p_i x_{ij} \qquad (5)$$
$$\text{s.t.} \sum_{j\in J} x_{ij} \le 1 \text{ for } i\in I,$$
$$\sum_{i\in I} w_i x_{ij} \le W y_j \text{ for } j\in J,$$
$$x_{ij} \le y_j \text{ for } i\in I, j\in J,$$

where $p_i = r_i\theta_i$. Here, we have introduced the binary variable y_j that can be set to $y_j = 1$ whenever the constraint $\sum_{i\in I} w_i x_{ij} \le W$ is satisfied; in contrast, when $y_j = 0$ then no category is included in the aisle since $x_{ij} \le y_j$. Problem (5) resembles the multiple knapsack problem, with the added challenge that the retailer also determines how many knapsacks to use. It is well known that even the single knapsack problem is a NP-hard problem (Martello and Toth 1990).

Despite the generic hardness of this problem, we can establish some structural properties of the optimal layout (i.e., optimal number of knapsacks to open), under different conditions. Lemma 1 specifies the optimal number of aisles to use, based on consumers' fixed costs of aisle entry.

LEMMA 1. *(i) When $W \le 0$, the optimal store layout has a single aisle.*
(ii) Otherwise, when $W > 0$, there is an optimal solution with at least $\left|\left\{ i | 0 \le w_i \le W \right\}\right|$ *aisles.*

The lemma provides guidance about the optimal structure of the store. We are now ready to characterize the related question of optimal aisle composition.

Case: $W < 0$**.** When shoppers face a cost to enter an aisle ($W < 0$), it is optimal to bundle all categories into a single aisle. It follows immediately that the problem can be simplified to a standard single Knapsack solution and equation (5) simplifies to

$$\max_{x_i=0,1} \sum_{i \in I} p_i x_i$$
$$s.t. \sum_{i \in I} w_i x_i \le W. \tag{6}$$

Two distinct types of product categories emerge: *anchor categories* ($w_i < 0$), which draw shoppers into the aisle, and which will be included in the optimal solution; and *costly categories* ($w_i > 0$), which incur an inspection cost higher than the utility they provide. The optimal solution will thus include all categories with negative w_i, together with the most profitable combination of categories with positive w_i. In this case, FPTAS algorithms have been well known for a long time (Ibarra and Kim 1975). Clearly in the special case when $p_i = p$, one can obtain exact solutions in $O(n)$ times by simply sorting w_i in increasing order.

Case: $W > 0$**.** Conversely, when shoppers derive utility from browsing new aisles ($W > 0$), costly categories are further divided into *exploration categories* ($0 \le w_i \le W$), which are sufficiently attractive to be inspected, and *impulse categories* ($w_i > W$), which encourage spontaneous purchases but are insufficiently attractive to be browsed on their own. Lemma 1 states that *exploration categories* should be placed in separate aisles, as they are engaging enough to warrant inspection. The remaining problem then is to decide how to allocate $w_i < 0$ and $w_i > W$. This problem has been studied with the restriction that all costs are positive (Kellerer et al. 2004). In our case, costs may be negative, which means that existing techniques need to be adapted for negative weights. As noted in Chapter 6 page 158 of Martello and Toth (1990), "There is no easy way, instead, of transforming an instance so as to handle negative weights." In fact, the decision problem has two conceptually different challenges. First, it needs to distribute items with $w_i < 0$ into different aisles (which are always included because they lift revenues and provide more capacity to insert other items with $w_i > 0$), so as to expand aisle capacity from W to a higher value. Second, it must choose which items with $w_i > 0$ to include, in which aisle. The second step is relatively easy, as PTAS exist (Chekuri and Khanna 2005) even for the general problem with arbitrary p_i. In contrast, the first step is difficult, and techniques like shifting (see Chekuri and Khanna 2005 for details) do not work; one can only imagine approximation schemes with bounded performance guarantees.

The latter also turns out to be a very hard task. Consider the following example. Consider n items with $w_i \leq 0$ and $p_i = 0$, and two items with $w_{n+1} = w_{n+2} = w > W$ and $p_{n+1} = p_{n+2} = p > 0$. If $\sum_{i=1}^{n} w_i + w \leq W$, then revenue p can be achieved, by putting all the items $\{1, \ldots, n\}$ and one of $\{n+1, n+2\}$ in one aisle. To achieve revenue $2p$ (which would be optimal), then we would need to find a partition A of $\{1, \ldots, n\}$ such that

$$\sum_{i \in A} w_i + w \leq W \text{ and } \sum_{i \notin A} w_i + w \leq W$$

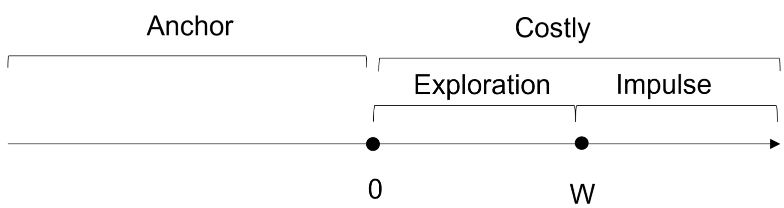
or, in other words,

$$\sum_{i \in A} |w_i| \geq w - W \text{ and } \sum_{i \notin A} |w_i| \geq w - W$$

If we set $w - W = 1/2 \left(\sum_{i=1}^{n} |w_i|\right)$, this is a partition problem, which is NP-hard. Hence, if there was a PTAS for this problem with a guaranteed performance better than 1/2, then this PTAS would be able to find a solution to the partition problem in polynomial time, which would imply P = NP. In summary, it is clear that problem (5) is hard to solve in general, and specifically it is not possible to generate approximate solutions with near-optimal performance. We end this section with some intuition.

The guiding principle that underpins the optimal aisle design is straightforward yet insightful: retailers should exploit the attractiveness of *anchor categories* ($w_i \leq 0$) to reduce the effective cost of entering the aisle, thereby increasing aisle capacity. This, in turn, enables the retailer to present *costly categories*—those where the search cost is higher than the expected value, i.e., $w_i > 0$—that would otherwise be too costly to display. These categories, while initially overlooked, can generate substantial profits once they are inspected. This mechanism mirrors the use of *loss leaders* in retail, where low-margin items attract shoppers who then make additional, higher-margin purchases. Here, the dynamic operates at the aisle level: *anchor categories* attract foot traffic into the aisle, allowing the retailer to benefit from cross-selling other products. This principle holds regardless of the sign of W, but its implications differ. When $W \leq 0$, and shoppers face a cost to enter aisles, minimizing the number of aisles is optimal, as this reduces search costs from the consumer's perspective. Conversely, when $W > 0$, and shoppers enjoy browsing, *exploration categories* ($0 \leq w_i \leq W$) should be placed in separate aisles, as they are engaging enough to warrant inspection. Importantly, they do not subsidize *impulse categories*, which should be grouped together with *anchor categories*.

Figure 4 **Category types.**



For example, categories like dairy products, typically with lower margins but high attractiveness ($w_i < 0$), should be positioned adjacent to higher-margin *impulse categories*, such as snacks ($w_i > W$). These pairings are not inherently complementary but can create synergies through co-location, encouraging unplanned purchases. In contrast, when shoppers enjoy visiting aisles, exploration categories such as craft beer or international foods ($0 \leq w_i \leq W$), are best placed in separate aisles to increase the likelihood of consideration. Aisle layout decisions play a critical role in shaping a shopper's consideration set. By strategically grouping categories, retailers influence whether shoppers inspect all the categories in an aisle or bypass them altogether.

5. Heterogeneous Consumers

The analysis so far shows that, with homogeneous shoppers, we can transform the probabilistic quantity $P(\sum_{i \in A_j} w_i^k \leq W^k)$ into a constraint in the assignment problem (5). The resulting problem is an integer program that is hard to solve or approximate, but is still amenable to solving with standard optimization solvers. With heterogeneous shoppers, there is an added difficulty: the transformation of probability into constraint is no longer valid. In this section, we provide solutions for this more general case.

5.1. Discrete number of consumer classes

When shoppers fall into well-defined types or classes, it is possible to extend the transformation of §4. Specifically, when we have a finite number of classes $k \in K$, with $|K| < \infty$, we can introduce auxiliary binary variables y_{jk} that capture whether class k chooses to enter aisle j or not, and x_{ijk} that denote whether class k inspects product i within aisle j. If $y_{jk}=0$, then we must have that $x_{ijk}\le y_{jk}=0$. On the other hand, if $y_{jk}=1$, we have that $x_{ijk}-x_{ij}\ge -1+y_{jk}=0$ and $x_{ijk}-x_{ij}\le 0$, hence $x_{ijk}=x_{ij}$. Finally, a necessary condition for visiting aisle j is that $\sum_{i\in I} w_i^k x_{ij}\le W^k$.

These auxiliary variables allow us to cast problem (4) into an integer program with $O(|I|\cdot|J|\cdot|K|)$ variables, as follows:

$$
\begin{aligned}
\max_{x_{ij},y_{jk},x_{ijk}} \quad & \sum_{j\in J}\sum_{i\in I}\sum_{k\in K} r_i\theta_i^k x_{ijk} \\
\text{s.t.} \quad & \sum_{j\in J} x_{ij}\le 1 \text{ for } i\in I, \\
& \sum_{i\in I} w_i^k x_{ijk}\le W^k y_{jk} \text{ for } j\in J, k\in K, \\
& x_{ijk}\le y_{jk} \text{ for } i\in I, j\in J, k\in K, \\
& x_{ijk}\le x_{ij} \text{ for } i\in I, j\in J, k\in K, \\
& x_{ijk}\ge x_{ij}+y_{jk}-1 \text{ for } i\in I, j\in J, k\in K, \\
& x_{ij}, y_{jk}, x_{ijk}\in\{0,1\} \text{ for } i\in I, j\in J, k\in K.
\end{aligned}
\tag{7}
$$

This formulation is able to numerically solve the category-aisle allocation problem with standard solvers. For practical purposes, it implies that we can easily solve the problem even with multiple shopper classes. There are two limitations though. First, we do not have structural results anymore that characterize the optimal solution – except when $W^k\le 0$ for all k, in which case a single aisle will of course be optimal. Second, the problem will scale poorly with the size of K. In particular, when shoppers are distributed over continuous types, this approach will require an infinite number of decision variables, thereby invalidating the idea. This case questions whether we can develop a tractable approach with continuous types. This is possible under certain technical conditions on the distribution of w_i^k and W^k, as shown next.

5.2. Continuous consumer types

When shopper types are continuous, formulating Equation (4) in a tractable way is possible depending on the dimension on which they differ.

There are two main sources of heterogeneity, namely product preferences (u_i^k) and search costs (c_i^k). Variability in product preferences will affect both θ_i^k (conversion contingent on aisle visit) and

w_i^k (budget used in aisle visit), while variability in search costs will affect w_i^k and W^k (total budget available). Let us consider the effects of heterogeneity on θ_i^k, w_i^k and W^k separately.

Heterogeneity in conversion rate θ_i^k. When θ_i^k is a random variable, but $w_i^k = w_i$ and $W^k = W$, then the problem turns out to be simple. One can just use $\theta_i = E[\theta_i^k]$ and solve the problem for homogeneous shoppers with this average parameter. In other words, the feasible solution space is unchanged with this type of uncertainty, and hence it is sufficient to solve (4) with the average objective value.

Heterogeneity in total budget available W^k. Let $\theta_i^k = \theta_i$ and $w_i^k = w_i$, i.e., the ex-ante utility and search cost of each category are deterministic and common to all. Consider shopper heterogeneity in terms of total budget W^k. Then the problem can be expressed as follows. We let $\bar{H}$ be the survival function, (i.e., one minus the c.d.f. of W^k), and introduce y_{ij} be the probability that a shopper is exposed to i – as opposed to a binary variable as in §4. We then have

$$\begin{aligned}
\max_{x_{ij}, y_{ij}} & \sum_{j \in J} \sum_{i \in I} r_i \theta_i y_{ij} && (8)\\
\text{s.t.} & \sum_{j \in J} x_{ij} \leq 1 \text{ for } i \in I, \\
& y_{ij} \leq x_{ij} \text{ for } i \in I, j \in J, \\
& y_{ij} \leq \bar{H}\left(\sum_{i' \in I} w_{i'} x_{i'j} \right) \text{ for } i \in I, j \in J, \\
& x_{ij} \in \{0,1\}, 0 \leq y_{ij} \leq 1 \text{ for } i \in I, j \in J.
\end{aligned}$$

In this mixed integer program (MIP) formulation, y_{ij} will be maximized since $r_i \theta_i \geq 0$. Hence it will take value $x_{ij} \times \bar{H}\left(\sum_{i' \in I} w_{i'} x_{i'j}\right)$, that is, $Pr[\sum_{i' \in I} w_{i'} x_{i'j} \leq W]$ when $x_{ij} = 1$ and zero otherwise. Note that the MIP now includes a nonlinear constraint $\bar{H}\left(\sum_{i' \in I} w_{i'} x_{i'j}\right) - y_{ij} \geq 0$; under certain circumstances it may be easy to incorporate (Kortanek and Evans 1967, Hillestad and Jacobsen 1980).

Besides the advantage of writing the problem in a compact way, this formulation also allows us to obtain some insight on the optimal layouts, albeit with limited scope. The optimal structure – optimal number of aisles – depends on the specific form of $\bar{H}$, which is a non-increasing positive function, but could be convex or concave, e.g., it is S-shaped (convex then concave) when the distribution of W^k is unimodal.

Specifically, it is optimal to use a single aisle when one can prove that merging two separate aisles into one provably increases the aisle visit probability. This condition is met when, for any sets A_1 and A_2, letting $W_1 = \sum_{i\in A_1} w_i$ and $W_2 = \sum_{i\in A_2} w_i$, we have $\bar{H}(W_1 + W_2) \geq \max\left\{\bar{H}(W_1), \bar{H}(W_2)\right\}$. This is never the case when $w_i > 0$ (in which case it is best to keep each item in a separate aisle), but it will always hold when $w_i \leq 0$ for all i (in which case all items should be included in a single aisle). When some items have $w_i \leq 0$ and some others have $w_i > 0$, and similar revenues per category, then multiple aisles may be optimal if these aisles consume a similar budget $\sum_{i\in A_j} w_i$; if this was not the case, then it would be possible to move some items from higher budget aisles into lower budget aisles and increase revenue.

Heterogeneity in budget consumed by category w_i^k. Finally, let $\theta_i^k = \theta_i$ and $W^k = W$, i.e., the ex-ante utility and the total budget per aisle are deterministic and common to all. Consider shopper heterogeneity in the budget used to inspect category i within an aisle, w_i^k. Moreover, let us assume that the I-dimensional vector $\{w_i^k\}$ is normally distributed with average equal to the I-dimensional vector μ, and variance-covariance matrix $\mathbf{\Sigma}$. In this case, the probability $P[\sum_{i\in I} w_i x_{ij} \leq W] = \Phi\Big((W - \mu_j)/\sigma_j\Big)$ with $\mu_j = \mu' \cdot x_{\cdot j}$, $\sigma_j^2 = x'_{\cdot j} \cdot \mathbf{\Sigma} \cdot x_{\cdot j}$, and Φ the standard Normal c.d.f. As a result, we can use a formulation similar to (8):

$$\begin{aligned}
&\max_{x_{ij}, y_{ij}} \sum_{j\in J} \sum_{i\in I} r_i \theta_i y_{ij} \qquad (9)\\
&\text{s.t.} \sum_{j\in J} x_{ij} \leq 1 \text{ for } i \in I,\\
&y_{ij} \leq x_{ij} \text{ for } i \in I, j \in J,\\
&y_{ij} \leq \Phi\left(\frac{W - \sum_{i'\in I} \mu_{i'} x_{i'j}}{\sqrt{\sum_{i',i''\in I} \Sigma_{i'i''} x_{i'j} x_{i''j}}}\right) \text{ for } i \in I, j \in J,\\
&x_{ij} \in \{0,1\}, 0 \leq y_{ij} \leq 1 \text{ for } i \in I, j \in J.
\end{aligned}$$

Problem (9), in comparison with (8), includes a complex nonlinear constraint. When adding a new category within an aisle j, we not only modify the expected budget consumed $W - \sum_{i'\in I} \mu_{i'} x_{i'j}$, but we also change its variance. This is a new effect that was not present when heterogeneity was contained to W^k. Indeed, one may now want to have larger aisles, so they are less variable for the shoppers. In a way, larger aisles may exploit the natural variability in shopper's search costs, and bundling items with high c_i^k and low c_i^k may result in making shoppers explore more aisles within the store, everything else being equal.

6. Numerical Experiments

Through the various formulations, we have demonstrated that the general problem (4) can be expressed in compact forms amenable to standard optimization packages. Unfortunately, since the problem remains hard to solve (cf. §4), it is not possible to extract insights about the optimal layouts from these formulations, except from some selected, simple cases. Having said this, we can still obtain insights from numerical experiments, as shown in this section.

We are interested in understanding how different aspects of the shopping process affect the optimal layout structure. When shoppers are homogeneous, we have established that when the fixed costs of search are positive ($W \leq 0$), race-track structures are optimal, while open-floor structures are optimal when search provides intrinsic benefits ($W > 0$). But even in the homogeneous case, the optimal category mix is not tractable analytically. Specifically, one may be interested in the characteristics of the different optimal aisles. Are they similar in attractiveness and margin? Or, in contrast, some of them are high footfall, low profit and some others low footfall, high profit? Furthermore, it is important to extend these insights to contexts in which shopper heterogeneity is significant.

To show this, we first provide an extensive numerical study, where we vary the average differences across categories (high vs. low cost w_i^k, and high vs. low conversion θ_i^k), as well as the heterogeneity of these categories across shoppers, with synthetically generated data. We then present a real case study based on two hypermarkets, in which we illustrate how model parameters can be estimated from data and then fed into our optimization framework; we discuss what the optimal aisle structures are in this real context.

6.1. Synthetic Instances

Instances with homogeneous shoppers. We begin our numerical study by focusing on the homogeneous shopper case analyzed in Section 4. Despite the NP-hard nature of the problem, we demonstrate that it can be solved within a reasonable time using commercial solvers. Motivated by store layout designs in supermarkets, which typically carry between 40–50 categories and 200–400 subcategories, we conduct our analysis with $|I| = 100$ categories and up to $|J| = 100$ aisles. We focus on the case where $W > 0$, setting $W = 2$ without loss of generality. The parameters for each category are generated as follows:

- *Revenues:* For each category, revenues r_i are randomly drawn from a uniform distribution $U[0.1, 10]$.

- *Purchase likelihood:* The purchase likelihood θ_i is also randomly drawn from $U[0,1]$, such that $p_i = r_i \cdot \theta_i \in [0,10]$.
- *Attractiveness w_i:* We consider cases where w_i are either negative ($w_i \in [-5,0]$) or greater than $W = 2$ ($w_i \in [2,5]$). Half of the categories fall within each of these two distributions. This ensures, with probability nearly one, that a single aisle cannot accommodate all categories.

The mixed-integer linear program (MILP) corresponding to Equation (5) is implemented using the R Optimization Infrastructure (ROI) with Gurobi as the solver. All experiments were executed on a standard laptop, equipped with an Intel Core i7 CPU and 32GB of RAM, the limits are set to 2% of optimality or 3000 seconds run time (the latter limit was never reached for the deterministic case).

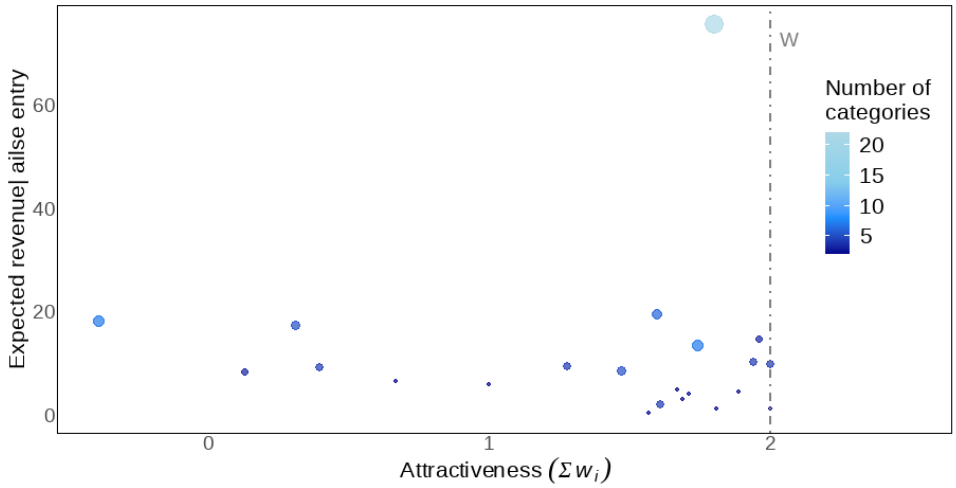
The resulting optimal store layout for one particular instance is presented in Figure 5. In this instance, the optimal layout consists of 22 aisles displaying 96 out of the 100 categories. The number of categories per aisle ranges from 2 to 22; there is a single such aisle with 22 categories included, with an attractiveness of 1.8 and the highest revenue conditional on visiting at 75.4. Among these 22 aisles, total weight $\sum_{i \in A_j} w_i$ values vary between -0.39 and 2.00 (recall that lower $\sum_{i \in A_j} w_i$ indicates higher attractiveness of aisle j), and expected conditional revenues per aisle in the 0.4-19.4 range. Each aisle typically contains a combination of anchor categories with at least one and at most 11 costly categories; for instance, the large aisle made of 22 categories carries 11 anchors and 11 costly categories. Furthermore, the four excluded categories have either a low conditional revenue p_i or a very high net cost w_i, which hinders shoppers from visiting the aisle if it had been included. For instance, the second most expensive category ($w_i = 4.92$) has been included, while the third most expensive ($w_i = 4.85$) is excluded, because the former category provides higher revenue, with 0.72 versus 0.68, respectively.

Instances with heterogeneous shoppers. We focus on the case where shoppers are *heterogeneous in total budget available*, as expressed in Equation (8). To address the challenge posed by the non-linear constraint

$$y_{ij} \leq \bar{H}\left(\sum_{i' \in I} w_{i'} x_{i'j}\right),$$

we reformulate the problem using a *piecewise linear approximation* of the survival function $\bar{H}(\cdot)$, following the convex-combination method as outlined in Geißler et al. (2011). To this end, we approximate $\bar{H}(\cdot)$ with a piecewise linear function $\tilde{H}(\cdot)$. The approximation is defined over a set of breakpoints z_t, where $t \in \{1, \ldots, T\}$ represents segment indices, and each breakpoint z_t

Figure 5 **Optimal aisle composition in one selected instance.**



corresponds to a specific value of the survival function $\bar{H}(z_t)$. We introduce auxiliary continuous variables $\lambda_{t,j} \in [0,1]$, which represent the weights in a convex combination of the survival function values $\bar{H}(z_t)$, that is $\tilde{H}(\cdot) = \sum_{t=1}^{T} \lambda_{t,j} \bar{H}(z_t)$. Binary variables $b_{t,j} \in \{0,1\}$ are introduced to ensure that at most two adjacent segments are active for a given x. These binary variables interact with $\lambda_{t,j}$ to enforce adjacency through constraints described below. The reformulated linear constraints are:

$$y_{ij} \leq \sum_{t=1}^{T} \lambda_{t,j} \bar{H}(z_t), \quad i \in I, j \in J,$$

$$\sum_{t=1}^{T} \lambda_{t,j} z_t = \sum_{i \in I} w_i x_{ij}, \quad j \in J,$$

$$\sum_{t=1}^{T} \lambda_{t,j} = 1, \quad j \in J,$$

$$\lambda_{t,j} \leq b_{t,j} + b_{t+1,j}, \quad t = 2, \ldots, T-1, \ j \in J,$$

$$\lambda_{1,j} \leq b_{1,j}, \quad j \in J,$$

$$\lambda_{T,j} \leq b_{T,j}, \quad j \in J,$$

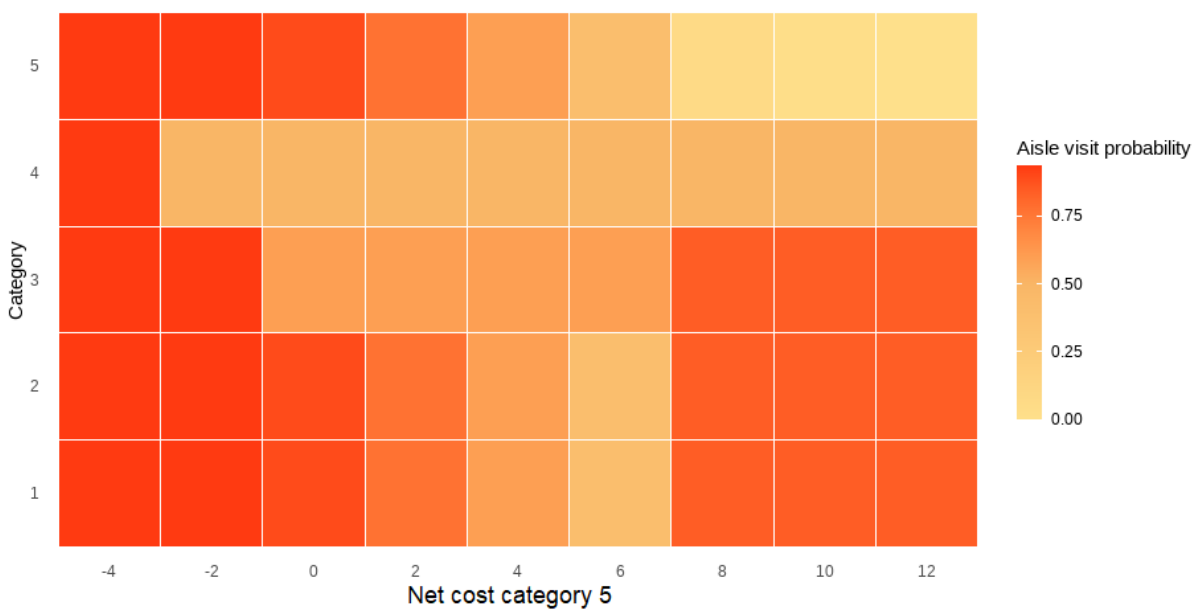
$$\sum_{t=1}^{T} b_{t,j} = 1, \quad j \in J,$$

This reformulation replaces the original non-linear constraint with a set of linear constraints, enabling the problem to remain solvable as a Mixed-Integer Linear Program (MILP).

To build intuition, we first examine small instances with $|I| = 5$ categories. Specifically, we set $r_i = 1$, $\theta_i = 1$, and $w_i = \{-2, -1, 1, 2\}$ for $i = 1, \ldots, 4$, while setting $r_5 = 5$ and allowing w_5 to vary with $\theta_5 = 1$. Additionally, we assume that $W \sim \mathcal{N}(2, 4)$. With these parameters, category 5 is the most profitable of all, but it can be quite expensive to inspect when w_5 is high.

Figure 6 provides a heat-map visualization of the optimal aisle composition under varying levels of category attractiveness w_5. The heat-map intensity (color gradient) represents the decision variable y_{ij}, which captures the probability of entering an aisle. Note that identical colors indicate that categories are assigned to the same aisle.

Figure 6 Aisle composition when varying w_5, the net cost of category 5.



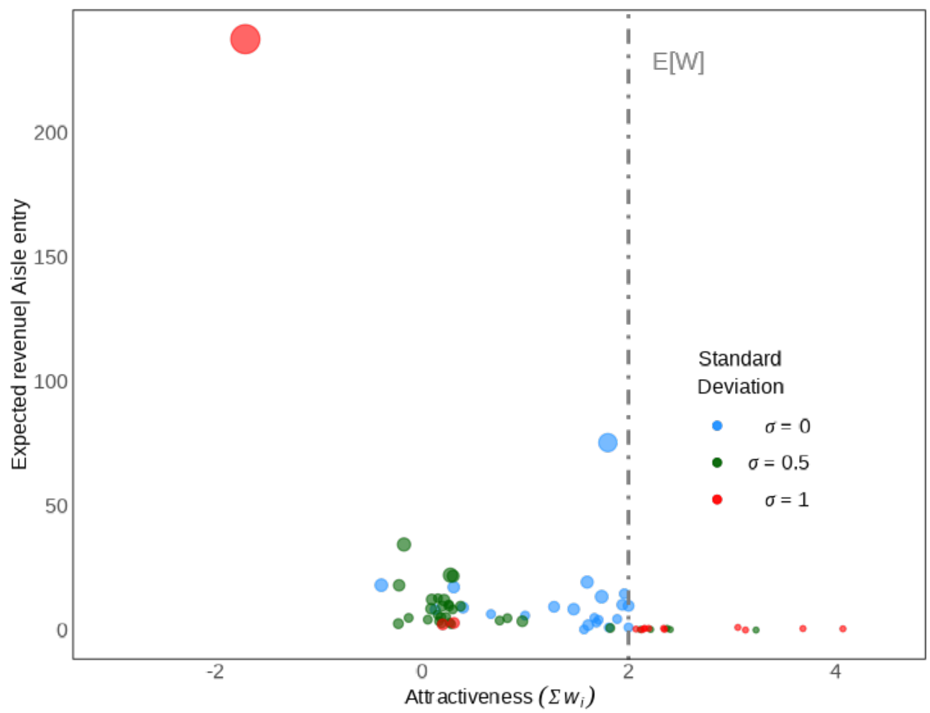
The figure shows that when $w_5 \leq -4$, the optimal layout consolidates all categories into a single aisle. As w_5 increases, categories that negatively impact traffic are put into separate aisles, e.g., category 4 which has the highest w_i, and eventually category 5 is also removed from the common aisle. Note that when $w_5 \geq 8$, this category, placed in its own aisle, is so expensive to inspect that it receives a nearly 0 probability of aisle entry.

Our next set of experiments examines the impact of the standard deviation of W on the optimal store layout. Specifically, for the instance with $|I| = 100$ categories discussed earlier, we now analyze the optimal layout when W is random rather than deterministic. To ensure comparability, we set the mean of W to 2 and vary the standard deviation $\sigma = \{0, 0.5, 1\}$, assuming a normal distribution. The resulting aisle compositions are shown in Figure 7.

In the deterministic case ($\sigma = 0$, blue), the maximum number of categories grouped into a single aisle is 22 as discussed before. However, as the standard deviation of the total budget increases, this number rises to 72 for $\sigma = 1$, respectively. Interestingly, this behavior is not monotone as for $\sigma = 0.5$ the maximum number of categories in an aisle shrinks to 10.

In the deterministic case, aisle attractiveness ($\sum w_i$) was capped at $W = 2$, limiting the number of products assigned. In contrast, the stochastic setting allows higher cumulative attractiveness due to randomness in the budget, enabling all products to be assigned. The number of single-category aisles also rises—from 0 in the deterministic case to 16 in the stochastic scenarios. As uncertainty in W increases, high-revenue aisles are protected by assigning them conservative attractiveness values (e.g., -1.62 when $W = 1$), ensuring near-certain visitation.

Figure 7 Aisle composition for different standard deviations.



We extend the computational exercise to 10 instances for each setting. Table 2 summarizes the resulting aisle compositions. As the standard deviation of W increases, we observe a non-monotonic pattern: the number of aisles, as well as the number of single-category aisles, generally increases. In some instances, extremely large aisles emerge—e.g., with over 70 categories—serving as ”race-tracks” visited by nearly all shoppers. At the same time, more single-category aisles appear, isolating categories with high attractiveness (w_i) to avoid degrading the performance of others.

Table 2 Summary of aisle composition across 10 instances. The reported metrics are the total number of aisles, the number of single-category aisles, the number of aisles with 2 to 10 categories, and the number of categories in the largest aisle.

Instance	Deterministic $\sigma = 0$				$\sigma = 0.5$				$\sigma = 1$				$\sigma = 1.5$			
	Total	**Single**	**2-10**	**Largest**	**Total**	**Single**	**2-10**	**Largest**	**Total**	**Single**	**2-10**	**Largest**	**Total**	**Single**	**2-10**	**Largest**
1	22	0	21	22	32	5	27	10	14	11	2	77	18	16	0	72
2	23	2	20	34	27	5	21	16	24	10	13	21	34	21	11	16
3	28	1	26	32	9	2	4	28	14	5	7	49	19	15	3	70
4	23	1	21	33	26	6	18	13	21	11	7	22	24	17	4	27
5	29	2	26	16	33	7	26	9	22	15	4	33	19	14	3	54
6	20	0	18	21	32	5	27	9	24	15	6	33	26	20	3	36
7	26	0	24	24	35	4	31	6	20	12	4	25	22	16	4	49
8	23	0	22	23	29	4	25	9	22	15	3	34	24	17	4	34
9	23	0	22	25	19	1	17	18	19	10	7	35	21	16	3	42
10	23	0	22	24	16	1	14	11	22	13	7	19	30	21	6	15

6.2. Case study on grocery retail

While the numerical study on synthetic data provides valuable insights into the optimal store layout under known input parameters, we now turn to the practical challenge of estimating these parameters in real-world settings. This empirical study does not aim to establish causal claims about shopper behavior. Instead, it serves to demonstrate how the model can be applied in practice. A fully causal application would require exogenous variation in aisle composition—such as experimental shocks—which is not available in this case.

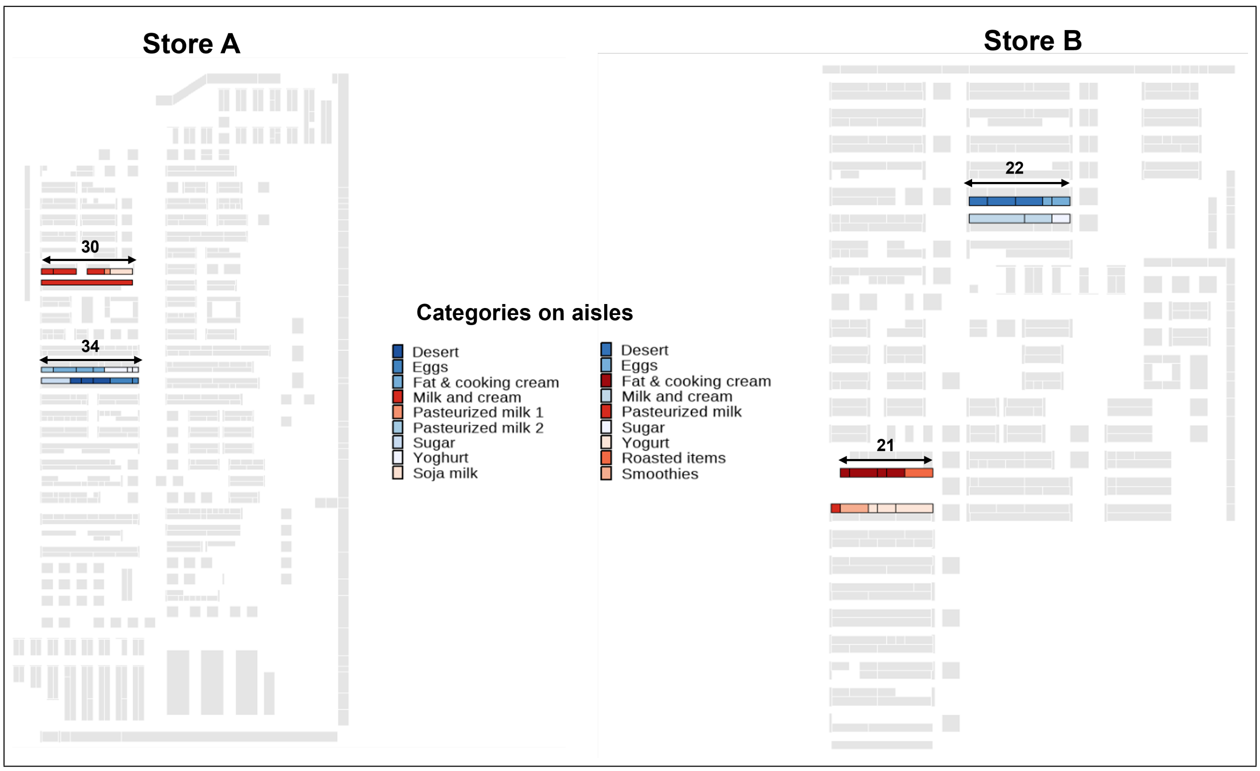
We partnered with a European grocery retailer that provided two types of data: (i) layouts for two of its largest stores, and (ii) scanner data capturing all shopper transactions in 2022. These two stores are hypermarkets of around 10,000 square meters, in the suburbs of the same large city. They are currently organized with open spaces, although with aisles that are quite long, and contain multiple categories. Our model helps assess whether aisles are too short—warranting a race-track redesign—or too long—creating excessive search costs that deter shoppers.

Ideally, empirically estimating our model would require data capturing the paths shoppers take through the store, from entry to checkout. However, such detailed data is challenging to obtain, with rare exceptions using radio-frequency identification tags on shopping carts (Hui et al. 2009b,c, Seiler and Pinna 2017, Seiler and Yao 2017) or video-tracking technologies (Jain et al. 2020). As our partner company does not collect this level of detail, the following analysis demonstrates the estimation process based on the available data, in the spirit of Ziaei et al. (2023). Although these estimates may be less precise than those derived from richer datasets, this exercise illustrates how our model can be broadly applied to empirically estimate unknown parameters in real-world settings.

To capture spatial variation, we compare aisle compositions across the two stores. Both place *Dessert, Eggs, and Sugar* in the same aisle. Store A adds *Fat Cooking Cream, Yogurt, and Pasteurized Milk* there, while these are separate in Store B. Store B instead co-places *Milk Cream* with the former trio, unlike Store A. *Soy Milk* appears in one of Store A's aisles; *Roasted Items and Smoothies* appear in Store B's. Aisles are longer in Store A (34m, 30m) than in Store B (22m, 21m).

Figure 8 details the aisle composition in each store. These differences in co-placement and aisle length provide the variation needed to identify the model. These differences in co-placement and aisle length provide the variation provide the variation necessary for identification.

Figure 8 Store layout in two hypermarkets.



Estimation procedure for purchase decisions. For the estimation procedure, we sampled approximately 100.000 unique transactions from shoppers who entered either one of the four target aisles or their surrounding areas. To estimate shoppers' purchase decisions conditional on aisle entry, as specified in Equation (1), we construct a panel dataset at the category-shopper level. For each category within the aisle, we record whether a shopper who entered the aisle purchased the category. To estimate $\theta_{i,t,s}$, we allow it to vary by day of the week t and store s and assume that the random shock $\epsilon_{i,t,s}$ follows a logistic distribution, which is equivalent to assuming a Gumbel shock that is compared against another Gumbel shock representing an outside option (Wagner and Martínez-de Albéniz 2020). This leads to the following expression:

$$\theta_{i,t,s} = \frac{\exp(u_{i,t,s})}{1+\exp(u_{i,t,s})},$$

where the latent utility $u_{i,t}$ is modeled as

$$u_{i,t,s} = \mathrm{FE}_i + \mathrm{FE}_{\text{day of week}_t} + \mathrm{FE}_s.$$

In this specification, FE_i represent category-specific fixed effects, while $\mathrm{FE}_{\text{day of week}_t}$ and FE_s capture time and store fixed effects. The results from the second stage purchase decision is shown in Table 3.

Table 3 Estimation of purchase model parameters.

	P(Purchase)
(Intercept)	-1.33^{***}
	(0.01)
Store A	0.12^{***}
	(0.01)
Eggs	0.94^{***}
	(0.01)
Fat & cooking cream	0.78^{***}
	(0.01)
Milk and cream	1.82^{***}
	(0.01)
Pasteurized milk	-0.90^{***}
	(0.01)
Roasted items	-1.35^{***}
	(0.03)
Soja milk	0.21^{***}
	(0.02)
Sugar	1.76^{***}
	(0.01)
Smoothies	-8.16^{***}
	(0.71)
Yogurt	2.37^{***}
	(0.01)
Monday	-0.12^{***}
	(0.01)
Tuesday	-0.10^{***}
	(0.01)
Wednesday	-0.10^{***}
	(0.01)
Thursday	-0.07^{***}
	(0.01)
Saturday	0.02^{*}
	(0.01)
Sunday	-0.02^{*}
	(0.01)
Num. obs.	$602,694$
Deviance	640345.41
Log Likelihood	-320172.70
Pseudo R^2	0.21

$^{***}p<0.001$; $^{**}p<0.01$; $^{*}p<0.05$

This approach reveals strong heterogeneity in purchase propensity across categories. For example, the predicted probability of purchasing *Yoghurt* upon entering the aisle in Store A on a Saturday is:

$$P(\text{Purchase}_{\text{Yogurt}}) = \frac{1}{1+e^{-(-1.33+0.12+0.02+2.37)}} \approx 76.5\%$$

In contrast, the predicted probability of purchasing *Pasteurized milk* in the reference store (Store B) on a Monday is:

$$P(\text{Purchase}_{\text{Smoothies}}) = \frac{1}{1+e^{-(-1.33-0.12-0.9)}} \approx 8.71\%.$$

This reflects nearly a ninefold difference in baseline purchase probabilities. Notably, the likelihood of purchasing *Smoothies* in Store B is close to zero, underscoring the substantial variation in conversion rates across categories. Temporal variation is less pronounced. Conversion increases significantly on Saturdays (coefficient 0.02, p-value < 0.05), while purchase probabilities on other weekdays are lower than on the reference day, Friday.

Estimation procedure for aisle entry decisions. Using the results from the second-stage regression, we can obtain utility estimates $u_{i,t,s}$ for each observation, and can infer the ex ante utility provided by category i, i.e., $v_{i,t,s} = E[\max(u_{i,t,s} + \epsilon_{i,t,s}, 0)]$, which will be compared to the cost of doing so $\sum_i c_i$, to determine the decision to enter the aisle. In the absence of detailed shopper path data, we assume c_i is proportional to the shelf area occupied by category i, making c_i deterministic. Alternative formulations of c_i could also be considered, depending on the data and context. Using total searchable shelf length as a proxy for shopper search cost presents challenges, because it may induce multicollinearity. Fortunately, the variables $\sum_i c_i$, $\sum_i v_{i,t,s}$, and FE_s (the fixed effect one for store A) exhibit low to moderate multicollinearity, with VIFs of 1.78, 1.08, and 1.89, respectively. These levels are well below conventional concern thresholds.

Thus, we next estimate the probability of aisle visits by modeling shopper variability driven by heterogeneity in available budget W_k. The goal is to estimate the aisle visit probability from the parameters, i.e.,

$$P\left(W_k \geq \sum_i w_i\right) = \phi\left(\alpha_0 + \alpha_1 \sum_i c_i + \alpha_2 \sum_i v_{i,t,s} + FE_s\right). \tag{10}$$

The resulting model is a probit model in line with the assumptions presented for the synthetic instances. Note that it is possible however to use any other specification, and then generate a counterfactual prediction of aisle visit probability when we alter aisle composition.

Table 4 Estimation of aisle visit model parameters.

	P(Aisle entry)
(Intercept)	-0.742^{***}
	(0.031)
Store A	-0.134^{***}
	(0.012)
Search cost ($\sum c_i$)	-0.001
	(0.001)
Expected utility ($v_{i,t}$)	0.598^{***}
	(0.006)
Num. obs.	$100,000$
AIC	111446.386
BIC	111484.438
Log Likelihood	-55719.193
Deviance	111438.386

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

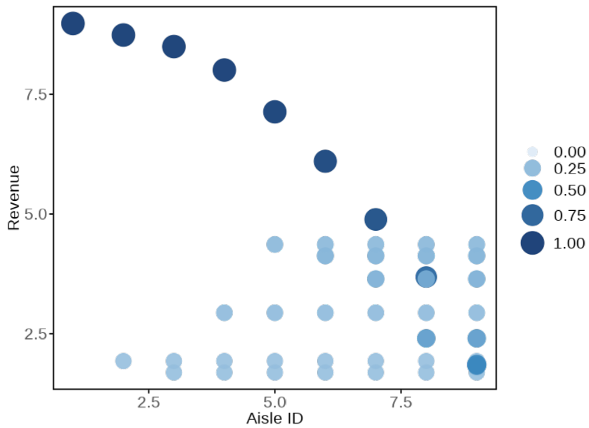
The estimation is based on binary visit decisions from 100,000 shoppers—those who entered one of the two stores and purchased categories in thee focal aisles or its adjacent aisles. Results are shown in Table 4. We find that expected utility within the aisle positively influences entry (coefficient = 0.598, p-value < 0.001), while search cost has a small negative effect (coefficient = -0.001, p-value ≥ 0.05), albeit non significant. Finally, the average value of W for stores A and B is negative, implying that $c_0^k > 0$ on average. This suggests that shoppers generally dislike exploring, consistent with economies of scale in search behavior.

Reengineering the layout. Building on the calibrated estimates in Table 4, we apply our prescriptive layout optimization model to the existing configurations. With nine categories in each store, this yields 3,282 feasible allocations, respectively. These include race-track designs with a single aisle, fully disaggregated layouts with single-category aisles (8 or 9 of them, depending on the store, and all hybrid combinations in between.

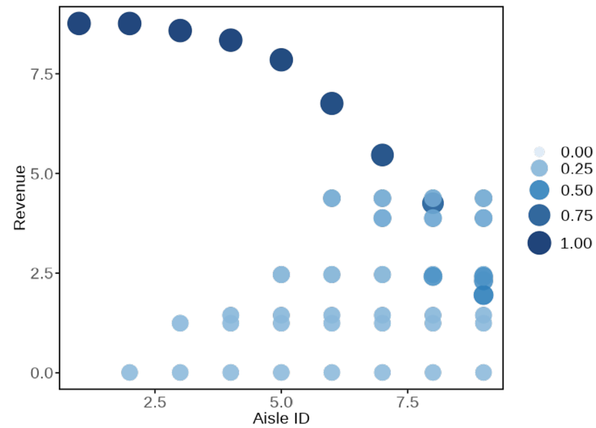
Figure 9 shows the highest expected revenue and profit outcomes by varying the total number of aisles in Stores A and B (variable in the x-axis). Panels (a) and (b) report expected revenue conditional on aisle entry for the best assignment with a certain total number of aisles. Bubble size and color indicate the entry probability. In both stores, the race-track layout yields high conditional revenue (€8.98 and €8.76, respectively) and a high visit probability. Splitting aisles generally reduces conditional revenue per aisle but may increase the likelihood of entry.

Panels (c) and (d) display expected total profit across layouts. In Store A, the race-track design dominates, while the fully disaggregated layout with nine aisles underperforms. In Store B, the race-track performs well but is outperformed by a two-aisle configuration where Smoothies are isolated and all other categories are grouped. This result is intuitive: search costs are modest,

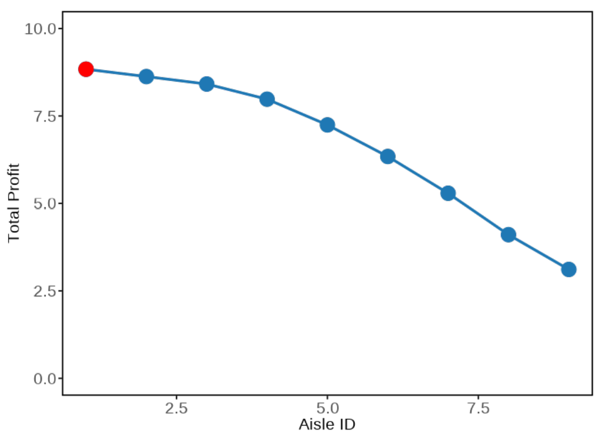
Figure 9 Revenue and profit per aisle layout for Stores A and B, for different total number of aisles.



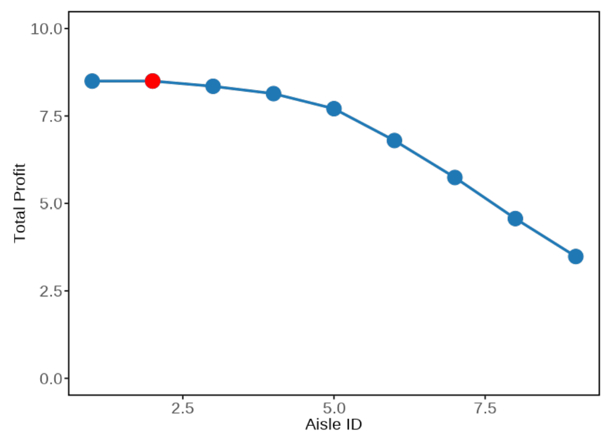
(a) Store A – Expected revenue per aisle



(b) Store B – Expected revenue per aisle



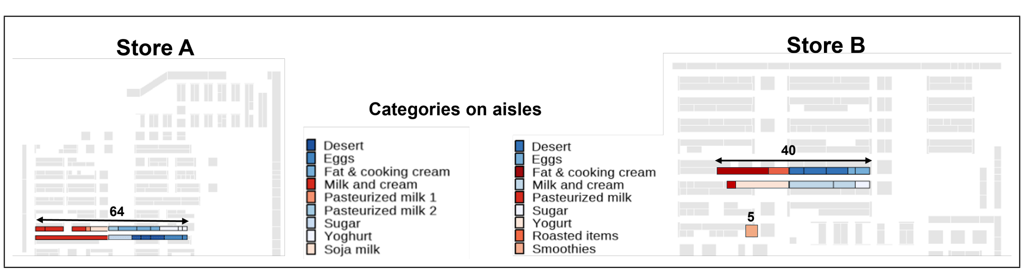
(c) Store A – Expected profit



(d) Store B – Expected profit

which favors longer aisles, and shoppers avoid navigating multiple aisles. Yet, separating Smoothies increases the chance of visiting the remaining aisle, which contains categories with high conditional purchase probabilities. For completeness, the optimal counterfactual layout is shown in Figure 10.

Figure 10 Counterfactual: optimized store layouts.



In summary, we see that, in practice, we should expect model parameters to vary across stores. The direct consequence of this finding is that the optimal store structure should also depend on local store characteristics, and hence we should not expect one store design to dominate across all conditions.

7. Conclusions and Future Research

Store layout design heavily affects how shoppers navigate a store and inspect products. This paper has developed a framework for the design of effective layouts. Our model internalizes how shoppers choose to enter aisles, based on the comparison of the expected utility derived from them vs. the cost of search. We formulate the decision as an assignment of categories into aisles, and show that it is generally a hard problem. We recognize the shape of the search cost as the key driver of optimal layouts. In the homogeneous shopper case, when there are diseconomies of scale in search ($W \leq 0$), then it is best to adopt a race-track design, while when visit entry provides a natural benefit ($W > 0$), then open spaces are preferable. In general, it is optimal to bundle anchor categories with impulse ones, so that the former subsidize the search cost of the latter. Our numerical experiments further reveal that when there is shopper heterogeneity, large aisles – similar to a race track – emerge as a structure that guarantees very high aisle entry probability despite uncertainty. Finally, we illustrate how our model can be applied in practice to a case study of a European retailer, in which we show that, depending on store characteristics, it may be better to favor race tracks or open spaces, depending on the average and standard deviation of W.

This study would not be complete without mentioning promising future research opportunities. First, our case study should be expanded into a full-fledged empirical investigation, with the triple objective of validating the shopper search model (including the search cost shape, i.e., sub- vs. super-modularity), of identifying the best performing store layout structures, and finally of linking optimal structure to store features, and possibly to shopper profiles in the store catchment area. Such research will need to access digitalized store layouts, which are unfortunately not readily available at most retailers in 2025. Second, the ideas behind our framework have implications that also apply in digital settings. For instance, we have shown that co-locating two categories in the same aisle creates a positive correlation between sales of those categories; it would be interesting to relate category correlation with the (digital) distance that an online shopper needs to overcome. This may expose the difference between intrinsic product complementarity and simple proximity during the shopping process. Third, since co-location induces correlation in sales, one should in theory be able to reverse engineer this link, so as to infer in-store location from transactional data, when layouts are not digitalized. These research directions are important not only for academics seeking to better understand consumer shopping experiences, but also for practitioners (retailers) in search of effective store guiding principles.

References

Adams WJ, Yellen JL (1976) Commodity bundling and the burden of monopoly. *The quarterly journal of economics* 90(3):475–498.

Aouad A, Deshmane A, Martínez-de Albéniz V (2022) Designing layouts for sequential experiences: Application to cultural institutions. *Management Science* Forthcoming.

Aouad A, Farias V, Levi R (2021) Assortment optimization under consider-then-choose choice models. *Management Science* 67(6):3368–3386.

Bakos Y, Brynjolfsson E (1999) Bundling information goods: Pricing, profits, and efficiency. *Management science* 45(12):1613–1630.

Caro F, Martínez-de-Albéniz V, Apaolaza B (2021) The value of online interactions for store execution. *Manufacturing & Service Operations Management* Forthcoming.

Chekuri C, Khanna S (2005) A polynomial time approximation scheme for the multiple knapsack problem. *SIAM Journal on Computing* 35(3):713–728.

Fang H, Norman P (2006) To bundle or not to bundle. *The RAND Journal of Economics* 37(4):946–963.

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

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

Geißler B, Martin A, Morsi A, Schewe L (2011) Using piecewise linear functions for solving minlps. *Mixed integer nonlinear programming*, 287–314 (Springer).

Hillestad RJ, Jacobsen SE (1980) Reverse convex programming. *Applied Mathematics and Optimization* 6(1):63–78.

Honhon D, Pan XA (2017) Improving profits by bundling vertically differentiated products. *Production and Operations Management* 26(8):1481–1497.

Hui SK, Bradlow ET, Fader PS (2009a) Testing behavioral hypotheses using an integrated model of grocery store shopping path and purchase behavior. *Journal of consumer research* 36(3):478–493.

Hui SK, Fader PS, Bradlow ET (2009b) Path data in marketing: An integrative framework and prospectus for model building. *Marketing Science* 28(2):320–335.

Hui SK, Fader PS, Bradlow ET (2009c) Research note—the traveling salesman goes shopping: The systematic deviations of grocery paths from tsp optimality. *Marketing science* 28(3):566–572.

Hui SK, Inman JJ, Huang Y, Suher J (2013) The effect of in-store travel distance on unplanned spending: Applications to mobile promotion strategies. *Journal of Marketing* 77(2):1–16.

Ibarra OH, Kim CE (1975) Fast approximation algorithms for the knapsack and sum of subset problems. *Journal of the ACM (JACM)* 22(4):463–468.

Jagabathula S, Mitrofanov D, Vulcano G (2023) Demand estimation under uncertain consideration sets. *Operations Research* .

Jain A, Misra S, Rudi N (2020) The effect of sales assistance on purchase decisions: An analysis using retail video data. *Quantitative Marketing and Economics* 18(3):273–303.

Kellerer H, Pferschy U, Pisinger D (2004) *Multiple Knapsack Problems*, 285–316 (Berlin, Heidelberg: Springer Berlin Heidelberg), ISBN 978-3-540-24777-7, URL `http://dx.doi.org/10.1007/978-3-540-24777-7_10`.

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

Kortanek KO, Evans J (1967) Pseudo-concave programming and lagrange regularity. *Operations Research* 15(5):882–891.

Lemon KN, Verhoef PC (2016) Understanding customer experience throughout the customer journey. *Journal of marketing* 80(6):69–96.

Li G, Rusmevichientong P, Topaloglu H (2015) The d-level nested logit model: Assortment and price optimization problems. *Operations Research* 62(2):325–342.

Ma W, Simchi-Levi D (2021) Reaping the benefits of bundling under high production costs. *International Conference on Artificial Intelligence and Statistics*, 1342–1350 (PMLR).

Martello S, Toth P (1990) *Knapsack problems: algorithms and computer implementations* (John Wiley & Sons, Inc.).

Martínez-de Albéniz V, Planas A, Nasini S (2020) Using clickstream data to improve flash sales effectiveness. *Production and Operations Management* 29(11):2508–2531.

Pashigian BP, Gould ED (1998) Internalizing externalities: the pricing of space in shopping malls. *The Journal of Law and Economics* 41(1):115–142.

Seiler S, Pinna F (2017) Estimating search benefits from path-tracking data: measurement and determinants. *Marketing Science* 36(4):565–589.

Seiler S, Yao S (2017) The impact of advertising along the conversion funnel. *Quantitative Marketing and Economics* 15:241–278.

Stremersch S, Tellis GJ (2002) Strategic bundling of products and prices: A new synthesis for marketing. *Journal of marketing* 66(1):55–72.

Tarazano DL, Daemmrich A (2020) One-way supermarket aisles. https://invention.si.edu/one-way-supermarket-aisles, URL https://www.theatlantic.com/business/archive/2017/05/rana-plaza-four-years-later/525252/.

Underhill P (2009) *Why we buy: The science of shopping–updated and revised for the Internet, the global consumer, and beyond* (Simon and Schuster).

Wagner L, Martínez-de Albéniz V (2020) Pricing and assortment strategies with product exchanges. *Operations Research* 68(2):453–466.

Wang R, Sahin O (2018) The impact of consumer search cost on assortment planning and pricing. *Management Science* 64(8):3649–3666.

Wen CH, Koppelman FS (2001) The generalized nested logit model. *Transportation Research Part B: Methodological* 35(7):627–641.

Ziaei Z, Mersereau A, Emadi SM, Gargeya V (2023) Inferring consideration sets from heatmap data, available at SSRN 4363486.

Appendix

Proof of Lemma 1 Part (i). By contradiction, assume that it is optimal to activate $n \geq 1$ aisles, such that the constraints $\sum_{i \in I} w_i x_{ij} \leq W$ for $j \in \{1, \ldots n\}$. Since $W \leq 0$, then $\sum_{j=1}^{n} \sum_{i \in I} w_i x_{ij} \leq nW \leq W$, and hence combining all these aisles together is possible, thereby keeping revenues at least the same.

Part (ii). Consider an optimal solution. If it does include a product i with $0 \leq w_i \leq W$, move it to a new aisle, which will be visited since $w_i \leq W$, while the old aisle will also be visited since $\sum_{i' \in A_j, i' \neq i} w_{i'} \leq \sum_{i' \in A_j} w_{i'} \leq W$. Thus, profit will not reduce. Hence, by repeating this procedure, we obtain an optimal solution with at least $\left| \left\{ i | 0 \leq w_i \leq W \right\} \right|$ aisles. $\square$