Les Cahiers du GERAD ISSN: 0711–2440

Bundling or discounting? Field experiments for healthy and unhealthy food

N. Bandi, M. Cohen, S. Ray

G–2025–21

January 2025

Citation suggérée : N. Bandi, M. Cohen, S. Ray (Janvier 2025). Bundling or discounting? Field experiments for healthy and unhealthy food, Rapport technique, Les Cahiers du GERAD G– 2025–21, GERAD, HEC Montréal, Canada.

Avant de citer ce rapport technique, veuillez visiter notre site Web (https://www.gerad.ca/fr/papers/G-2025-21) afin de mettre à jour vos données de référence, s'il a été publié dans une revue scientifique.

Suggested citation: N. Bandi, M. Cohen, S. Ray (January 2025). Bundling or discounting? Field experiments for healthy and unhealthy food, Technical report, Les Cahiers du GERAD G–2025–21, GERAD, HEC Montréal, Canada.

Before citing this technical report, please visit our website (https://www.gerad.ca/en/papers/G-2025-21) to update your reference data, if it has been published in a scientific journal.

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2025
– Bibliothèque et Archives Canada, 2025

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2025
– Library and Archives Canada, 2025

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Bundling or discounting? Field experiments for healthy and unhealthy food

Nymisha Bandi [a, b]

Maxime Cohen [a, b]

Saibal Ray [a, b]

[a] *GERAD, Montréal (Qc), Canada, H3T 1J4*

[b] *Desautels Faculty of Management, McGill University, Montréal (Qc), Canada, H3A 1G5*

nymisha.bandi@mail.mcgill.ca
maxime.cohen@mcgill.ca
saibal.ray@mcgill.ca

January 2025
Les Cahiers du GERAD
G–2025–21

Abstract : Retailers routinely face the strategic decision of how to promote their products. Two of the most common promotion strategies are bundling and discounting. Determining which of these strategies is more effective, and how their effectiveness varies across product categories, is an open question in the world of retail. To address this, we conducted a multi-store field experiment with a global convenience store chain that regularly uses bundle promotions (particularly for certain unhealthy snacks). Our study focused on healthy and unhealthy food categories to understand how these promotions influence consumer behavior and financial outcomes. Specifically, we tested these strategies—discounts and bundles—applied to (i) healthy snacks, (ii) unhealthy snacks, and (iii) both categories simultaneously. For healthy snacks, we found that while discounting was more effective at boosting sales, bundling was more effective at increasing revenue and profit. Additionally, both promotion strategies for healthy snacks increased the sales of other unhealthy items, as consumers may engage in moral licensing, using their healthy purchase as justification to indulge in less healthy options. In contrast, for unhealthy snacks, neither promotion strategy had an effect on sales. Based on this surprising finding, we convinced the retailer to conduct a follow-up experiment to investigate the impact of discontinuing a long-standing bundle promotion on unhealthy snacks. The results of this second experiment revealed no change in sales but an increase in revenue and profit, confirming our initial conclusion. These insights provide valuable guidance for retailers seeking to optimize promotion strategies for different product categories.

Keywords: Retail promotions, bundling, discounting, field experiments, food choices

1 Introduction

Price promotions are a critical component of retailers' marketing strategies, designed to drive consumer engagement and boost sales and profits. Retailers employ a variety of price tactics or promotion vehicles (Baardman et al. 2019), including price discounts, bundling, buy-one-get-one-free offers, free samples, and coupons, all to influence consumers' price perceptions and shape their purchase decisions. When promoting a product, retailers face two primary decisions: determining the scale of the promotion and deciding how to frame and communicate the offer to consumers. Both decisions are crucial in influencing how consumers will react to promotions (Berkowitz and Walton 1980). While advanced dynamic algorithms can be used to set the optimal level of promotion prices (e.g., Cohen et al. 2017, 2021), the psychological impact of framing also plays an important role in purchasing decisions of consumers (Darke and Chung 2005). Studies like Alvarez and Casielles (2005) and Chenreddy et al. (2019) leverage historical sales and promotion data to identify patterns in consumer responses and help retailers design effective promotion campaigns. Two of the most commonly used promotion strategies are bundling (i.e., combining several products into a deal) and discounting (i.e., temporarily decreasing the price), each serving to influence consumer behavior and enhance profitability (Balachander et al. 2010).

More precisely, bundling combines multiple products at a single discounted price, simplifying consumers' purchase decisions and boosting the average transaction size by encouraging complementary purchases (Stremersch and Tellis 2002). In this paper, we consider the special case of add-on bundling, which consists of selling a complementary product as an add-on to a main product (e.g., customers have the option to pay a small additional amount to purchase a second product). Discounting, on the other hand, offers a direct price cut, hence enhancing the perceived value and boosting short-term sales by catering to price-sensitive consumers and creating a sense of urgency (Blattberg 1990, Nunes and Park 2003). Discounting is considered the most common type of sales promotion.[1] This is due to its ease of implementation and its historical success across various product categories (JRTech 2023). Bundling and discounting have become essential tools in retail across a wide range of sectors, from fast-moving consumer goods to technology products and even services (Bandi et al. 2024b).

Researchers have compared discounting to several other promotion types, such as bonus pack (Mishra and Mishra 2011), coupon promotions (Chen et al. 1998), and free product offers (Darke and Chung 2005). While discounting is popular among retailers, studies have shown that it is not always the most effective strategy (e.g., Palazon and Delgado-Ballester 2009). To our knowledge, there is limited research on comparing the effects of discounting and bundling on consumer preferences for different product categories. Previous studies suggest that the effectiveness of promotion strategies depends heavily on the product type and the context. For instance, Mishra and Mishra (2011) found that discounts may outperform bonus packs for unhealthy snacks in a controlled lab environment. Similarly, in another lab experiment, Tripathi and Pandey (2019) observed that green products are more effectively promoted using bonus packs, whereas non-green hedonic goods benefit more from discounts. Despite these insights, there remains a significant gap in the literature—and a lack of field evidence from physical retail—examining the comparative impact of bundling and discounting, within specific product contexts. Promotions often reflect broader trends, with unhealthy products frequently bundled or discounted, raising dietary concerns (Ravensbergen et al. 2015). In contrast, healthy food products face the challenge of overcoming perceived cost barriers (Haws et al. 2017), making tailored promotion strategies critical. Bandi et al. (2024a) found that using add-on bundling to promote healthy food increases the consumption of healthy snacks while simultaneously decreasing the consumption of unhealthy snacks.

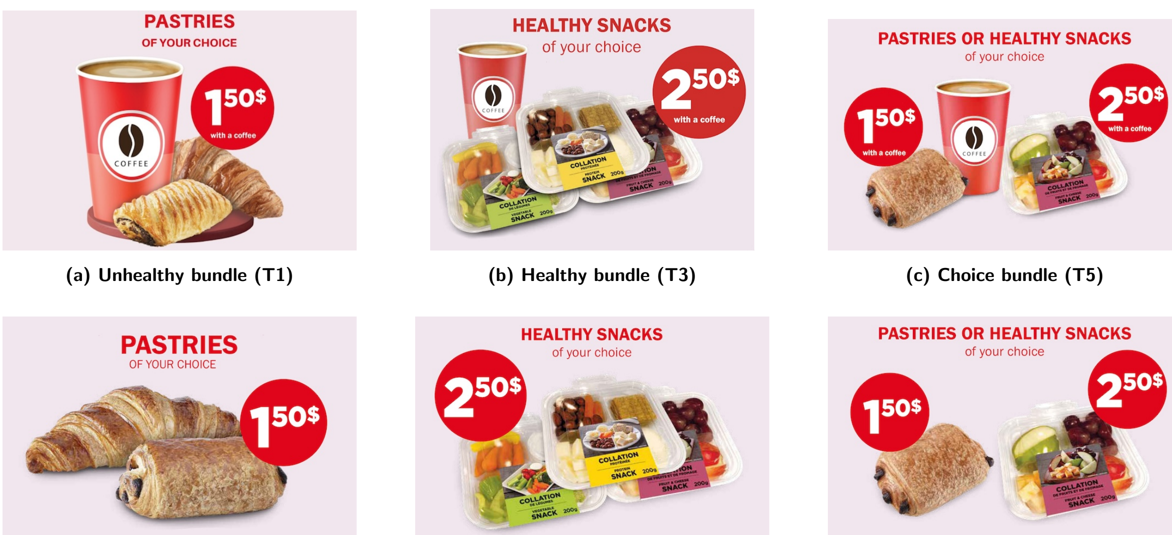
As mentioned, despite their widespread use, there is limited research on how consumers respond to bundling versus discounting across different product categories. There is a growing interest in using field experiments to derive causal insights and optimize strategies in real-world settings (Gao et al.

[1] https://www.fool.com/the-ascent/small-business/retail-management/retailer-promotions/

2023). This study aims to fill this gap by examining how those two promotion strategies influence both consumer behavior and financial outcomes, focusing on healthy and unhealthy food categories through a field experiment in physical retail. We concentrate on add-on bundling and discounting as the two types of promotions. The former involves a discounted product (e.g., A) tied to the purchase of another full-priced, typically high-margin, product (e.g., B, coffee in our case) (Figures 1a, b). For discounting, we consider a promotion design in which a product (say, A) is unconditionally discounted (Figures 1d, e). We also analyze two choice-based promotions, namely the choice bundle and the choice discount, in which consumers have the choice between healthy and unhealthy snacks (Figures 1c, f).

Consumers shop in various stores, including grocery chains and convenience stores (C-stores), as well as dollar stores, gas stations, and pharmacies, which differ significantly in product offerings, shopping experience, and responses to promotions (Pai et al. 2017). Despite their smaller size and limited range, C-stores are integral to shopping routines, with nearly half of consumers having weekly visits and 43% of Gen Z and Millennials making daily purchases in C-stores according to a Canadian survey.[2] Still, research on promotion effectiveness in C-stores remains limited, despite their unique consumption patterns and retail dynamics. Sharpening our understanding on these behaviors is essential for designing effective promotion strategies.

Our research aims to provide real-world evidence on the interplay between promotion strategies (bundling vs. discounting) and food categories (healthy vs. unhealthy) in C-stores. To do so, we conducted a multi-store field experiment in collaboration with a leading global C-store chain located in a major North American city. Analysis of existing promotion strategies for all products sold revealed that 61% were discounts, while 18% were bundles. These two types of promotions clearly dominated the landscape, though others, such as free-product offers and buy-one-get-one deals, were also occasionally used.



(a) Unhealthy bundle (T1)

(b) Healthy bundle (T3)

(c) Choice bundle (T5)

(d) Unhealthy discount (T2)

(e) Healthy discount (T4)

(f) Choice discount (T6)

Figure 1: Promotion posters used in the stores during the experiment.

In our field experiment, we implemented six promotions (or treatments), denoted T1 to T6: three bundle-based and three discount-based (see Figure 1). As discussed, we considered both healthy snacks (fruits, vegetables, and protein items) and unhealthy snacks (a variety of pastries) to study the effects across different product categories (more details can be found in Section 3). These six treatments were tested simultaneously across six stores, with a seventh store serving as a control. The promotions were carefully structured to maintain an equivalent percentage price decrease across

[2] https://ccentral.ca/c-store-iq-2023-purchasing-report

both types of strategies to ensure fair comparisons. To mitigate interferences, we utilized a rigorous five-period crossover design with alternating control periods. This design minimizes between-store variations and allows us to systematically evaluate the treatment effects of each promotion strategy on consumer behavior and on financial performance of the stores. The control store further helps account for unobserved temporal heterogeneity.

Our research offers several key contributions, detailed below and summarized in Table 1.

Bundling vs. discounting for healthy snacks. We found that both bundling and discounting were effective at increasing sales for the promoted healthy products, and overall revenue and profit for the three product categories used in the experiment. Interestingly, we found that bundling was more effective at increasing revenue and profit, while discounting was more effective at increasing sales. Indeed, bundling not only encouraged consumers to make healthier food choices but also strongly enhanced the retailer's revenue and profit, due to increasing the sales of the high-margin add-on item (coffee). Specifically, the sales of healthy snacks increased by 27.12%, 93.48%, 80.39%, and 105.44% under the healthy bundle (T3), healthy discount (T4), choice bundle (T5), and choice discount (T6), respectively, relative to the control condition. Revenue (resp. profit) increased by 79.41% (72.28%), 44.46% (38.47%), 112.46% (160.46%), and 55.06% (46.17%) under T3, T4, T5, and T6 respectively.

Bundling vs. discounting for unhealthy snacks. Surprisingly, we found no effect on the sales of unhealthy snacks either under bundling or under discounting (T1 and T2). This is contrary to our expectation that promoting a product would naturally provide a boost in sales. A plausible explanation is that the demand for unhealthy snacks may be relatively inelastic, with consumers exhibiting low sensitivity to promotions. To provide additional context, our retail partner has been offering an add-on bundling promotion on unhealthy snacks for several years preceding our field experiment. This promotion may have firmly anchored consumers' habits, making their purchase behavior less responsive to new promotions. Actually, we observed a significant decline in revenue and profit for those two treatments. Specifically, revenue (resp. profit) decreased by 19.78% (17.46%) and 27.63% (30.19%) under T1 and T2 relative to the control condition.

Impact on other product categories. We found that discount-based promotions significantly increased the sales of other healthy products that were not part of the promotion (namely, salads, fruits, and yogurt). The sales of these products increased by 36.34% and 41.91% under the healthy discount (T4) and the choice discount (T6) relative to the control condition (no effect for T2, as expected). Consumers might be using the savings from discount promotions to try healthier alternatives. In contrast, bundle-based promotions (T3 and T5) substantially reduced the sales of other beverages (namely, juice, energy drinks, and various types of water). Specifically, the sales of other beverages decreased by 7.25% and 8.32% for the unhealthy (T1) and healthy (T2) bundles (the choice bundle had no effect). This suggests that bundle-based promotions cannibalize the sales of other beverages as consumers substitute their usual beverage with a coffee (coffee being the primary product in the bundle promotions). We also found that healthy promotions (T3 and T4) increased the sales of sweets/candies (10.52% and 15.03%) and cigarettes (8.33% and 6.18%), implying that promotions on healthy products may paradoxically encourage indulgence in less healthy habits. This phenomenon may stem from moral licensing, where consumers perceive their healthy purchase as a virtuous act, thereby justifying indulgence in less healthy options (Khan and Dhar 2006).

Table 1: Summary of our results.

	Sales	Revenue and Profit
Healthy	Discounting is better	Bundling is better
Unhealthy	No effect	No promotion is better
Choice	Discounting is better	Bundling is better

Since we clearly observed that unhealthy promotions had no effect on the sales of unhealthy products but were lowering revenue and profit, we managed to convince our retail partner to run a follow-up experiment to investigate the impact of discontinuing the long-lasting bundle promotion on unhealthy snacks. As mentioned, the retailer had used this unhealthy add-on bundle for the past several years across many of its stores. More precisely, our follow-up field experiment examines the effect of discontinuing the unhealthy bundle promotion in 10 randomly-selected stores for 8 weeks. We found that discontinuing this promotion had no effect on the sales of unhealthy products while significantly increasing revenue and profit. This supplementary experiment further validated our earlier conclusion that both bundling and discounting had no impact on the sales of unhealthy products. The results suggest that prolonged exposure to consistent promotions may diminish the effectiveness of new promotion strategies.

Our main empirical strategy relied on a series of A/A tests to select the stores, and on regression analysis and difference-in-differences (DID) method to compute the treatment effects. For the follow-up experiment, we combined the DID method with propensity score matching (PSM) to identify a set of control stores that closely resemble our treatment stores. We used several control variables, such as time-fixed effects, store-fixed effects, and product stockouts. Finally, we conducted a series of robustness tests to showcase the stability of our results.

The remainder of the paper is structured as follows. Section 2 outlines the development of our hypotheses, followed by presenting the design, data, and metrics used in our field experiment in Section 3. Section 4 presents the results of our main experiment, including detailed analyses of revenue and profit implications. Section 5 provides additional analyses such as DID analysis, impact on store performance metrics, and the impact on the products outside our interventions, while Section 6 describes the design and findings of our follow-up experiment. Finally, Section 7 offers managerial insights and conclusions. Comprehensive robustness tests, supplementary figures, and detailed estimation results are relegated to the appendix.

2 Hypotheses development

As discussed, since consumer responses are significantly shaped by how promotions are framed and communicated, bundling and discounting offer distinct psychological appeals, even when they are financially equivalent in terms of discount percentage. According to the mental accounting framework (Thaler 1985), consumers tend to perceive bundled goods as a single transaction rather than a series of smaller, individual purchases, hence leading to a reduced price sensitivity. This can make bundles more appealing, especially when the perceived value of the bundle exceeds the sum of its individual parts (Yadav and Monroe 1993). Discounts, on the other hand, tap into consumers' desire to get a good deal, leveraging the power of loss aversion by framing the promotion as an opportunity to avoid paying the full price (Kahneman and Tversky 2013). Importantly, the effectiveness of these promotions varies based on individual consumer traits, shopping habits, and preferences. For example, price-sensitive consumers may be more drawn to discounts, whereas those seeking convenience may prefer bundles (Meyer and Johnson 1995).

We next develop our three hypotheses on the impact of the two promotion types (bundling and discounting) while providing support from the literature.

Hypothesis 1.

a) Both bundling and discounting increase sales relative to not having a promotion, irrespective of the product category.

b) Discounting is more effective than bundling at increasing sales assuming the same percentage reduction in price is applied in both cases, irrespective of the product category.

In bundling promotions, consumer behavior is influenced by the anchoring effect, a principle from behavioral economics. More precisely, the purchase of Product A serves as a cognitive anchor, framing the reduced price of Product B as a gain. Thus, consumers perceive the price reduction on B as a reward for their commitment to buying A, hence enhancing their sense of value (Yadav 1994). Consequently, bundling often encourages multi-product purchases, thus increasing the basket size and potentially boosting sales, irrespective of the product category (Stremersch and Tellis 2002). This theory strongly supports Hypothesis 1a that bundling is effective at increasing sales relative to not promoting.

On the other hand, discounting a product as a standalone promotion directly appeals to price-sensitive consumers, leveraging the principle of reference price adjustment (Kopalle and Lindsey-Mullikin 2003). Consumers tend to compare the discounted price to the original price, interpreting the price cut as an immediate saving. This type of promotion triggers heuristic-based decision-making, where the salience of the discount often outweighs the perceived need or utility of the product. Blattberg (1990) and Chandon et al. (2000) demonstrate that discounts consistently drive an increase in purchase volume by appealing to price-sensitive consumers. This effect is observed across a range of product categories, from essential goods to discretionary items, underscoring the universal appeal of price reductions, hence supporting Hypothesis 1a that discounting a product increases its sales relative to not promoting. Indeed, for the vast majority of retail products (with the exception of luxury and niche products for example), customer demand is generally assumed to be a decreasing function of the price (Cohen et al. 2022). Offering a promotion lowers the effective price and, thus, increases sales. At the same time, it is possible to observe a high level of heterogeneity in terms of the effect of such price reductions. For example, Hoch et al. (1995) shows that price elasticity estimates vary significantly across product categories. It is even possible that the effect (i.e., how much sales increase when a promotion is offered) is close to zero for some products. In this paper, we will investigate whether the effect is present for two types of products: healthy and unhealthy snacks.

Unlike bundling, discount promotions do not rely on an anchoring effect tied to another purchase, leading consumers to evaluate the offer solely based on the perceived value of the discounted product. This limits the promotion's ability to upsell or cross-sell additional items, but at the same time, it appeals to a broader audience by offering flexibility to consumers who may not be interested in bundled products (Hähnchen and Baumgartner 2020). Discounts are generally easier to evaluate than bundles, as they provide immediate, transparent savings without requiring consumers to assess the combined value of multiple products. Additionally, discounts align with consumer preferences for lower cognitive effort in decision-making, whereas bundles may obscure perceived value, especially if not all items are equally desirable. Because bundling is shown to outperform not promoting, and discounting is a less restrictive promotion than bundling, we may hypothesize that discounting is the most effective strategy for increasing sales. These factors, combined with framing effects that favor straightforward price reductions, suggest that discounts drive higher sales than bundles, even when the percentage of price reduction is identical, once again, supporting Hypothesis 1b. In this paper, we rigorously test this hypothesis across two different product categories, further evaluating the relative effectiveness of bundling and discounting in influencing consumer behavior.

Hypothesis 2.

a) Bundling generates higher revenue and profit relative to either offering a discount promotion or not having a promotion.
b) Offering a choice bundle (T5) generates higher revenue and profit relative to all other conditions (C, T1, T2, T3, T4, and T6).

Bundling, as a strategic promotion mechanism, allows retailers to appeal to consumers by combining multiple products into a single offer at an attractive price. This approach is particularly effective at increasing the consumer surplus by targeting individuals who may otherwise purchase only one product or buy both products at prices below their willingness-to-pay (Stremersch and Tellis 2002). Consumers often display an asymmetry (i.e., a negative correlation) in their willingness to pay for complementary products. By bundling multiple products, retailers can take advantage of this asymmetry and increase the sales, the revenue, and the consumer surplus. For example, pricing each product separately at a relatively low price may increase the sales but will leave money on the table from the customers who are willing to pay a higher price for one of the products. By carefully pricing the bundle, one can mitigate this issue and capture a higher revenue and profit. Moreover, bundling can increase the perceived value for consumers, reduce transaction costs, and simplify the customer decision-making process, which all enhance the appeal of the bundle deal compared to a simple discounting strategy. Research has shown that bundling can significantly improve revenue and profit in various settings, including digital subscriptions, software packages (Bakos and Brynjolfsson 2000), and physical retail, where the complementarity between the bundled products strengthen consumer purchase intentions (Hanson and Martin 1990). The above arguments strongly support Hypothesis 2a.

Bundling also facilitates cross-selling opportunities by encouraging consumers to purchase combinations of products they might not have otherwise considered. For instance, the strategic inclusion of products in a bundle can offset the lower margins of other products, hence optimizing the overall profitability of the bundle promotion (Guiltinan 1987). When designed carefully, bundles not only increase revenue and profit but also strengthen customer perceptions of value and satisfaction. Hypothesis 2a aligns with our prior discussion arguing that bundling is more profitable for the retailer than either discounting or not having a promotion.

While the inclusion of healthy snacks in the bundle promotion increases their visibility and accessibility, the hedonic appeal of unhealthy snacks often dominates consumer decisions. Dual-process theories of decision-making (Kahneman 2011) explain this behavior as a conflict between impulsive, pleasure-driven decisions and rational, goal-oriented ones. For many consumers, the immediate gratification provided by unhealthy snacks outweighs the long-term benefits of choosing a healthy option, particularly when both options are presented concurrently in a promotion. Research by Bandi et al. (2024a) has shown that choice-based bundling allows retailers to appeal to a broader spectrum of consumers - both who are attracted by bundling and those who are attracted by hedonic products. This results in increased bundle sales and, consequently, higher revenue and profit for the choice bundle, ultimately supporting Hypothesis 2b.

Hypothesis 3. Discontinuing a long-lasting add-on bundling promotion on unhealthy products does not have a negative impact on the sales of the unhealthy products, and hence increases revenue and profit.

Interestingly, we could find arguments both in favor and against Hypothesis 3 in the literature.

On the one hand, studies in behavioral economics suggest that long-lasting consistent promotions foster habitual purchasing behavior, particularly for hedonic products like unhealthy snacks (Aaker 2009). This behavior may even persist when promotions are discontinued, as both the convenience and the emotional satisfaction associated with the product are likely to outweigh price considerations. Discontinuing a long-lasting promotion often is not likely to significantly alter consumer purchasing patterns, and hence may not have a negative impact on the sales. This could be attributed to inertia in consumer habits (Cui et al. 2021), wherein existing consumption patterns persist due to psychological anchoring from prior promotions or from the habitual nature of the purchase (Gupta 1988). In essence, after a certain time, the promotion is not necessary anymore to sustain the sales, since the customers are already hooked and loyal to the product.

On the other hand, some studies suggested that discontinuing long-term promotions can have a negative effect. For example, Melnyk and Bijmolt (2015) found that ending a loyalty program that offered discounts to consumers significantly reduced loyalty among long-term members, who became less likely to continue shopping at the retailer. Similarly, Carlson (2017) investigated the effect of discontinuing a bonus promotion campaign (e.g., buy-one-get-one free) for Walmart, CVS, Target, and dollar stores after one year and found that 75% of consumers stopped purchasing the bonus products. Nevertheless, this study was limited to a single type of promotion and to specific types of products.

The literature thus presents conflicting perspectives, and it remains unclear how consumers respond when long-term promotions are terminated. Our second field experiment offers a unique opportunity to formally test Hypothesis 3 for a long-lasting add-on bundling promotion on unhealthy snacks to advance our understanding on this topic.

3 Experiment design, data, and metrics

Our field experiment was conducted in collaboration with a global C-store chain operating thousands of stores worldwide. The experiment took place in a metropolitan North American city, where our retail partner has more than 90 stores, most operating 24 hours a day, seven days a week. Our experimental framework consists of three key components: the treatments, the design and timeline of the experiment, and the criteria for selecting the stores where the experiment was conducted. The following sections provide a detailed discussion of each aspect.

3.1 Treatments

(a) Fruit snack

(b) Vegetable snack

(c) Protein snack

(d) Pastry

Figure 2: Healthy and unhealthy snacks used in our promotions.

The treatment design involves key promotional elements, such as selecting the products to be promoted and determining their promotional prices. Since our experiment includes both healthy and unhealthy products, categorizing items into these groups serves as a crucial initial step.

Product categorization: To categorize products as healthy and unhealthy, we adhere to common conventions supported by the nutrition literature (Lachance and Fisher 1986, Scheidt and Daniel 2004, Lobstein and Davies 2009). Products categorized as healthy typically contain fresh fruits or vegetables, low fats, low sugar, and low carbohydrates, while being high in fiber and in essential nutrients. Our retail partner sells snack boxes that consist of an assortment of healthy items including fruits, protein, nuts, and vegetables. Specifically, we identified three types of healthy snack boxes: fruits, vegetables, and protein (Figure 2a–c) to be used in our experiment. Product descriptions are provided in Table 2. Conversely, all pastry items, croissants, cinnamon rolls, apple turnovers, fruit Danishes, chocolate avalanches, and chocolate muffins (Figure 2d), are classified as unhealthy aligning with general consumer understanding and the nutrition literature.

Promotion price: Prior to our experiment, the original price of all the healthy snack boxes was \$4.69, while the price of the unhealthy snacks was \$2.85, on average (ranging from \$2.09 to \$3.19). After

Table 2: Details of healthy and unhealthy snacks used in the experiment, including prices and discounted prices.

Snack type	Contents	Product images	Price	Discounted price (discount percentage)
Vegetable box	Celery, broccoli, carrots, pepper, and a dip	Figure 2b	\$4.69	\$2.50 (46.7%)
Fruit box	Apple slices, grapes, and cheese	Figure 2a	\$4.69	\$2.50 (46.7%)
Protein box	Hard boiled egg, almonds, cheese, and crackers	Figure 2c	\$4.69	\$2.50 (46.7%)
Pastry	Croissant, cinnamon roll, apple turnover, fruit danish, chocolate avalanche, chocolate muffin	Figure 2d	\$2.85	\$1.50 (47.3%)

extensive discussions with the pricing team at the C-store chain, we were instructed to adhere to certain pricing guidelines. First, all promotion prices were required to end in zero or five cents to align with the company's pricing policy. Second, we aimed to ensure that the discount percentages for both product categories were as close as possible.

As mentioned, prior to our experiment, the C-store chain had been running a successful add-on bundle promotion on unhealthy snacks for several years. Just before our experiment started, this promotion allowed customers to add a pastry for an additional \$1.50 when they purchased a coffee (or any other hot beverage). Thus, the average discount on unhealthy snacks in this promotion was 47.3%. We used this existing promotion framework for T1 in our experiment. The promotion banner located outside the stores for T1 is shown in Figure 1a. To establish the unhealthy discount for T2, we maintained the same discount of \$1.50 but removed the requirement of purchasing a coffee. As a result, customers exposed to T2 could purchase a variety of unhealthy snacks at the reduced price of \$1.50, without the requirement to buy a coffee.

To create comparable treatments for the healthy snacks, we took several factors into account to align with the existing promotion framework while ensuring both customer appeal and cost effectiveness. Our objective was to offer a discount that matched the perceived value as in the unhealthy snack promotion, thereby encouraging healthier choices. A perceived discount of 47.3% on healthy snacks resulted in a discounted price of \$2.47 for the snack boxes. To adhere to the company's pricing guidelines, we set a discounted price of \$2.50 for the healthy snacks. According to Lehtimäki et al. (2019), any price variation within $\pm 2\%$ falls within the range of price indifference, meaning consumers are almost insensitive to such changes when evaluating a promotion. In summary, consumers exposed to T3 could add a healthy snack box for an additional \$2.50 when purchasing a coffee (see Figure 1b). Similarly, those exposed to T4 could buy a healthy snack box at the reduced price of \$2.50 without having to purchase a coffee (Figure 1e).

The final set of interventions (T5 and T6) aimed to assess customer preferences between healthy and unhealthy snacks when both options were offered simultaneously, giving customers the freedom to choose. We referred to them as "choice promotions". We maintained the same promotion prices as in T1–T4. The choice bundle promotion implemented in T5 is illustrated in Figure 1c. Specifically, when customers purchased a coffee, they had the option to add either a healthy snack box for an additional \$2.50 or an unhealthy snack for \$1.50. This promotion allowed consumers to choose between a healthy or an unhealthy snack as an add-on while keeping the percentage discount consistent across both product categories. Similarly, we introduced a choice discount option (T6), where both healthy and unhealthy snacks were promoted simultaneously, without the requirement to buy a coffee, as seen in Figure 1f. This approach also enabled consumers to choose between both types of products while receiving the same percentage discount.

3.2 Design and timeline

Having designed the six promotions outlined above, our objective was to design a field experiment in physical retail stores while controlling for any source of variation across observations. In an ideal world, the effect of the six different promotions on the sales of different products could be best identified if the promotions were all implemented simultaneously in the same store. However, a physical store can run only one promotion at a time, which prevents the ability to assess multiple promotions concurrently. To address this, we deploy different promotions across several stores at the same time, allowing us to observe product sales under various promotions simultaneously. At the same time, this strategy may introduce differences between stores that are unrelated to the promotions. To minimize this between-store variation, we implemented a five-period crossover design with alternating control (or washout) periods to rigorously evaluate the treatment effects of the different promotion strategies across multiple stores. In this design, each store alternates between a control period, where no new promotion is offered (serving as a baseline or washout period), and a treatment period, where a specific promotion is implemented. Specifically, Stores 1 and 2 have T1 and T2 that crossover with a washout period (C) running between them as seen in Table 3. This careful design ensures that each promotion is implemented in exactly one store at any given moment. In addition, the use of a crossover design combined with incorporation of store fixed effects in the model enhances the robustness of the estimates by alleviating the biases stemming from the specific sequencing of the promotions and from store-specific factors. While this method controls for differences between stores, it introduces temporal variation within the stores that offer the promotions. To control for this temporal variation, we use an additional control store where the no-promotion strategy remains in effect throughout the entire experiment. This allows us to compare the sales within the same store over time, isolating the impact of each promotion while controlling for other factors. By using this alternating design, we ensure that each promotion strategy is tested within each store while providing sufficient time for any carryover effects from previous periods to diminish, thereby enhancing the robustness and validity of our findings. We also start and end all the promotions at the same time to ensure a time-balanced experiment design.

Table 3: Experiment design (C indicates Control and T_i represent the different treatments).

Time period	Number of days	Store 1	Store 2	Store 3	Store 4	Store 5	Store 6	Store 7
11/06/23 to 11/26/23	21	C	C	C	C	C	C	C
11/27/23 to 12/17/23	21	T1	T2	T3	T4	T5	T6	C
12/18/23 to 01/07/24	21				Excluded			
01/08/24 to 01/28/24	21	C	C	C	C	C	C	C
01/29/24 to 02/18/24	21	T2	T1	T4	T3	T6	T5	C
02/19/24 to 03/10/24	21	C	C	C	C	C	C	C

In total, our experiment involves seven stores, with one store designated as the control store and the remaining six as treatments. As discussed above, each treatment store alternates between treatment and control periods to allow for a robust analysis of the treatment effects. The experiment lasted for 18 weeks as shown in Table 3. Out of these 18 weeks, we drop the data collected between December 18, 2023 and January 7, 2024 (end-of-year holiday) as this period is likely unrepresentative. The detailed timeline is summarized in Figure 3.

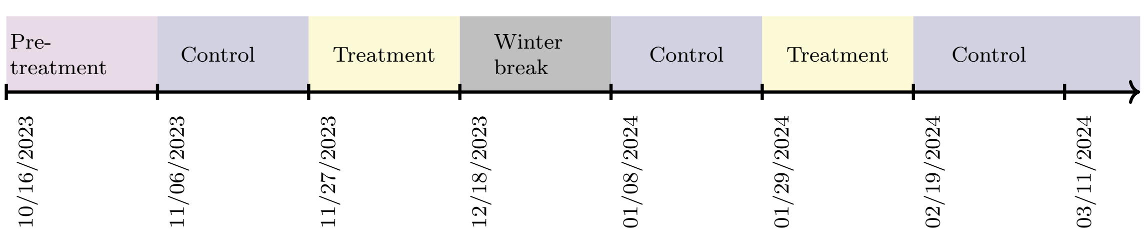


Figure 3: Timeline of our field experiment.

As mentioned, our industry partner has been running a long-lasting add-on bundle promotion on unhealthy snacks (the same as T1) across all stores for the past several years. To reset consumer perceptions and create a neutral baseline with no promotion, we discontinued this promotion in all seven experiment stores starting October 16, 2023, namely, three weeks before our experiment, which officially launched on November 6, 2023, under the control condition (i.e., no promotion). This reset (pre-experiment) period allows us to form a neutral baseline of having no promotion, and compare our six promotions relative to this baseline. Interestingly, this reset baseline remained stable over time, showing no statistically significant pairwise differences at three, six, nine, twelve, and fifteen weeks for Store 7 (the details can be found in Appendix B).

3.3 Stores selection

When selecting the seven experimental stores, the primary objective was to ensure that customer behavior across these stores was similar. All the stores in the city (around 90) operated by the C-store chain can be categorized into two groups depending on whether the store is co-located with a gas station or not. For the purpose of this experiment, we excluded the stores located in gas stations, reducing the list to 46 stores. We next eliminated six stores that did not offer all the products included in our experiment (i.e., healthy snacks, unhealthy snacks, and coffee). This resulted in a final pool of 40 stores, from which we randomly selected seven stores for the experiment, one for control and six for treatments. Prior to the start of our experiment, a formal A/A analysis was conducted to confirm the comparability of these seven stores. Our analysis allowed us to draw the clear conclusion that all seven stores chosen for our experiment have very similar trends in terms of customer behavior (across several variables) and can be used as a solid experiment framework to make meaningful comparisons. The details of the analysis can be found in Appendix A.

3.4 Data and metrics

In this section, we provide an overview of the data collected during the experiment. The dataset comprises transaction data including transaction time, quantity of each product sold, price, promotion discount, and end-of-day (EOD) inventory for each product in the store. The inventory data is important to help us identify stockouts, which can have an impact on our inference. However, the recorded inventory levels were not perfectly accurate, as they were generated using estimation algorithms rather than actual physical counts, which are impractical for large organizations due to logistical challenges. To address the inaccuracy in inventory records and accurately determine when specific products were out of stock, we used both the sales data and the reported EOD inventory. Specifically, if there were no recorded sales for a particular product on a given day and the EOD inventory was non-positive, we inferred that the product was not available on that day. This inference was captured by a binary variable (denoted SO) as a control in our empirical models to account for potential stock-related issues. We also routinely visited the seven experimental stores in person to carefully ensure that inventory availability was not a problem.

To evaluate the performance of the different promotions, we considered the following three key metrics validated by our retail partner:

Sales: The total sales units of a specific product (healthy or unhealthy snack) in each store on a given day. This metric captures the direct impact of the promotions on the sales volume of the two product categories.

Revenue: The total revenue from the sales of the products used in our experiment (healthy snacks, unhealthy snacks, and coffee) in each store on a given day. This metric measures the financial performance of the store (for the focal products) and the effectiveness of the promotions in boosting the top line.

Profit: The net earnings from the daily sales of the products used in our experiment (healthy snacks, unhealthy snacks, and coffee) in each store. This metric evaluates the effectiveness of the promotions in terms of profitability.

In addition to these three performance metrics, we also analyzed the number of transactions, the average basket size, and the overall store revenue. We also examined the effect on the sales of the products outside the scope of the experiment. These additional metrics provide further insight into whether the different promotions attracted more shoppers to the store and how they affected customer spending.

To enhance the reliability of our analyses, we refined the dataset by discarding transactions with abnormally high sales or revenue. Specifically, we excluded outliers at the top 1% of observations for each key variable, based on their respective distributions. For instance, when examining the sales volume during the experiment, we analyzed the data distribution and removed the highest 1% of transactions involving exceptionally high sales. This data filtering aims to make our dataset more representative by removing rare abnormal transactions. To validate the robustness of our conclusions, we tested alternative thresholds, incrementally increasing the cutoff from 1% to 3%. The results remained consistent across all tested thresholds, hence reinforcing the reliability of our findings. In addition, we also obtained consistent results when not removing outliers at all.

4 Main experiment

In this section, we present the results of our field experiment. We first present our A/A test showing the validity of the control conditions and the experimental design. We then report the results from the experiment on the sales, revenue, and profit. Our main econometrics methods are ANOVA and regression analyses.

4.1 A/A test

In a field experiment such as ours, a formal A/A test is essential to ensure the validity of the control periods across the different stores. Recall that our experiment involves alternating periods of control and treatment across multiple stores, with one store consistently acting as an additional control. Before estimating the treatment effects, it is crucial to verify that the control periods are stable and not subject to systematic differences that could confound the results. We performed an A/A test as a diagnostic tool to confirm that the outcome variables behaved similarly across control periods, both within the same store and across stores. This allowed us to confidently attribute the observed effects to the treatments. The details of the A/A test are provided in Appendix C.1.

4.2 Bundling vs. discounting – Aggregate level analysis

In this section, we analyze the effects of both promotion strategies—bundling and discounting—at the aggregate level without differentiating between product types. More precisely, we examine the impact of bundling (combining T1, T3, and T5) and discounting (combining T2, T4, and T6) promotions individually, relative to having no promotion (C), for healthy and unhealthy snacks together. This approach allows us to quantify the overall impact of each promotion type on the sales, revenue, and profit. We use the following model specification:

$$Y_{ds} = \alpha_s + \beta_B \cdot I_B + \beta_D \cdot I_D + \mu_d + SO_{ds} + \epsilon_{ds},$$

where Y_{ds} denotes the total value of the dependent variable (DV) (sales, revenue, profit in our case) on day d in store s. For count-based variables (sales), we use a Negative Binomial (NB) regression model. In contrast, for continuous variables (revenue and profit), we apply a logarithmic transformation and use a Two-Way Fixed Effects (TWFE) regression model. I_B and I_D represent the type of promotion applied on day d in store s, specifically bundling and discounting, respectively, α_s represents the store fixed effects for store s, μ_d represents the time fixed effects for day d, and SO_{ds} is a binary variable that indicates whether or not a stockout occurred on day d for promoted products in store s. We considered various types of time-fixed effects, including day-of-week effects, as well as the week number, month, and year during the treatment period. Finally, the parameters β_B and β_D capture the impact of each promotion type.

Figure 4 plots the sales of the promoted products (healthy + unhealthy snacks) for each promotion type along with a 95% confidence interval.[3] Table 4 reports the estimation results and shows that both bundling and discounting lead to a significant increase in sales relative to having no promotion, namely, 13.88% and 20.92%, respectively. This result strongly supports Hypothesis 1a (at the aggregate level). While both promotion strategies are effective at increasing sales, discounting demonstrates a notably higher impact, suggesting that price discounts might be a more successful driver in influencing customer purchase decisions. On the other hand, we find that bundling is more effective than discounting at boosting revenue and profit. Specifically, bundling increases revenue (resp. profit) by 16.36% (resp. 15.03%) while discounting has no statistically significant effect on neither revenue nor profit, hence supporting Hypothesis 2a.

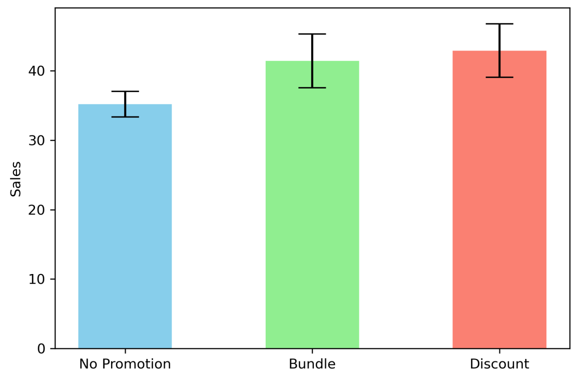


Figure 4: Comparing bundling vs. discounting (sales of the promoted products).

Table 4: Impact of the two promotion types on sales, revenue, and profit.

	Sales	Revenue	Profit
Bundle (B)	0.13 **	0.15 **	0.14 **
	(0.04)	(0.06)	(0.06)
Discount (D)	0.19 ***	0.09	0.02
	(0.04)	(0.06)	(0.06)
No. Obs.	735	735	735
Log likelihood	-8478.9	-611.18	-593.20
Model	NB	TWFE	TWFE

Note: $^{*}p < 0.05$;$^{**}p < 0.01$;$^{***}p < 0.001$.
The standard errors are reported in parentheses.
No. Obs. = 21 days per treatment ×5 treatment periods ×7 stores = 735.

In summary, while both bundling and discounting are effective at increasing sales, only bundling has an impact on revenue and profit. Overall, the bundling strategy strikes a more balanced approach: it drives substantial sales increase (albeit less than discounting) while simultaneously enhancing revenue and profit. Conversely, offering discounts substantially boosts sales, which could be more beneficial for short-term goals such as clearing out inventory. The above analysis takes an aggregate approach and does not account for the potential differences across product categories (healthy and unhealthy). In the next analysis, we will analyze each category separately to provide deeper insights into how the effectiveness of promotion strategies differs across product categories.

4.3 Bundling vs. discounting – Sales analysis across product categories

We now examine the effects of bundling and discounting on healthy and unhealthy snacks separately. In Section 4.3, we focus on sales as the dependent variable and defer the analysis on revenue and profit

[3] In this paper, all confidence intervals are reported at the 95% level.

to Section 4.4. We estimate a regression model to assess the effect of each promotion from Figure 1. As before, we use the following NB regression model:

$$Y_{dps} \sim \alpha_s + \sum_{I \in \{T_1, T_2, T_3, T_4, T_5, T_6\}} \beta_I \cdot I + \mu_d + SO_{dps} + \epsilon_{dps}, \quad (1)$$

where Y_{dps} corresponds to the sales of the different product categories p={healthy snacks, unhealthy snacks} on day d in store s. I represents the different treatments $T_1, T_2, T_3, T_4, T_5, T_6$ on day d (see Table 3), so that β_I is the coefficient for treatment $I \in \{T_1, T_2, T_3, T_4, T_5, T_6\}$. The key parameters in Eq. (1) are $\beta_{T_1}, \beta_{T_2}, \beta_{T_3}, \beta_{T_4}, \beta_{T_5}, \beta_{T_6}$, which capture the effect of each promotion on the sales. The estimates are reported in Table 5. Next, we provide a detailed explanation and interpretation of the results for each promotion strategy: unhealthy promotions (T1 and T2), healthy promotions (T3 and T4), and choice-based promotions (T5 and T6).

Table 5: Estimates of the effect of the different promotions on sales.

Product promoted	Promotion type	Unhealthy sales	Healthy sales
Unhealthy	Bundle (T1)	0.01 (0.05)	0.35 (0.24)
	Discount (T2)	0.06 (0.05)	0.40 (0.24)
Healthy	Bundle (T3)	-0.09 ** (0.05)	0.24 ** (0.08)
	Discount (T4)	0.03 (0.04)	0.66 *** (0.07)
Choice	Bundle (T5)	0.03 (0.05)	0.59 *** (0.14)
	Discount (T6)	0.06 (0.05)	0.72 *** (0.13)
No. Obs.		735	735
Log likelihood		-3,185.0	-1,360.9
Model		NB	NB

Note: $^{*}p < 0.05$;$^{**}p < 0.01$;$^{***}p < 0.001$.
The standard errors are reported in parentheses.
No. Obs. = 21 days per treatment ×5 treatment periods ×7 stores = 735.

Unhealthy promotions (T1, T2) vs. Control (C): We found that both unhealthy promotions, namely unhealthy bundle (T1) and unhealthy discount (T2), did not lead to a noticeable increase in the sales of unhealthy snacks when compared to the control setting. This finding is surprising as promotions are typically expected to boost sales. This result rejects Hypothesis 1a for unhealthy snacks, which posits that both types of promotions should increase sales compared to having no promotion. Additionally, it also does not support Hypothesis 1b, which claims that discounting is more effective than bundling in boosting sales. Instead, we found that neither promotion was effective for unhealthy snacks. A potential explanation could be that demand is inelastic for unhealthy snacks since consumers seem to purchase them even when not promoted (this may be partially due to, as discussed before, the long-lasting past promotion offered by the retailer). A second explanation could be that for hedonic products like pastries, a promotion is not needed to sustain high sales and consumers might purchase them as a matter of habit. We will elaborate on this important finding in Section 6. In addition, we found no effect on the sales of healthy snacks as expected, as the promotions were specifically targeted at unhealthy snacks and did not include any incentive for healthy snacks.

Healthy promotions (T3, T4) vs. Control (C): We observed a significant increase in the sales of healthy snacks under the healthy bundle (T3) and healthy discount (T4), namely 27.12% and 93.48%, respectively. These results align with our expectations, as these promotions focused on healthy snacks, thereby supporting Hypothesis 1a that both promotion strategies should boost sales. More importantly, we found that discount-based promotions were more effective than bundling at increasing sales for healthy snacks, hence supporting Hypothesis 1b. This outcome can be attributed to the inherent flexibility of discount promotions, as they are not tied to the purchase of a coffee to redeem the price reduction. By removing the dependency on an additional purchase, the discount promotion appeals to a broader range of consumers, including those who are not interested in the bundled product. Interestingly, we observed a decrease of 8.61% in the sales of unhealthy snacks under the healthy bundle (T3). This result is largely explained by the behavior of customers who routinely purchased the unhealthy bundle during the control period (without any promotion) since the combination of coffee + pastry is very common and appreciated by customers. When the healthy bundle (T3) was offered, some of those "coffee-lover" customers substituted the unhealthy snack with a healthier alternative, hence leading to a significant change in their purchasing habits—which can be seen as a socially positive result. In contrast, the healthy discount (T4) neither reduced nor increased the sales of unhealthy snacks, as anticipated. This outcome can be attributed to the promotion's focus on healthy snacks, without any connection to coffee or unhealthy snacks. Unlike T3, where customers were offered an appealing alternative in the form of a healthy bundle that included coffee, T4 did not provide a comparable option to encourage habitual unhealthy bundle buyers to modify their purchasing behavior. As a result, the sales of unhealthy snacks remained unchanged, while the promotion mainly attracted customers seeking a healthier option.

Choice promotions (T5, T6) vs. Control (C): We found that the sales of healthy snacks increased by 80.39% and 105.44% for the choice bundle (T5) and choice discount (T6), respectively. Notably, under T5 and T6, the sales of unhealthy snacks remained unaffected. It implies that the increase in healthy snack sales is likely to be driven by customers who leverage the choice-based promotions to either purchase the healthy bundle (T5) or buy only healthy snack boxes (T6). At the same time, the steady sales of unhealthy snacks suggest that a segment of customers continued to purchase these items out of habit or preference, unaffected by the availability of a healthier promotion. By including both healthy and unhealthy snacks in the promotion, the choice-based strategy took advantage of habitual indulgent behavior, maintaining unhealthy snack sales, while simultaneously boosting healthy snack sales. This dual impact highlights the key advantage of offering such an option. We also observed that the choice discount (T6) resulted in a greater effect than the choice bundle (T5). As discussed before, this can be attributed to the inherent flexibility of discount promotions. Unlike bundles, discounts do not require consumers to purchase an additional item (coffee), and hence appeal to a broader audience.

Choice promotions (T5, T6) vs. Unhealthy promotions (T1, T2): When comparing choice-based promotions (T5 and T6) to unhealthy promotions (T1 and T2), we observed a noticeable increase in the sales of healthy snacks while the sales of unhealthy snacks remained unchanged. This outcome is intuitive, as choice-based promotions specifically offer the choice of (discounted) healthy snacks, hence appealing to health-conscious consumers. The simultaneous promotion of unhealthy snacks in choice-based promotions did not change their sales. As discussed earlier for T1 and T2, this result can be explained by the relatively stable demand for unhealthy snacks, where habitual buyers consistently maintained their purchasing behavior, regardless of the promotion.

Choice promotions (T5, T6) vs. Healthy promotions (T3, T4): When comparing choice-based promotions (T5 and T6) to healthy promotions (T3 and T4), we found that the former ones had a greater effect on the sales of healthy snacks. Additionally, the choice bundle (T5) resulted in higher sales of unhealthy snacks compared to the healthy bundle (T3). These results can be explained in two ways. First, unlike healthy promotions, which effectively "forced" consumers to only consider healthy options, choice-based promotions offered the flexibility to indulge, hence leading to an increase in the sales of unhealthy snacks. Second, the inclusion of unhealthy snacks alongside healthy snacks in the choice-based promotions significantly influenced how consumers perceived the healthy snacks. Facing a choice between both options may have triggered feelings of guilt associated with choosing the unhealthy option, making the healthy snacks appear more virtuous and appealing. This behavior aligns with the literature, which suggests that when utilitarian options (e.g., healthy food) and hedonic alternatives (e.g., unhealthy food) are presented together, consumers are more likely to favor the utilitarian choice as a way to justify their decision (Okada 2005). This dynamic enhanced the appeal of both healthy and unhealthy options, ultimately driving higher sales during choice-based promotions compared to healthy promotions.

We also repeated the analysis by excluding the weekend observations (Appendix D.1) and excluding the days where stockouts occurred (Appendix D.2) and observed consistent results.

4.4 Impact on revenue and profit

In this section, we analyze the impact of the different promotion strategies on daily revenue and profit, focusing on the transactions that involve the three product categories included in our experiment, namely coffee, healthy snacks, and unhealthy snacks. Revenue refers to the total dollar amount generated from the sales transactions within the store on a specific day. Profit represents the financial gain for the retailer, calculated as the revenue minus the total costs associated with the products sold.[4] The cost values for all the products remained constant throughout our experiment. Table 6 presents the results of the pairwise t-tests on revenue and profit across the seven promotions implemented in our experiment (Xu et al. 2024). For each pairwise comparison (A–B), we report the percentage change in revenue and profit for B relative to A along with the p-value. We naturally focus on the pairwise comparisons that are relevant for our analysis.

Table 6: Pairwise comparisons of the revenue and profit for the different promotions.

		Revenue		Profit	
		% change	p-value	% change	p-value
No Promotion - Promotion	C-T1	-19.78%	0	-17.46%	0
	C-T2	-27.63%	0	-30.19%	0
	C-T3	79.41%	0	72.28%	0
	C-T4	44.46%	0	38.47%	0
	C-T5	112.46%	0	160.46%	0
	C-T6	55.06%	0	46.17%	0
Unhealthy bundle - Unhealthy discount	T1-T2	-	1.00	-	0.06
Unhealthy bundle - Healthy bundle	T1-T3	123.65%	0	108.73%	0
Unhealthy bundle - Choice bundle	T1-T5	164.86%	0	215.57%	0
Unhealthy discount - Healthy discount	T2-T4	99.60%	0	98.34%	0
Unhealthy discount - Choice discount	T2-T6	114.25%	0	109.38%	0
Healthy bundle - Healthy discount	T3-T4	-19.48%	0	-19.62%	0
Healthy bundle - Choice bundle	T3-T5	18.42%	0.01	51.19%	0
Healthy discount - Choice discount	T4-T6	-	0.24	-	0.35
Choice bundle - Choice discount	T5-T6	-27.02%	0.01	-43.88%	0

Note: We report the Bonferroni adjusted p-value here.

No Promotion vs. Promotions: We found that the unhealthy promotions (T1 and T2) resulted in lower revenue and profit than the no-promotion condition (C). This is explained by the fact that these promotions did not significantly increase the sales of unhealthy snacks (as discussed in §4.3), yet the discounts offered led to a considerable reduction in both revenue and profit. Specifically, revenue (resp. profit) declined by 19.78% (17.46%) and 27.63% (30.19%) under T1 and T2, respectively, relative to the no-promotion condition (C). In contrast, both healthy promotions (T3 and T4) and choice promotions (T5 and T6) led to a significantly higher revenue and profit. Specifically, revenue (resp. profit) increased by 79.41% (72.28%) and 44.46% (38.47%) under T3 and T4, respectively, relative to the no-promotion condition (C). In the case of T3, we can

[4] We obtained the cost information directly from the retailer who informed us that it includes both purchase and transportation costs.

explain the increase in revenue and profit by the increase in the sales of healthy snacks. Although there is a slight decline in the sales of unhealthy snacks (see Table 5), the substantial increase in healthy snack sales combined with that of the coffee sales, more than compensates for this decline, ultimately boosting both revenue and profit. For T4, the higher sales of healthy snacks combined with the unaffected sales of unhealthy snacks led to a notable increase in both revenue and profit. Similarly, for the choice promotions, revenue (resp. profit) increased by 112.46% (160.46%) and 55.06% (46.17%) under T5 and T6, respectively, relative to the no-promotion condition (C). This result is driven by the significant increase in the sales of healthy snacks under both choice promotions, while the sales of unhealthy snacks remained stable, hence leading to higher revenue and profit. We also found that the choice bundle (T5) generated the highest percentage revenue and profit increases compared to all other conditions, providing support to Hypothesis 2b, which states that the choice bundle will generate higher revenue and profit compared to having no promotion and to all other promotions (C, T1, T2, T3, T4, T6). The superior performance of the choice bundle can be attributed to its ability to boost the sales of bundled items. This option drives a significant increase in revenue by encouraging customers to purchase additional products—whether healthy or unhealthy—that they may not have originally intended to buy. Furthermore, the strategic inclusion of a high-margin product (in our case, coffee), enhances profitability, hence making the choice bundle the most effective strategy to increase profit.

Unhealthy bundle (T1) vs. Other promotions (T2, T3, T5): We found no significant differences in revenue and profit between the unhealthy bundle (T1) and the unhealthy discount (T2). This does not support Hypothesis 2a, which states that bundling generates higher revenue than discounting. This is likely because there was no substantial change in the sales for either unhealthy or healthy snacks under these promotions. In contrast, both the healthy bundle (T3) and the choice bundle (T5) performed significantly better in terms of revenue and profit than the unhealthy bundle (T1). Specifically, revenue (resp. profit) increased by 123.65% (108.73%) and 164.86% (215.57%) under T3 and T5, respectively, relative to T1. For T3, while the sales of healthy snacks exceeded those under T1, the sales of unhealthy snacks were lower. As discussed earlier, this is likely due to the increase in the number of healthy bundles sold, resulting in higher revenue and profit. For T5, this result is expected as this promotion significantly increased the sales of healthy snacks while not affecting the sales of unhealthy snacks.

Unhealthy discount (T2) vs. Other discount promotions (T4, T6): We observed that both the healthy discount (T4) and the choice discount (T6) outperformed the unhealthy discount (T2) in terms of revenue and profit. Specifically, revenue (resp. profit) increased by 99.60% (98.34%) and 114.25% (109.38%) under T4 and T6, respectively, relative to T2. This is due to the significantly higher sales of healthy snacks under T4 and T6, while keeping the sales of unhealthy snacks relatively unchanged.

Healthy bundle (T3) vs. Other promotions (T4, T5): We observed that the healthy discount (T4) generated lower revenue and profit than the healthy bundle (T3). Specifically, the revenue (resp. profit) decreased by 19.48% (19.62%) under T4 relative to T3. While T4 resulted in higher sales of both healthy and unhealthy snacks, the bundling strategy in T3 drove additional coffee sales, thereby leading to a higher revenue and profit. This supports Hypothesis 2a that bundling promotions generate higher revenue and profit than discounting, and underscores the importance of strategic choices of products, e.g., coffee, in bundling. We also found that the choice bundle (T5) outperformed the healthy bundle (T3) in terms of revenue and profit. Specifically, the revenue (resp. profit) increased by 18.42% (51.19%) under T5 relative to T3. This result can be attributed to the higher sales of both healthy and unhealthy snacks under T5. As discussed before, the flexibility of the choice bundle allowing consumers to select between healthy and unhealthy options plays a key role in its success. By appealing to a broader range of consumers, the choice bundle attracts both health-conscious individuals and those seeking indulgent treats. This wider reach led to higher sales and, ultimately, resulted in greater revenue and profit.

Healthy discount (T4) vs. Choice discount (T6): We found no significant differences in revenue and profit between the healthy discount (T4) and the choice discount (T6). This outcome can be explained by the similar increase in the sales of healthy snacks under both promotions, combined with the stable sales of unhealthy snacks. Consequently, the revenue and profit for T4 and T6 were also similar. Interestingly, despite the higher flexibility offered by the choice discount, it did not lead to a notable advantage in terms of revenue or profit, hence showcasing the comparable effectiveness of these two promotion strategies.

Choice promotions (T5 and T6): We found that the choice discount (T6) generated lower revenue and profit compared to the choice bundle (T5). Specifically, the revenue (resp. profit) decreased by 27.02% (43.88%) under T5 relative to T6. While T6 achieved a higher percentage increase in the sales of healthy snacks, the bundling strategy in T5 offered a distinct advantage by significantly boosting the bundle sales, hence leading to a higher revenue. Once again having the high-margin, full priced coffee as part of the bundle strongly contributed to profit. This supports Hypothesis 2a that bundling promotions generate higher revenue and profit than discounting.

In summary, unlike our findings in Section 4.3, bundling—especially the choice bundle (T5) and, to a lesser extent, the healthy bundle (T3)—was more effective at increasing revenue and profit than discounting. Although discount promotions (T4 and T6) also led to an increase in revenue and profit relative to the control, bundling promotions (T3 and T5) were clearly superior. Interestingly, unhealthy promotions (T1 and T2), which did not significantly affect the sales, resulted in substantial declines in revenue and profit. This finding will be further explored in Section 6.

5 Additional analyses

In this section, we conduct three supplementary analyses to showcase the robustness of our findings and uncover additional insights. First, we apply the DID methodology to validate our results. Second, we analyze the impact of promotions on three key store performance metrics. Finally, we assess potential spillover effects by examining how the promotions influenced the sales of product categories beyond those directly included in the experiment.

5.1 Difference-in-differences analysis

To showcase the robustness of our estimates from Section 4.3, we employ a DID regression model, which quantifies the impact of each promotion strategy on the sales by comparing the changes over time between treatment and control stores. The model specification is as follows (Cui et al. 2019):

$$Y_{dps} = \alpha_s + \beta_T Treat_s + \sum_{I \in \{T_1,T_2,T_3,T_4,T_5,T_6\}} (\beta_I \cdot I + \gamma_I \cdot I \cdot Treat_s) + \mu_d + SO_{dps} + \epsilon_{dps}. \quad (2)$$

In Eq. (2), $Treat_s = 1$ for the treated stores and $Treat_s = 0$ for the control store. Since we are interested in quantifying the treatment effects, we focus on the interaction coefficients $\gamma_{T_1}, \gamma_{T_2}, \gamma_{T_3}, \gamma_{T_4}, \gamma_{T_5}, \gamma_{T_6}$, measuring the differential impact of each promotion on the treated stores relative to the control one. These coefficients allow us to isolate the effect of each intervention, providing insights into how each promotion influences the sales of healthy and unhealthy products. This ensures that we account for both temporal and cross-sectional variations, thereby robustly estimating the effect of each promotion used in our experiment. The validity of the DID method itself relies on the parallel trends assumption, which states that in the absence of treatment, the treated and control groups should exhibit similar trends over time. To confirm this, we visually examined the average weekly sales for both product categories in treated and control stores prior to the experiment and noted that there are no significant deviations (see Appendix F.1).

The estimates for both unhealthy and healthy snacks are reported in Table C3. The results align closely with those discussed in Section 4.3, confirming the conclusion that discount-based promotions are more effective than bundling in terms of increasing the sales of healthy snacks, while unhealthy promotions are ineffective. For additional robustness check, we also repeated the DID analysis by excluding the weekend observations (Appendix D.1) and excluding the days where stockouts occurred (Appendix D.2) and observed consistent results.

5.2 Store performance metrics

In this section, we examine the impact of the promotion strategies on three store performance metrics: average number of daily transactions, average basket size, and average daily total revenue (from all products). We estimate the model from Eq. (1) and report the treatment effect estimates in Table 7. First, we note that the average number of daily store transactions remained unchanged across all

Table 7: Estimates of the effect of the different promotions on store performance metric: Preliminary analysis.

Product promoted	Promotion type	Number of transactions	Basket size	Store revenue
Unhealthy	Bundle (T1)	-0.02 (0.06)	-0.014 (0.04)	-0.091 *** (0.02)
	Discount (T2)	0.03 (0.06)	0.03 (0.04)	-0.10 *** (0.02)
Healthy	Bundle (T3)	-0.04 (0.06)	0.47 *** (0.04)	0.05 ** (0.02)
	Discount (T4)	0.02 (0.06)	0.42 *** (0.04)	0.06 ** (0.02)
Choice	Bundle (T5)	0.002 (0.06)	0.54 *** (0.04)	0.05 *** (0.02)
	Discount (T6)	0.01 (0.06)	0.50 *** (0.04)	0.04 ** (0.02)
No. Obs.		735	735	735
Log likelihood		-4410.6	162.77	457.59
Model		NB	TWFE	TWFE

Note: $^{*}p < 0.05$;$^{**}p < 0.01$;$^{***}p < 0.001$.
The standard errors are reported in parentheses.
No. Obs. = 21 days per treatment $\times 5$ treatment periods $\times 7$ stores = 735.

promotions. This outcome can likely be attributed to the fact that the promotions were primarily noticed by regular customers who frequented the stores during our experiment, as the promotions were not advertised on broader platforms like social media or via large-scale campaigns. It is worth noting that C-store customers tend to be habitual customers who shop regularly to meet specific needs.[5] Since these promotions were not advertised to a wider audience, it makes sense that they did not draw additional traffic. Instead, they altered the spending patterns of the regular customers (e.g., by encouraging them to purchase promoted products or additional items).

The average basket size, defined as the average number of items purchased per transaction, increased during both healthy promotions (T3 and T4) and choice promotions (T5 and T6). Specifically, we found an increase of 11.3%, 10.16%, 13%, and 12.26% under T3, T4, T5, and T6 respectively, compared to the no-promotion condition (C). The increase in basket size during bundle-based promotions (T3 and T5) is expected since the whole idea of bundling is to encourage customers to purchase multiple items. More interesting is the significant increase in basket size under discount-based promotions (T4 and T6). This suggests that discounting a product may encourage customers to use those savings to make cross-category purchases and increase their basket. Thus, by lowering the price of targeted items, the retailer can incentivize customers to explore and purchase additional items within the store (we will elaborate on this point in Section 5.3). However, it is worth noting that unhealthy promotions (T1 and T2) had no significant effect on the average basket size. This suggests that consumers visiting the store to buy unhealthy products already have a clear idea of what they intend to purchase. As

[5] https://aytm.com/post/convenience-store-shopper-insights

a result, the promotions neither influenced the sales of the targeted unhealthy items (as discussed in Section 4.3) nor altered the overall basket composition.

Lastly, when evaluating the average daily total store revenue, calculated as the overall number of units sold multiplied by their respective selling prices for all products, we observed a significant decline of 8.70% and 9.83% under unhealthy promotions (T1 and T2) compared to the no-promotion condition (C). These results align with the trends discussed in Section 4.4, where T1 and T2 also showed a substantial decrease in revenue from the three product categories in our experiment. The lack of any compensatory revenue from cross-category purchases (as discussed above) means that these promotions fail to drive additional sales, hence resulting in an overall net deficit. In contrast, both healthy and choice promotions (T3–T6) led to a notable increase in the average daily store revenue. Specifically, it rose by 5.13%, 6.18%, 5.12%, and 4.08% under T3, T4, T5, and T6, respectively, compared to the no-promotion condition (C). These findings also align with the results of Section 4.4. The revenue growth observed during bundle promotions (T3 and T5) can be attributed to customers opting for bundles, which naturally increased the transaction value. The revenue increase during discount promotions (T4 and T6) most likely stem from the the positive impact of these promotions on cross-category sales as explained before. Since, in this section, we examine the revenue impact for the entire store (rather than only focusing on the products included in our experiment), the percentage increase or decrease is naturally much lower.

An interesting insight from this section is the fact that the promotions did not only affect the purchasing behavior for the promoted products but also drove sales from alternative product categories. We dive deeper into such cross-category spillover effects in Section 5.3 below.

5.3 Impact on sales of other product categories

This section examines the possible spillover effects of the six promotion strategies, specifically how they affect the sales of products beyond the three focal categories included in our experiment (healthy snacks, unhealthy snacks, and coffee). Understanding these dynamics is critical for retailers when designing effective holistic promotion strategies. For this analysis, we focused on the subset of transactions that included at least one product from one of the three focal categories. We analyzed these transactions to determine the frequency of purchase of the alternative categories alongside the focal products in our experiment. The top three most frequently purchased product categories were beverages other than coffee (e.g., juices, energy drinks, flavored and plain water), candies, and cigarettes (see Table D8). Interestingly, none of the other healthy products offered in the store, such as salads, fruits, and yogurt, appeared on the most frequently purchased list. However, since one of the objectives of this study is to explore the impact of promotions on healthy food purchasing patterns, we included the “other healthy products” category in our analysis. This resulted in four product categories being considered for our spillover analysis, as shown in Table 8. We used the model from Eq.(2) to assess the impact of the various promotion strategies (T1–T6) on the sales of the four additional product categories, namely, other healthy products, other beverages, candies, and cigarettes. Table 8 reports the estimates for each category.

Three main insights emerge from our analysis. First, offering discounts on healthy options encourage customers to use these savings to purchase other healthy (non-promoted) products. Specifically, we found that healthy and choice discounts (T4 and T6) increased the sales of other healthy products by 36.34% and 41.91%, respectively. Discounts on healthy snacks likely enhanced the perceived value of other healthy products, hence making customers more inclined to explore and purchase such products. In contrast, we observed no significant effect on the sales of other healthy products due to either unhealthy promotions (T1 and T2) or bundle promotions (T3 and T5). Indeed, unhealthy promotions are geared towards a customer segment that is set in its purchase intentions (as discussed before), while bundle promotions focus on predefined product assortments, hence limiting their ability to increase the sales of products outside the bundle.

Table 8: DID estimates for the different promotions using the sales of other products.

Product promoted	Promotion type	Other healthy products	Other beverages	Candies	Cigarettes
Unhealthy	Bundle (T1)	0.33 (0.12)	-0.04 (0.03)	-0.08 (0.04)	0.06 (0.03)
	Discount (T2)	0.37 (0.12)	-0.002 (0.03)	-0.04 (0.04)	0.02 (0.03)
Healthy	Bundle (T3)	0.04 (0.09)	-0.07 ** (0.03)	0.1 *** (0.03)	0.08 ** (0.03)
	Discount (T4)	0.31 *** (0.07)	-0.01 (0.02)	0.14 *** (0.03)	0.06 * (0.03)
Choice	Bundle (T5)	0.08 (0.17)	-0.08 ** (0.03)	-0.11 (0.04)	-0.04 (0.03)
	Discount (T6)	0.35 ** (0.14)	-0.03 (0.03)	-0.03 (0.04)	0.03 (0.03)
No. Obs.		735	735	735	735
Log likelihood		-624.76	-2,842.6	-2,477.9	-2,176.8
Model		NB	NB	NB	NB

Note: $^{*}p < 0.05$;$^{**}p < 0.01$;$^{***}p < 0.001$.
The standard errors are reported in parentheses.
No. Obs. = 21 days per treatment $\times 5$ treatment periods $\times 7$ stores = 735.

Second, we found that bundle (coffee) promotions significantly cannibalized the sales of other beverages (juices, energy drinks, flavored and regular water). Specifically, the sales decreased by 7.25% and 8.32% under the healthy bundle (T3) and the choice bundle (T5), respectively. In other words, the promoted coffee bundles encouraged customers to select coffee as their preferred beverage by taking advantage of the promotion, thereby reducing their inclination to buy other drinks. In contrast, we observed no significant effect on the sales of other beverages when coffee was not part of the promotion (T2, T4, and T6) and when the customers' purchase behaviors were not influenced by the promotion (T1). These findings highlight the inherent trade-offs associated with bundling strategies. While these promotions are effective at driving sales within the promoted bundle, they can cannibalize sales from substitute product categories.

Third, and interestingly, we found that healthy promotions incentivize customers to purchase unhealthy products. Specifically, we found that the sales of candies increased by 10.52% and 15.03% under the healthy bundle (T3) and the healthy discount (T4), respectively. Similarly, the sales of cigarettes increased by 8.33% and 6.18% under T3 and T4, respectively. In other words, by purchasing healthy products first, consumers may reduce their guilt, effectively giving themselves "moral license" to indulge in other less healthy options (Khan and Dhar 2006, Blanken et al. 2015). This dynamic underscores the unintended consequences of promoting healthy products, as it may inadvertently boost the sales of indulgent products. Moreover, we observed no significant effect on the sales of both candies and cigarettes either under unhealthy promotions (T1 and T2) or choice promotions (T5 and T6). This lack of impact for T1 and T2 could be due to the absence of a behavioral offset, as consumers purchasing unhealthy products likely already considered them indulgent and, thus, do not experience moral licensing effect. Similarly, choice promotions (T5 and T6), which simultaneously offer both healthy and unhealthy alternatives, prompt consumers to evaluate their selections more thoughtfully. By giving the choice between healthy and unhealthy products, choice promotions eliminate the sequential psychological process of guilt alleviation. Namely, consumers are prompted to make a conscious decision between healthy and indulgent items, rather than using one category to justify the consumption of the other. For robustness, we also estimated the treatment effects using only the weekday transactions and report the estimates in Table D9 and observed consistent results.

In summary, our analyses in this section led to three main findings: (1) coffee bundling promotions substantially cannibalize the sales of other beverages; (2) discount promotions that involve healthy products positively reinforce customers' tendency to buy other healthy products; (3) healthy promotions (both discounts and bundles) incentivize consumers to buy unhealthy products due to a moral licensing effect.

6 Follow-up experiment

One surprising and unexpected result from the two previous sections is the clear lack of impact of unhealthy promotions. In this section, we dig deeper into the potential rationale behind this finding. A likely contributing factor is the C-store chain's long-standing practice of bundling unhealthy products (T1; see Figure 1a), a strategy that has been in place for several years across many of its stores. While we strove to reset the baseline by removing the promotion before the start of our experiment, it is plausible that the long-term effects of this promotion are still lingering. Thus, to evaluate how prolonged exposure to such promotions influences consumer behavior, we conducted an additional field experiment on 40 stores, as described in Section 3.3, while excluding the seven stores used in the initial experiment. From this subset of 33 eligible stores, we randomly selected 10 stores to evaluate the impact of discontinuing long-term promotions on purchasing behavior. This analysis aims to explore whether the prolonged unhealthy promotion had resulted in consumer habituation and, consequently, alleviated the impact of subsequent (unhealthy) promotions.

Table 9: Follow-up experiment: dates and promotions.

Experiment phase	Date range	Number of weeks	Promotion offered
Promotion (pre-intervention)	06/10/24 to 08/05/24	8	Unhealthy bundle (Figure 1a)
Promotion disruption (T)	08/07/24 to 10/02/24	8	No promotion
Promotion (post-intervention, PI)	10/03/24 to 11/28/24	8	Unhealthy bundle (Figure 1a)

The follow-up experiment spanned 24 weeks, as outlined in Table 9. We first paused the long-standing promotion (T1), which had been active in all the stores for several years prior to the experiment, for eight weeks from August 7, 2024 to October 2, 2024. After this eight-week reset period, we resumed the unhealthy bundle promotion (T1) in all 10 treated stores to conduct a post-treatment analysis. Our main econometrics specification is DID combined with matching methods (Cui et al. 2020). For robustness purposes, we also use ANOVA and regression analyses (see Appendix E.1).

We estimated a DID model to assess the impact of discontinuing the unhealthy bundle promotion on all our key performance metrics: sales, average number of daily transactions, average basket size, revenue, and profit. To enhance the validity of the DID model, we first applied Propensity Score Matching (PSM) to identify a control group of stores that closely resembled the treated stores across relevant characteristics (Rosenbaum and Rubin 1983, Han et al. 2019). The detailed methodology and results of the matching process are provided in Appendix E.2. We then used the matched stores to estimate the DID model based on the following specification:

$$Y_{dps} = \alpha_s + \beta_1 \cdot T + \beta_2 \cdot PI + \beta_T Treat_s + \gamma_1 \cdot T \cdot Treat_s + \gamma_2 \cdot PI \cdot Treat_s + \mu_d + SO_{dps} + \epsilon_{dps}, \quad (3)$$

where $T = 1$ when the promotion was discontinued and $T = 0$ for the remaining periods. Similarly, $PI = 1$ for the post-intervention periods and $PI = 0$ for the remaining periods. The key parameters in Eq. (3) are γ_1 and γ_2, which capture the effect of discontinuing the long-standing unhealthy bundle promotion and the effect of reactivating the unhealthy promotion post-intervention, respectively. Table 10 presents the DID estimates when using the matched dataset. Our analysis clearly confirms that discontinuing the unhealthy bundle promotion had no significant effect on key metrics, including the sales of unhealthy snacks, the average number of transactions, and the average basket size in the treated stores. Additionally, reintroducing the unhealthy promotion post-intervention also showed no significant impact on those metrics. These findings align with our earlier results, which indicated no significant effect on neither the sales of unhealthy snacks (Section 4.3) nor on the overall store performance metrics (Section 5.2), supporting Hypothesis 3. This consistency across multiple analyses underscores the inelastic nature of demand for unhealthy products, suggesting that habitual purchasers of such products are sticky regardless of the offered promotion. Furthermore, the absence of impact on the number of transactions and on the basket size implies that the promotion discontinuation did not disrupt the broader shopping behavior in these stores. This reinforces the intuition that a long-term exposure to a promotion may create consumer habits that persist even in the absence of promotion incentives. Moreover, reactivating the promotion post-intervention did not result in any significant changes, hence highlighting the limited efficacy of such promotions.

We also observed that revenue and profit increased by 10.62% and 8%, respectively, in the treated stores following the discontinuation of the unhealthy bundle promotion. This finding is consistent with our earlier result (Section 4.4) and supports Hypothesis 3, which highlighted that the no-promotion scenario led to higher revenue and profit relative to the unhealthy bundle promotion. This increase in revenue and profit can be attributed to the fact that while the sales level remained unchanged, the prices of unhealthy snacks reverted to their regular, non-promotional levels. Thus, without the discounts associated with the promotion, each sale of unhealthy snacks generated higher revenue and profit per unit, thereby improving the store's overall financial performance. In addition, we found that the revenue and profit after reactivating the promotion remained unchanged with respect to the pre-intervention, highlighting once again the ineffectiveness of unhealthy promotions.

To validate the parallel trends assumption, we plotted the average weekly sales for the unhealthy snacks in Appendix F.2 and observed a clear parallel trend, ensuring the validity of the DID estimates. Overall, the findings from our follow-up experiment greatly align with the earlier analyses. These results collectively show the lack of impact of unhealthy promotions on purchase behavior after long-standing exposure and, consequently, the negative financial impact of such promotions.

For additional robustness check, we repeated the above analysis by excluding the weekend observations and excluding the days where stockouts occurred and observed consistent results. We report the results in Appendix E.3 and E.4.

Table 10: DID regression with PSM – Impact of discontinuing the long-term bundle promotion.

	Unhealthy sales	Number of transactions	Basket size	Revenue	Profit
Promotion discontinuation (T)	0.06 (0.09)	-0.01 (0.03)	-0.01 (0.01)	14.27 *** (4.12)	13.76 *** (2.69)
Post Intervention (PI)	-0.01 (0.10)	0.04 (0.03)	-0.02 (0.01)	2.62 (4.92)	4.82 (3.21)
No. Obs.	3,360	3,360	3,360	3,360	3,360
Log likelihood	-12,957.	-17,579.	-2,888.2	-15,307.	-14,059.
Model	NB	NB	TWFE	TWFE	TWFE

Note: $^{*}p < 0.05$;$^{**}p < 0.01$;$^{***}p < 0.001$.
The standard errors are reported in parentheses.
No. Obs. = 24 weeks of experiment ×7 days per week ×20 stores = 3,360.

7 Conclusion

This paper examines the comparative effectiveness of the two of the most commonly used retail promotion strategies—discounting and bundling—and explores how their impact varies across two product categories (healthy and unhealthy snacks). To study this, we conducted a large-scale field experiment in collaboration with a leading convenience store chain. The experiment employed a five-period crossover design, implementing six distinct promotion strategies across seven stores (one for control and six for treatment) over 18 weeks. These strategies included both discounts and bundling, targeting healthy and unhealthy snacks, as well as a choice promotion allowing the consumers to choose between the two product categories. The experiment aimed to rigorously assess the impact of these interventions on consumer behavior and on the store's revenue and profit.

Our findings provide actionable recommendations for retailers seeking to optimize their promotion strategies to drive sales, revenue, and profit, particularly in the context of healthy and unhealthy product categories. First, for healthy products, we found that bundling was more effective than discounting in terms of revenue and profit. Bundling not only encouraged healthier purchases but also enhanced profitability, particularly when paired with a high-demand, high-margin item like coffee. Nonetheless, discounting was favorable when the objective is to primarily boost sales. Therefore, retailers should use discounts sparingly, focusing on short-term objectives such as inventory clearance, while leveraging bundling as a strategic tool to balance healthier purchases and higher revenue and profit. Additionally, our findings indicated that healthy promotions can paradoxically boost the sales of unhealthy items such as candies and cigarettes. This suggests that consumers who purchase healthy snacks may engage in moral licensing, perceiving their healthy choice as justification to subsequently indulge in less healthy options.

Interestingly, the impact of the promotions was quite different for unhealthy products. Our analysis revealed no increase in the sales of unhealthy snacks when these items were promoted—under both bundling and discounting. In fact, these promotions were detrimental to the retailer's revenue and profit. This was unexpected given the typical positive impact of promotions on sales. A potential explanation relies on the fact that the sales of unhealthy snacks is relatively inelastic, with consumers showing no responsiveness to promotions. This may be especially true when they were exposed to a long-standing promotion on these products before the experiment, which may have created purchasing habits. To better understand this finding, we convinced the retailer to conduct a follow-up experiment. For this experiment, we expanded the study to cover 10 new stores over an eight-week period. This follow-up experiment strongly confirmed our result that promotions have no impact on the sales of unhealthy snacks and adversely affect revenue and profit.

While our research provides valuable insights into the effectiveness of bundling versus discounting for healthy and unhealthy products, there are admittedly limitations to note. Our analysis primarily focused on short-term outcomes, without examining potential long-term effects such as customer loyalty and retention. Moreover, while we focused on specific categories of healthy and unhealthy snacks, other product categories might respond differently, hence warranting further exploration (we noted spillover effects on other products). In conclusion, future research could build upon our findings by considering a broader range of product categories and investigating both short-term and long-term effects on consumer behavior.

A Stores selection

To ensure a robust selection of the stores for the experiment, our primary objective was to validate that the customer behavior across all selected experimental stores was as similar as possible. Initially, we had a pool of 90 stores from our C-store chain partner all located in the same city. To create a more uniform comparison basis, we first excluded the stores located in a gas station, reducing the pool to 46 stores. We then eliminated six stores that lacked sufficient inventory of the focal products in our experiment (healthy snacks, unhealthy snacks, and coffee), leaving us with a pool of 40 stores. From this final pool, we randomly selected seven stores, assigning one as the control store and six as treatment stores.

Before initiating the experiment, we conducted an A/A analysis to rigorously assess whether the selected stores were comparable across several key consumer-related metrics such as number of transactions, average basket size, and revenue. This analysis relied on data from the 46 weeks leading up to the experiment. Each week during this pre-experiment period was virtually considered as a fictitious treatment, and we estimated weekly treatment effects to evaluate the similarity between the control and treatment stores. The analysis was performed using the following model (Jalali et al. 2023):

$$Y_{ds} = \alpha_s + Time_d + \beta \cdot Treat_s + \gamma \cdot Time_d \cdot Treat_s + \mu_d + \epsilon_{ds},$$

where Y is the one of the following: number of transactions, average basket size, sales and revenue from the three main product categories used in the experiment (healthy snacks, unhealthy snacks, and coffee). $Time_d$ is the number of days leading up to the experiment during the pre-treatment period from day d, and $Treat_s$ takes the value 1 for the treated stores and 0 for the control store. The interaction term γ quantifies the difference between treated and control stores, capturing any systematic effects arising from the fictitious treatments. We also accounted for time-fixed effects (μ_d), including day-of-week variations and week numbers during the treatment period.

We estimated the model for the different key metrics using different approaches based on the nature of the data. For count-based metrics, such as the average daily number of coffee transactions, healthy snacks, unhealthy snacks, and total number of transactions in the store, we employed a Negative Binomial (NB) regression. For revenue, a log transformation was applied to ensure normal distribution of the data. We used Two-way fixed effects (TWFE) regression to estimate revenue and basket size.

The rolling analysis spanned the period from December 19, 2022, to November 5, 2023, utilizing a three-week window before and after each test event. This design led to a total of 40 placebo tests. In Figure C1, each point represents one of the 40 placebo tests (in chronological order) across different metrics. For each test, we report the estimate for γ along with the corresponding 95% confidence interval.

Our results show that, for the majority of the tests, the confidence intervals include zero, implying that there is no statistically significant evidence to suggest that the treated stores differ from the control store in terms of the examined metrics. This supports the point that the selected stores are comparable, thereby providing a solid foundation for the subsequent experimental phase.

B Reset period analysis

In the main field experiment, we used three weeks as a reset period to form a neutral baseline of having no promotion and compared our six promotions relative to this baseline. To ensure that the number of weeks we selected is sufficient for the consumers to form an accurate baseline and reset their preferences, we analyze how different reset periods, namely, three, six, nine, 12, and 15 weeks affect consumer preferences for the product that is not promoted anymore. To do so, we use the data from Store 7, which has no promotions for a period of 15 weeks. We use C1 to correspond to the interval from the start of the experiment to three weeks, C2 from three to six weeks, C3 from six to nine weeks, C4 from nine to 12 weeks, and C5 from 12 to 15 weeks. We found that there is no significant change in consumer purchase behavior during any of the five intervals as seen in Table C1.

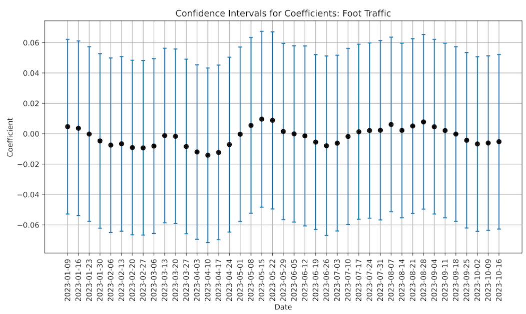
C Main experiment

C.1 A/A test

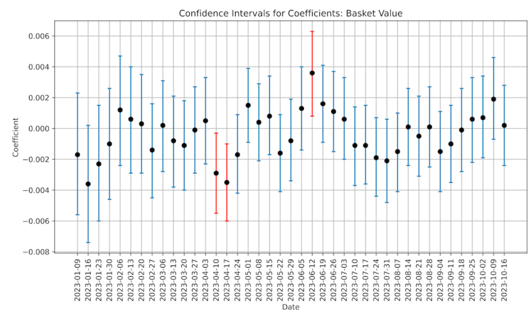
The A/A tests in this section are conducted using the three control (C) periods of the experiment from both the control and treatment stores. It is important to note that there is no promotion running during the control period (C). The dataset represents a panel containing daily observations in the seven stores for three control periods of 21 days each, totaling 63 days. We use the following model specification:

$$Y_{ds} = \alpha_s + \beta_1 Time_d + \beta_2 Treat_s + \gamma_C \cdot Time_d \times Treat_s + \mu_d + \epsilon_{ds}, \tag{4}$$

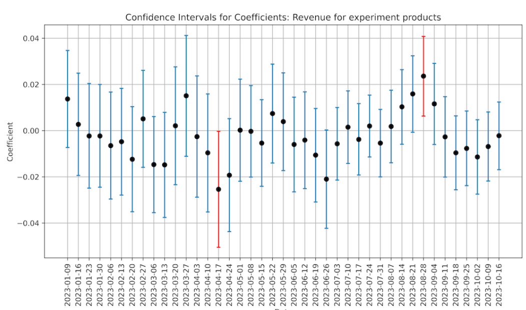
where Y is the basket size, revenue, number of transactions, coffee sales, healthy sales, and unhealthy sales (daily per store). We transform revenue using the natural logarithm of the average dollar sales from a store on a given day. We then use a TWFE regression for basket size and revenue variables and a NB regression for the number of transactions, coffee sales, healthy sales, and unhealthy sales. $Time_d$ counts the number of days from the experiment start date to day d, $Treat_s$ takes the value 1 for the treated stores and 0 for the control store. We also included other time-fixed effects, denoted by μ_d, and store-fixed effects, represented by α_s. In Eq. (4), if γ is statistically significant, it would imply that the trends in our metrics are different between the treated stores and the control store so that there is a potential stickiness effect after discontinuing the promotion. We report the estimates of γ in Table C2. We found that all the estimates are not statistically significant, thus indicating that the A/A tests passed successfully. This demonstrates that the treated and control stores have similar characteristics, such as basket value, revenue, number of transactions, and number of coffee, healthy and unhealthy transactions during the three control periods as shown in Figure 3. This also indicates that there is no significant stickiness behavior once the promotion is discontinued in the stores.



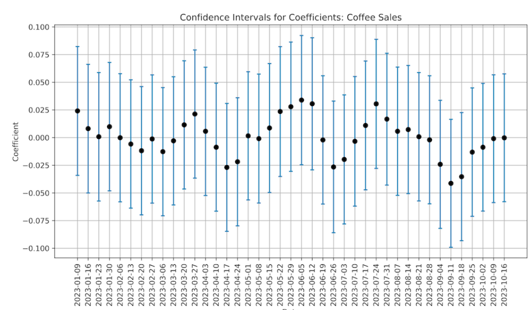
(a) Number of transactions



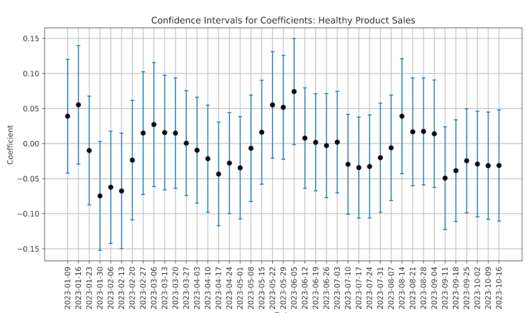
(b) Basket Value



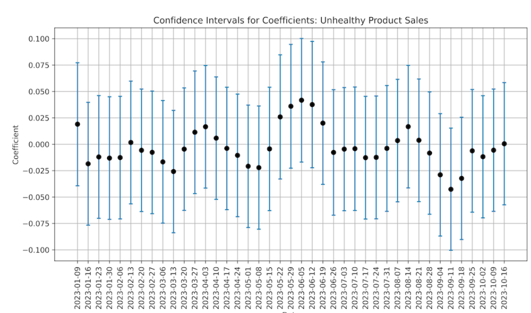
(c) Revenue



(d) Coffee sales



(e) Healthy product sales



(f) Unhealthy product sales

Figure C1: Placebo test across different metrics.

Table C1: Comparing the effect of different reset periods.

	Unhealthy snacks	
	Mean difference	p-value
C1-C2	3.91	1.00
C1-C3	6.33	0.33
C1-C4	0.57	1.00
C1-C5	6.62	0.25
C2-C3	2.43	1.00
C2-C4	-3.33	1.00
C2-C5	2.71	1.00
C3-C4	-5.76	0.81
C3-C5	0.29	1.00
C4-C5	6.05	0.65

Table C2: Estimates for A/A tests.

	Number of transactions	Basket value	Revenue	Coffee sales	Healthy prod. sales	Unhealthy prod. sales
Time X Treat	0.0004 (0.001)	0.0002 (0.00)	0.0004 (0.001)	0.0017 (0.001)	-0.0047 (0.003)	0.0014 (0.001)
No. Obs.	441	441	441	441	441	441
Log likelihood	-2,640.7	851.28	395.96	-2,039.9	-763.21	-1,997.9
Model	NB	TWFE	TWFE	NB	NB	NB

Note: $*p < 0.05$;$**p < 0.01$;$***p < 0.001$. The standard errors are reported in parentheses.
No. Obs. = 21 days per Control period ×3 Control periods ×7 stores=441.

C.2 Difference-in-differences

Table C3: DID estimates of the different promotions on sales.

Product promoted	Promotion type	Unhealthy sales	Healthy sales
Unhealthy	Bundle (T1)	-0.01 (0.02)	0.17 (0.12)
	Discount (T2)	0.04 (0.02)	0.20 (0.12)
Healthy	Bundle (T3)	-0.05 ** (0.02)	0.12 ** (0.04)
	Discount (T4)	0.01 (0.02)	0.33 *** (0.03)
Choice	Bundle (T5)	0.01 (0.02)	0.30 *** (0.07)
	Discount (T6)	0.03 (0.02)	0.36 *** (0.07)
No. Obs.		735	735
Log likelihood		-3,270.8	-1,360.9
Model		NB	NB

Note:$*p < 0.05$;$**p < 0.01$;$***p < 0.001$.
The standard errors are reported in parentheses.
No. Obs. = 21 days per treatment ×5 treatment periods ×7 stores = 735.

D Robustness tests: Main experiment

In this section, we conduct a series of robustness tests by repeating the analyses from Sections 4.3, 5.1, and 5.2, while excluding both weekend data and days with product stockouts. Shopping behavior may differ significantly between weekdays and weekends, as weekend shoppers often have different purchasing patterns and motivations. By restricting the analysis to weekdays, we ensure that our results are not biased by these behavioral differences. Additionally, we exclude days when products were out of stock to prevent distortions in the findings, as stockouts could artificially suppress sales and misrepresent the true impact of the promotion strategies. These adjustments help verify the robustness of our results and confirm that our conclusions remain valid more generally. We find that, in nearly all cases, the results remain consistent with our main analysis. This reinforces the robustness of our conclusions, suggesting that our key insights hold even when excluding weekend data and accounting for stockouts.

D.1 Analyses on weekday sales

In this section, we repeat the analysis performed in Sections 4.3 and 5.1 using only weekday transactions.

Table D4: Impact of different interventions using only weekday transactions.

Product promoted	Promotion type	Unhealthy sales	Healthy sales	Number of transactions	Basket size	Store revenue
Unhealthy	Bundle (T1)	-0.04 (0.04)	0.25 (0.28)	-0.02 (0.07)	-0.02 (0.04)	-0.10 *** (0.03)
	Discount (T2)	0.04 (0.04)	0.30 (0.28)	0.02 (0.07)	0.05 (0.04)	-0.12 *** (0.03)
Healthy	Bundle (T3)	-0.11 *** (0.03)	0.22 ** (0.09)	-0.04 (0.07)	0.50 *** (0.04)	0.09 *** (0.03)
	Discount (T4)	0.05 (0.03)	0.67 *** (0.08)	0.02 (0.07)	0.41 *** (0.04)	0.08 *** (0.03)
Choice	Bundle (T5)	0.01 (0.04)	0.59 *** (0.15)	-0.01 (0.07)	0.56 ** (0.04)	0.02 *** (0.03)
	Discount (T6)	0.05 (0.04)	0.71 *** (0.14)	0.01 (0.07)	0.54 *** (0.04)	0.04 *** (0.03)
No. Obs.		525	525	525	525	525
Log likelihood		-2,503.5	-1,033.2	-3,148.7	151.06	350.77
Model		NB	NB	NB	TWFE	TWFE

Note: $*p < 0.05$;$**p < 0.01$;$***p < 0.001$. The standard errors are reported in parentheses.
No. Obs. = 15 weekdays per treatment $\times 5$ treatment periods $\times 7$ stores = 525.

Table D5: DID estimates for different interventions using only weekday transactions.

Product promoted	Promotion type	Unhealthy sales	Healthy sales	Number of transactions	Basket size	Store revenue
Unhealthy	Bundle (T1)	-0.03 (0.02)	0.10 (0.14)	-0.01 (0.03)	-0.01 (0.02)	-0.05 *** (0.01)
	Discount (T2)	0.01 (0.02)	0.11 (0.14)	0.01 (0.03)	0.02 (0.02)	-0.06 *** (0.01)
Healthy	Bundle (T3)	-0.03 * (0.02)	0.13 ** (0.04)	-0.02 (0.03)	0.25 *** (0.02)	0.04 ** (0.01)
	Discount (T4)	0.01 (0.02)	0.35 *** (0.04)	0.01 (0.03)	0.20 *** (0.02)	0.042 * (0.01)
Choice	Bundle (T5)	-0.01 (0.02)	0.31 *** (0.07)	-0.01 (0.03)	0.28 ** (0.02)	0.01 * (0.01)
	Discount (T6)	0.01 (0.02)	0.38 *** (0.07)	0.01 (0.03)	0.27 *** (0.02)	0.02 * (0.01)
No. Obs.		525	525	525	525	525
Log likelihood		-2,517.5	-1,048.3	-3,148.7	151.06	350.77
Model		NB	NB	NB	TWFE	TWFE

Note: $*p < 0.05$;$**p < 0.01$;$***p < 0.001$. The standard errors are reported in parentheses.
No. Obs. = 15 weekdays per treatment $\times 5$ treatment periods $\times 7$ stores = 525

D.2 Analyses excluding stockouts

In this section, we repeat the analysis performed in Sections 4.3 and 5.1 by removing the days on which product stockouts occurred. For the analysis on number of transactions and basket value, we remove the days on which stockouts occurred for either product category.

Table D6: Impact of different interventions using transactions without stockout days.

Product promoted	Promotion type	Unhealthy sales	Healthy sales	Number of transactions	Basket size	Store revenue
Unhealthy	Bundle (T1)	0.01 (0.05)	0.46 (0.26)	0.04 (0.10)	-0.01 (0.07)	-0.06 * (0.04)
	Discount (T2)	0.02 (0.05)	0.43 (0.26)	-0.01 (0.09)	0.10 (0.06)	-0.16 *** (0.04)
Healthy	Bundle (T3)	-0.06 * (0.03)	0.20 ** (0.08)	-0.03 (0.07)	0.50 *** (0.04)	0.09 ** (0.03)
	Discount (T4)	-0.01 (0.04)	0.66 *** (0.07)	-0.04 (0.09)	0.71 *** (0.06)	0.05 * (0.04)
Choice	Bundle (T5)	0.03 (0.04)	0.57 *** (0.14)	-0.04 (0.08)	0.58 ** (0.05)	0.01 (0.03)
	Discount (T6)	0.06 (0.04)	0.69 *** (0.13)	0.06 (0.08)	0.56 *** (0.05)	0.07 * (0.03)
No. Obs.		492	641	437	437	437
Log likelihood		-2,509.0	-1,258.8	-2,613.6	79.63	287.84
Model		NB	NB	NB	TWFE	TWFE

Note:* $p < 0.05$;** $p < 0.01$;*** $p < 0.001$. The standard errors are reported in parentheses.

Table D7: DID estimates for different interventions using transactions without stockout days.

Product promoted	Promotion type	Unhealthy sales	Healthy sales	Number of transactions	Basket size	Store revenue
Unhealthy	Bundle (T1)	0.01 (0.05)	0.23 (0.13)	0.02 (0.05)	-0.01 (0.03)	-0.03 * (0.02)
	Discount (T2)	0.09 (0.05)	0.21 (0.13)	-0.01 (0.05)	0.05 (0.03)	-0.08 *** (0.02)
Healthy	Bundle (T3)	-0.03 (0.03)	0.10 * (0.04)	-0.01 (0.03)	0.25 *** (0.02)	0.04 ** (0.01)
	Discount (T4)	-0.004 (0.05)	0.33 *** (0.04)	-0.02 (0.04)	0.36 *** (0.03)	0.02 * (0.02)
Choice	Bundle (T5)	0.02 (0.04)	0.29 *** (0.07)	-0.02 (0.04)	0.29 *** (0.03)	0.01 * (0.02)
	Discount (T6)	0.03 (0.04)	0.34 *** (0.07)	0.03 (0.04)	0.28 *** (0.03)	0.03 * (0.02)
No. Obs.		492	641	437	437	437
Log likelihood		-2,168.1	-1,258.8	-2,613.6	80.17	288.83
Model		NB	NB	NB	TWFE	TWFE

Note: * $p < 0.05$;** $p < 0.01$;*** $p < 0.001$. The standard errors are reported in parentheses.

D.3 Impact on other products

In this section, we repeat the analysis from Section 5.2 using only weekday transactions. We also show the frequencies of other product purchases during the experiment.

Table D8: Frequency of other products purchases.

Product category	C	T1	T2	T3	T4	T5	T6
Other healthy products	0.86%	0.49%	0.48%	1.72%	2.29%	0.66%	1.09%
Other beverages	26.82%	14.02%	18.13%	40.74%	37.07%	27.14%	26.32%
Candies	15.52%	8.10%	9.12%	30.14%	28.19%	14.28%	16.17%
Cigarettes	15.56%	15.18%	14.97%	14.94%	12.95%	21.47%	24.97%

Table D9: DID estimates for the different promotions for sales of other products using weekday transactions.

Product promoted	Promotion type	Other healthy products	Other beverages	Candies	Cigarettes
Unhealthy	Bundle (T1)	0.28	-0.04	-0.08	0.05
		(0.24)	(0.03)	(0.05)	(0.03)
	Discount (T2)	0.41	-0.002	-0.07	0.01
		(0.23)	(0.03)	(0.05)	(0.04)
Healthy	Bundle (T3)	0.09	-0.06 **	0.16 ***	0.13 **
		(0.13)	(0.02)	(0.03)	(0.03)
	Discount (T4)	0.39 **	0.01	0.17 ***	0.10 *
		(0.11)	(0.02)	(0.03)	(0.04)
Choice	Bundle (T5)	0.13	-0.13 ***	-0.06	-0.03
		(0.19)	(0.03)	(0.04)	(0.04)
	Discount (T6)	0.38 *	0.001	-0.03	0.05
		(0.17)	(0.03)	(0.04)	(0.03)
No. Obs.		525	525	525	525
Log likelihood		-493.81	-1,944.9	-1,864.3	-1,549.2
Model		NB	NB	NB	NB

Note: $^*p < 0.05$;$^{**}p < 0.01$;$^{***}p < 0.001$. The standard errors are reported in parentheses.
No. Obs. = 15 days per treatment ×5 treatment periods ×7 stores = 525.

E Follow-up experiment

E.1 Preliminary analysis

In this section, we estimate a slightly modified version of the model from Eq. (1), where we aim to estimate the effect of discontinuing the long-standing bundle promotion on unhealthy snacks (T1). We use the following NB regression specification, which is appropriate for count data like sales (Q_{dps}):

$$Y_{ds} = \alpha_s + \beta_1 \cdot T + \beta_2 \cdot PI + \mu_d + SO_{dps} + \epsilon_{ds}, \tag{5}$$

where

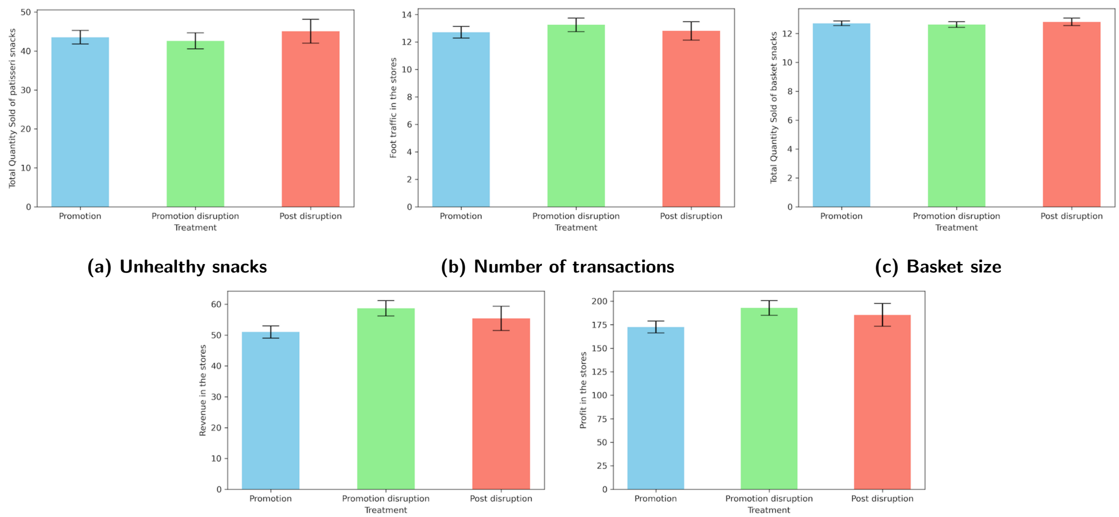
$$T = \begin{cases} 1, & \text{if the promotion is discontinued,} \\ 0, & \text{otherwise} \end{cases}$$

and

$$PI = \begin{cases} 1, & \text{if it is post intervention periods,} \\ 0, & \text{otherwise.} \end{cases}$$

In Eq. (5), β_1 measures the average treatment effect of discontinuing the long-standing bundle promotion and β_2 measures the average treatment effect of reactivating the unhealthy bundle promotion. Y represents the sales of unhealthy snacks, unhealthy bundles, number of transactions, basket size, revenue, and profit from unhealthy snacks and coffee. The key parameters in Eq. (5) are β_1 and β_2.

Figure E2 plots the sales of unhealthy snacks, number of transactions, average basket value, revenue, and profit from coffee and unhealthy snacks, each with a 95% confidence interval. These metrics are



(a) Unhealthy snacks **(b) Number of transactions** **(c) Basket size**

(d) Revenue **(e) Profit**

Figure E2: Average daily values before, during, and after the promotion discontinuation.

shown for each of the three experimental phases. Figures E2a, E2b, and E2c reveal statistically non-significant differences in unhealthy snack sales, number of transactions, and average basket size across the experimental phases. This stability suggests that discontinuing the unhealthy bundle promotion did not alter consumer purchasing patterns, as indicated by stable basket sizes and number of transactions. On the contrary, Figures E2d and E2e show a statistically significant increase in revenue and profit for the period when the promotion was discontinued.

Table E10 presents the estimated β coefficients from Eq. (5) for all our key variables. The results suggest that discontinuing the promotion did not significantly impact the sales of unhealthy snacks, as quantities sold remained consistent. Similarly, there is no noticable effect on either the number of transactions or the average basket size, indicating that consumers maintained similar purchasing behavior even in the absence of the promotion. Thus, this supports Hypothesis 3 and is aligned with our results from Section 4.

Table E10: Impact of discontinuing the long-standing bundle promotion.

	Unhealthy sales	Unhealthy bundle sales	Number of transactions	Basket size	Revenue	Profit
Treatment	0.05	-0.06	-0.09	-0.001	22.70 ***	19.01 ***
	(0.09)	(0.09)	(0.08)	(0.01)	(6.07)	(3.77)
Post Treatment	0.18	0.23	-0.13	0.014	11.92	3.72
	(0.13)	(0.15)	(0.13)	(0.01)	(9.43)	(5.85)
No. Obs.	1,680	1,680	1,680	1,680	1,680	1,680
Log likelihood	-4,920.2	-4,276.6	-9,850.1	-1,326.4	-7,763.3	-7,070.6
Model	NB	NB	NB	TWFE	TWFE	TWFE

Note:*$p < 0.05$;**$p < 0.01$;***$p < 0.001$. The standard errors are reported in parentheses.
No. Obs. = 24 weeks of experiment ×7 days per week ×10 stores = 1,680.

Interestingly, both revenue and profit significantly increased when the promotion was discontinued. More precisely, the revenue (resp. profit) increased by 9.63% (resp. 6.46%) with respect to the pre-intervention phase, hence supporting Hypothesis 3. This suggests that the long-standing bundle promotion used by the retailer for many years was likely not effective and was even detrimental in terms of revenue and profit. The steady sales during the no-promotion phase, along with an increased profitability, indicates that the unhealthy bundle promotion does not seem to cause any significant behavioral change among customers.

E.2 Propensity score matching

In this section, we discuss the propensity score matching (PSM) combined with the DID specification from Section 6. PSM allows us to match the 10 treated stores with the remaining stores to ensure a more robust comparison by aligning the characteristics of both groups. Our dataset comprises 33 stores, of which 10 were treated, and 23 served as control. To perform the matching, we utilized 12 weeks of pre-experiment data,[6] ensuring that the control stores closely resemble the treated stores across all the relevant variables. A key assumption behind PSM is conditional independence, which posits that only the observed variables influence both the likelihood of receiving the treatment and the outcomes related to the treatment (Rosenbaum 1987). Importantly, as illustrated in Figure E3, the distributions of treated and control stores were similar and statistically indistinguishable from each other, hence supporting the validity of the conditional independence assumption. We use the following logistic regression specification to compute the propensity scores based on store-level characteristics:

$$ln\left(\frac{P(Treat_s = 1|X)}{1 - P(Treat_s = 1|X)}\right) = \alpha + \beta X,$$

where $Treat_s = 1$ for the treated stores (i.e., the stores where the promotion was discontinued) and $Treat_s = 0$ for the control stores. The explanatory variables (X) include store area (in square feet), assortment size (i.e., the number of distinct products sold) during the 12-week period, average weekly number of transactions, average basket size of the transactions in the store, average basket value of the transactions in the store, average total weekly store revenue (in dollars) during the 12-week period, average total weekly store profit from unhealthy product sales (in dollars) during the 12-week period, average number of weekly unhealthy bundle (i.e., coffee and pastries) transactions during the 12-week period, and average number of weekly unhealthy snacks (i.e., pastries) transactions during the 12-week period for all the stores in the city where the experiment was conducted. We repeated the analysis for the number of weeks ranging from nine to 15 weeks and observed that the control stores obtained using the matching algorithms provided consistent results.

Following the calculation of the propensity scores, we performed a nearest-neighbor matching between the treated stores and the untreated stores, without replacement, to identify the best possible control matches for each treated store. This matching process is designed to improve the balance between treated and control groups, thus reducing selection bias. We verified the covariate balancedness by examining standardized differences. All the paired t-tests show no statistically significant differences between treated and control stores on all matched variables, as detailed in Table E11.

Table E11: Summary statistics and balancedness of covariates before and after matching.

	Unmatched				Matched			
	Treated	Control	Difference in means	t-stat	Treated	Control	Difference in means	t-stat
Store area (sqft)	2,645.50	2,215.91	429.59	1.64	2,645.50	2,521.60	123.9	0.64
Assortment size	2,379.20	2,203.77	175.43 *	-2.10	2,379.20	2,371.90	7.3	0.20
Number of transactions	454.78	608.75	-153.96 ***	-4.23	454.78	430.51	24.27	1.19
Basket size	4.01	10.89	-6.88 ***	-8.53	4.01	4.04	-0.033	-0.51
Basket value	12.73	17.39	-4.66 ***	-4.63	12.73	12.27	0.46	0.63
Revenue	53.72	56.62	-2.90	-0.25	53.72	48.34	5.38	0.58
Profit	352.23	349.19	3.04	0.04	352.23	299.85	52.38	0.95
Unhealthy sales	84.50	78.23	6.27	0.38	84.50	67.97	16.52	1.38
Unhealthy bundle sales	10.52	10.41	0.11	0.05	10.52	8.46	2.06	0.98

Note: Matching was performed using nearest neighbors without replacement method. $^{*}p < 0.05$;$^{**}p < 0.01$;$^{***}p < 0.001$.

[6] We used a time window of 12 weeks from 2024-03-17 to 2024-06-09 in which the sales for all stores are representative.

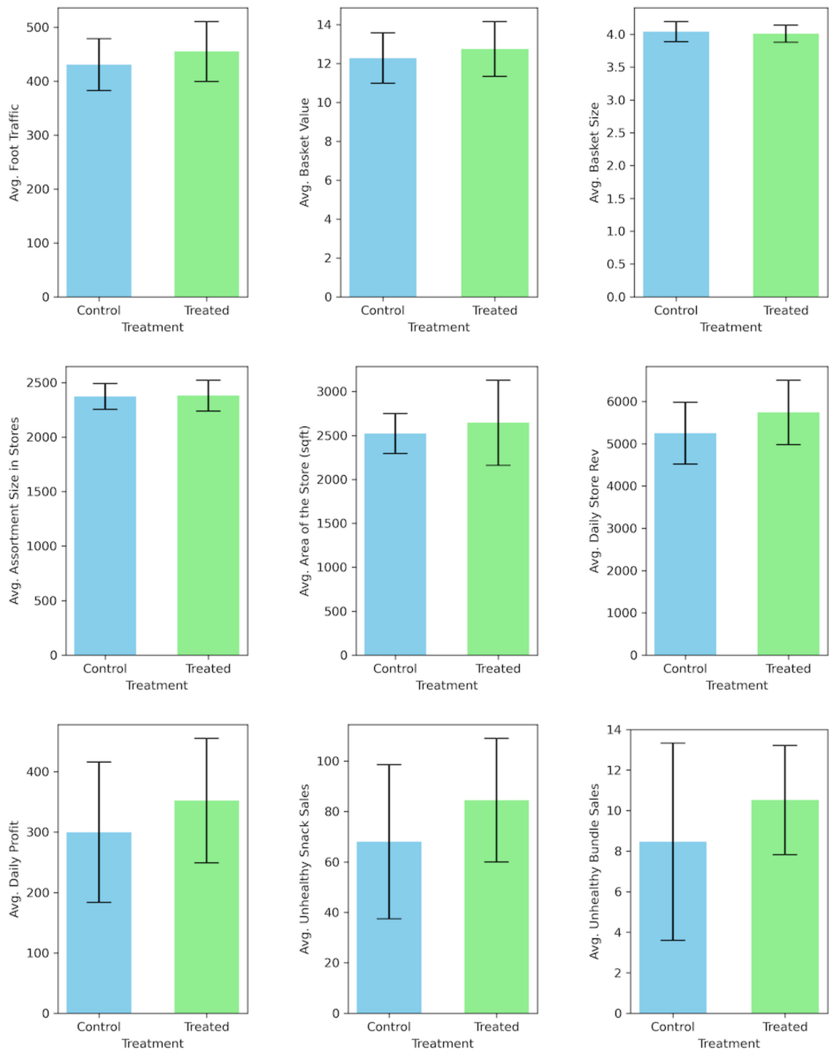


Figure E3: Conditional independence assumption showing that the matched samples are well balanced.

E.3 Robustness test: Analyses on weekday sales

In this section, we repeat the analysis performed in Section E.1 and Section 6 using only weekday transactions.

Table E12: Impact of discontinuing the long-standing promotion using only weekday transactions.

	Unhealthy sales	Number of transactions	Basket size	Revenue	Profit
Treatment	0.05	-0.04	-0.02	10.61 *	8.21 *
	(0.14)	(0.04)	(0.01)	(6.54)	(4.29)
Post Treatment	0.12	-0.05	-0.01	4.39	2.21
	(0.24)	(0.07)	(0.02)	(10.94)	(7.17)
No. Obs.	1,200	1,200	1,200	1,200	1,200
Log likelihood	-9,346.3	-12,576.0	2,015.4	-10,880.0	-10,011.0
Model	NB	NB	TWFE	TWFE	TWFE

Note: $^{*}p < 0.05$;$^{**}p < 0.01$;$^{***}p < 0.001$.
The standard errors are reported in parentheses.
No. Obs. = 24 weeks of experiment ×5 days per week ×10 stores = 1,200.

Table E13: DID with PSM – Impact of discontinuing the long-standing promotion using only weekday transactions.

	Unhealthy sales	Number of transactions	Basket size	Revenue	Profit
Treatment	0.02 (0.10)	-0.01 (0.03)	-0.02 (0.01)	11.26 ** (4.81)	12.55 *** (3.15)
Post Treatment	0.004 (0.12)	0.05 (0.04)	-0.02 (0.01)	5.75 (5.74)	6.10 (3.75)
No. Obs.	2,400	2,400	2,400	2,400	2,400
Log likelihood	-9,346.2	-12,575.0	2,018.1	-10,877.0	-10,002.0
Model	NB	NB	TWFE	TWFE	TWFE

Note: $^{*}p < 0.05$;$^{**}p < 0.01$;$^{***}p < 0.001$.
The standard errors are reported in parentheses.
No. Obs. = 24 weeks of experiment ×5 days per week ×20 stores = 2,400.

E.4 Robustness test: Analyses excluding stockouts

In this section, we repeat the analysis performed in Sections E.1 and 6 by removing the days on which product stockouts occurred. For the analysis on number of transactions and basket value, we remove the days on which stockouts occurred for either product category.

Table E14: Impact of discontinuing the long-standing promotion using transactions without stockout days.

	Unhealthy sales	Number of transactions	Basket size	Revenue	Profit
Treatment	0.08 (0.12)	-0.10 (0.11)	-0.01 (0.02)	15.96 * (8.14)	13.46 ** (5.24)
Post Treatment	0.10 (0.23)	-0.08 (0.22)	0.02 (0.03)	6.88 (16.30)	5.77 (10.50)
No. Obs.	1,337	1,337	1,337	1,337	1,337
Log likelihood	-6,171.5	-9,422.3	1,268.6	-7,439.9	-6,828.0
Model	NB	NB	TWFE	TWFE	TWFE

Note: $^{*}p < 0.05$;$^{**}p < 0.01$;$^{***}p < 0.001$.
The standard errors are reported in parentheses.

Table E15: DID with PSM – Impact of discontinuing the long-standing promotion using transactions without stockout days.

	Unhealthy sales	Number of transactions	Basket size	Revenue	Profit
Treatment	0.03 (0.09)	-0.02 (0.03)	-0.01 (0.01)	13.37 *** (4.25)	13.75 *** (2.75)
Post Treatment	-0.02 (0.10)	0.04 (0.03)	-0.02 * (0.01)	2.42 (4.88)	4.74 (3.17)
No. Obs.	3,233	3,233	3,233	3,233	3,233
Log likelihood	-12,414.0	-16,825.0	2,743.4	-14,637.0	-13,440.0
Model	NB	NB	TWFE	TWFE	TWFE

Note: $^{*}p < 0.05$;$^{**}p < 0.01$;$^{***}p < 0.001$.
The standard errors are reported in parentheses.

F Parallel trends assumption

F.1 Main experiment

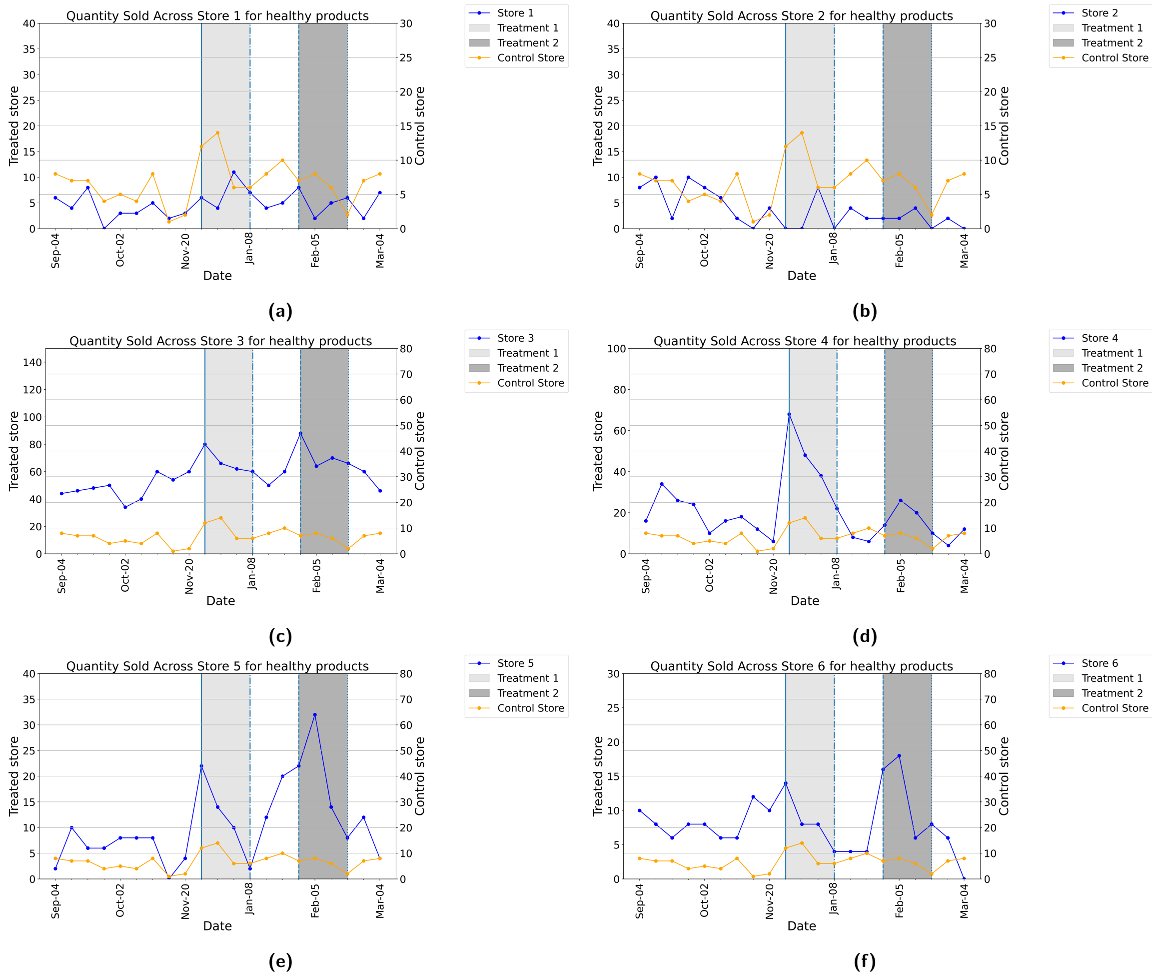


Figure F4: Parallel trends assumption for the main experiment: Healthy products.

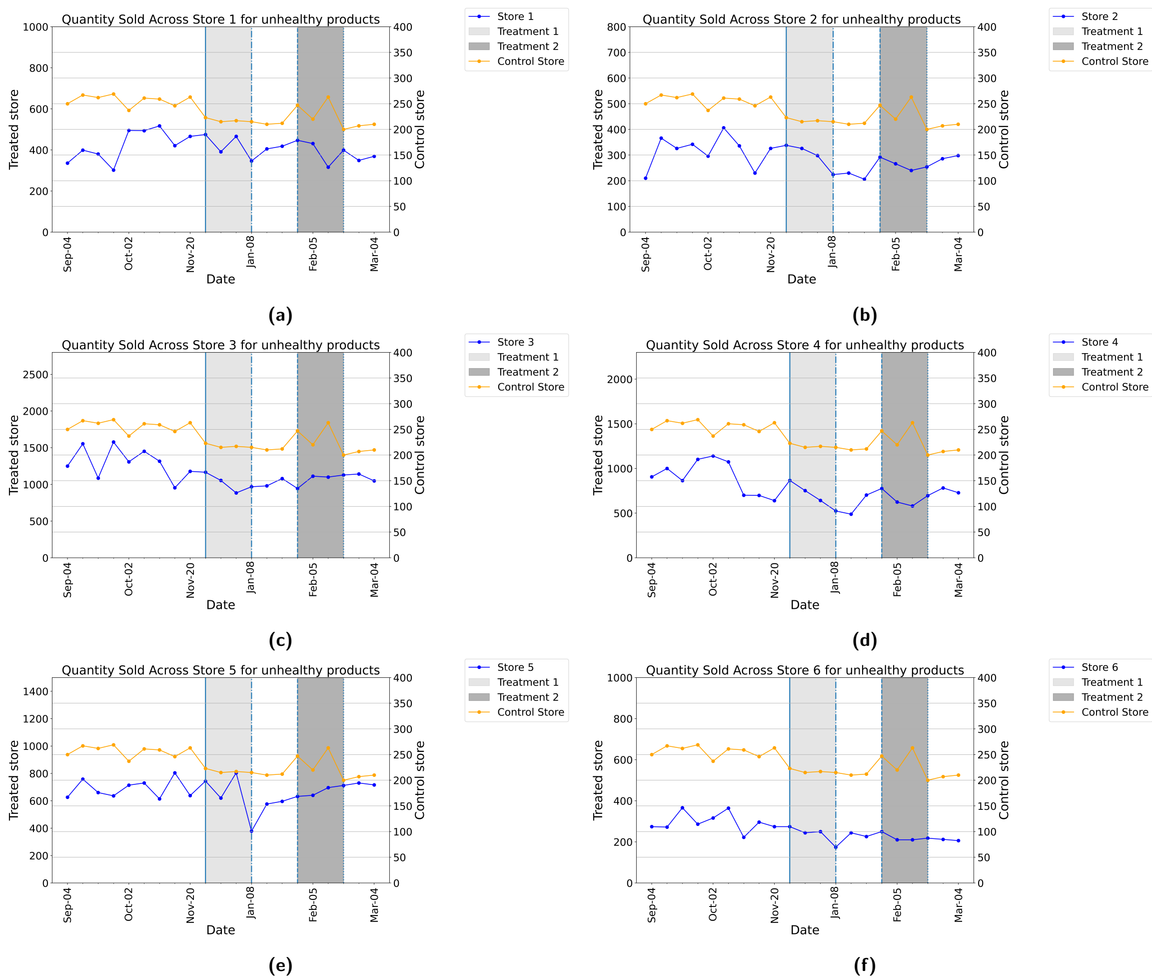


Figure F5: Parallel trends assumption for the main experiment: Unhealthy products.

F.2 Follow-up experiment

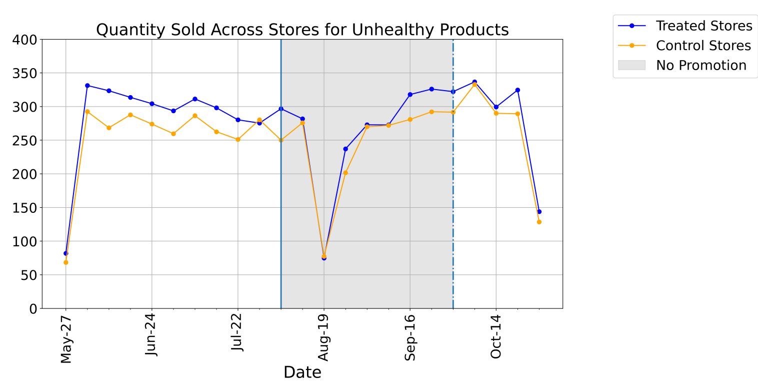


Figure F6: Parallel trends assumption for the follow-up experiment.

References

Aaker, D. A. (2009). Managing brand equity: Capitalizing on the value of a brand name. simon and schuster.

Alvarez, B. A. and Casielles, R. V. (2005). Consumer evaluations of sales promotion: the effect on brand choice. european Journal of Marketing, 39(1/2):54–70.

Baardman, L., Cohen, M. C., Panchamgam, K., Perakis, G., and Segev, D. (2019). Scheduling promotion vehicles to boost profits. Management Science, 65(1):50–70.

Bakos, Y. and Brynjolfsson, E. (2000). Bundling and competition on the internet. Marketing science, 19(1):63–82.

Balachander, S., Ghosh, B., and Stock, A. (2010). Why bundle discounts can be a profitable alternative to competing on price promotions. Marketing Science, 29(4):624–638.

Bandi, N., Cohen, M. C., and Ray, S. (2024a). Incentivizing healthy food choices using add-on bundling: A field experiment. Manufacturing & Service Operations Management.

Bandi, N., Cohen, M. C., Ray, S., et al. (2024b). Behavioral retail operations: Tactics to win customers. Foundations and Trends® in Technology, Information and Operations Management, 18(3-4):214–420.

Berkowitz, E. N. and Walton, J. R. (1980). Contextual influences on consumer price responses: An experimental analysis. Journal of Marketing Research, 17(3):349–358.

Blanken, I., Van De Ven, N., and Zeelenberg, M. (2015). A meta-analytic review of moral licensing. Personality and Social Psychology Bulletin, 41(4):540–558.

Blattberg, R. C. (1990). Sales promotion: Concepts, methods, and strategies. Prentice Hall google schola, 2:982–1003.

Carlson, J. (2017). A content analysis of bonus pack promotions. Journal of Promotion Management, 23(6):930–950.

Chandon, P., Wansink, B., and Laurent, G. (2000). A benefit congruency framework of sales promotion effectiveness. Journal of marketing, 64(4):65–81.

Chen, S.-F. S., Monroe, K. B., and Lou, Y.-C. (1998). The effects of framing price promotion messages on consumers' perceptions and purchase intentions. Journal of retailing, 74(3):353–372.

Chenreddy, A. R., Pakiman, P., Nadarajah, S., Chandrasekaran, R., and Abens, R. (2019). Smoile: A shopper marketing optimization and inverse learning engine. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pages 2934–2942.

Cohen, M. C., Gras, P.-E., Pentecoste, A., and Zhang, R. (2022). Demand prediction in retail: A practical guide to leverage data and predictive analytics. Springer.

Cohen, M. C., Kalas, J. J., and Perakis, G. (2021). Promotion optimization for multiple items in supermarkets. Management Science, 67(4):2340–2364.

Cohen, M. C., Leung, N.-H. Z., Panchamgam, K., Perakis, G., and Smith, A. (2017). The impact of linear optimization on promotion planning. Operations Research, 65(2):446–468.

Cui, R., Li, M., and Li, Q. (2020). Value of high-quality logistics: Evidence from a clash between sf express and alibaba. Management Science, 66(9):3879–3902.

Cui, R., Xin, S., and Li, Z. (2021). Interrogating and redefining the concept of consumer inertia. Journal of Consumer Behaviour, 20(1):21–31.

Cui, R., Zhang, D. J., and Bassamboo, A. (2019). Learning from inventory availability information: Evidence from field experiments on amazon. Management Science, 65(3):1216–1235.

Darke, P. R. and Chung, C. M. (2005). Effects of pricing and promotion on consumer perceptions: it depends on how you frame it. Journal of retailing, 81(1):35–47.

Gao, Y., Li, M., and Sun, S. (2023). Field experiments in operations management. Journal of Operations Management, 69(4):676–701.

Guiltinan, J. P. (1987). The price bundling of services: A normative framework. Journal of marketing, 51(2):74–85.

Gupta, S. (1988). Impact of sales promotions on when, what, and how much to buy. Journal of Marketing research, 25(4):342–355.

Hähnchen, A. and Baumgartner, B. (2020). The impact of price bundling on the evaluation of bundled products: Does it matter how you frame it? Schmalenbach business review, 72(1):39–63.

Han, B. R., Sun, T., Chu, L. Y., and Wu, L. (2019). Connecting customers and merchants offline: Experimental evidence from the commercialization of last-mile stations at alibaba. This article has been accepted for publication in MIS Quarterly on behalf of the Regents of the University of Minnesota and the MIS Research Center.

Hanson, W. and Martin, R. K. (1990). Optimal bundle pricing. Management Science, 36(2):155–174.

Haws, K. L., Reczek, R. W., and Sample, K. L. (2017). Healthy diets make empty wallets: The healthy= expensive intuition. Journal of Consumer Research, 43(6):992–1007.

Hoch, S. J., Kim, B.-D., Montgomery, A. L., and Rossi, P. E. (1995). Determinants of store-level price elasticity. Journal of marketing Research, 32(1):17–29.

Jalali, Z., Cohen, M. C., Ertekin, N., and Gumus, M. (2023). Vertical product location effect on sales: A field experiment in convenience stores.

JRTech (2023). 5 types of retail sales promotions and how to apply them. JRTech Solutions.

Kahneman, D. (2011). Thinking, fast and slow. Farrar, Straus and Giroux.

Kahneman, D. and Tversky, A. (2013). Prospect theory: An analysis of decision under risk. In Handbook of the fundamentals of financial decision making: Part I, pages 99–127. World Scientific.

Khan, U. and Dhar, R. (2006). Licensing effect in consumer choice. Journal of marketing research, 43(2):259–266.

Kopalle, P. K. and Lindsey-Mullikin, J. (2003). The impact of external reference price on consumer price expectations. Journal of retailing, 79(4):225–236.

Lachance, P. and Fisher, M. (1986). Educational and technological innovations required to enhance the selection of desirable nutrients. Clinical nutrition (USA).

Lehtimäki, A.-V., Monroe, K. B., and Somervuori, O. (2019). The influence of regular price level (low, medium, or high) and framing of discount (monetary or percentage) on perceived attractiveness of discount amount. Journal of Revenue and Pricing Management, 18:76–85.

Lobstein, T. and Davies, S. (2009). Defining and labelling 'healthy'and 'unhealthy'food. Public health nutrition, 12(3):331–340.

Melnyk, V. and Bijmolt, T. (2015). The effects of introducing and terminating loyalty programs. European Journal of Marketing, 49(3/4):398–419.

Meyer, R. and Johnson, E. J. (1995). Empirical generalizations in the modeling of consumer choice. Marketing Science, 14(3_supplement):G180–G189.

Mishra, A. and Mishra, H. (2011). The influence of price discount versus bonus pack on the preference for virtue and vice foods. Journal of Marketing Research, 48(1):196–206.

Nunes, J. C. and Park, C. W. (2003). Incommensurate resources: Not just more of the same. Journal of Marketing Research, 40(1):26–38.

Okada, E. M. (2005). Justification effects on consumer choice of hedonic and utilitarian goods. Journal of marketing research, 42(1):43–53.

Pai, F.-Y., Chen, C.-P., Yeh, T.-M., and Metghalchi, M. (2017). The effects of promotion activities on consumers' purchase intention in chain convenience stores. International Journal of Business Excellence, 12(4):413–432.

Palazon, M. and Delgado-Ballester, E. (2009). Effectiveness of price discounts and premium promotions. Psychology & Marketing, 26(12):1108–1129.

Ravensbergen, E. A., Waterlander, W. E., Kroeze, W., and Steenhuis, I. H. (2015). Healthy or unhealthy on sale? a cross-sectional study on the proportion of healthy and unhealthy foods promoted through flyer advertising by supermarkets in the netherlands. BMC Public Health, 15:1–10.

Rosenbaum, P. R. (1987). Sensitivity analysis for certain permutation inferences in matched observational studies. Biometrika, 74(1):13–26.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70(1):41–55.

Scheidt, D. M. and Daniel, E. (2004). Composite index for aggregating nutrient density using food labels: ratio of recommended to restricted food components. Journal of nutrition education and behavior, 36(1):35–39.

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

Thaler, R. (1985). Mental accounting and consumer choice. Marketing Science, 4(3):199–214.

Tripathi, A. and Pandey, N. (2019). Promotional framing of green products: Mediating role of sale proneness. Journal of Promotion Management, 25(7):983–1008.

Xu, Y., Dai, H., and Yan, W. (2024). Identity disclosure and anthropomorphism in voice chatbot design: a field experiment. Management Science.

Yadav, M. S. (1994). How buyers evaluate product bundles: A model of anchoring and adjustment. Journal of Consumer Research, 21(2):342–353.

Yadav, M. S. and Monroe, K. B. (1993). How buyers perceive savings in a bundle price: An examination of a bundle's transaction value. Journal of Marketing Research, 30(3):350–358.