Les Cahiers du GERAD ISSN: 0711–2440

Incentivizing healthy food choices using add-on bundling: A field experiment

N. Bandi, M. Cohen, S. Ray

G–2023–49

November 2023

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, 2023
– Bibliothèque et Archives Canada, 2023

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, 2023
– Library and Archives Canada, 2023

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

Incentivizing healthy food choices using add-on bundling: A field experiment

Nymisha Bandi [a]

Maxime Cohen [a, b]

Saibal Ray [a]

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

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

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

November 2023
Les Cahiers du GERAD
G–2023–49

Abstract : How can retailers incentivize customers to make healthier food choices? Price, convenience, and taste are known to be among the main drivers behind such choices. Unfortunately, healthier food options are often expensive and infrequently promoted. Recent efforts in deploying healthy nudges to incentivize customers toward healthier food choices have been observed. In this paper, we conducted a field experiment with a global convenience store chain to better understand how different add-on bundle promotions influence healthy food choices. We considered three types of add-on bundles: (i) an unhealthy bundle (when customers purchased a coffee, they could add a pastry for $1), (ii) a healthy bundle (offering a healthy snack as an add-on), and (iii) choice bundle (offering either a pastry or a healthy snack). In addition to our field experiment, we conducted an online lab study to strengthen the validity of our results. We found that offering healthy snacks as part of an add-on bundle significantly increased healthy purchases (and decreased unhealthy purchases). Surprisingly, this finding continued to hold for the choice bundle, that is, even when unhealthy snacks were concurrently on promotion. Unfortunately, we did not observe a long-term stickiness effect, meaning that customers returned to their original (unhealthy) purchase patterns once the healthy or choice bundle was discontinued. Finally, we show that offering an add-on choice bundle is also beneficial for retailers, who can earn higher revenue and profit.

Keywords: Healthy eating, field experiment, nudging, bundling

1 Introduction

Food habits have changed considerably in the last few decades, with a shift toward high-calorie and high-sugar dishes, frequent eating out, and larger food portions along with a reduced intake of fruits, vegetables, and high-fiber items.[1] Diets with higher amounts of salt, sugar, and trans fats and lower amounts of fruits, vegetables, and fibers are typically categorized as unhealthy (Lobstein & Davies 2009, Hersey et al. 2013). One of the main consequences of this diet change is a higher incidence of obesity and chronic non-communicable diseases, such as diabetes, heart disease, and stroke (Muhammad et al. 2017). For example, the global prevalence of diabetes nearly doubled from 4.7% in 1980 to 8.5% in 2014 in the adult population (Roglic 2016).

In this context, there is an interest in incentivizing customers to make healthier food choices. Research suggests that price, convenience, and taste are the main drivers behind these choices, whereas listing nutritional benefits and dietary guidelines has very little effect (Sogari et al. 2018). It is thus unsurprising that promotions are one of the most popular strategies used by food retailers (see, e.g., Neslin 2002, Cohen et al. 2021). Ravensbergen et al. (2015) and Bennett et al. (2020) showed that most grocery stores use price promotions—especially on unhealthy food items—to attract price-sensitive consumers, thus exacerbating the aforementioned health problems.

Although supermarkets are popular places to buy food items, customers also buy such items from small stores and other limited-service establishments, such as gas stations, dollar stores, and pharmacies (Ver Ploeg et al. 2015). Given the primary benefit offered by such establishments, they are often categorized as convenience stores (C-stores). C-stores are notorious for carrying a large assortment of unhealthy food (Farley et al. 2009). Studies have shown that the ratio of healthy to unhealthy food available and purchased in C-stores is lower relative to supermarkets (Larson et al. 2009, Stern et al. 2016). Moreover, Bennett et al. (2020) found that the prevalence of promotions for products with high fat, sugar, and salt was also quite high (56%) in convenience stores.

Given that an unhealthy diet has adverse health consequences, there have been concerted efforts in recent years to encourage healthy eating through various interventions or nudges (Reisch et al. 2017, Hinnosaar 2023). Researchers have explored the effectiveness of popular interventions, such as descriptive nutrition labeling (Nikolova & Inman 2015), visibility enhancement (Kroese et al. 2016), increased assortment and availability of healthy items (Van Kleef et al. 2012), healthy eating calls (Salmon et al. 2015), and price promotions for healthy food (Afshin et al. 2017). Hawkes (2009) surveyed the literature on using promotions as an intervention to affect food choices (especially related to healthy eating) and observed that some interventions were successful at influencing consumers' purchasing choices. Many of these studies have focused on the effectiveness of price discounts. For example, An (2013) reviewed 24 field experiments and observed that subsidies for healthier food options significantly increased the purchases of the promoted products.

Among the various promotional tactics, bundling, whereby consumers need to purchase more than one item to receive a price discount,[2] is quite popular. According to the UK's Competition Commission (2000), bundle offers typically increase sales more than price promotions. Notably, Honhon et al. (2017) showed that around 25% of the promotions in supermarkets and convenience stores involve bundling. Unfortunately, it seems that the majority of bundling promotions focus on products with low nutritional content (Furey et al. 2019). Exum et al. (2014) revealed that around 35% of bundling promotions, identified from 10 weeks of U.S. supermarket flyers, were focusing on the food and beverage categories designated as "empty calories" (i.e., high-calorie items with no nutrients) using the MyPlate nutritional classification system.[3]

However, anecdotal evidence suggests that healthy food bundles can be an effective nudge (Gordon & ICF International 2014). For example, the retailer can nudge customers by bundling a healthy

[1] https://ncdalliance.org/why-ncds/risk-factors-prevention/unhealthy-diets-and-malnutrition

[2] Bundling may involve buying multiple units of the same product or a combo offer that includes several products.

[3] https://www.usda.gov/media/blog/2017/09/26/back-basics-all-about-myplate-food-groups

product with a popular one. Bundling also allows the retailer to counterbalance any profit loss due to the discounted bundle offer by including a product with a high profit margin into the bundle. Given its popularity and potential effectiveness, it is surprising that, to our knowledge, there is limited research on the impact of bundling on healthy food purchases. Carroll et al. (2018) performed a lab study in which participants shopped via a grocery display while a cognitive load (i.e., mental strain) was induced. Participants were given one of six bundle treatments,[4] with differences examined among the proportion of items selected from three categories: fruits and vegetables, junk food, and protein/dairy/grains. Specifically, a cognitive load was induced by asking participants to solve an arithmetic task and a memorization task, leading to consumers' decision fatigue and, ultimately, affecting their shopping behavior. The authors found that discounted bundles could successfully nudge consumers toward a higher fruit and vegetable consumption in the absence of a cognitive load, that is, when consumers were more attentive to what they were buying and were conscious about making healthier choices. Carroll et al. (2018) focused on understanding how cognitive load affects consumers' healthy food choices in a bundling setting, making their study different from ours in several respects. Most importantly, they relied on a lab study and hence could not analyze the impact on the retailer's revenue and profit. Our field experiment was conducted in a real store, thus providing us with the ability to examine the impact of bundling interventions both on consumers and on the retailer. Additionally, we use a different bundling technique—called add-on bundling (see below for a formal definition)—as the context to incentivize healthy food choices while simultaneously disincentivizing unhealthy food choices.

Add-on bundling refers to the retail practice of offering a product (say, B) at a discounted price when consumers purchase another product (say, A) at the regular price, with A often being a popular, high-margin product. Examples of add-on bundles are presented in Figure 1. Most add-on bundles are offered in fast-food chains and C-stores and often combine an unhealthy item with a popular one. For example, Tim Hortons, a Canadian multinational fast-food chain, offers consumers the option to add a sausage biscuit for 99 cents (originally priced at $3.29) when they buy a coffee. Another such promotion is offered by 7-Eleven, where consumers can get a 32-oz Big Gulp for 99 cents (originally priced at $1.79) when they buy nachos. Our main research goal is to provide strong experimental evidence regarding the impact of offering add-on bundles on healthy and unhealthy food choices. Can add-on bundles successfully nudge customers toward healthier choices, especially when there may still be unhealthy add-on bundles available? In addition, what is the impact of such offers on the retailer's revenue and profit?

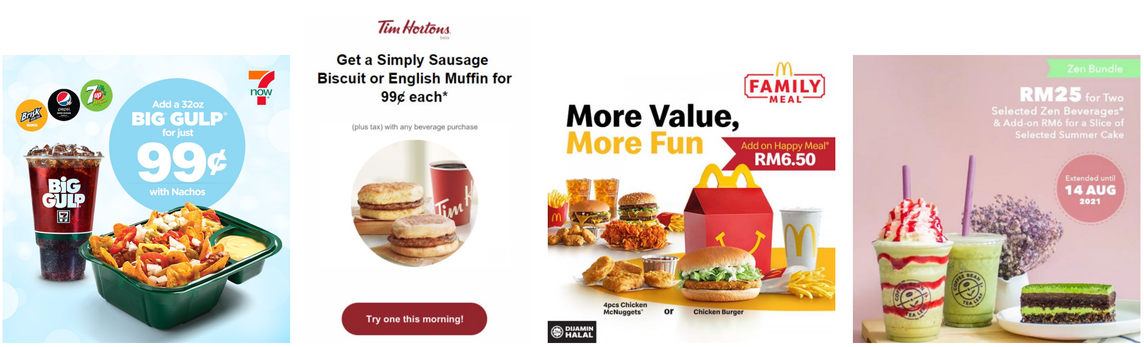


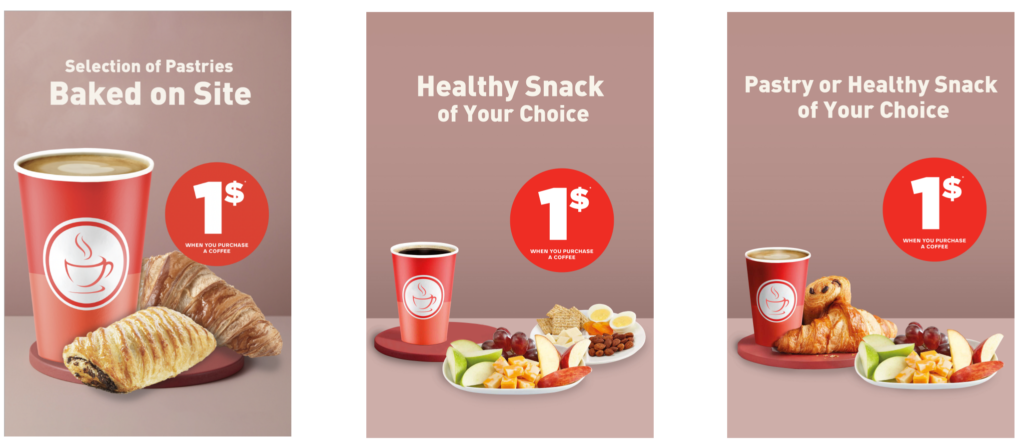
Figure 1: Examples of add-on bundles.

To address the above questions, we conducted a field experiment in a branch store of a leading global C-store chain in a major North American city. Like many C-stores, this store carries a disproportionately low volume of healthy products,[5] and healthy products are promoted much less often (around 28%). In this paper, we use three common nutrient profiling techniques (CFN, RRR, FSA) to

[4] Pay $X when you buy both products A and B.

[5] In our data, less than 15% of the overall transactions involved healthy items. The healthy and unhealthy categorization was based on the Food Standard Agency (FSA) technique. See Appendix A for more details.

classify food as healthy or unhealthy (for more details, see Section 3.1 and Appendix A). The primary unhealthy foods are bakery items, chocolates, salty snacks, and sugary beverages, accounting for 68% of all sales. One of the most popular unhealthy promotional add-on bundles offered by the store is the option to add a bakery item for $1 (average price: $2.36) when purchasing a hot beverage, such as coffee, tea, or hot chocolate (see Figure 2a for an illustration of the promotion). This successful promotion is available in all the stores in the city and has been in place for more than three years. We highlight that 30% of all coffee transactions avail of this promotion and that it accounts for 70% of the sale of bakery items in the store. We consider this setting the status quo (called Control 1) for our experiment. We then tested two different add-on bundles. In Treatment 1, we replaced the unhealthy item in the bundle with a healthy alternative. Specifically, customers were offered the option to add a healthy snack for $1 (average price: $3.99) when purchasing a coffee (see Figure 2b).[6] We call this a healthy add-on bundle. Subsequently, in Treatment 2, customers were offered the choice of an add-on bundle that involves either adding a healthy snack or an unhealthy bakery item for $1 when purchasing a coffee (see Figure 2c). We call this a choice add-on bundle. Lastly, we returned to the status quo by offering the original unhealthy add-on bundle in Control 2. We use the same add-on price of $1 irrespective of the original price of the add-on product to ensure a fair and standardized comparison of the different promotions offered. Each of the above interventions was executed for three consecutive weeks. We used several other control stores in the same city to account for unobserved time heterogeneity. To further support and validate the results of our field experiment, we also ran a lab experiment on the Mechanical Turk (MTurk) platform with a survey that mimicked our field experiment without suffering from the temporal split among the different interventions.



(a) Control (unhealthy bundle) **(b) T1 (healthy bundle)** **(c) T2 (choice bundle)**

Figure 2: Promotion banners used.

Our main empirical methods relied on a difference-in-differences (DID) approach and on a new methodology known as synthetic difference-in-differences (SDID). The latter was recently developed (Arkhangelsky et al. 2019) and fits well our empirical setting. Our control conditions in both methods leveraged the data from all the other stores in the same city by identifying comparable stores. We also used several control variables, such as time fixed effects, to control for city-wide unobserved temporal trends or shocks (e.g., seasonality) and product stockouts. Finally, we conducted a series of robustness tests to showcase the stability of our results. The fact that we found consistent results both in our field experiment (under several model specifications) and in our online MTurk survey experiment enhances our confidence in the validity of our results.

[6]Using a slight abuse of notation, we used coffee to represent all hot beverages in the rest of the paper since coffee purchases accounted for 93% of the total hot beverage sales.

Summary of Results.

As discussed, our field experiment compared three add-on bundles: Controls 1 and 2 (unhealthy bundles), Treatment 1 (healthy bundle), and Treatment 2 (choice bundle). Our results are summarized below:

- **Analyzing the impact of add-on bundling on food choices.**
  - **Healthy add-on bundling.** When comparing Control 1 and Treatment 1, we found that replacing the unhealthy bundle with the healthy one resulted in a significant number of customers substituting unhealthy bakery items with healthy snacks when buying a coffee. Sales of healthy snacks increased by 1,107.69%, whereas sales of unhealthy snacks decreased by 36.52%.
  - **Choice add-on bundling.** More importantly, under Treatment 2 (i.e., the choice bundle), many customers persisted with the healthy option as an add-on instead of opting for the discounted bakery item. Specifically, while sales of the healthy add-on bundle (i.e., coffee + healthy snack) were naturally lower in Treatment 2 than in Treatment 1, overall sales of healthy snacks under Treatment 2 were significantly higher than Control 1. Sales of the healthier alternative increased by 817.5% relative to Control 1 and decreased by 31.63% relative to Treatment 1. Also, sales of the unhealthy add-on bundle (i.e., coffee + bakery item) were higher in Treatment 2 than in Treatment 1 but remained almost the same as in Control 1. Lastly, sales of both bakery items and healthy snacks were similar in Control 1 and Control 2.

  The main takeaway from our experiment is that a healthy add-on bundle can nudge customers toward healthier purchases even in the presence of a concurrent unhealthy bundle. We thoroughly established the robustness of these results by considering a multitude of models and settings.
- **Revenue and profit analysis.** Our experiment also revealed interesting insights into the retailer's profit implications of health nudging. Retailers may not be keen to promote healthy items if doing so might result in a loss. Our results suggest that it is possible to achieve a win-win situation for both customers (who will be more likely to choose a healthy food option) and the retailer (who will earn a higher revenue and profit). Specifically, when comparing Treatment 1 to Control 1, the extra profit earned from the bakery items sold at full price compensated for the loss from the discount offered on healthy snacks, hence maintaining a similar profit level. When comparing Treatment 2 to Control 1 (Control 2), however, we observed a profit increase of 23.93% (28.54%). We also found a 27.41% profit increase from Treatment 1 to Treatment 2. This result is somewhat counter-intuitive since the retailer was offering a discount for both healthy and unhealthy items in Treatment 2. Nevertheless, by offering an add-on choice bundle, the total sales of bundles increased significantly relative to when only one type of bundle was offered. Since the common product in the two bundles was a high-margin product (coffee), the loss incurred due to the discount on healthy snacks was offset by the additional margin accrued from the coffee purchases. In conclusion, when the add-on bundle is carefully designed, an outcome that is both profitable for the retailer and encourages healthy food choices for consumers can be achieved.

The rest of the paper is organized as follows. We develop our hypotheses in Section 2 followed by the design of our field experiment in Section 3. We present the results of our experiment in Section 4, including the revenue and profit analysis. Section 5 reports the results of the MTurk survey. Finally, we conclude in Section 6. Several additional analyses are relegated to the Appendix.

2 Hypotheses development

In this section, we develop hypotheses to study how consumer choices regarding healthy and unhealthy items are affected when exposed to three different promotional bundles (healthy, unhealthy, and choice).

In this paper, the focus is on mixed bundling, namely, the strategy in which a firm sells both the bundle and each of the products separately. The attractiveness of a bundle depends on the products included in the bundle and on whether the purchase is driven by hedonic or utilitarian considerations (Khan & Dhar 2010). Hedonic goods (e.g., designer clothes, luxury watches, and unhealthy food) provide more fun, pleasure, and excitement, whereas utilitarian goods (e.g., microwaves, personal computers, and healthy food) are instrumental and functional (see Hirschman & Holbrook 1982). Consumers' purchase decisions depend on their reservation price for the products in the bundle, which is equal to the sum of the conditional reservation prices of the separate products (Stremersch & Tellis 2002).[7]

We next present our first hypothesis on the impact of the healthy bundle on the sales of both healthy and unhealthy snacks.

Hypothesis 1. *Offering a healthy bundle with a healthy snack as an add-on to a popular item—instead of an unhealthy bundle—will have the following effects:*
a. increase sales of healthy snacks and
b. decrease sales of unhealthy snacks.

We focused on two product categories for the above hypothesis–healthy and unhealthy snacks (more details can be found in Section 3.1)–whereas the popular, common item was coffee. Stremersch & Tellis (2002) showed that a mixed bundle increases sales of the constituent products in the bundle relative to the unbundled scenario. After all, this is the main motivation behind using a bundling strategy. Bundling can be viewed as a price discrimination technique. By properly setting the bundle price, the retailer can capture different customer segments with heterogeneous valuations for the individual products in the bundle. In this case, by replacing the unhealthy snacks in the bundle with healthy ones, we can expect an increase in sales of healthy items (Hypothesis 1a). Recall that the second product category (unhealthy snacks) was part of the bundle prior to our intervention. Consumer choice research states that hedonic items are associated with greater guilt and, thus, require greater justification. Hence, a bundle promotion with a hedonic item (in our case, an unhealthy snack) will be more effective at increasing the purchases of the bundled items than the unbundled items (Khan & Dhar 2010). We thus expect a decrease in sales of unhealthy snacks when the unhealthy bundle is not offered (Hypothesis 1b).

Our second hypothesis is on the impact of the choice bundle on the sales of both healthy and unhealthy snacks.

Hypothesis 2. *Offering a choice bundle—which includes both a healthy and an unhealthy snack option as an add-on to a popular item—will have the following effects:*
a. no effect on sales of healthy snacks relative to the unhealthy bundle setting and
b. no effect on sales of unhealthy snacks relative to the unhealthy bundle setting.

The choice bundle lets the consumers choose between a healthy or an unhealthy snack as the add-on item based on their inherent preferences, both emotional and cognitive. This choice can be seen as being between consumption for immediate pleasure and consumption for long-term benefits and well-being. It is well-known that consumers tend to assign disproportionate weight to short-term benefits and costs (Ainslie 1975). For example, when contemplating a future meal, one may plan to consume healthy options, but when consumption is imminent, one is more likely to prioritize immediate appeal and temptation and opt for unhealthy items. This is driven by temporally inconsistent preferences. This type of choice is also related to the conflict between desire and willpower. People often choose the short-term easy, gratifying option (Shiv & Fedorikhin 2004). The few who can resist this impulse are the ones who make decisions based on a rigorous assessment of the long-term repercussions behind these choices. Researchers have found that emotions—rather than logic—tend to have a greater impact on choice (Khan et al. 2005). Consequently, when presented with a choice bundle, we hypothesize that

[7] The reservation price of a product is the maximum price a consumer is willing to pay for the product. The conditional reservation price is the reservation price of a product conditional on the consumer buying another product.

consumers will opt for the unhealthy snack option, hence implying that the sales of both healthy and unhealthy snacks will not be affected (relative to the setting with the unhealthy bundle).

Hypothesis 3. *Offering a healthy bundle or a choice bundle—instead of an unhealthy one—will not have any effect on sales of unhealthy snacks purchased outside the bundle.*

Each individual develops a reference price for products based on historical prices and other context variables about the product. Consumer purchase behavior is influenced explicitly or implicitly by this reference price (Putler 1992). For each product, there are some individuals who are willing to pay the full price without leveraging any bundled promotions since they have a reference price that is equal to or higher than the price of the product. In the context of our field experiment, the customers who purchased unhealthy snacks at the full price (even when there was an offer to leverage the coffee + unhealthy snack bundle) fell into that category. Such individuals are not likely to alter their purchase behavior when the bundle is modified to include healthy snacks. Alternatively, we can consider the main motivations behind the bundling strategy, which include market segmentation, new product introduction, and cross-selling (Stremersch & Tellis 2002). At the same time, not every customer will be influenced by this strategy. In the unhealthy bundle in our experiment, the focal product was coffee, and the unhealthy snack was the discounted add-on product. There will naturally be some customers who are only interested in the unhealthy snack and will not be influenced by the bundle promotion. These consumers are loyal to the product (in our case, unhealthy snacks) irrespective of the promotions bundled with coffee (since they are most likely not interested in purchasing a coffee). Consequently, we hypothesize that sales of unhealthy snacks purchased outside the bundle will not be affected throughout.

Hypothesis 4. *Our bundling strategies do not have a long-term stickiness effect on sales of healthy and unhealthy items.*

There is a limited understanding on how consumers behave once promotional offers are discontinued. Several researchers have observed in various settings that nudges may only have a short-term impact. For example, in a field experiment to conserve energy, Allcott and Rogers (2014) found that the nudges had no long-term effects on energy consumption once they were discontinued. They observed only short-term effects immediately after the experiment. Similarly, Ni Mhurchu et al. (2010) studied the long-term impact of promoting healthy items and found that there was no significant effect on sales of healthy and unhealthy items once the promotion was discontinued. Motivated by them, we hypothesize that our interventions will not have a long-lasting effect. Hence, we do not expect to observe a stickiness effect in the increased sales of healthy items.

3 Field experiment

To formally test our hypotheses, we conducted a live field experiment in a physical C-store located in the city center of a North American metropolitan city. We then ran a lab study based on an online survey to further showcase the validity and robustness of our findings. In this section, we present the experimental design of our field experiment.

3.1 Design

One of the key steps in our experimental design was to determine the healthy and unhealthy categorization of the offered products. In this paper, we rely on the common convention used by most consumers to define healthy and unhealthy products, while providing support from the nutrition literature. Specifically, products that include fresh fruits or vegetables with low fats, low sugar, and low carbohydrates while having high fiber and other necessary nutrients are considered healthy. To complement this healthy versus unhealthy categorization, we use three common nutrient profiling methods: the ratio of recommended to restricted (RRR) food score (Scheidt & Daniel 2004), the calories-for-nutrient (CFN) score (Lachance & Fisher 1986), and the score developed by the Food Standard Agency.[8] The nutritional information on the various snacks considered in our experiment is reported in Table 1, and the scores computed using the above three methods are presented in Table 2. One can clearly see that the average fats and carbohydrates for the last six items were significantly higher relative to the first three. Similarly, the beneficial nutrients, including fiber, potassium, iron, and calcium, were higher for the first three items. The same can be observed regarding the scores computed using the nutrient profiling methods. More details on this topic can be found in Appendix A. The products with the highest score in all categories were picked as the healthy snacks, which are assortments of healthy items combined and sold as a snack box (Figure 3). More specifically, we considered three types of healthy snacks, namely, fruits, vegetables, and protein (Figure 3a–c). The product descriptions can be found in Table 2. Similarly, all bakery items, namely, croissants, cinnamon rolls, apple turnovers, fruit Danishes, chocolate avalanches, and chocolate muffins (Figure 3d), are unanimously classified as unhealthy by all nutrient profiling methods (as well as based on common sense).

Table 1: Nutrient information for healthy and unhealthy snacks.

Snack type	Total energy (kCal)	Fats (g)	Sugar (g)	Carbohydrates (g)	Fiber (g)	Protein (g)	Potassium (mg)	Calcium (mg)	Iron (mg)
Vegetable box	190	0.5	6	10	3	2	500	40	0.75
Fruit box	230	10	22	27	3	8	250	225	0.4
Protein box	430	27	16	32	4	18	500	125	3
Croissant	272	14	7.5	31	1.7	5.5	79	0	0
Cinnamon roll	184	16	17	32	0.8	3.1	0	18	0
Apple turnover	285	13	25	41	1.6	2.4	0	18	0.75
Fruit Danish	263	20	20	34	1.3	3.8	59	33	1.26
Chocolate avalanche	320	16	11	37	3	5	0	40	1.8
Chocolate muffin	318	14	27	45	0.8	3.8	0	38	0

Table 2: Nutrient profiling scores and contents for healthy and unhealthy snacks used in our experiment.

Snack type	Contents	CFN	RRR	FSA	Product images
Vegetable box	Celery, broccoli, carrots, pepper, and a dip	452.38	0.23	-8	Figure 3b
Fruit box	Apple slices, grapes, and cheese	184.74	0.92	-7	Figure 3a
Protein box	Hard boiled egg, almonds, cheese, and crackers	206.24	0.32	-1	Figure 3c
Pastry	Croissant, cinnamon roll, apple turnover, fruit danish chocolate avalanche, chocolate muffin	722.88	0.1	12	Figure 3d

As discussed, the C-store chain has been running a successful promotion campaign involving an add-on bundle offer. The promotion went as follows: When customers purchased a coffee (aka hot beverage), they could add a pastry for an additional \$1. This was the Control condition for our experiment. The banner used in the store to promote this bundle offer can be seen in Figure 2a. We call this the unhealthy bundle. Our field experiment involved two alternative interventions to test our hypotheses. The first intervention was to study the effect of replacing the unhealthy snack (pastries) in the bundle with a healthy alternative (snack boxes), resulting in Treatment 1 (T1). We call this the healthy bundle. The banner used to promote this intervention can be found in Figure 2b. During this intervention, customers could only add a healthy snack (and not unhealthy pastries) for an additional \$1 when they purchased a coffee. The second intervention was to examine the preference between a healthy and an unhealthy snack when the choice was left to the customers. We call this the choice bundle. The promotion offered during Treatment 2 (T2) is shown in Figure 2c. In other words, when customers purchased a coffee, they could add either a healthy or an unhealthy snack for an additional \$1. As explained before, we strategically designed these three promotion bundles around

[8] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/216094/dh_123492.pdf

(a) Fruit snack

(b) Vegetable snack

(c) Protein snack

(d) Pastry

Figure 3: Healthy and unhealthy snacks used in the bundle promotions.

coffee purchases since it was by far the most sold item in the store (around 45% of purchases). Per our retail partner's suggestion, we opted to use a uniform promotional value of $1 for all bundles in order to keep the deals simple and equally appealing.

As we discuss in Section 4.1, the primary outcome variable in our analyses is the number of transactions per day in each category (healthy and unhealthy snacks). We then examine the impact of the healthy bundle (T1) and the choice bundle (T2) relative to the unhealthy bundle (C1). We also study the post-experiment effect when the promotion reverts to the unhealthy bundle (C2). We emphasize that the C2 condition enables us to check whether the effects observed in our experiment are not merely attributable to an awareness increase of the healthy products but are causally linked to our intervention.

We next discuss the implementation timeline of our field experiment. Since the experiment was conducted at a specific time of the year and each treatment was at a different period, we also conducted a complementary lab study (based on an online survey, presented in Section 5) to verify that our results were not influenced by unobserved temporal heterogeneity.

3.2 Implementation timeline

The experiment lasted for a total of 14 weeks as shown in Table 3 along with the promotions offered. Each intervention was in place for a period of three consecutive weeks. To ensure uniform conditions throughout all phases, we excluded the data collected for two consecutive weeks between 02/21/22 and 03/06/22. During this period, the store faced technical issues while switching from the unhealthy bundle to the healthy one, and there were inventory shortages for the healthy snacks (also, the second week had very few transactions since it coincided with a vacation period). During the experiment, the promotions shown in Figure 2 were displayed near the store entrance for visibility and awareness purposes. There were no other promotions on the same products to alleviate interference effects. The store employees were informed of the promotions to ensure they could answer any customer questions about the products or on the nature of the promotions.

Table 3: Live experiment: dates and promotions offered.

Experiment phase	Date range	Number of days	Promotion offered	Promotion banner
Control 1 (C1)	01/31/22 to 02/20/22	21	Unhealthy add-on bundle	Figure 2a
Excluded (E)	02/21/22 to 03/06/22	14	–	–
Treatment 1 (T1)	03/07/22 to 03/27/22	21	Healthy add-on bundle	Figure 2b
Treatment 2 (T2)	03/28/22 to 04/17/22	21	Choice add-on bundle	Figure 2c
Control 2 (C2)	04/18/22 to 05/08/22	21	Unhealthy add-on bundle	Figure 2a

The timeline for the different interventions in the experiment can be summarized as follows:

- The first three-week period was considered as the baseline phase in which the (usual) unhealthy bundle was offered (Figure 2a). We call this phase C1.
- The next two weeks were excluded (E) due to technical issues beyond our scope.
- During the next three weeks, the healthy bundle was offered (Figure 2b). We ensured that no other promotions were offered for the unhealthy products included in the experiment and that everything else remained the same for these two product categories. We call this phase T1.
- During the next three weeks, the choice bundle was offered (Figure 2c). We call this phase T2.
- Finally, after the T2 phase, the promotion bundle reverted to the default unhealthy bundle offered in C1. Although this promotion continued throughout the rest of the year, we only considered the first three weeks as the C2 phase.

Our goal was to rigorously analyze the customers' purchase patterns of healthy and unhealthy snacks under each of the four conditions (C1, T1, T2, and C2). We investigated the impact of the different treatments (T1 and T2) as well as C2 relative to the control (C1) by running two types of empirical analyses: (a) treatment effect using a regression specification (Section 4.2) and (b) DID (Section 4.3). The first analysis established the effect of the treatment by only considering the sales from the treated store, whereas the second analysis relied on variation in the time series by analyzing the trend changes using 88 other stores of the same chain from the same city. Since the treated store was located in the city center, we saw a clear drop in sales during weekends. This was due to offices being closed on weekends and general footfall being significantly lower. We thus conducted our analyses by using only the weekday sales (that said, the vast majority of our results continued to hold when we included the weekends, as shown in Appendix B).

4 Data and results

In this section, we present the data collected and report our results. Our main econometrics methods are DID and SDID. Nevertheless, we also use ANOVA and regression analyses to showcase the robustness of our estimates.

4.1 Data and metrics

In this section, we provide an overview of the data collected in the treated store during our field experiment. There were two types of data: point-of-sale (POS) data and end-of-day (EOD) inventory data. The POS data provide us with detailed information on all the transactions. For each transaction, we had access to several features, such as transaction time, the total amount spent, the total discount amount, discount details, payment method, and individual items purchased along with quantities and prices. Similarly, the EOD inventory data recorded the inventory level at the end of the day (i.e., midnight) for each item in the store. Unfortunately, this number was often not accurately recorded since it was computed internally in the system based on approximation rules. A physical inventory count would typically result in more reliable ending inventory numbers. However, for large organizations, this is obviously unfeasible. To mitigate this inaccuracy in inventory records and accurately identify when specific products were out of stock, we used both the sales data on a given day and the value of EOD inventory in the system. Specifically, if there were no sales recorded for a particular item on a given day and the EOD inventory was not positive, we could safely conclude that the item was not available on that day. We stored this information as a binary variable called *Stockouts* and used this as a control variable in our empirical models.

In all our analyses, we used data aggregation at the daily level. The values of the average daily transactions for each phase were as follows: 545.71 (SD = 206.14) for C1, 576.95 (SD = 237.82) for T1, 689.65 (SD = 179.03) for T2, and 542.15 (SD = 212.41) for C2. The majority of the transactions were coffee purchases, with a daily average of 267.0 (SD = 101.61) for C1, 258.81 (SD = 107.56) for T1, 326.0 (SD = 93.93) for T2, and 220.45 (SD = 113.97) for C2. Coffee transactions accounted for 46% of the overall store transactions and 34% of those transactions included one of the bundle promotions used in our experiment.

Data filtering.

To ensure that our results are representative, we carefully applied basic filtering rules. We eliminated the top 1% of observations based on the distribution of each key metric. For example, to analyze the total sales during the different phases of the experiment, we first looked at all the transactions and eliminated the top 1% that had unusually large basket sizes. Similarly, we eliminated the transactions with the top 1% highest sales amounts. Finally, we considered the total transactions recorded on each day during the experiment and eliminated the one day with the highest value (i.e., the top 1%, assuming that this value was exceedingly high). To showcase the robustness of our results, we varied this filtering threshold between 1% and 3%. We also considered outlier removal using three standard deviations away from the mean. We observed consistent results under each of these outlier removal approaches.

Key metrics.

We used the folllowing two metrics to capture customer preferences toward healthy and unhealthy food choices:

1. *Number of add-on bundles sold.* The total number of add-on bundles purchased as well as the number of healthy and unhealthy add-on bundles purchased during the experiment period.
2. *Number of transactions that included specific types of items.* The total number of transactions in which either a healthy or an unhealthy snack was purchased.

We then aggregated these metrics at the day level to guide our empirical analyses and estimate the various treatment effects.

4.2 Preliminary results

We examined the impact of the different treatments T1 (healthy bundle), T2 (choice bundle), and C2 (unhealthy reverted bundle) on the average daily number of bundles (all, healthy, and unhealthy) purchased and the average daily transactions containing items related to our experiment (healthy and unhealthy snacks). As mentioned, each treatment lasted for three consecutive weeks (i.e., 21 days). As discussed, we focused on the data from weekdays to reflect a more representative picture (nevertheless, our results remained consistent when including weekend observations) and removed the day with the largest number of transactions (outlier). We thus had a sample of 60 days. We let Y_{ip} denote the values of the two metrics (bundles sold and quantity sold) on day i for different product groups p. The number of bundles sold is analyzed for bundle group $p = \{\text{all, healthy, unhealthy}\}$, whereas the quantity sold is analyzed for product group $p = \{\text{healthy snacks, unhealthy snacks}\}$. We used the following regression model to estimate the treatment effects:

$$Y_{ip} = \alpha + \boldsymbol{\beta} \cdot \boldsymbol{T_i} + \beta_s \cdot StockOuts_{ip} + \beta_t \cdot \mu_i + \epsilon_{ip}, \tag{1}$$

where $\boldsymbol{T_i}$ is the treatment indicator for day i (see Table 3), $StockOuts_{ip}$ indicates whether a stockout has occurred on day i for product p, and μ_i represents time fixed effects to capture any unobserved time-specific demand shocks. We considered various types of time fixed effects, such as day-of-week effects and the week number during the treatment period. The key parameters in Equation (1) are $\boldsymbol{\beta}$, which captures the causal effect of each type of promotion bundle on the snack preferences of customers.

Bundles sold.

We estimated Equation (1) for healthy, unhealthy, and total add-on bundles sold. Specifically, we considered four different models for each case. Model (1) reports the treatment effect estimates ($\boldsymbol{\beta}$) without controlling for stockouts and without including time fixed effects. Models (2)–(4) explicitly account for stockout occurrences and time fixed effects (all possible combinations). As mentioned, we considered three types of time fixed effects: (i) day-of-week effects, (ii) week number during the treatment period, and (iii) a combination of both. We only report the results for day-of-week time fixed effects, but we found consistent results in all three cases. Table 4 shows that the results are consistent across all model specifications (i.e., with and without time fixed effects and with and without controlling for stockouts) for healthy, unhealthy, and all add-on bundles.

Table 4: Impact of different interventions using weekday sales on bundles sold.

	Healthy add-on bundle				Unhealthy add-on bundle			
Treatment	Model (1)	Model (2)	Model (3)	Model (4)	Model (1)	Model (2)	Model (3)	Model (4)
T1	54.07*** (3.73)	54.07*** (3.73)	54.32*** (4.01)	54.90*** (4.03)	-54.87*** (9.51)	-54.87*** (8.74)	-54.74*** (9.74)	-55.65*** (8.97)
T2	38.65*** (3.79)	38.80*** (3.80)	38.84*** (3.96)	39.41*** (3.97)	-10.54 (9.68)	-9.95 (8.91)	-10.48 (9.80)	-10.24 (9.00)
C2	2.53 (3.73)	2.53 (3.73)	2.79 (4.01)	3.37 (4.03)	-10.53 (9.51)	-10.53 (8.74)	-10.53 (9.60)	-10.53 (8.81)
No. obs	59	59	59	59	59	59	59	59
Time FE	No	Yes	No	Yes	No	Yes	No	Yes
Stockouts	No	No	Yes	Yes	No	No	Yes	Yes
R^2	0.85	0.86	0.85	0.86	0.43	0.57	0.43	0.57

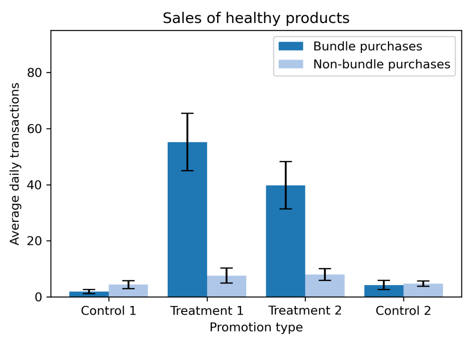
(a) Healthy and unhealthy bundles

	All coffee add-on bundles			
Treatment	Model (1)	Model (2)	Model (3)	Model (4)
T1	-3.67 (10.44)	-3.67 (9.33)	-3.67 (10.44)	-3.67 (9.33)
T2	25.67* (10.63)	26.40*** (9.51)	25.67* (10.63)	26.40*** (9.51)
C2	-8.27 (10.44)	-8.27 (9.33)	-8.27 (10.44)	-8.27 (9.33)
No. obs	59	59	59	59
Time FE	No	Yes	No	Yes
Stockouts	No	No	Yes	Yes
R^2	0.18	0.39	0.18	0.39

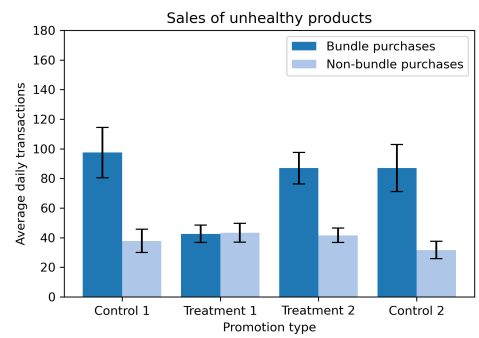
(b) All coffee add-on bundles

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

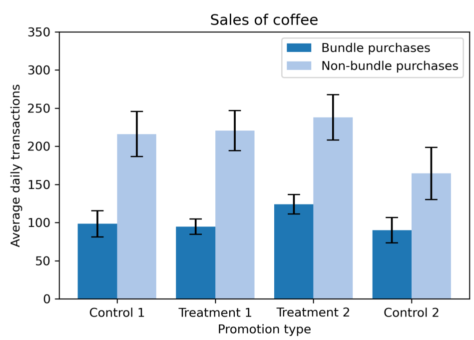
For illustration purposes, Figure 4 plots the sales of coffee add-on bundles and the healthy and unhealthy add-on bundles for each treatment along with a 95% confidence interval.[9] Figure 4c suggests that T2 increased the sales of coffee add-on bundles by 26.08%. This indicates a strong statistically significant positive effect on the sales of healthy snacks during T2 (Table 4b). Recall that these add-on bundles consist of either the healthy or the unhealthy bundle. Figure 4a suggests that T1 (resp. T2) increased the sales of the healthy bundles by 4,784.96% (resp. 3,420.35%). The high percentages are due to low purchases observed in the control period. In other words, we observed a strong statistically significant positive effect on the sales of healthy snacks during both T1 and T2 (Table 4a). Finally, Figure 4b suggests that T1 decreased the sales of unhealthy bundles by 56.29% while T2 did not have a significant effect on the bundle sales. We next proceed to analyze the impact of the different treatments on the number of transactions of healthy and unhealthy snacks.



(a) Healthy snacks



(b) Unhealthy snacks



(c) Coffee

Figure 4: Average daily transactions for different types of products under each phase of the experiment.

Number of transactions.

We estimated Equation (1) for the average daily number of transactions containing healthy snacks, unhealthy snacks, and unhealthy snacks without coffee. We report the treatment effect estimates ($\boldsymbol{\beta}$) in Table 5 and find that the results are consistent across all model specifications (i.e., with and without time fixed effects and with and without controlling for stockouts) for all three cases. In addition, when we changed the control condition to C2 (instead of C1), the results for both healthy and unhealthy snacks remained consistent.

For illustration purposes, Figure 4 plots the average daily sales of healthy and unhealthy snacks for each treatment along with the 95% confidence interval. Figure 4a suggests that T1 (resp. T2) increased the sales of healthy snacks by an impressive 1,107.69% (resp. 817.5%). In other words, we observed a strong statistically significant positive effect on the sales of healthy snacks during both

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

Table 5: Impact of different interventions using weekday sales on quantity sold.

	Healthy snacks				Unhealthy snacks			
Treatment	Model (1)	Model (2)	Model (3)	Model (4)	Model (1)	Model (2)	Model (3)	Model (4)
T1	57.60*** (4.25)	57.60*** (4.17)	57.29*** (4.57)	58.16*** (4.51)	-49.40*** (12.83)	-49.40*** (11.79)	-48.47*** (13.13)	-49.35*** (12.12)
T2	42.51*** (4.32)	42.61*** (4.24)	42.29*** (4.52)	43.02*** (4.45)	-6.77 (13.06)	-6.20 (12.02)	-6.33 (13.20)	-6.18 (12.17)
C2	3.20 (4.25)	3.20 (4.17)	2.89 (4.57)	3.76 (4.51)	-16.67 (12.83)	-16.67 (11.79)	-16.67 (12.93)	-16.67 (11.91)
No. obs	59	59	59	59	59	59	59	59
Time FE	No	Yes	No	Yes	No	Yes	No	Yes
Stockouts	No	No	Yes	Yes	No	No	Yes	Yes
R^2	0.83	0.85	0.83	0.85	0.24	0.40	0.24	0.40

(a) Healthy and unhealthy snacks

	Unhealthy without coffee			
Treatment	Model (1)	Model (2)	Model (3)	Model (4)
T1	5.47 (4.60)	5.47 (4.45)	6.27 (4.67)	6.30 (4.53)
T2	3.77 (4.68)	3.75 (4.53)	4.14 (4.69)	4.06 (4.54)
C2	-6.13 (4.60)	-6.13 (4.45)	-6.13 (4.60)	-6.13 (4.45)
No. obs	59	59	59	59
Time FE	No	Yes	No	Yes
Stockouts	No	No	Yes	Yes
R^2	0.12	0.24	0.14	0.25

(b) Unhealthy snacks without coffee

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

T1 and T2 (Table 5a), which supports Hypothesis 1a but rejects Hypothesis 2a. This implies that although customers were offered a choice between a healthy and an unhealthy snack, there was still a significant demand for healthy snacks relative to the case when there was no promotion on healthy snacks. We also observed that the treatment effect was not significant for C2, indicating that there was no stickiness in the effect once the promotion was discontinued. This supports Hypothesis 4. Figure 4b shows that T1 led to a 36.52% drop in the sales of unhealthy snacks, whereas T2 and C2 did not have significant effects. Similarly, Table 5a reports a statistically significant decrease in the average daily sales of unhealthy snacks during T1 but no effect during T2, hence supporting Hypotheses 1b and 2b. This suggests that by replacing unhealthy snacks with healthy snacks in the bundle, consumers' purchases can be significantly shifted toward healthy choices. It is important to note that during T2, the unhealthy snack sales did not decrease, but the healthy snack sales increased significantly. This is because there were more individuals who were interested in the bundle. Thus, we tapped into consumer groups of both healthy and unhealthy snacks by providing a discount on both. Table 5b also reveals that the sales of unhealthy snacks purchased without coffee were not affected by the experiment, hence supporting Hypothesis 3. This implies that customers who usually purchase unhealthy snacks irrespective of the offered promotion were not affected by the type of bundle.

As discussed, all our results remained consistent when we included the weekend observations (for more details, see Appendix B).

4.3 Difference-in-differences

In the previous section, we estimated various regression models without accounting for a possible treatment selection bias. To address this concern, we compared the sales in the treated store to other untreated stores in the same city during the same period. We relied on a DID specification to quantify the impact of the treatment conditions by contrasting the treated group's performance relative to an untreated group. We specify our DID model as follows:

$$\begin{aligned} Q_{ips} =& \alpha + \boldsymbol{\beta} \cdot \boldsymbol{T_i} + \beta_T Treated_s + \boldsymbol{\beta_d} \cdot \boldsymbol{T_i} \times Treated_s \\ &+ \beta_s \cdot StockOuts_{ips} + \beta_t \cdot \mu_i + \beta_s t \cdot Store_s + \epsilon_{ips}, \end{aligned} \tag{2}$$

where Q_{ips} is the quantity sold on day i for product group $p = \{\text{healthy snacks, unhealthy snacks}\}$ in store s, $Store_s$ represents store fixed effects, $\boldsymbol{T_i}$ is the treatment indicator, and

$$Treated_s = \begin{cases} 1, & \text{if observation occurs in the treated store s,} \\ 0, & \text{otherwise.} \end{cases}$$

Since we were interested in quantifying the impact of the treatments on the treated group relative to the control group, we focused on the interaction coefficients $\boldsymbol{\beta_d}$.

The validity of the DID technique is based on the parallel trends assumption, namely, that no time-varying differences exist between the treatment and control groups. We tested the parallel trends assumption graphically by plotting the sales of both the treated and untreated stores and comparing the trends before the experiment. An alternative statistical technique to test the parallel trends assumption is by using the following equation (O'Neill et al. 2016, Han et al. 2019, Cui et al. 2020):

$$Q_{ips} = \alpha + \beta_1 \cdot d_i + \beta_2 Treated_s + \beta_3 \cdot d_i \times Treated_s + \epsilon_{ips}, \tag{3}$$

where d_i represents the day counter, counting up to the start of the experiment during the pre-treatment period. The above equation measures the effect of time on sales in a difference-in-differences fashion and was run using 12 weeks of data prior to the experiment. If the estimated coefficient $\beta_3 = 0$, then both groups would have the same slope before the experiment started and, hence, the parallel trends assumption would be satisfied.

In this paper, we considered various combinations of stores as the control group and showed consistency in our results. We had a pool of $N = 88$ stores from the retail chain in the same metropolitan city to choose from. We considered the following two approaches to select our control group: (i) stores based on a close geographical distance from the treated store (see Appendix C.1), and (ii) stores based on similar coffee purchasing patterns since coffee was our focal product (see Appendix C.2). The data used to estimate our various DID specifications were the historical POS data and EOD inventory data from the 88 stores starting from September 27, 2021. For better representation, we removed the days between December 20, 2021, and January 30, 2022, due to end-of-year holidays and city-wide COVID-19 restrictions.

We found consistent results using either of the two control groups in the DID analysis. The sales of healthy snacks, unhealthy snacks, and unhealthy snacks without coffee all follow the same trend as we reported in the preliminary results.

4.4 Synthetic DID

In the previous approach, we implemented the DID method by selecting control stores based on two logical rules. It is still possible, however, that the data from the selected control stores are not an appropriate control group. In such scenarios, synthetic control methods can be used as they are generally considered more flexible and more robust than traditional DID estimators. Synthetic control methods can handle cases where the number of available comparison units is small, multiple treatment periods exist, and the treatment effect may vary over time (Bekkerman et al. 2021, Yilmaz et al. 2022). This method still has its limitations, and a new method called Synthetic Difference-in-Differences (SDID) was introduced in Arkhangelsky et al. (2019). SDID is a promising method for estimating causal effects. In the original paper, the authors use the classic example of the California smoking cessation program to compare their technique to the traditional synthetic control method and showcase its benefits.

SDID combines the strengths of DID and synthetic control methods while minimizing their weaknesses. SDID re-weights and matches pre-exposure trends to weaken the reliance on parallel trend-type assumptions present in synthetic control and is invariant to additive unit-level shifts like DID. The SDID estimators can be obtained as follows:

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{d}} = \arg\min \Big\{ \sum_{s=1}^{N=88} \sum_{i=1}^{T} (Q_{ips} - \alpha - \boldsymbol{\beta} \cdot \boldsymbol{T_i} - \beta_T Treated_s - \boldsymbol{\beta_d} \cdot \boldsymbol{T_i} \times Treated_s - \beta_s \cdot StockOuts_{ips} - \beta_t \cdot \mu_i - \beta_{st} \cdot Store_s)^2 \hat{w}_s \hat{\lambda}_i \Big\}, \tag{4}$$

where $\hat{w}_s$ are the unit weights and $\hat{\lambda}_i$ are the time weights. The unit weights are the weights assigned to each of the $N = 88$ stores, whereas the time weights are the weights assigned to each day in both the pre-treatment and post-treatment periods. The details on how to compute these weights are provided in Arkhangelsky et al. (2019). This paper considered the case of a single treatment effect. Since our field experiment includes multiple treatment effects, we had to slightly adapt the algorithm proposed in Arkhangelsky et al. (2019). Specifically, we generate three T_i matrices, one for each treatment, and apply the SDID estimator to them independently. The parameters we need to choose are T_{pre} (number of pre-treatment periods) and N (number of untreated control units). For this analysis, we consider a pre-treatment period of 15 weeks and all $N = 88$ stores as untreated control units. Varying the number of pre-treatment periods will affect both $\hat{w}_s$ and $\hat{\lambda}_i$ and, hence, will also affect the estimate $\hat{\boldsymbol{\beta}}_{\boldsymbol{d}}$. We varied the pre-treatment period from three weeks to 15 weeks with increments of three weeks for robustness purposes and obtained the same qualitative insights. For conciseness, we only report the results when using a pre-treatment period of 15 weeks and including time-varying covariates such as $StockOuts_{ips}$. The results are presented in Table 6. We can see that the sales of healthy, unhealthy snacks, and unhealthy snacks without coffee are consistent with the previous analyses. This strengthens the validity of our results. We highlight that the parallel trends assumption is not a strong requirement for SDID. Nonetheless, it is still satisfied as shown in Table C.6.

Table 6: SDID estimates using 15 weeks of pre-treatment.

	Healthy			Unhealthy			Unhealthy w/o coffee		
Treatment	T1	T2	C2	T1	T2	C2	T1	T2	C2
	49.48***	28.05***	-5.53	-49.45***	0.93	-13.01	3.66	3.44	-7.64
	(0.46)	(0.49)	(2.49)	(3.32)	(3.48)	(3.86)	(1.47)	(1.53)	(2.06)
No. obs	9345	9345	9345	9345	9345	9345	9345	9345	9345
R^2	0.77	0.65	0.29	0.19	0.24	0.19	0.16	0.14	0.05

4.5 Revenue and profit analysis

We investigated the impact of the different interventions tested in our field experiment on the retailer's revenue and profit. More precisely, we examined the impact of the different bundles on the revenue and profit from the three product categories (healthy snacks, unhealthy snacks, and coffee). From the retail chain's perspective, a negative effect on revenue or on profit would reduce the incentive to deploy this type of intervention at scale. The profits are computed using the difference between the selling price and the purchasing cost of each product. The profit value remained constant for all the products throughout our experiment. We then aggregated the profit values from the three product categories for the four phases of the experiment for further analysis.

The pairwise t-tests on revenue and profit between the four phases of our experiment are reported in Table 7. We found that the revenue and profit in T1 were not significantly different when compared to either C1 or C2. This is driven by the fact that the retailer earned a higher revenue and profit by charging the full price on unhealthy snacks (pastries) during T1. Our analysis shows that this additional revenue and profit approximately offset the loss from the promotion on healthy snacks (snack boxes). This ultimately led to similar average revenue and profit levels. In addition, as we can see from Figure 4c, there was no increase or decrease in coffee sales between T1 and C1 (or C2), ensuring that there was no overall negative impact on revenue or profit. However, the revenue and profit in T2 had a statistically significant positive effect relative to either C1 or C2 (and even T1). Specifically, we observed a 23.93% (resp. 28.54%) profit increase and a 28.31% (resp. 38.45%) revenue increase during T2 relative to C1 (resp. C2). Since the sales of unhealthy snacks during T2 were not significantly different relative to C1 and C2 (from our results in Section 4), we attribute the increase in revenue and profit to the increase in sales of coffee bundles during T2 (which is found to be 25.21%). Indeed, the profit margin on coffee happened to be much higher than the profit margins on healthy and unhealthy snacks, so the profit increase from coffee clearly more than compensate the loss incurred from healthy snacks. It is interesting to highlight that by offering both healthy and unhealthy snacks via a choice bundle in T2, the retailer can generate higher revenue and profit relative to only offering the healthy bundle in T1. The revenue (resp. profit) in T2 was 17.91% (resp. 29.1%) higher than in T1.

Table 7: Pairwise comparisons of revenue and profit between the different interventions.

	Revenue		Profit	
	Mean difference	p-value	Mean difference	p-value
C1-T1	-76.88	0.33	22.12	0.65
C1-T2	-246.69	**0.01**	-130.02	**0.04**
C1-C2	63.81	0.51	19.47	0.75
T1-T2	-169.81	**0.05**	-152.14	**0.02**
T1-C2	140.69	0.15	-2.65	0.97
T2-C2	310.50	**0.01**	149.49	**0.04**

In summary, the above findings bear the following practical implications:

1. Offering an add-on bundle with only healthy items does not have a significant positive impact on revenue and profit.
2. Offering an add-on choice bundle with either a healthy or an unhealthy item leads to a significant increase in revenue and profit relative to a separate bundle (healthy or unhealthy).

5 Mechanical Turk survey

The findings from the previous section were highly consistent using multiple model specifications. Admittedly, there is still a possibility that the results were driven by the specific time period when the experiment was conducted and were affected by the time differences between treatments. To test our interventions' effects without time-related biases, we conducted a custom online survey with 2,000 individuals using the Mechanical Turk (MTurk) platform, an online labor market of U.S. workers who complete online tasks and surveys (see, e.g., Katok 2018, Mohan et al. 2020). Lab experiments are often used to test hypotheses in carefully controlled environments (see, e.g., Carroll et al. 2018, Devlin et al. 2022). We performed a cross-sectional study of American adults through the MTurk platform. Participants completed a Qualtrics survey and received a unique ID for data tracking. The survey took 3–4 minutes to complete, and participants received a $1 compensation. Informed consent was not required due to the fact that the data was anonymized.

We divide the rest of this section into two parts. First, we discuss the study design and the survey questions. Second, we present the results from the data analysis and relate them to the findings of our field experiment.

5.1 Study design

The survey was designed to closely replicate the decision-making process in the physical store. Recall that the items offered in the store in the context of our field experiment were coffee, pastries, and healthy snack boxes. We thus included the same items in the lab study. Each participant was given a certain budget to spend and was shown a specific promotion, as depicted in Figure 2. The participants were then provided with the range of products they could select from along with their prices before starting the survey. Once the participants were allocated a budget and assigned to a specific promotion, we asked them a series of questions to understand their preferences. We asked the participants to select at least one item to be eligible for the $1 compensation.

The first aspect of the survey was deciding the budget amount. We made this decision by looking at the coffee-based transactions in the treated physical store and observed an average transaction amount of $3.45. We decided to round up the budget (i.e., to $4) to be used as a low budget for half of the participants in the lab study. This allowed participants to purchase a coffee while also being able to take advantage of the promotion (i.e., add a pastry or a healthy snack for an additional $1). Similarly, a high budget of $8 was used for half of the participants to allow them to purchase all three items offered in the survey if they desired. After randomly assigning a budget value to the participants, we further randomly assigned them to one of three promotions (Figure 2): the Control group viewed the promotion in Figure 2a, the T1 group viewed Figure 2b, and the T2 group viewed Figure 2c. The first question in the survey asked the participants to select their preferred hot beverage (if any), as shown in Figure D.8 in Appendix D. If they did not select a hot beverage, the survey redirected to ask whether they would like to purchase a pastry or a healthy snack at the full price (Figure D.12). If they selected a hot beverage from the options provided, they could opt to add one of the following items depending on their assigned condition:

- A pastry (i.e., unhealthy snack) for an additional $1 if assigned to the Control (Figure D.9).
- A snack box (i.e., healthy snack) for an additional $1 if assigned to T1 (Figure D.10).
- Either a healthy or an unhealthy snack for an additional $1 if assigned to T2 (Figure D.11).

Answering these questions brought the survey to an end. The survey was designed to be dynamic, and the questions depended on the participants' previous responses. The survey flow diagram is shown in Figure 5. At each survey stage, participants could see the remaining budget and the prices of the items available for selection. This ensured that participants could plan their purchases. The survey included two attention-check questions (Figures D.6a and D.6b) placed at the beginning and end of the survey to ensure data quality. The complete transcript of the survey is provided in Appendix D. Table 8 outlines the assignment of participants who successfully completed the survey and made at least one purchase.

Table 8: Allocation of participants to the different budget values and promotions.

	Promotion offered		
Budget	Control	Treatment 1	Treatment 2
LOW ($4)	242	217	249
HIGH ($8)	216	246	224

The main benefit of this online survey was that it fully eliminated any time-dependent effects that could have been present in our field experiment. Since all three promotions were run simultaneously

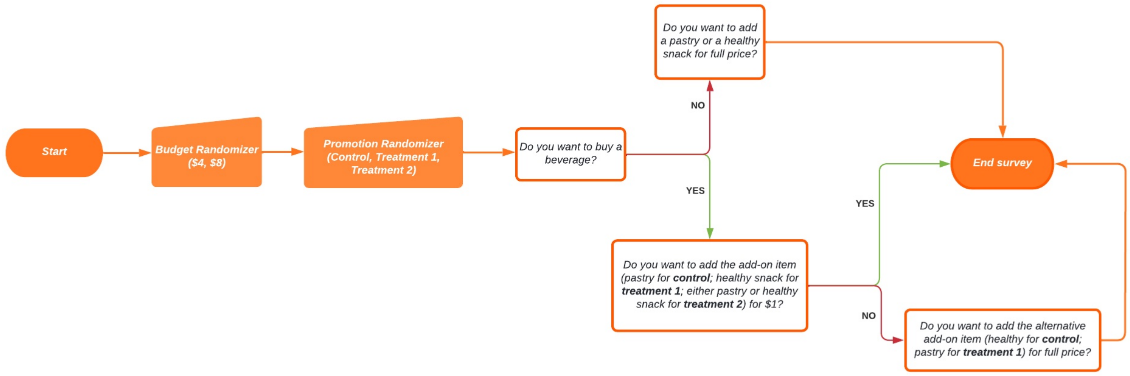


Figure 5: Flow of our online survey conducted on MTurk.

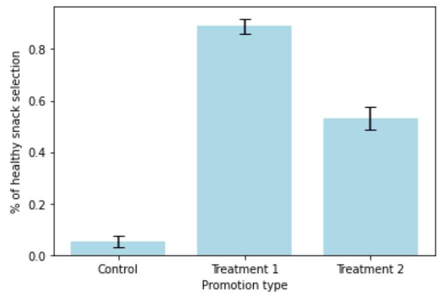
in the form of a survey, it provides the perfect data to strengthen our results and overcome the shortcoming of having different intervention timings.

5.2 Results

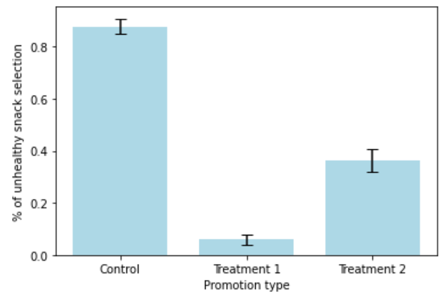
The survey was set up to be filled by 2,000 respondents on the MTurk platform. Once participants completed the survey on Qualtrics, they received a code that was recorded on MTurk. We were able to match this unique ID for 1,870 responses. We filtered out the responses where participants spent $0 in the virtual store, resulting in 1,582 remaining responses. Subsequently, we removed the survey responses that failed to pass the attention check questions, leaving us with 1,435 responses. Finally, we applied a filter based on the amount of time the respondents spent to complete the survey. Specifically, we removed the responses for which it took longer than three standard deviations from the mean to complete the survey. Ultimately, we retained a final sample with 1,394 records. We note that all our results and insights still held even when we did not apply the last filtering rule on the response time (i.e., when we used the sample with 1,435 observations instead of 1,394). We investigated the balancedness of the experimental groups to make sure that the differences between the groups originated solely from the treatment and were unaffected by other factors. Using a chi-squared test, we assessed the uniform distribution of samples across groups, as shown in Table 8. Our analysis indicated that the percentage of participants assigned to different promotions did not vary significantly based on budget, $\chi^2(2, N = 1,394) = 4.27, p = 0.12$, hence confirming that the sample is properly balanced. Thus, we can safely attribute the observed differences between groups to the treatment effect.

Figure 6 displays the selection percentages for healthy snacks, unhealthy snacks, and unhealthy snacks without coffee under the different interventions (Control, T1, and T2). Figure 6a illustrates the higher preference for healthy snacks in T1 and T2 compared to the Control group. It is also interesting to highlight that the effect dropped for T2 relative to T1, hence replicating the effect observed in our field experiment (see Figure 4a). Figure 6b shows a lower preference for unhealthy snacks in T1 relative to the Control, consistent once again with our field experiment. However, T2's lower preference for unhealthy snacks seems to contradict our earlier result (Figure 4b). In fact, this discrepancy makes total sense. Due to the fixed number of respondents per intervention in the survey (as opposed to our field experiment), we could not measure the increase in affinity for the choice bundle as observed in the field experiment. Lastly, Figure 6c shows no significant differences in the selection percentages of unhealthy snacks without coffee across the different promotions. These trends were consistent with the ones observed in our field experiment (see Figure 4). To formally support the above observations, we estimate the following model specification:

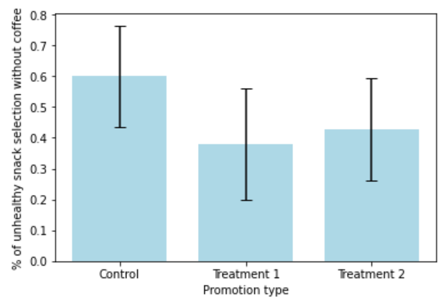
$$R_{ip} = \alpha + \boldsymbol{\beta} \cdot \boldsymbol{T_i} + \gamma B_i^H + \epsilon_{ip}, \tag{5}$$



(a) Healthy snacks



(b) Unhealthy snacks



(c) Unhealthy snacks without coffee

Figure 6: Summary statistics of the lab experiment (online survey).

where R_{ip} is a binary variable that indicates participant i's preference for product group $p = \{$healthy snacks, unhealthy snacks, coffee$\}$, $\boldsymbol{T_i}$ is the treatment indicator (i.e., the promotion offered) to participant i, and

$$B_i^H = \begin{cases} 1, & \text{if participant } i\text{'s budget is \$8,} \\ 0, & \text{otherwise.} \end{cases}$$

The estimated coefficients from Equation (5) are reported in Table 9. Model (1) reports the treatment effect ($\boldsymbol{\beta}$) for healthy snacks, unhealthy snacks, and unhealthy snacks without coffee while not controlling for the budget allocated to the participants, whereas Model (2) explicitly controls for the budget allocated to each participant. The results with and without controlling for the budget were consistent. Reassuringly, except for the low observed preference for unhealthy snacks during T2 in the survey setting, all other results were perfectly consistent with the results from our field experiment.

Under T1, the preference for healthy snacks was 1,598% higher compared to the Control, while under T2, it was 917% higher. In contrast, the preference for unhealthy snacks under T1 was 93% lower compared to the Control, and under T2, it was 59% lower. The lower preference for unhealthy snacks under T1 aligns with our field experiment results. As discussed, however, the lower preference for unhealthy snacks under T2 initially appears to contradict our field experiment findings. This discrepancy is explained by the equal probability assignment of the three promotions (Control, T1, and T2) to survey participants.

In summary, the results from our online lab experiment strongly support the findings from our field experiment, hence reinforcing the conclusion that both a healthy add-on bundle and a choice add-on bundle significantly increase the likelihood of selecting healthy food choices. To conclude, we note that analyzing the revenue and profit metrics for the online participants is not relevant.

Table 9: Effect of different interventions while controlling for the budget value.

	Healthy		Unhealthy		Unhealthy without coffee	
	Model (1)	Model (2)	Model (1)	Model (2)	Model (1)	Model (2)
Treatment 1	0.84***	0.84***	-0.82***	-0.82***	-0.02	-0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Treatment 2	0.48***	0.48***	-0.51***	-0.51***	-0.01	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
No. of observations	1,394	1,394	1,394	1,394	1394	1394
Budget	No	Yes	No	Yes	No	Yes
R^2	0.47	0.47	0.46	0.46	0.002	0.003

6 Conclusion

According to Thaler and Sunstein (2016), nudging is becoming a key method to positively influence people's behavior. Nudging for social good often involves assisting individuals in adopting healthier and more sustainable lifestyles by leveraging their mental shortcuts, emotions, and surroundings (Chaurasia et al. 2022). In this context, private firms can also have a social impact when interacting with their customers. The study by Kroese et al. (2016) is a good example of nudging for social good, where visibility enhancement was used to nudge customers toward a healthier food alternative. A second example is Cohen et al. (2021), who used nudging to encourage a more environmentally sustainable carpooling behavior for daily commuting. Another example is Drake et al. (2016), which discusses how regulatory interventions can be used to drive positive environmental change.

In this paper, we focused on incentivizing retail customers to make healthier food choices by offering add-on bundles with healthy snacks. We investigated the impact of these bundling strategies on customers' snack purchases. We conducted two studies—a field experiment in a physical store and an MTurk-based online lab study—to study this question. We considered three bundle combinations: (i) an unhealthy bundle (status quo), (ii) a healthy bundle, and (iii) a choice bundle. We found strong evidence that healthy snacks are purchased much more frequently when offered as part of a bundle. At the same time, the sales of unhealthy snacks are significantly reduced when they are not part of the add-on bundle. Unfortunately, however, there was no long-term stickiness; the preferences reverted back to the original levels when we stopped offering promotions on healthy snacks. Ultimately, we found that strategic add-on bundling incentivizes healthy food choices even when unhealthy items are included in a choice bundle. We also convey that well-designed bundles can increase the retailer's revenue and profit. Specifically, offering a choice bundle boosted the revenue and profit by 28.31% and 23.93%, respectively. Thus, offering such an add-on choice bundle is beneficial for both customers (who can enjoy healthy food at a discount) and retailers (who can earn higher revenue and profit).

We then conducted an online MTurk survey to showcase the robustness of our results and, in particular, that they are not driven by the time differences between treatments. Each participant was shown one of the three bundles (unhealthy, healthy, and choice) and asked a series of questions to infer their snack preferences. The data gathered from our online survey readily confirmed the findings from our field experiment.

To conclude, it is interesting to examine how consumers' food preferences vary under the different interventions of our field experiment. In Figure 7, we plot the proportions of sales for the different types of purchases.[10] We readily observed that the number of customers who purchased healthy snacks (C1 = 0.54%, T1 = 0.92%, T2 = 0.92%, C2 = 0.62%) and unhealthy snacks (C1 = 7.63%, T1 = 7.50%, T2 = 6.51%, C2 = 6.27%) outside the bundle, and the number of customers who purchased coffee together

[10] "Others" refers to all purchases outside the categories in our experiment (coffee, healthy, and unhealthy snacks). "Coffee only" corresponds to transactions where only coffee was purchased. "Coffee + Unhealthy" corresponds to coffee purchases with pastry items. "Coffee + Healthy" corresponds to coffee purchases with healthy snacks."Unaffected categories" correspond to purchases of unbundled healthy or unhealthy snacks or Coffee + other products.

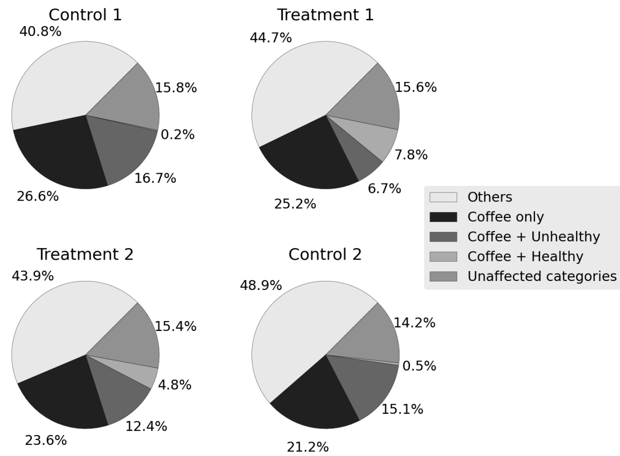


Figure 7: Customers' preferences for various product categories under the different interventions.

with items outside the experiment (C1 = 7.61%, T1 = 7.22%, T2 = 7.93%, C2 = 7.33%) remained roughly the same throughout all four phases of our experiment. These transactions were aggregated and represented as "Unaffected categories." For simplicity, let us consider that 100 customers entered the store to make a purchase during each intervention. Under T1, we observed an increase of 7.6% (= 7.8 − 0.2) in healthy bundle purchases compared to C1, which is likely coming from customers who were only purchasing a coffee before (i.e., adapters) and from some of the customers who were purchasing the unhealthy bundle (i.e., switchers). The proportion of sales from other product categories actually increased by a slight 3.9% (= 44.7 − 40.8), hence indicating that there is no cannibalization effect from other product categories. We also observed that the decrease in unhealthy bundles was not entirely compensated by the increase in healthy bundles. That is, a portion of the customers who stopped purchasing unhealthy bundles switched to healthy bundles, whereas the remaining switched to "Others." Similarly, under T2, the increase in healthy bundles amounted to 4.6% (= 4.8 − 0.2) compared to C1. The proportion of sales from other product categories increased by 3.1% (= 43.9 − 40.8). Thus, the primary reason behind the increase in healthy bundles came from customers adopting the healthy bundle instead of only purchasing a coffee and from customers switching from the unhealthy to the healthy bundle. As discussed, we found that more than half of the customers continued to purchase the healthy bundle even when they were offered a choice between healthy and unhealthy snacks. When the promotion reverted back to the original unhealthy bundle (C2), the preferences for healthy and unhealthy snacks reached similar levels as in C1; hence, there was no long-term stickiness effect.

A potential limitation of our study is the fact that we cannot disentangle whether the healthy buying effect comes from the add-on bundling or if we would find the same effect by just discounting the products. Our intuition suggests that the effect we observe (with such a large magnitude) is due to the add-on bundling mechanism, and we would have not seen such a strong effect under just a price discount. To validate our intuition, we run two new parallel lab experiments: one on add-on bundling only, and one where we compare add-on bundling with price discounting. In the first lab experiment, we simply re-run the exact same survey as in Section 5. In the second lab experiment, we expose consumers to one of three different deals: control, discount, and bundle. In the first deal (control group), we offer consumers the baseline unhealthy add-on bundle (Figure 2a). In the second deal, we offer two independent promotions, namely, the baseline unhealthy add-on bundle as well as a price discount for the healthy snacks. We carefully ensured that the price discount offered for the healthy snacks is equal to the price discount one would obtain by purchasing the unhealthy snack under the add-on bundle (47.35%). Finally, in the third deal, we offer the same healthy bundle as before (Figure 2b). To make an apples-to-apples comparison, we set the add-on price to $2.25 to ensure that the percentage discount for the healthy snacks is equal to the percentage discount in the add-on bundle.

This new lab experiment allows us to examine whether bundling is indeed a more successful strategy than price discounting to incentivize healthy food choices. The results from the survey can be found in Appendix E. Interestingly, we found a 10.16% higher likelihood to purchase healthy snacks when they were promoted as part of a bundle relative to being offered at a discounted price. In addition, the likelihood to purchase unhealthy snacks was 38.34% lower when they were not included in the add-on bundle, namely during the healthy bundle intervention. These results highlight the effectiveness of bundling as a more compelling strategy than price discounting in incentivizing healthy food choices. As discussed, bundling has also the added advantage to boost revenue and profit when it is designed strategically.

Appendix A Definition of healthy food items

The classification of items as healthy or unhealthy is an important stage in our experimental design. Most individuals use the nutrient information on the packaging to classify products as healthy or unhealthy. One of the most common techniques used to categorize a food item as healthy or unhealthy relies on the amounts of nutrients and fats, carbohydrates, and sugar per kCal of food. In addition, there exist several profiling techniques that are used to categorize food items. At a high level, healthy food items are considered to be nutrient-dense, namely, they provide substantial levels of vitamins and minerals while containing relatively few calories (Drewnowski & Fulgoni 2008). The nutrient composition of the products under consideration in our field experiment is listed in Table 1 of the paper. We can then use this information to compute the score from nutrient profiling methods. Specifically, we use the calories for nutrient (CFN) score, the ratio of recommended to restricted (RRR) score, and the food standard agency (FSA) rating. The three nutrient profiling models used to classify products as healthy or unhealthy can be summarized as:

1. **CFN score** (Lachance & Fisher 1986) – The lower the CFN value, the lower the cost in calories to obtain the nutrients associated with a given food (and hence the healthier the food is). This is equivalent to computing how densely packed with nutrients a particular food is. However, this metric does not consider the nutrients that can be harmful when excessively consumed, such as sugar and carbohydrates. The CFN score can be computed as follows:

   $$CFN = \frac{ED}{\sum_{i=1}^{13} \%DV_i/13},$$

   where ED is the energy density of the food item measured in kcal and the denominator corresponds to the average daily value percentages of 13 nutrients, namely, protein, vitamins A, C, B6, and B12, thiamin, riboflavin, niacin, folate, calcium, iron, zinc, and magnesium. To compute the CFN score, one needs to scale the nutrients available in 100g of the food item.

2. **The RRR score** (Scheidt & Daniel 2004) – A higher value translates into a healthier food item. This metric computes the ratio of recommended nutrient values (e.g., vitamins) with the restricted ones (e.g., sugars, fats, carbohydrates). This metric is more comprehensive since it relies on both the recommended and non-recommended nutrients. The RRR score can be computed by using the following formula:

   $$RRR = \frac{\sum_{i=1}^{6} Nutrient\_recommended_i/6}{\sum_{i=1}^{5} Nutrient\_restricted_i/5}.$$

   The recommended nutrients are protein, fiber, vitamins A and C, calcium, and iron. The restricted nutrients are energy, saturated fats, sugar, cholesterol, and sodium. The score is computed per serving of the item.

3. **The FSA rating**[11] – This rating provides an integer value, and any food item with a value below four is considered as healthy. It also accounts for both the recommended and non-recommended

[11] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/216094/dh_123492.pdf

nutrients. In addition, it explicitly considers whether a particular food item contains fruits, vegetables, and nuts. To compute the FSA score, one needs to scale the nutrients available in 100g of the food item. The FSA scoring algorithm can be divided into the following three steps:

(a) Compute the total 'A' points = (points for energy) + (points for saturated fats) + (points for sugar) + (points for sodium).

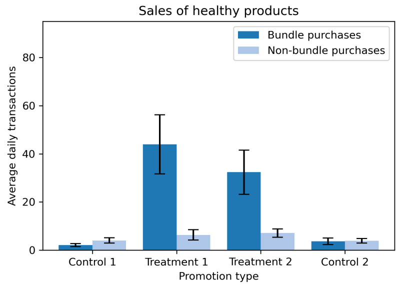
(b) Compute the total 'C' points = (points for % fruits, vegetables, and nut content) + (points for fiber) + (points for protein).

(c) Final score = Total 'A' points - Total 'C' points, if Total 'A' points are lower than 11. Otherwise, we do not count points towards protein, unless Total 'C' points are higher than 5.

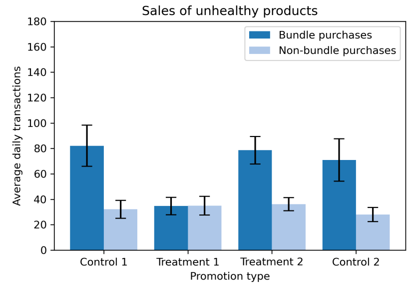
The computed values for the healthy snacks and the average values for the pastry items are reported in Table 2 of the paper. For the pastry items, we use an average value for simplicity. Overall, it is clear that the pastry items have a much higher score relative to the three healthy snacks. One exception is the fruit snack box that contains natural sugar, as opposed to the pastry items that have artificial sweeteners. This key difference is accounted for in the FSA score but not in the other profiling methods.

Appendix B Robustness tests

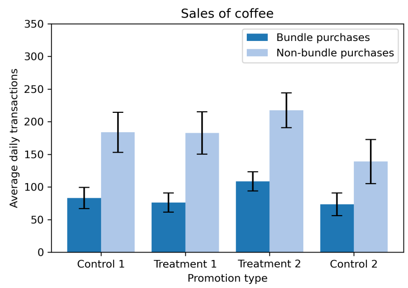
In this section, we show that the results presented in Section 4.2 are robust to the inclusion of weekend observations. Accordingly, we include the data from all the weekends and re-estimate the treatment effects for both the bundles sold and the quantity sold as we did before. Figure B.1 shows the average daily transactions of both bundled and unbundled purchases of healthy, unhealthy snacks, and coffee under the different treatments. As we can see, the trends are very similar to the ones observed without including the weekend observations. Table B.1 supports our previous results on coffee bundle preferences being higher during T2. The sales of healthy and unhealthy coffee add-on bundles align as well. Similarly, Table B.2 reports the treatment effects for the overall products purchased with and without controlling for stockouts and time fixed effects. All the results are consistent with the estimates presented in Table 5 from Section 4.2. When you consider the consumer preference for coffee add-on bundles, T2 has a positive effect whereas T1 and C2 have no significant impact on it (Table B.1b). T1 and T2 have a strong positive effect on the number of healthy add-on bundles sold while C2 has no effect on it. Finally, T1 has a negative effect on the number of unhealthy add-on bundles purchased, and T2 and C2 had no significant effect on it (Table B.1a). T1 and T2 have a positive effect, whereas C2 has no effect on the number of transactions with healthy snacks (Table B.2a). This confirms Hypotheses 1(a), and 4 but rejects Hypothesis 2(a) as we observed in the previous analysis. When we include weekend observations, we find that T1 (resp. T2) increased the number of transactions with healthy snacks by 926.12% (resp. 705.12%). Table B.2a indicates that T1 reduced the overall purchases of unhealthy snacks, whereas T2 and C2 have no statistically significant effect. This confirms Hypotheses 1(b), 2(b), and 4. The sales of unhealthy snacks decreased by 38.97% during T1. None of the interventions had a significant effect on the unhealthy snacks purchased without coffee (Table B.2b). This confirms Hypothesis 3. Overall, we are not introducing data selection bias by excluding weekends from the data in the main analysis. The analysis in this section confirms our findings from the main paper.



(a) Healthy snacks



(b) Unhealthy snacks



(c) Coffee

Figure B.1: Average daily transactions for different types of products under each phase of the experiment (including weekend observations). The error bars indicate the 95% confidence intervals.

Table B.1: Impact of different interventions using weekday and weekend sales on bundles sold.

	Healthy add-on bundle				Unhealthy add-on bundle			
Treatment	Model (1)	Model (2)	Model (3)	Model (4)	Model (1)	Model (2)	Model (3)	Model (4)
T1	42.67*** (4.35)	42.67*** (3.72)	41.21*** (4.59)	42.76*** (3.96)	-47.43*** (9.65)	-47.43*** (7.15)	-45.86*** (9.74)	-47.71*** (7.29)
T2	31.06*** (4.41)	31.40*** (3.77)	30.37*** (4.46)	31.44*** (3.85)	-3.54 (9.77)	-2.53 (7.24)	-2.80 (9.78)	-2.63 (7.30)
C2	1.52 (4.35)	1.52 (3.72)	0.31 (4.52)	1.60 (3.90)	-11.29 (9.65)	-11.29 (7.15)	-11.29 (9.64)	-11.29 (7.19)
No. obs	83	83	83	83	83	83	83	83
Time FE	No	Yes	No	Yes	No	Yes	No	Yes
Stockouts	No	No	Yes	Yes	No	No	Yes	Yes
R^2	0.65	0.76	0.65	0.76	0.28	0.64	0.29	0.64

(a) Healthy and unhealthy bundles

	All coffee add-on bundles			
Treatment	Model (1)	Model (2)	Model (3)	Model (4)
T1	-7.00 (11.42)	-7.00 (7.26)	-7.00 (11.42)	-7.00 (7.26)
T2	25.46* (11.56)	26.80** (7.36)	25.46* (11.56)	26.80** (7.36)
C2	-9.90 (11.42)	-9.90 (7.26)	-9.90 (11.42)	-9.90 (7.26)
No. obs	83	83	83	83
Time FE	No	Yes	No	Yes
Stockouts	No	No	Yes	Yes
R^2	0.13	0.67	0.13	0.67

(b) All coffee add-on bundles

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

Table B.2: Impact of different interventions using weekday and weekend sales on quantity sold.

	Healthy snacks				Unealthy snacks			
Treatment	Model (1)	Model (2)	Model (3)	Model (4)	Model (1)	Model (2)	Model (3)	Model (4)
T1	45.38*** (4.90)	45.86*** (4.11)	43.13*** (5.13)	45.32*** (4.33)	-44.52*** (13.30)	-44.52*** (9.37)	-41.90*** (13.36)	-44.33*** (9.55)
T2	34.55*** (4.96)	35.37*** (4.17)	33.47*** (4.99)	35.15*** (4.23)	0.51 (13.47)	1.80 (9.49)	1.76 (13.42)	1.87 (9.57)
C2	1.76 (4.90)	2.24 (4.11)	-0.12 (5.05)	1.80 (4.27)	-15.38 (13.30)	-15.38 (9.37)	-15.38 (13.23)	-15.38 (9.43)
No. obs	83	83	83	83	83	83	83	83
Time FE	No	Yes	No	Yes	No	Yes	No	Yes
Stockouts	No	No	Yes	Yes	No	No	Yes	Yes
R^2	0.62	0.73	0.62	0.73	0.16	0.61	0.18	0.61

(a) Healthy and unhealthy snacks

	Unhealthy snacks without coffee			
Treatment	Model (1)	Model (2)	Model (3)	Model (4)
T1	2.90 (4.60)	2.90 (3.54)	2.90 (4.60)	2.90 (3.54)
T2	4.05 (4.66)	4.33 (3.58)	4.05 (4.66)	4.33 (3.58)
C2	-4.10 (4.60)	-4.10 (3.54)	-4.10 (4.60)	-4.10 (3.54)
No. obs	83	83	83	83
Time FE	No	Yes	No	Yes
Stockouts	No	No	Yes	Yes
R^2	0.04	0.48	0.04	0.48

(b) Unhealthy snacks without coffee

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

Appendix C DID results

C.1 Stores within a 1-km radius

The first control group for the DID analysis was to use all the (untreated) stores within a radius of 1 km from the treated store. There were two such stores that sold all the products used in our field experiment. We estimated the model in Equation (2) and report the interaction estimated coefficient $\boldsymbol{\beta_d}$ in Table C.3. Model (1) reports the treatment effect without including any control variables, whereas Models (2)–(4) explicitly account for stockouts and time fixed effects. We can see that the sales of healthy snacks were significantly higher in the treated store during both T1 and T2 but did not show any significant change during C2. Similarly, the sales of unhealthy snacks decreased significantly for the treated store during T1 with no effect during T2 and C2. Finally, the sales of unhealthy snacks purchased without coffee remained unaffected by the different interventions. All these results are perfectly aligned with the results obtained in the previous section.

To test the parallel trends assumptions, we used the model in Equation (3) and inspected the estimated value of β_3. Our goal was to check whether $\beta_3 = 0$ with a p-value exceeding 0.05. The estimated β_3 values are reported in Table C.6, hence satisfying the assumption. In Figure C.2, we also graphically convey that the parallel trends assumption was satisfied. Specifically, we plotted the average weekly sales for the different product categories and observed a clear parallel trend.

Table C.3: DID estimates using the stores within a 1-km radius.

	Healthy snacks				Unhealthy snacks			
Treatment	Model (1)	Model (2)	Model (3)	Model (4)	Model (1)	Model (2)	Model (3)	Model (4)
T1	57.27*** (3.00)	57.27*** (2.97)	57.54*** (3.00)	57.48*** (2.98)	-55.03*** (9.40)	-52.40*** (9.17)	-55.03*** (9.40)	-52.40*** (9.17)
T2	42.27*** (3.00)	42.27*** (2.97)	42.67*** (3.01)	42.59*** (2.99)	-4.70 (9.40)	-5.86 (9.17)	-4.70 (9.40)	-5.86 (9.17)
C2	2.17 (3.00)	2.17 (2.97)	2.64 (3.01)	2.55 (2.99)	-5.70 (9.40)	-4.78 (9.17)	-5.70 (9.40)	-4.78 (9.17)
No. obs	177	177	177	177	177	177	177	177
Time FE	No	Yes	No	Yes	No	Yes	No	Yes
Stockouts	No	No	Yes	Yes	No	No	Yes	Yes
R^2	0.90	0.91	0.90	0.91	0.86	0.87	0.86	0.87

(a) Healthy and unhealthy snacks

	Unhealthy snacks without coffee			
Treatment	Model (1)	Model (2)	Model (3)	Model (4)
T1	2.58 (6.83)	2.74 (6.63)	2.52 (6.83)	2.64 (6.63)
T2	1.59 (6.83)	1.59 (6.63)	1.54 (6.83)	1.50 (6.63)
C2	-6.93 (6.83)	-6.87 (6.63)	-7.00 (6.83)	-6.98 (6.63)
No. obs	177	177	177	177
Time FE	No	Yes	No	Yes
Stockouts	No	No	Yes	Yes
R^2	0.86	0.87	0.86	0.87

(b) Unhealthy snacks without coffee

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

Robustness tests with a larger radius.

We next report the DID estimates when using all the untreated stores within a 2-km radius of the treated store as a control group. We have 7 stores within a 2-km radius. For conciseness, we only report the results without including weekend observations. As we can see from Tables C.4, the results are consistent with the estimates of the main specification. Ultimately, using stores from a wider radius strengthens our confidence in our results. For each model, we also ensure that the parallel trends assumption is satisfied (the details are omitted for conciseness). We repeated this analysis for 3-km, which includes 18 stores, and obtained consistent results.

C.2 Clustering using coffee sales

The second control group we considered was determined by clustering stores based on historical coffee sales. Since coffee was the primary product in all the add-on bundles used in our experiment, it seemed natural to select control stores with a similar level of coffee sales. To perform the clustering, we used the weekly aggregated sales of coffee during the 12 weeks prior to the experiment for each store. We constructed a dataset that represents the weekly coffee sales patterns in all stores. This data was scaled before clustering to account for the data variability across the different stores. We implemented the K-means method to cluster the scaled data. Using the elbow method, we determined the optimal number of clusters to be three. We then identified the cluster that contained the treated store and used all the other stores in the same cluster (four of them) as our control stores for the DID analysis. We report the interaction estimated coefficient $\boldsymbol{\beta_d}$ in Table C.5 for healthy snacks, unhealthy snacks, and unhealthy snacks without coffee. The results for healthy, unhealthy snacks and unhealthy snacks

Table C.4: DID estimates using the stores within a 2-km radius.

	Healthy snacks				Unealthy snacks			
Treatment	Model (1)	Model (2)	Model (3)	Model (4)	Model (1)	Model (2)	Model (3)	Model (4)
T1	57.48*** (28.61)	57.40*** (28.76)	57.39*** (27.52)	57.48*** (27.74)	-55.98*** (-4.65)	-56.05*** (-4.65)	-56.72*** (-4.72)	-56.79*** (-4.73)
T2	43.05*** (20.80)	43.01*** (20.87)	42.99*** (20.47)	43.05*** (20.58)	-15.37 (-1.23)	-15.80 (-1.27)	-15.73 (-1.27)	-16.15 (-1.30)
C2	2.34 (1.16)	2.22 (1.11)	2.25 (1.08)	2.30 (1.11)	-12.03 (-0.92)	-12.13 (-0.93)	-12.54 (-0.96)	-12.66 (-0.97)
No. obs	649	649	649	649	649	649	649	649
Time FE	No	Yes	No	Yes	No	Yes	No	Yes
Stock out	No	No	Yes	Yes	No	No	Yes	Yes
R^2	0.90	0.90	0.90	0.90	0.43	0.44	0.44	0.44

(a) Healthy and unhealthy snacks

	Unhealthy without coffee			
Treatment	Model (1)	Model (2)	Model (3)	Model (4)
T1	2.36 (0.42)	2.31 (0.42)	2.09 (0.38)	2.04 (0.38)
T2	-0.28 (-0.05)	-0.49 (-0.09)	-0.43 (-0.08)	-0.65 (-0.11)
C2	-9.06 (-1.50)	-9.10 (-1.50)	-9.24 (-1.53)	-9.29 (-1.54)
No. obs	649	649	649	649
Time FE	No	Yes	No	Yes
Stockouts	No	No	Yes	Yes
R^2	0.23	0.23	0.23	0.23

(b) Unhealthy snacks without coffee

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

without coffee were consistent with those observed in all previous analyses. The results of testing the parallel trends assumption are reported in Table C.6. A visual representation is presented in Figure C.3.

C.3 Parallel trends assumption

This section discusses the parallel trends assumption required for the DID analysis from Sections C.1 and C.2. Table C.6 reports the estimated values of β_3 when estimating Equation (3). The goal is to check whether $\beta_3 = 0$ with a statistically significant p value. Figures C.2 and C.3 plot the number of healthy, unhealthy, and unhealthy without coffee transactions as a function of the time for the different treatments in the treated store along with the average over the control stores. As we can see, the parallel trends assumption is satisfied based on the results from both Figures C.2 and C.3 and Table C.6.

Table C.5: DID estimates using clustering based on coffee sales.

	Healthy snacks				Unhealthy snacks			
Treatment	Model (1)	Model (2)	Model (3)	Model (4)	Model (1)	Model (2)	Model (3)	Model (4)
T1	57.62*** (2.49)	57.62*** (2.48)	57.73*** (2.46)	57.73*** (2.45)	-64.67*** (10.36)	-67.72*** (10.26)	-64.67*** (9.78)	-64.67*** (9.55)
T2	43.33*** (2.49)	43.33*** (2.48)	43.13*** (2.46)	43.12*** (2.45)	-16.62 (10.36)	-19.44 (10.25)	-18.35 (9.78)	-18.25 (9.56)
C2	3.31 (2.49)	3.31 (2.48)	3.10 (2.46)	3.10 (2.45)	-17.63 (10.36)	-16.46 (10.25)	-17.37 (9.78)	-17.26 (9.56)
No. obs	236	236	236	236	236	236	236	236
Time FE	No	Yes	No	Yes	No	Yes	No	Yes
Stockouts	No	No	Yes	Yes	No	No	Yes	Yes
R^2	0.90	0.91	0.90	0.91	0.58	0.6	0.58	0.6

(a) Healthy and unhealthy snacks

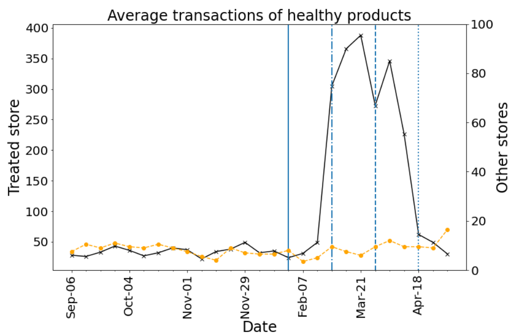
	Unhealthy without coffee			
Treatment	Model (1)	Model (2)	Model (3)	Model (4)
T1	-5.73 (6.84)	-6.14 (6.65)	-5.73 (6.84)	-6.14 (6.65)
T2	-12.16 (6.84)	-9.82 (6.65)	-12.16 (6.84)	-9.82 (6.65)
C2	-1.71 (6.84)	-3.22 (6.65)	-1.71 (6.84)	-3.22 (6.65)
No. obs	236	236	236	236
Time FE	No	Yes	No	Yes
Stockouts	No	No	Yes	Yes
R^2	0.66	0.69	0.66	0.69

(b) Unhealthy snacks without coffee

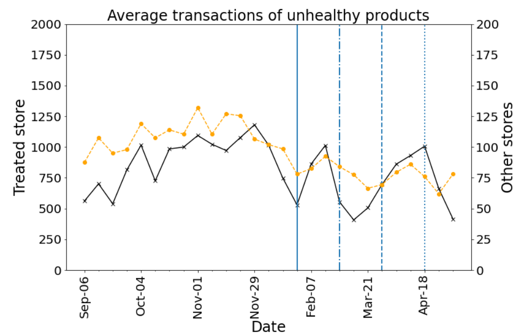
Note: $*p < 0.05$;$**p < 0.01$;$***p < 0.001$. The standard errors are reported in parentheses.

Table C.6: Testing the parallel trends assumption for the DID analysis.

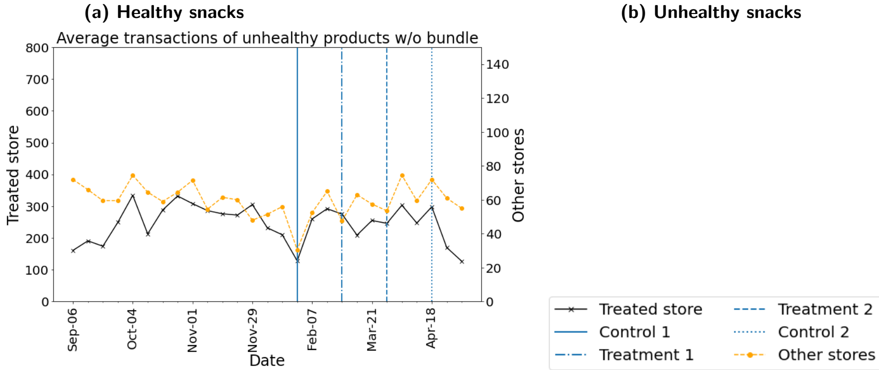
	Healthy snacks		Unhealthy snacks		Unhealthy w/o coffee	
	β_3	p-value	β_3	p-value	β_3	p-value
Stores within a 1-km radius	0.0098	0.312	0.1465	0.393	-0.0197	0.761
Clustering based on coffee sales	0.023	0.051	0.1970	0.216	-0.0086	0.914
SDID	0.0054	0.732	0.35	0.138	0.024	0.787



(a) Healthy snacks

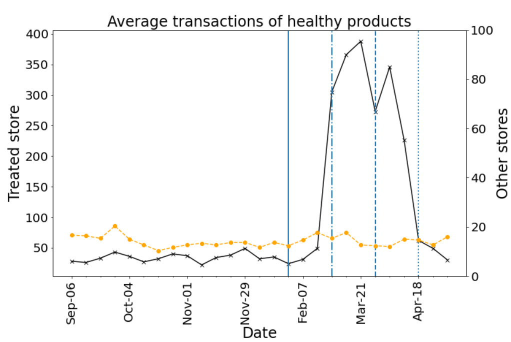


(b) Unhealthy snacks

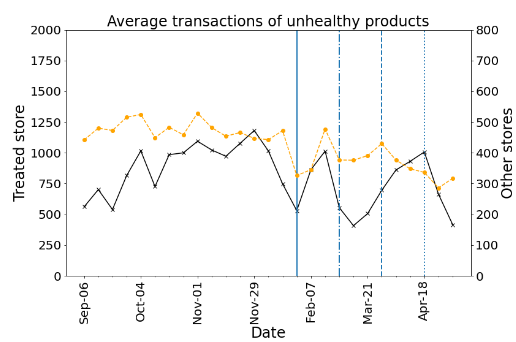


(c) Unhealthy snacks without coffee

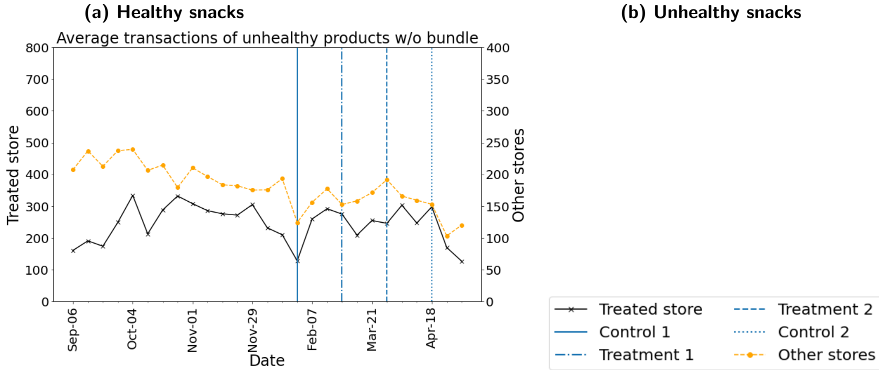
Figure C.2: Parallel trends assumption for stores within a 1-km radius.



(a) Healthy snacks



(b) Unhealthy snacks



(c) Unhealthy snacks without coffee

Figure C.3: Parallel trends assumption using clustering based on coffee sales.

Appendix D Details on online survey

The online survey was designed in Qualtrics and hosted on Amazon MTurk. The survey is anonymous and voluntary and the first page of the survey serves as a consent form (see Figure D.4). Once the survey is completed, the participants cannot withdraw their responses. The Qualtrics survey can be directly accessed using the following link.[12] The link is left active for reference purposes but the new responses are not used in the analysis. The screenshots of the questions are also provided in the sequel. The order in which these questions have appeared in the survey is reported in Figure 5.

Welcome to the research study!

Title of Project: Convenience store shopping survey

Purpose of the Study: This is an invitation to participants in a research study to understand consumers' beverage and snack preferences in a convenience store.

Study Procedures: Each participant will be given a specific budget to spend in a virtual convenience store to buy beverages and snacks. They will be asked a series of questions that replicate their decision making in a store. The response from the participant is used to build their beverage and snack preferences.

Voluntary Participation: As participation is anonymous, withdrawal is not possible after the study session is concluded.

Potential Risks: There are no anticipated risks to you by participating in this research.

Potential Benefits: Participating in the study will have no direct benefits; however we hope to learn consumer's preferences and improve the assortment in the store.

Compensation: Each eligible participant will receive a compensation of $1.

If you have any ethical concerns or complaints about your participation in this study, and want to speak with someone not on the research team, please contact the Associate Director, Research Ethics at 514-398-6831 or **lynda.mcneil@mcgill.ca** citing REB file number 20-05-069.

Please read this document before continuing to the survey. **Submitting your study responses indicates that you consent to participate in this study.** Please save or print a copy of this document to keep for your own reference.

- I consent, begin the study
- I do not consent, I do not wish to participate

Figure D.4: Survey consent form.

[12] https://qfreeaccountssjc1.az1.qualtrics.com/jfe/form/SV_eYe35OBzLabXEEu

Thank you for taking part in this survey.
Please read the following instructions carefully. You will **not** be allowed to go back at any point in the survey.

You are given **a budget of $8** to buy a beverage and a snack at a convenience store. You will be asked a series of questions to select your desired food options. The available options in the store are:

Beverage options

	Small	Medium	Large
House blend coffee (medium or dark roast)	$1.85	$2.09	$2.29
Hot chocolate	$1.89	$2.09	$2.29
Iced coffee	-	$2.79	$2.99
Latte Mokaccino Capuccino	$2.49	$2.99	-
Espresso	$1.99		
Tea	$1.65		

Pastry items

	Price
Muffin	$1.89
Chocolate croissant	$2.59
Croissant	$2.59
Apple Turnover	$2.79
Danish	$2.79
Cinnamon roll	$2.79
Chocolate avalanche	$2.79

Healthy snacks (choice of fruits, vegetables, proteins) - **$3.99 each**

The survey will take **~3 minutes** to complete. When ready, click the arrow below to start the survey.

Figure D.5: Survey introduction page.

Please select "agree" to show you are paying attention to this question.

- Strongly agree
- Agree
- Somewhat agree
- Disagree
- Strongly disagree

(a) Attention question 1

Please pay attention to the image below.

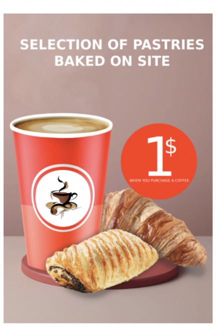
How many pastry items are there in the above image?

- 6
- 8
- 12

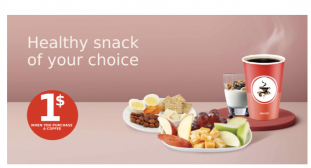
(b) Attention question 2

Figure D.6: Attention questions.

Current promotion: When you buy a beverage, you can get a pastry (e.g., muffin, chocolate croissant, croissant) by paying an **additional $1**.



Current promotion: When you buy a beverage, you can get a healthy snack (choice of fruits, vegetables, proteins) by paying an **additional $1**.



Current promotion: When you buy a beverage, you can either get a pastry (e.g., muffin, chocolate croissant, croissant) **OR** a healthy snack (choice of fruits, vegetables, proteins) by paying an **additional $1**.

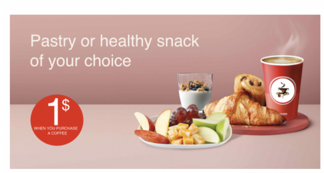


Figure D.7: The three different promotions (treatments) assigned to the respondents.

Remaining budget: $8

Please select the beverage you would like. If you don't want any beverage, select "None".

Below are the available options:

	Small	Medium	Large
House blend coffee (medium or dark roast)	$1.85	$2.09	$2.29
Hot chocolate	$1.89	$2.09	$2.29
Iced coffee	-	$2.79	$2.99
Latte Mokaccino Capuccino	$2.49	$2.99	-
Espresso	$1.99		
Tea	$1.65		

Which beverage do you want to buy?

- House blend coffee (medium or dark roast)
- Hot chocolate
- Iced coffee
- Latte/ Mokaccino/ Capuccino
- Espresso
- Tea
- None

What size do you want?

- Small
- Medium
- Large

← →

Figure D.8: Hot beverage selection question.

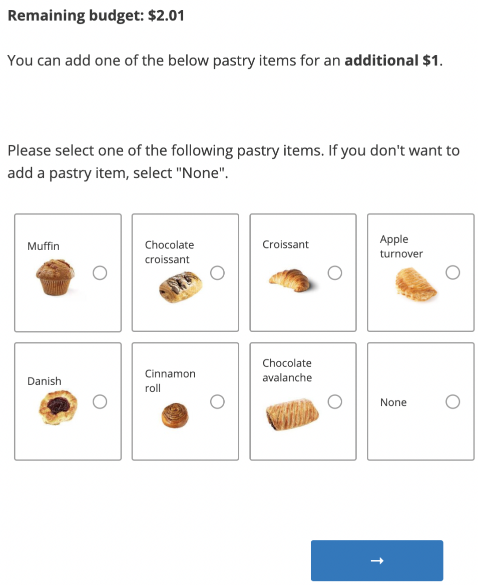


Figure D.9: Pastry selection question for respondents in the Control condition (assuming that a hot beverage was selected).

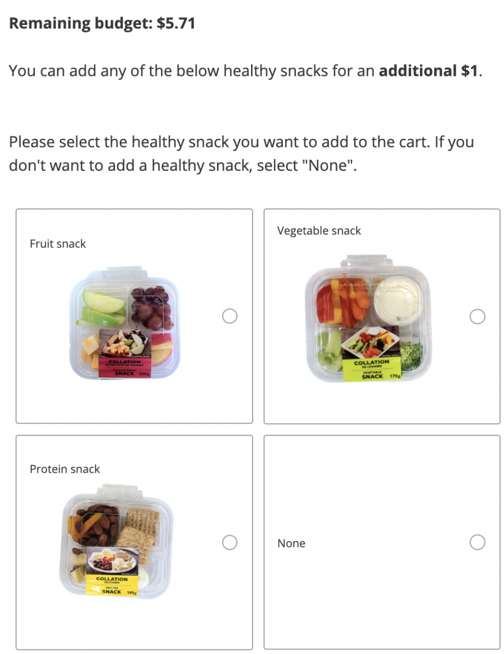
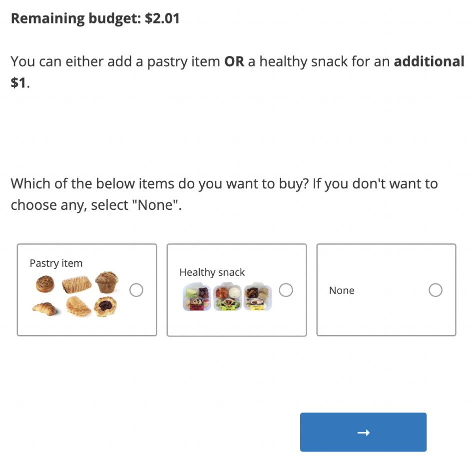


Figure D.10: Healthy snack selection question for respondents in Treatment 1 (assuming that a hot beverage was selected).



Remaining budget: $2.01

You can either add a pastry item **OR** a healthy snack for an **additional $1**.

Which of the below items do you want to buy? If you don't want to choose any, select "None".

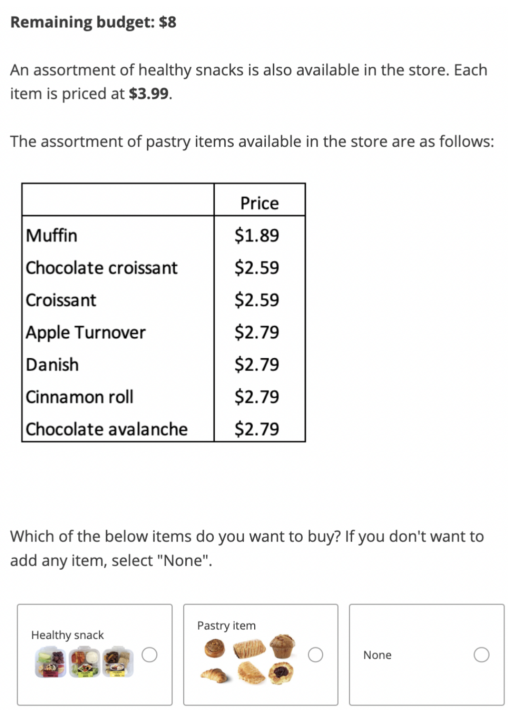
Pastry item

Healthy snack

None

→

Figure D.11: Snack selection question for respondents in Treatment 2 (assuming that a hot beverage was selected).



Remaining budget: $8

An assortment of healthy snacks is also available in the store. Each item is priced at **$3.99**.

The assortment of pastry items available in the store are as follows:

	Price
Muffin	$1.89
Chocolate croissant	$2.59
Croissant	$2.59
Apple Turnover	$2.79
Danish	$2.79
Cinnamon roll	$2.79
Chocolate avalanche	$2.79

Which of the below items do you want to buy? If you don't want to add any item, select "None".

Healthy snack

Pastry item

None

Figure D.12: Snack selection question for respondents who did not select any hot beverage.

Appendix E Lab experiment to compare add-on bundling and price discounting

Objective. We simultaneously ran two new lab experiments to shed light on a potential limitation of our field experiment. Specifically, we are aiming to address the limitation of separating the effect of add-on bundling from the effect of just offering a price discount. The first lab experiment replicated the same survey as in Section 5, whereas the second lab experiment compared add-on bundling with price discounting for healthy snacks. To make an apples-to-apples comparison, we made a modification to the add-on price for the healthy snacks. Instead of using the same price for both the add-on healthy and unhealthy snacks (i.e., extra \$1), which results in a higher discount for the healthy snacks, we now impose the same percentage discount (approximately 47%) for both types of products. This resulted in a new price equal to \$2.25 for both the add-on offer and the discount. This adjustment was done to mitigate the potential influence of the difference in discount rates on consumer decision-making.

Experiment design. The second lab experiment includes three conditions: control, discount, and bundle (Figure E.13). The control condition is the same as in the previous experiment, namely, the unhealthy bundle. The only difference is that the add-on price for the unhealthy snacks is now set to \$1.25 instead of \$1 (Figure E.13a).[13] In the second condition (discount), we simultaneously offer two independent promotions: the usual unhealthy add-on bundle where consumers can add a pastry item for \$1.25 (Figure E.13a) and a discount on healthy snacks, which are sold at \$2.25 instead of the\$4.29 regular price (Figure E.13b). As discussed, we selected a discounted price of \$2.25 to ensure that the percentage discount on the healthy snacks (approximately 47%) matches with the discount on unhealthy snacks. In the third condition (bundle), we offer the healthy bundle while setting the add-on price to \$2.25 (Figure E.13c). By doing so, we ensure that the percentage discount on healthy snacks is the same as the percentage discount on unhealthy snacks (and the same as in the discount condition). The survey can be accessed using the following link.[14]

(a) Unhealthy bundle **(b) Discount on healthy snacks** **(c) Healthy bundle**

Figure E.13: Promotion banners used in the new lab experiment.

The survey closely replicated the decision-making process from the physical store. We randomly assigned one of the three promotions to each participant as seen in Table E.7. Likewise, we randomly assigned each participant a budget of either \$6 (low) or \$10 (high) to be spent in the store. We slightly increased the budget values relative to the previous experiment to ensure that participants can afford the products in light of the price increase. We ran both new lab experiments simultaneously on the Prolific platform.[15] We recruited a total of 2,000 participants to each of the two surveys.

[13] The add-on price was increased by 25¢ due to price inflation across all food categories in 2023.

[14] https://qfreeaccountssjc1.az1.qualtrics.com/jfe/form/SV_ai68S2Rri9259Gu

[15] Prolific is an online survey platform similar to MTurk. Prolific is gaining a reputation for more attentive survey respondents, so we decided to use this platform for our new lab experiments.

Table E.7: New experiment: Promotions offered.

Experiment phase	Promotions offered	Discount offered on unhealthy snacks	Discount offered on healthy snacks	Promotion banner
Control 1 (C1)	Unhealthy add-on bundle	47.35%	0%	Figure E.13a
Treatment 1 (T1)	1. Unhealthy add-on bundle 2. Discount on healthy snacks	47.35%	47.55%	Figure E.13a Figure E.13b
Treatment 2 (T2)	Healthy add-on bundle	0%	47.55%	Figure E.13c

Results. The first result is a sanity check showing that repeating the same survey as in Section 5 on the Prolific platform at a different time period led to consistent results. This provides us reassurance that the participants across both platforms (MTurk and Prolific) are similar. This also conveys that consumer preferences are stable and have not changed between the time the first survey was conducted (March 2022) and the time of the second one (June 2023).

For the second survey, we performed a balancedness test to ensure that the differences across groups are only due to the difference in treatment and not to other factors. Our analysis indicated that participants were indeed assigned in a properly balanced fashion, $\chi^2(2, N = 1,986) = 0.81, p = 0.67$. We performed pairwise t-tests on the likelihood of selecting healthy and unhealthy snacks by the participants and summarized the results in Table E.8. Our analysis revealed a statistically significant increase of 10.16% in the likelihood of selecting healthy snacks when they are offered as part of a bundle, relative to being offered under a price discount. In addition, the likelihood of selecting unhealthy snacks is 38.34% (resp. 45.71%) lower when they are not part of the bundle compared to T1 (resp. C1). These findings strongly support the intuition that using a bundling strategy to incentivize consumers toward healthy food choices is more successful relative to just offering a price discount.

Table E.8: Pairwise comparisons of consumer preferences between the different interventions.

	Healthy snacks		Unhealthy snacks	
	Mean difference	p-value	Mean difference	p-value
C-T1	-0.37	0.00	0.11	0.00
C-T2	-0.45	0.00	0.42	0.00
T1-T2	-0.074	0.001	0.31	0.00

References

Afshin, A. et al. (2017). The prospective impact of food pricing on improving dietary consumption: a systematic review and meta-analysis. PloS One, 12(3).

Ainslie, G. (1975). Specious reward: a behavioral theory of impulsiveness and impulse control. Psychological Bulletin, 82(4):463–496.

Allcott, H., Rogers, T. (2014). The short-run and long-run effects of behavioral interventions: experimental evidence from energy conservation. American Economic Review, 104(10):3003–3037.

An, R. (2013). Effectiveness of subsidies in promoting healthy food purchases and consumption: a review of field experiments. Public Health Nutrition, 16(7):1215–1228.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2019). Synthetic difference in differences (No. w25532). National Bureau of Economic Research.

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

Bekkerman, R., Cohen, M. C., Liu, X., Maiden, J., & Mitrofanov, D. (2021). The Impact of the Opportunity Zone Program on Residential Real Estate. Available at SSRN 3780241.

Bennett, R., Zorbas, C., Huse, O., Peeters, A., Cameron, A. J., Sacks, G., Backholer, K. (2020). Prevalence of healthy and unhealthy food and beverage price promotions and their potential influence on shopper purchasing behaviour: a systematic review of the literature. Obesity Reviews, 21(1).

Carroll, K. A., Samek, A., Zepeda, L. (2018). Food bundling as a health nudge: investigating consumer fruit and vegetable selection using behavioral economics. Appetite, 121:237–248.

Chaurasia, S., Pati, R. K., Padhi, S. S., Jensen, J. M., & Gavirneni, N. (2022). Achieving the United Nations Sustainable Development Goals-2030 through the nutraceutical industry: A review of managerial research and the role of operations management. Decision Sciences, 53(4), 630–645.

Cohen, M. C., Fiszer, M. D., Ratzon, A., Sasson, R. (2021). Incentivizing commuters to carpool: a large field experiment with Waze. Manufacturing & Service Operations Management.

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

Competition Commission (2000). Supermarkets: A report on the supply of groceries from multiple stores in the United Kingdom. Stationery Office.

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

Devlin, A. G., Elmaghraby, W. J., & Hamilton, R. W. (2022). Partitioning cash flows to overcome retailer aversion to stocking new products. Decision Sciences, 53(6), 1048–1067.

Drake, D. F., & Just, R. L. (2016). Ignore, avoid, abandon, and embrace: what drives firm responses to environmental regulation?. Environmentally responsible supply chains, 199–222.

Drewnowski, A., Fulgoni III, V. (2008). Nutrient profiling of foods: creating a nutrient-rich food index. Nutrition reviews, 66(1):23–39.

Exum, B., Thompson, S. H., Thompson, L. (2014). A pilot study of grocery store sales: do low prices= high nutritional quality? Nutrition & Food Science, 44(1):64–70.

Farley, T. A., Rice, J., Bodor, J. N., Cohen, D. A., Bluthenthal, R. N., Rose, D. (2009). Measuring the food environment: shelf space of fruits, vegetables, and snack foods in stores. Journal of Urban Health, 86(5):672–682.

Furey, S. et al. (2019). What's on offer? The types of food and drink on price promotion in retail outlets in the Republic of Ireland (Safefood, the Food Safety Promotion Board, Cork).

Gordon, E., ICF International. (2014). Approaches for promoting healthy food purchases by SNAP participants: ICF International. US Department of Agriculture, Food and Nutrition.

Han, B. R., Sun, T., Chu, L. Y., & Wu, L. (2019). Connecting customers and merchants offline: Experimental evidence from the commercialization of last-mile stations at Alibaba. Available at SSRN 3452769.

Hawkes, C. (2009). Sales promotions and food consumption. Nutrition Reviews, 67(6):333–342.

Helson, H. (1964). Adaptation-level theory: an experimental and systematic approach to behavior. Harper and Row: New York.

Hersey, J. C., Wohlgenant, K. C., Arsenault, J. E., Kosa, K. M., Muth, M. K. (2013). Effects of front-of-package and shelf nutrition labeling systems on consumers. Nutrition Reviews, 71(1):1–14.

Hinnosaar, M. (2023). The persistence of healthy behaviors in food purchasing. Marketing Science, 42(3), 521–537.

Hirschman, E. C., Holbrook, M. B. (1982). Hedonic consumption: emerging concepts, methods and propositions. Journal of Marketing, 46(3):92–101.

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

Katok, E. (2018). Designing and conducting laboratory experiments. Katok, E., Leider, S., & Donohue, K. (Eds.). The Handbook of Behavioral Operations (John Wiley & Sons), 1–33.

Khan, U., Dhar, R., Wertenbroch, K. (2005). A behavioral decision theory perspective on hedonic and utilitarian choice. Pham, M. T., Higgins, E. T., Ratneshwar, S., & Glen Mick, D. (Eds.). Inside Consumption: Frontiers of Research on Consumer Motives, Goals, and Desires (Routledge), 166–187.

Khan, U., Dhar, R. (2010). Price-framing effects on the purchase of hedonic and utilitarian bundles. Journal of Marketing Research, 47(6):1090–1099.

Kroese, F. M., Marchiori, D. R., De Ridder, D. T. (2016). Nudging healthy food choices: a field experiment at the train station. Journal of Public Health, 38(2):e133–e137.

Lachance, P. A., Fisher, M. C. (1986). Educational and technological innovations required to enhance the selection of desirable nutrients. Clinical Nutrition (USA), 5(6):257–264.

Larson, N. I., Story, M. T., Nelson, M. C. (2009). Neighborhood environments: disparities in access to healthy foods in the US. American Journal of Preventive Medicine, 36(1):74–81.

Lobstein, T., Davies, S. (2009). Defining and labelling "healthy" and "unhealthy" food. Public Health Nutrition, 12(3):331–340.

Mensah, G. A. et al. (2015). Mortality from cardiovascular diseases in sub-Saharan Africa, 1990–2013: a systematic analysis of data from the Global Burden of Disease. Cardiovascular Journal of Africa, 26(2):S6–S10.

Mohan, B., Buell, R. W., & John, L. K. (2020). Lifting the veil: The benefits of cost transparency. Marketing Science, 39(6):1105–1121.

Muhammad, A., D'Souza, A., Meade, B., Micha, R., Mozaffarian, D. (2017). The influence of income and prices on global dietary patterns by country, age, and gender. Report, US Department of Agriculture.

Neslin, S. A. (2002). Sales promotion. Weitz, B. A., & Wensley, R. (Eds.). The Handbook of Marketing (Sage), 311–338.

Ni Mhurchu, C., Blakely, T., Jiang, Y., Eyles, H. C., Rodgers, A. (2010). Effects of price discounts and tailored nutrition education on supermarket purchases: a randomized controlled trial. The American Journal of Clinical Nutrition, 91(3):736–747.

Nikolova, H. D., Inman, J. J. (2015). Healthy choice: the effect of simplified point-of-sale nutritional information on consumer food choice behavior. Journal of Marketing Research, 52(6):817–835.

O'Neill, S. et al. (2016). Estimating causal effects: considering three alternatives to difference-in-differences estimation. Health Services and Outcomes Research Methodology, 16(1):1–21.

Putler, D. S. (1992). Incorporating reference price effects into a theory of consumer choice. Marketing Science, 11(3):287–309.

Ravensbergen, E. A., Waterlander, W. E., Kroeze, W., 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):1–10.

Reisch, L. A., Sunstein, C. R., Gwozdz, W. (2017). Beyond carrots and sticks: Europeans support health nudges. Food Policy, 69:1–10.

Roglic, G. (2016). WHO global report on diabetes: a summary. International Journal of Noncommunicable Diseases, 1(1):3–8.

Salmon, S. J. et al. (2015). Social proof in the supermarket: promoting healthy choices under low self-control conditions. Food Quality and Preference, 45:113–120.

Scheidt, D. M., 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.

Shiv, B., Fedorikhin, A. (2004). Heart and mind in conflict: the interplay of affect and cognition in consumer decision making. Journal of Consumer Research, 26(3):278–292.

Sogari, G., Velez-Argumedo, C., Gómez, M. I., Mora, C. (2018). College students and eating habits: A study using an ecological model for healthy behavior. Nutrients, 10(12):1823–1839.

Stern, D., Ng, S. W., Popkin, B. M. (2016). The nutrient content of US household food purchases by store type. American Journal of Preventive Medicine, 50(2):180–190.

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

Thaler, R. H., Sunstein, C. R. (2021). Nudge. Yale University Press.

Van Kleef, E., Otten, K., & van Trijp, H. (2012). Healthy snacks at the checkout counter: a lab and field study on the impact of shelf arrangement and assortment structure on consumer choices. BMC Public Health, 12(1):1–10.

Ver Ploeg, M., Mancino, L., Todd, J. E., Clay, D. M., Scharadin, B. (2015). Where do Americans usually shop for food and how do they travel to get there? Initial findings from the National Household Food Acquisition and Purchase Survey. Report, US Department of Agriculture.

Yilmaz, O., Son, Y., Shang, G., & Arslan, H. A. (2022). Causal Inference under Selection on Observables in Operations Management Research: Matching Methods and Synthetic Controls. Available at SSRN 4310241.