The effect of nudges on food purchasing behavior: a field experiment at a deli chain

Timo Klein
ANR: 727015, SNR: 2040880
Date of submission: June 19, 2023
Program: MSc Thesis Economics, TiSEM
Supervisor: Ben Vollaard
Number of words: 13,937

Abstract

This study investigates the effect of salience and priming nudges on the purchasing behavior of customers in the context of food retail. To investigate this effect, a field experiment is performed on naturally occurring customers at ‘t Smoske, a deli chain in the Netherlands. Three different nudges are implemented and a difference-in-differences analysis, with the use of three ordinary least squares regressions, is performed to analyze the effects. Results show that the placement of a sign displaying the most popular sandwiches significantly increased their daily proportion of sales by 8.6%. The placement of a basket with oranges and a bowl with donuts and waffles out of sight of the customers led to an insignificant decrease of 17.2% in the daily proportion of orange juice sold and a significant decrease of 78.3% in the daily number of donuts and waffles sold. Controlling for week of the day fixed effects did not alter the main effects. No evidence is found for an upward or downward trend in one of the treatment effects over time nor is there found evidence for treatment heterogeneity based on waiting time. Based on the results, the cost-effective nudges could yield benefits within and possibly outside of the context of ‘t Smoske. However, due to the context-specific nature of the nudges and potential infringement with social values, caution should be applied when implementing them.

Keywords: Nudges, deli chain, field experiment, difference-in-differences, proportion/number of sales, treatment heterogeneity, time-varying treatment effects

Table of Contents

1. Introduction *3*

2. Literature Review *6*

2.1 Literature Review **6**

2.2 Hypotheses **14**

3. Data Description and Descriptive Statistics *15*

3.1 Data Description **15**

3.2 Descriptive Statistics **17**

4. Method *24*

4.1 Experimental Design **24**

4.2 Empirical Analysis **26**

5. Results *29*

5.1 Results **29**

5.2 Time-varying Treatment Effects **34**

6. Conclusion *37*

6.1 Discussion and Conclusion **37**

6.2 Policy Implications **40**

6.3 Limitations and Recommendations for Future Research **41**

References *43*

Appendix *46*

1. Introduction

"*The first misconception is that it is possible to avoid influencing people's choices.*"(Thaler & Sunstein, 2008, p. 10)

According to Thaler and Sunstein (2008), and indicated by their quote above, there are many situations in which there is no way of avoiding nudging people in some direction, whether it is intended or not. Thaler and Sunstein (2008) describe nudging as "any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives." They tried to show that it is possible to design a choice architecture to influence people's behavior and that, in many cases, it is really easy to do so. Their proof that small subtle changes can easily and significantly influence people's behavior was the motivation for this study.

There already is a growing body of research motivated by Thaler and Sunstein (2008) suggesting that nudges can lead to significant changes in behavior. However, there are a lot of different kinds of nudges and various contexts in which these nudges can be applied. One nudge is more effective than the other in one context compared to the other. Wilson et al. (2016) categorize nudges into six different categories; priming nudges, salience nudges, default nudges, incentive nudges, commitments and ego nudges, and norms and messenger nudges. Such nudges have been applied in various contexts including healthcare, finance, education, and politics (Benartzi & Thaler, 2007; Damgaard & Nielsen, 2018; Imai et al., 2007; Voyer, 2015). Because of the variety in effects between nudges and contexts, this study specifically investigates the effect of salience and priming nudges in the context of food retail, by means of a field experiment at a deli chain. Salience nudges are described by Wilson et al. (2016) as "novel, personally relevant or vivid examples and explanations used to increase attention to a particular choice." Priming nudges are described by Wilson et al. (2016) as "subconscious cues which may be physical, verbal or sensational, and are changed to nudge a particular choice."

The goal of the study is to further improve the external validity of the effect of nudging and to fill small research gaps. A lot of existing studies in the context of priming and salience nudges in food retail namely have a relatively short treatment period or a small and skewed sample group (Wisdom et al., 2010; Bredie et al., 2019; Olstad et al., 2014; Kroese et al., 2016; Chapman and Ogden, 2012; Cheung et al., 2019). The field experiment in this study ran, on the other hand, for a relatively long time and gathered data on a relatively large and representative sample group compared

to the previous literature. This allows for the identification of possible time-varying treatment effects. Besides this, this is the first study looking into treatment heterogeneity between customers with a short or long waiting time. With these futures of the study, it can make an important contribution to the already existing literature.

The study does however have more advantages and motivation for it can be based on more than the academic literature alone. The relevance of research on priming and salience nudges also lies in their potential to shape behavior in a predictable and desirable way in favor of the nudged person (Thaler and Sunstein, 2008). Libertarian paternalism even sees this as a criterion of nudging, while it says that a nudge can only be justified when it increases the nudged person's well-being (Schmidt & Engelen, 2020). However, by understanding how priming and salience nudges work and when they are most effective, policymakers, researchers, and other practitioners can design cost-effective interventions that not only encourage people to make better choices for themselves. Schmidt & Engelen (2020) contradict with libertarian paternalism by mentioning that these nudges can also lead to more sustainable, or more socially responsible choices. They even mention that nudges can increase a person's well-being while simultaneously increasing the firm's profit. The manager of 't Smoske could thus nudge customers into buying healthier products, while simultaneously increasing its profits.

This paper thus looks into priming and salience nudges. More precisely, this paper attempts to investigate the effect of highlighting specific options on a menu and placing items out of sight of customers on the purchasing behavior of customers at a deli chain. The two research questions of this paper are as follows: "What is the effect of the salience nudge of highlighting specific options of a menu on the number of sales of these options?" and "What is the effect of the priming nudge of placing specific items out of sight of the customers on the number of sales of these items?".

To answer these research questions, this paper conducted a field experiment at a deli chain while it allows for naturalistic observations in a real-world setting. It is an ideal setting for conducting a field experiment on nudging because you can directly alter the choice architecture such as the choices about what to eat or drink. Besides this, small changes in the environment can have a big impact on purchase behavior. Due to the experimental design, it was feasible to allocate treatments to different restaurants instead of to different individuals within the same restaurant, which mitigated the risk of treatment migration. Treatments included placing a sign next to the cash register displaying the most popular sandwiches and placing a basket with oranges as well as a bowl with donuts and waffles out of sight of the customer. A difference-in-differences approach, with the use of three ordinary least squares regressions, is used in the empirical analysis to estimate the effects of the

treatments on the number of sales. It is tested whether the post-treatment trend in the proportion and number of sales in the control restaurant is different from the post-treatment trend in the proportion and number of sales in the treatment restaurants. Data is gathered via the electronic cash register on the daily number of sales per establishment and product sold between January 1 and May 21, 2023. To look if there is treatment heterogeneity between customers with a short or a long waiting time a dummy variable interacting with all other terms is included in the regression. It is equal to one if the number of products sold on a given day is above the median of the entire research period and zero otherwise. This variable then works as a proxy for the presence of a long waiting time. Additional regressions are run and graphs are plotted to look if there is any variation in the treatment effect over time.

The findings of this study show that the placement of a sign next to the cash register displaying the most popular sandwiches led to a significant increase of 8.6% in the daily proportion of sales of these sandwiches compared to the baseline average proportion. A difference in the outcome ranging from a 4.3% to a 12.9% increase is also reasonably compatible with the data. They also show that the placement of a basket with oranges and a bowl with donuts and waffles out of sight of the customers led to respectively an insignificant decrease of 17.2% in the daily proportion of orange juice sold and a significant decrease of 78.3% in the daily number of donuts and waffles sold compared to the baseline average. A difference in the proportion of orange juice ranging from a 36.4% decrease to a 2.0% increase is also reasonably compatible with the data as is a difference in the number of donuts and waffles ranging from a 46.2% to a 110.4% decrease. Controlling for week of the day fixed effects did not alter the main effects. No evidence is found for an upward or downward trend in the treatment effects over time for any of the three nudges nor did a proxy for waiting time indicate any evidence for treatment heterogeneity between customers with a short or a long waiting time.

The remainder of the paper is structured as follows. Chapter 2 provides an overview of the most relevant literature on priming and salience nudging within food retail. Next, the data and its descriptive analysis are presented in chapter 3. Chapter 4 describes the experimental design as well as the empirical analysis used in this paper. In chapter 5, the results of this empirical analysis are documented. Chapter 6 discusses the results and limitations of the study, gives policy implications, and concludes.

2. Literature Review

This chapter reviews the literature that is relevant to the aforementioned research questions. There is focused on studies within the domain of food purchasing behavior and focused on studies looking into the effect of salience and priming nudges in the context of food retail, such as (fast food) restaurants, supermarkets, or canteens. Specific search terms were used such as “nudging”, “nudges”, “priming nudges”, “prime nudges”, “salience nudges”, “nudging in restaurants”, “nudging in food retail”, “field experiment”, “behavioral economics” and “difference-in-differences”.

2.1 Literature Review

Thaler and Sunstein (2008) can be seen as the founding fathers of the concept of nudging. Since the publication of their book “Nudge: Improving Decisions About Health, Wealth, and Happiness” there have been numerous studies and empirical evidence supporting the use of nudges in various contexts. For example, research has shown that providing feedback on energy consumption can significantly reduce household energy usage (Allcott, 2011). Another study suggests that changes in defaults might result in thousands more people donating annually in the United States (Johnson & Goldstein, 2003). This literature review will thus focus on studies investigating the effect of salience and priming nudges in the context of food retail.

2.1.1 Salience Nudges

A study by Wisdom et al. (2010), published in the American Economic Journal, writes about the effect of four different salience nudges on the total calorie intake of customers of a fast-food sandwich restaurant. For the first two nudges, information is given about the daily calorie recommendation and the number of calories of each menu item. For the third and fourth nudge, five of the ten sandwich choices are labeled as "featured" sandwiches on the first page of the menu. These sandwiches were either the ones with the fewest calories, the ones with the most calories, or a combination of both. At the third nudge, the remaining sandwiches can easily be seen on the second page of the menu, while you can only see the other sandwiches at the fourth nudge when you opened a package. The third nudge resembles the treatment in this study where a sign is placed next to the cash register displaying the most popular sandwiches. This treatment namely also features specific sandwiches, while they can also easily choose from other sandwiches displayed on a menu on a large canvas posted above the cashier. Wisdom et al. (2010) gather data on the selected menu items of 638 participating dinners and calculate the impact of the nudges on total caloric intake using an OLS model. The regression includes controls for demographic factors and dummy variables to test for the effect of the featured low-calorie sandwiches on calorie intake.

According to the findings of Wisdom et al. (2010), the third nudge made customers 35% more likely to order low-calorie sandwiches. It did however not affect the total calorie intake, while customers indulged in side dishes and drinks that were higher in calories as a reward for forgoing the unhealthy sandwiches. The stronger fourth nudge did have a significant impact on total calorie intake, while people could only compare with the higher-calorie sandwiches when opening a package. It can however be concluded from both nudges that the menu page with "featured" low-calorie sandwiches did create incentives for customers to select these specific sandwiches.

Another study that examines the salience nudge of featuring certain items from a restaurant menu is a study by Bredie et al. (2019). More specifically, Bredie et al. (2019) use a between-subject design to look at the effect of featuring vegetable dishes as "dish of the day" on the number of sales of these specific dishes. They hypothesize that labeling items as "dish of the day" may make these items more salient and increase the choice of these nudged dishes compared to the alternatives. They also hypothesize that the effect will be stronger for an unappealing dish compared to an appealing dish and that the effect will increase as the number of alternative options increases. The latter is mainly interesting for this study because the number of alternative sandwiches besides the featured sandwiches on the sign is very large. To test these hypotheses, Bredie et al. (2019) conducted an experiment in an experimental restaurant in which a total of 294 customers participated. At the different treatments, specific dishes were featured as "dish of the day". At the point of purchasing, some customers were first exposed to this "dish of the day", after which they could look at the alternatives on the menu board. The other customers were shown the menu board directly. This experimental procedure is very similar to the procedure in this study. Here, people at the treatment restaurants are also exposed to the featured sandwiches as well as the other sandwiches on the canvas above the cashier, while people at the control restaurant are merely exposed to the sandwiches on the canvas. Bredie et al. (2019) then use Pearson's Chi-squared test to compare the number of vegetarian dishes sold as a share of total dishes sold across all treatments. The results from Bredie et al. (2019) show that featuring a vegetarian dish as "dish of the day" significantly increased the number of sales of these dishes by 25.2%. With a non-vegetarian dish as "dish of the day" the increase in the number of sales was still significant but much smaller, 7.6%. The results also show that more alternative dishes increased the number of sales of the featured dishes by 30%. The type of dish and the number of alternative dishes thus influence the size of the effect and are two important moderators for the investigated effect.

Olstad et al. (2014) published another paper that closely resembles the treatment with the sign that is placed next to the cash register displaying the most popular sandwiches. This study promoted

healthy products at an outdoor community pool in Canada by means of a sign, taste testing, and a price reduction. The intervention with the price reduction does not count as a nudge while this intervention significantly changes the economic incentives of the pool guests (Thaler & Sunstein, 2008). The treatment with the sign does however count as a nudge and it closely resembles the treatment with the sign in this study. For this treatment, Olstad et al. (2014) enlarged the signage for healthy products at one of the two establishments and strategically positioned it near their cash register. No modifications were made to the signage for unhealthy products and no intervention took place at the other establishment. The treatment period of the sign had a duration of eight days after which the taste testing and price reduction intervention were added sequentially, both for the same duration of eight days. After this, a control period of once more eight days started. Data on the number of sales of healthy products were collected from the cash register during the entire period. Afterwards, a maximum likelihood analysis and an analysis of covariance were used to examine the data and there was adjusted for air temperature and number of pool patrons. During the signage intervention, sales of healthy products initially rose by 12.7%, but this rise was not statistically significant. When the taste testing intervention and price reduction were added, the effect became significant and led to an increase in the sales of healthy products of respectively 30.4% and 28.7%. The combination of the signage and taste testing nudge thus enhanced the total nudge effect.

2.1.2 Priming Nudges

Thorndike et al. (2017) published an influential paper in Public Health Nutrition exploring the impact of priming nudges in restaurants. The study examines whether increasing the visibility of fresh produce in participating corner stores in Chelsea, America, would affect their monthly sales. The interventions included placing baskets of fresh fruits and vegetables at the store's front or adding a new refrigerator to better showcase the produce. This intervention is comparable to the intervention in this research, which examines the effects of placing a basket of orange juice out of sight of the customers. The paper uses generalized least-squares models with store random effects for their statistical analysis to compare changes in monthly sales trends. The results show that fruit and vegetable sales dropped during the control period in both the treatment and control stores. However, during the treatment period, fruit and vegetable sales rose at the treatment stores while they dropped in the control stores. Increasing the visibility of fresh produce thus increases the number of sales according to Thorndike et al. (2017).

Forwood et al. (2015) is one of the few papers, within the branch of research concerned with nudges, that uses a controlled experiment instead of a field experiment. While Forwood et al. (2015) also look into priming nudging methods, the main focus of the study is not on the impact of priming

nudges on individuals' consumption behavior. Instead, the authors choose to focus on potential moderators of the primary effect and predict that self-reported hunger and cognitive load influence the effect of the priming nudge. To evaluate the impact of these moderators, questionnaire respondents were not only randomly exposed to a prime nudge or no prime nudge, they were also randomly exposed to a cognitive load condition or a control condition and had to report their level of hunger and thirst. For the prime nudge, some of the participants were exposed to advertisements about fruit and a banner featuring a wide selection of fruit, while the other participants in the control group saw no advertisements or banners. Although this exposure takes place in a controlled setting, it still resembles the nudge with the basket of oranges. Afterward, the participants were offered a repeated choice between either a healthy fruit snack or an unhealthy sweet snack. The total quantity of fruits chosen was the outcome of interest. Participants in the cognitive load condition had to keep a series of random letters in mind while making these choices, while the people in the control condition did not have to do so. Already before the exposure to the priming nudge, all participants were asked about their level of hunger and thirst. A logistic regression model was then used to analyze the 143 survey responses. According to the results of Forwood et al. (2015), hunger significantly reduced the odds of fruit snacks being chosen, while there was no moderating effect of thirst and cognitive load. However, while the sample group in this study was not comparable to the population, the external validity of the study is low. Therefore, a second study, with a more representative sample group of 764 respondents, replicated the previous one. This study did not support the previous one and rejected the aforementioned significant effect of hunger.

Two studies on priming nudges that are comparable to the treatment of placing the bowl with donuts and waffles out of sight of the customers are studies by Kroese et al. (2016) and Chapman and Ogden (2012). The first paper conducts a field experiment at three different kiosks at a train station, while the later paper uses a university canteen setting. Both papers look at the effect of placing food items at the cash register on the number of sales of these items, but the findings of the papers are contradictory. Kroese et al. (2016) gathered data for two weeks on a sample of kiosk customers and used an analysis of covariance to examine their results. As hypothesized by Kroese et al. (2016), the nudged products were sold significantly more often in the two shops where these products were placed at the checkout counter compared with the control kiosk. In the nudge period, the sales of the nudged products rose by 78.26%, while it rose by 52.17% in the nudge plus disclosure period. Chapman and Ogden (2012) use a chi-square analysis on the data generated by 180 customers to look at differences in sales. Their results are not in line with their hypothesis nor with the hypothesis and results of Kroese et al. (2016). They namely found that the number of sales of fruits and confectionary decreased by respectively 26.1% and 14.2% when they were placed next to the cash register. The

study of Kroese et al. (2016) was later replicated by the same author and some co-authors, where the treatment period was extended considerably (Van Gestel et al., 2018). The nudge was now only implemented in one kiosk but for a duration of eight weeks instead of two. The first four weeks served as the control period, while the last four weeks served as the treatment period. The replication study used the same analysis of covariance as the paper from Kroese et al. (2016) to analyze the sales data, which again led to a significant increase of nudged products sold, this time of 46.51%. However, Van Gestel et al. (2018) performed an additional analysis of covariance with the week number included as independent variable to allocate the presence of a varying treatment effect over time. This additional analysis shows no variation in treatment effect over time and shows that each week in the control period is significantly different from each week in the treatment period. This replication study from Van Gestel et al. (2018) thus supports the previously obtained results from Kroese et al. (2016) and also showed that the effect persisted over time. The paper by Van Gestel et al. (2018) is one of the few papers that deals with time-varying treatment effects and emphasizes the importance of it. This paper will also explore such dynamic treatment effects.

2.1.3 Salience and Priming Nudges

There are two studies that resemble the salience as well as the priming nudge from this study, which include a paper from Foster et al. (2014) published in The American Journal of Clinical Nutrition and a paper from Cheung et al. (2019) published in the BMC Public Health. Both studies try to promote the sales of specific food items, Foster et al. (2014) in a supermarket environment and Cheung et al. (2019) at a take-away food vendor inside an academic hospital. To influence the sales, Foster et al. (2014) use a wide range of marketing strategies including two nudges to increase the placement and visibility of the targeted products and a nudge where signs with the targeted products are placed in the supermarket. These nudges correspond to the nudges in this study where donuts, waffles, and oranges are placed out of sight of the customers and where a sign featuring specific items is placed on the counter. Cheung et al. (2019) implemented three different nudges; an accessibility nudge where certain products were placed next to the cash register, a salience nudge where the appearance of certain products had been improved and a social norm nudge where products were placed on a sign and described as the "bestselling choice". These nudges correspond to the nudges where the donuts and waffles and a basket with oranges are placed out of sight of the customers and the nudge with the sign displaying the most popular sandwiches.

The study from Foster et al. (2014) consisted of a three-month control period directly followed by a six-month intervention period and was conducted in eight urban supermarkets located in Philadelphia and Wilmington. The stores were divided into four matched pairs based on similar

characteristics. Four stores implemented the wide range of marketing strategies including the aforementioned nudges, while the other matched stores served as control stores. To analyze the weekly sales per supermarket, a linear mixed-effect model was used with fixed effects for the treatment group, the treatment period, and the interaction between these two. The results of this analysis show that nudges can positively affect the sales of the targeted items, but that the strength of the effect varies strongly between different product categories. The sales at the treatment and control stores were significantly different for some milk products, some frozen meals, and water. There were however no significant differences between the treatment and control supermarkets for other frozen meals, regular and diet soda, and cereal. As also shown by Bredie et al. (2019), the type of product being nudged is an important moderator for the effect, and even within the same category, the effect differs per product.

The study from Cheung et al. (2019) consisted of a week-long control period, a week-long treatment period, and a week-long treatment plus disclosure period. In this disclosure period, the purpose of the nudge was disclosed to all customers by means of a sign with a one-sentence message. Between the treatment period and the treatment plus disclosure period, a four-week-long interim period was implemented to prevent any effects being carried over from the treatment week to the treatment plus disclosure week. A Chi-square test was then used to compare the number of sales of the targeted products to the number of sales of comparable products across the control period, the treatment period, the interim period, and the nudge plus disclosure period. The results from Cheung et al. (2019) show that the accessibility nudge led to a significant increase in the proportion of sales of the targeted products of 73.3% in the treatment period and 82.2% in the treatment plus disclosure period. The salience nudge on the other hand, led to an insignificant increase in the proportion of sales of the targeted products of 9.3% in the nudge week and 12.4% in the nudge plus disclosure week. For the social norm nudge, it was impossible to conduct a statistical analysis due to the dataset being too small.

Various studies in this literature review (Bredie et al., 2019; Chapman & Ogden, 2012; Cheung et al., 2019; Wisdom et al., 2010) randomly assign treatments on the individual level in the same restaurants and thus make use of a randomized field experiment. However, the treatments at the deli chain of this study were too obvious and unavoidable. It would, therefore, have been impossible to rule out the possibility of treatment migration where customers in the control group are also exposed to the treatments. A regression discontinuity design would also have been suboptimal, while ‘t Smoske has a lot of regular customers. This would lead to a high chance of customers in the control group obtaining the experimental treatment if treatments are turned on on different days. To reduce

the chance of treatment migration treatments are assigned at restaurant level to three establishments located in different cities in the province of Zeeland. The difference-in-differences design in combination with a natural field experiment is therefore the most applicable design to estimate the causal effect of the treatments in a real-world food choice setting.

Other limitations of some studies in this literature review are the homogeneous as well as the small sample groups. The setting of the field experiment can lead to customers having the same characteristics and a sample group not being representative of the entire population. For example, a field experiment in a university canteen or a hospital cafeteria (Chapman & Ogden, 2012; Cheung et al., 2019) can lead to a skewed proportion being highly educated or having the same profession. While the restaurants in this study are accessible to all kinds of customers, the risk of having a skewed dataset is minimized and it is thus way more likely that the sample group is representative of the entire population. A small sample group or a short treatment period can also lead to some severe issues. For example, the study by Kroese et al. (2016) only lasted for two weeks, while the study by Forwood et al. (2015) only included 143 subjects. Such a short treatment period or small sample size may not provide enough time or information to adequately assess the effectiveness of a treatment. This study examines a period of twenty weeks for the salience nudge with the sign and a period of ten weeks for the priming nudge with the donuts and waffles and the basket of oranges. This can provide a more comprehensive picture of the treatment effect. The design and the setting of this study thus ensure a valuable contribution to the already existing literature.

2.1.4 Overview of the Effect Sizes

Table 1. Results from previous literature for the effects of nudges similar to the salience nudge of placing a sign next to the cash register displaying the most popular sandwiches on the proportion of sales of these nudged sandwiches

Paper	Specification	Sign of the effects	Size of the effects (compared to baseline)	Significant
Wisdom et al. (2010)	Effect of featured sandwiches on sales of low-calorie sandwiches	+	35%	Yes
Bredie et al. (2019)	Effect of "dish of the day" on sales of non-vegetarian dishes	+	7.6%	Yes
	Effect of "dish of the day" on sales of vegetarian dishes	+	25.2%	Yes
	Effect of "dish of the day" on sales of vegetarian dishes, with two alternatives instead of one	+	30%	Yes
Olstad et al. (2014)	Effect of sign near cash register on sales of healthy products	+	12.7%	No

Note: The "Significant" column indicates self-reported significance, the level can thus differ between papers. The paper from Foster et al. (2014) is excluded while the effect per type of nudge is not reported.

Table 2. Results from previous literature for the effects of nudges similar to the priming nudge of placing a basket that is filled with oranges out of sight of the customers on the proportion of sales of orange juice

Paper	Specification	Sign of the effects	Size of the effects (compared to baseline)	Significant
Thorndike et al. (2017)	Effect of increased visibility of fresh products on their sales	+	19.1%	Yes
Forwood et al. (2015)	Effect of exposure to fruit on the number of fruit snacks chosen (study 1)	+	11%	No
	Effect of exposure to fruit on the number of fruit snacks chosen (study 2)	+	5%	No
Cheung et al. (2019)	Effect of improved appearance of products on sales proportion	+	9.3%	No
	Effect of improved appearance of products on sales proportion plus disclosure of the nudge	+	12.4%	No

Note: The size of the effect in the paper of Thorndike et al. (2017) is estimated by eyeballing a graph. The size of the effect in the paper of Forwood et al. (2015) is not the main outcome effect of the paper but they are computed by converting odds ratios. The "Significant" column indicates self-reported significance, the level can thus differ between papers. The paper from Foster et al. (2014) is excluded while the effect per type of nudge is not reported.

Table 3. Results from previous literature for the effects of nudges similar to the priming nudge of placing a bowl with donuts and waffles out of sight of the customers on the number of sales of donuts and waffles

Paper	Specification	Sign of the effects	Size of the effects (compared to baseline)	Significant
Kroese et al. (2016)	Effect of placing products next to check out on their sales	+	78.26%	Yes
	Effect of placing products next to the cash register on their sales plus awareness of the nudge	+	52.17%	Yes
Van Gestel et al. (2018)	Effect of placing products next to the cash register on their sales	+	46.51%	Yes
Chapman & Ogden (2012)	Effect of positioning fruit by the checkout on their sales	-	26.1%	Yes
	Effect of moving confectionery away from the checkout on their sales	+	14.2%	Yes
Cheung et al. (2019)	Effect of placing products next to checkout on sales proportion	+	73.3%	Yes
	Effect of placing products next to checkout on sales proportion plus disclosure of the nudge	+	82.2%	Yes

Note: The "Significant" column indicates self-reported significance, the level can thus differ between papers. The paper from Foster et al. (2014) is excluded while the effect per type of nudge is not reported.

2.2 Hypotheses

Based on previous literature, the following hypotheses can be formulated for this paper:

1. Placing a sign next to the cash register displaying the most popular sandwiches significantly increases the proportion of sales of these nudged sandwiches.
2. Placing a basket that is filled with oranges out of sight of the customers significantly decreases the proportion of sales of orange juice.
3. Placing a bowl with donuts and waffles out of sight of the customers significantly decreases the number of sales of donuts and waffles.

3. Data Description and Descriptive Statistics

In this chapter, the data used for this study is explained and summarized. In section 3.1 the data description, and the data cleaning process is given. In section 3.2 descriptive statistics are shown to simplify and understand the large datasets.

3.1 Data Description

The data used in this study is data on the daily number of sales per establishment of all the sandwiches, waffles, donuts, soup, and sodas and fresh juices sold between January 1 and May 21, 2023. Data on the daily number of in-store sales were gathered via the system of Vectron Systems AG used for the electronic cash register of Eijsink Afrekensystemen B.V. Data on the daily number of online sales was gathered via the software platform Booq developed by Eijsink Afrekensystemen B.V. (German Wholesaler METRO Acquires Eijsink|Booq, 2023). The obtained raw data was loaded into R-Studio and cleaned thoroughly in order to use it for analysis in this study.

The data were collected from January 1, 2023, to May 20, 2023, covering a period of 140 days (twenty weeks) and a sample of three participating establishments. In this period, 167,856 products were sold to 79,257 individual customers across all establishments. However, these numbers came about after missing data had been added. In-store sales data was namely missing due to technical problems at the cash register due to the malfunctioning of the system of Vectron Systems AG. Data were missing for Goes on January 1, January 30, February 6 through February 10, and April 9. Sales data was also missing for Middelburg but only on January 1. This missing data has been supplemented with sales data from a comparable day in the previous or the following week. Besides this, due to another technical problem, the in-store sales data of March 8 at all three establishments also contained the sales of March 7. The same holds for the sales data of Goes on February 3 which contained the sales of February 2. While it was impossible to tell how the sales were divided between the two days they were just equally divided between them. This method of filling up the missing data did not lead to any implications for the final results of the difference-in-differences analysis due to the careful selection of comparable dates as well as the relatively long research period.

The dataset of the online sales contained rows with payment status “Not paid” and with the location “Boulevard”. Rows with the payment status “Not paid” were removed while the products in this order did not actually get sold. Rows with the location “Boulevard” were also removed while the sales at this establishment did not fall within the scope of this study. Lastly, some products like “Frozen yogurt” or “Soup of the day” were removed from the in-store datasets as well as the online

dataset while they were irrelevant within this study. These removals also did not influence the results of the statistical analysis.

After proper data cleaning, three datasets were created which showed respectively the proportion of nudged sandwiches sold, the proportion of orange juice sold, and the number of donuts and waffles sold per day and establishment. The proportion of nudged sandwiches is equal to the number of cheese and ham sandwiches, chicken curry sandwiches, and shoarma salad sandwiches as a share of the total number of sandwiches sold and the proportion of orange juice is equal to the number of orange juices sold as a share of the total number of drinks sold. For the donuts and waffles, the number of sales instead of the proportion of sales is used. This is done because ‘t Smoske does not sell any substitutes for these products. A preview of the dataset with the proportion of nudged sandwiches sold is shown in table 4. The column “Location” indicates the establishment, while the column “Day” indicates the day of the research period. The column “Treatment group” is one if the column “Location” is equal to “Middelburg & Goes” and zero if the column is equal to “Vlissingen”. The column “Treatment period” is zero for the control period and one for the treatment period and the column “Treatment indicator” multiplies the column “treatment group” and “treatment period”. This column is equal to one for the treatment group during the treatment period and zero otherwise. The column “Weekday” shows the day of the week while the column “Wait” is equal to one if the number of products sold on a given day is above the median of the entire research period and zero otherwise. Lastly, the column “Week” indicates the week of the research period. This dataset was sufficient for the statistical difference-in-differences analysis. A more detailed description of the data-cleaning process is provided in the appendix.

Table 4. A preview of the dataset with the proportion of nudged sandwiches sold used for difference-in-differences analysis

Proportion	Location	Day	Treatment group	Treatment period	Treatment indicator	Weekday	Wait	Week
0.3171355	Middelburg & Goes	1	1	0	0	Sunday	0	1
0.2213115	Vlissingen	1	0	0	0	Sunday	1	1
0.2857143	Middelburg & Goes	2	1	0	0	Monday	1	1
0.3016241	Vlissingen	2	0	0	0	Monday	1	1
0.2810591	Middelburg & Goes	3	1	0	0	Tuesday	0	1
0.2706767	Vlissingen	3	0	0	0	Tuesday	1	1
0.2510204	Middelburg & Goes	4	1	0	0	Wednesday	1	1
0.2592593	Vlissingen	4	0	0	0	Wednesday	1	1
0.2821369	Middelburg & Goes	5	1	0	0	Thursday	1	1
0.2943548	Vlissingen	5	0	0	0	Thursday	1	1

3.2 Descriptive Statistics

Table 5. Descriptive statistics of the number and proportion of sales

	All			Restaurant: Control restaurant			Restaurant: Treatment restaurants		
Variable	Mean no. sales	Mean share	Min-max	Mean no. sales	Mean share	Min-max	Mean no. sales	Mean share	Min-max
Control period									
Sandwiches	132.03 (40.34)	0.281 (0.024)	0.221-0.350	116.51 (24.58)	0.283 (0.025)	0.221-0.331	147.54 (46.75)	0.280 (0.024)	0.237-0.350
Orange juice	18.94 (5.51)	0.156 (0.045)	0.058 - 0.303	18.2 (5.31)	0.162 (0.044)	0.081-0.264	19.69 (5.67)	0.149 (0.045)	0.058-0.303
Donuts & waffles	5.91 (2.95)	-	1-12	6.43 (2.80)	-	2-12	5.40 (3.05)	-	1-12
Treatment period									
Sandwiches	133.56 (33.81)	0.292 (0.029)	0.221-0.366	113.67 (22.65)	0.282 (0.027)	0.221–0.345	153.46 (31.42)	0.303 (0.027)	0.237-0.366
Orange juice	18.26 (5.71)	0.142 (0.045)	0.008 - 0.248	20.23 (4.92)	0.161 (0.042)	0.084-0.248	16.29 (5.82)	0.122 (0.040)	0.008-0.222
Donuts & waffles	3.86 (3.43)	-	0-12	6.49 (2.70)	-	1-12	1.23 (1.54)	-	0-5

Note: standard errors in parenthesis.

Descriptive statistics on the number and proportion of nudged products sold are shown in table 5. As can be seen under the descriptive statistics of all establishments during the control period, the daily average number of nudged sandwiches sold is slightly higher than 130. As stated in section 3.1 and as shown in figure 12 in the appendix, these nudged sandwiches are the cheese and ham sandwich, the chicken curry sandwich, and the shoarma salad sandwich. On average, slightly less than 20 orange juices and slightly less than 6 donuts and waffles get sold each day during the control period. It can be seen that the mean of the combined number of nudged sandwiches sold at the treatment restaurants in the control period is higher than the mean of the nudged sandwiches sold at the control restaurant (147.54 > 116.51). While there is no significant difference between the share of nudged sandwiches sold (0.280 and 0.283) it can be concluded that the daily combined customer population of the treatment restaurants is in general higher than the daily customer population of the control restaurant. The mean number of orange juice sold in the control period at the treatment restaurants is, however, only slightly higher compared to the mean at the control restaurant (19.69 > 18.2). The mean number of donuts and waffles sold is even slightly lower at the treatment restaurants during the control period compared to the control restaurant (5.40 < 6.43).

The minimum and maximum proportion of nudged sandwiches sold, the minimum and maximum proportion of orange juice sold, and the minimum and maximum number of donuts and waffles sold are also shown in table 5. The daily proportion of nudged sandwiches sold ranges from 22.1% to 35.0% in the control period and does not vary considerably between the treatment restaurants and the control restaurant. The daily proportion of orange juice sold in the control period does vary between the treatment restaurants and the control restaurant. The minimum at the control restaurant is equal to 8.1% while the maximum is equal to 26.4%. The range of the combined proportion at the treatment restaurants is somewhat wider and ranges from 5.8% to 30.3%. Lastly, the control period included a day on which only 1 donut or waffle was sold and a day where a number of 12 donuts and waffles got sold.

To get a first glimpse of the difference-in-differences treatment effect we need to look at the differences in the means between the treatment period and the control period for the treatment restaurants as well as the control restaurant. For the nudged sandwiches and the orange juice, we need to look at the mean proportion, while we look at the mean absolute numbers for the donuts and waffles. These means are stated in table 5 but are also shown graphically with the bar plots in figures 1, 2, and 3 and with the dotted lines in the difference-in-differences graphs in figures 4, 5, and 6. As the figures clearly show, there is almost no difference between the means of all nudged products between the control period and treatment period in the untreated restaurant in Vlissingen. The mean proportion only changes from 28.3% to 28.2% and from 16.2% to 16.1% for respectively the proportion of nudged sandwiches sold and the proportion of orange juice sold. The mean number of donuts and waffles sold only changed from 6.43 to 6.49 between the control and treatment period. The bar plots and dotted lines in figure 1 to 6 show, however, an apparent difference between the means in the control and treatment period at the treatment restaurants. The placement of the sign displaying the most popular sandwiches led to an increase in the proportion of nudged sandwiches sold from 28.0% to 30.3%. Placing the basket with oranges as well as the bowl with donuts and waffles out of sight of the customers led to a decrease in the proportion of orange juice sold from 14.9% to 12.2% and a decrease in the mean number of donuts and waffles sold from 5.40 to 1.23. Figures 1, 2, and 3 also show the 95%-confidence intervals belonging to the means. It can be seen that those intervals do not overlap at the treatment restaurants for the proportion of nudges sandwiches sold and the number of donuts and waffles sold. They do however overlap at the treatment restaurants for the proportion of orange juice sold. This means that the difference in the means is significant for the proportion of nudges sandwiches sold and for the number of donuts and waffles sold, but we cannot conclude that the difference in the mean between the control and treatment period at the treatment restaurants are significant for the proportion of orange juice sold.

Figure 1. Proportion of nudged sandwiches sold per control and treatment group and period

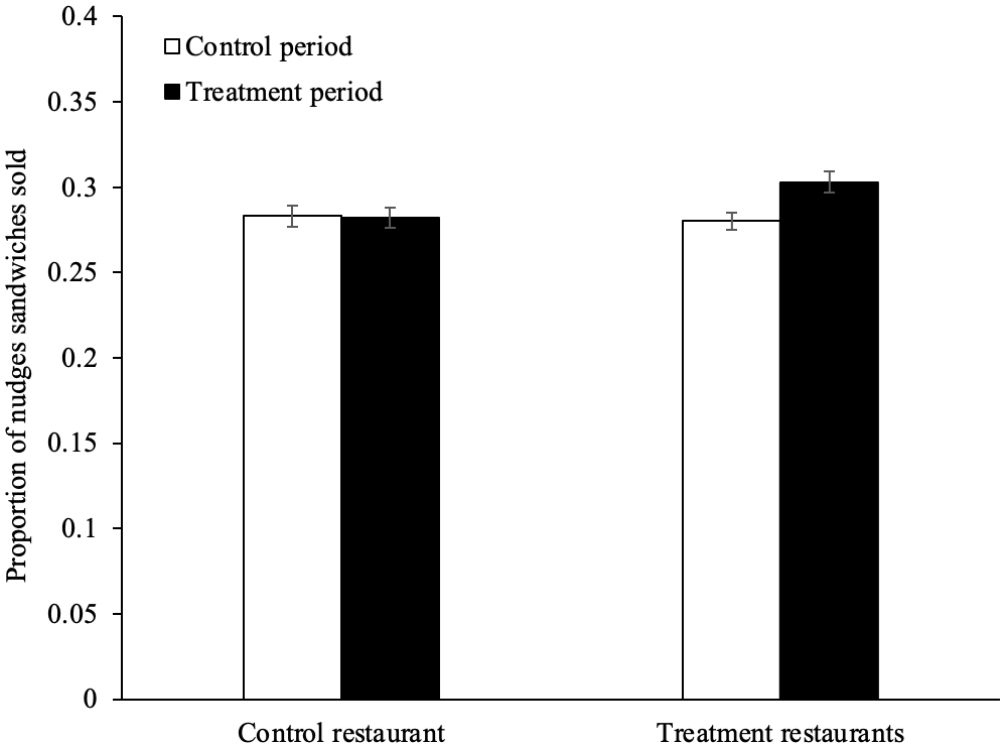


Figure 2. Proportion of orange juice sold per control and treatment group and period

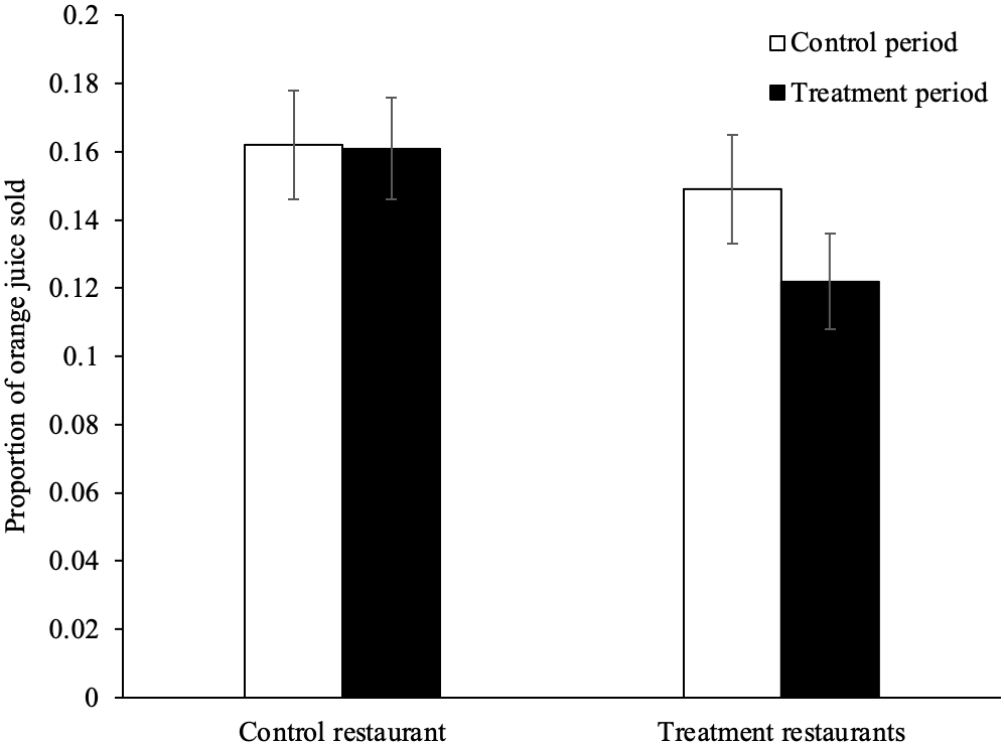
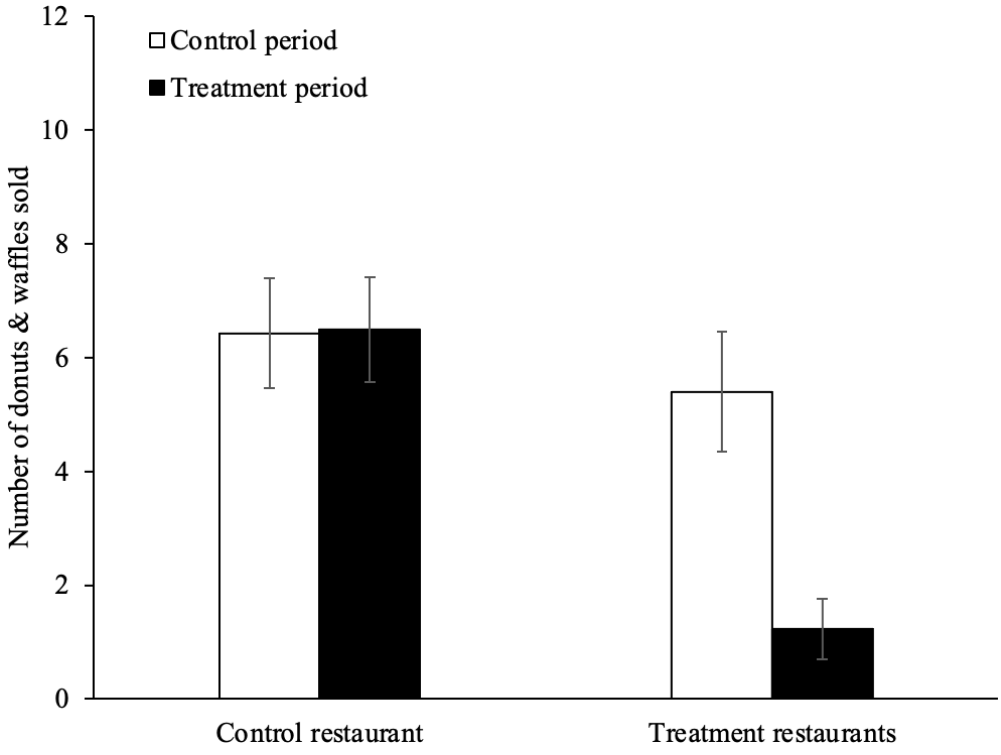
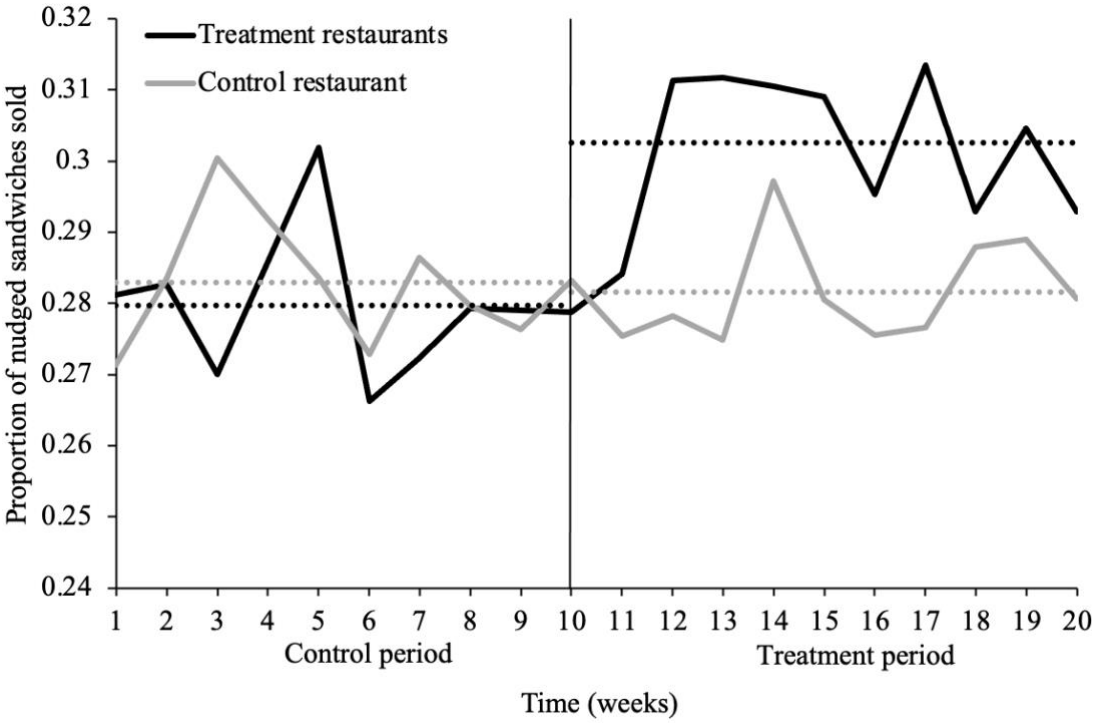


Figure 3. Number of donuts & waffles sold per control and treatment group and period



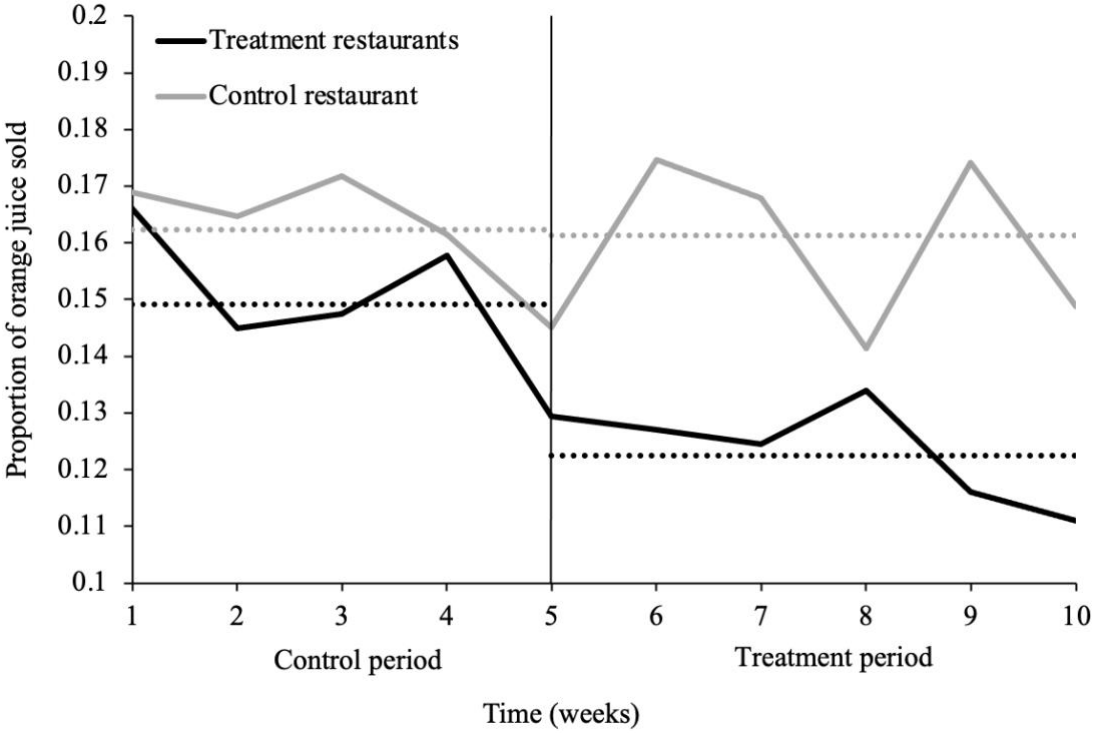
Figures 4, 5, and 6 can also be used to investigate the common trend assumption. Without graphical evidence, we can already name several factors that support the assumption of common trends. Each establishment namely belongs to the same manager, has the same interior and exterior to a certain extent, and also has the exact same menu. The establishments are also located in the same area which ensures the same weather conditions and the same seasonal effects on purchasing behavior. We can now look at the trends in the proportion of nudged sandwiches and orange juice sold and at the trends in the number of donuts and waffles sold. The trends in the control period in figures 4, 5, and 6 namely provide additional support for the common trend assumption. Although there is a lot of variability in the weekly proportion and number of nudged products sold, there is no clear upward or downward trend before the introduction of the nudges and the lines run roughly the same. The introduction of the treatment immediately causes an upward trend in the number of sandwiches sold and a downward trend in the proportion of orange juice and the number of donuts and sandwiches sold. The lines deviate from each other in the treatment period which is not the case in the control period. We can therefore assume this is due to the nudges and assume that the pretreatment trend would have carried on had it not been for these nudges. Figures 4, 5, and 6 thus provide support for the assumption of common trends.

Figure 4. Proportion of nudges sandwiches sold per week



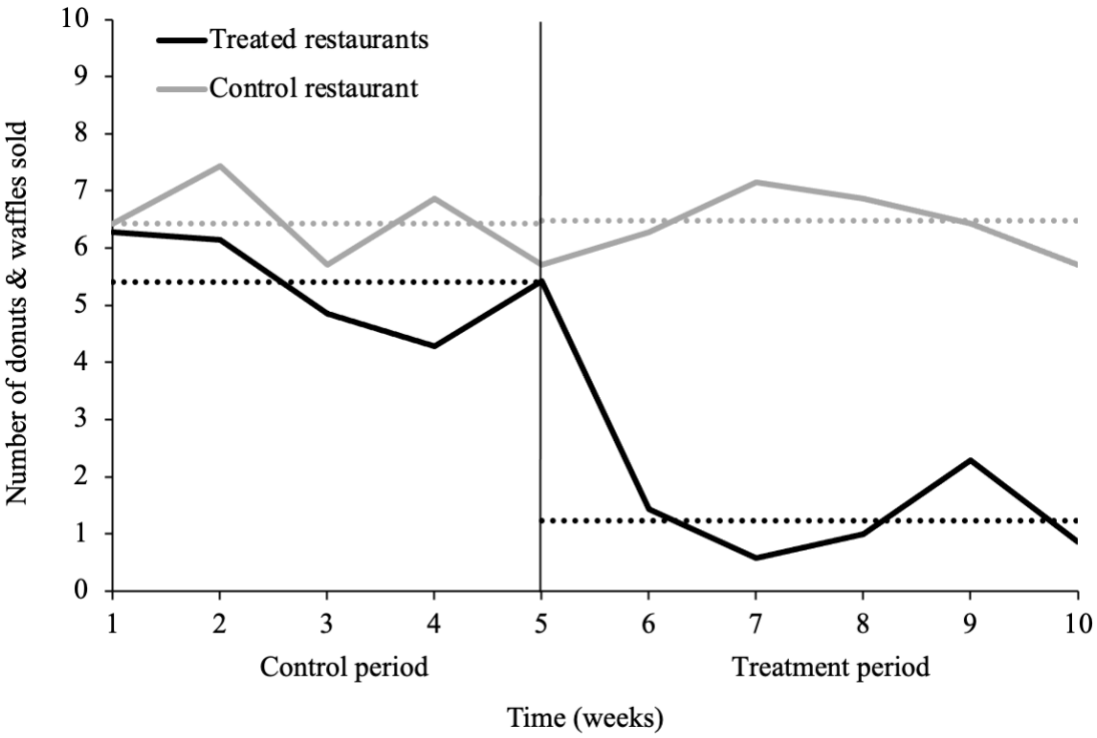
Note: treatment starts after week 10. The dotted lines indicate the mean proportion of nudged sandwiches sold in the control and treatment group and period.

Figure 5. Proportion of orange juice sold per week



Note: treatment starts after week 5. The dotted lines indicate the mean proportion of orange juice sold in the control and treatment group and period.

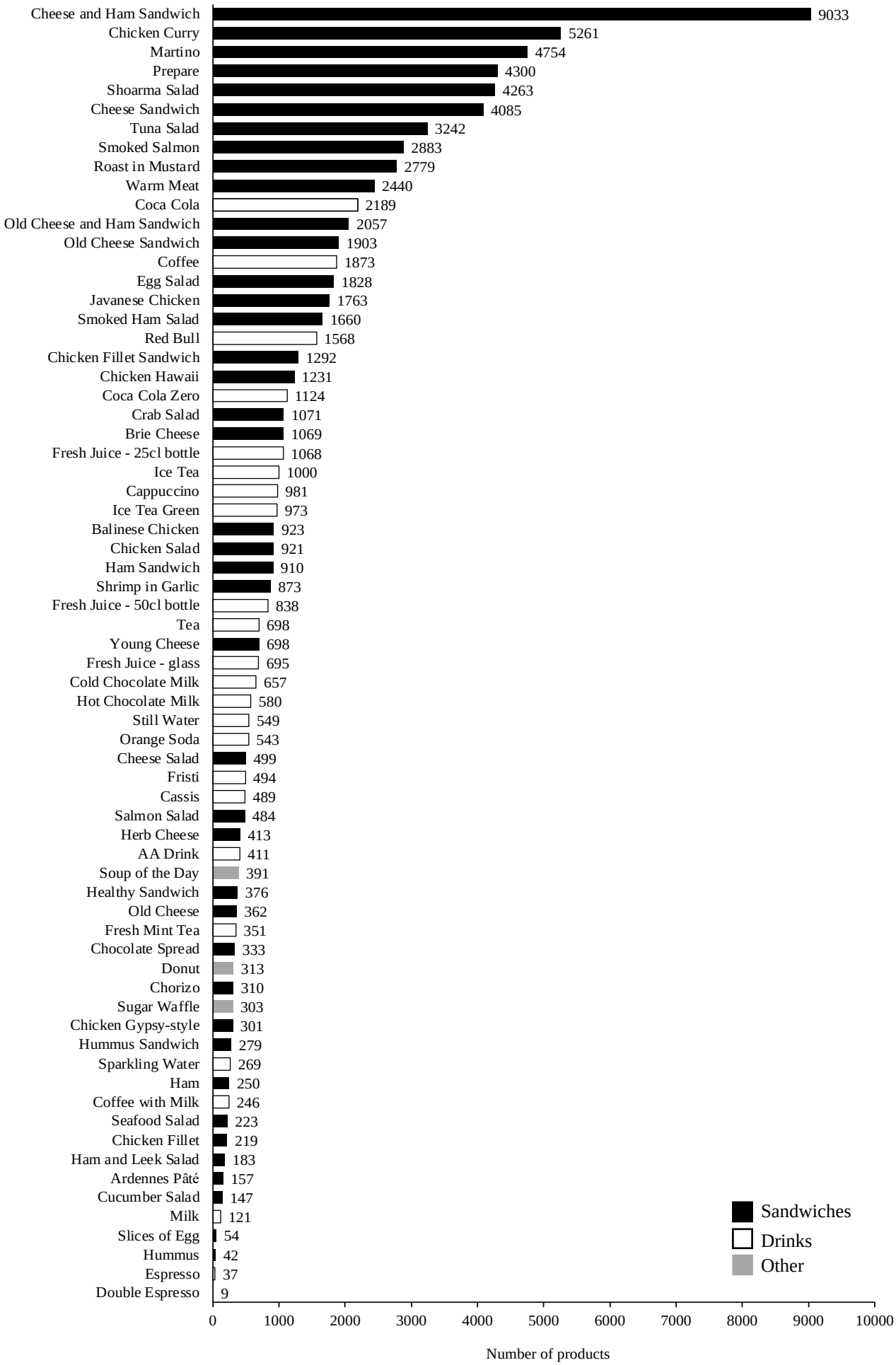
Figure 6. Number of donuts and waffles sold per week



Note: treatment starts after week 5. The dotted lines indicate the mean number of donuts and waffles sold in the control and treatment group and period.

Figure 7 shows an overview of the number of sales between January 1, 2023, and March 11, 2023, for all products offered by ‘t Smoske. The products are categorized into three different product categories, which are sandwiches, drinks, and other products. The figure clearly shows that the sandwich category is the main focus for ‘t Smoske, while the top ten most sold products only contain sandwiches. The cheese and ham sandwich, the chicken curry sandwich, and the shoarma salad sandwich all fall within the top five most sold sandwiches and are therefore chosen to be featured on the sign with the most popular sandwiches. It can also be seen that fresh juice is sold less than Coca-Cola, coffee, and Red Bull and that fresh juice in a bottle of 25 centiliters is sold more than in a bottle of 50 centiliters. Donuts are sold slightly more than waffles, but both products are not the main focus of ‘t Smoske.

Figure 7. Number of products sold between 01/01/2023 and 11/03/2023, without any intervention



4. Method

In this chapter, the experimental design and the empirical analysis are described. Section 4.1 describes the experimental setting as well as the implemented nudges. In section 4.2 the estimation equations for this study are given and the variables in the equations are explained in words.

4.1 Experimental Design

4.1.1 Experimental Setting

To answer the research questions stated in the introduction, a field experiment was performed on 39.444 naturally occurring customers at three different establishments from 't Smoske. The experiment continued for ten weeks and took place from March 12 to May 21. 't Smoske is a deli chain in the south of the Netherlands and the three different establishments are located in different cities in the province of Zeeland. One establishment is located in Vlissingen, the other in Middelburg, and the last one in Goes. These establishments, where the experiment is conducted, are on average twenty kilometers apart from each other. As said before, each establishment belongs to the same manager, has the same interior and exterior to a certain extent, and also has the exact same menu. This menu includes forty different sandwiches, waffles, donuts, soup, frozen yogurt, and sodas and fresh juices. The average price of the sandwiches is equal to €5.90, the waffles cost €2.25, the donuts €1.40, and the sodas and fresh juice are on average €2.59. At all three restaurants, the customers need to order at the counter where the menu is displayed on a large canvas posted above the cashier. Before they order they can either choose to order at the take-away counter and take their order with them, or to order at the restaurant counter and eat their food in the restaurant of 't Smoske. Besides this, a substantial share of the orders at 't Smoske comes from online orders.

The three different establishments are split into two groups, a control group, and a treatment group. The customers of the establishment in Vlissingen serve as a control group, while the customers of the establishment in Middelburg and Goes serve as the treatment group. This distribution is chosen based on the number of sales at each establishment; Vlissingen has the highest number of visitors, but less than the number of visitors in Middelburg and Goes combined. As a result, the treatment group is slightly larger than the control group, but they are comparable in size.

Before the experiment in these three restaurants could be started, the research idea needed to be explained to the manager of the deli chain to convince him to participate in the study. In a first meeting with the manager, the research idea was discussed and it was explained why it was so useful to conduct the research within the settings of his restaurants. In a second meeting, the manager was

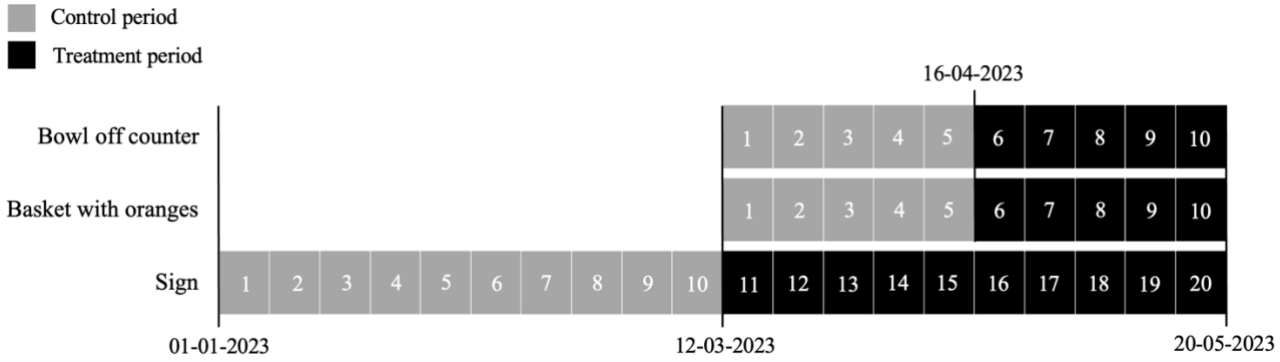
informed about the exact treatments, the date and duration of the treatment period, and the data required to analyze the effect of the treatments. Shortly after this meeting, the treatment period started.

A few measurements had been taken to make sure that the experiment ran properly and to make sure that the treatment was always on during the treatment period. First of all, all employees had been informed about the experiment and their full cooperation had been requested. They were ordered not to inform customers about the experiment, to make sure that they were unaware that they were taking part in an experiment. Besides this, a list had been posted at each establishment indicating the things the employees needed to pay attention to. During the experiment period, there were also signs placed throughout the restaurants telling the employees what they should and should not do. It was made sure that these signs were not visible to the customers, but only to the employees. For each different establishment, one employee had been contacted to ask if he/she wanted to be in charge of the experiment and wanted to make sure that it ran properly. Lastly, I regularly checked for myself if the treatment was still running.

4.1.2 Experimental Treatments

Figure 8 provides a graphical representation of the experimental design and shows the start and end dates of the different treatments.

Figure 8. Timeline of the experimental treatments



Note: the numbers on the timeline show the number of weeks in the control and treatment period. "Bowl off counter" indicates the treatment where a bowl with donuts and waffles is placed out of sight of the customers, "Basket with oranges" indicates the treatment where a basket with oranges is placed out of sight of the customers, and "Sign" indicates the treatment where a sign with most popular sandwiches is placed next to the cash register.

For the first nudge, a sign is placed on the counter next to the cash register where customers order their sandwiches. The sign displays the text: "Most popular sandwiches!" ("Populairste broodjes!") and shows three of the five most sold sandwiches according to historical sales from the period January 1 to March 12. All three sandwiches are also shown in a picture on the sign and a description of the sandwiches is given as well. The sandwiches displayed on the sign are not changed and the sign is not modified in any other way during the treatment period. For this nudge, the ten

weeks before the start of the experiment, January 1 to March 12, serve as the control period. The period from March 12 to May 21 serves as the treatment period. The treatment was implemented on the morning of March 12 before the opening of the restaurants. The sign is shown in figure 12 in the appendix and the setting before and after the start of the treatment are shown in figure 13 in the appendix.

For the second nudge, a basket that is filled with oranges is taken out of sight of the customer. For the control period, the basket of oranges is filled to the brim and placed prominently in the sight of the customers. After this period, the treatment period starts where the basket of oranges is taken away out of sight of the customer. The juicer is also completely empty during this period. For this nudge, the first 5 weeks of the experiment, March 12 to April 16, serve as the control period. The last five weeks, April 16 to May 21, serve as the treatment period. On the morning of March 12 before the opening of the restaurants, a sign was placed next to the juicer. As shown in figure 14 in the appendix, this sign stated that the basket with oranges as well as the juicer must be filled to the brim. On the morning of April 16 before the opening of the restaurants, the basket with oranges was removed from the restaurant. The sign was also replaced with another sign stating that the juicer must now be left completely empty (Appendix figure 15). Again, these signs were not visible to the customers. The setting before and after the start of the treatment is shown in figure 16 in the appendix.

For the third nudge, a bowl with donuts and waffles is taken out of sight of the customer. For the control period, the bowl is filled to the brim and placed prominently in the sight of the customers. After this period, the treatment period starts where the bowl is taken away out of sight of the customers. For this nudge, also the first five weeks of the experiment served as the control period, while the last five served as the treatment period. On the morning of March 12 before the opening of the restaurants, a sign was placed on the counter. As shown in figure 17 in the appendix, this sign indicated where the bowl with donuts and waffles should be placed and stated that it must be filled to the brim. On the morning of April 16 before the opening of the restaurants, the bowl was placed in the refrigerator with a sign on top of it stating that it should stay there (Appendix figure 18). Also these signs were not visible to the customers. The setting before and after the start of the treatment is shown in figure 19 in the appendix.

4.2 Empirical Analysis

4.2.1 Difference-in-differences Regressions

For the empirical analysis, a difference-in-differences design, with the use of three ordinary least squares regressions, is used to estimate the causal impact of highlighting specific options of a

menu and placing items out of sight of customers on the number of sales of certain products. It will be tested whether the post-treatment trend in the proportion and number of nudged products sold at the control restaurant is different from the post-treatment trend in the proportion and number of nudged products sold at the treatment restaurants.

For the estimation equation of the first hypothesis, let SS_{st} be the dependent variable indicating the proportion of nudged sandwiches sold at each establishment s on day t. Let $(Treat_s * Sign_t)$ be the independent variable and the main variable of interest. It is a treatment dummy which is 1 for the treatment group during the treatment period and 0 otherwise. In this regression, it is thus 1 for customers in Middelburg and Goes when the sign is placed on the counter and 0 otherwise. The empirical model used to test the first hypothesis can now be described as follows:

$$SS_{st} = \alpha + \delta(Treat_s * Sign_t) + \beta Treat_s + \gamma Sign_t + Weekday_{st} + Wait_{st} + \varepsilon_{st} \quad (1)$$

Where $Treat_s$ is the group-fixed effect, which is 1 for the treatment group and 0 otherwise. In this regression, it is thus 1 for customers in Middelburg and Goes and 0 for Vlissingen. $Sign_t$ is the time-fixed effect, which is 1 for the treatment period and 0 for the control period. In this regression it is thus equal to 0 before the sign was placed on the counter and equal to 1 after the sign was placed on the counter. To control for unobservable factors that may vary across weekdays, fixed effects for the day of the week are included. $Weekday_{st}$ is thus shorthand for a series of binary control variables for six of the seven weekdays. To look if there is treatment heterogeneity between customers with a short or a long waiting time a dummy variable $Wait_{st}$ is included in the regression, shorthand for a dummy variable interacting with all other terms. This dummy variable is equal to one if the number of products sold on a given day is above the median of the entire research period and zero otherwise. This variable then works as a proxy for the presence of a long waiting time.

For the estimation equation of the second hypothesis, let OS_{st} be the dependent variable indicating the proportion of orange juice sold to customers i at each establishment s at day t. Let $(Treat_s * Basket_t)$ be the independent variable and the main variable of interest. It is a treatment dummy which is 1 for the treatment group during the treatment period and 0 otherwise. In this regression, it is thus 1 for customers in Middelburg and Goes when the basket with oranges was out of sight of the customers and 0 otherwise. The empirical model used to test the second hypothesis can now be described as follows:

$$OS_{st} = \alpha + \delta(Treat_s * Basket_t) + \beta Treat_s + \gamma Basket_t + Weekday_{st} + Wait_{st} + \varepsilon_{st} \quad (2)$$

Where $Treat_s$ is the group-fixed effect, which is 1 for the treatment group and 0 otherwise. In this regression, it is thus 1 for customers in Middelburg and Goes and 0 for Vlissingen. $Basket_t$ is the time-fixed effect, which is 1 for the treatment period and 0 for the control period. In this regression, it is thus equal to 0 before the basket was out of sight of the customers and equal to 1 after the basket was out of sight of the customers. $Weekday_{st}$, again, indicates fixed effects for the day of the week. The interaction variable $Wait_{st}$ is included again to check for treatment heterogeneity between customers with a short and long waiting time.

For the estimation equation of the third hypothesis, let DWS_{st} be the dependent variable indicating the number of donuts and waffles sold to customers i at each establishment s at day t. As said in section 3.1, the number of sales instead of the proportion of sales is used, because 't Smoske does not sell any substitutes for donuts and waffles. Let $(Treat_s * Bowl_t)$ be the independent variable and the main variable of interest. It is a treatment dummy which is 1 for the treatment group during the treatment period and 0 otherwise. In this regression, it is thus 1 for customers in Middelburg and Goes when the bowl with donuts and waffles was out of sight of the customers and 0 otherwise. The empirical model used to test the third hypothesis can now be described as follows:

$$DWS_{st} = \alpha + \delta(Treat_s * Bowl_t) + \beta Treat_s + \gamma Bowl_t + Weekday_{st} + Wait_s + \varepsilon_{st} \quad (3)$$

Where $Treat_s$ is the group-fixed effect, which is 1 for the treatment group and 0 otherwise. In this regression, it is thus 1 for customers in Middelburg and Goes and 0 for Vlissingen. $Bowl_t$ is the time-fixed effect, which is 1 for the treatment period and 0 for the control period. In this regression, it is thus equal to 0 before the bowl was out of sight of the customers and equal to 1 after the bowl was out of sight of the customers. The fixed effect $Weekday_{st}$ and the interaction variable $Wait_{st}$ are also included again.

4.2.2 Time-varying Treatment Effects

After estimating the difference between the treatment and the control over the complete treatment period, multiple treatment-control comparisons will be made in order to explore dynamic treatment effects. To do this, the difference-in-differences regressions in equations (1), (2), and (3) are run again, but now include multiple interactions between the independent variable and treatment dummies for every week. This way, the difference between treatment and control is estimated for every week within the research period.

5. Results

In this chapter, the difference-in-differences results are shown and interpreted. Section 5.1 shows the results of a difference-in-differences calculation by hand as well as the results of a statistical test of the treatment effects. Section 5.2 gives a more comprehensive look at the treatment effect while time-varying treatment effects are explored.

5.1 Results

5.1.1 Sign With Most Popular Sandwiches

The result of the difference-in-differences calculation by hand for the salience nudge where a sign with the most popular sandwiches is placed on the counter is shown in table 6. The mean proportion of nudged sandwiches is given for the control and treatment period for the treatment restaurants as well as the control restaurant. The "Difference" column shows the change in the mean between the control and treatment period per treatment and control restaurants. It can be seen that the placement of the sign in the treatment restaurants increased the daily proportion of nudged sandwiches sold by 2.3 percentage points. No sign was placed at the control restaurant, which made it possible to perceive the natural change in the proportion. This restaurant experienced a next-to-nil decrease in the proportion of nudged sandwiches sold by 0.1 percentage points. Assuming that the common trend assumption holds, as proven in section 3.2, it can be expected that the treatment group would have shown the same nil change had it not been subjected to the treatment. The difference-in-differences effect is therefore equal to a 2.4 percentage points increase in the proportion of nudged sandwiches sold.

The regression results in column (I) in table 7 show a similar and moreover significant treatment effect, presenting a point estimate of 0.0242. This estimation result suggests a 2.42 percentage point higher proportion of nudged sandwiches sold at the treatment restaurants compared to the control restaurants due to the placement of the sign. If this point estimate is compared to the baseline average proportion in the control period, it can be concluded that the results suggest an 8.6% increase in the daily proportion of nudged sandwiches sold given that a sign with the most popular sandwiches is placed next to the cash register. A difference in the outcome ranging from a 4.3% to a 12.9% increase is also reasonably compatible with the data. While the effect is significant at the 1% level, the adjusted R-squared is low. Only 10.75% of the total variability in the proportion of nudged sandwiches sold is explained by the placement of the sign. There are thus presumably numerous other factors influencing the proportion due to the inherent complexity of human behavior.

Column (II) in table 7 shows the regression results when there is controlled for unobservable factors that may vary across weekdays. Including weekday fixed effects does not alter the main effect and does not substantially increase the performance of the model.

Column (III) shows the results of the fully interacted regression model, hence when the interaction variable "Wait" is included. This variable does alter the main effect and the effect becomes insignificant. The interaction effect itself is only significant at 10%. It can therefore not be concluded that the relationship between placing a sign next to the cash register displaying the most popular sandwiches and the proportion of nudged sandwiches sold varies significantly depending on the length of the waiting time.

Table 6. Effect of placing a sign next to the cash register displaying the most popular sandwiches on the proportion of sales of these nudged sandwiches: difference-in-differences calculation by hand

	Control period	Treatment period	Difference
Treatment restaurants	0.280	0.303	0.303 – 0.280 = 0.023
Control restaurant	0.283	0.282	0.282 – 0.283 = -0.001
		Difference-in-differences:	0.023 + 0.001 = 0.024
		Difference relative to control period:	0.024 / 0.280 = 8.6%

Table 7. Regression results for the effect of placing a sign next to the cash register displaying the most popular sandwiches on the proportion of sales of these nudged sandwiches

Independent variable	Nudged sandwich proportion		
	(I)	(II)	(III)
Sign	0.0242*** (0.0061)	0.0242*** (0.0061)	0.0135 (0.0093)
Treatment period	-0.0013 (0.0043)	-0.0013 (0.0043)	0.0015 (0.0052)
Treatment group	-0.0032 (0.0043)	-0.0032 (0.0043)	0.0017 (0.0069)
Wait × Sign	-	-	0.0217* (0.0131)
Weekday FE	No	Yes	No
Baseline average	0.280 (0.024)	0.280 (0.024)	0.280 (0.024)
Adjusted R^2	0.1075	0.1135	0.1111
Number of observations	280	280	280

*Note: standard errors in parenthesis. * denotes significance at 10 percent, ** denotes significance at 5 percent, and *** denotes significance at 1 or less than 1 percent. "Sign" is an interaction between "Treatment period" and "Treatment group" and indicates the treatment where a sign displaying the most popular sandwiches is placed next to the cash register*

5.1.2 Basket With Oranges Out Of Sight Of The Customer

The result of the difference-in-differences calculation by hand for the priming nudge where a basket that is filled with oranges is placed out of sight of the customers is shown in table 8. It can be seen that the placement of the basket with oranges out of sight of the customers in the treatment restaurants decreased the daily proportion of orange juice sold by 2.7 percentage points. The control restaurant experienced a next-to-nil decrease in the proportion of orange juice sold of 0.1 percentage points between the control and treatment period. Assuming that the treatment group would have shown the same change had it not been subjected to the treatment, the difference-in-differences effect is equal to a 2.6 percentage points decrease in the proportion of orange juice sold.

The regression results in column (I) in table 9 show a similar but insignificant treatment effect, presenting a point estimate of -0.0256. This estimation result suggests an insignificant decrease in the proportion of orange juice sold at the treatment restaurants compared to the control restaurant of 2.56 percentage points, due to the replacement of the basket with oranges out of sight of the customers. These results constitute a 17.2% decrease in the daily proportion of orange juice sold compared to the baseline average. Nonetheless, a difference in the outcome ranging from a 36.4% decrease to a 2.0% increase is also reasonably compatible with the data. The effect is only significant at the 10% level and the adjusted R-squared is low. Only 10.73% of the total variability in the proportion of orange juice sold is explained by the replacement of the basket with oranges.

Just like at the results for the effect of placing a sign next to the cash register displaying the most popular sandwiches, including weekday fixed effects in this regression does not alter the main effect and does not substantially increase the performance of the model. Column (III) shows the results of the fully interacted regression model, when the interaction variable "Wait" is included. This variable almost doubles the main effect and makes it significant at the 5% level. The interaction effect itself is however insignificant. It can therefore not be concluded that the relationship between placing a basket filled with oranges out of sight of the customers and the proportion of orange juice sold varies significantly depending on the length of the waiting time.

Table 8. Effect of placing a basket that is filled with oranges out of sight of the customers on the proportion of sales of orange juice: difference-in-differences calculation by hand

	Control period	Treatment period	Difference
Treatment restaurants	0.149	0.122	0.122 – 0.149 = -0.027
Control restaurant	0.162	0.161	0.161 – 0.162 = -0.001
		Difference-in-differences:	-0.027 + 0.001 = -0.026
		Difference relative to control period:	-0.026 / 0.149 = -17.4%

Table 9. Regression results for the effect of placing a basket that is filled with oranges out of sight of the customers on the proportion of sales of orange juice

Independent variable	Orange juice proportion		
	(I)	(II)	(III)
Basket with oranges	-0.0256* (0.0145)	-0.0256* (0.0144)	-0.0474** (0.0212)
Treatment period	-0.0010 (0.0102)	-0.0010 (0.0102)	0.0072 (0.0128)
Treatment group	-0.0133 (0.0102)	-0.0133 (0.0102)	-0.0083 (0.0133)
Wait × Basket with oranges	-	-	0.0148 (0.0308)
Weekday FE	No	Yes	No
Baseline average	0.149 (0.045)	0.149 (0.045)	0.149 (0.045)
Adjusted R^2	0.1073	0.1207	0.1628
Number of observations	140	140	140

*Note: standard errors in parenthesis. * denotes significance at 10 percent, ** denotes significance at 5 percent, and *** denotes significance at 1 or less than 1 percent. "Basket with oranges" is an interaction between "Treatment period" and "Treatment group" and indicates the treatment where a basket that is filled with oranges is placed out of sight of the customers.*

5.1.2 Bowl With Donuts and Waffles Out Of Sight Of The Customer

The result of the difference-in-differences calculation by hand for the priming nudge where a bowl with donuts and waffles is placed out of sight of the customers is shown in table 10. It can be seen that the placement of the bowl with donuts and waffles out of sight of the customers in the treatment restaurants decreased the daily number of donuts and waffles sold by 4.17. The control restaurant experienced a next-to-nil increase in the number of donuts and waffles sold of 0.06 between the control and treatment period. Assuming that the treatment group would have shown the same change had it not been subjected to the treatment, the difference-in-differences effect is equal to a decrease in the daily number of donuts and waffles sold of 4.23.

The regression results in column (I) in table 11 show a similar treatment effect of -4.229, significant at the 5% level. This estimation result suggests 4.23 less number of donuts and waffles sold per day for the treatment restaurants compared to the control restaurant. If this point estimate is compared to the baseline average number in the control period, it can be concluded that the results suggest a 78.3% decrease in the daily number of donuts and waffles sold given that a bowl with donuts and waffles is placed out of sight of the customers. A difference in the outcome ranging from a 46.2% to a 110.4% decrease is also reasonably compatible with the data. Looking at the adjusted

R-squared, 40.34% of the total variability in the number of donuts and waffles sold is explained by the replacement of the bowl with donuts and waffles.

Column (II) in table 11 shows again that including weekday fixed effects in the regression does not alter the main effect and does not substantially increase the performance of the model. Column (III) shows the results of the fully interacted regression model, when the interaction variable "Wait" is included. This variable does alter the main effect substantially while it keeps being significant. The interaction effect itself is however insignificant. It can therefore not be concluded that the relationship between placing a bowl with donuts and waffles out of sight of the customers and the number of donuts and waffles sold varies significantly depending on the length of the waiting time.

Table 10. Effect of placing a bowl with donuts and waffles out of sight of the customers on the number of sales of donuts and waffles: difference-in-differences calculation by hand

	Control period	Treatment period	Difference
Treatment restaurants	5.40	1.23	1.23 – 5.40 = -4.17
Control restaurant	6.43	6.49	6.49 – 6.43 = 0.06
		Difference-in-differences:	-4.17 – 0.06 = -4.23
		Difference relative to control period:	-4.23 / 5.40 = -78.33%

Table 11. Regression results for the effect of placing a bowl with donuts and waffles out of sight of the customers on the number of sales of donuts and waffles

Independent variable	Donut & waffle sales		
	(I)	(II)	(III)
Bowl off counter	-4.229*** (0.8754)	-4.229*** (0.8407)	-2.7088** (1.2734)
Treatment period	0.0571 (0.6190)	0.0571 (0.5945)	-1.1912 (0.7661)
Treatment group	-1.0286* (0.6190)	-1.0286* (0.5945)	-1.6500** (0.7972)
Wait × Bowl off counter	-	-	-2.5417 (1.8471)
Weekday FE	No	Yes	No
Baseline average	5.4 (3.05)	5.4 (3.05)	5.4 (3.05)
Adjusted R^2	0.4034	0.4498	0.4477
Number of observations	140	140	140

*Note: standard errors in parenthesis. * denotes significance at 10 percent, ** denotes significance at 5 percent, and *** denotes significance at 1 or less than 1 percent. "Bowl off counter" is an interaction between "Treatment period" and "Treatment group" and indicates the treatment where a bowl with donuts and waffles is placed out of sight of the customers.*

5.2 Time-varying Treatment Effects

While the rest of this paper looks at the difference between the treatment and control group over the complete time period, this section estimates the difference between the treatment and control group for every week. The average treatment effects provided in section 5.1 only give a good idea of the average effect for the entire treatment period, while there is a chance that the treatment effect grows or falls over time. An increase or decrease in the treatment effect over time may occur due to several reasons. Kurz (2018) mentions that it can for example be the case that the effect increases over time due to network effects. If people consume a nudged product as a result of the nudge they can recommend others to also consume it which can lead to a constantly increasing consumption and thus increasing treatment effect over time. On the other hand, Kurz (2018) mentions that it is possible that nudges only have a high immediate effect and no long-term effect if people get nudged into the consumption of certain products but return to their original choice quickly after.

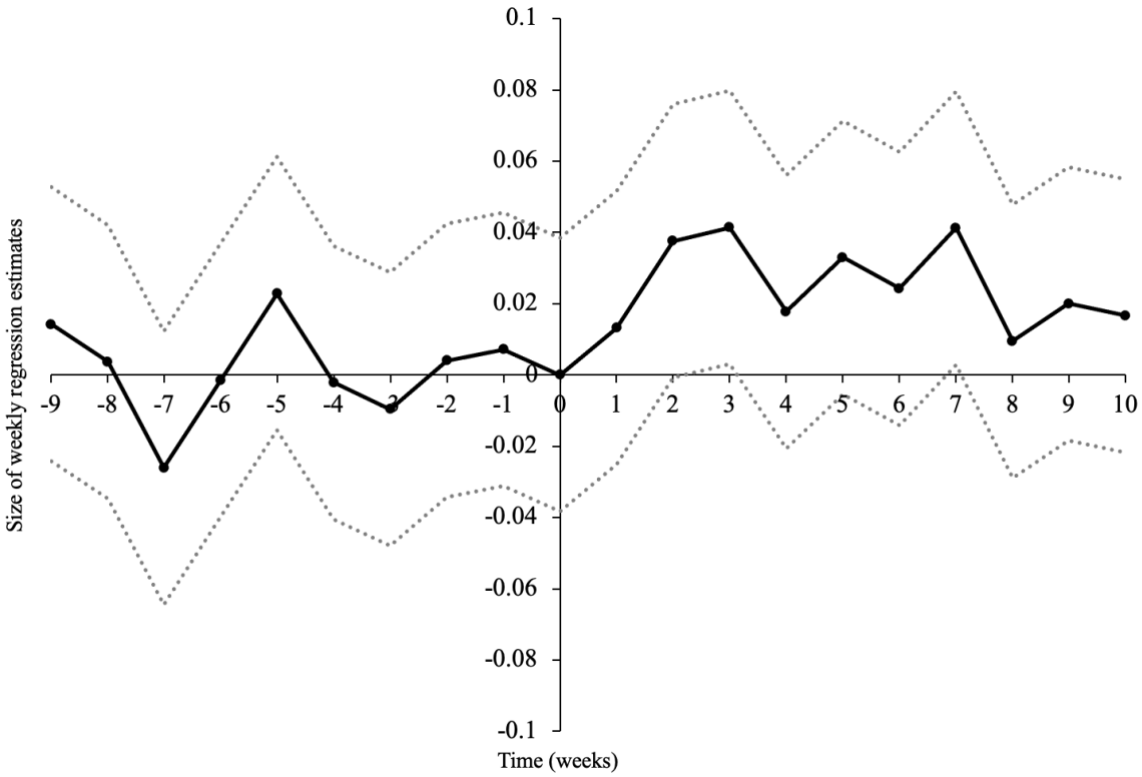
To look if any of this is the case in this study, weekly point estimates are calculated by including multiple interactions between the independent variable and treatment dummies for every week. These estimates are then plotted in figures 9, 10, and 11. Figures 9, 10, and 11 thus show the average treatment effects of the salience nudge where a sign with the most popular sandwiches is placed on the counter and the priming nudges where a basket with oranges and a bowl with donuts and waffles are placed out of sight of the customers for each week in the research period. The black dots in figures 9, 10, and 11 show the exact regression estimates while the grey dotted lines indicate the 95%-confidence intervals of these estimates.

The point estimates in the pre-treatment period should be zero or close to zero and there should be no trend, while the point estimates in the treatment period should not be zero and there can be an upward or downward trend in the point estimates in the case of time-varying treatment effects. It can be seen that not all point estimates in the three graphs are exactly equal to zero but that the black line through the point estimates moves around the zero line without any sign of a trend. This thus shows additional proof of the common trend assumption. There is, however, also no clear upward or downward trend in the points estimates in the treatment period. There is lots of noise in these weekly coefficients which causes some variety over time, but there is no sign of an upward or downward trend. There is thus no evidence of any time-varying treatment effects, just like in the paper from Van Gestel et al. (2018).

The confidence intervals during the treatment period should also not include zero to be able to conclude that the weekly treatment effects are statistically significantly different from zero. It can

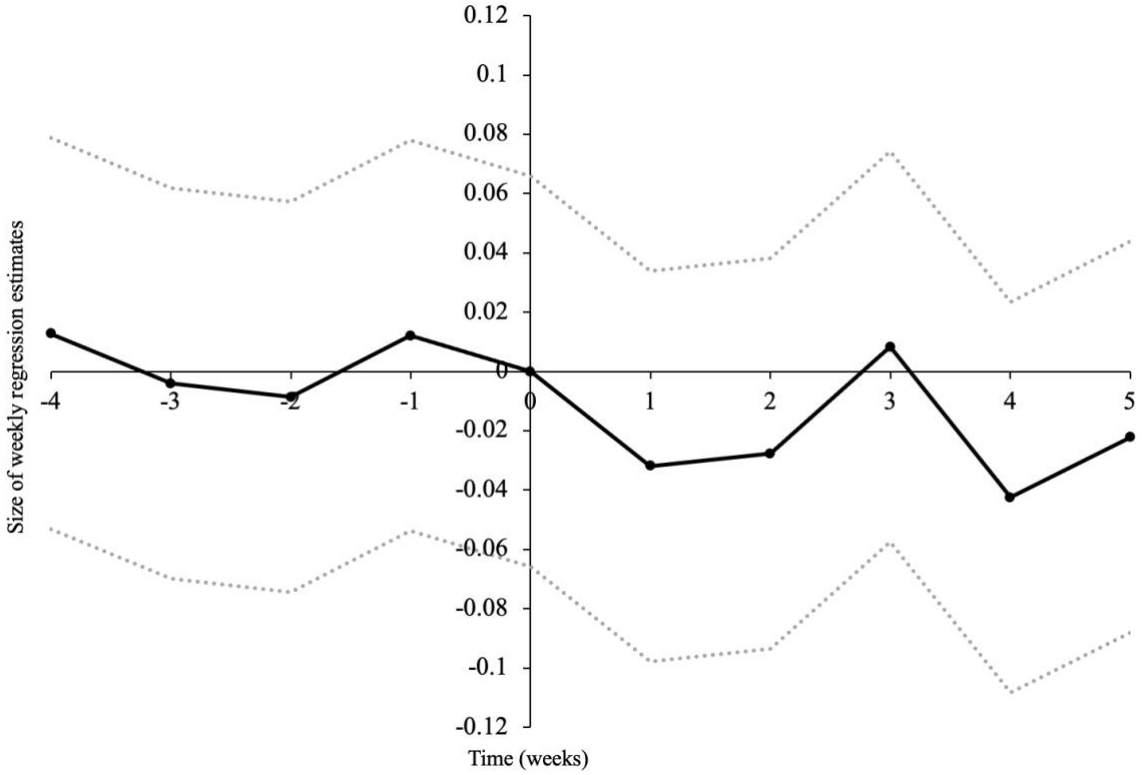
however be seen that the confidence interval of the nudge where a sign with the most popular sandwiches is placed on the counter includes zero in all weeks but weeks three and seven. The treatment effect in most weeks is thus not statistically significantly different from zero. However, the confidence bound is close to zero which still leads to a treatment effect with a 5% significance level for the point estimate of the entire treatment period (section 5.1). The confidence interval of the nudge where a basket with oranges is placed out of sight of the customers even includes zero for all weeks and is far from close to zero. As also shown in section 5.1, this leads to an average treatment effect that is only significant at the 10% level. The confidence interval of the nudge where a bowl with donuts and waffles is placed out of sight of the customers does not include zero for all weeks but week four. Most weekly treatment effects are thus significantly different from zero, which is due to the high magnitude of the effect.

Figure 9. Weekly regression estimates for the effect of placing a sign next to the cash register displaying the most popular sandwiches on the proportion of sales of these nudged sandwiches



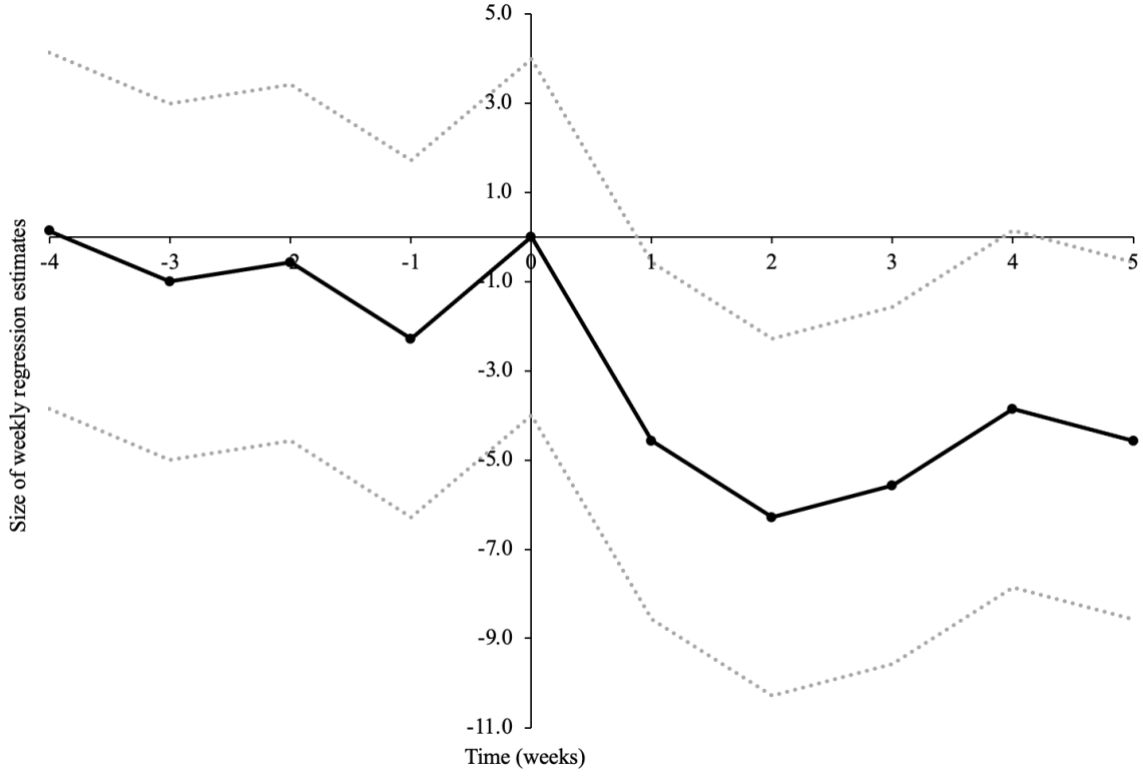
Note: treatment starts after week 0. The dotted lines indicate the 95% confidence interval.

Figure 10. Weekly regression estimates for the effect of placing a basket that is filled with oranges out of sight of the customers on the proportion of sales of orange juice



Note: treatment starts after week 0. The dotted lines indicate the 95% confidence interval.

Figure 11. Weekly regression estimates for the effect of placing a bowl with donuts and waffles out of sight of the customers on the number of sales of donuts and waffles



Note: treatment starts after week 0. The dotted lines indicate the 95% confidence interval.

6. Conclusion

Chapter 6 discusses the results of chapter 5 and concludes. Section 6.1 summarizes the paper and analyzes the results with respect to the hypotheses and literature review in chapter 2. In section 6.2 the policy implications of this study are given while the limitations of this study will be discussed and recommendations will be made for future research in section 6.3.

6.1 Discussion and Conclusion

Despite some compelling evidence of the effect of nudges within food retail, such as the positive effect of featuring specific items on the orders of low-calorie sandwiches (Wisdom et al., 2010), the impact of the product type, number of alternative options and self-reported hunger on the effectiveness of nudges (Bredie et al., 2019; Forwood et al., 2015) and the effect of increased visibility of fresh produce on the number of sales, further research into nudging is still of great importance due to the context-specific nature of nudging. This study thus aimed to investigate the effects of salience and priming nudges in the context of food retail, conducting a field experiment at a deli chain. More specifically, the paper focuses on the effects of highlighting specific menu options and placing items out of sight of customers on the purchasing behavior of customers. The field experiment conducted in this study covers a sample of naturally occurring customers visiting one of three establishments of a deli chain in the Netherlands during the first twenty weeks of 2023. This field experiment is motivated by Thaler and Sunstein (2008) and their book "Nudge: Improving Decisions About Health, Wealth, and Happiness" in which they introduce the concept of nudging. Since then, several studies analyzed the concept in the context of food retail, but only a few of them explored variety in the effect over time (Kurz, 2018; Van Gestel et al., 2018). To my knowledge, none of them looked into the moderating effect of the length of the waiting time on the effect of nudges on the purchasing behavior of customers. Hence, this study contributed to the literature by conducting a field experiment with a relatively long treatment period in order to look if there is any variation in treatment effects over time and contributes to the literature by looking if there is any treatment heterogeneity between customers with a short or a long waiting time.

This study thus conducted a field experiment where three priming and salience nudges are implemented within a deli chain. The nudges included placing a sign next to the cash register displaying the most popular sandwiches and placing a basket with oranges as well as a bowl with donuts and waffles out of sight of the customer. A difference-in-differences approach, with the use of three ordinary least squares regressions, is used to analyze the effect of these nudges on the sales. The main outcome measures in the regressions are the proportion of nudged sandwiches sold, the proportion of orange juice sold, and the number of donuts and waffles sold. The independent variables

are dummy variables which are 1 for the treatment group during the treatment period and 0 otherwise. Moreover, in each regression, the total daily number of products sold is used as a proxy for the length of the waiting time. Weekday fixed effects are included to control for unobservable factors that may vary across weekdays.

In accordance with previous literature and in accordance with the aforementioned hypotheses, this study finds a significant positive relationship between the placement of a sign displaying the most popular sandwiches and the proportion of sales of these sandwiches as well as a significant negative relationship between the placement of a bowl with donuts and waffles out of sight of the customers and the number of sales of donuts and waffles. However, no evidence is found for a significant effect of the placement of a basket with oranges out of sight of the customers on the proportion of orange juice sold. Controlling for week of the day fixed effects did not alter these main effects. No evidence is found for an upward or downward trend in the treatment effects over time for any of the three nudges nor did a proxy for waiting time indicate any evidence for treatment heterogeneity between customers with a short or a long waiting time.

Placing a sign next to the cash register displaying the most popular sandwiches increased the daily proportion of sales of these sandwiches by 2.42 percentage points, which constitutes an 8.6% increase in the daily proportion of nudged sandwiches sold compared to the baseline average proportion of 28%. A difference in the outcome ranging from a 4.3% to a 12.9% increase is also reasonably compatible with the data. These results are in accordance with the first hypothesis and in accordance with the effect sizes in the previous literature ranging from a 7.6% to a 35% increase (Bredie et al., 2019; Wisdom et al., 2010) as shown in table 1. Based on these previous effect sizes we can assess the type S-error, the probability of the effect being in the wrong direction, and the type M-error, the factor by which the magnitude of an effect might be overestimated (Gelman & Carlin, 2014). Based on the computations of Gelman and Carlin (2014), the previous results, and a standard error of 2.17% (0.0061/0.280) from the results of this study, we can expect the type S-error to range from 2.71×10^{-27} to 6.14×10^{-73}. The chance that the statistically significant finding in this study has the opposite sign of the hypothesized effect is thus really small. The type M-error ranges from 1.04 to 0.99, meaning that the statistically significant finding does not exaggerate the hypothesized true effect size in magnitude. Based on the single-minded results described in the literature review, my prior belief that the treatment worked was quite high and equal to 80%. With this prior, a Bayes factor is obtained of 1.23, which suggests that the estimates are 1.23 times more likely to have occurred under the alternative hypothesis that there is an effect of placing a sign displaying the most popular sandwiches on the proportion of these sandwiches sold than under the null hypothesis of no effect. If

we use Jeffrey's scale to assess the strength of evidence provided by the data, we can conclude that the Bayes factor indicates that there is only anecdotal evidence for the alternative hypothesis. Due to the high prior, the study thus contributes little to the already existing literature.

placing a basket that is filled with oranges out of sight of the customers led to an insignificant decrease in the daily proportion of sales of orange juice of 2.56 percentage points, which constitutes a 17.2% decrease in the daily proportion of orange juice sold compared to the baseline average proportion of 14.9%. Nonetheless, a difference in the outcome ranging from a 36.4% decrease to a 2.0% increase is also reasonably compatible with the data. The effect is, however, only significant at the 10% level. These results are in accordance with the results of Forwood et al. (2015) and Cheung et al. (2019) who also found the effect to be insignificant. The results, however, contradict with the results from Thorndike et al. (2017) who did find a significant effect. Based on these results, my prior belief that the treatment worked was only equal to 20%. With this prior, a Bayes factor is obtained of 3.49, which suggests that the estimates are 3.49 times more likely to have occurred under the alternative hypothesis that there is an effect of placing a basket with oranges out of sight of the customers on the proportion of orange juice sold than under the null hypothesis of no effect. If we use Jeffrey's scale to assess the strength of evidence provided by the data, we can again conclude that the Bayes factor indicates that there is moderate evidence for the alternative hypothesis.

Placing a bowl with donuts and waffles out of sight of the customers decreased the daily number of sales of donuts and waffles by 4.23, which constitutes a 78.3% decrease in the daily number of donuts and waffles sold compared to the baseline average number of 5.4. A difference in the outcome ranging from a 46.2% to a 110.4% decrease is also reasonably compatible with the data. These results are in accordance with the third hypothesis and in accordance with the effect sizes in the previous literature ranging from a 26.1% decrease to an 82.2% increase as shown in table 3. Based on these previous effect sizes we can assess the type S-error and the type M-error again (Gelman & Carlin, 2014). Based on the computations of Gelman and Carlin (2014), the previous results, and a standard error of 16.21% (0.8754/5.4) from the results of this study, we can expect the type S-error to range from 4.91×10^{-4} to 1.03×10^{-12}. The chance that the statistically significant finding in this study has the opposite sign of the hypothesized effect is thus really small. The type M-error ranges from 1.65 to 1.00, implying that the statistically significant results in this study can be 1.65 times larger than can be expected given the hypothesized true effect, but also that it can be that the statistically significant finding does not exaggerate the hypothesized true effect size in magnitude. Based on the results described in the literature review, my prior belief that the treatment worked was also quite high and equal to around 70%. With this prior, a Bayes factor is obtained of 1.40, which

suggests that the estimates are 1.40 times more likely to have occurred under the alternative hypothesis that there is an effect of placing a bowl with donuts and waffles out of sight of the customers on the number of donuts and waffles sold than under the null hypothesis of no effect. If we use Jeffrey's scale to assess the strength of evidence provided by the data, we can conclude that the Bayes factor indicates that there is only anecdotal evidence for the alternative hypothesis. Due to the high prior, the study thus contributes little to the already existing literature.

We can now answer the two research questions of this paper: "What is the effect of the salience nudge of highlighting specific options of a menu on the sales of these options?" and "What is the effect of the priming nudge of placing specific items out of sight of the customers on the sales of these items?". Within the specific context of the deli chain in this study, this study finds a significant positive effect of the salience nudge of highlighting specific options of a menu on the proportion of sales of these options. The effectiveness of the priming nudge of placing specific items out of sight of the customers on the sales of these items varies depending on the item. There is a significant negative effect for donuts and waffles, but an insignificant effect for orange juice.

6.2 Policy Implications

The manager of 't Smoske could thus adopt these cost-effective nudges at all establishments and use them on a wider range of products. With the right implementation and the proper selection of products, the cost-effective nudges could lead to more sustainable and socially responsible choices while simultaneously increasing the customers' well-being and the firm's profit (Schmidt & Engelen, 2020). However, while there could be treatment heterogeneity between products and effects can differ between contexts (Olstad et al., 2014), the manager should always closely monitor the effects. The same holds for other food providers who can adopt the effective nudges, but always have to take into account the range of contextual factors that can impact the effectiveness of the nudge (Olstad et al., 2014). Besides this, critics say that nudges can infringe on important social values such as the individual autonomy and dignity of customers and that nudges can even be perceived as manipulative (Schmidt & Engelen, 2020). Furthermore, it can be argued that it is wrong to use nudges without being transparent about those nudges (Kuyer & Gordijn, 2023). Many studies have however shown that nudges can be transparent while maintaining or even gaining effectiveness (Cheung et al., 2019; Kroese et al., 2016). All things considered, the manager of 't Smoske as well as other food providers should apply caution when implementing the priming and salience nudges from this study.

6.3 Limitations and Recommendations for Future Research

This study is subject to several limitations. Firstly, there is a small chance of treatment dilution and treatment migration. There could be treatment dilution while there is a small chance that customers do not get exposed to the nudges and thus do not get treated. It is possible that customers do not see the sign displaying the most popular sandwiches, the basket with oranges, or the bowl with donuts and waffles. However, this study simply compares all the customers at the control restaurant with the customers at the treatment restaurants, including those who are not actually treated. We, therefore, actually estimate the intention-to-treat effect instead of the average treatment effect. The chance of treatment migration is minimized when treatments are assigned at restaurant level instead of the individual level within one restaurant. However, there could still be treatment migration if customers visit the control restaurant as well as the treatment restaurants. While the establishments are located in different cities and are on average twenty kilometers apart from each other, it is not unlikely that people visit different establishments in the treatment period of ten weeks.

Secondly, 't Smoske has a lot of loyal returning customers who always order the same. Kurz (2018) and Cheung et al. (2019) state that customers with predetermined choices and a strong habitual nature are probably not affected by the nudges while they stuck to their predetermined choice throughout the entire treatment period. If this choice includes any of the nudged products the customers can be described as "always-takers", which means that they buy the nudged products regardless of the presence of the nudge. However, if this choice does not include any of the nudged products the customers can be described as "never-takers", which means that they never buy the nudged products and thus do not get influenced by the nudge. Potential "Always-takers" and "never-takers" thus influence the treatment effects.

Thirdly, a limitation of this study is that there is a small probability that the execution of the nudges did not go as planned in one of the treatment restaurants or in both treatment restaurants. It could be the case that the sign with the most popular sandwiches was not always placed next to the cash register or that the basket with oranges or the bowl with donuts and waffles was placed in sight of the customers while they needed to be placed out of sight during the treatment period. However, while Foster et al. (2014) describe it as a risk of field experiments, the risk in this study was quite low due to the measurements taken described in section 4.1. These measurements made sure that the experiment ran properly and that the treatment was always on during the treatment period.

While this study found significant effects for two of the three nudges implemented in the context of this study, section 6.2 already mentioned the context-specific nature of the nudges. There is a chance that the effects diminish when the nudges are implemented in another context or on a higher scale. Nudges that are effective in one situation or for one population may well have null or even opposite effects in another situation or for another population with different properties and

characteristics (Olstad et al., 2014). The effectiveness of the nudges implemented in this study can for example be different in a situation with other products, a situation with more alternative options, or a population with a higher average level of hunger (Bredie et al., 2019; Forwood et al., 2015; Foster et al., 2014). Implementation on a higher scale can also influence the treatment effect while it gets increasingly harder to ensure a good implementation and to ensure that the experiment runs properly. Even in the implementation goes well and the treatment always keeps on, there is a chance that scale simply acts as a moderator decreasing the treatment effect.

A first recommendation for feature research is to also look at treatment heterogeneity between different kinds of products as has been done in the studies of Bredie et al. (2019) and Foster et al. (2014). This study only looked at the overall effect on nudged sandwiches, orange juice, and donuts and waffles combined. However, the effect may differ between a cheese and ham sandwich, a chicken curry, or a shoarma salad or between white or whole grain bread. The effect can also differ between orange juice in a 25-centiliter bottle, a 50-centiliter bottle, and a glass or between donuts and waffles separately. The reported treatment effects in this study are now equal to the weighted averages of the products and product types mentioned above. Insight into the effect per product could increase the external validity of this study which makes it interesting for future studies to look into this.

A second recommendation for future research is to look if the effect of the nudges is persistent after the end of the treatment period. While the research period in this study is relatively long, it provides a comprehensive picture of the three different treatment effects. However, we have no indication of the development in the sales of the nudged sandwiches, orange juice, and donuts and waffles after the end of the research period on May 20. Kurz (2018) emphasizes the importance of this knowledge while a concern of nudging is that it might only have short-term and rapidly diminishing effects.

Acknowledgments

I am grateful to Ton, the manager of ‘t Smoske, for giving me the opportunity to conduct my research within his restaurants and for the data he provided. I am especially grateful to Sil for her help and support without whom this research would have been impossible. She has made a major contribution to the implementation of the various nudges within 't Smoske and has provided all the necessary data for this research. I would also like to thank Kaj and Sandra for their help at the establishments in Middelburg and Goes and for making sure that the experiment ran properly. Lastly, I would like to thank my supervisor, Ben Vollaard, for his guidance and advice throughout the entire writing proses of this paper.

References

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, *95*(9–10), 1082–1095. https://doi.org/10.1016/j.jpubeco.2011.03.003

Benartzi, S., & Thaler, R. H. (2007). Heuristics and Biases in Retirement Savings Behavior. *Journal of Economic Perspectives*, *21*(3), 81–104. https://doi.org/10.1257/jep.21.3.81

Bredie, W. L., Massey, C., Perez-Cueto, F. J., Appleton, K. M., Dinnella, C., Monteleone, E., Depezay, L., Hartwell, H., & Giboreau, A. (2019). When are "Dish of the Day" nudges most effective to increase vegetable selection? *Food Policy*, *85*, 15–27. https://doi.org/10.1016/j.foodpol.2019.04.003

Chapman, K., & Ogden, J. (2012). Nudging Customers towards Healthier Choices: An Intervention in the University Canteen. *Journal of Food Research*, *1*(2). https://doi.org/10.5539/jfr.v1n2p13

Cheung, T., Gillebaart, M., Kroese, F. M., Marchiori, D., Fennis, B. M., & De Ridder, D. T. D. (2019). Cueing healthier alternatives for take-away: a field experiment on the effects of (disclosing) three nudges on food choices. *BMC Public Health*, *19*(1). https://doi.org/10.1186/s12889-019-7323-y

Damgaard, M. T., & Nielsen, H. S. (2018). Nudging in education. *Economics of Education Review*, *64*, 313–342. https://doi.org/10.1016/j.econedurev.2018.03.008

Forwood, S. E., Ahern, A. L., Hollands, G. J., Ng, Y., & Marteau, T. M. (2015). Priming healthy eating. You can't prime all the people all of the time. *Appetite*, *89*, 93–102. https://doi.org/10.1016/j.appet.2015.01.018

Foster, G. D., Karpyn, A., Wojtanowski, A. C., Davis, E. E., Weiss, S. E., Brensinger, C. M., Tierney, A., Guo, W., Brown, J. M., Spross, C., Leuchten, D., Burns, P., & Glanz, K. (2014). Placement and promotion strategies to increase sales of healthier products in supermarkets in low-income, ethnically diverse neighborhoods: a randomized controlled trial. *The American Journal of Clinical Nutrition*, *99*(6), 1359–1368. https://doi.org/10.3945/ajcn.113.075572

Gelman, A., & Carlin, J. B. (2014). Beyond Power Calculations. *Perspectives on Psychological Science*, *9*(6), 641–651. https://doi.org/10.1177/1745691614551642

German wholesaler METRO acquires Eijsink|Booq. (2023, January 17). Booq. https://www.booqsolutions.com/en/blog/german-wholesaler-metro-acquires-eijsink/

Imai, K., Goldstein, D. A., Göritz, A. S., & Gollwitzer, P. M. (2007). Nudging Turnout: Mere Measurement and Implementation Planning of Intentions to Vote. *Social Science Research Network*. https://doi.org/10.2139/ssrn.977000

Johnson, E., & Goldstein, D. A. (2003). Do Defaults Save Lives? *Science*, *302*(5649), 1338–1339. https://doi.org/10.1126/science.1091721

Kroese, F. M., Marchiori, D., & De Ridder, D. T. D. (2016). Nudging healthy food choices: a field experiment at the train station. *Journal of Public Health*, *38*(2), e133–e137. https://doi.org/10.1093/pubmed/fdv096

Kurz, V. (2018). Nudging to reduce meat consumption: Immediate and persistent effects of an intervention at a university restaurant. *Journal of Environmental Economics and Management*, *90*, 317–341. https://doi.org/10.1016/j.jeem.2018.06.005

Kuyer, P., & Gordijn, B. (2023). Nudge in perspective: A systematic literature review on the ethical issues with nudging. *Rationality and Society*, 104346312311550. https://doi.org/10.1177/10434631231155005

Olstad, D. L., Goonewardene, L. A., McCargar, L. J., & Raine, K. D. (2014). Choosing healthier foods in recreational sports settings: a mixed methods investigation of the impact of nudging and an economic incentive. *International Journal of Behavioral Nutrition and Physical Activity*, *11*(1), 6. https://doi.org/10.1186/1479-5868-11-6

Schmidt, A., & Engelen, B. (2020). The ethics of nudging: An overview. *Philosophy Compass*, *15*(4). https://doi.org/10.1111/phc3.12658

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Yale University Press.

Thorndike, A. N., Bright, O. M., Dimond, M. A., Fishman, R., & Levy, D. E. (2017). Choice architecture to promote fruit and vegetable purchases by families participating in the Special Supplemental Program for Women, Infants, and Children (WIC): randomized corner store pilot study. *Public Health Nutrition*, *20*(7), 1297–1305. https://doi.org/10.1017/s1368980016003074

Van Gestel, L. C., Kroese, F. M., & De Ridder, D. (2018). Nudging at the checkout counter – A longitudinal study of the effect of a food repositioning nudge on healthy food choice. *Psychology & Health*, *33*(6), 800–809. https://doi.org/10.1080/08870446.2017.1416116

Voyer, B. (2015). 'Nudging' behaviours in healthcare: Insights from behavioural economics. *British Journal of Healthcare Management*, *21*(3), 130–135. https://doi.org/10.12968/bjhc.2015.21.3.130

Wilson, A. A., Buckley, E., Buckley, J. D., & Bogomolova, S. (2016). Nudging healthier food and beverage choices through salience and priming. Evidence from a systematic review. *Food Quality and Preference*, *51*, 47–64. https://doi.org/10.1016/j.foodqual.2016.02.009

Wisdom, J., Downs, J. S., & Loewenstein, G. (2010). Promoting Healthy Choices: Information versus Convenience. *American Economic Journal: Applied Economics*, *2*(2), 164–178. https://doi.org/10.1257/app.2.2.164

Appendix

The data used in this study is data on the daily number of sales per establishment of all the sandwiches, waffles, donuts, soup and sodas and fresh juices sold between January 1 and May 21, 2023. The data of the in-store sales consisted of a separate Excel file per day and establishment belonging to the study, while the data of the online sales consisted of one big file containing all dates within the research period and all products sold on these dates.

All the code in the data cleaning process used to clean the in-store sales data was inserted into for loops to iterate over each dataset in the R environment and apply the code to each separate dataset. All the files of the in-store data consisted of five columns of which only the column with the product names of the products sold on that day and the column with the number of sales on that day were kept. After this, some products were renamed and unnecessary products and rows were removed. Duplicate names were then grouped together while the number of sales of these duplicates were summed up. Data were missing for Goes on January 1, January 30, February 6 through February 10, and April 9 and for Middelburg on January 1. The datasets of the in-store sales for Goes from January 8, January 23, February 13 to February 17 and April 2 and for Middelburg from January 8 have therefore been duplicated to fill up the missing data. Besides this, the dataset for March 8 in Goes, Middelburg en Vlissingen also contained the sales of March 7 and the dataset for February 3 in Goes also contained the sales of February 2. These datasets were therefore split up into two datasets while the number of sales was divided equally between them.

The file of the online sales consisted of 7483 rows each representing an order and 57 columns which respectively included a column with the row number, the order date, the order number, the location of the establishment, the status of the payment, 26 columns with the 1st unique product sold in one order up to the 26th unique product sold in that same order and 26 columns with the number of sales within one order of the unique products sold in that order. If there were sold less than 26 unique products in one order, the remaining columns contained “NA” values which were all converted to zeros. Also in this dataset, some products were renamed to match the product names of the in-store data. Furthermore, rows with payment status “Not paid” and with the location “Boulevard” were removed. Then, three subsets were created, where each one only contained the rows with the location “Vlissingen”, the rows with the location “Middelburg” or the rows with the location “Goes”. These datasets were split up into even more subsets, one per establishment and per date. A for loop was then used again to transform all these datasets from a wide format to a long format by gathering the columns into two columns. The 26 columns with the unique products were put below each other in one column and the 26 columns with the sales numbers were put below each other in one column.

Again, unnecessary products and rows were removed and duplicate product names were grouped together while the number of sales of these duplicates were summed up.

The R environment now contained various datasets with two columns, one column with the product names and one column with the number of products sold. These datasets consisted for each date within the research period, each establishment that participated in the study, and for in-store sales as well as online sales. The datasets for the in-store sales and the online sales were then merged together after which the number of online sales was subtracted from the number of in-store sales. After this, the proportion of nudged sandwiches sold and the proportion of orange juice sold were calculated for each day and establishment by dividing the number of nudged products sold by the total number of nudged products sold in each dataset. The number of donuts and waffles sold was calculated for each day and establishment by summing up the number of donuts and waffles sold in each dataset. All these proportions and quantities were inserted into a column in a new dataset. Then, a column "Location" was created indicating the establishment, a column "Day" indicating the day of the research period, a column "Treatment group" which is one if the column "Location" is equal to "Middelburg & Goes" and zero if the column is equal to "Vlissingen", a column "Treatment period" which is zero for the control period and one for the treatment period and a column "Treatment indicator" which multiplies the column "treatment group" and "treatment period". This column is equal to one for the treatment group during the treatment period and zero otherwise. Also, a column called "Wait" was created which is equal to one if the number of products sold on a given day is above the median of the entire research period and zero otherwise, a column "Weekday" was created which showed the day of the week and multiple "Week" columns were created which are one if the row belongs to that specific week and zero for all other weeks.

Figure 12. The sign that is placed next to the cash register displaying the most popular sandwiches

Figure 13. The setting in the establishment in Middelburg before (left) and after (right) the start of the treatment where a sign is placed next to the cash register displaying the most popular

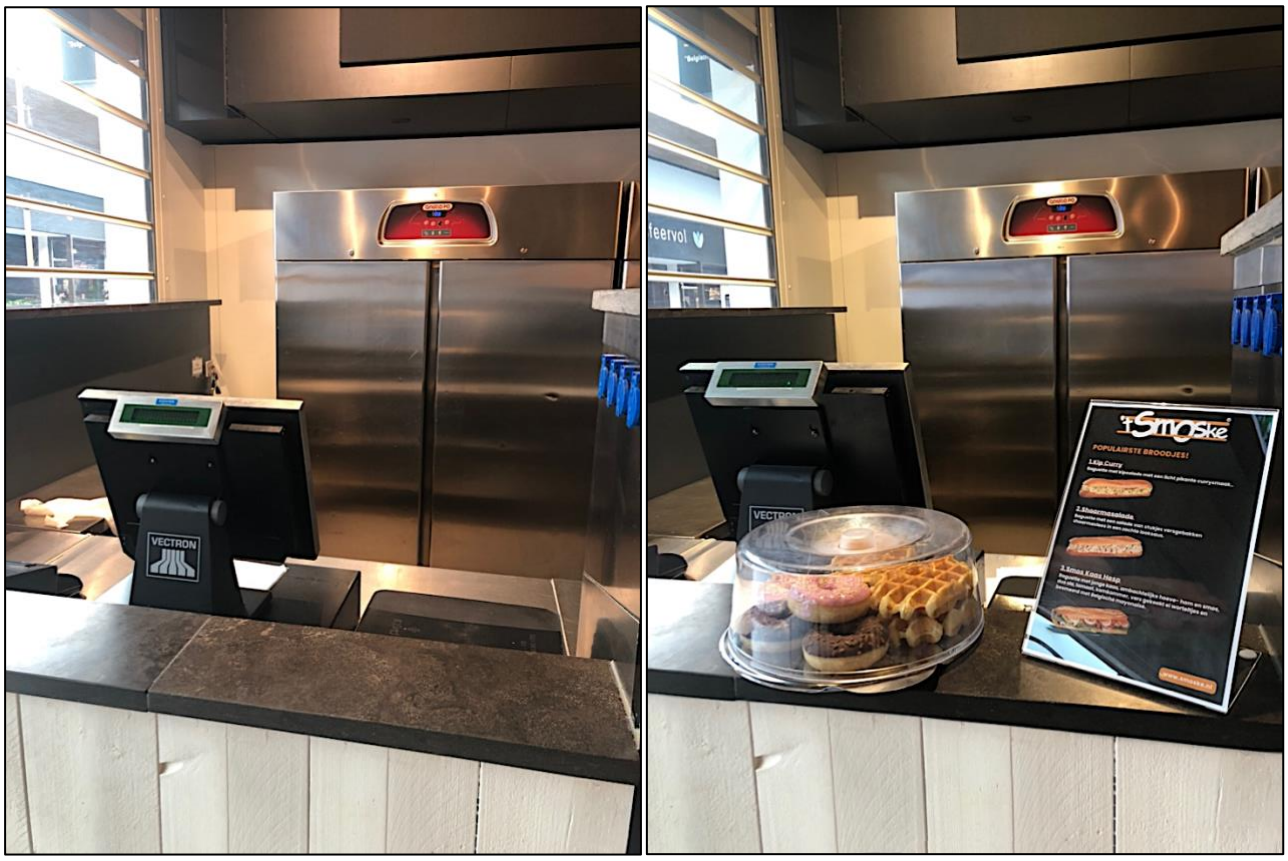


Figure 14. The sign stating that the basket with oranges must be filled to the brim: “Keep this basket fully stocked!”

Figure 15. The sign stating that the juicer must be left completely empty: “Do not fill the juicer with oranges, leave it empty!”

Figure 16. The setting in the establishment in Goes before (left) and after (right) the start of the treatment where a basket that is filled with oranges is placed out of sight of the customers

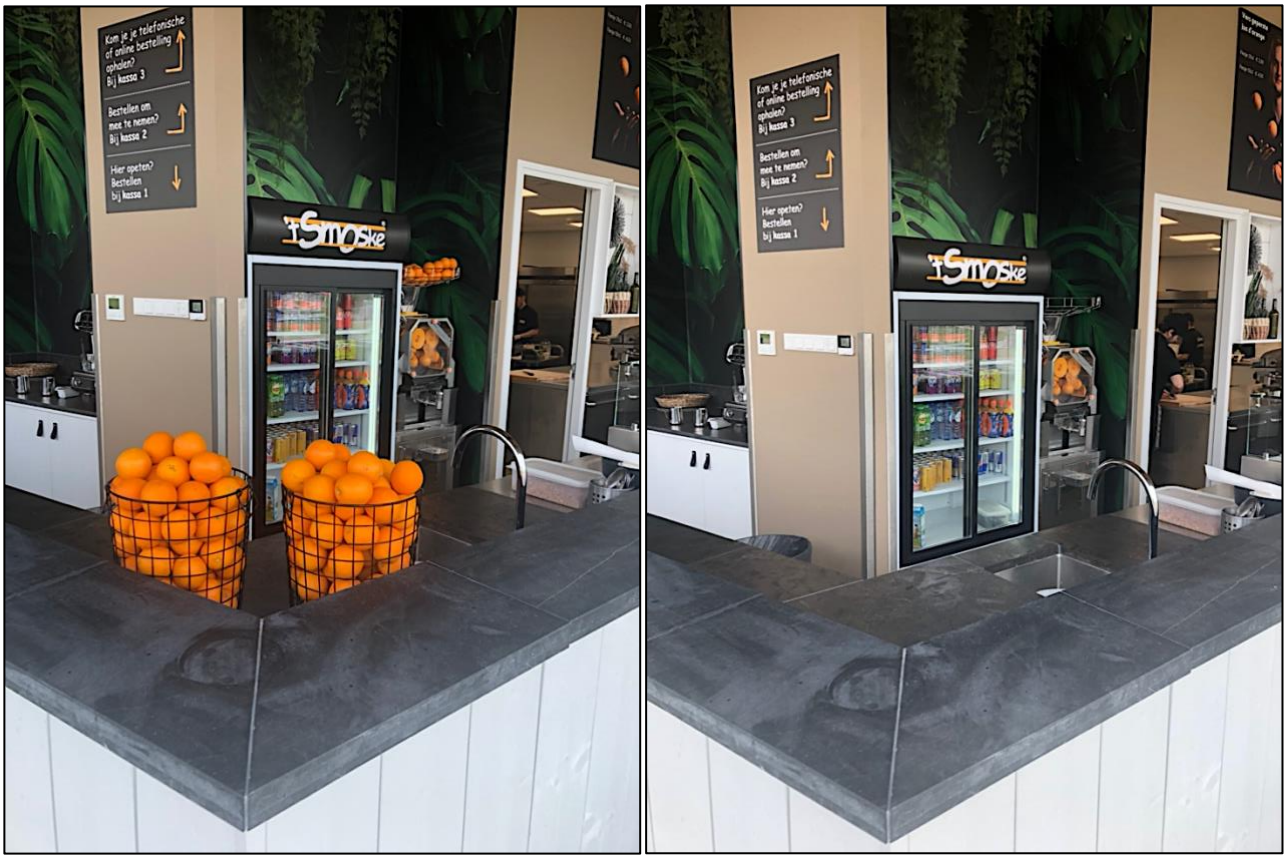


Figure 17. The sign indicating where the bowl with donuts and waffles should be placed and stating that it must be filled to the brim: “Place the bowl with donuts and waffles here, on top of the red dot! Make sure the bowl is completely filled!

Figure 18. The bowl with donuts and waffles with the sign on top of it stating that it should stay in the refrigerator: “Leave the bowl with waffles and donuts here in the refrigerator!”

Figure 19. The setting in the establishment in Middelburg before (left) and after (right) the start of the treatment where a bowl with donuts and waffles is placed out of sight of the customers