Before the Lunch Line: Effectiveness of Behavioral Economic Interventions for Pre-Commitment on Elementary School Children's Food Choices

Ozturk, Orgul and Frongillo, Edward and Blake, Christine and Turner-McGrievy, Gabrielle

University of South Carolina, Darla More School of Business, Economics Department, University of South Carolina, Arnold School Public Health Department of Health Promotion, Education and Behavior.

2019

Before the Lunch Line:
Effectiveness of Behavioral Economic Interventions for Pre-Commitment on Elementary School Children's Food Choices

Orgul D. Ozturk[1, 2]

Edward A. Frongillo[3]

Christine Blake[3]

Gabrielle Turner-McGrievy[3]

Abstract

In this study, we intervened in elementary schools on lunch entrée selection using some of the behavioral economic methods shown to be effective in earlier food choice studies. Unlike many earlier behavioral interventions, which were mostly done in controlled environments and smaller café type settings for one-off interactions, we conducted our interventions in a real-world environment in twelve elementary schools in one school district in South Carolina over nine school weeks. By increasing salience and prominence of the healthy entrée of the day through visual and verbal tools, we nudged students towards selecting healthier options in treatment schools. We estimated the treatment effects using a difference-in-differences setup, comparing changes in the share of students selecting nudged entrées during the treatment period relative to the shares before the treatment period in treatment and comparison schools. Our estimates show that the nudges are effective when present. They increase selection of the healthy option by thirteen to thirty-five percent on the days the entrée is treated. Effects disappear when the nudge is removed, however, and there is evidence for reduced effectiveness of nudges in repeat instances. There is no evidence of habit formation.

JEL Codes: C93, D91, I12
Keywords: Nudge, behavioral economics, healthy eating, school lunch, salience, prominence, difference-in-differences

---

[1] Corresponding Author. odozturk@moore.sc.edu.
This research was supported by a grant from Cornell BEN Center Small Grants Program.

[2] Darla Moore School of Business, University of South Carolina. 1014 Greene Street. Columbia, SC 29205

[3] Arnold School of Public Health, Department of Health Promotion, Education and Behavior. University of South Carolina

I. Introduction

Behavioral factors have been shown to affect choice in and out of laboratory environments on issues ranging from retirement savings (Choi et al., 2003) to coaching decisions in the National Basketball Association (Lefgren, Platt and Price, 2015). Perhaps not surprisingly, most common applications of behavioral economic research are on issues related to public health. Behavioral economic methods are shown to be effective by themselves or when paired with incentive-based interventions to encourage healthy behaviors such as medication adherence (Viswanathan et al. 2012), exercise (Royer, et al 2015), smoking-cessation (Volpp 2011, Gine et al. 2010), or take-up on other wellness initiatives such as health risk assessments (Haisley 2012). Over the last decade or so, there is a stream of literature studying the role of behavioral factors in health-related decision making and designing behavioral interventions to improve health decisions (see Hanoch, Barnes and Rice, 2017 for an extensive review).

This is especially true for nutrition-related decisions. A growing body of research looks at behavioral interventions aimed at altering food choice. The evidence points to the importance of some factors affecting food choice, other than available options or cost of options, such as the presentation of items and ease of access. These factors constitute what behavioral economists call the choice architecture (Thaler, and Sunstein, 2008). Specifically, in the context of the school lunches, the choice architecture includes the presentation order, the choices available, the default option, the speed of the lunch line, the relative convenience for selecting an item, the social context, and many other factors that subtly shape individual choices (Hanks et al., 2012). Thaler and Sunstein (2008) suggest we can redesign the choice architecture to "nudge" boundedly rational consumers to make better choices. Nudges and modification of school lunch-line environments based on behavioral economic theory has been shown to significantly increase the likelihood that children make healthy choices. Schwartz (2007) found that elementary school children significantly increase fruit consumption in response to verbal nudges of "would you like fruit or juice with your lunch." Changing the default from offering a fruit to serving a fruit increased fruit consumption, particularly when coupled with a small reward (Just, D. and Price, J. 2011). Providing a vegetable to elementary students while they waited in the lunch line significantly increased vegetable consumption (Elsbernd et al, 2016). Perry (2004) reported similar results from a cafeteria based randomized control trial. Other studies have shown that high school and college students respond to small changes, such as restricting debit cards to healthy foods or repositioning the salad bar (Hanks et al., 2012; French et al. 2017; Thomas, Desai, and Seenivasan, 2011). There are also papers showing effectiveness of nudges in food choices in the work place and other non-school environments (for example, Wisdom, Downs and Lowenstein, 2010 and Cioffi et al., 2015. See also Thompson and Ravia, 2011 for a systematic review of the earlier work).

In this paper, we contribute to this literature on effectiveness of behavioral methods in increasing choice of healthy food alternatives by intervening in a real-world setting. Validating laboratory findings in uncontrolled environments is crucial for the policy relevance for this type of work. We also contribute to a more general literature on habit formation. Behavioral interventions, if effective, can be an important policy tool in the quest to change nutritional choices made and habits formed, as they are easier and, in most cases, inexpensive to implement (Kessler, 2016). Thus, it is important to test if they are effective in habit formation.

We intervened in a randomly chosen group of elementary schools within a school district in South Carolina with intention to alter students' lunch entrée selection. We utilized some of the methods shown to be effective in earlier food choice studies which were not always tested in real world settings. Most of the earlier purely behavioral interventions provide evidence over a short period of time, thus not providing evidence of habit formation. We collected data from 12 elementary schools for over 75 school days (over 100,000 entrée/class/day observations) with data on pre- and post-intervention periods in addition to 8 weeks of treatment data. This setup enables us to not only test effectiveness of nudges in increasing healthy food selection by children in a real-world environment, but also allows us to test the persistence of their effectiveness when they are repeated and their potential to lead to formation of healthy eating habits.

II. Background

Elementary school children are actively forming tastes (Birch, 1999), so there is an opportunity to improve their nutrition and health by introducing new foods and encouraging preferences for more nutrient-dense foods that are less processed and have lower fat and sodium content. Existing research suggests nudges to nutritional decisions can constitute low cost interventions that can change behavior significantly. It is not clear, however, if these effects can be sustained when the nudge is repeated or if they persist in the absence of these nudges. Nutrition decisions are different than many other decisions individuals make due to their frequency. Nudges may induce the desired behavior when an individual is choosing retirement plans, but they may fail to work consistently when one is choosing what to eat over and over again. Moreover, in the case of repeated decisions, effectiveness of the nudge may fade over time, effects may not persist when not the nudge is not present and these interventions, no matter how low cost they are, may not lead to habit formation and be a waste of money and effort.

Habit formation results from the general literature are mixed. Even when researchers find effects persisting beyond the study period, they find the persistence is short-lived. (for example, Acland and Levy, 2015, with gym attendance; John et al., 2008, with weight loss; Volpp et al., 2009, for smoking cessation). Some recent nutrition studies tackled the question of habit formation with interventions that are extended over longer time periods, with data collected over the period post intervention, with mixed results. For example, Belot, James and Nolen (2016) and Just and Price (2013) find no evidence of habit formation, though the latter has a very short intervention period. With a longer study period, similar to Belot, James and Nolen (2016), List and Samek (2015) leverage behavioral tools with the goal of increasing effectiveness of incentives in increasing healthy food consumption in a large-scale field experiment. They show that incentives were not only effective in the short run in increasing the consumption of healthy snacks (though there was no differential effect due to loss/gain framing), but also there was some evidence of habit formation. Lowenstein, Price and Volpp (2016) provide the strongest evidence of habit formation with a large-scale intervention over a long study period. They do, however, not use behavioral interventions. Our study will contribute to this discussion with a pure behavioral economic intervention – at multiple sites and over a long period echoing these latest studies – and provide analysis of persistence of interventions and habit formation in children's food choices.

In the schools we are studying, we found from exploratory work that children are asked to preorder their lunch in the morning, before they get to the lunch line later in the day. Pre-ordering meals allows the food-service personnel to more efficiently prepare lunch and serves to speed the lunch line by reducing the number of decisions children must make while in line. This is important in interpreting our results as students are pre-committing to their entrée choice.[1] Pre-commitment has also been shown to significantly alter choices. Individuals may choose healthier food when they pre-order/commit, because the decision is guided by more self-control and less temptation. This may be particularly of relevance when the choice environment is as fast moving as the school lunch line. When individuals pre-commit, they are also not likely to switch as the pre-committed food item becomes their "default" and individuals are shown to stay with defaults (Wisdom, Downs and Lowenstein, 2010; Just, D. and Price, J. 2011). Effects of pre-ordering on healthy eating have been studied for high school students (Smith, 2012), for middle school students (Ferro, Gupta and Kropp, 2013) and elementary school students (Hanks et al., 2013; Miller et al, 2016), but our study is the first to use an existing pre-commitment system and complementing it with visual and verbal cues to nudge elementary school children towards selecting healthier entrées. Thus, we can consider our estimates as lower bounds to treatment effects by nudges, as it may be harder to nudge choice which is already been improved by pre-commitment.

III. Research Design and Methodology

a. Choice Architecture Survey

We used the survey instrument designed by Ozturk et al. (2016) to identify the components of the school lunch environment that could be nudged. This survey instrument was designed based on observations of lunch lines in 16 elementary schools in the same district where we collected the entrée choice data used in the current paper. The main observation from this survey instrument was that the food choice architecture extended beyond the lunchroom into the classroom. Students were required to pre-commit to a main entrée each morning in their home classroom and the teacher relayed their selections to the cafeteria staff in advance. Another observation was that the teachers played an important role in the selection of other foods once in the cafeteria as well and the influence varied by teacher, age of students, and school culture. Default options were determined idiosyncratically by lunch-line staff, and interaction between teachers and lunch-line staff determined whether children had autonomy to select healthier options.

[1] They get to choose their drinks and side items in the lunch line.

Based on our survey, we identified the pre-commitment phase as the most promising and least intrusive target for behavioral economic intervention. This enabled us to use a pre-existing mechanism for data collection. This also led us to concentrate on the entrée choice decision, which was least affected by the lunchroom influences.

b. Study Setting

For this project, we collaborated with the food service provider for one of the school districts in Columbia, South Carolina. The food service provider was following the rules for the National School Lunch and School Breakfast Programs. Our survey showed that the lunch environment across schools was homogenous in terms of menu, quality of food, and presentation; there is great diversity, however, in the school environment in terms of percent free and reduced lunch (11%-89%), ethnicity, and age of facility.

The district-wide menu featured a choice of 4 entrées, 2 vegetables, and 2 fruits. Students pre-order only their entrée choice. Our team ranked the entrée options in terms of their nutritional quality and identified the healthier default menu items to be promoted. Healthier menu items were defined as those that have greater nutrient density and lower amounts of nutrients that should be limited. Specifically, school lunch menu items were rated on meeting the 2012 Nutrition Standards for School Meals (USDA, 2012) for whole grains (using ≥ 6 g of fiber per entrée as a threshold), sodium (2012 nutrition standard suggest overall meals contain ≤1,230 mg so we used <1000 mg to rate each entrée), saturated fat (< 10% energy), and energy (<600 kilocalories). In addition, meeting one-third of the Recommended Dietary Allowance for calcium and protein and having menu items that contained less than one-third of the 300 mg/day maximum recommendation for cholesterol was used. Each school lunch menu choice was evaluated on these seven criteria with each criterion receiving one point each (for a range of 0 to 7 points). For example, a default meal of peanut butter and jelly on whole wheat bread, steamed corn, and fresh apple (6 out of 7 points for <600 kcals, <1000 mg sodium, <10% saturated fat, >6 g of fiber, >10 g of protein, and < 50 mg of cholesterol) would be promoted against cheese pizza, tater tots, and canned diced pears (2 points for >10 g of protein and > 50 mg of calcium). Appendix A details the algorithm used and the tie-breaking rules.

c. Decision Points

Our intervention to the choice architecture was through existing channels already in use by the food service provider, including menus that are posted and sent home, morning announcements promoting healthy eating, and pre-commitment in the classroom. Because children were making choices about what to eat in advance of the lunch line, the intervention focused on nudges aimed at these earlier choice points. Students were asked to select entrées in their classrooms in the morning, and, while it was possible to make a different choice on the line, the speed of the lunch line makes this unlikely. Ozturk et al (2016) notes that children can choose their drink and side items in line and there usually are verbal nudges by lunch room staff and teachers towards healthier options on these dimensions.[2] In addition, it was possible that some students made their choices before they arrived at school. Monthly menus were sent home giving parents a role in choosing whether to purchase a lunch and what to eat on a given day.

d. Nudges

The nudges used in this study were based on literature showing that **defaults** and **salience and prominence** of selected food items can be manipulated to increase consumption (Choi et al.,2003; Johnson and Goldstein, 2003; Wisdom, Downs and Lowenstein, 2010; Just and Price, 2011). The specific types of interventions used were inspired by the earlier work showing theme-related food names can increase the appeal of the default (highlighted) menu items and visual cues about what to place on the tray may increase consumption by establishing healthy norms (Reicks,et al, 2012). To optimize our promotional materials, we consulted with a consumer behavior expert and a graphic artist. Dinosaurs and mystery/detective cartoon characters were identified as two age-appropriate themes from which we chose illustrations and food *nicknames*. We used 4 different combinations of interventions. We had two themes (Dinosaurs and Mystery/Detective) and two ways to increase the salience/prominence of the preferred food items (Highlight or Names). Examples of the art used and materials distributed are given in Appendix B. These menus

[2] To the best of our knowledge there was not any changes to entrée choices in the lunch line. We believe if the teachers had any intention to alter students' entrée choices they did so in the morning in the classroom. Moreover, we are only tracking entrée counts and cannot speak to treatment / spillover effects to other side item options highlighted in the treatment materials.

were sent home weekly and a slideshow featuring the visual nudges was shown in the morning in each class in the treatment schools. A sample menu from the control schools is also given in this appendix.

e. Design and Data Collection

The 18 elementary schools in the district were divided into 9 groups of 2 based on percent free and reduced lunch. We randomly assigned one of the schools in each group to be a treatment school with a random number generator, assigning the school with the lower number to the treatment. The unit of observation was the class (teacher/grade). The menu repeated every 10 days, so over a 3-month study period we could observe a class' choice for a given menu up to 6 times depending on holidays and special events. In this district, each teacher communicates the entrée counts to the café manager, and these records were obtained as the primary data source.

Data (via production sheets from the food service providers at the schools) for 26 school days before treatment were collected. There was then a 9-week treatment period. During this period, we used modified menus and promotional items in the treated schools. In the middle of this 9-week period, we had one week of no treatment giving us a reset period between different treatment combinations. We continued to collect data for 2 more weeks following the treatment period. The treatment timing and schedule is provided in Appendix Table B1.[3]

Over the course of the study period, data were collected on entrée choices made by students in 14 schools (we started with 18 schools, but 4 schools dropped out of the study; 2 of which dropped out early on and 2 more during the treatment period. Two other schools did not have sufficient pre-treatment data so we excluded them from our analysis. As a result we have 12 schools in our analysis, 7 of which are treated)[4] representing 6 grade levels (K thru 5) in 533 unique classrooms. Data from the pre-treatment and first treatment periods are used for our main analysis where there are over 39,000 observations. Then, the post treatment period data were added where there are about 8,000 observations, bringing our sample size to almost 47,000. There were three main outcomes of interest: 1) share of the *treated* healthy

---

[3] Day count is based on actual school days when schools are open and data is collected. Not all school weeks are 5-day weeks.

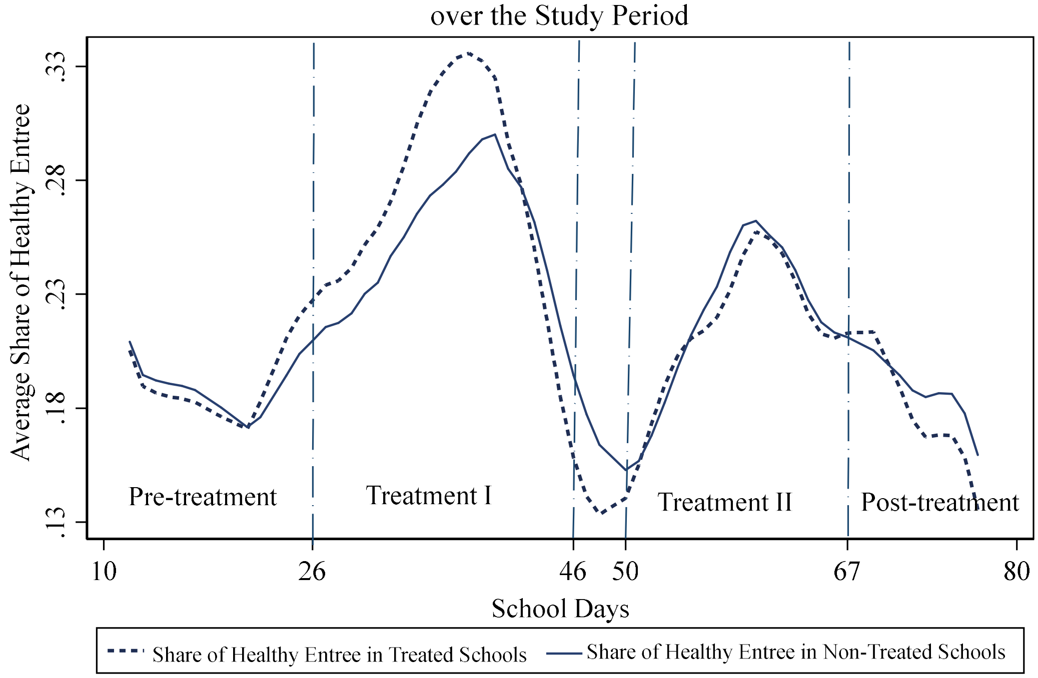
[4] We use school characteristics to control for possible confounding factors, to factor in the possibility of non-random attrition, in addition to unconditional comparisons of control and treatment schools and periods. Values for these school characteristics are given in Appendix Table C1.

entrées in the daily entrée counts when the treatment is present and 2) persistence of these treatment effects when the nudges are repeated and 3) effects on shares in the post-treatment period, i.e. evidence of habit formation.

IV. Unconditional Treatment Effects of Nudges - Comparison of Means

Figure 1 visualizes the pre-treatment trends and the treatment effects by showing daily shares of healthy (treated) entrées over the study period in treatment and control schools. It provides visual evidence that before our intervention, shares of the treated entrées evolved in a similar fashion in treatment and control schools. According to this figure, the treatment effect is present only in the first treatment period. During the break between treatment periods the effect vanishes, and the second period interventions do not increase the selection of healthy entrée significantly.

Figure 1: Average Share of Healthy Entrees in Treatment and Control Schools over the Study Period



Notes: This figure captures the differences in shares of ever-treated entrées in treatment and control schools over the course of the study period. Vertical lines indicate end points of each treatment period as labelled. X-axis is in days of the semester. Created using *lowess* command in Stata 15 with bandwith of 0.4 points.

Table 1 cross tabulates the average shares for treated and non-treated entrées by location and time, comparing pre-treatment period shares to first treatment period shares. The upper panel of the table reports the data for the treatment schools and the lower panel shows them for the control schools. Each panel's rows further divide the data by time dimension: pre-treatment (first three weeks of our study period) versus the initial intervention period (first round of intervention – weeks 4 thru 8). Columns, on the other hand, split the data by the treatment status of the entrée. Each cell reports the average share for the food group in a given location and time period with standard errors in parentheses and the number of observations in the square brackets.

Table 1: Treatment Effects - Difference in Difference in Differences using Mean Entrée Shares

		Non-treated Entrees	Treated Entrees	*Difference*
treated schools	**Before**	0.292	0.191	-0.100
		(0.004)	(0.004)	(0.005)
		[6,855]	[4,215]	[11,070]
	During	0.253	0.278	0.025
		(0.003)	(0.004)	(0.005)
		[7,752]	[5,755]	[13,507]
	Change over time	*-0.039*	*0.086*	*DD_{TS}=0.126*
		(0.005)	*(0.006)*	*(0.008)*
control schools	**Before**	0.300	0.187	-0.113
		(0.005)	(0.006)	(0.008)
		[3,437]	[2,040]	[5,477]
	During	0.265	0.251	-0.014
		(0.004)	(0.005)	(0.006)
		[5,132]	[3,847]	[8,979]
	Change over time	*-0.035*	*0.064*	*DD_{NTS}=0.099*
		(0.007)	*(0.008)*	*(0.011)*
DDD = DD_{TS} - DD_{NTS}				***0.026***
				(0.013)

The first difference between shares of treated entrées and non-treated entrées before and during the treatment period in treated schools gives us the difference-in-differences treatment effect estimate for the treated schools. We get a similar difference-in-differences estimate for the non-treated schools also. Comparing the difference-in-differences in treatment schools to the corresponding change in the control schools gives us the triple-differences estimate that we seek. This calculation gives a triple-differences estimator for the treatment effect of 2.6 percentage points, which is statistically significant.

Identification of unbiased treatment effect from a difference-and-differences setup requires that in the absence of treatment the shares of the entrées in each group (treated and non-treated) would have evolved similarly in treatment and control schools (common trends assumption; Meyer, 1995). This is not obvious from Table 1 since the share of entrées designated as treated food increase in both treated and control schools over time. Figure 1 provides visual evidence, however, and next we will provide evidence from regression analysis.

We will deliberate the role of preexisting factors such as teachers as influencers on the existence of these trends later in the *Discussion* section. If there were no shocks that affected treated food differentially in treated schools compared to in control schools, then the triple-differences estimate is unbiased. In our regression analysis in the next section we will reproduce these estimates and assess their robustness using further controls for entrée, grade, and teacher fixed effects and school-level characteristics.

V. Econometric Analysis of Treatment Effects

The main focus of our analysis is the effectiveness of the treatment on the selection of the entrées designated as the healthiest option that day. We are analyzing the data to determine if children are choosing the healthier options they are *nudged* towards. If this is the case, on the treatment days the shares of the treated entrées should be higher than what they would have been in the absence of a treatment at the treated schools. Since this counterfactual cannot exist, in our analytical setup we will obtain a treatment effect using a difference-in-differences (-in-differences) method and compare changes of the shares of treated entrées to changes in *other* shares. In order to specify the *other shares* there are several things worth clarifying about our intervention: To repeat an obvious point, before our intervention period there were no treated entrées. During the two treatment periods, each day there was only one nudged/treated entrée in the treatment schools. Thus, the same entrée could have been a non-nudged entrée some other days during the treatment period even in treatment schools, although it was never a treated entrée in control schools. Hence, we can identify the effect of treatment using *differences* in 4 dimensions: within schools between entrées, within schools between time periods between (treated/non-treated) schools, and for each treated entrée by treatment presence, and by incidence or *order* of the nudge. Specifically, we calculate the changes in the mean share of the entrées that were (ever) treated compared to never-treated foods (1) in schools where they are nudged during treatment period (treated schools) and in control schools where no treatment was ever present (2) over time, i.e., during the treatment period relative to the pre-treatment, (3) the difference in the evolution of these two sets of differences between control and treatment schools and (4) for any ever treated entrée the difference in shares when the treatment is present, first time it is treated vs. all other times after the first-ever treatment.

Baseline Model: Treatment Effect of Being Ever Treated

In our intervention, our goal was to be minimally intrusive which limits the changes we could make. This provided the study with a realistic environment and easily replicable intervention, but also a more complex treatment design. We did not change anything in the food offerings, but chose the one that best fits within the nutritional parameters set by USDA as described above. As a result, we did not necessarily have a pre-treatment observation for all of our treated entrées. Each entrée has a different baseline share, and in order to identify the treatment effect without any possible bias resulting from entrée combinations, we restricted our sample to the entrées that were observed both before and during the treatment period. For both the treated entrée and the non-treated entrée, as a result, we have class-level shares from the period prior to the treatment and during the treatment in both treatment and control schools. We also drop a couple of schools from our data which only had treatment period data but no pre-treatment counts reported. The production reporting by the lunch-room managers in the first couple of weeks was not consistent while the study was starting, and we dropped these early days. We ended up with 3 weeks of pre-treatment and 8 weeks of treatment.[5] We estimated the following model:

$$\begin{aligned}\text{Share}_{itscd} = {} & \beta_0 + \beta_1\text{TreatedSchool}_c + \beta_2\text{EverTreatedFood}_i \\ & + \beta_3\text{TreatmentPeriod}_d + \beta_4\text{EverTreatedFood}_i \,\mathbf{x}\, \text{TreatedSchool}_c \\ & + \beta_5\text{TreatmentPeriod}_d \,\mathbf{x}\, \text{TreatedSchool}_c \\ & + \beta_6\text{EverTreatedFood}_i \,\mathbf{x}\, \text{TreatmentPeriod}_d \,\mathbf{x}\, \text{TreatedSchool}_c \\ & + \gamma'\text{X}_s + \delta_c + \eta_i + \varepsilon_{itscd}\end{aligned}$$

[5] Second treatment period had many special lunch events such as “Thanksgiving feasts” for the families and field trips with many sandwich only/bagged lunch days. Thus, we exclude those observations from our analysis.

where *Share* indicates share of entrée *i* by treatment status *t* (treated or non-treated) in school *s* class *c* (a unique teacher/grade combination) on day *d*. *TreatedSchool* is an indicator for classes in schools where nudges were in effect during the treatment period which is captured by indicator, *TreatmentPeriod*. *EverTreatedFood* is an indicator for entrées which are ever designated as the healthiest option. Thus, β_6 is a triple-difference estimator which captures the average difference-in-difference-in-differences in shares of healthy entrées vs non-healthy entrées in treated school versus non-treated schools in treatment period compared to the pre-treatment period. This measure does not distinguish the effect of being the treated entrée of the day from the effect of being ever treated (puts equal weight on treated entrée shares for all days of treatment period), thus it may not directly and fully capture the effect of the nudge. We will build on this specification to address this distinction and differentiate the first incidence of the nudge from subsequent incidences to test persistence of treatment effects. Lastly by including post treatment data we will directly tackle the question of effectiveness of nudges in habit formation in a more standard model.

Our preferred model specification includes a vector of school characteristics (X_s) as controls, as well as teacher(classroom) and food fixed effects (δ_c and η_i, respectively). In presenting our regression results, we also provide model specifications with no entrée or teacher fixed effects and no school level controls to gauge the role of these factors in choice of entrée.

Average Difference in Differences Estimates for Effect of Nudges on Healthy Entrée Selection

Our setting is not a perfect treatment and control environment, and it is possible that some non-random distribution of school or student level component artificially generates the treatment effect we observe. For this reason, we also provide the regression analysis with controls for school level socioeconomic and demographic characteristics, teacher and food fixed effects.

The first column of the Table 2 regenerates the numbers for different components of the DDD estimates in Table 1 with subcluster wild bootstrap p-values.[6] We add to this simple

[6] In our analysis, we only have 12 clusters, 7 of which are treated. It has been shown in the literature (see for example Cameron, Gelbach and Miller, 2008; Conley and Taber 2011) when there are few (treated) clusters

Table 2: Average Treatment Effects for Ever-Treated Entrees - Conditional Difference in Difference Estimates

Ever-treated entrée	-0.113**	-0.113**	-0.113*	-0.113**	-0.053**	-0.053**	-0.053**
	(0.005)	(0.006)	(0.005)	(0.005)	(0.008)	(0.008)	(0.009)
Treated school	-0.008	-0.012+	-0.007	-0.003	-0.000	-0.004	0.006
	(0.219)	(0.062)	(0.283)	(0.601)	(0.978)	(0.578)	(0.547)
Treatment period	-0.035**	-0.035**	-0.036**	-0.036**	0.009	0.009+	0.009+
	(0.004)	(0.001)	(0.002)	(0.002)	(0.109)	(0.090)	(0.088)
Ever-treated entrée x Treated school	0.013	0.013	0.013	0.013	0.001	0.001	0.001
	(0.268)	(0.267)	(0.267)	(0.267)	(0.934)	(0.911)	(0.911)
Treated school x Treatment period	-0.004	-0.004	-0.004	-0.004	-0.007	-0.008	-0.007
	(0.472)	(0.458)	(0.494)	(0.494)	(0.236)	(0.221)	(0.237)
Ever-treated entrée x Treatment period	0.099*	0.100*	0.100*	0.100*	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.929)	(0.902)	(0.899)
Ever-treated entrée x Treated school x Treatment period	0.026+	0.026+	0.026+	0.026+	0.030*	0.029*	0.029*
	(0.067)	(0.069)	(0.070)	(0.070)	(0.027)	(0.028)	(0.028)
School characteristics	NO	YES	NO	YES	NO	YES	YES
Teacher fixed effects	NO	NO	YES	YES	NO	NO	YES
Food fixed effects	NO	NO	NO	NO	YES	YES	YES
Observations	39,033	39,033	39,033	39,033	39,033	39,033	39,033
R-squared	0.014	0.014	0.014	0.014	0.638	0.638	0.639

Notes: Coefficients are estimated using OLS regressions. Subcluster wild bootstrap p-values (reported in parenthesis) are obtained using *boottest* command (Roodman, 2015) in Stata 15 with 5000 replications and Webb weights. Share of an entrée in total number of entrees ordered by a class is the outcome variable. Ever treated entrée is an entrée that is at some point designated as "the healthy entrée of the day" during treatment period and treated school indicates if the school had any treatment. Standard errors are clustered at school level. **,*, + indicate significance at 1%, 5% and 10% respectively using subcluster wild bootstrap p-values.

specification different combinations of teacher and food fixed effects along with controls for socio-demographic composition of the school in subsequent specifications. The most informative additions are the food fixed effects. Our healthy entrées are of a wide variety, from PB and J to Broccoli Chicken Alfredo and are treated at different frequencies (once to 10 times, with a mode of 4 times). Untreated entrées are equally diverse. Note that

robust standard errors can be severely biased as the large-sample assumptions do not hold. Our cluster number and near balance between the number of treated and non-treated clusters enables us to use wild cluster bootstrap p-values, which is commonly used as an unbiased alternative. However, MacKinnon and Webb (2017) show that wild cluster bootstrap also can fail in difference-in-differences settings as not all observations in treated clusters are always treated and imbalance in numbers of treated vs untreated observations can result in over- or under-rejection. MacKinnon and Webb (2018) proposed sub cluster bootstrap as a way to reduce this problem we generated sub-cluster wild bootstrap p-values using Stata 15. These sub-cluster wild bootstrap p-values are reported in the tables in brackets. Wild bootstrap p-values are reported along with cluster robust standard errors from regressions in Appendix Table D1 for Table 2 estimates in addition to subcluster wild bootstrap p-values reported in the main table. Same results are available for all other tables upon request.

coefficients of variables *Ever-treated entrée* and *Treatment Period* (capturing effect on untreated entrées) and the interaction of the two change significantly when food fixed effects are included. We believe this is simply capturing that the menu is cyclical and baseline selection rates for treated and untreated entrées alike have a range. Our treatment period seems to have coincided with a specific set of healthy foods with higher baselines. Thus it is important to control for entrée-specific fixed effects.

Several important observations can be made from these regression results. First the entrées designated as healthy entrées are on average less popular than the other entrées (from the coefficient of *Ever-treated entrée*). On average, the share of students selecting these entrées are about 5 percentage points lower in our preferred models with food fixed effects. There are no significant differences in shares of healthy entrées in treatment schools versus control schools in pre-treatment period (*Ever-treated entrée* X *Treated School)*. In models with no food fixed effects there appears to be a statistically and economically significant increase in shares of ever-treated entrées and a decrease in shares of all other entrées in all schools (coefficient estimates for "*Treatment period*" and "*Ever-treated entrée* X *Treatment period*" and for both treated and control schools). Once we control for food fixed effects however, size of these coefficients shrink significantly, signs reverse, and only a minimal increase in the share of all other entrées in the control schools remain marginally significant.

Most importantly, regression results are consistent with the unconditional difference in difference estimates in terms of effectiveness of nudges. In our preferred models where we control for school characteristics, teacher and food fixed effects, treatment effect is an increase of about 3 percentage points on average capturing the differential change in the share of treated entrées in treated schools relative to pre-treatment period compared to the change in these shares in non-treated schools over the same two periods. This is about a 13% increase in the shares of treated entrées due to being ever-nudged.[7]

The main outcome of interest is the changes in the share of entrées. Before we provide a deeper analysis of the nudge treatment effects we also want to know if there is any change in lunch participation due to treatment. That is, we ask if the introduction of nudge materials increased the number of children eating the lunch provided in the cafeteria instead of bringing

[7] Using fitted values using estimates of our preferred model we also recreated Figure 1. It is given in Appendix D.

lunch from home. This effect is of interest for a couple of reasons. Students who bring lunch from home may be making healthier choices on average and our treatment effect can be due to just the compositional changes in the student body who is eating lunch at school. Also increased participation may make the lunch line longer and result in limited time to eat. This negates the benefits from increasing pick up of health entrée by reducing opportunity to consume it. In order to understand if our treatment effect is robust and is not a byproduct of a change in composition and size of the lunch crowd, we compared the total number of entrées ordered in a day with and without treatment (in classes in treated vs control schools) during treatment period relative to the pre-treatment period. Appendix Table D2 provides the results for the analysis of the effects of nudges on participation in school lunch program. Using a difference-in-differences setup, we show that there are no significant changes in participation level either at classroom or at the school level. Nudges do not seem to be affecting the external margin of lunch participation, but only changing lunch entrée selected for already participating students.

ii. Persistence of Treatment Effects

Spillover Treatment effects and treatment effects when the nudge is not present

In our intervention, unlike a controlled experimental environment or other interventions using school settings, not all treated foods are treated every day, though there is always a treated entrée during the treatment days in treated schools. We work with the existing school lunch environment with no structural changes and choose the healthiest option as the entrée to be nudged. As a result, Table 2 captures changes not only from the days these entrées are nudged, but also from the days when they are not, during the treatment period. Treated entrées, on average, have lower shares in the pre-period. If the nudge is only effective when present, even during the treatment period they may have lower shares on days they are not treated (but served as an option) and this may result in a downward bias in the effect of being the nudged entrée of the day. If there is, however, any lingering treatment effect in days subsequent to treatment day when this entrée is served (even when the nudge is not present) the shares may be larger than the pre-treatment period shares even without the daily nudge (we label this *spillover effect*).

We tackle this issue by defining treatment with more nuance. In the next set of regressions, we specify treatment period by entrée, in addition to timeline definition, specifically differentiating between the period before the first time an entrée is designated as treatment entrée and when nudged and any day before that. In addition, we have an indicator for the entrée that is nudged each day, capturing the presence of the nudge for the entrée in a given day.

Table 3 reports these results for our preferred model with full set of controls.[8] Highlighted rows report coefficients of interest; *spillover* treatment effect and treatment effect when the nudge is present. Treatment effect estimate from this table is almost double the previous estimate when food fixed effects are not included (See the unconditional estimates in

Table 3: Persistence of Treatment Effects of Nudges - Spillover Treatment Effects

Ever-treated entrée	-0.068**	[0.002]
Treated school	0.006	[0.547]
Treatment period	0.009+	[0.088]
Treated school x Treatment period	-0.007	[0.237]
Ever-treated entrée x Treated school	0.002	[0.871]
Ever-treated entrée x Treatment period	-0.005	[0.510]
Ever-treated entrée x Treated school x Treatment period	-0.012	[0.653]
Ever-treated entrée x After the first treatment	-0.017*	[0.036]
Ever-treated entrée x After the first treatment x Treated school	0.022	[0.228]
Treated entrée	0.035*	[0.045]
Treated entrée x Treated school	0.034*	[0.047]
Observations	39,033	
R-squared	0.641	

Notes: Model also includes school characteristics, food fixed effects and teacher fixed effects. *Treated entrée* indicates the entrée designated as the nudged entrée on a given day. *After the first treatment* is an indicator for an ever-treated entrée for all days when it is not the treated entree after it is treated at least once. See also notes for Table 2.

Appendix Table D1 and the full set of estimates on Appendix Table D2), but only slightly larger in the preferred model. Using this model, the nudge effect is estimated to be about 3.5 percentage points. This effect corresponds to about 18 percent additional increase in the consumption share of the treated food when it is the treated entrée of the day relative to the

[8] Unconditional Treatment effect for this setup is calculated in Appendix Table D3 and regression results with different set of controls are provided in Appendix Table D4.

control schools. Estimates for the differential change in share of these entrées in days they are not treated – coefficient of *Ever-treated entrée x After the first treatment x Treated school* interaction – is about 2/3 of the size of the coefficient of *Treated entrée* x *Treated school* interaction, but not statistically different from zero.

Another interesting observation from this table is about the changes in the control schools. There is a significant change in the selection of the treated entrée of the day in control schools as well. This indicates pre-commitment due to pre-ordering may be effective in having children choose the healthiest entrée of the day. It is likely that teachers or parents nudge students to try these entrées. In control schools after the first incidence when this entrée is offered but is not the healthiest option, however, the share is significantly lower compared to the pre-treatment period by about 1.7 percentage points. Children may be trying the healthier options when they pre-commit, but only once and not in repeated instances. Though insignificant, large and positive differential effects we found in treatment schools may indicate some persistent treatment effects.

Effectiveness of Repeated Nudges

In order to more directly measure persistence of treatment effects we next differentiate the effect for the first incidence of the nudge for a specific entrée from the subsequent incidences for the same entrée. We achieve this by introducing two new dummy indicators to our model: *First treatment* and *Repeat treatment.* Indicator *First treatment* captures the day when an ever-treated food is the nudged entrée for the first time. All other times it is treated the indicator *Repeated treatment* turns on, instead. *Ever-treated entrée* x *Treatment period* interaction captures all other days when the entrée is offered, but is not the healthiest option of the day. Interaction of these three variables with the *Treated school* dummy captures the differential effect for the treated schools. Specifically, *Treated school x First Treatment* is the treatment effect of the nudge when it is nudged for the first time. Estimate for the coefficient of this variable indicates on average 6.8 percentage points differential change in the selection of the treated entrées relative to their pre-treatment shares in treated schools compared to the corresponding change in control schools when they are the healthiest option for the first time. This is a treatment effect of almost 35 percent increase. If the nudges were persistently effective the coefficient of the *Treated school x Repeat Treatment* would have been of similar size and significance. The estimate, however, is indicating a statistically zero effect.

Table 4: Persistence of Treatment Effects of Nudges - Initial vs Subsequent Nudges

Ever-treated entrée	-0.067**	[0.002]
Treated school	0.004	[0.674]
Treatment period	0.009+	[0.090]
Treated school x Treatment period	-0.008	[0.229]
Ever-treated entrée x Treated school	0.002	[0.848]
Ever-treated entrée x Treatment period	-0.020+	[0.127]
First treatment	0.052*	[0.032]
Repeat treatment	0.013	[0.403]
Ever- treated entrée x Treated school x Treatment period	0.004	[0.834]
Treated school x First Treatment	0.068**	[0.003]
Treated school x Repeat Treatment	0.001	[0.958]
School characteristics	YES	
Teacher fixed effects	YES	
Entrée fixed effects	YES	
Observations	39,033	
R-squared	0.643	

Notes: F*irst treatment* is an indicator for an ever-treated entree for the first day it is the treated entree and *Repeat treatment* is an indicator for all subsequent times it is the treated entree. See also Table 2 notes and Table 3 notes for all other variable definitions.

Persistence analysis for treatment period effects all point to fading treatment effects and therefore a lack of evidence for habit formation. Next we test the habit formation with a setup more in the spirit of earlier studies, by introducing data from the post-treatment period to our analysis.

Post-Treatment Period Effects: Habit formation

The goal of any intervention and most important outcome of interest is the sustainability of the treatment effects or incidence of habit formation. Does the treatment effect persist beyond the treatment period in treated schools? Do children still select more of the entrées that were promoted during the treatment period when the nudges are no longer present? Do we see healthier food choice habits formed? In our earlier analysis, we estimated no significant differential change in consumption of treated foods in treated schools during the treatment period on days when they were not the treated entrée. However, the lack of effect during treatment period is likely due to the existence of another nudged entrée or another ever-treated option. To overcome this issue, we compare the average shares of ever-promoted entrées after the treatment to their shares before and during the treatment. In our analytical setup we distinguish between three time periods (pre-treatment, treatment and post-treatment). Table 5 reports the estimates for this model. Highlighted row reports the estimate of interest indicating no persistence for the treatment effects estimated for nudges when they were present during the treatment period. There is no evidence of habit formation.

Table 5: Persistence of Treatment Effects of Nudges - Habit Formation/Post Treatment Effects

Ever-treated entrée	-0.055**	[0.006]
Treated school	0.012	[0.999]
Treatment period	0.004	[0.502]
Post-treatment period	-0.004	[0.802]
Treated school **x** Treatment period	-0.005	[0.406]
Treated school **x** Post-treatment period	-0.015	[0.421]
Ever-treated entrée **x** Treated school	0.001	[0.940]
Ever-treated entrée **x** Treatment period	0.008	[0.356]
Ever-treated entrée **x** Post-treatment period	0.015	[0.291]
Ever-treated entrée **x** Treated school **x** Treatment period	0.029*	[0.039]
Ever-treated entrée **x** Treated school **x** Post-treatment period	-0.002	[0.930]
Observations	46,668	
R-squared	0.628	

Notes: *Post-Treatment period* is an indicator for the two-week period following the end of the second treatment period. See also Table 2 notes.

VI. Conclusion and Discussion

In most applications of behavioral economic interventions, where nudges are shown to be effective, a decision process that is nudged is not a repeating event and decisions are often one-off. This is not the case, however, for nutrition-related food choice decisions. For this reason, it is particularly important to establish effectiveness of these interventions in real-world environments as the decisions are made repeatedly. Many of the nutrition-related behavioral interventions in literature are done in laboratory or controlled environments and are not repeated. Even though findings of these studies are supportive of effectiveness of nudges, it is hard to conclude that they can be effective in altering daily decisions in real-world environments.

In this paper, we provide evidence on effectiveness of nudges in changing lunch food choice of elementary school age children in a real-world environment. We show that during the treatment period, selection of the healthy entrées increased on average by about 15 percent. This change was not due to changes in composition of the school lunch crowd, i.e. our treatment did not differentially affect school lunch participation. We find, however, no evidence of *persistent* treatment effects. We measure persistence in several ways, testing existence of spillover treatment effects in days when a previously treated entrée is offered but not targeted, comparing treatment effects in the first incidence of the nudge vs. the subsequent incidence for an entrée and, in a more traditional way, by testing habit formation with the addition of post-treatment data. Though treatment effect at the first incidence of the nudge was twice the size of the average treatment effects in models without this distinction, subsequent nudges failed to differentially change the selection of a given healthy entrée. There was also no evidence of habit formation when selection rates for healthy entrées were compared pre-, during and post-treatment periods; differential change in treatment schools post treatment is a statistical zero.

There are several possible reasons why our interventions did not result in persistent treatment effects that do not fade with repetition over the course of the treatment period and do not disappear post-treatment. Literature shows habit formations takes many repetitions (Cooke, 2007; Wardle et al, 2003; Laureati et al, 2014; Skinner et al, 2002). For habits to form with the aid of nudges they may have to be repeated multiple times with the same food. This was the case for some entrées, but not all in our setting. It is doubtful, however, that this would have altered our findings. We find, despite large treatment effect for the first-time nudge for the treated entrées, there is no significant treatment effect on average for the subsequent times the entrée is nudged. We cannot measure if the selected entrée is indeed consumed. Literature highlights the role of frequency of tasting a food in developing preference for it (Birch, 1999 and Laureati et al, 2014 among others). Some of the earlier studies, that document habit formation, use incentive-based intervention and reward consumption (Lowenstein et al 2016, List and Price, 2015). Persistent treatment effects they document may be due to existence of the incentives, but habit formation is likely achieved by repeat consumption of the rewarded food.

Another important aspect to consider is the preexistence of pre-ordering system in the schools we study. Pre-commitment has been shown to lead to healthier choices. If this was the case already in our schools, we might have been working with a limited bandwidth of possible treatment effects. Even though we do not observe persistent treatment effects, our work shows that there may be room to improve any given treatment by leveraging secondary methods.

Moreover, we are trying to nudge entrées, not snacks (such as fruits and vegetables) as in many earlier studies. Risk aversion is shown to be important in choosing familiar foods and not trying new or healthier options (Daniel, 2015). Children may be more risk-averse regarding the choice of their entrée compared to selection of snacks. It may be more important to choose something they want to consume. They may be enticed by nudges once in selecting the entrées, but not in the subsequent incidence if they did not like the entrée they chose. Lack of persistence in treatment effect may be due to the complexity of factors involved in selection of the main food item in comparison to selection of snacks or side items.

Behavioral interventions we utilized are low-cost and can be adjusted at school and classroom level to target student body interests or seasonal events. Increase in selection on the first instance of a nudge for an entrée is quite promising, but without persistent effects these treatments fall short. Future work is needed to test effectiveness and persistence with low-stakes foods as treatment targets. Moreover, incentives can be built in to increase persistence of treatment effects and encourage consumption of the treatment food, not only selection, to form tastes and potentially change habits.

References

Birch LL. 1999. “Development of food preferences.” *Annual Review of Nutrition.* Vol. 19 pp.41–62.

Belot, Michele, Jonathan James, Patrick Nolan. 2016. "Incentives and Dietary Choices – A Field Experiment in Primary Schools" *Journal of Health Economics* 50, 213-29

Cameron, A. C., Gelbach, J. B., Miller, D. L. 2008. Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics* 90: 414–427.

Choi, J.J., D. Laibson, B.C. Madrian and A. Metrick. 2003. “Optimal Defaults.” *American Economic Review* Vol. 93(2) pp.180-185.

Cioffi, Catherine, E.David, A.Levitsky, Carly R.Pacanowski and Fredrik Bertz. 2015. “A nudge in a healthy direction. The effect of nutrition labels on food purchasing behaviors in university dining facilities”. *Appetite.* Vol. 92. pp 7-14

Conley, T. G., Taber, C. R. 2011. "Inference with "difference in differences" with a small number of policy changes". *Review of Economics and Statistics* 93: 113–125.

Cooke, L. (2007), The importance of exposure for healthy eating in childhood: a review. *Journal of Human Nutrition and Dietetics*. Vol. 20. pp. 294-301.

Daniel, Caitlin. 2015. "Economic constraints on taste formation and the true cost of healthy eating." *Social Science & Medicine V.* 148 pp.34-41. doi:10.1016/j.socscimed.2015.11.025

Elsbernd, S.L., M.M. Reicks, T.L. Mann, J.P. Redden, E. Mykerezi, Z.M. Vickers. 2016. "Serving vegetables first: A strategy to increase vegetable consumption in elementary school cafeterias." *Appetite*. Vol.96. pp. 111-115.

Ferro, G., S. Gupta, and J. Kropp. 2013. "The Effect of Pre-Selection and Visual Cues on Food Item Selection by Middle School Children", Annual Meeting, Washington, D.C., Agricultural and Applied Economics Association, August.

FNS, USDA. 2012. "Nutrition Standards in the National School Lunch and School Breakfast Programs." *Federal Register* 77(17). pp.4088-167.

French, S.A., Rydell, S.A., Mitchell, N.R. et al. Financial incentives and purchase restrictions in a food benefit program affect the types of foods and beverages purchased: results from a randomized trial. Int J Behav Nutr Phys Act 14, 127 (2017). https://doi.org/10.1186/s12966-017-0585-9

Haisley E, K.G. Volpp, T. Pellathy , G. Loewenstein. 2012. "The impact of alternative incentive schemes on completion of health risk assessments". Am J Health Promot. Jan-Feb;26(3):184-8. doi: 10.4278/ajhp.100729-ARB-257.

Hanoch, Y, A. J. Barnes, & T. Rice (Eds.), <u>Behavioral economics and healthy behaviors: Key concepts and current research</u> (p. 109–126). Routledge/Taylor & Francis Group.

Johnson, Eric J. and Daniel G. Goldstein. 2003. "Do Defaults Save Lives?" *Science.* Vol. 302. pp. 1338-1339.

Just, D. and Price, J. 2013. "Default options, incentives and food choices: evidence from elementary-school children". *Public Health Nutrition* Vol.16(12). pp.2281-8.

Gine X, D Karlan, J Zinman. 2010. "Put Your Money Where Your Butt Is: A Commitment Contract for Smoking Cessation", *American Economic Journal: Applied Economics*, 2(4):213-235.

Hanks AS, Just DR, Smith LE, et al. Healthy convenience: nudging students toward healthier choices in the lunchroom. J Public Health. 2012;34:370–376.

Kessler, Holly. 2016. "Simple interventions to improve healthy eating behaviors in the school cafeteria." *Nutrition Review*. Vol. 74(3). pp. 198–209.

Laureati, Monica, Valentina Bergamaschi, Ella Pagliarini. 2014. "School-based intervention with children. Peer-modeling, reward and repeated exposure reduce food neophobia and increase liking of fruits and vegetables." *Appetite*. Vol. 83. pp. 26-32.

List, J.A. and S. Samek. 2015. "The behavioralist as nutritionist: leveraging behavioral economics to improve child food choice and consumption" *Journal of Health Economics*, 39, pp. 135-146

Levitt, Steven D., John A. List, Susanne Neckermann and Sally Sadoff. 2016. "The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance" *American Economic Journal: Economic Policy* Vol. 8, No. 4, pp. 183-219

Loewenstein, G., J. Price and K. Volpp. 2016. "Habit formation in children: evidence from incentives for healthy eating" *Journal of Health Economics*, 45, pp. 47-54

MacKinnon, J. G., Webb, M. D. 2018. "The wild bootstrap for few (treated) clusters." Econometrics Journal 21: 114–135.

MacKinnon, J. G., Webb, M. D. 2017a. "Pitfalls when estimating treatment effects using clustered data." Political Methodologist 24: 20–31.

Marteau, Theresa M; Ogilvie, David; Roland, Martin; Suhrcke, Marc; Kelly, Michael P. 2011. "Judging nudging: can nudging improve population health?" *BMJ : British Medical Journal (Online)*; London Vol. 342.

Miller, GF. S. Gupta, JD Kropp, KA Grogan, A Mathews. 2016. "The effects of pre-ordering and behavioral nudges on National School Lunch Program participants' food item selection" Journal of Economic Psychology. Vol.55 pp 4-16

Ozturk, Orgul, Melayne M. McInnes, Christine Blake, Edward A. Frongillo and Sonya Jones. 2016. "Development of a structured observational method for the systematic assessment of school food choice architecture" *Ecology of Food and Nutrition.* Vol.55 No.2 pp 119-140.

Perry C.L., Bishop D.B., Taylor G.L., Davis M., Story M., Gray C., Bishop S.C., Mays R.A., Lytle L.A., Harnack L. 2004. "A randomized school trial of environmental strategies to encourage fruit and vegetable consumption among children." *Health Education and Behavior*. Vol.31(1). pp.65-76.

Reicks M., Redden J.P., Mann T, Mykerezi E., Vickers Z. 2012. "Photographs in lunch tray compartments and vegetable consumption among children in elementary school cafeterias." *JAMA*. Vol.307(8). pp. 784-5.

Roodman, D., Nielsen, M. Ø., MacKinnon, J. G., & Webb, M. D. (2019). Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal*, *19*(1), 4–60.

Roodman, David. 2015. "BOOTTEST: Stata module to provide fast execution of the wild bootstrap with null imposed," Statistical Software Components S458121, Boston College Department of Economics, revised 05 Sep 2019.

Schwartz, M. B. 2007. "The influence of a verbal prompt on school lunch fruit consumption: a pilot study." *International Journal of Behavioral Nutrition and Physical Activity* Vol.4: 6

Skinner, Jean D., Betty Ruth Carruth, Wendy Bounds, and Paula J. Ziegler. 2002. "Children's Food Preferences: A Longitudinal Analysis." *Journal of the American Dietetic Association*. Vol. 102(11). pp. 1638-1647.

Smith, Tosha. 2012. "Taste Texting: A Pre-Order System for Fresh and Healthy High School Lunch." Dissertation. Dept. of Nutrition, University of North Carolina –Chapel Hill

Thaler, Richard H. and Sunstein, Cass R. 2008. Nudge: Improving decisions about health, wealth and happiness. Yale University Press.

Thomas, Manoj, Kalpesh Kaushik Desai, and Satheeshkumar Seenivasan. "How Credit Card Payments Increase Unhealthy Food Purchases: Visceral Regulation of Vices." Journal of Consumer Research 38, no. 1 (2011): 126-39.

Thomson, Cynthia A. and Jennifer Ravia. 2011 "A Systematic Review of Behavioral Interventions to Promote Intake of Fruit and Vegetables" Journal of American Dietary Association. 111. pp.1523-1535

Viswanathan M, Golin CE, Jones CD, et al. 2012. Interventions to Improve Adherence to Self-administered Medications for Chronic Diseases in the United States: A Systematic Review. Ann Intern Med. 157:785–795. doi: https://doi.org/10.7326/0003-4819-157-11-201212040-00538

Volpp KG, D.A. Asch, R. Galvin R, G. Loewenstein. 2011. "Redesigning employee health incentives – lessons from behavioral economics". N Engl J Med. 4;365(5):388-90. doi: 10.1056/NEJMp1105966.

Wardle, J., M. L. Herrera, L. Cooke, and E.L. Gibson. 2003. "Modifying children's food preferences: the effects of exposure and reward on acceptance of an unfamiliar vegetable." *European Journal of Clinical Nutrition*. Vol.57:341.

Wisdom J., J.S. Downs, G. Loewenstein. 2010. "Promoting healthy choices: information versus convenience" *American Economic Journal: Applied Economics*. Vol.2. pp. 164-178

Appendix A: Algorithm for Choosing Foods to Target

Entrées received a point if they are:

<600 kcal,

<1000 mg sodium,

<10% sat fat

> 6 g fiber or whole grain

> 10 g protein,

< 50 mg cholesterol, and

> 50 mg calcium.

Giving us 7 possible points for each entrée. Total score is calculated for each entrée and highest scoring one is chosen. In event of a tie, food with the lower calories is chosen.

We targeted 2 entrées a week, 2 sandwiches, and 1 salad. We chose the highest scoring hot entrées first, then picked the sandwiches, then the salad.

For fruit we always chose the fresh one (over canned). For vegetables, we chose the one with fewer calories (for salads this include the caloric content of the dressing packet). We made sure to not pick a salad entrée and a salad side on the same day.

Appendix B: Treatment Materials and Treatment Schedule

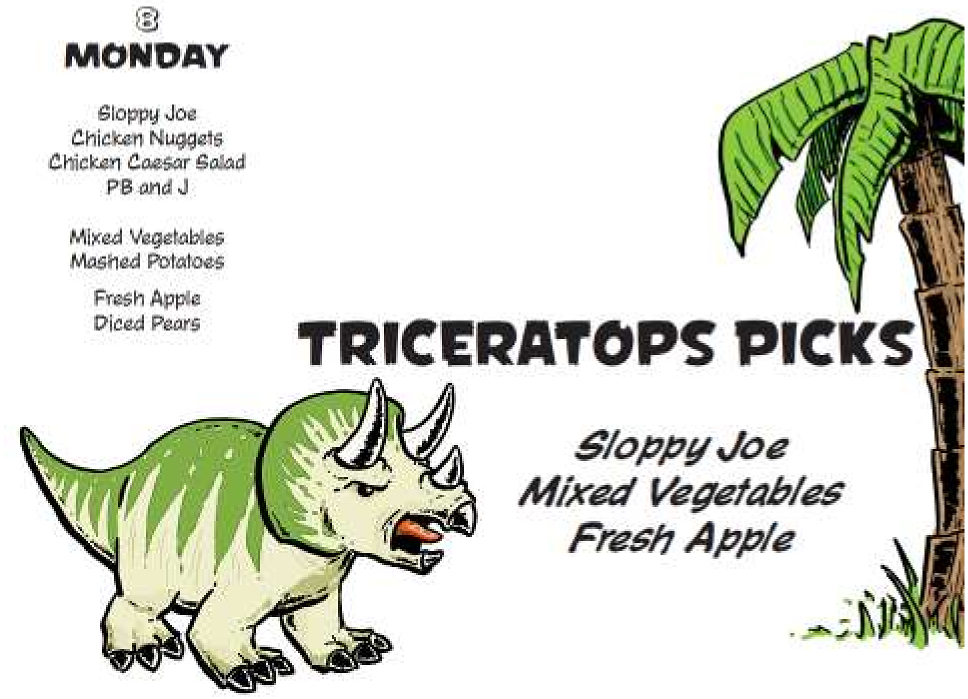
Appendix Table B1: Intervention Timeline and Schedule for Treated Schools

Weeks	1 to 5	6	7	8	9	10	11	12	13	14	15 to 16
Days	(1 to 26)	(27-31)	32-36	37-41	42-45	46-50	51-54	55-59	60-61	62-66	67-78
Intervention											
Theme											
Dinosours											
Detective											
Emphasis											
Creative Naming											
Highlighting											

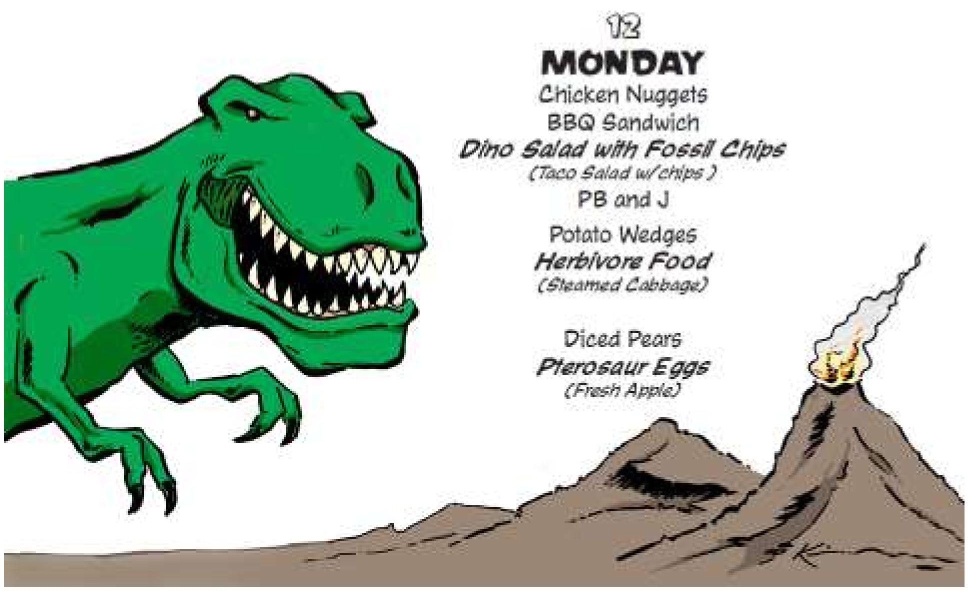
Materials:

Samples from Morning Slide Shows and Menus

Morning Slide Show. Dinosaur Theme (highlight)



Morning Slide Show. Dinosaur Theme (names)



Morning Slide Show. Mystery Detective Theme (names)

Morning Slide Show. Mystery Detective Theme (highlight)

Sample menu in the control schools:

KID'S WAY CAFE by sodexo

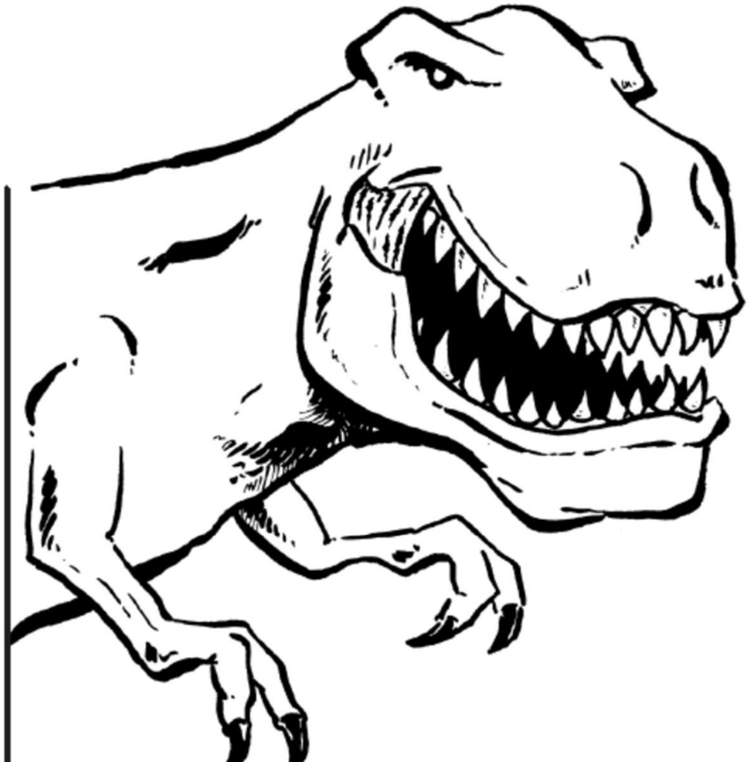
SEP 2012

MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY
3 Chicken Nuggets, BBQ Sandwich, Taco Salad with Chips, PB&J Sandwich Mashed Potatoes, Green Beans, Fresh Apple, Diced Pears, Coleslaw Milk	**4** Baked Ziti, Cheeseburger, Chef Salad, Cheese Sandwich Potato Wedges, Sweet Peas, Fresh Banana, Apple Sauce, Garden Salad Milk	**5** Beef Tacos, Hot Ham and Cheese, Minnie Mouse Salad, PB &J Sandwich Black Beans & Rice, Mixed Vegetables, Fresh Pear, Diced Peaches, Broccoli Salad Milk	**6** Soup with 1/2 Grilled Cheese, Chicken Sandwich, Chef Salad, Cheese Sandwich Candied Yams, Collard Greens, Oraneg Wedges, Apple Sauce, Garden Salad Milk	**7** Cheese Pizza, Corn Dog, Peppi Salad, PB&J Sandwich Tator Tots, Golden Corn, Fresh Apple, Diced Pears, Cherry Tomatoes Milk
10 Chicken Nuggets, Sloppy Joe, Taco Salad with Chips, PB&J Sandwich Rice and Gravy, Steamed Broccoli, Fresh Apple, Diced Pears, Pasta Salad Milk	**11** Chili Macaroni, Chicken Sandwich, Chef Salad, Cheese Sandwich Tator Tots, Green Beans, Fresh Banana, Apple Sauce, Garden Salad Milk	**12** Soft Chicken Taco, Hot Dog, Minnie Mouse Salad, PB&J Sandwich Spanish Brown Rice, Pinto Beans, Fresh Pear, Diced Peaches, Baby Carrots Milk	**13** Lasagna, Hot Turkey & Cheese, Chef Salad, Veggie Wrap Lima Beans, Steamed Spinach, Orange Wedges, Apple Sauce, Garden Salad Milk	**14** Cheese Pizza, Hamburger, Peppi Salad, PB&J Sandwich Potato Wedges, Glazed Carrots, Fresh Apple, Diced Pears, Cucumber Salad Milk
17 Chicken Nuggets, BBQ Sandwich, Taco Salad with Chips, PB&J Sandwich Mashed Potatoes, Steamed Cabbage, Fresh Apple, Diced Pears, Garden Salad Milk	**18** Baked Spaghetti, Cheeseburger, Chef Salad, Cheese Sandwich Potato Wedges, Sweet Peas, Fresh Banana, Apple Sauce, Garden Salad Milk	**19** Nachos and Beef, Hot Ham & Cheese, Minnie Mouse Salad, PB&J Sandwich Black Beans & Rice, Mixed Vegetables, Fresh Pear, Diced Peaches, Broccoli Salad Milk	**20** Shredded BBQ, Cheeseburger, Chef Salad, Chicken Salad Wrap Fresh Fruit, Chilled Fruit, Candied Yams, Lima Beans, Cole Slaw Milk	**21** Cheese Pizza, Corn Dog, Peppie Salad, PB&J Sandwich Fresh Fruit, Chilled Fruit, Tator Tots, Steamed Cabbage, Garden Salad Milk
24 Chicken Nuggets, Sloppy Joe, Chef Salad, PB&J Sandwich Rice with Gravy, Steamed Broccoli, Fresh Apple, Diced Pears, Pasta Salad Milk	**25** Macaroni and Cheese, Chicken Sandwich, Chef Salad, Cheese Sandwich Tator Tots, Green Beans, Fresh Banana, Apple Sauce, Garden Salad Milk	**26** Beef Tacos, Hot Dog, Minnie Mouse Salad, PB&J Sandwich Pinto Beans, Spanish Rice, Fresh Pear, Diced Pears, Baby Carrots Milk	**27** Breakfast for Lunch, Hot Turkey and Cheese, Chef Salad, Cheese Sandwich Black-eyed Peas, Steamed Spinach, Orange Wedges, Apple Sauce, Garden Salad Milk	**28** Cheese Pizza, Hamburger, Peppi Salad, PB&J Sandwich Potato Wedges, Glazed Carrots, Fresh Apple, Diced Pears, Sliced Cucumbers Milk
				

sodexo

KID'S WAY CAFE BY SODEXO — RICHLAND SCHOOL DISTRICT 2

OCTOBER 2012 LUNCH MENU



1 MONDAY

Dino Salad with Fossil Chips
(Taco Salad with Chips)
Chicken Nuggets
BBQ Sandwich
PB and J

Herbivore Food
(Steamed Cabbage)
Potato Wedges

Pterosaur Eggs
(Fresh Apple)
Diced Pears

2 TUESDAY

Pasta-raptor'n Broccoli
(Broccoli Chicken Alfredo)
Cheeseburger
Chef Salad
Cheese Sandwich

Dinosaur Sticks
(Celery Sticks)
Sweet Peas

Triceratops Horn
(Banana)
Apple Sauce

3 WEDNESDAY

T-Rex Tacos
(Beef Tacos)
Hot Ham and Cheese
Minnie Mouse Salad
PB and J

Baby Carrotops
(Baby Carrots)
Pinto Beans

Alxasaurus Eggs
(Fresh Pear)
Diced Peaches

4 THURSDAY

Dino-mite Sandwich
(Chicken Sandwich)
Roast Turkey w/rice
Chef Salad
Cheese Sandwich

Triceratops Frills
(Collard Greens)
Candied Yams

Stegosaurus Spikes
(Orange Wedges)
Apple Sauce

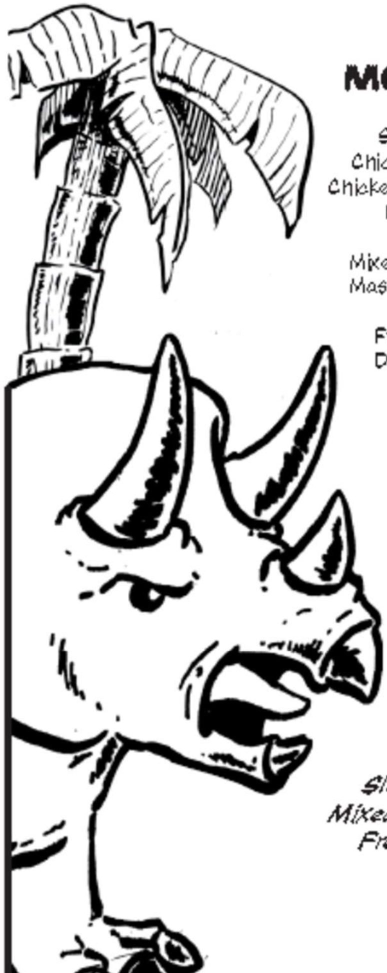
5 FRIDAY

PB-and-J-asaurus
(PB&J Sandwich)
Cheese Pizza
Corn Dog
Peppi Salad

Allosaurus Corn
(Golden Corn)
Tator Tots

Pterosaur eggs
(Fresh Apples)
Diced Pears

OCTOBER 2012 LUNCH MENU



8 MONDAY	9 TUESDAY	10 WEDNESDAY	11 THURSDAY	12 FRIDAY
Sloppy Joe Chicken Nuggets Chicken Caesar Salad PB and J	Chili Macaroni Chicken Sandwich Chef Salad Cheese Sandwich	PB & J Soft Chicken Tacos Hot Dog Minnie Mouse Salad	Hot Turkey & Cheese BBQ Chicken w/rice Chef Salad Cheese Sandwich	PB & J Cheese Pizza Hamburger Peppi Salad
Mixed Vegetables Mashed Potatoes	Cherry Tomatoes Steamed Broccoli	Steamed Spinach Black Beans	Sweet Potatoes Golden Corn	Glazed Carrots Potato Wedges
Fresh Apple Diced Pears	Fresh Banana Apple Sauce	Fresh Pear Diced Peaches	Orange Wedges Apple Sauce	Fresh Apple Diced Pears

From Today's Menu
TRICERATOPS PICKS

Sloppy Joe Mixed Vegetables Fresh Apple	*Chili Macaroni Cherry Tomatoes Fresh Banana*	*PB & J Steamed Spinach Fresh Pear*	*Hot Turkey & Cheese Sweet Potatoes Orange Wedges*	*PB & J Glazed Carrots Fresh Apple*

KID'S WAY CAFE BY SODEXO **RICHLAND SCHOOL DISTRICT 2**

OCTOBER 2012 LUNCH MENU

15 MONDAY	16 TUESDAY *Half Day*	17 WEDNESDAY	18 THURSDAY	19 FRIDAY
Chicken Nuggets BBQ Sandwich Taco Salad w/chips PB & J	Corn Dog Ham and Cheese Sandwich	Nachos & Beef Hot Ham and Cheese Minnie Mouse Salad PB & J	Sliced Ham w/roll Chicken Sandwich Chef Salad Cheese Sandwich	Cheese Pizza Corn Dog Peppi Salad PB & J
Potato Wedges Steamed Cabbage Potato Wedges Diced Pears Fresh Apple	Baby Carrots Fresh Apple Milk	Red Beans Mixed Vegetables Diced Peaches Fresh Pear	Candied Yams Collard Greens Apple Sauce Orange Wedges	Tator Tots Golden Corn Diced Pears Fresh Apple
Taco Salad w/chips Steamed Cabbage Fresh Apple	Corn Dog Baby Carrots Fresh Apple	Nachos & Beef Mixed Vegetables Fresh Pear	Chicken Sandwich Collard Greens Orange Wedges	PB and J Golden Corn Fresh Apple

Appendix Tables C: School Level Controls

Appendix Table C1: Demographic Characteristics for Treatment and Control

Site	Number of Students	Free and Reduced	Female	African American	Asian	Hispanic	White	Pair	Treated
9	770	11.17	0.53	0.32	0.06	0.05	0.56	1	1
16	598	18.06	0.49	0.28	0.04	0.02	0.65	1	0
2	560	24.46	0.47	0.4	0.04	0.03	0.53	2	0
1	606	33.33	0.47	0.36	0.02	0.03	0.58	2	1
13	**588**	**45.41**	**0.47**	**0.48**	**0.1**	**0.18**	**0.24**	**3**	**0**
12	725	46.07	0.51	0.71	0.03	0.05	0.21	3	1
17	**671**	**48.14**	**0.47**	**0.75**	**0.03**	**0.06**	**0.16**	**4**	**0**
4	**428**	**49.53**	**0.46**	**0.45**	**0.04**	**0.19**	**0.31**	**4**	**1**
15	**771**	**51.36**	**0.49**	**0.74**	**0.04**	**0.06**	**0.16**	**5**	**1**
14	661	51.74	0.49	0.52	0.05	0.1	0.32	5	0
10	489	52.15	0.47	0.47	0.01	0.05	0.47	6	1
11	**559**	**52.95**	**0.48**	**0.71**	**0.03**	**0.05**	**0.21**	**6**	**0**
3	573	59.34	0.49	0.69	0.04	0.09	0.18	7	1
6	**578**	**61.42**	**0.48**	**0.7**	**0.03**	**0.07**	**0.2**	**7**	**0**
8	**631**	**67.35**	**0.49**	**0.85**	**0.02**	**0.07**	**0.07**	**8**	**0**
18	612	73.2	0.48	0.69	0.01	0.19	0.1	8	1
5	726	82.64	0.54	0.72	0.02	0.22	0.04	9	1
7	**635**	**87.24**	**0.53**	**0.92**	**0.01**	**0.06**	**0.01**	**9**	**0**

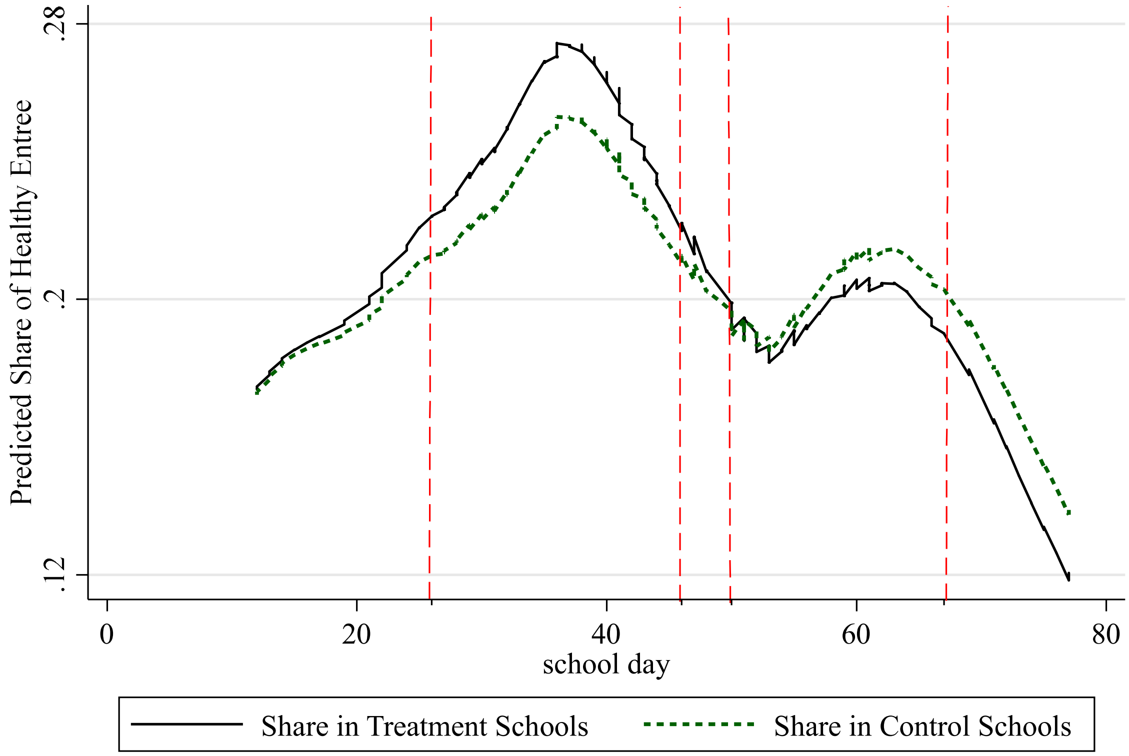
Notes: Column *Free and Reduced* reports the share of students who qualify for free and reduced price lunch. *Pair* shows the schools paired as control and treatment. Schools in each pair has similar demographic and socioeconomic characteristics. *Treated* indicates the school that is the treatment school in each pair. Treatment was randomly assigned in each pair. Red colored schools do not have enough data. Yellows have fewer data from pre period but are included in the analysis. Blue site never reported.

Appendix D: Additional Tables and Figures

Appendix Table D1: Treatment Effects - Regression Estimates

Ever-treated entrée	-0.113	-0.113	-0.113	-0.113	-0.053	-0.053	-0.053
cluster robust std error	0.007	0.007	0.007	0.007	0.020	0.020	0.020
wild bootstrap p-value	0.005	0.005	0.005	0.005	0.017	0.018	0.018
subcluster wild bootsrap p-value	0.005	0.006	0.005	0.005	0.008	0.008	0.009
Treated school	-0.008	-0.012	-0.007	-0.003	0.000	-0.004	0.006
cluster robust std error	0.005	0.005	0.004	0.004	0.006	0.006	0.005
wild bootstrap p-value	0.213	0.071	0.137	0.429	0.981	0.582	0.381
subcluster wild bootsrap p-value	0.219	0.062	0.283	0.601	0.978	0.578	0.547
Treatment period	-0.035	-0.035	-0.036	-0.036	0.009	0.009	0.009
cluster robust std error	0.004	0.004	0.004	0.004	0.004	0.004	0.004
wild bootstrap p-value	0.014	0.013	0.013	0.013	0.067	0.069	0.078
subcluster wild bootsrap p-value	0.004	0.001	0.002	0.002	0.109	0.090	0.088
Ever-treated entrée*Treated school	0.013	0.013	0.013	0.013	0.001	0.001	0.001
cluster robust std error	0.010	0.010	0.010	0.010	0.010	0.010	0.010
wild bootstrap p-value	0.256	0.258	0.256	0.256	0.931	0.914	0.914
subcluster wild bootsrap p-value	0.268	0.267	0.267	0.267	0.934	0.911	0.911
Treated school*Treatment period	-0.004	-0.004	-0.004	-0.004	-0.007	-0.008	-0.007
cluster robust std error	0.005	0.005	0.005	0.005	0.006	0.006	0.006
wild bootstrap p-value	0.459	0.452	0.497	0.497	0.216	0.200	0.226
subcluster wild bootsrap p-value	0.472	0.458	0.494	0.494	0.236	0.221	0.237
Ever-treated entrée*Treatment period	0.099	0.100	0.100	0.100	0.001	0.001	0.001
cluster robust std error	0.009	0.010	0.010	0.010	0.007	0.007	0.007
wild bootstrap p-value	0.014	0.014	0.014	0.014	0.926	0.901	0.897
subcluster wild bootsrap p-value	0.001	0.001	0.001	0.001	0.929	0.902	0.899
Ever-treated entrée *Treated	0.026	0.026	0.026	0.026	0.030	0.029	0.029
cluster robust std error	0.012	0.012	0.013	0.013	0.012	0.012	0.012
wild bootstrap p-value	0.082	0.084	0.084	0.084	0.034	0.036	0.036
subcluster wild bootsrap p-value	0.067	0.069	0.070	0.070	0.027	0.028	0.028
School characteristics	NO	YES	NO	YES	NO	YES	YES
Teacher fixed effects	NO	NO	YES	YES	NO	NO	YES
Food fixed effects	NO	NO	NO	NO	YES	YES	YES
Observations	39,033	39,033	39,033	39,033	39,033	39,033	39,033
R-squared	0.014	0.014	0.014	0.014	0.638	0.638	0.639

Appendix Figure D1: Predicted Healthy Entree Shares



Appendix Table D2: Effect of Treatment on Lunch Participation

	Classroom Level Participation		School Level Participation
Treatment period	0.01	0.008	-0.011
	[0.007]	[0.010]	[0.009]
Treatment school	0.033	0.009	0.109**
	[0.027]	[0.007]	[0.023]
Treatment school x Treatment period	0.011	0.011	0.019
	[0.010]	[0.012]	[0.012]
Observations	37,478	37,478	312
R-squared	0.167	0.601	0.8586
School Characteristics	YES	YES	YES
Teacher Fixed Effects	NO	YES	NO

Notes: School level participation is calculated as the ratio of total count of entrees to the number of students enrolled in a school. We know the school level enrollment but not classroom level enrollment. In order to calculate class level paticipation we took the ratio of count of entrees in a given day to the maximum count of entrees ever observed in that classroom through our study period.

Appendix Table D3: Unconditional Triple Differences - Only When Treated

		Non-treated Entrees	Treated Entrees	*Difference*
treated schools	**Before**	0.292	0.191	-0.100
		(0.004)	(0.004)	(0.005)
		[6,855]	[4,215]	[11,070]
	During -Treated entrée only	0.250	0.303	0.054
		(0.003)	(0.006)	(0.006)
		[10,039]	[3,468]	[13,507]
	Change over time (Treated Entrée only)	*-0.042*	*0.112*	*$DD_{TSwt=}$ 0.176*
		(0.005)	*(0.007)*	*(0.008)*
control schools	**Before**	0.300	0.187	-0.113
		(0.005)	(0.006)	(0.008)
		[3,437]	[2,040]	[5,477]
	During -Treated entrée only	0.260	0.255	-0.005
		(0.004)	(0.006)	(0.007)
		[6,679]	[2,300]	[8,979]
	Change over time (Treated Entrée only)	*-0.040*	*0.069*	*DD_{NTS} =0.127*
		(0.008)	*(0.012)*	*(0.012)*
$DDD_{wt} = DD_{TSwt} - DD_{NTSwt}$				***0.045***
				(0.015)

Appendix Table D4: Persistence of Treatment Effects of Nudges- Spillover Treatment Effects

Ever-treated entrée	-0.113**	-0.113**	-0.113**	-0.113**	-0.067**	-0.068**	-0.068**
	[0.005]	[0.006]	[0.005]	[0.005]	[0.002]	[0.002]	[0.002]
Treated school	-0.008	-0.012+	-0.007	-0.003	-0.000	-0.004	0.006
	[0.219]	[0.066]	[0.509]	[0.783]	[0.978]	[0.576]	[0.547]
Treatment period	-0.035**	-0.035**	-0.036**	-0.036*	0.009	0.009+	0.009+
	[0.004]	[0.001]	[0.002]	[0.016]	[0.109]	[0.090]	[0.088]
Treated school x Treatment period	-0.004	-0.004	-0.004	-0.004	-0.007	-0.008	-0.007
	[0.472]	[0.461]	[0.509]	[0.509]	[0.236]	[0.220]	[0.237]
Ever-treated entrée x Treated school	0.013	0.013	0.013	0.013	0.001	0.002	0.002
	[0.268]	[0.267]	[0.267]	[0.267]	[0.887]	[0.871]	[0.871]
Ever-treated entrée x Treatment period	0.003	0.003	0.003	0.003	-0.006	-0.005	-0.005
	[0.714]	[0.708]	[0.707]	[0.707]	[0.477]	[0.502]	[0.510]
Ever-treated entrée x Treated school x	-0.015	-0.015	-0.015	-0.015	-0.011	-0.012	-0.012
Treatment period	[0.632]	[0.630]	[0.629]	[0.629]	[0.666]	[0.655]	[0.653]
Ever-treated entrée x After the first	0.119*	0.119*	0.119*	0.119*	-0.016*	-0.016*	-0.017*
treatment	[0.014]	[0.015]	[0.014]	[0.014]	[0.035]	[0.037]	[0.036]
Ever-treated entrée x After the first	0.015	0.015	0.015	0.015	0.022	0.022	0.022
treatment x Treated school	[0.448]	[0.458]	[0.452]	[0.452]	[0.222]	[0.232]	[0.228]
Treated entrée	-0.021+	-0.021+	-0.021+	-0.021+	0.035*	0.035*	0.035*
	[0.055]	[0.051]	[0.052]	[0.052]	[0.043]	[0.045]	[0.045]
Treated entrée x Treated school	0.045**	0.045**	0.045**	0.045**	0.033*	0.034*	0.034*
	[0.009]	[0.008]	[0.009]	[0.009]	[0.048]	[0.046]	[0.047]
School characteristics	NO	YES	NO	YES	NO	YES	YES
Teacher fixed effects	NO	NO	YES	YES	NO	NO	YES
Entrée fixed effects	NO	NO	NO	NO	YES	YES	YES
Observations	39,033	39,033	39,033	39,033	39,033	39,033	39,033
R-squared	0.019	0.019	0.02	0.02	0.640	0.64	0.641

Notes: Coefficients are estimated using OLS regressions. Subcluster wild bootstrap p-values (reported in brackets) are obtained using *boottest* command (Roodman, 2015) in Stata 15 using 5000 replications with Webb weights. Share of an entrée in total number of entrées ordered by a class is the outcome variable. *Ever-treated entrée* is an entree that is at some point designated as healthy entree of the day during treatment period and treated school indicates if the school had any treatment. *Treated entree* indicated the entree designated as the nudged entree on a given day. *After the first treatment* is an indicator for an ever-treated entree for all days when it is not the treated entree after it is treated at least once. **,*, + indicate significance at 1%, 5% and 10% respectively using subcluster wild bootstrap p-values.