Changing Dietary Habits Early in Life: A Field Experiment with Low Income Families

Michèle Belot*, Noémi Berlin†, Jonathan James‡ and Valeria Skafida§

January 19, 2022

Abstract

We evaluate the malleability of dietary habits early on in life with two interventions targeting low income families with very young children. Both interventions were conducted over a three-month period. Families were asked to either prepare healthy meals with ingredients delivered to their door or to avoid snacking and eat at regular times. Families in the control group could carry on as usual. We find that both interventions led to a fall in children's body mass index, an effect that persisted for three years for the first intervention, but faded away after two years for the second. Using a range of measures of dietary intake, we find that children reduced their consumption of food with added sugar. We do not find evidence that the interventions shifted children's tastes towards healthier foods however.

JEL Classification: I12, I14, I18
Keywords: Dietary Habits, Early Life, Healthy eating, Field Experiments, Obesity

Acknowledgments: This project was funded by the European Union's Seventh Framework Programme, under grant agreement no. 607310. The experiment is registered under trial no. AEARCTR-0003281. We also thank local authorities, early life centres, and schools in Edinburgh and Colchester; Ilenia Paciarotti, Hazel Blakemore and Gaja Lescinar for their help with the design and implementation of the interventions, Philipp Kircher, Andrew Kelloe, Lise van Wijngaarden, Kritsina Enger, Hasna Fadhilah, Dragos Gruia, Yvonne Laird, Agnès Charpin, Chiara Serra, Zoey Verdun and Miguel Alquezar for their help with assistance during the study, ISER and the EssexLab for allowing us to use their facilities, and nutritionists from Alison Clark Health and Nutrition. We thank Gareth Leng and John Menzies for their help with the blood samples collection. We thank Joe Doyle, Gareth Leng, John Menzies and Philipp Kircher for very useful comments. We also thank the participants of the seminars at Aberdeen, Amsterdam, ALISS, Bath, Bolzano, Copenhagen, Essex, GATE (Lyon), Glasgow, Gothenburg, INRA, LIRAES, Lucerne, Mainz, Nottigham, Paris 5, Paris School of Economics, Stellenbosch, Stirling, UCL, Utrecht, Hamburg; and participants to the ESRC workshop in Health Economics 2018, Hamburg workshop in Risky Behaviors 2019, ESA 2016, ESPE 2018, EALE 2018, Royal Economic Association 2019, Australian Health and Human Capital Seminar Series 2021. The paper was also circulated with the title "The Formation and Malleability of Dietary Habits: A Field Experiment with Low Income Families"

*Corresponding author: mb2693@cornell.edu, Department of Economics, Cornell University, 109 Tower Road 404 Uris Hall, Ithaca, New York 14853-2501

†noemi.berlin@parisnanterre.fr, CNRS, EconomiX, University of Paris Nanterre

‡jj412@bath.ac.uk, University of Bath

§valeria.skafida@ed.ac.uk, University of Edinburgh

1 Introduction

Poor diet is a major issue in most developed and developing countries. While there are many policies targeting diet such as prices and information campaigns[1] and, more recently, a series of interventions based on insights from behavioural economics, the evidence on medium or longer term effects of interventions remains limited.

The most promising interventions with a potential for long term effects appear to be those targeting children, because dietary habits are believed to form early in life (Birch, 1999). Also, recent research in biology has shown that the number of adipocytes ("fat cells") - a strong predictor of obesity in adulthood - is set during childhood and early adolescence (Spalding et al., 2008).

In this study, we propose and evaluate two comprehensive interventions, each lasting for three months, aimed at changing eating habits. The interventions target families with young children (below the age of six) and treat the entire family. We follow them over the course of three consecutive years. The first intervention (*Meal treatment*) consists of providing families with ingredients and recipes directly at home to cook five meals a week over a period of 12 weeks. This intervention is of greater intensity than most large scale policy interventions, and can therefore give an idea of what is achievable.[2] The second intervention (*Snack treatment*) targets snacking and erratic eating habits. Again for a period of 12 weeks, families are instructed to eat three meals a day, at regular times, avoiding eating (snacking) between meals.

Our sample consists of 285 low income families located in two different locations in the UK – Colchester (England) and Edinburgh (Scotland). We focus on low income families because of the well documented evidence of a strong socio-economic gradient in chronic diseases and in obesity (Dalstra et al., 2005). We collected different measures to evaluate the impact of the interventions: some objective (such as BMI), others based on incentivized choices and finally, methods used in nutrition research based on in-depth surveys of participants. We collect information before, immediately after, and once a year for 3 years after the interventions.

We find evidence of significant sustained changes in the children's BMI (z-score) following

[1] Capacci et al. (2012) and Frisvold and Lansing (2021) propose recent reviews of relevant studies.
[2] It fits to the spirit of Ludwig et al. (2011)'s suggestion to study the problem of "food deserts".

the meal treatment. Children in both treatment groups appear to have moved down in the distribution of BMI. These effects are large (between 0.2 and 0.36 standard deviations) and are sustained in the longer run for the Meal treatment, while they attenuate over time for the Snack treatment and are no longer significant in the third year. Parallel to these changes in BMI, we find evidence of significant changes in diet (measured with different methods). The changes are more sustained in the Meal treatment, in line with the results on BMI. Notably, we find consistent evidence of a reduction in the intake of added sugar of treated children, more pronounced in the Meal treatment than in the Snack treatment. Using an incentivized measure of dietary choices, we find that the parents of children in the Meal treatment purchase less sugary foods post intervention.

Notably, we find no evidence that the interventions altered preferences in favor of healthier foods. In fact, measures of food preferences indicate that children in the treated groups seem to like sweets more after the intervention, which goes against predictions of habit formation models (Becker and Murphy, 1988). The most plausible mechanism for the effects we observe is that parents reduced access to unhealthy foods, in particular sweets, and that this was successful in limiting weight gain.

Our primary focus in this study is children and the paper focuses on them. But we note that for adults, we find little evidence for malleability of dietary habits. Neither subjective nor objective measures show any evidence of improvement in the healthiness of their diet or their preferences for healthier foods, in the short or longer run. Overall, our results confirm that interventions later on in life are much more challenging.

Our study makes several contributions to the existing literature (discussed in Section 2). First, the study focuses on families with very young children (aged between 2 and 6 at the beginning of the study). The interventions also target the home-setting rather than schools, and the parents are treated along with their children. To our knowledge, this is the first experimental evaluation of interventions targeting children and their parents at home.[3]

Second, we evaluate the effects on the whole diet rather than on isolated one-shot purchase/eating decisions. Most studies rely on partial measures of dietary choices, such as

[3]Charness et al. (2019) find in a school-based experiment that involving the parents is key to the long-term success of an intervention aiming at encouraging healthy choices at school. The interventions tested are one-shot, they show that using a grading system for informing children about the healthiness of foods is effective in encouraging healthy choices, with effects carrying on weeks after the information was given.

consumption of specific items. Downs and Loewenstein (2012) identify this as a key shortcoming of existing studies, writing that "the true success of such measures will remain unclear until researchers are able to measure an individual's total food intake – not only calories at a single meal or in a single episode of snacking."

Third, we evaluate effects on a much longer horizon than other studies. We follow the families for *three* consecutive years. Most studies are typically short run or, if they include a follow-up at a later point in time, the horizon considered is usually shorter than a year.

Fourth, the scope of the interventions tested is unprecedented. Particularly for the Meal treatment, we would argue that it is the most invasive intervention one could probably think of implementing with families over a period of time.

The remainder of the paper is organized as follows: Section 2 reviews the recent relevant literature. Section 3 presents the experimental design, Section 4 describes the main outcome measures and Section 5 presents the analysis and results. In Section 6, we discuss the robustness of the results. Section 7 concludes with a discussion of our findings.

2 Related Literature

The paper fits within the recent literature testing and evaluating interventions aimed at fostering healthier habits. Charness and Gneezy (2009) were the first to show that an intervention of temporary nature (presumably long enough to form new habits) could lead to a sustained change in a lifestyle behavior - in their case, exercise. In their study, students were incentivized to exercise for a period of one month. They find that these were then more likely to exercise even when the incentives were removed.

The idea of using temporary interventions to encourage healthier habits has also found encouraging support in the domain of nutrition. A number of studies (Just and Price, 2013, List and Samek, 2015, Belot et al., 2016, Loewenstein et al., 2016) have shown promising evidence that children's dietary choices can be influenced over sustained periods, but the evidence relates to isolated single eating decisions, and the evidence on long run effects is very limited. Also, these studies do not evaluate the short or long run effects on the overall diet or on other relevant measures such as body mass index (BMI).

There are very few studies focusing on very young children (below the age of 6). A recent study by Griffith et al. (2018) studies the effects of a fruit and vegetables voucher scheme targeting low income families with children below the age of 3 in the UK. They find a significant impact on purchases of fruit and vegetables for eligible families.[4] Daniels et al. (2012) conduct a RCT to evaluate an educational intervention targeting parents of infants and aiming at encouraging healthy feeding practices. They find that the intervention had an impact on the Body Mass Index of the treated children measured 6 months after the intervention. Neither study investigates effects on the longer term though.

More broadly and not targeting children specifically, Cawley et al. (2016) and List et al. (2015) conduct field experiments in collaboration with a supermarket and test different types of interventions targeting the prices of nutritious and less nutritious foods (subsidy, taxes, information) in order to increase the consumption of the former. These experiments show that these interventions affect purchasing behavior (low income families purchased more of both nutritious and less-nutritious food under the subsidy framing in Cawley et al., 2016), and that incentives can lead to sustained changes in the purchase of fruit and vegetables even when the incentives are removed, suggesting habit formation has taken place (List et al., 2015). However, overall spending in the store involved in the experiment was low, suggesting that most other food purchases were taking place elsewhere. Brownback et al. (2019) also targets grocery shoppers (with a low income) and examine the effects of combining food subsidies in combination with other behavioral interventions (giving shoppers greater agency on the choice of subsidies, and introducing a waiting period before the shopping trip). They find that these behavioral interventions substantially increase the effectiveness of food subsidies. While providing encouraging evidence, one challenge in these studies is that they focus on purchasing behavior rather than consumption, they only have a partial picture of what people buy or eat and they do not evaluate effects in the longer term either.

Finally, our paper also contributes to a recent literature on the effects of the quality of food environments one is exposed to. Allcott et al. (2019) use a structural model to simulate the effects of a hypothetical experiment where low-income households would have access to the same products and prices as high-income households. They find that this would reduce

[4]They use as a comparison group families with a pregnant mother or with children slightly older than 3.

nutritional inequality by only about 10 percent. This is because the demand for healthy foods is lower among low-income households. Here we provide experimental evidence where families are actively encouraged to try different foods, and the goal is therefore to evaluate whether it is possible to act on the demand.

3 Experimental Design

The experiment was reviewed and approved by the Ethics Committee at the University of Edinburgh and was pre-registered on the AEA RCT Registry.[5] Full details of the experimental design are provided in the online Appendix Section A.1.

Eligibility and recruitment

Families were required to satisfy the following criteria to participate in the study: have a child aged between 2 and 6 years old at the start of the study; own a fridge and a hob (cooktop); live in Edinburgh or Colchester, have a household income below the median income £26,426 for Scotland, £26,600 for England. Families with severe chronic health issues were excluded.

Families were recruited from areas around Edinburgh (Scotland) and Colchester (England). Various recruiting methods were used such as adverts in public spaces, community centers, nurseries, buses and shopping malls; letters sent to school principals, advertisements on radio, individual letters targeting low income areas. Samples of our recruitment materials (leaflet and poster) are shown in Appendix Figure A.1a and A.1b.

For ethical reasons, participants needed to provide informed consent before being included in the study. To mitigate potential self selection concerns, we were careful in not disclosing our research question and providing only general information about the study - such as the study being related to health and lifestyle choices, and the study duration of 3 years. We did not emphasize the focus on nutrition and dietary habits. Also, participants received a relatively large monetary compensation (between £350 and £400 in total), which should

[5]Details of the registered trials can be found on AEA RCT Registry: Belot, Michèle et al. 2018. "The formation and malleability of dietary habits: A Field Experiment with Low Income families." AEA RCT Registry. October 04. https://doi.org/10.1257/rct.3281-1.0. Pre-registration was not common practice at the start of the data collection, so the experiment was registered after the start of the experiment but before the end of data collection. The trial was however described in detail in the ethical approval application at the University of Edinburgh, which was approved before the start of the experiment. The main deviation from the original plan is in the fractions allocated to each treatment, which we discuss below.

mitigate potential self selection effects. This said, we can of course not completely alleviate all concerns regarding external validity.

Randomization

When registering to take part, families were asked to indicate several dates when they would be available to come to our facilities for the first session of measurements. All initial sessions' dates had been randomly pre-assigned to a specific treatment (control or one of the two treatments), and families were randomly assigned to one of their selected dates, without knowing these corresponded to the two different treatments and/or the control group.

Treatments

Treatment 1 - "Meal Treatment"

The first treatment, referred to as "Meal Treatment", consists of providing (free of charge) ingredients and recipe booklets every week, for twelve weeks, directly at participants' homes for five main meals for the *whole* family.[6].

The protocol has been designed to ensure convenience and limit non-financial costs that could be important obstacles in adopting a healthy diet: Families do not have to plan for these meals, i.e. they do not have to search for suitable recipes, organize the shopping, etc. The food is delivered at home and families receive a weekly booklet of recipes for the ingredients delivered. The recipes have been chosen by a nutritionist for their low cost and simplicity of execution, which ensured that the food families were exposed to would be part of the usual British cuisine.[7] Most of them were borrowed from a government website associated with the largest UK public health campaign around healthy eating at the time ("Change4Life"). This protocol has also been deliberately chosen over stricter protocols that would impose constraints on families on all meals and food consumed to ensure its feasibility.

Also, by providing the food for free, we aimed at maximizing compliance during the 12 weeks of treatment. The costs of the meals have however been calibrated to the average

[6]Families could select between a regular or vegetarian food basket. For a list of recipes by week see Appendix Tables A.2 and A.3

[7]An isocaloric comparison (fixed at 365 calories, the average calories of the meals) between the recommended nutritional guidelines and our recipes shows that our recipes are overall consistent with the recommendations, and are lower than the maximum thresholds on sugar and fat (and saturated fat), compensating for these calories via higher carbohydrate and protein contents. This can be seen in Appendix Table A.12, by comparing the second and the third columns.

weekly budget of low-SES families in the UK so families can afford continuing buying the ingredients and recipes once the interventions are over.[8]

Of course, the subsidizing of food may have spillover effects on other expenditures through income and substitution effects. This is a key reason why collecting measures of the whole diet and of body mass index is important.

Treatment 2 - "Snack Treatment"

The second treatment, referred to as "Snack Treatment", consists of regulating eating patterns and avoiding unhealthy snacking, again for twelve consecutive weeks. Adults in the family were asked to eat three meals per day at regular times (selected by participants) and to avoid consuming any food or calorific drinks between meals. For the children, the protocol included an additional morning and afternoon snack (to be consumed at regular times). The snacks were provided by us and determined by a nutritionist. The list of snacks can be found in Appendix Table A.4.

This intervention draws on a body of evidence in economics and nutrition suggesting a link between snacking and obesity. A review on changes in childhood food consumption patterns by St-Onge et al. (2003) suggests that the rising proportion of calories coming from snack foods, which are in turn associated with higher sugar and fat consumption, may be a contributor to rising overweight and obesity in children. A review paper by Bellisle (2014) suggests that snacking often contributes to the calorie count but little nutrition, especially among obese children and adults. Although snacking is often blamed for rising obesity rates (Cutler et al., 2003), research on the effects of snacking on BMI is not conclusive (Field et al., 2004; Larson and Story, 2013).

A number of studies have also shown an association between meal irregularity and poor dietary outcomes more generally (Laska et al., 2015; Leech et al., 2015), and there may be metabolic advantages to eating at more regular and structured intervals (Alhussain et al.,

[8]Based on ONS household expenditure data for 2015 (see https://tinyurl.com/Familyfooddatasets and https://tinyurl.com/householdexpend), average food expenditure for the 5 first income deciles was £28.81 per person per week (average taking into account both expenditure for adults and children). Our sample has an average household composition of 1.65 adults and 1.9 children per household. Based on this we calculate that our families would spend on average £102.28 per week on all in-house food and non-alcoholic drink expenditure. No data exists to capture the amount of this spent on the evening meal, though we hypothesize that this will be circa one third of this budget. Based on this assumption we calculate that £34.09 per week is spent on 7 evening meals, and thus £24.35 is spent on 5 evening meals per week on average. This final figure was our benchmark for calibrating the cost of the meals.

2016). In the case of adults, people appear more likely to choose healthier foods when they select them in advance than when they select them at the moment when they will be consumed (Read and van Leeuwen, 1998). A related body of literature in biology hypothesizes that irregularity of food intake could have a significant impact on diet and total calories, although this hypothesis is not supported in non-animal experiments (Hume et al., 2016).

Aside from providing healthy snacks for children, families were not instructed to alter the content of other meals. The focus remained on limiting and regulating snacking for the family so as to create a source of exogenous variation across groups, and study how that has an impact on diet. This protocol is of course difficult to enforce, and so the main goal of this intervention is to create a source of exogenous variation in snacking behavior across groups, and study how that has an impact on diet.

Control group

Control group participants were instructed to carry on as usual with their daily routines.

All groups received one-to-one instructions on the protocols during their first visit at our facilities.

Compensation

Families were compensated for each visit at our facilities. The compensation added up to £350 in Edinburgh and £400 in Colchester for all visits, spread over the 6 visits.[9]

Compliance

Compliance was assessed in various ways, such as asking families to provide pictures of the meals prepared and writing feedback information (see Appendix Section A.1.3).

To make sure that families fully understood what was expected from them, we met with each of them (treated and control) one-to-one at baseline and provided face-to-face instructions about the study and the protocols they should follow. Families in both treatment groups were told that we were interested in learning how easy the protocols were to follow and would value feedback on the difficulties they have encountered. In Appendix A.1.3, we present all the details on the compliance measures. Overall our compliance measures show relatively high compliance to the protocols. The Meal protocol appears to have been

[9]The total amount was increased for the Colchester arm of the study to increase sample size.

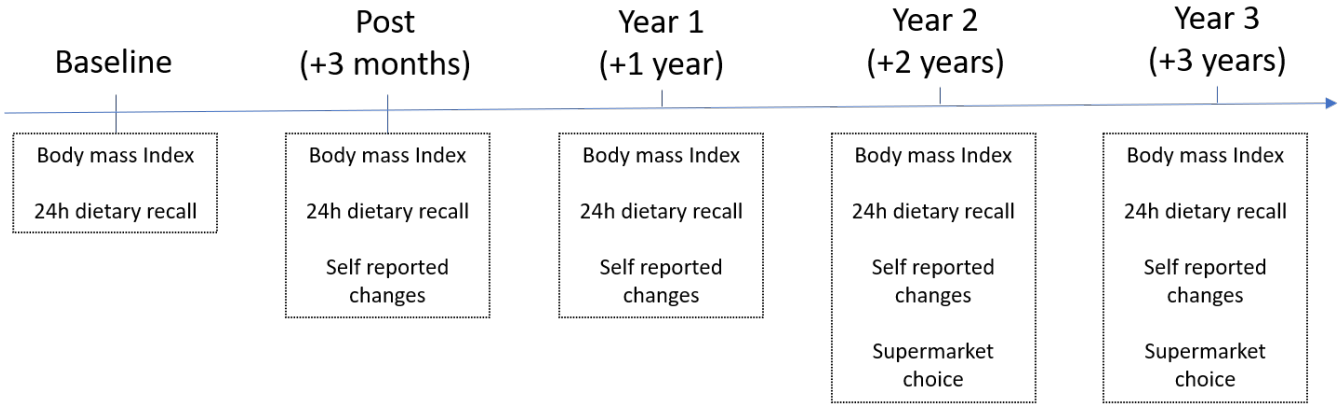
easier to follow compared to the Snack treatment. Of course the estimates we will report are Intention-to-Treat estimates and the compliance tasks are part of the treatments we evaluate.

4 Main outcome measures

We evaluate the effects of the interventions on BMI and dietary habits. Note that both treatments were not designed as weight-loss programmes and did not impose any calorie restriction. Nevertheless, these protocols were expected to have longer term effects on BMI through a change in dietary habits.

We collected measures at five different points in time: At baseline, after completion of the twelve week treatments, as well as one, two and three years after the interventions had taken place (see Figure 1).[10] All measures were collected on site at the Universities of Edinburgh and Essex.

Figure 1: Study timeline



4.1 Anthropometric measurements and BMI

Participants were weighed and measured by a trained member of the team. Height and weight data were used to create age-adjusted BMI for children using BMI cut-offs (based on the percentiles) recommended by the Childhood Obesity Working Group of the International Obesity Taskforce (Vidmar et al. 2004).

[10]The baseline and post experiment measurements were collected during 2-week time windows before and after the twelve week treatment period.

4.2 Measures of Diet

We collected data on diet using three methods. For each method, we focus on variables that are most indicative of the healthiness or unhealthiness of a diet: total calorie intake, fruit and vegetable consumption, saturated fat, and added sugar (otherwise known as Non-Milk Extrinsic Sugar (NMES) or free sugars).[11] Importantly, these should not be considered as outcomes we wish to study independently. As we noted earlier, one major drawback of existing studies is that they usually rely on a single measure of dietary choices, often based on an isolated food consumption event. The goal here is to identify a consistent and robust pattern across these different measures and understand the mechanisms behind the changes observed.

4.2.1 24h dietary recall

The first measure of dietary intake relies on method commonly used in nutrition research, called the "24-hour diet recall". The method consists of asking participants to recall in detail what they have eaten in the last 24 hours. The data was recorded face-to-face with nutritionists at baseline and immediately after the intervention, and using a computer-based British software called Intake24 for the three subsequent years. A parent (most often the mother) provided the information on behalf of the child and for themselves. The data was then entered into a nutritional analysis software (called WISP) which computes measures of dietary intake based on a large database of food items available in the UK. The software then provides direct information on calorie intake and macro-nutrient composition. More information on the method is provided in Appendix Section A.2.1.

One important challenge with our implementation of the 24-hour dietary recall method is that participants could anticipate having to report their dietary intake and could adjust their consumption the day before their visit. Obviously, this challenge applies to all groups and it is unclear whether the risk of bias is greater for treated or control groups.

[11] See a report from the National Diet and Nutrition survey (https://tinyurl.com/NDNS20122014) claiming that diets low in saturated fats and sugars and high in fruit and vegetables are typically recommended for disease.

4.2.2 Self-reported changes

The second measure is based on a computer survey where parents were also asked to report changes in the consumption of specific food items (for their child and for themselves). We then group these items in three categories: fruit and vegetables, foods high in fat and foods high in added sugar. We calculate an index equal to the number of items for which an increase is reported, minus the number of items for which a decrease is reported.

Obviously, this measure has only been collected after the treatments, immediately after the 12 weeks, and then in the 3 subsequent years. Participants were also asked about other lifestyle changes (such as smoking and drinking) and their perceived health.

4.2.3 Supermarket Choice Tool

The third measure is an incentivized measure based on a one-shot post-experimental choice of a supermarket basket worth GBP 30, using the tool developed by Spiteri et al. (2019). Parents are asked to select food items among a choice set of 120 popular items from a supermarket, organized in categories that resemble typical aisles in a supermarket (fruit and vegetables, meat and fish, dairy, etc.). Screenshots of the tasks can be found in Appendix Figure A.5.

Participants are instructed to shop as they normally would. This means that the shopping basket is intended for the whole family rather than for themselves individually. The basket was delivered to 1 in every 10 families selected at random, two weeks after the intervention (to avoid that choices would depend on current food stocks).

The tool allows us to extract information on calories and nutrients such as sugar and saturated fat. We also construct a measure of percentage spent on "unhealthy" items. Unhealthy items are identified through the nutrient profiling technique developed by the UK's Food Standards Agency (FSA).

This task was implemented in years 2 and 3 to avoid contamination with the interventions conducted in the treatment groups.

5 Analysis and Results

5.1 Descriptive characteristics

Our baseline sample consists of 285 families (91 families in Edinburgh and 194 in Colchester).[12] We deliberately assigned a higher proportion of the families to the Control and Meal Treatment groups: 109 families were originally assigned to the Snack treatment, 141 to Control and 146 to Meal. We did so after realizing challenges of recruiting such a large sample. We decided we would drop one treatment arm if we did not succeed in recruiting at least 50 families for each treatment group. To prevent having to cancel the entire study if the sample was to small for two treatment arms, we decided to allocate proportionally more families to one of the treatments (Meal treatment). The show up rate at baseline was above 71% in the Meal group, it was a bit higher in the Snack group with a show up rate of 72%, and it 82% in the control group.

Table 1 presents descriptive statistics of our sample for the different groups.[13] Overall, there are no statistically significant differences between the control and the treatment groups at baseline.

Table A.7 provides descriptive statistics on attendance and attrition. The attrition rate after the twelve-week intervention was very low (3.85%). The overall show-up rate remained high in the subsequent data collection years (87% in 2016, 84% in 2017 and 82% in 2018 relative to the first "baseline" session). However, the attrition rate differs slightly between treatments. There is a slightly higher rate of attrition in the snack treatment at year 3 relative to baseline (20%) compared to both the Meal treatment (11%) and the Control group (16%). We discuss how attrition affects our results in Section 6. The main conclusion is that attrition is unlikely to drive the significant treatment effects we identify.

Table 2 shows the mean BMI (z-score), as well as the percentage of children underweight, overweight or obese. The z-score is a the standard deviation (SD) above or below the mean. We see that the BMI z-score of the snack group is lower but the differences between the groups are not statistically significant. The obesity rate of our sample is in line with national

[12] 377 families originally registered for the study.

[13] Appendix Table A.9 presents a comparison of our sample with representative surveys.

statistics (16.8% of the children and 33% of adults). We do not find significant differences in the distribution of weight categories between the three groups at baseline.

The table also shows baseline measurements for children for different diet intakes (total calorie intake, added sugar, saturated fat and fruit and vegetables) and food preferences. No significant differences are found for those measurements at baseline.

To get a sense of how the meals in the Meal treatment compare to the pre-intervention diet of participants, we compare the nutritional content holding calories constant (benchmark of 365 calories). The average diet of participants contained twice the amount of fat (15g versus 8g) and twice the amount of sugars (20g versus 10g) found in the meals (see Appendix Table A.12). Participants' diets at baseline were lower in carbohydrates and protein relative to the meals.

5.2 ITT estimates of the Impact on BMI and Diet

We now turn to the core results of the paper and present ITT estimates of the effects of the treatments. For measures with baseline information, the estimates are difference-in-differences estimates. For measures with only post-intervention information, the estimates relate to the difference between the treated and control groups. All estimates are from linear models, with standard errors clustered at the household level.

As mentioned earlier, we are looking for consistent evidence across measures. For this reason we report all treatment effect estimates in one Table for each treatment: Table 3 for the Meal Treatment and Table 4 for the Snack Treatment.

Starting with the meal treatment (Table 3), we find a precisely estimated and negative treatment effect on BMI, of the order of 0.2 standard deviations in the early years and equal to 0.34 in year 3. The size of the effects is initially very similar across both treatments. The effect is remarkably robust and persistent.

Turning to the dietary intake measures, we present the individual coefficients as well as the tests for joint significance across measures (in a given year and across years). Not all coefficients of the treatment effects are precisely estimated, but there are robust significant effects. We find evidence from all measures pointing at a reduction in sugar consumption in years 1 and 2. In year 3, the effects are less consistent across measures. We also have

evidence (albeit weaker) of a reduction in fat consumption. While all except one of the point estimates is negative, only two of the ten coefficients are precisely estimated, jointly across all the coefficients on fat we find a significant effect. On the other hand, we do not find consistent evidence of a change in fruit and vegetable consumption. We also do not see a statistically significant drop in calories, but the reduction in added sugar consumption for the Meal group could be consistent with a drop in calories (corresponding to 66% of the point estimate change in calories at -101). After 3 years, we do not see a significant difference in added sugar intake, and the point estimate of the difference in calorie intake becomes positive (but remains insignificant).

For the Snack treatment (Table 4), we see less consistent effects. We find a negative treatment effect on BMI immediately after the intervention and in year 1 but the effect fades out such that, by year 2, the point estimate is no longer precisely estimated although it is still negative.

There is also evidence of a drop in fat and sugar consumption in the first year (post and/or year 1) but no consistent evidence emerges of significant changes in year 2 and beyond. These results do fit well with the observed changes in BMI.[14]

Figure 2 presents the distributions of percentile ranks in each group and year. The top three panels show the change for the control group and the two treatments before and immediately after the intervention. The bottom three panels show the changes after three years. We see that the distribution of the percentile rank shifts from the top of the distribution for both the Meal and the Snack treatments, albeit more pronounced in the Meal treatment. For both treatments, the intervention has moved children from the top of the BMI distribution more towards the bottom half of the distribution.

Altogether, these results provide evidence that children reduced their consumption of foods high in sugar, and to a lesser extent fat, which in turn affected their BMI. The effects are persistent for the Meal treatment, but fade away after a year for the Snack treatment.

In Section 6 we examine the robustness of the results to alternative inference procedures, multiple hypotheses testing and attrition.

[14]As an alternative analysis (available upon request), we have also computed two healthy eating indices. In a first index, we use the recommendation threshold for each macronutrient and create a dummy variable equal to 1 if the participant is within the recommendations' ranges, 0 otherwise. We then sum those dummies to create an index. We find no significant treatment effects on any of these two indices, for children or adults.

We present additional results in the Appendix. First, we present results including individual fixed effects (see Appendix Section A.4), which are almost identical. In Appendix Section A.5, we report the results of the analysis on the adult sample, for which we find no consistent evidence of improvements in dietary habits or changes in BMI. We also collected blood samples from a sub-sample adults to identify a number of key biomarkers, but again did not find evidence of improvements in diet (see Appendix Section A.6).

5.3 Food preferences

One question is whether the treatments significantly affected children's tastes, in line with a Becker-Murphy (1988) habit formation type of model. To investigate this possibility, we collected direct measures of food preferences.

Due to the young age of the children at the beginning of the study, our main measure is a simple non-incentivized measure of preferences for a range of foods. We collected additional incentivized measures of food preferences based on single one-shot decisions in year 1 for adults and year 3 for children (see Appendix Section A.7).

Children and adults were asked independently to rate their liking of a set list of foods. The survey was computer-based for parents, while for children it was administered on a one-to-one basis in a separate room by a member of our team. The questionnaire included 20 food items aimed at capturing a range of different food groups and 5 recipes that featured in the Meal treatment (see Appendix Table A.23 for the full list of items). For each item, participants had to answer on a 4-point scale how much they liked the item (really dislike to really like, illustrated with smiley emoticons), with the additional possibility of an 'allergic' or 'never tried' option. Items were then grouped into food categories following the "eatwell plate" food categories (fruit, vegetables, meat/fish/eggs, cheese, bread, unhealthy processed food and sweets).[15]

The ITT estimates are presented in Table 5. We find that children in the Meal treatment report a weaker preference for foods high in saturated fat, immediately after the intervention. The impact remains negative although not precisely estimated in the years after the inter-

[15]The eatwell plate is a policy tool used to define the British government recommendations on eating healthily and achieving a balanced diet.

vention. In contrast, children report liking sweets more. The point estimates for the Meal treatment effects are all positive, but they are neither jointly nor individually statistically significant. In the Snack treatment the point estimate is initially positive and significant but then the sign changes in subsequent years - overall, however, we find a positive and significant increase in the preference for sweets for those in the Snack treatment, although neither of the estimates for the effect on preferences for sweets remain statistically significant with more conservative inference procedures (Appendix Table A.28).

An increase in the liking of sweets seems a priori inconsistent with a habit formation model, but is consistent with some evidence in nutrition research, which shows that restricting access to certain foods can increase their valuation (Fisher and Birch, 1999).

In addition to the self-reported measure, we also collected a measure based on a one-shot incentivized choice in year 3 for children, when they were presumably old enough to fully understand an experimental protocol with incentives. The task was conducted one-to-one with each child and with the same investigator.

Children were offered a choice between a low calorie food item and a high calorie item. They were presented with four different pairs of items involving either sweet or savoury food, fresh or storable. To ensure that the low calorie item was attractive, we chose to impose a price on each item. The simplest price we thought we could implement was to attach a risk of not obtaining it. More precisely, if the child chose the low calorie item, she would have two chances out of three of actually getting it, while if she chose the high calorie item, she would only have one chance out of three to get it. The child had to indicate which item she chose among four different pairs of low-high calorie items (see Appendix Picture A.6)

The outcome measure of interest is a simple count of the number of times the child picks the low calorie (healthy) item. Great care was taken to ensure that the child understood the instructions.[16]

We find no statistically significant difference across treatments (see Table 6). That is, we have no indication that children in either treatment group developed a preference for "healthier" foods.

Altogether, these results provide little evidence that the interventions resulted in stronger

[16]Since the price involves risk, we also collected a measure of risk preference, inspired by the bomb task (Crosetto and Filippin, 2013, see Appendix for further details).

preferences for healthier foods. Thus, the most plausible mechanism for the changes observed is that children were restricted access to certain foods, such as foods high in added sugar.

6 Robustness and Attrition

In this section we document various robustness checks of the main effects. We focus on alternative inference procedures as well the issue of attrition. We include individual fixed effects for all outcome measures that were taken pre and post intervention. The relevant tables are presented in Appendix Section A.8.

6.1 Alternative inference procedures

We perform a randomization inference procedure set out in Young (2019). This involves a test of a sharp null (all participants of a particular treatment have a zero treatment effect rather than an average treatment effect of zero). These results are shown in Table A.27 for BMI and Table A.29 for dietary intake. Point estimates that are statistically significant when estimated with standard errors clustered at the household level remain so under the randomization inference procedure.

6.2 Multiple Hypotheses Correction

Recall that adjustments to p-values are made in Tables 3 and 4 for multiple hypothesis testing on the joint tests for coefficients for each of the domains we analysis (i.e. BMI, calories, sugar, fat and fruit and vegetables). Here we now document adjustments made within analysis of dietary recall separately. Table A.29 also shows those adjustments to the p-values when taking into consideration adjustment for the family wise error rate (FWER) as proposed by Romano and Wolf (2005a,b) - this shows us the chance that at least one of our outcomes within the family of outcomes is significant when the null hypothesis of no effect is true. When we carry out this procedure, the effect of the Snack treatment is less precisely estimated - fading out by year 1, which is more in line with the estimates on BMI. The effect of the Meal treatment on added sugar remains statistically significant even with

the most conservative approach.

There is a reduction in calories in the Snack treatment in the longer run, however, this estimate is not statistically significant when we use more conservative inference procedures (see Appendix Table A.29).

6.3 Attrition

As mentioned earlier, there is some attrition in our sample. To evaluate how attrition has an impact on our main results, we carry out three exercises. First, we compare the baseline characteristics of all those who started the experiment with the baseline characteristics of those who attended the final session in year 3. The idea is to see whether those who do and do not drop out differ based on the baseline measures of interest (our outcomes). There are no statistically significant differences - the means are remarkably similar (Appendix Tables A.30, A.31 and A.32). The biggest differences occur in the Snack treatment, as might be expected due to the larger attrition rate, however the differences for both adults and children are in absolute size small.

A second way to examine the impact of attrition, and specifically to see whether differences in attrition from session to session affect our estimates, is to compare the treatment effects of the sample of participants who attended every session (or who attended the final session) with the baseline estimates. We focus on the two outcomes with significant treatment effects and for which we have repeated measures: children's body mass index and consumption of added sugar.

The estimated treatment effects on a balanced sample or on a sample of those who attended the last session are slightly smaller but are very similar to the baseline estimates (Appendix Tables A.33 and A.34), suggesting that attrition was not systematically related to improvements in outcomes.

The third exercise is to estimate bounds on the treatment effects taking attrition into account in the spirit of Lee (2009). We describe the procedure in detail in Appendix Section A.8.3. Here again we focus on children's body mass index and consumption of added sugar. For each session, we sort the difference in outcomes comparing the baseline to the particular session within each treatment group from highest to lowest. Then for that particular session

the sample is trimmed from above and below to produce groups that have had equal rates of attrition. We then re-estimate the treatment effects based on the these trimmed samples (Appendix Tables A.35, A.36, and A.37). We do not trim to produce equal groups because we did not begin with equal groups by design as described above. Our conclusions are robust to this exercise.

These three exercises show that attrition cannot explain the significant treatment effects we have identified for children's body mass index and added sugar consumption.

7 Discussion and Conclusion

This paper presents the evaluation of two interventions targeting dietary habits. The first changes what families eat (Meal treatment), while the second targets regularity of eating and snack choices and behaviors (Snack treatment).

We find that, relative to the control group, children in both treatments moved down the distribution in terms of weight and BMI, particularly those from the top half of the distribution. This result persists for three years for the children exposed to healthy meals, but does not persist for those in the Snack treatment.

The most plausible explanation is that parents purchased less unhealthy foods (e.g. foods high in sugar) and that resulted in a lower BMI in the treated groups. The difference in consumption of added sugar disappears in year 3 though, but the effect on BMI persists. So it could be that the temporary reduction in sugar intake (for a couple of years), has led to a sustained decrease in BMI. The effects are weaker and less consistent across measures in the Snack intervention, and hence the lack of persistence of the BMI effect.

Is the change in body mass index observed for children large? A recent Cochrane review by Brown et al. (2019) summarizes the findings from 16 RCTs evaluating interventions targeting diet and physical activity among children between the age of 0 and 5 and find a mean effect on BMI z-score of -0.11 (95% confidence interval -0.21 to 0.01). They only report one study targeting dietary habits, by Daniels et al. (2012). The intervention tested is a universal educational intervention to promote protective feeding practices that commence in infancy. They find a reduction in BMI z-score of 0.19 six months after the intervention. In

that light, we conclude that the effects we find here (0.34 decrease in BMI z-score for the Meal treatment) are large.

The changes we observe do not seem to affect preferences for healthier foods. Thus we do not find evidence of a habit-formation mechanism in the spirit of habit formation models (Becker and Murphy, 1988). Of course it is possible that the interventions are not long enough to alter preferences.

As with any field experiment, one may be concerned about the external validity of the results. For ethical and practical reasons, participants had to be informed that participation to the study could involve having to follow a protocol for 12 weeks. One may perhaps expect these families to be more receptive than the average to policies encouraging changes. It could be that the results we find for children are an upper bound, if we believe the parents taking part in this study are more motivated than the average parent in encouraging healthy eating in their children. The baseline data doesn't provide much indication of a positive selection, in particular the diet of children doesn't meet the UK recommendations. So we do not have indication of positive selection on baseline characteristics.

Finally, even though these interventions are not meant to be policy proposals, the effects we find provide encouraging evidence that interventions aimed at changing children's diet early on in life can have a longer term impact.

References

Alhussain, Maha H., Ian A. Macdonald, and Moira A. Taylor. 2016. "Regular meal-pattern effects on energy expenditure, metabolism, and appetite regulation: a randomized controlled trial in healthy normal-weight women." *The American Journal of Clinical Nutrition*, 104(1), 21–32.

Allcott, Hunt, Rebecca Diamond, Jean-Pierre Dubé, Jessie Handbury, Ilya Rahkovsky, and Molly Schnell. 2019. "Food Deserts and the Causes of Nutritional Inequality." *Quarterly Journal of Economics*, 134(4): 1793–1844.

Becker, Gary S., and Kevin M. Murphy. 1988. "A Theory of Rational Addiction." *The Journal of Political Economy.* 96(4): 675–700

Belot, Michèle, Jonathan James, and Patrick Nolen, 2016. "Incentives and children's dietary choices: A field experiment in primary schools", *Journal of Health Economics*, 50: 213–229.

Bellisle, France. 2014. "Meals and snacking, diet quality and energy balance." *Physiology and Behavior* 134: 38–43.

Birch, Leann L. 1999. "Development of food preferences." *Annual Review of Nutrition*, 19(1): 41–62.

Brown T., Moore T.H.M., Hooper L., Gao Y., Zayegh A., Ijaz S., Elwenspoek M., Foxen S.C., Magee L., O'Malley C., Waters E., Summerbell C.D. 2019. Interventions for preventing obesity in children. *Cochrane Database of Systematic Reviews*, Issue 7. Art. No.: CD001871.

Capacci, S., M. Mazzocchi, B. Shankar, J. Brambila Macias, W. Verbeke, F Ja Pérez-Cueto, A. Koziol-Kozakowska, Beata Piorecka, B. Niedzwiedzka, D. D'Addesa, A. Saba, A. Turrini, J. Aschemann-Witzel, T. Bech-Larsen, M. Strand, L. Smillie, J. Wills, W. B. Traill. 2012. "Policies to promote healthy eating in Europe: a structured review of policies and their effectiveness." *Nutrition Reviews*, 10(3): 288–200.

Charness, Gary, Ramon Cobo-Reyez, Gabriel Katz, A.ngela Sanchez, Matthuias Sutter. 2019. "Improving Healthy Eating: Experimental Evidence", *mimeo.*

Crosetto, Paolo, and Antonio Filippin. 2013. "The "bomb" risk elicitation task" *Journal of Risk and Uncertainty*, 47(1): 31–65.

Cutler, David M., Edward L. Glaeser, and Jesse M. Shapiro. 2003. "Why have Americans become more obese?" *The Journal of Economic Perspectives*, 17(3): 93–118.

Dalstra, A.A., A.E. Kunst, C. Borrell, E. Breeze, E. Cambois, G. Costa, J.J.M. Geurts, E. Lahelma, H. Van Oyen, N.K. Rasmussen, E. Regidor, T. Spadea, and J.P. Mackenbach. 2005. "Socioeconomic differences in the prevalence of common chronic diseases: an overview of eight European countries." *International Journal of Epidemiology*, 34(2): 316–326.

Field, Alison E., Sydney Bryn Austin, Matthew W. Gillman, Bernard Rosner, Helaine R. Rockett, and Graham A. Colditz. 2004. "Snack food intake does not predict weight change among children and adolescents." *International Journal of Obesity*, 28(10): 1210–1216.

Frisvold, D. and Lansing, C. 2021. "Economics of Obesity" in K. F. Zimmermann (ed.), *Handbook of Labor, Human Resources and Population Economics*, Springer Nature Switzerland AG 2021

Hut, Stefan. 2020. "Determinants of dietary choice in the US: Evidence from consumer migration." *Journal of Health Economics*, 72: 102327.

Just, David R., and Joseph Price. 2013. "Using Incentives to Encourage Healthy Eating in Children." *Journal of Human Resources*, 48(4): 855–872.

Griffith, Rachel, Stephanie von Hinke, and Sarah Smith. 2018. "Getting a healthy start: The effectiveness of targeted benefits for improving dietary choices." *Journal of Health Economics*, 58, 176–187.

Hume, Catherine, Barbara Jachs, and John Menzies. 2016. "Homeostatic responses to palatable food consumption in satiated rats." *Obesity*, 24(10): 2126–2132.

Larson, Nicole, and Mary Story. 2013. "A review of snacking patterns among children and adolescents: what are the implications of snacking for weight status?" *Childhood Obesity*, 9(2): 104–115.

Laska, Melissa N., Mary O. Hearst, Katherine Lust, Leslie A. Lytle, and Mary Story. 2015. "How we eat what we eat: identifying meal routines and practices most strongly associated with healthy and unhealthy dietary factors among young adults." *Public Health Nutrition*, 18(12): 2135–2145.

Leech, Rebecca M., Anthony Worsley, Anna Timperio, and Sarah A. McNaughton. 2015. "Understanding meal patterns: definitions, methodology and impact on nutrient intake and diet quality." *Nutrition Research Reviews*, 28(1): 1–21.

List, John A., and Anya Savikhin Samek. 2015. "The behavioralist as nutritionist: Leveraging behavioral economics to improve child food choice and consumption." *Journal of Health Economics*, 39: 135–146.

Loewenstein, George, Joseph Price, and Kevin Volpp. 2016. "Habit formation in children: Evidence from incentives for healthy eating", *Journal of Health Economics*, 45: 47–54.

Ludwig, Jens, Jeffrey R. Kling, and Sendhil Mullainathan. 2011. "Mechanism Experiments and Policy Evaluations." *Journal of Economic Perspectives*, 25(3): 17–38.

Read, Daniel, and Barbara Van Leeuwen. 1998. "Predicting Hunger: The Effects of Appetite and Delay on Choice." *Organizational Behavior and Human Decision Processes*, 76(2): 189–205.

Spalding, K. L., and Amer, E. Westermark P.O. Bernard S, Buchhoiz BA, Bergmann O, Blomqvist L, Hofl'stedt J, Naslund E, Britten T, Concha H, Hassan M, Ryden M, Frisen J, Amer P., 2008. "Dynamics of fat cell turnover in humans." *Nature*, 453(7196): 783–787.

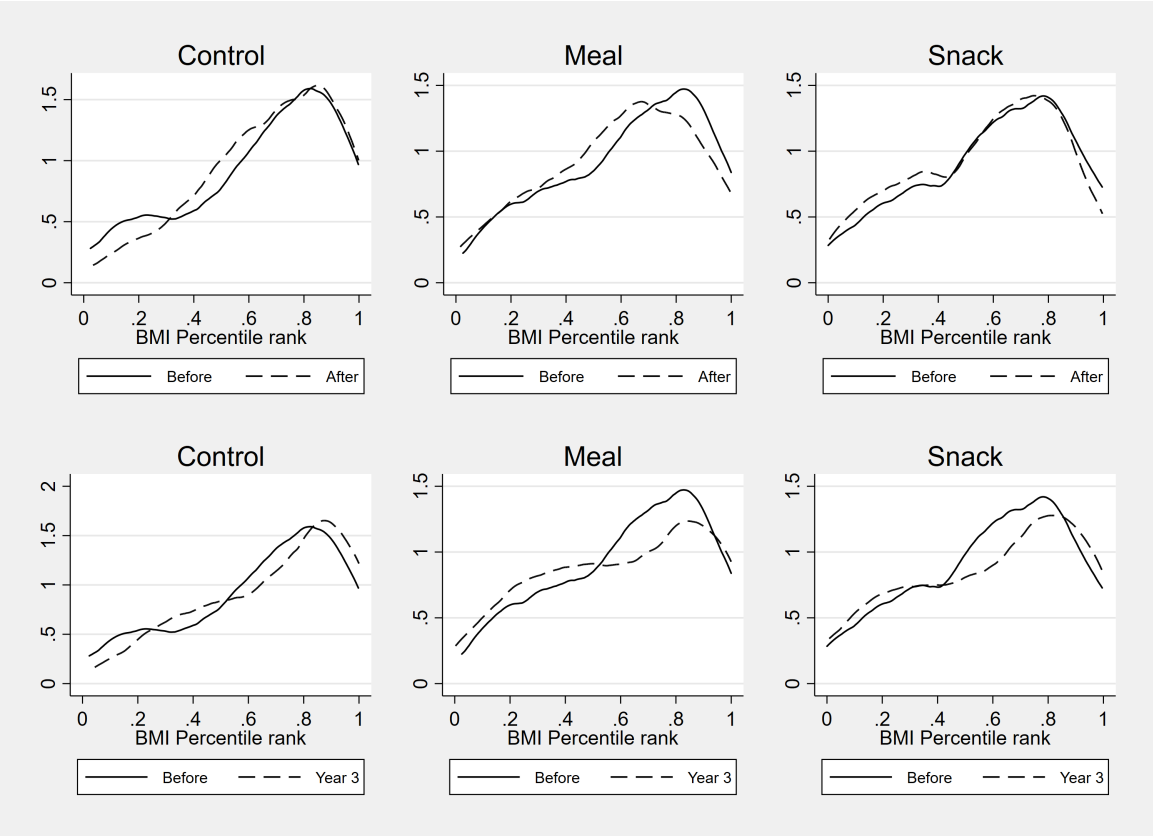
St-Onge, Marie-Pierre, Kathleen L. Keller, and Steven B. Heymsfield. 2003. "Changes in childhood food consumption patterns: a cause for concern in light of increasing body weights." *American Journal of Clinical Nutrition.* 78(6): 1068–1073.

Spiteri, Jonathan, Jonathan James, and Michèle Belot. 2019. A computer-based incentivized food basket choice tool: presentation and evaluation. *PLOSOne*, 14(1): e0210061.

Vidmar, Suzanna, John Carlin, Kylie Hesketh, and Tim Cole. 2004. "Standardizing anthropometric measures in children and adolescents with new functions for egen." *The Stata Journal*, 4(1): 50–55.

Tables and Figures

Figure 2: Changes in the distribution of the BMI percentile rank in children



Note: Each line shows the kernel density function for the distribution for the percentile ranks of BMI of children. The top three panels show the change for the control group and the two treatments before and immediately after the intervention. The bottom three panels show the changes after three years

Table 1: Demographic characteristics at baseline and across groups

	Control	Meal	Snack	P-value	P-value
		Mean (std)			
	(1)	(2)	(3)	(1)=(2)	(1)=(3)
Sample Size (Families)					
Start of the study	111	103	71	-	-
After of the intervention	109	101	64	-	-
3 years later	92	90	51	-	-
Demographics					
% Female adults among all sample	72.2	79.6	75.3	0.15	0.59
# Adults in household	1.7	1.61	1.7	0.43	0.85
	(0.85)	(0.6)	(0.7)		
# Children in household	1.9	1.8	1.9	0.47	0.9
	(0.9)	(1.0)	(1.0)		
Age (adults)	35.1	34.7	34	0.67	0.23
	(7.5)	(6.5)	(6.9)		
Age (study child)	3.9	4	3.9	0.99	0.75
	(1.7)	(1.7)	(1.4)		
Mean annual household income	20,855	21,167	23,928	0.87	0.15
	(10,056)	(19,227)	(21,844)		
% Receiving housing benefits	37.8	41.7	38.0	0.56	0.98
% Receiving income support	22.5	17.5	22.5	0.36	0.99
% No qualifications	2.7	3.1	3.2	0.85	0.81

Note: Means with standard deviations in parentheses. 4th and 5th columns report the p-value of a t-test of equality of estimated parameters in Col. (1) and (2) and in Col. (1) and (3) respectively. *p-values* from Kolmogorv-Smirnov test of distribution are reported to compare the BMI categories distribution between groups, rank-sum tests were performed to compare BMI levels. A descriptive statistics table for Edinburgh and Colchester sample separately can be found in Appendix Table A.9.

Table 2: Baseline measurements (children)

	Control	Meal	Snack	P-value	P-value
		Mean (std)			
	(1)	(2)	(3)	(1)=(2)	(1)=(3)
Body Size					
BMI z-score	0.51	0.49	0.33	0.591	0.251
% Underweight	0.0	0.0	1.47	*1.000*	*0.999*
% Normal weight	71.0	75.5	75.0		
% Overweight	12.2	6.86	8.82		
% Obese	16.8	17.7	14.7		
Diet Intake					
Total calorie intake (Kcal)	1438.9	1463.8	1383.2	0.34	0.93
	(538.6)	(475.4)	(378.0)		
Added sugar (g)	18	25.9	18.1	0.21	0.41
	(22.6)	(34.8)	(20.1)		
Saturated Fat (g)	23.9	25.9	23.8	0.23	0.94
	(11.9)	(12.9)	(11.6)		
Fruit and veg (g)	101.5	122.4	123.5	0.23	0.27
	(124.7)	(126.8)	(141.1)		

Note: Means with standard deviations in parentheses. 4th and 5th columns report the p-value of a t-test of equality of estimated parameters in Col. (1) and (2) and in Col. (1) and (3) respectively. *p-values* from Kolmogorv-Smirnov test of distribution are reported to compare the BMI categories distribution between groups, rank-sum tests were performed to compare BMI levels. A descriptive statistics table for Edinburgh and Colchester sample separately can be found in Appendix Table A.9. Self-reported preferences: An item that has never been tried or for which the participants declares to be allergic to is considered missing. 1 corresponds to not liking at all, 4 to liking very much.

Table 3: Impact of Meal Treatment on Body Mass Index and Diet of Children

	Estimate	Post	Year 1	Year 2	Year 3
Body Mass Index (z-score for children)	DD	-0.21 (0.06)***	-0.23 (0.08)***	-0.21 (0.09)***	-0.34 (0.12)***
Joint significance of all coefficients			0.003 [0.004]		
Total calories (kcal)					
24h-dietary recall	DD	-54.8 (72.4)	-101.0 (108.3)	-156.5 (115.0)	86.6 (116.6)
Supermarket parental choice	Treat-Control			-20.3 (21.0)	3.8 (22.4)
Joint significance coefficients by year				0.244	0.75
Joint significance of all coefficients			0.389 [0.389]		
Sugar					
24h-dietary recall (added sugar, g)	DD	-6.0 (5.3)	-21.9 (8.1)***	-26.7 (10.2)***	2.7 (8.9)
Self-reported changes (foods high in sugar)	Treat-Control	-0.3 (0.4)	-1.4 (0.4)***	-0.8 (0.4)**	-0.7 (0.4)**
Supermarket parental choice (g)	Treat-Control			-4.4 (1.6)***	-2.0 (1.6)
Joint significance coefficients by year		0.4427	0.0000	0.0008	0.1275
Joint significance of all coefficients			0.000 [0.001]		
Fat					
24h-dietary recall (saturated fat, g)	DD	-2.8 (1.9)	-3.1 (2.5)	-3.2 (2.5)	-0.7 (2.6)
Self-reported changes (foods high in fat)	Treat-Control	-1.2 (0.3)***	-1.2 (0.4)***	-0.4 (0.3)	0.1 (0.4)
Supermarket parental choice (saturated fat in g)	Treat-Control			-1.0 (0.9)	-0.5 (0.9)
Joint significance coefficients by year		0.0007	0.0014	0.2345	0.9263
Joint significance of all coefficients			0.0006 [0.002]		
Fruit and vegetables					
24h-dietary recall (g)	DD	-16.7 (21.1)	-6.5 (39.6)	-26.1 (26.8)	-37.8 (26.2)
Self-report changes	Treat-Control	0.5 (0.1)***	0.4 (0.1)***	0.2 (0.1)*	0.2 (0.1)
Supermarket parental choice (% spent on F&V)	Treat-Control			0.3 (0.6)	-0.2 (0.5)
Joint significance coefficients by year		0.0000	0.0194	0.2326	0.1983
Joint significance of all coefficients			0.0009 [0.002]		

Note: Standard errors in parenthesis clustered at the household level. p<0.01 ***, p<0.05 **, p<0.1 *. Joint significance of all coefficients is the p-value of the test as to whether all coefficients are jointly significant for that particular outcome, in square brackets are the FDR q-value calculated using the method from Benjamini and Hochberg (1995) and Anderson (2008). Estimate indicates the estimation strategy with "DD" being difference-in-differences and "Treat-Control" tests the difference between the treatment and control group at for that particular time period. For self-reported changes for foods high in sugar and fat these range from -8 to +8, and for for fruit and vegetables these are reported as the number of changes out of 2 categories.

Table 4: Impact of Snack Treatment on Body Mass Index and Diet of Children

	Estimate	Post	Year 1	Year 2	Year 3
Body Mass Index	DD	-0.21 (0.09)***	-0.18 (0.11)*	-0.12 (0.14)	-0.16 (0.19)
Joint significance of all coefficients			0.078 [0.098]		
Total calories (kcal)					
24h-dietary recall	DD	-48.4 (74)*	-133.8 (101.3)	-174.4 (110.5)	-84.0 (106.7)
Supermarket parental choice	Treat-Control			8.4 (25.8)	25.8 (29.1)
Joint significance coefficients by year				0.277	0.454
Joint significance of all coefficients			0.626 [0.626]		
Sugar					
24h-dietary recall (added sugar, g)	DD	-0.5 (4.7)	-19.0 (8.6)**	-18.7 (10.6)*	-7.2 (9.0)
Self-reported changes (foods high in sugar)	Treat-Control	-0.9 (0.5)*	-0.7 (0.4)*	-0.1 (0.4)	0.1 (0.4)
Supermarket parental choice (g)	Treat-Control			0.7 (1.9)	2.1 (2.3)
Joint significance coefficients by year		0.1626	0.0276	0.3292	0.6238
Joint significance of all coefficients			0.0615 [0.098]		
Fat					
24h-dietary recall (saturated fat, g)	DD	-1.5 (2.0)	-1.5 (2.4)	-3.0 (2.4)	-1.7 (2.5)
Self-reported changes (foods high in fat)	Treat-Control	-1.3 (0.4)***	-0.4 (0.4)	0.03 (0.4)	0.8 (0.4)
Supermarket parental choice (saturated fat in g)	Treat-Control			0.1 (1.2)	0.4 (1.1)
Joint significance coefficients by year		0.0015	0.5609	0.6647	0.2104
Joint significance of all coefficients			0.0007 [0.003]		
Fruit and vegetables consumption					
24h-dietary recall (g)	DD	-23.4 (25.6)	-31.7 (32.6)	-30.1 (30.8)	-24.0 (29.8)
Self-report changes	Treat-Control	0.6 (0.1)***	0.5 (0.1)***	0.1 (0.2)	0.3 (0.2)*
Supermarket parental choice (% spent on F&V)	Treat-Control			0.5 (0.8)	-0.4 (0.8)
Joint significance coefficients by year		0.0000	0.0055	0.5800	0.309
Joint significance of all coefficients			0.0009 [0.003]		

Note: Standard errors in parenthesis clustered at the household level. p<0.01 ***, p<0.05 **, p<0.1 *. Joint significance of all coefficients is the p-value of the test as to whether all coefficients are jointly significant for that particular outcome, in square brackets are the FDR q-value calculated using the method from Benjamini and Hochberg (1995) and Anderson (2008). Estimate indicates the estimation strategy with "DD" being difference-in-differences and "Treat-Control" tests the difference between the treatment and control group at for that particular time period. For self-reported changes for foods high in sugar and fat these range from -8 to +8, and for for fruit and vegetables these are reported as the number of changes out of 2 categories.

Table 5: Meal and Snack treatment effects on food preferences (children)

	Fruits	Vegetables	Meat Fish Eggs	Processed Food	Sweets	Bread	Cheese
After	0.13**	-0.06	0.00	0.07	-0.14**	0.05	0.20**
	(0.06)	(0.07)	(0.10)	(0.05)	(0.07)	(0.09)	(0.10)
1-year	0.05	-0.01	0.16	0.06	-0.09	0.01	0.02
	(0.07)	(0.08)	(0.10)	(0.07)	(0.08)	(0.10)	(0.11)
2-year	0.09	-0.09	0.02	0.08	-0.16**	0.03	0.03
	(0.09)	(0.09)	(0.12)	(0.07)	(0.08)	(0.11)	(0.13)
3-year	0.03	-0.11	0.11	-0.07	-0.29***	-0.14	-0.24*
	(0.09)	(0.09)	(0.11)	(0.07)	(0.09)	(0.10)	(0.14)
Meal x After	-0.10	0.01	-0.07	-0.21***	0.17*	-0.17	-0.33**
	(0.09)	(0.10)	(0.14)	(0.08)	(0.09)	(0.13)	(0.13)
Meal x 1-year	-0.01	0.03	-0.08	-0.09	0.10	-0.11	-0.08
	(0.10)	(0.11)	(0.15)	(0.09)	(0.13)	(0.13)	(0.16)
Meal x 2-year	-0.06	0.04	-0.08	-0.13	0.23*	-0.21	-0.19
	(0.12)	(0.12)	(0.16)	(0.09)	(0.13)	(0.14)	(0.18)
Meal x 3-year	-0.04	0.08	-0.20	-0.03	0.20	-0.13	0.03
	(0.12)	(0.12)	(0.16)	(0.09)	(0.13)	(0.14)	(0.18)
Snack x After	0.03	0.13	-0.14	0.02	0.21**	0.24*	-0.18
	(0.11)	(0.11)	(0.14)	(0.09)	(0.10)	(0.15)	(0.16)
Snack x 1-year	0.15	0.13	-0.26	0.04	-0.04	0.13	-0.12
	(0.12)	(0.12)	(0.19)	(0.10)	(0.15)	(0.14)	(0.17)
Snack x 2-year	0.16	0.26**	-0.19	0.15	0.18	0.12	-0.21
	(0.13)	(0.12)	(0.18)	(0.12)	(0.11)	(0.17)	(0.20)
Snack x 3-year	0.09	0.06	-0.17	0.07	-0.10	-0.04	-0.23
	(0.15)	(0.11)	(0.18)	(0.12)	(0.14)	(0.17)	(0.21)
Constant	3.19***	2.58***	2.66***	3.35***	3.52***	3.47***	3.29***
	(0.03)	(0.03)	(0.04)	(0.02)	(0.03)	(0.04)	(0.04)
# Obs	1,243	1,254	1,254	1,257	1,251	1,248	1,244
R-squared	0.01	0.01	0.01	0.02	0.04	0.03	0.03
# individuals	288	289	289	289	289	289	288

Note: Each column is from a separate regression. All regressions include individual fixed effects. Standard errors in parentheses are clustered at the household level, *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Meal and Snack treatment effects on children's incentivized food preferences

	Number of unhealthy items chosen		
Meal	0.048	0.061	0.085
	(.164)	(.168)	(.170)
Snack	-0.184	-0.15	-0.132
	(.192)	(.196)	(.197)
Risk preference		0.006	0.003
		(.031)	(.031)
Controls for gender and age	N	N	Y
Observations	212	212	212

Note: Each column in each panel represents a separate regression. Data was collected on children's incentivized food preferences in year 3 only.