Incentives and Children's Dietary Choices: A Field Experiment in Primary Schools*

Michèle Belot[†], Jonathan James[‡], and Patrick Nolen[§]

January 16, 2015

Abstract

We conduct a field experiment in 31 primary schools in England to test the effectiveness of different temporary incentive schemes, an individual based incentive scheme and a competitive scheme, on increasing the choice and consumption of fruit and vegetables at lunchtime. The individual scheme has a weak positive effect whereas all pupils respond to positively to the competitive scheme. For our sample of interest, the competitive scheme increases choice of fruit and vegetables by 33% and consumption of fruit and vegetables by 48%, twice and three times as much as the individual incentive scheme, respectively. The positive effects generally carry over to the week immediately following the treatment but we find little evidence of any effects six months later. Our results show that incentives can work, at least temporarily, to increase healthy eating but there are large differences in effectiveness between schemes and across demographics such as age and gender.

JEL Classification: J13, I18, I28, H51, H52

Keywords: Incentives, Health, Habits, Child nutrition, Field experiments

*We thank the Esme Fairbairn Foundation for their financial support (grant number 10-2206). This paper was previously circulated as "Changing Eating Habits: A Field Experiment in Primary Schools." We thank seminar and workshop participants at Aarhus, Bath, Bristol, Ghent, Glasgow, Edinburgh, Essex, Queen's University Belfast, Oslo, Cologne, 'The Determinants of Dietary Choices: An Interdisciplinary Approach' in Edinburgh (2013), the Royal Economic Society annual conference (2013), the Florence Workshop on Behavioural and Experimental Economics (2013), the International Conference in Public Policy (2013), the European Economic Association annual conference (2013) and the European Society of Population Economics annual congress (2014). Finally, we especially thank all the local authorities, schools, lunch monitors, parents, and children who helped and participated in the experiment.

[†]University of Edinburgh, email: Michele.Belot@ed.ac.uk

[‡]Corresponding Author. University of Bath, email: jj412@bath.ac.uk

[§]University of Essex, email: pjnolen@essex.ac.uk

1. Introduction

Poor nutrition is a primary cause behind the rising cost of health care in many developed countries.[1] According to the World Health Organization (2009) poor nutrition is related to three of the five highest risks for morality in the world: high blood pressure; high blood glucose; and overweight and obesity. In response, policy makers have been pushing information interventions, such as the "5-a-day" campaign in the UK, to encourage people to develop better eating habits. However, the success of these campaigns has been moderate.[2]

This paper investigates how to incentivise school age children to consume healthier food. Recent evidence shows that incentives can motivate people to exercise (Charness and Gneezy (2009), Acland and Levy (2013)), stop smoking (Volpp et. al (2009) and Giné et. al. (2011)) and eat more fruit and vegetables (Just and Price (2003)). While the evidence is encouraging, it remains an open question which incentives work best and for whom. We are particularly interested in changing the behaviour of two key groups: boys and children from low socioeconomic status, both of which have been shown to have a less healthy diet and are particularly resistant to change (see Muller et al. (2005), Perry et al. (1998) and Kelder et al. (1995)). We use insights from behavioural economics to investigate whether we can improve the intake of healthy foods overall and for these groups in particular by providing incentives to select fruit and vegetables during school lunches.

Using incentives to encourage the healthy eating is a controversial idea. Indeed, there is evidence showing that rewarding children for eating fruit and vegetables can lead to those items being less preferred (using self-reports as a measure of preference (Birch et. al. (1982), Birch et. al. (1984), and Newman and Taylor. (1992)). We test the effectiveness of two different incentive schemes: an individual based incentive and a competitive incentive. The idea of using a competition rather than an individual incentive is inspired by the recent evidence in behavioural economics showing that men tend to be more competitive than women (see Gneezy et. al (2003), Gneezy and Rustichini (2004), and Booth and Nolen (2012)). Competitive incentives have not yet been studied in exercise or health. While this might have potential, it also has the threat of being effective only for boys or competitive kids while discouraging others. We are primarily interested in the effects for immediate food intake, but also look at the build-up of short and long-run health habits

[1] See Bhattacharya and Sood (2011) for an overview of the costs of obesity.

[2] See Ciliska et al. (2000) for a review of many community based interventions. They appear to have been successful at informing people but have had less success in changing actual behaviour (see Robertson (2008) and Verplanken and Wood (2006)).

once incentives are removed.

We conduct a randomised field experiment in 31 primary schools across the United Kingdom and implement the incentive schemes for four weeks. In each school it was implemented in a class in year 2 (pupils aged 6–7) and in year 5 (pupils aged 10–11) to be able also investigate effects by age. We find that the competitive scheme works well overall, with no negative effects on any subgroup. The results of individual incentives are more mixed, and the scheme is overall less effective. The competition treatment is more effective across the board and is overall nearly three times as effective at getting children to consume a portion of fruit or vegetables at lunch. If we focus on the specific group of children who did not consume fruit and vegetables every day before the intervention started, we find that the competitive scheme increases their likelihood of trying a fruit or vegetable at lunch by 48%.

Our second important finding is that incentives do not work in the same way for everyone. We find that, in general, females, pupils from poorer socio-economic backgrounds, and younger children respond more positively to competition than the individual based incentive. The individual based incentive even appears to have a negative effect on younger children. Other subgroups, such as boys, older children, and pupils from wealthier socio-economic backgrounds respond positively to both the competitive and the individual incentive scheme, though, the estimated effect is larger for the competition treatment in nearly every case. This suggests that using a competitive incentive could improve effectiveness by increasing the choice and consumption of those already responding to the individual scheme *and* those groups that typically do not respond to health interventions.

The results presented in this paper are directly relevant for policy. We show that incentives do work in encouraging healthy dietary choices and that the results of a short term intervention can have lasting effects after the intervention period but that a "one-size-fits-all" reward scheme will not likely work. The differential effects by subgroup suggest that health incentives need to be evaluated at the individual level and, consequently, different policies may have to be developed for different subgroups or an incentive scheme other than the standard individual scheme may have to be considered. Furthermore, increasing the length of time an intervention is taking place is not the only way policy makers can increase the likelihood that positive behaviours are adopted: for instance, competitions could be more effective than individual-based schemes at changing behaviour in the same time period.

The remaining part of the paper is structured as follows. In Section 2 we discuss the related literature. Section 3 presents the experimental design and Section 4 presents a simple conceptual framework and hypotheses that guide the analysis of the results. We present the results in Section 4 and conclude in Section 5.

2. Background and related literature

The most related paper to our work is by Just and Price (2013), who tested various individual incentive schemes in fifteen schools in two districts in Utah. The treatment period was 2-3 weeks. They found that short term rewards are given for eating fruit and vegetables does lead to an increase in the proportion of children consuming a serving of fruits or vegetables at lunch time. They find no lasting change in the amount of fruits and vegetables consumed two weeks after the incentive has been removed. The lack of longer term effects could be due to the intervention period being too short or the incentive scheme not being effective enough.

More generally, our paper relates to the literature on behavioural anomalies underlying 'unhealthy' behaviours. Present-biased (hyperbolic) preferences, such as those discussed in Laibson (1997) and O'Donoghue and Rabin (1999), can explain unhealthy dietary choices despite an individual being fully aware of (having all the information about) the effects of poor nutrition and the benefits of healthy eating: individuals may over-weight the initial costs of eating healthier and (or) under-weight the longer term benefits. In that context, using a temporary and effective incentive scheme to encourage healthier eating among children could lead to long term dietary habit changes.[3] Interestingly for our study, recent work has shown that boys and children from poorer socio-economic backgrounds are more impatient than other children[4] and those differences could explain why these children are less likely to make healthy dietary choices. In that context, providing immediate incentives to eat healthily may prove a powerful tool to get these groups to respond.

The effectiveness of different interventions on changing behavior has not been widely examined.[5] There is a well-established literature showing that boys tend to be more competitive than girls (see Geenzy et. al (2003), Gneezy and Rustichini (2004), and Booth and Nolen (2012)) yet competitive incentives have not yet been studied in exercise or health.

[3]Work by Kelder et. al. (1994), Resnicow et. al. (1998), and Singer et. al. (1995) suggest that dietary habits appear to form in childhood and track into adulthood.

[4]See Delaney and Doyle (2012) for children from poorer socio-economic backgrounds and Bettinger and Slonim (2007) for boys versus girls.

[5]Some work has looked at the effect of information only campaigns versus interventions with individual based incentives with small prizes (see List and Samek (2014) for example) but we know of no study that has looked at two reward based schemes.

3. Experimental Design

To examine the effect of two incentive schemes on the choice and consumption of fruit and vegetables we conducted a field experiment in England. We recruited schools in a three step process.[6] First we approached all 150 Local Education Authorities (LEAs) in England to ask if they would be interested in participating; 22 responded positively. Second, we provided more information about the project to LEAs that responded and set-up meetings with them to answer questions and discuss how to recruit schools. We indicated to LEAs that we were interested in testing and comparing the effectiveness of incentives schemes in increasing choice and consumption of fruit or vegetables at lunchtime and that the interventions were specifically designed to target children who were generally considered unresponsive to health interventions. After the meetings 12 LEAs agreed to let us approach their schools and provided a list of at least three schools that would consider being involved. Finally we approached all 46 schools suggested by the LEAs; 31 of them agreed to participate.

We recruited children from year two (aged 6 and 7) and year five (aged 9 and 10) in participating schools. Parents were provided with information about the study, asked to fill out a questionnaire, and were required to give consent to have data collected about their child. As agreed with the schools, all children in years two and five were included in the project. However, data about choice and consumption of fruit or vegetables were only recorded for children whose parents gave permission. Therefore, we have data on 638 children for the main part of the analysis.

Randomisation

We randomly allocated schools to one of three groups: control; competition; or individual incentive. We were particularly careful to make sure that, *ex ante*, the average school in each group had roughly the same number of children and looked the same in terms of school characteristics.

Within LEA schools were randomly assigned to treatment arms such that the overall sample was balanced based on observables. For the purpose of balancing the three groups we used the following characteristics: (i) proportion of female pupils; (ii) number of pupils; (iii) number of pupils in class groups (year 2 and year 5); (iv) proportion of children eligible for free school meals; (v) proportion of children eating free school meals; (vi) per pupil expenditure; (vii) per pupil expenditure on catering; (viii) percent of children

[6] A companion paper, Belot and James (2013), documents the selection process of which schools choose to participate in this experiment. In particular they find that selection on observables and unobservables is unlikely to drive the results.

achieving level 4 in both English and Mathematics; (ix) average point scores of children on level 4 exams; (x) average percent of children absent on a given day; (xi) percent of children absent from the level 4 exams; (xii) school type (religious or comprehensive); (xiii) whether a school was involved in the "Food for Life" Programme; (xiv) Ofsted School Categorization; and (xv) Ofsted Health Categorization (OfHealth).

The variables listed above were used to make sure that the average school in each treatment arm was similar in ways that could have influenced whether the treatment scheme worked: socio-economic background of the student body; school quality; student quality; and school type.[7] Using a random number generator, schools were assigned to one of the three treatment arms. We then checked to see if the sample was balanced based on the 15 observable characteristics. If it was not, we re-started the randomization. This ensures that, *ex ante*, at the *school level*, our sample was balanced by treatment arm.

Treatments

The two treatments we designed incentivise *choice* (rather than *consumption*) of fruit or vegetables at lunch. We decided to incentivise choice for a few reasons. First, the health literature highlights how making rewards contingent on consumption of a particular food can cause children to have a lower preference for that item (see Birch et. al. (1982, 1984) and Newman and Taylor (1992) for examples). We wanted to minimise the potential for negative effects on healthy eating. Second, we wanted the experiment to be something that was relevant to policy and simpler to implement. Rewarding for choice removes any subjective judgement of the monitor to decide what constitutes an adequate amount of food consumed to be rewarded. Furthermore, schools can require children to take a fruit or vegetable at lunch but are unlikely to be able to force them to eat the item. Therefore the results of our study are likely to be more relevant to policies that are being considered at the school level now.[8] Third, we also wanted the program to involve minimal costs. Finally, rewarding for choice rather than actually consuming an item negates the possibility of cheating. For example, if rewards were based on eating, pupils may have an incentive to dispose of the fruit or vegetable, hide it, give it to a friend or try to mislead

[7]Variables (i), (ii), and (iii) relate to the demographic characteristics of the schools involved. Variables (iv) and (v) relate to the economic background of the children. Variables (vi) and (vii) relate to the financial expenditure at the school level. Variables (viii) - (xi) relate to the quality of the student body at each school. Variable (xii) denotes if a school has a religious affiliation. Variable (xiii) denotes whether the school voluntarily chose to be part of the "Food for Life" programme which involves schools agree to teach children about healthy eating (See http://www.foodforlife.org.uk/ for further information). Variable (xiv) is the overall classification of the school based on its Ofsted results: 1 = outstanding; 2 = good; 3 = requires improvement; and 4 = inadequate. Variable (xv) relates to the extent to which the pupils adopt a healthy lifestyle.

[8]Indeed the results of our study are especially relevant to determine if providing (or requiring a pupil to take) a fruit or vegetable at lunchtime has any follow through effect on consumption behaviour.

monitors regarding actual consumption. For this reason, monitoring consumption is more reliable when choice is incentivised and we will be able to check if children eat healthier options or not.

In both of our experimental schemes, the standard individual and competitive, the pupils were given a sticker for choosing or bringing in a fruit or vegetable at lunch.[9] The individual incentive scheme was chosen because it is similar to many of the other individual based incentive schemes used in the healthy eating and habit formation literature (for instance, see Charness and Gneezy (2009), Just and Price (2013), or List and Samek (2014)). The competition was chosen because the literature on gender and competition suggests that boys respond more to competition than girls (see Gneezy and Rustichini (2004), Gneezy et. al. (2003), and Booth and Nolen (2012)). Given that boys tend not to respond to traditional healthy eating interventions, the competition was seen as an incentive scheme that could get boys to respond. However, gender differences in competition can vary by task (see Iriberri and Rey-Biel (2011)). Therefore if the task of choosing a healthily item is viewed as a 'favouring females' then even the competitive scheme might not get boys to choose or consume fruit or vegetables.

In both schemes children received a sticker every day they chose or brought in a fruit or vegetable at lunchtime [10] Then, at the end of the week (Friday afternoon after lunch), each pupil had the opportunity to pick a larger prize depending on the incentive scheme in which the pupil was enrolled. In the individual incentive scheme, if a pupil collected four stickers in the week she or he was allowed to choose a prize such as an item of stationery or a small toy from a reward box. If the pupil had three or less stickers, though, the pupil could not pick a prize and the stickers did not count to earning an award next week. In the competition, children were assigned to random groups of four, and only the pupil with the most stickers in each group was able to select a prize from the reward box.[11] In the case of a tie all children with the highest number of stickers in the group were eligible for a prize. The groups were revealed at the end of the week after lunch so children would not engage in strategic behaviour, such as making choices based on other group member's actions or absenteeism. For example, if a pupil was absent on Monday then the others in their group would know that that pupil could only collect a maximum of four stickers. The groups were changed each week so the children could not anticipate with whom they

[9]Examples of the stickers can be seen in the appendix. All children were given a list of fruits and vegetables that would be rewarded if they were included in packed lunches; the list is also included in the appendix.

[10]Monitors, who recorded whether children were choosing and consuming fruit and vegetables at lunch time, were either canteen staff working in the school or parents of children occasionally hired by the school for help at lunch time.

[11]See appendix for pictures of some of the rewards from which children were allowed to choose.

would be competing and, in this treatment as well, unused stickers did not carry over to the next week.

Timing

Before the interventions began a background survey was sent to the parents that covered information on age, gender, ethnicity, primary language, height, weight, and typical dietary habits. Then, starting the second week of October, we monitored what children ate at lunch in all 31 schools. Lunch monitors[12] recorded if a pupil chose a fruit or vegetable or brought a fruit or vegetable in with a packed lunch and if the pupil consumed none, some, or more than half the item. On Friday that week children took a food knowledge test and a "spot-the-difference" test.[13] The food knowledge test required pupils to identify seven pictures of different items (e.g. celery or snickers bar) and mark if each item was healthy or not. The "spot-the-difference" test was designed to test a pupil's concentration and required a pupil to compare two sets of 30 dice that were arranged in a six-by-five square. There were five differences between the two sets of dice; the pupil was asked to circle the five differences. Children had 10 minutes to complete each test.

The children went on half-term break for one week after the baseline data was collected. Upon returning to school the children were reminded of the project and children were monitored for the next five weeks. At control schools, the lunch monitors continued to monitor children in the same way they did during the week in October: they collected data on whether a pupil choose or consumed a fruit or vegetable. At the competition and individual incentive schools children were incentivised to choose a fruit or vegetable for a period of four weeks[14]. Each day a pupil choose or brought in a fruit or vegetable with a packed lunch[15] the pupil received a sticker. Furthermore, as discussed above, at the end of each week, children would get a large prize based on the type of incentive scheme in which they were enrolled.

On the fourth Friday of the treatment, the children completed another food knowledge and "spot-the-difference" test and were reminded that it was the last day of incentives. The following week, immediately after the treatment, the choices and consumption of children were still monitored. This allows us to see if there was any effect on choice and consumption after the incentives were removed. To examine the longer term effects of

[12]Lunch monitors were dinner ladies who worked in the cafeteria or school assistants who were normally present at lunch time and sat with the children as usual during the lunch period.

[13]Examples of both can be seen in the appendix.

[14]Just and Price (2013) incentivised children for a period of 2-3 weeks and found no longer run effects. Therefore, we chose to incentivise children for a longer period of time; 1-2 weeks longer.

[15]With the questionnaire and again at the start of the five weeks of monitoring, the parents of all children received lists of what items would count as healthy if they were included with packed lunches.

the incentives we also went back to schools six months later, in June, and monitored the choice and consumption of the same children.

4. Conceptual Framework & Hypotheses

We designed our field experiment to test the three hypotheses laid out below, to examine whether there were heterogeneous effects of incentives, and to compare the two incentive schemes.

Hypothesis 1: Children will choose more fruit or vegetables when they are rewarded for taking a fruit or vegetable at lunchtime.

By providing a reward for choosing a healthy option, the benefit of taking a fruit or vegetable at lunchtime will have increased for each pupil. Therefore we would expect that, while the incentive scheme is running, children are more likely to choose a fruit or vegetable. This would be consistent with the work by Gneezy and Charness (2009), Just and Price (2013), and List and Samek (2014). Furthermore, the effect is likely to differ by subgroups. Since boys and children from poorer socio-economic backgrounds have been shown to be more impatient (see Delany and Doyle (2012) and Bettinger and Slonim (2007)) then they may respond more positively to the immediate reward. The literature has also shown that there are gender differences in responses to information only campaings (see Muller et al. (2005), Perry et al. (1998) and Kelder et al. (1995)). The health literature highlights age effects with regards to food preferences and tastes (see Birch (1999) and the references therein); suggesting that there is likely to be differences in the effect of the incentive by age as well.

Hypothesis 2: Children will consume more fruit or vegetables when they are rewarded for taking a fruit or vegetable at lunchtime.

The behavioural literature has shown us that the default option can affect choices made by individuals (see Keller et. al. (2011), Choi et. al. (2003), and Johnson and Goldstein (2003) for examples) and even help reduce calorie consumption (Wisdom et. al. (2010)). As a result health initiatives at schools have started to require children to have a fruit or vegetable on their plate.[16] By incentivizing children to take a fruit or vegetable our

[16]See Dillon and Lane (1989) for an evaluation of the differences between offering and serving a fruit or vegetable and Just and Price (2013a) for the effect of requiring schools to serve fruit and vegetables.

experiment is likely to have a follow-through effect on consumption. Furthermore, unlike previous studies, our children have no incentive to lie or cheat regarding the amount of the fruit or vegetable they consumed; the rewards are only based on choice. This means that we can estimate the causal effect of how an increase in having a fruit or vegetable on one's lunch tray effects consumption. As with choice, there is reason to expect that the effect on consumption will vary with gender, age, and socio-economic background.

Hypothesis 3: Children will choose and consume more fruit or vegetables after the incentive is removed than before.

Given how food preferences develop, if children have been eating more fruit or vegetables during the intervention period they may have developed a preference for fruit or vegetables or developed a habit of eating fruit or vegetables at lunch time.[17] Becker and Murphy (1988) and Becker (1992) develop a model of habit formation where the marginal utility of today's consumption is correlated with historical consumption. Therefore a small change in today's behaviour - caused by an exogenous increase in the benefit of consuming a fruit or vegetable for instance - could lead to long term changes in consumption. More recently theory on present-bias (hyperbolic) preferences such as that in Laibson (1997) and O'Donoghue and Rabin (1999) suggest that providing incentives to overcome the initial costs of switching to healthy behaviour may have long lasting effects (see Acland and Levy (2013) for instance). Of course, if the extrinsic incentive replaced the intrinsic motivation that children had to eat healthily before the intervention, then after the prizes are removed we may see a decrease in the amount of fruit and vegetables chosen and consumed. Therefore, to see if there is a lasting effect (positive or negative) of the two schemes we examine choice and consumption of fruits and vegetables in the week immediately following the intervention and six months later.

5. Results

5.1 Summary Statistics

We begin by comparing our treatment and control schools in the baseline period. The upper half of Table 1 presents the means of the outcome variables and other covariates by control and both treatment groups. The final three columns show the p-values for differences between the treatments and control and between the two treatments. The p-

[17] There is some evidence that dietary habits appear to form in childhood and track into adulthood. See Kelder et. al. (1994), Resnicow et. al. (1998), and Singer et. al. (1995) for discussions.

value were calculated, to account for intra-school correlation, by regressing each baseline variable on one of the treatment indicators, and clustering the standard errors at the school level. We have 31 schools in our sample but, when looking at sub-samples, our analysis may contain less than 30 schools. Therefore, the standard clustering methods might not be appropriate. To deal with this we correct for the potential clustering problems using the the Cameron, Gelbach, and Miller (2008) wild bootstrap method with 1000 replications. The p-values shown in Table 1 are based on this cluster correction method, though, in this case, the standard clustering method gives nearly identical results.

The upper half of Table 1 shows that, for the whole sample, there are no statistically significant differences between the control group and either treatment group. We do have one significant difference when we compare the two treatments but that is far less than the seven at the 10% level we would randomly expect from conducting the 69 tests in this panel. This suggests that, based on observables, the randomization worked as expected. Furthermore, even though they are insignificant, the size of the differences (in most cases) is less than one standard deviation, suggesting that the control and treatment groups are close to being observationally equivalent in the baseline.

The lower part of the Table 1 shows the summary statistics for the sample of pupils who chose a fruit or vegetable at lunch less than 100% of the time in the baseline week. This group is of interest because they are the ones who were most able to change their behaviour due to the treatment, as opposed to those who already chose a fruit or vegetable every day. Of the 69 tests presented in this panel we only find four significant differences at the 10% level; again, this is far below the seven significant differences one would expect to occur randomly. Furthermore, as with the whole sample, the size of the differences are generally less than one standard deviation suggesting that, again, the control and treatment groups are close to being observationally equivalent in the baseline.

5.2 Descriptive Figures

We will examine the effects of the incentive schemes on both choice and consumption. The "choice" variable is a dummy equal to one if a pupil choose a fruit or vegetable on a given day. To get at consumption we will use a "try" variable which will equal one if the pupil eats at least some of a fruit or vegetable on that day.[18] Since the incentive was based on the total amount of healthy choices made in a week, we provide a descriptive

[18]We also examined the intensity of consumption by looking at whether pupils ate more than half their fruit or vegetable. The results are broadly similar to our findings with 'try' and there is the possibility of subjectivity due to lunch monitors judging what is more than half. Therefore, we include those results in the appendix for the interested reader

overview of the weekly mean outcomes for choice and consumption in Figures 1 and 2.

Figure 1 shows the effect of our treatments on choosing a fruit or vegetable. Panel (a) shows the full sample. During the baseline, pupils in control and treatment schools were choosing a fruit or vegetables with their lunch, roughly, 83% of the time. In the individual incentive scheme, to earn a small prize at the end of the week a pupil would have to choose a fruit or vegetable four times, 80% of the time. Therefore, on average, pupils already qualified for a prize in the individual incentive scheme. However, with the introduction of the incentives in week one, pupils in both treatments began to choose significantly more fruit and vegetables. Over time, though, the control group improves their eating habits and catches up to the treatment groups. In panel (b) of Figure 1 we see the effect of the treatment on pupils who did not choose fruit and vegetables 100% of the time in baseline, those with room to improve their behaviour. During baseline there is no difference in behaviour for pupils between the treatments or the control. In week one pupils who received an incentive choose fruit and vegetables more but the control group catches up quicker in this sample. Overall, this figure shows that pupils would gradually begin to make healthier choices after returning from a mid-term break, since the intervention started after the autumn holiday, but that the intervention can speed the return to healthier behaviour by getting pupils to make better choices immediately upon return to school.

Figure 2 shows the effect of the treatments on trying a fruit or vegetable. In panel (a) we again see the full sample. In the baseline there is no significant differences between the treatment and the control (refer to Table 1). The control group is much slower to improve their consumption of fruit or vegetables upon returning to school in comparison to choosing a one; they only show a small increase in week three that seems to persist in week four and the week after the treatment. However the treatments have an immediate and significant effect: pupils increase their consumption of fruit and vegetables by, roughly, 12%. After two weeks, though, the effect of the individual incentive appears to dissipate while the effect of the competition stays constant. Panel (b) shows the effects for the sample that did not choose fruit and vegetables 100% of the time in the baseline. Here we see roughly the same results as we did with choice. The interventions increase consumption immediately but the control group catches up quicker than in the overall sample. Here, though, competition may be working better and still having an effect in the last two weeks of the experiment. Overall, this figure shows that pupils are much less likely to improve their consumption of fruit and vegetables when returning from a mid-term break and that at least the competitive incentive scheme can have a positive and consistent effect in increasing consumption of fruit and vegetables.

5.3 Short and Medium Term Effects

We begin by reporting the average treatment effects for the main outcome variables of interest: *choice* and *try*. We discuss the results for the short-term (while the intervention is taking place) and the medium term (the week immediately after the intervention finishes). Our primary estimation method is a linear probability model (LPM) with pupil fixed effects (FE). This technique allows us to examine within-subject treatment effects and the comparison to the control group allows us to control for any day and week effects that might be present over the course of our field experiment.

Since the randomization was conducted at the school level it is important to cluster standard errors by school. In the overall sample, when we do not look at subgroups, we have 31 schools so standard clustering methods are possible. However, when we look at subgroups, especially age, the number of schools in our sample may drop below 30.[19] Therefore, standard clustering methods might not be appropriate. To calculate appropriate standard errors we use the Cameron, Gelbach, and Miller (2008) wild bootstrap method. In all of our result tables we report both the standard errors clustered at the school level using standard methods and the p-value from the wild bootstrap. There are very few instances where the results are different.

The dependent variable in our regressions is bounded upwards (at 1); children who choose and consumed a fruit or vegetable every day at baseline have an outcome variable equal to one and no improvement is possible for this group. Therefore, we estimate the LPM with pupil FE on the whole sample and on the sample of children who are not bounded upwards in their response, i.e. those who did not have a mean outcome equal to one in the baseline (referred to later as "Less than 100%" group). We are particularly interested in the latter group because those who are not choosing or consuming a fruit or vegetable every day is the subgroup that could most benefit from the intervention - they could be encouraged to make healthier choices.

Average treatment effects on choice

We start with the results on the whole sample in Table 2, including children who were already at the upper bound in week 1. We find little effects of either incentive scheme on choice overall (Column [1]). The point estimates for competition and the individual

[19] Some schools did not have both year two and year five or would only let one of the years participate in the field experiment.

incentive are positive but small and imprecisely estimated. When we break the sample up by gender and whether a pupil qualified for a free school meal (FSM)[20] we also find no significant effect: columns [2] and [3] split the sample by gender; columns [4] and [5] by FSM. However when we look at the results by age in columns [6] and [7] we find significant results. Column [6] shows that younger children, those in year two, respond negatively to the individual incentive: pupils *decrease* their choice of fruit and vegetables by 8% at lunchtime. Furthermore, in the week immediately after the incentive is taken away, younger pupils continue to choose less fruit and vegetables. This significantly negative effect does not show up in the overall effect because the older pupils, those in year five, respond positively to the individual incentive: they choose fruit and vegetables 16% more often than the control group.

Table 2A allows us to test whether the estimates of the effects in Table 2 are significantly different by gender, FSM status, and age. As would be expected, when we examine if the estimates for the individual incentives in column [6] are equal to those in column [7] we find that they are significantly different; older pupils respond more positively to the individual incentive than younger pupils. The comparisons by gender and FSM status, though, show no significant difference. Therefore, Tables 2 and 2A show us that the overall average treatment effect of the individual incentive on choice is masking a significant heterogeneous effect by age.

Table 2 also allows us to examine if there are differential responses to the treatment type. At the bottom of Table 2 we present the p-values for whether the estimated effect from competition equals that of the individual incentive. We find that for two groups - poorer pupils and younger pupils - the competitive incentive works better: pupils who qualify for FSM and those in Year 2 choose more fruit and vegetables in the competitive setting. These results carry over to the medium term as well. This suggests competition may be more effective at getting pupils to choose healthier items than an individual based incentive scheme.

When we consider the restricted sample - those who did not choose a fruit or vegetable every day during the baseline and, thus, have room to improve their nutritional habits - in Table 3 we find large positive and significant effects for competition in both the short and medium term but small and imprecise estimates for the individual incentive scheme. Column [1] shows that the competition increased the probability of choosing a fruit and vegetables by 17.5 percentage points and we find evidence that the effect was sustained

[20]Pupils from poorer households qualify for free school meals. Therefore, to examine the effect of the treatment on children from poorer socio-economic backgrounds, we break the sample into pupils who qualify for FSM and those that do not.

to some extent in week 6, immediately after the incentive is removed, although the size of the effect is halved to 9.6 percentage points. This means that the competition, roughly, led to pupils choosing one more fruit or vegetable per week during the intervention and one more fruit or vegetable every two weeks even after the intervention finished. The results for the individual incentive are positive but not significant in the short term.

Looking at subgroups we find that competition significantly increased the likelihood of consuming fruit or vegetable for nearly everyone (the point estimate for females is large but not significant). However, the effect of the individual incentive is mixed; there is evidence males responded positively to the incentive but we again have that younger children responded negatively and older children responded positively. Therefore, we observe the same pattern for choice with this sample as we did with the whole sample: there is a stark heterogeneous effect of the individual incentive by age. However, in this case we have the fact that the negative effect on younger children carries over into the medium term. The significance of the heterogeneous effect by age is shown in Table 3A.

When we compare the two treatments, looking at the results at the bottom of Table 3, we find that females and younger pupils responded significantly more positively to the competition than then the individual incentive.

These results suggest that competition is working well on incentivising pupils who have room to improve their choice of healthier items at lunchtime. While, even for pupils with poorer diets, the individual incentive is causing some groups to choose fruit or vegetables less often. Furthermore the positive effect of competition seems to have a lasting effect at least into the medium term by causing males and younger pupils (two key groups) along with non-FSM pupils to choose healthier items even after the incentive has been removed.

Average treatment effects on trying

We now examine our consumption variable that we call "trying" which equals one if a child ate at least part of a portion of the fruit or vegetable at lunchtime.[21] We do not condition the consumption variable or the regressions on whether a pupil choose a fruit or vegetable. Therefore the estimates in the tables below show the causal effect of the incentives on the probability that any given pupil tries a fruit or vegetable in the short and medium term.

Table 4 shows the effects on the overall sample, including those at the upper bound at baseline. Focusing first on the short term effects, we find that the competitive incentive scheme increases trying by 11.2 percentage points during the intervention (Column [1]).

[21]We also monitored whether the children at more than half the portion they were served. We report these in Tables B1 and B2, the results are very similar to what we report for trying.

We find no evidence of positive effects for the individual incentive scheme. Splitting by gender and FSM status (columns [2]-[5]) gives a similar picture as the one observed with choice: we find positive significant effects for the competitive scheme for all groups except, somewhat notably, males and we do not find significant effects for the individual incentive scheme. Similarly, when breaking the sample by age, we find positive effects of the competitive scheme on both subgroups, albeit somewhat imprecisely estimated. However, for the individual incentive, there are stark differences in the response by age. Table 4A shows that the differences we find by age are significant for the individual incentive. We estimate an increase of around 20 percentage points for the Year 5 children and a decrease of about 7 percentage points for the Year 2 children. These results provide evidence for Hypothesis 2, but the hypothesis is strongly rejected for young children. We find little evidence of persistence in week 6, except for girls and Year 2 children in the competition treatment as well as for Year 2 children in the individual incentive treatment (the latter being an adverse effect). There is evidence that the competitive incentive led to a significantly more positive response, both during the period when the incentive was in place and when the week after it was removed, among females, FSM pupils, and the younger children.

Table 5 shows the effects on trying when we restrict the sample (excluding those bounded upwards in terms of *choice* behaviour). The results are much larger but similar in nature to the results reported in Table 4. We find an overall significant increase of 21 percentage points due to the competition intervention and no significant effects of the individual incentive in the overall sample. Again, the imprecisely estimated positive effect of the individual incentive masks strong differences in response between younger and older children, with younger children responding negatively and older children responding positively. These differential effects by age are significant as seen in Table 5A. While the differences by age for competition are not significantly different.

We find stronger evidence of persistence once the incentive is removed, at least for the competitive incentive. Except for girls and Year 5 children, all effects are positive and significant. They are also quite large in magnitude: overall, the probability of trying a fruit or vegetable at lunch has increased by 14 percentage points in week 6 for children in the competition treatment. In contrast, the only persistent effect we find with the individual incentive is the adverse negative effect on Year 2 children. Comparing the two treatments we again find that female and younger pupils respond more to the competitive incentive scheme, both during the incentive period and once it had been taken away. This means that the competitive scheme, on average, caused children to choose and try more than one additional fruit or vegetable per week both during and after the treatment.

These results provide stark evidence regarding the three hypotheses by incentive scheme. There is weak and imprecise evidence that the individual incentive increases choice and consumption of fruit and vegetables (Hypotheses 1 and 2). The only significant evidence with regards to the individual incentive regarding Hypothesis 3 (the effect after the incentive is removed) is that the individual effect appears to have a lasting negative effect on younger children. Indeed the overall imprecise positive effect of the individual incentive masks the differential effect that the individual incentive has by age. However, there is a strong positive evidence that the competitive incentive encourages all pupils to choose and consume fruit and vegetables (Hypotheses 1 and 2) and that, for most groups, those effects are present when the incentive is removed (Hypothesis 3). Furthermore males and FSM pupils do respond positively to the competitive scheme (unlike under other interventions) while females, FSM pupils, and Year 2 pupils also generally respond better to the competitive scheme than the individual incentive.

Cost Effectiveness

To understand the implication of these results and what they mean for policy makers we now want to look at the costs of getting a pupil to try an additional fruit or vegetable under each scheme. Furthermore we compare the results to one other commonly used intervention to understand how each scheme compare to currently implemented programs.

The prizes for both schemes cost, in total, £3,727 and we had 413 pupils in the treatment schools. That means we spent £9 per pupil over the course of the intervention. When looking at the individual incentive for our group of interest (the less than 100% group) we find that, during the intervention, pupils increased the likelihood of trying a fruit or vegetable by 7 percentage points, though, this was imprecisely measured, and there were no medium term effects. That means that, over the first five weeks of our experiment (including medium term), pupils ate 1.5 more fruit and vegetables because of the intervention or, that it cost, roughly, £6 to get a pupils to eat an additional fruit or vegetable.

The competition scheme was more effective than the individual scheme; it increased the likelihood that, for our group of interest, the probability of trying a fruit or vegetable increased by 21 percentage points during the intervention and by 14 percentage points immediately after the incentive was removed. Thus, for the first five weeks of our experiment pupils ate 5 additional fruit or vegetables. That means it cost £1.8 to get a pupil to eat an additional fruit or vegetable. Looking at the overall sample, competition increased trying by 11 percentage points during the intervention period and 7 percentage points during the medium term. That means that, with the competition scheme, it costs £3.5 to get an average pupil (not just one from our group of interest) to eat an additional fruit or vegetable.

Are these costs large or small? To determine this we compare the results to the "Food Dudes" intervention that has been implemented in many countries (e.g. the UK, Ireland, Italy, and the USA). There have been many experimental studies done showing the effectiveness of the program but we will focus on the Horne et. al. (2009) study from Ireland because Ireland is one of the few countries to have released cost data. In Ireland the Food Dudes program had two main parts: (1) during an intervention period of four weeks schools provided fruits and vegetables[22] and showed six minute videos[23] of 'Food Dudes' eating and extolling the virtues of fruit and vegetables to save the world from the 'Junk Punks;' (2) prizes and 'Food Dude' lunchboxes were provided for bringing in and eating fruits and vegetables. The prizes were given out throughout the school year. According to the Irish government[24] implementing the programme for 60,000 children would cost €658,000 for the prizes and €503,550 for the fruit and vegetables or, roughly, €20 per pupil.

Horne et. al. (2009) find that during the intervention period (when food was being provided) pupils consumed, roughly 22 grams more of fruits and vegetables per week. Using the NHS living well proportion of 40g as a measure, this means that, over the nine month school year, pupils would have consumed nearly 9.7 more fruits and vegetables or that it costs at least £1.9 per additional fruit or vegetable consumed. This is a lower bound as these costs do not include licensing, organizational costs, etc. Indeed the Irish government puts the cost of the whole program for 60,000 pupils at over €2 million; nearly double the costs we are considering here. Therefore the upper bound on costs is £3.8 per additional fruit or vegetable consumed.

What does this comparison tell us? It shows that our competitive scheme has the potential to be as cost effective as a commonly used, multifaceted, individual incentive scheme that had to be augmented by videos, food provision, and teachers taking time to discuss the goals of the programme.[25] Indeed, this implies, that augmenting the competitive scheme with the same additions that the 'Food Dudes' programme uses with its

[22]In Ireland, generally, there is no provision of food by schools. Pupils are expected to bring in a packed lunch.

[23]See http://www.fooddudes.co.uk for examples of the videos.

[24]See "Strategy for School Fruit Scheme" submitted by Ireland for the 2012/2013 school year.

[25]While our 'trying' variable does not equate to the actual eating of fruits and vegetables as examined by Horne et. al. (2009) our 'eating more than half' results are likely to be comparable. Those results would predict the same cost effectiveness as looking at 'trying' (refer to tables B1 and B2 in the appendix).

individual incentive could have even larger results and be more cost effective.

5.4 Choice and Consumption Dynamics

Having established that there are differences in the effectiveness of the incentive schemes we now move onto explain why it might be the case the competitive scheme appears to work better in comparison to individual incentive scheme. In this section we will analyse the dynamics of choice and consumption throughout the week and as such we exclude the post incentive period. In particular we will look at if there are different dynamics during the intervention based on the two types of treatments.

First when looking at choice, the children who were most responsive to the treatments were those who had not chosen a fruit or vegetable 100% of the time during the baseline. Column [1] in Table 6 shows the effect for that sample of children.[26]

We find that competition had a large and significant effect on choice during treatment weeks; children assigned to the competition group were 17 percentage points more likely to choose a fruit or vegetable. There was a large imprecisely estimated effect due to individual incentive. Columns [2]-[6] show the effect of the treatments for each day of the week. The effect of the competitive scheme started off very strong at the beginning of the week; on Mondays and Tuesdays children were 24 and 25 percentage points, respectively, more likely to choose a fruit or vegetable. As the week went on the effect dissipated, though; the point estimate decreased from 18 percentage points on Wednesday to 6 percentage points on Friday (the latter estimate not being significant). The individual incentive had the opposite effect; children were more likely to choose their fruit or vegetable at the end of the week. The only significant increase in choice due to the individual incentive treatment took place on Friday when children were 27 percentage points more likely to choose a fruit or vegetable.

In the competitive scheme children did not know how many fruit or vegetables they would have to choose to get a prize at the end of the week; if they choose five fruit or vegetables, though, they were guaranteed a prize. Since children did not know who was in their group and some children did not choose a fruit or vegetable every day, a pupil could assign a subjective probability to winning given how many items she had chosen during the week.[27] Based on a pupil's subjective probability one could calculate the number of

[26]There was no effect - either positive or negative - on the sample of children that had chosen a fruit or vegetable 100% of the time during the baseline week. The effect on all children is just a weighted average of these two groups.

[27]In fact there was an increasing probability of winning the prize based on the number of fruit and vegetables one chose. There was a small probability (under 5%) chance of winning if a pupil had chosen zero or one item, a 6.7% chance of winning if a pupil chose two items, a 21% chance of winning if a pupil

fruit or vegetables that a pupil would ideally want to consume each week to maximise her benefit from getting a prize subject to her disutility from having to choose a fruit or vegetable. Once a pupil has reached that number of fruit or vegetables she could switch back to her preferred unhealthy item. This type of pattern would explain why the effect of competition tapered off during the week.

In the individual scheme the threshold to obtain the weekly prize was known and fixed. Given the exogenous pre-determined goal a pupil had to reach there was room for discouragement to take place; if a pupil had not eaten a fruit or vegetable on Monday and Tuesday then there was zero probability the pupil would get a prize that week. Besides having no external incentive from Wednesday onwards, a pupil might also feel discouraged and choose not to select a healthy option. Therefore, to examine this discouragement effect we break the sample into two groups in columns [7] and [8]. Column [7] contains children who had 'missed' the prize as of Wednesday, i.e. they had not chosen a fruit or vegetable on Monday and Tuesday. Column [8] contains those children who had chosen at least one fruit or vegetable before Wednesday. The effect of individual incentive is large and significant for those who still have a chance of getting a prize, i.e. those in column [8]. However, for those that have missed the chance of getting a prize the effect of individual incentive is estimated to be negative, though, it is insignificant. This means that as the week goes on the incentive to choose a fruit or vegetable wears off for those that miss the goal in the individual incentive scheme. However, this is not the case in the competition treatment because there is always a positive probability of winning the prize no matter how many items the pupil has consumed during the week.[28]

These results speak to the intrinsic incentive differences between the two treatments. The external, known goal in the individual scheme can lead to a lack of incentive because of previous choice patterns. However, there is always a positive chance of winning in the competition treatment because the goal is unknown and endogenous to the system. In the habit formation literature with regards to healthy eating the goals have all been exogenous and known. Therefore, there is room to design rewards like the competitive scheme that can have a greater effect (than an individual scheme) over the same period of time.

The effect of the competitive scheme on consuming at least part of a fruit or vegetable is similar to what we found for choice. Table 7, Columns [2]-[6] shows again a large

chose three items, and a 39% chance of winning if a pupil chose 4 items.

[28]Indeed we cannot reject that the point estimates for competition are the same in columns [7] and [8] showing that the choice pattern before Wednesday does not change the effect that the competition treatment has from Wednesday onwards. However we can reject that the point estimates in columns [7] and [8] are the same in the individual incentive scheme.

positive effect of competition that is relatively constant but drops off slightly on Friday. The individual incentive only has a significant effect on Friday, and again when comparing children who missed the chance to win a prize and those who are still eligible (columns [7] and [8]), we find that the individual incentive has a positive significant effect only for the latter group. Also, the point estimate for competition is not significantly different between columns [7] and [8]. This means that previous choice patterns in the week do not effect consumption choices later in the week systematically, unlike for the individual incentive treatment.

Summarising, we find that each incentive scheme is associated with different dynamics of choice and consumption behaviour. The competition works throughout the week, while the individual incentive only has an end-of-the-week effect. This effect is particularly pronounced for children who still have the chance to win a prize, while it is basically zero for those who know they have already forgone the chance to win a prize by Wednesday. These differences is choice and consumption are, thus, likely due to the way the goals are defined; the known constant goal of the individual incentive causing discouragement and the unknown endogenous goal of the competitive treatment providing at least some positive incentive to choose a fruit or vegetable every day.

5.5 Long term effects

To evaluate whether the effects we find lead to permanent changes in habits, we contacted the schools again 6 months later and asked them to conduct an additional week of monitoring; 21 out of the 31 schools agreed to conduct an additional week of monitoring.[29] To get at the longer run effects we redid the analysis presented in the section 5.3 on that selected sample only. In creating those tables we included an additional interaction term of the treatment with an indicator denoting 6 months later. For brevity, in Tables 8 (choice) and 9 (trying), we only present this additional interaction term. In both tables panel A shows is for the overall sample and panel B is for the restricted sample.

We find little evidence of any persistence 6 months later on the overall sample or in the restricted sample. In Table 8 for choice, the largest positive point estimates for both samples occur for the free school meal registered pupils in the competition scheme (column

[29]To be sure that the sample used for the long-term analysis is not a positively selected sample (of schools that have had a positive experience with the incentive schemes in particular) we ran the previous analysis on the subset of 21 schools to check the selection. The results are very similar in nature to the ones found with the whole sample (Tables 2 - 5), so we are confident that the long-term results are not driven by selection. We also recreated the descriptive table, Table 1, and found similar results, i.e. no significant differences between treatments and control or the treatments. The results are not reported here but are available upon request.

[4]). However, this is a small group and the estimates are imprecise. We do not find any significant differences across groups and only one significant difference across treatments; the wild p-value is not significant for any estimate, though. Turning to trying in Table 9, again the largest point estimates we find are for the free school registered group, but again they imprecisely estimated. We do find a significant difference for the overall sample (Panel A) between the treatments for the year 5 pupils. With the individual incentive scheme having a larger effect than in the long run than the competitive scheme. We also find a significant estimate for FSM pupils in the less than 100% group for the individual incentive scheme. However, given the wild p-value for the estimate is 0.651 and that the individual incentive scheme never had a significant effect or a positive point estimate above 0.027 for FSM pupils in the previous analysis, this estimate does not provide any strong evidence for a longer term effect. Overall, we find little, if any, evidence for long run effects as a result of either of the treatments. This means there is little evidence for Hypothesis 3 with regards to the longer term.

5.6 Learning: Food Knowledge

One question is whether the intervention triggered a response only through the incentives, or also through learning. It could be that the intervention taught children that fruit and vegetables are healthy and that they respond to that information rather than the incentives. We are able to test for this possibility by comparing the results on a knowledge test that was conducted just before and at the end of the intervention. The test shows pictures of seven food items, including three or four fruit or vegetables and unhealthy items (such as sweets, chips, ice cream, crisps, fish fingers). On the test children were asked to identify what the item was and whether the item was healthy or not (see Figure A2 for an example). On average, we find that children described 92% items correctly as healthy or not and were able to identify 83% of the items correctly before the intervention.

We estimate a simple linear model with the change in the test score of identifying items correctly as the dependent variable and include indicators for the two treatment groups. The results are presented in Table 10 for the whole sample and in Table 11 for the sample of children who chose less than 100% in the first week. The effects across group are not consistent and we fail to find evidence that the scores improved more in the treated schools than in the control schools. If anything, we find negative effects for the children in the individual incentive group (restricted sample). We only find a positive significant effect for the Year 5 children in the competition treatment. These results indicate that knowledge was very high before the intervention and that the positive effects we find on choice and trying are not due to improvement in knowledge. Children know very well that fruit and vegetables are healthy and we can safely rule out the hypothesis that the responses to the intervention are driven by learning.

5.7 Effects on other outcomes

An additional exercise we propose is to check whether the interventions affected other relevant outcomes that could partially explain the treatment effects we found. These results are reported in Appendix B.

A first outcome of interest is attendance. One concern could be that the prospect of receiving (or not) a reward may affect attendance rates. We investigate this possibility in Tables B3 and B4. Table B3 reports results for the whole sample, while Table B4 reports results for the less than 100% sample). We do not find any significant effect on attendance overall or by sub-group. We do find positive and significant effects on attendance for males in the individual incentive scheme for the restricted sample. However, in the main results we do not find positive and significant effects of the individual incentive for boys when looking at either choice or try. Thus, these effects appear to be difficult to reconcile with the treatment effects we found. We conclude that changes in attendance rates are unlikely to drive the treatment effects on choice and consumption.

A second outcome that seems worth considering is whether children are more or less likely to bring a packed lunch as a result of the intervention. This would not be a confounding factor though. But it would provide some information regarding how children adjusted to the introduction of the incentive schemes. For example, pupils may have put pressure on their parents to provide a packed lunch if they do not like the fruits or vegetables on offer at school. Table B5 and B6 report the results. We find no evidence that children were more or less likely to bring a packed lunch overall. In the restricted sample, we find a positive and significant effect for males in the competitive scheme for week 6 but not while the intervention is actually taking place. This means that the treatment effects we find are driven by children changing their behaviour within the meal context they started with (packed lunch or school meal).

6. Conclusion

This paper provides field evidence on how two incentive scheme change how children choose and consume fruit and vegetables at lunchtime. We conducted a large scale field experiment in 31 primary schools in England testing for the effects of two different incentive schemes: a competition and an individual incentive scheme. Both schemes lasted 4 weeks and we monitored choice and consumption of fruit and vegetables by children made over that period, as well as one week before, one week after and 6 months later.

We find two main results. First, competitive and individual incentives have very different effects and one cannot draw a unique conclusion on whether incentives work or not. The competitive incentive is overall more effective and more robust. Children respond positively to the competition and increase their choice and consumption of fruit and vegetables. The individual incentive, in contrast, has very heterogeneous effects. Older children respond positively, while younger children are affected negatively. Second, we do find evidence that the intervention continues to affect behaviour after the incentives are removed. However, we find little evidence of behaviour change six months later; the effects of the temporary incentive appear to be short lived.

Overall our results show the need to study various forms of incentive schemes as it is not clear that incentives will work in the same way for different subgroups of the population. In particular, an exogenous, know incentive can lead some groups to become discouraged. In terms of policy implications, our findings suggest that the competitive incentive is more effective overall, while the individual incentive can have adverse effects on some subgroups of children. But we also advocate for more research, particularly using field experiments, to investigate in more detail how incentive schemes work and for whom.

References

Acland, D. and Levy, M. (2013). 'Naivete, Projection Bias, and Habit Formation in Gym Attendance', Available at SSRN: http://ssrn.com/abstract=2233004

Becker, G.S. (1992). 'Habits, Addictions, and Traditions', *Kyklos*, vol. 45(3), pp. 327-45

Becker, G.S. and Murphy, K.M. (1988). 'A Theory of Rational Addiction', *Journal of Political Economy*, vol. 96, pp. 675-700

Behavioral Insights Team. (2010). 'Applying behavioural insight to health', UK Cabinet Office Behavioural Insights Team.

Belot M., and James J. (2013). 'Partner Selection into Policy Relevant Field Experiments'. Mimeo

Bettinger, E. and Slonim, R. (2007). 'Patience in children: Evidence from experimental economics', *Journal of Public Economics*, vol. 91(1-2), 343-363

Bhattacharya, J. and Sood, N. (2011). 'Who Pays for Obesity?' *Journal of Economic Perspectives*, vol. 25(1), pp. 139-58.

Birch, L.L. (1999) 'Development of food preferences', *Annual Review of Nutrition*, vol 19, pp. 41-62.

Birch, L.L., Birch, D., Marlin, D.W., and Kramer, L. (1982). 'Effects of instrumental consumption on children's food preference', *Appetite*, vol. 3(2), pp. 125-134.

Birch, L.L., Gunder, L., and Grimm-Thomas, K., and Laing, D.G. (1998). 'Infants' consumption of a new food enhances acceptance of similar foods', *Appetite*, vol. 30(3), pp. 283-295.

Birch, L.L. and Marlin, D.W. (1982). 'I don't like it; I never tried it: effects of exposure on two-year-old children's food preferences', vol. 3(4), pp. 353-360.

Birch, L.L., McPhee, L., Shoba, B.C., Pirok, E., and Steinberg, L. (1987). 'What kind of exposure reduces children's food neophobia? Looking vs. tasting', *Appetite*, vol. 9(3), pp. 171-178.

Birch, L.L., Marlin, D.W., and Rotter, J. (1984). 'Eating as the "Means" Activity in a Contingency: Effects on Young Children's Food Preference', *Child Development*, vol. 55(2), pp. 431-439.

Booth, A. and Nolen, P. (2012). 'Choosing to compete: How different are girls and boys?', *Journal of Economic Behavior & Organization*, vol. 81(2), pp. 542-555.

Cameron, A.C., Gelbach, J.B., and Miller, D.L. (2008). 'Bootstrap-Based Improvements for

Inference with Clustered Errors' *The Review of Economics and Statistics*, vol. 90(3), pp. 414-427.

Charness, G. and Gneezy, U. (2009). 'Incentives to Exercise', *Econometrica*, vol 77(3), pp. 909-931.

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

Ciliska, D., Miles, E., O'brien, M.A., Turl, C., Tomasik, H.H., Donovan, U., and Beyers, J. (2000). 'Effectiveness of Community-Based Interventions to Increase Fruit and Vegetable Consumption', *Journal of Nutrition Education*, vol. 32(6), pp. 341-352.

Cooke, L., Wardle, J., and Gibson, E.L. (2003). 'Relationship between parental report of food neophobia and everyday food consumption in 2-6-year-old children', *Appetite*, 41, 205-206.

Delaney, L. and Doyle, O.M. (2012). 'Socioeconomic differences in early childhood time preferences', *Journal of Economic Psychology*, vol. 33(1), pp. 237-247.

Dillon, M.S., Lane, H.W. (1989). 'Evaluation of the offer vs. serve option within self-serve, choice menu lunch program at the elementary school level', *Journal of the American Diet Association*, vol. 89(12), pp. 1780-1785.

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

Gneezy, U. and Rustichini, A. (2000). 'A Fine is a Price', *The Journal of Legal Studies*, vol. 29(1), pp. 1-17

Gneezy, U. and Rustichini, A. (2004). 'Gender and Competition at a Young Age', *American Economic Review*, vol. 94(2), pp. 377-381.

Gneezy, U., Niederle, M., and Rustichini, A. (2003). 'Performance in Competitive Environments: Gender Differences', *Quarterly Journal of Economics*, vol 118(3), pp. 1049-1074.

Gneezy, U., Meier, S., and Rey-Biel, P. (2011). 'When and Why Incentives (Don't) Work to Modify Behavior', *Journal of Economic Perspectives*, vol. 25(4), pp. 191-210.

Horne, P.J., Lowe, C.F., Bowdery, M., and Egerton, C. (1998). 'The way to healthy eating for children', *British Food Journal*, vol. 100(3), pp. 133-140.

Horne, P.J., Lowe, C.F., Fleming, P.F., and Dowey, A.J. (1995). 'An effective procedure for changing food preferences in 5-7 year-old children', *Proceedings of the Nutrition Society*, vol. 54, pp. 441-452.

Iriberri, N., and Rey-Biel, P. (2011). ‘Let’s (not) talk about sex: The effect of information provision on gender differences in performance under competition’, Economics Working Papers 1288, Department of Economics and Business, Universitat Pompeu Fabra.

Johnson, E.J. and Goldstein, D. (2003). ‘Do Defaults Save Lives?’, *Science*, vol. 302(5649), pp. 1338-1339.

Just, D.R. and Price, J. (2013). ‘Using Incentives to Encourage Healthy Eating in Children’, *Journal of Human Resources*, vol. 48(4), pp. 855-872.

Just, D., and Price, J. (2013a). ‘Default options, incentives and food choices: evidence from elementary-school children’, *Public Health Nutrition*, vol. 16(12), pp. 2281-2288.

Kelder, S.H., Perry, C.L., Klepp, K.I., and Lytle, L.L. (1994). ‘Longitudinal tracking of adolescent smoking, physical activity, and food choice behaviors’, *American Journal of Public Health*, vol. 84(7), pp. 1121-1126.

Kelder, S.H., Perry, C.L., Lytle, L.A., and Klepp, K.I. (1995). ‘Community-wide youth nutrition education: long-term outcomes of the Minnesota Heart Health Program’, *Health Education Research*, vol. 10(2), pp. 119-31.

Keller, P.A, Harlam, B., Loewenstein, G., and Volpp, K.G. (2011). ‘Enhanced active choice: A new method to motivate behavior change’, *Journal of Consumer Psychology*, vol. 21, Issue 4, pp. 376-383.

Laibson, D. (1997). ‘Golden Eggs and Hyperbolic Discounting’, *Quarterly Journal of Economics*, vol. 112(2), pp. 443-477.

List, J.A. and Samek, A.S. (2014). ‘The Behavioralist as Nutritionist: Leveraging Behavioral Economics to Improve Child Food Choice and Consumption’, NBER Working Paper No. 20132.

Muller, M.J., Danielzik, S., and Pust S. (2005). ‘School- and family-based interventions to prevent overweight in children’, *The Proceedings of the Nutrition Society*, vol. 64(2), pp. 249-54.

Newman, J. and Taylor, A. (1992). ‘Effect of a means-end contingency on young children’s food preferences’, *Journal of Experimental Child Psychology*, vol. 53(2), pp. 200-216.

NICE Citizens Council Meeting, (2010), The use of incentives to improve health,
http://www.nice.org.uk/media/9AF/56/CCReportIncentives.pdf

O’Donoghue, T. and Rabin, M. (1999). ‘Doing It Now or Later’, *American Economic Review*, vol. 89(1), pp. 103-124.

Perry, C.L., Bishop, D.B., Taylor, G., Murray, D.M., Mays, R.W., Dudovitz, B.S., Smyth, M., and Story, M. (1998). 'Changing fruit and vegetable consumption among children: the 5-a-Day Power Plus program in St. Paul, Minnesota', *American Journal of Public Health*, vol. 88(4), pp. 603-609.

Rabin, M. (2011). 'Healthy Habits: Some Thoughts on The Role Of Public Policy in Healthful Eating and Exercise Under Limited Rationality', mimeo.

Resnicow, K., Smith, M., Baranowski, T., Baranowski, J., Vaughan, R., and Davis, M. (1998), '2-year tracking of children's fruit and vegetable intake', *Journal of the American Dietetic Association*, vol 98(7), pp. 785-789.

Roberson, R., (2008). 'Using Information to Promote Healthy Behaviours', King's Fund Report.

Singer, M.R., Moore, L.L., Garrahie, E.J., and Ellison, R.C. (1995). 'The tracking of nutrient intake in young children: the Framingham Children's Study', *American Journal of Public Health*, vol. 85(12), pp. 1673-1677.

Sullivan, S.A. and Birch, L.L. (1990). 'Pass the sugar, pass the salt: Experience dictates preference', *Developmental Psychology*, vol. 26(4), pp. 546-551.

Verplanken, B. and Wood, W. (2006). 'Helping Consumers Help Themselves: Improving the Quality of Judgments and Choices', *Journal of Public Policy & Marketing*, vol. 25(1), pp. 90-103.

Volpp, K.G., Troxel, A.B., Pauly, M.V., Glick, H.A., Puig, A., Asch, D. A., Galvin, R., Zhu, J., Wan, F., DeGuzman, J., Corbett, E., Weiner, J., and Audrain-McGovern, J. (2009). 'A randomized, controlled trial of financial incentives for smoking cessation', *The New England Journal of Medicine*, vol. 360, pp. 699-709.

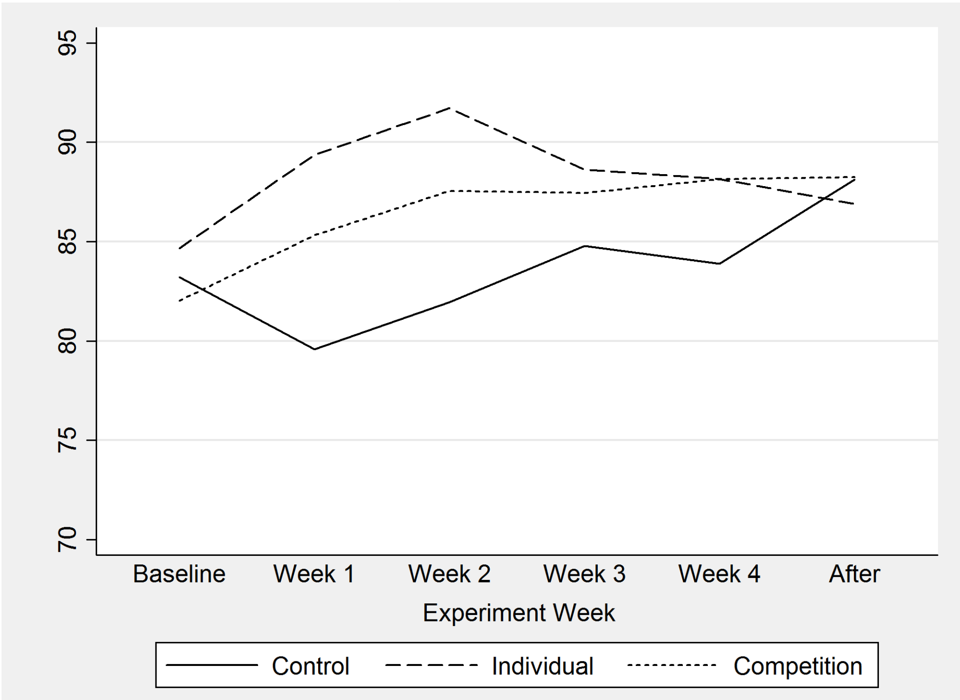
Wisdom, J., Downs, J.S., and Loewenstein, G. (2010). 'Promoting Healthy Choices: Information versus Convenience', *American Economic Journal: Applied Economics*, vol. 2(2), pp. 164-78.

World Health Organization. (2009). *Global health risks: mortality and burden of disease attributable to selected major risks.* Geneva: World Health Organization

Figures & Tables

Figure 1: Proportion of pupils choosing a fruit or vegetable

a) Full Sample



b) Sample with less than 100% Choice in Baseline

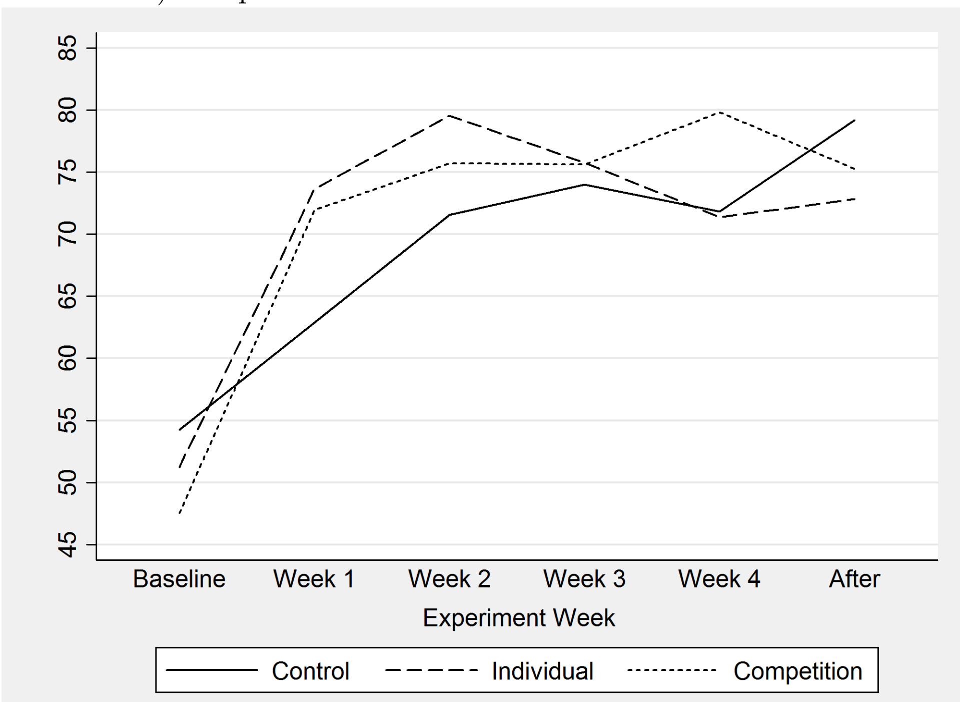


Figure 2: Proportion of pupils trying a fruit or vegetable

a) Full Sample

b) Sample with less than 100% Choice in Baseline

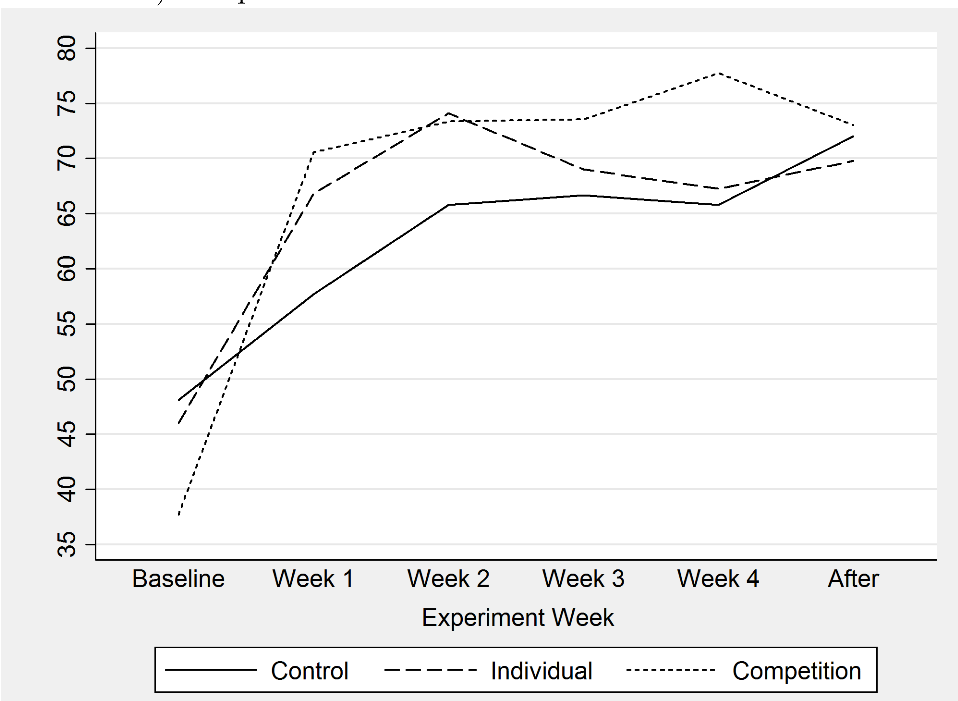


Table 1: Summary Statistics Control and Treatment Groups

	Control (C)	Individual (T1)	Comp. (T2)	Ctrl vs Ind C vs T1	Ctrl Vs Comp C vs T2	Comp Vs Ind T1 vs T2
Panel A: All Pupils						
Choice	0.841	0.847	0.821	0.925	0.769	0.713
Try	0.739	0.769	0.72	0.721	0.815	0.599
Eat more than half	0.554	0.618	0.614	0.352	0.571	0.985
Female %	0.513	0.438	0.558	0.188	0.414	0.040
English first language	0.977	0.983	0.931	0.945	0.244	0.152
White British	0.905	0.926	0.805	0.771	0.322	0.254
Year 2	0.5	0.537	0.619	0.835	0.286	0.647
Free School Meal %	0.206	0.197	0.154	0.901	0.406	0.515
School Dinner %	0.52	0.453	0.479	0.539	0.699	0.795
Packed Lunch %	0.479	0.547	0.521	0.531	0.671	0.795
Special dietary requirements %	0.053	0.097	0.128	0.162	0.132	0.699
Specific health cond. %	0.144	0.167	0.161	0.561	0.585	0.887
Ofsted overall score	2.066	1.875	2.206	0.418	0.569	0.244
Ofsted health score	1.396	1.536	1.424	0.633	0.971	0.667
Per pupil expenditure	4097	4126	3816	0.941	0.370	0.280
Catering costs	112.1	94.1	62.6	0.573	0.236	0.336
Food for life status	0.205	0.395	0.173	0.364	0.815	0.292
Headcount girls	106.4	122.1	122.8	0.667	0.362	0.979
Headcount boys	114.3	138.0	131.2	0.625	0.358	0.875
Average point score	0.288	0.28	0.283	0.144	0.272	0.731
Achieving Level 4 or > in Eng/Maths	0.815	0.789	0.751	0.607	0.200	0.571
Persistent Absence	0.024	0.017	0.021	0.671	0.831	0.693
Absence	0.054	0.051	0.054	0.569	0.959	0.677
Panel B: Restricted sample (Chose less than 100% Choice in baseline week)						
Choice	0.545	0.515	0.477	0.735	0.464	0.639
Try	0.455	0.458	0.375	0.977	0.388	0.300
Eat more than half	0.329	0.356	0.323	0.715	0.929	0.675
Female	0.396	0.419	0.575	0.769	0.064	0.084
1st Language English	0.961	0.965	0.946	0.889	0.777	0.659
White British	0.854	0.944	0.784	0.262	0.617	0.202
Year 2	0.382	0.303	0.624	0.771	0.048	0.348
Free School Meal %	0.154	0.102	0.162	0.635	0.947	0.533
School Dinner %	0.441	0.371	0.558	0.729	0.452	0.302
Packed Lunch %	0.556	0.629	0.442	0.723	0.456	0.302
Special dietary requirements %	0.028	0.108	0.177	0.104	0.072	0.350
Specific health cond. %	0.179	0.228	0.128	0.625	0.482	0.236
Ofsted overall score	2.169	2.079	2.263	0.613	0.759	0.422
Ofsted health score	1.346	1.485	1.468	0.815	0.749	0.965
Per pupil expenditure	3727	3919	3743	0.282	1.009	0.521
Catering costs	84.2	77.1	40.5	0.823	0.112	0.188
Food for life status	0.244	0.062	0.124	0.545	0.667	0.675
Headcount girls	111.1	120.0	119.1	0.603	0.671	0.947
Headcount boys	116.3	133.2	127.5	0.434	0.595	0.773

Continued on next page

Table 1 – *Continued from previous page*

	Control (C)	Individual (T1)	Comp. (T2)	Ctrl vs Ind C vs T1	Ctrl Vs Comp C vs T2	Comp Vs Ind T1 vs T2
Average point score	0.287	0.289	0.283	0.677	0.306	0.156
Achieving Level 4 or > in Eng/Maths	0.838	0.827	0.752	0.813	0.152	0.138
Persistent Absence	0.017	0.011	0.018	0.667	0.847	0.482
Absence	0.052	0.047	0.053	0.539	0.915	0.490

notes: All variables are evaluated for the first week, before the start of the treatment. The first column shows the means for the pupils in the control school in the, the second column for schools in the individual incentive scheme and the third column in the competition schools. The fourth and fifth columns show the p-value difference in the means of each treatment compared to the control group. The p-value were calculated, to account for intra-school correlation, by regressing each baseline variable on one of the treatment indicators, standard errors are clustered at the school level and due to the small number clusters we present wild bootstrapped p-values using 1000 replications which are estimated following Cameron, Gelbach, Miller (2008), the p-value is matched to the t-statistic on the treatment dummy.

Table 2: Effect on Choice for Overall Sample and Its Subgroups

	Dependent Variable (=1) if Student Chose a Healthy Item						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Competition (=1) * Week 2-5	0.045	0.059	0.026	0.071	0.045	0.057	0.023
	(0.031)	(0.036)	(0.049)	(0.065)	(0.032)	(0.043)	(0.048)
	[0.180]	[0.144]	[0.739]	[0.352]	[0.164]	[0.246]	[0.667]
Competition (=1) * Week 6	0.001	0.027	-0.030	0.002	0.003	0.040	-0.051
	(0.034)	(0.044)	(0.029)	(0.100)	(0.029)	(0.033)	(0.065)
	[0.955]	[0.595]	[0.390]	[1.00]	[0.889]	[0.294]	[0.492]
Individual Incentive (=1) * Week 2-5	0.024	0.010	0.037	-0.033	0.033	-0.066**	0.126*
	(0.050)	(0.045)	(0.061)	(0.052)	(0.053)	(0.027)	(0.072)
	[0.659]	[0.863]	[0.549]	[0.537]	[0.515]	[0.034]	[0.236]
Individual Incentive (=1) * Week 6	-0.045	-0.045	-0.051	-0.164	-0.027	-0.122***	0.048
	(0.059)	(0.058)	(0.063)	(0.114)	(0.059)	(0.036)	(0.083)
	[0.567]	[0.450]	[0.486]	[0.166]	[0.701]	[0.004]	[0.641]
Constant	0.821***	0.843***	0.798***	0.838***	0.819***	0.852***	0.788***
	(0.014)	(0.014)	(0.020)	(0.021)	(0.015)	(0.013)	(0.022)
Sample	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
P-Value: Comp = Ind Incentive Week 2-5	0.698	0.278	0.875	***0.088***	0.837	***0.012***	0.198
P-Value: Comp = Ind Incentive Week 2-5 (wild)	0.711	0.276	0.809	0.108	0.859	***0.020***	0.340
P-Value: Comp = Ind Incentive Week 6	0.415	0.218	0.733	***0.071***	0.606	***0.000***	0.273
P-Value: Comp = Ind Incentive Week 6 (wild)	0.396	0.222	0.755	***0.068***	0.627	***0.002***	0.364
Observations	15,338	7,986	7,352	2,664	12,256	8,033	7,305
R-squared	0.007	0.009	0.006	0.021	0.006	0.011	0.014
Number of pupils	638	328	310	114	509	343	295

*notes: Robust standard errors clustered at school level are included in parentheses; * sig at 10%, ** sig at 5%, *** sig at 1%. Wild P-Values are shown in brackets and are estimated following Cameron, Gelbach, Miller (2008) using 1000 replications. Column [2] includes only female and column [3] contains only males. Column [4] includes only pupils who are eligible for free school meals (FSM) and column [5] contains those pupils not eligible for FSM; there are 15 pupils for whom we are missing FSM status. Column [6] contains only pupils in Year 2 and column [7] contains only pupils in Year 5.*

Table 2A: Tests for Differences Between Subgroups			
	Column [2] = [3]	Column [4] = [5]	Column [6] = [7]
Competition (=1) * Week 2-5	0.577	0.686	0.611
Competition (=1) * Week 2-5 (wild-p)	0.595	0.681	0.687
Competition (=1) * Week 6	0.164	0.985	0.216
Competition (=1) * Week 6 (wild-p)	0.186	1.019	0.240
Individual Incentive (=1) * Week 2-5	0.543	0.316	***0.020***
Individual Incentive (=1) * Week 2-5 (wild-p)	0.571	0.316	***0.076***
Individual Incentive (=1) * Week 6	0.871	0.269	***0.067***
Individual Incentive (=1) * Week 6 (wild-p)	0.893	0.322	0.132
First Group in Column Heading	Female	FSM	Year 2
Second Group in Column Heading	Male	Non-FSM	Year 5

notes: The table contains the p-values for the tests of whether the coefficient on the variables in Table 2 for the two columns listed are equal. Wild p-values shown are estimated following Cameron, Gelbrach, Miller (2008) using 1000 replications.

Table 3: Effect on Choice for Sample with Week 1 less than 100% Choice and Its Subgroups

	Dependent Variable (=1) if Student Chose a Healthy Item						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Competition (=1) * Week 2-5	0.175***	0.108	0.214***	0.256*	0.165***	0.157*	0.160**
	(0.060)	(0.081)	(0.073)	(0.131)	(0.057)	(0.076)	(0.068)
	[0.018]	[0.302]	[0.002]	[0.112]	[0.016]	[0.176]	[0.042]
Competition (=1) * Week 6	0.096**	0.058	0.111**	0.085	0.094**	0.110*	0.060
	(0.043)	(0.064)	(0.053)	(0.152)	(0.037)	(0.057)	(0.068)
	[0.048]	[0.370]	[0.126]	[0.723]	[0.020]	[0.174]	[0.456]
Individual Incentive (=1) * Week 2-5	0.096	-0.014	0.173*	0.027	0.088	-0.194***	0.231***
	(0.080)	(0.095)	(0.095)	(0.188)	(0.071)	(0.068)	(0.076)
	[0.340]	[0.871]	[0.260]	[0.847]	[0.382]	[0.108]	[0.032]
Individual Incentive (=1) * Week 6	-0.035	-0.104	0.010	-0.298	-0.023	-0.389***	0.109
	(0.094)	(0.086)	(0.116)	(0.351)	(0.084)	(0.068)	(0.082)
	[0.687]	[0.200]	[0.961]	[0.727]	[0.765]	[0.000]	[0.212]
Constant	0.517***	0.540***	0.495***	0.459***	0.527***	0.511***	0.523***
	(0.024)	(0.026)	(0.030)	(0.054)	(0.022)	(0.025)	(0.025)
Sample	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
P-Value: Comp = Ind Incentive Week 2-5	0.371	0.170	0.721	0.260	0.348	***0.000***	0.383
P-Value: Comp = Ind Incentive Week 2-5 (wild)	0.428	0.168	0.755	0.490	0.346	***0.014***	0.468
P-Value: Comp = Ind Incentive Week 6	0.191	***0.069***	0.426	0.288	0.189	***0.000***	0.559
P-Value: Comp = Ind Incentive Week 6 (wild)	0.204	***0.050***	0.436	0.639	0.182	***0.000***	0.593
Observations	5,586	2,641	2,945	802	4,587	2,369	3,217
R-squared	0.054	0.067	0.046	0.089	0.047	0.065	0.061
Number of pupils	215	102	113	29	179	93	122

*notes: Robust standard errors clustered at school level are included in parentheses; * sig at 10%, ** sig at 5%, *** sig at 1%. Wild P-Values are shown in brackets and are estimated following Cameron, Gelbach, Miller (2008) using 1000 replications. Column [2] includes only female and column [3] contains only males. Column [4] includes only pupils who are eligible for free school meals (FSM) and column [5] contains those pupils not eligible for FSM; there are 15 pupils for whom we are missing FSM status. Column [6] contains only pupils in Year 2 and column [7] contains only pupils in Year 5.*

Table 3A: Tests for Differences Between Subgroups			
	Column [2] = [3]	Column [4] = [5]	Column [6] = [7]
Competition (=1) * Week 2-5	0.240	0.456	0.972
Competition (=1) * Week 2-5 (wild-p)	0.276	0.573	0.911
Competition (=1) * Week 6	0.473	0.951	0.570
Competition (=1) * Week 6 (wild-p)	0.529	0.907	0.637
Individual Incentive (=1) * Week 2-5	***0.072***	0.729	***0.000***
Individual Incentive (=1) * Week 2-5 (wild-p)	0.154	0.733	***0.002***
Individual Incentive (=1) * Week 6	0.205	0.444	***0.000***
Individual Incentive (=1) * Week 6 (wild-p)	0.252	0.611	***0.002***
First Group in Column Heading	Female	FSM	Year 2
Second Group in Column Heading	Male	Non-FSM	Year 5

notes: The table contains the p-values for the tests of whether the coefficient on the variables in Table 2 for the two columns listed are equal. Wild p-values shown are estimated following Cameron, Gelbrach, Miller (2008) using 1000 replications.

Table 4: Effect on Trying for Overall Sample and Its Subgroups

	Dependent Variable (=1) if Student Tried a Healthy Item						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Competition (=1) * Week 2-5	0.112**	0.142***	0.073	0.195**	0.099**	0.116*	0.105*
	(0.049)	(0.051)	(0.069)	(0.088)	(0.047)	(0.059)	(0.054)
	[0.022]	[0.012]	[0.456]	[0.080]	[0.036]	[0.084]	[0.114]
Competition (=1) * Week 6	0.067	0.099*	0.027	0.156	0.050	0.097*	0.032
	(0.050)	(0.052)	(0.062)	(0.107)	(0.043)	(0.047)	(0.069)
	[0.210]	[0.110]	[0.799]	[0.260]	[0.282]	[0.070]	[0.671]
Individual Incentive (=1) * Week 2-5	0.033	0.021	0.042	-0.024	0.043	-0.073*	0.199***
	(0.058)	(0.053)	(0.077)	(0.080)	(0.059)	(0.041)	(0.066)
	[0.587]	[0.707]	[0.623]	[0.763]	[0.557]	[0.124]	[0.0961]
Individual Incentive (=1) * Week 6	-0.025	-0.025	-0.028	-0.125	-0.012	-0.121**	0.130
	(0.072)	(0.069)	(0.085)	(0.131)	(0.068)	(0.044)	(0.096)
	[0.869]	[0.723]	[0.753]	[0.386]	[0.855]	[0.016]	[0.282]
Constant	0.736***	0.760***	0.711***	0.759***	0.734***	0.769***	0.692***
	(0.019)	(0.018)	(0.026)	(0.028)	(0.019)	(0.017)	(0.022)
Sample	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
P-Value: Comp = Ind Incentive Week 2-5	0.251	***0.041***	0.730	***0.010***	0.418	***0.002***	0.247
wild P-Value: Comp = Ind Incentive Week 2-5	0.286	***0.068***	0.807	***0.020***	0.464	***0.002***	0.378
P-Value: Comp = Ind Incentive Week 6	0.164	***0.054***	0.484	***0.012***	0.323	***0.000***	0.256
wild P-Value: Comp = Ind Incentive Week 6	0.220	***0.080***	0.565	***0.016***	0.326	***0.000***	0.328
Observations	14,714	7,719	6,994	2,495	11,838	7,916	6,798
R-squared	0.012	0.018	0.008	0.026	0.011	0.015	0.023
Number of pupils	638	328	310	114	509	343	295

*notes: Robust standard errors clustered at school level are included in parentheses; * sig at 10%, ** sig at 5%, *** sig at 1%. Wild P-Values are shown in brackets and are estimated following Cameron, Gelbach, Miller (2008) using 1000 replications. Column [2] includes only female and column [3] contains only males. Column [4] includes only pupils who are eligible for free school meals (FSM) and column [5] contains those pupils not eligible for FSM; there are 15 pupils for whom we are missing FSM status. Column [6] contains only pupils in Year 2 and column [7] contains only pupils in Year 5.*

Table 4A: Tests for Differences Between Subgroups

	Column [2] = [3]	Column [4] = [5]	Column [6] = [7]
Competition (=1) * Week 2-5	0.324	0.204	0.831
Competition (=1) * Week 2-5 (wild-p)	0.376	0.284	0.847
Competition (=1) * Week 6	0.229	0.202	0.299
Competition (=1) * Week 6 (wild-p)	0.248	0.316	0.338
Individual Incentive (=1) * Week 2-5	0.745	0.437	***0.001***
Individual Incentive (=1) * Week 2-5 (wild-p)	0.775	0.452	***0.020***
Individual Incentive (=1) * Week 6	0.965	0.364	***0.012***
Individual Incentive (=1) * Week 6 (wild-p)	0.969	0.378	***0.068***
First Group in Column Heading	Female	FSM	Year 2
Second Group in Column Heading	Male	Non-FSM	Year 5

notes: The table contains the p-values for the tests of whether the coefficient on the variables in Table 2 for the two columns listed are equal. Wild p-values shown are estimated following Cameron, Gelbrach, Miller (2008) using 1000 replications.

Table 5: Effect on Try for Sample with Week 1 less than 100% Choice and Its Subgroups

	Dependent Variable (=1) if Student Tried a Healthy Item						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Competition (=1) * Week 2-5	0.211***	0.158**	0.235**	0.275**	0.198***	0.171*	0.210***
	(0.066)	(0.073)	(0.086)	(0.097)	(0.067)	(0.086)	(0.066)
	[0.002]	[0.072]	[0.008]	[0.050]	[0.004]	[0.094]	[0.002]
Competition (=1) * Week 6	0.141**	0.101	0.154**	0.196**	0.120**	0.170***	0.090
	(0.054)	(0.080)	(0.059)	(0.088)	(0.051)	(0.057)	(0.073)
	[0.002]	[0.220]	[0.042]	[0.058]	[0.022]	[0.008]	[0.260]
Individual Incentive (=1) * Week 2-5	0.074	-0.023	0.140	0.019	0.074	-0.265***	0.245***
	(0.078)	(0.079)	(0.105)	(0.192)	(0.072)	(0.056)	(0.050)
	[0.364]	[0.821]	[0.374]	[0.879]	[0.414]	[0.008]	[0.008]
Individual Incentive (=1) * Week 6	-0.020	-0.081	0.018	-0.140	-0.026	-0.352***	0.123
	(0.095)	(0.091)	(0.119)	(0.322)	(0.091)	(0.057)	(0.081)
	[0.788]	[0.454]	[0.915]	[0.727]	[0.791]	[0.006]	[0.176]
Constant	0.436***	0.458***	0.414***	0.357***	0.449***	0.416***	0.452***
	(0.025)	(0.026)	(0.032)	(0.043)	(0.024)	(0.027)	(0.021)
Sample	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
P-value for Competition=Individual for Wks 2-5	0.167	***0.067***	0.463	0.239	0.192	***0.000***	0.662
wild P-Value: Comp = Ind Incentive Week 2-5	0.188	***0.092***	0.527	0.484	0.206	***0.004***	0.743
P-value for Competition=Individual for Wk 6	0.117	***0.047***	0.301	0.322	0.126	***0.000***	0.715
wild P-Value: Comp = Ind Incentive Week 6	0.134	***0.038***	0.326	0.521	***0.098***	***0.000***	0.779
Observations	5,466	2,583	2,883	799	4,476	2,360	3,106
R-squared	0.066	0.083	0.053	0.107	0.058	0.083	0.070
Number of pupils	215	102	113	29	179	93	122

*notes: Robust standard errors clustered at school level are included in parentheses; * sig at 10%, ** sig at 5%, *** sig at 1%. Wild P-Values are shown in brackets and are estimated following Cameron, Gelbach, Miller (2008) using 1000 replications. Column [2] includes only female and column [3] contains only males. Column [4] includes only pupils who are eligible for free school meals (FSM) and column [5] contains those pupils not eligible for FSM; there are 15 pupils for whom we are missing FSM status. Column [6] contains only pupils in Year 2 and column [7] contains only pupils in Year 5.*

Table 5A: Tests for Differences Between Subgroups

	Column [2] = [3]	Column [4] = [5]	Column [6] = [7]
Competition (=1) * Week 2-5	0.362	0.444	0.608
Competition (=1) * Week 2-5 (wild-p)	0.360	0.468	0.679
Competition (=1) * Week 6	0.528	0.441	0.292
Competition (=1) * Week 6 (wild-p)	0.601	0.513	0.324
Individual Incentive (=1) * Week 2-5	0.139	0.768	***0.000***
Individual Incentive (=1) * Week 2-5 (wild-p)	0.280	0.765	***0.000***
Individual Incentive (=1) * Week 6	0.322	0.727	***0.000***
Individual Incentive (=1) * Week 6 (wild-p)	0.362	0.695	***0.000***
First Group in Column Heading	Female	FSM	Year 2
Second Group in Column Heading	Male	Non-FSM	Year 5

notes: The table contains the p-values for the tests of whether the coefficient on the variables in Table 2 for the two columns listed are equal. Wild p-values shown are estimated following Cameron, Gelbrach, Miller (2008) using 1000 replications.

Table 6: Effects on Choice Over Treatment Weeks on Sample with Week 1 less than 100% Choice

	Dependent Variable (=1) if Student Chose a Healthy Item							
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Competition (=1) * Week 2-5	0.172***	0.243***	0.251*	0.177*	0.151	0.057	0.043	0.112
	(0.061)	(0.047)	(0.135)	(0.100)	(0.113)	(0.097)	(0.085)	(0.093)
	[0.024]	[0.002]	[0.150]	[0.156]	[0.236]	[0.607]	[0.649]	[0.330]
Individual Incentive (=1) * Week 2-5	0.099	0.033	0.056	0.073	0.064	0.266**	-0.044	0.176**
	(0.079)	(0.067)	(0.133)	(0.102)	(0.127)	(0.115)	(0.200)	(0.064)
	[0.336]	[0.643]	[0.785]	[0.557]	[0.663]	[0.254]	[0.799]	[0.162]
Constant	0.477***	0.440***	0.562***	0.587***	0.564***	0.431***	0.327***	0.546***
	(0.018)	(0.027)	(0.041)	(0.033)	(0.042)	(0.039)	(0.050)	(0.038)
Days of the Week Used	Mon-Fri	Mon	Tue	Wed	Thur	Fri	Wed-Fri	Wed-Fri
Sample Used	All	All	All	All	All	All	Missed	Not Missed
Day of Week Controls	Yes	No	No	No	No	No	Yes	Yes
P-Value: Comp = Ind Incentive	0.402	***0.006***	***0.084***	0.368	0.608	0.148	0.664	0.557
P-Value: Comp = Ind Incentive (wild)	0.432	***0.016***	***0.084***	0.384	0.621	0.348	0.677	0.661
Observations	4,745	910	977	952	975	931	876	1,982
R-squared	0.060	0.103	0.049	0.050	0.068	0.092	0.029	0.080
Number of pupils	215	212	214	215	213	213	158	202

*notes: Robust Standard Errors clustered at the school level are in brackets; * sig at 10%, ** sig at 5%, *** sig at 1%. Wild P-Values are shown in brackets and are estimated following Cameron, Gelbach, Miller (2008) using 1000 replications. The sample used in this regression are children who did not try at least some of a healthy option 100% of the time during the baseline week. The "Missed" sample in column [7] includes only those children who had not eaten any healthy times on Monday and Tuesday of the given week. The "Not Missed" sample in column [8] includes only those children who had eaten at least one fruit or vegetable on Monday or Tuesday during the given week.*

Table 7: Effects on Try Over the Week During Treatment on Sample with Week 1 less than 100% Choice

	Dependent Variable (=1) if Student Chose a Healthy Item							
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Competition (=1) * Week 2-5	0.212***	0.243**	0.241**	0.223	0.224**	0.132	0.120	0.182
	(0.069)	(0.097)	(0.100)	(0.136)	(0.104)	(0.079)	(0.110)	(0.111)
	[0.006]	[0.038]	[0.068]	[0.162]	[0.084]	[0.160]	[0.346]	[0.192]
Individual Incentive (=1) * Week 2-5	0.075	0.006	-0.060	0.047	0.121	0.240*	-0.044	0.185**
	(0.077)	(0.104)	(0.091)	(0.086)	(0.145)	(0.137)	(0.201)	(0.073)
	[0.342]	[0.955]	[0.569]	[0.595]	[0.547]	[0.348]	[0.873]	[0.242]
Constant	0.393***	0.341***	0.460***	0.497***	0.490***	0.392***	0.223***	0.589***
	(0.023)	(0.031)	(0.034)	(0.043)	(0.042)	(0.037)	(0.045)	(0.042)
Days of the Week Used	Mon-Fri	Mon	Tue	Wed	Thur	Fri	Wed-Fri	Wed-Fri
Sample Used	All	All	All	All	All	All	Missed	Not Missed
Day of Week Controls	Yes	No	No	No	No	No	Yes	Yes
P-value for Competition=Individual	0.176	***0.002***	***0.020***	0.241	0.552	0.489	0.435	0.984
P-value for Competition=Individual (wild)	0.204	***0.006***	***0.026***	0.292	0.591	0.595	0.490	1.007
Observations	4,639	884	944	935	956	920	887	1,924
R-squared	0.074	0.128	0.074	0.069	0.080	0.083	0.035	0.081
Number of pupils	215	211	213	215	212	213	157	203

*notes: Robust Standard Errors clustered at the school level are in brackets; * sig at 10%, ** sig at 5%, *** sig at 1%. Wild P-Values are shown in brackets and are estimated following Cameron, Gelbach, Miller (2008) using 1000 replications. The sample used in this regression are children who did not try at least some of a healthy option 100% of the time during the baseline week. The "Missed" sample in column [7] includes only those children who had not eaten any healthy times on Monday and Tuesday of the given week. The "Not Missed" sample in column [8] includes only those children who had eaten at least one fruit or vegetable on Monday or Tuesday during the given week.*

Table 8: Long Run Effect on Choice for Overall Sample and Its Subgroups							
	Dependent Variable (=1) if Student Tried a Healthy Item						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Panel A: Choice							
Competition (=1) * 6 months later	-0.058	-0.018	-0.104	0.045	-0.084*	-0.027	-0.102
	(0.057)	(0.055)	(0.069)	(0.127)	(0.047)	(0.057)	(0.097)
	[0.358]	[0.731]	[0.250]	[0.725]	[0.149]	[0.615]	[0.356]
Individual Incentive (=1) * 6 months later	-0.016	-0.004	-0.035	-0.121	-0.015	-0.081	0.035
	(0.070)	(0.053)	(0.091)	(0.133)	(0.067)	(0.060)	(0.100)
	[0.853]	[0.490]	[0.350]	[0.629]	[0.416]	[0.150]	[1.38]
P-Value: Comp = Ind Incentive 6 Months	0.492	0.806	0.360	***0.0943***	0.298	0.414	0.105
P-Value: Comp = Ind Incentive 6 Months (wild)	0.496	0.851	0.388	0.154	0.374	0.464	0.182
Observations	11,630	6,045	5,585	2,125	9,092	5,575	6,055
R-squared	0.013	0.013	0.015	0.023	0.014	0.012	0.023
Number of pupils	392	204	188	68	311	195	197
Panel B: Choice < 100% Choice in Week 1							
Competition (=1) * 6 months later	0.055	0.089	0.020	0.237	0.009	0.042	0.044
	(0.104)	(0.100)	(0.127)	(0.258)	(0.075)	(0.099)	(0.148)
	[0.629]	[0.394]	[0.923]	[0.432]	[0.903]	[0.677]	[0.775]
Individual Incentive (=1) * 6 months later	0.017	-0.015	0.037	0.078	-0.010	-0.040	0.044
	(0.066)	(0.064)	(0.082)	(0.186)	(0.061)	(0.138)	(0.110)
	[0.853]	[0.913]	[0.749]	[0.593]	[0.987]	[0.787]	[0.793]
P-Value: Comp = Ind Incentive 6 Months	0.695	0.297	0.888	0.402	0.825	0.625	0.996
P-Value: Comp = Ind Incentive 6 Months (wild)	0.753	0.406	0.885	0.424	0.847	0.659	1.027
Sample	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
Observations	5,072	2,321	2,751	679	4,197	1,794	3,278
R-squared	0.051	0.058	0.052	0.108	0.044	0.065	0.055
Number of pupils	168	78	90	21	141	62	106

*notes: Robust Standard Errors clustered at the school level are in brackets; * sig at 10%, ** sig at 5%, *** sig at 1%. Wild P-Values are shown in brackets and are estimated following Cameron, Gelbach, Miller (2008) using 1000 replications.*

Table 8A: Tests for Differences Between Subgroups

	Column [2] = [3]	Column [4] = [5]	Column [6] = [7]
Competition (=1) * 6 months later	0.152	0.223	0.490
Competition (=1) * 6 months later (wild-p)	0.206	0.282	0.484
Individual Incentive (=1) * 6 months later	0.601	0.406	0.332
Individual Incentive (=1) * 6 months later (wild-p)	0.587	0.478	0.448
First Group in Column Heading	Female	FSM	Year 2
Second Group in Column Heading	Male	Non-FSM	Year 5

notes: The table contains the p-values for the tests of whether the coefficient on the variables in Table 2 for the two columns listed are equal. Wild p-values shown are estimated following Cameron, Gelbrach, Miller (2008).

Table 9: Long Run Effect on Try for Overall Sample and Its Subgroups

	Dependent Variable (=1) if Student Tried a Healthy Item						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Panel A: Try							
Competition (=1) * 6 months later	-0.030	-0.009	-0.057	0.142	-0.072	-0.038	-0.022
	(0.079)	(0.059)	(0.113)	(0.151)	(0.061)	(0.067)	(0.107)
	[0.697]	[0.827]	[0.649]	[0.370]	[0.354]	[0.639]	[0.885]
Individual Incentive (=1) * 6 months later	-0.019	-0.017	-0.023	-0.023	-0.049	-0.118	0.099
	(0.092)	(0.067)	(0.127)	(0.172)	(0.080)	(0.076)	(0.111)
	[0.819]	[0.366]	[0.551]	[0.905]	[0.358]	[0.126]	[1.089]
P-Value: Comp = Ind Incentive 6 Months	0.867	0.899	0.679	0.162	0.727	0.244	***0.006***
P-Value: Comp = Ind Incentive 6 Months (wild)	0.875	0.911	0.681	0.168	0.759	0.304	***0.010***
Observations	11,021	5,796	5,224	1,974	8,673	5,504	5,517
R-squared	0.016	0.018	0.013	0.018	0.019	0.012	0.033
Number of pupils	392	204	188	68	311	195	197
Panel B: Try and <100% choice in baseline week							
Competition (=1) * 6 months later	0.029	0.020	0.035	0.159	-0.010	-0.006	0.036
	(0.110)	(0.108)	(0.129)	(0.175)	(0.091)	(0.106)	(0.157)
	[0.779]	[0.829]	[0.827]	[0.434]	[0.903]	[0.981]	[0.829]
Individual Incentive (=1) * 6 months later	-0.030	-0.060	-0.015	0.119*	-0.060	-0.130	0.023
	(0.074)	(0.080)	(0.086)	(0.061)	(0.081)	(0.125)	(0.113)
	[0.817]	[0.607]	[0.889]	[0.651]	[0.585]	[0.432]	[0.873]
Sample	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
P-Value: Comp = Ind Incentive 6 Months	0.547	0.412	0.693	0.809	0.582	0.406	0.907
P-Value: Comp = Ind Incentive 6 Months (wild)	0.523	0.513	0.711	0.817	0.581	0.468	0.913
Observations	4,944	2,258	2,686	678	4,076	1,793	3,151
R-squared	0.057	0.066	0.052	0.110	0.051	0.070	0.062
Number of pupils	168	78	90	21	141	62	106

*notes: Robust Standard Errors clustered at the school level are in brackets; * sig at 10%, ** sig at 5%, *** sig at 1%. Wild P-Values are shown in brackets and are estimated following Cameron, Gelbach, Miller (2008) using 1000 replications.*

Table 9A: Tests for Differences Between Subgroups

	Column [2] = [3]	Column [4] = [5]	Column [6] = [7]
Competition (=1) * 6 months later	0.581	***0.044***	0.865
Competition (=1) * 6 months later (wild-p)	0.631	0.144	0.887
Individual Incentive (=1) * 6 months later	0.940	0.843	***0.053***
Individual Incentive (=1) * 6 months later (wild-p)	0.927	0.859	0.112
First Group in Column Heading	Female	FSM	Year 2
Second Group in Column Heading	Male	Non-FSM	Year 5

notes: The table contains the p-values for the tests of whether the coefficent on the variables in Table 2 for the two columns listed are equal. Wild p-values shown are estimated following Cameron, Gelbrach, Miller (2008) using 1000 replications

Table 10: Food Knowledge

	Dependent Variable: Change in Food knowledge Test Score						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Competition (=1)	-0.041	-0.047	-0.035	-0.115**	-0.025	-0.059	-0.019
	(0.031)	(0.040)	(0.051)	(0.052)	(0.034)	(0.048)	(0.028)
	[0.230]	[0.256]	[0.589]	[0.076]	[0.521]	[0.204]	[0.551]
Individual Incentive (=1)	-0.018	-0.045	-0.005	0.005	-0.017	0.015	-0.048
	(0.041)	(0.053)	(0.057)	(0.061)	(0.041)	(0.062)	(0.043)
	[0.739]	[0.442]	[0.959]	[0.875]	[0.663]	[0.851]	[0.374]
Constant	0.045	0.038	0.055	0.109***	0.028	0.049	0.039
	(0.026)	(0.033)	(0.048)	(0.030)	(0.029)	(0.037)	(0.027)
1st Test Score	0.827	0.852	0.798	0.754	0.843	0.806	0.853
Mean of Dependent Variable	0.022	0.008	0.038	0.061	0.013	0.024	0.020
Sample	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
P-Value: Comp = Ind Incentive Week 2-5	0.516	0.965	0.388	***0.093***	0.818	0.220	0.418
P-Value: Comp = Ind Incentive Week 2-5 (wild)	0.507	1.003	0.426	0.172	0.801	0.234	0.494
Observations	302	162	140	45	247	164	138
R-squared	0.007	0.011	0.005	0.065	0.002	0.017	0.008

*notes: Robust standard errors clustered at school level are included in parentheses; * sig at 10%, ** sig at 5%, *** sig at 1%. Wild P-Values are shown in brackets and are estimated following Cameron, Gelbach, Miller (2008) using 1000 replications.*

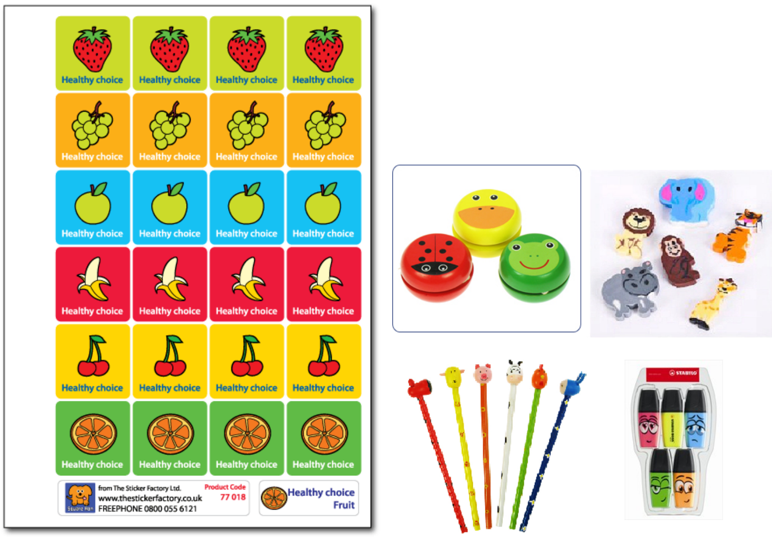
Table 11: Food Knowledge on Sample with Week 1 less than 100% Choice

	Dependent Variable: Change in Food knowledge Test Score						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Competition (=1)	-0.011	-0.032	0.017	-0.133	-0.003	-0.113	0.061***
	(0.039)	(0.040)	(0.074)	(0.182)	(0.044)	(0.097)	(0.018)
	[0.793]	[0.428]	[0.897]	[0.579]	[0.945]	[0.226]	[0.020]
Individual Incentive (=1)	-0.012	-0.076*	0.035	-0.103***	-0.017	0.044	-0.023*
	(0.038)	(0.038)	(0.063)	(0.009)	(0.044)	(0.125)	(0.011)
	[0.765]	[0.136]	[0.663]	[0.509]	[0.745]	[0.819]	[0.292]
Constant	0.023	0.035***	0.013	0.032**	0.022	0.052	0.005
	(0.027)	(0.006)	(0.046)	(0.009)	(0.035)	(0.080)	(0.005)
1st Test Score	0.847	0.872	0.821	0.848	0.854	0.798	0.874
Mean of Dependent Variable	0.015	0.001	0.030	-0.032	0.015	0.013	0.017
Sample	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
P-Value: Comp = Ind Incentive Week 2-5	0.963	0.431	0.802	0.875	0.730	0.178	***0.002***
P-Value: Comp = Ind Incentive Week 2-5 (wild)	0.987	0.484	0.751	0.935	0.753	0.222	***0.006***
Observations	118	60	58	12	99	42	76
R-squared	0.001	0.025	0.003	0.064	0.001	0.050	0.037

*notes: Robust standard errors clustered at school level are included in parentheses; * sig at 10%, ** sig at 5%, *** sig at 1%. Wild P-Values are shown in brackets and are estimated following Cameron, Gelbach, Miller (2008) using 1000 replications.*

Appendix A: Experimental Materials

Figure A1: Stickers and rewards



Appendix B: Additional Figures and Tables

Not for Publication

Figure A2: Example food knowledge test

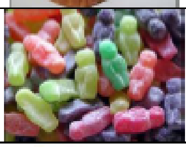
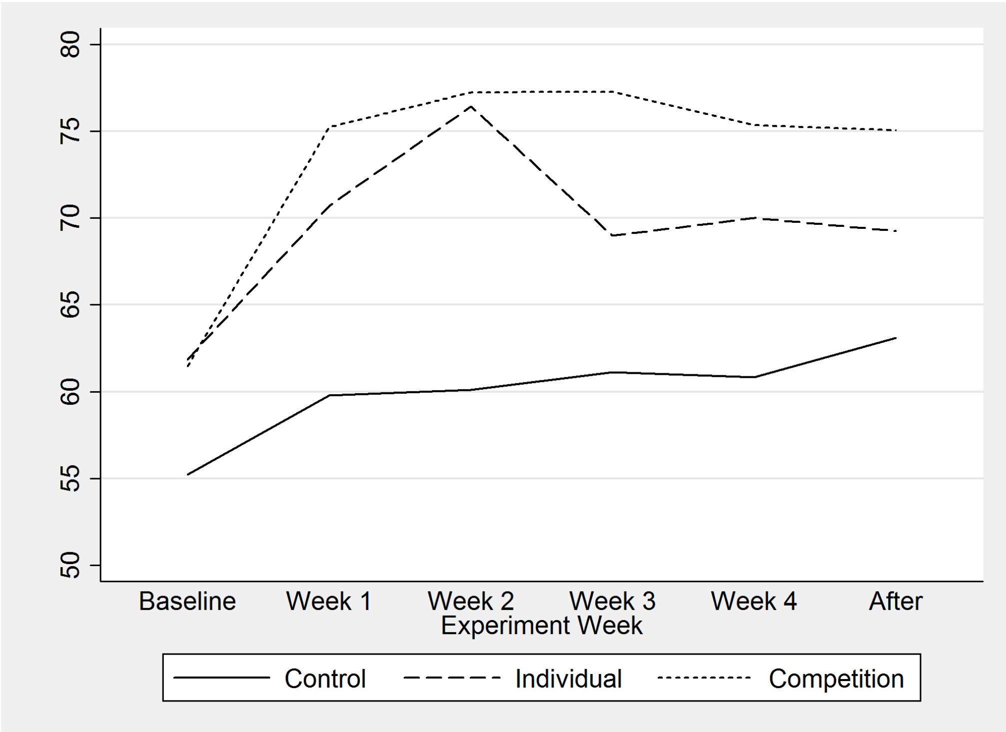
	What is it?	Is it healthy?
		Yes □ No □
		Yes □ No □
		Yes □ No □
		Yes □ No □
		Yes □ No □
		Yes □ No □

Figure B1: Proportion of pupils eating more than half a fruit or vegetable

a) Full Sample



a) Sample with less than 100% Choice in Baseline

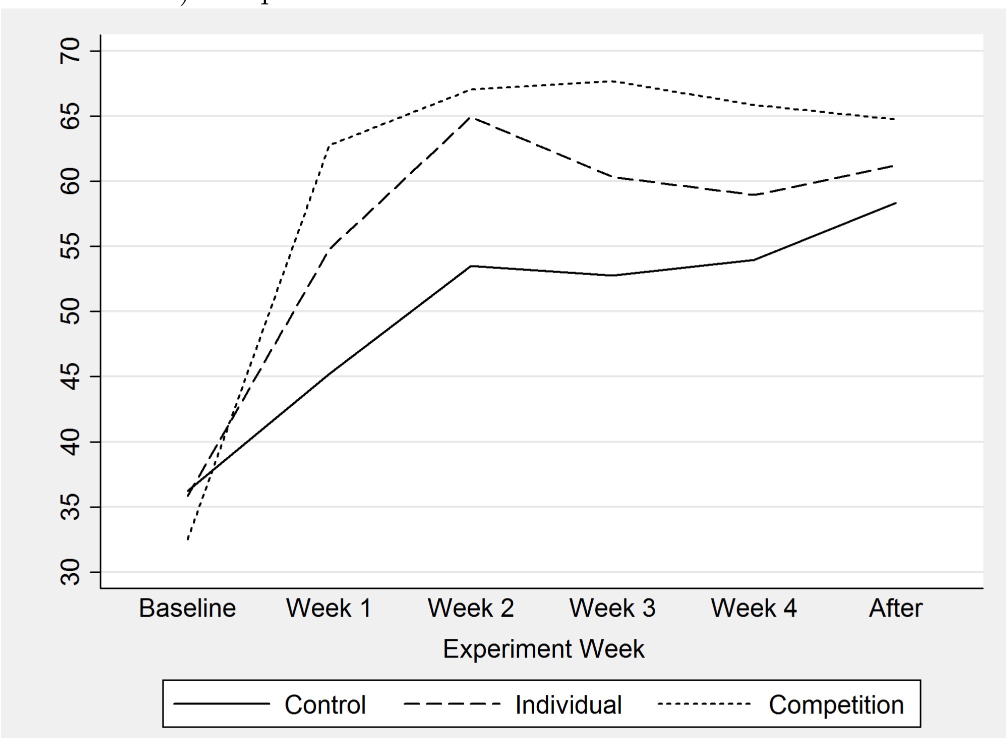


Table B1: Effect on Eating More than Half for Overall Sample and Its Subgroups

	Dependent Variable (=1) if Student Ate More than Half a Healthy Item						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Competition (=1) * Week 2-5	0.114*	0.129	0.096	0.107	0.120	0.096	0.133**
	(0.063)	(0.084)	(0.079)	(0.086)	(0.072)	(0.108)	(0.063)
	[0.194]	[0.178]	[0.288]	[0.272]	[0.144]	[0.438]	[0.070]
Competition (=1) * Week 6	0.082	0.099	0.061	0.124	0.078	0.108	0.062
	(0.073)	(0.104)	(0.073)	(0.086)	(0.088)	(0.111)	(0.083)
	[0.354]	[0.416]	[0.490]	[0.168]	[0.420]	[0.418]	[0.505]
Individual Incentive (=1) * Week 2-5	0.054	0.051	0.053	0.008	0.057	-0.054	0.219***
	(0.060)	(0.076)	(0.067)	(0.072)	(0.066)	(0.072)	(0.048)
	[0.464]	[0.561]	[0.438]	[0.927]	[0.452]	[0.498]	[0.014]
Individual Incentive (=1) * Week 6	0.008	0.040	-0.023	-0.010	0.005	-0.068	0.143
	(0.075)	(0.091)	(0.078)	(0.101)	(0.083)	(0.083)	(0.090)
	[0.893]	[0.695]	[0.813]	[0.915]	[0.989]	[0.488]	[0.172]
Constant	0.599***	0.628***	0.567***	0.592***	0.606***	0.602***	0.588***
	(0.022)	(0.029)	(0.026)	(0.029)	(0.025)	(0.032)	(0.021)
Sample	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
P-Value: Comp = Ind Incentive Week 2-5	0.410	0.356	0.638	0.320	0.437	0.109	0.193
P-Value: Comp = Ind Incentive Week 2-5 (wild)	0.488	0.428	0.687	0.360	0.460	0.164	0.256
P-Value: Comp = Ind Incentive Week 6	0.327	0.502	0.340	0.212	0.387	***0.049***	0.294
P-Value: Comp = Ind Incentive Week 6 (wild)	0.446	0.607	0.390	0.256	0.444	0.054	0.352
Observations	14,714	7,719	6,994	2,495	11,838	7,916	6,798
R-squared	0.012	0.012	0.013	0.010	0.011	0.009	0.025
Number of pupils	638	328	310	114	509	343	295

*notes: Robust standard errors clustered at school level are included in parentheses; * sig at 10%, ** sig at 5%, *** sig at 1%. Wild P-Values are shown in brackets and are estimated following Cameron, Gelbach, Miller (2008). Column [2] includes only female and column [3] contains only males. Column [4] includes only pupils who are eligible for free school meals (FSM) and column [5] contains those pupils not eligible for FSM; there are 15 pupils for whom we are missing FSM status. Column [6] contains only pupils in Year 2 and column [7] contains only pupils in Year 5.*

Table B2: Effect on Eating More Than Half for Sample with Week 1 less than 100% Choice and Its Subgroups

	Dependent Variable (=1) if Student Ate Mopre than Half a Healthy Item						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Competition (=1) * Week 2-5	0.190**	0.145	0.218**	0.268**	0.175**	0.141	0.203**
	(0.076)	(0.095)	(0.088)	(0.114)	(0.076)	(0.100)	(0.087)
	[0.024]	[0.178]	[0.042]	[0.104]	[0.038]	[0.230]	[0.036]
Competition (=1) * Week 6	0.117*	0.074	0.143**	0.245**	0.086	0.119	0.094
	(0.066)	(0.102)	(0.064)	(0.095)	(0.068)	(0.069)	(0.099)
	[0.126]	[0.501]	[0.052]	[0.058]	[0.288]	[0.172]	[0.404]
Individual Incentive (=1) * Week 2-5	0.078	0.001	0.130	0.096	0.061	-0.193***	0.216***
	(0.068)	(0.091)	(0.082)	(0.171)	(0.069)	(0.063)	(0.063)
	[0.318]	[0.973]	[0.292]	[0.695]	[0.466]	[0.016]	[0.008]
Individual Incentive (=1) * Week 6	-0.006	-0.024	0.003	0.049	-0.030	-0.326***	0.133
	(0.096)	(0.102)	(0.118)	(0.272)	(0.097)	(0.073)	(0.106)
	[0.979]	[0.795]	[0.979]	[0.617]	[0.773]	[0.004]	[0.270]
Constant	0.342***	0.372***	0.314***	0.231***	0.363***	0.291***	0.381***
	(0.025)	(0.030)	(0.029)	(0.047)	(0.025)	(0.031)	(0.027)
Sample	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
P-value for Competition=Individual for Wks 2-5	0.199	0.104	0.420	0.391	0.183	***0.001***	0.883
P-value for Competition=Individual for Wks 2-5 (wild)	0.220	0.134	0.513	0.511	0.228	***0.008***	0.879
P-value for Competition=Individual for Wk 6	0.166	0.121	0.274	0.507	0.156	***0.000***	0.692
P-value for Competition=Individual for Wk 6 (wild)	0.210	0.110	0.322	0.555	0.124	***0.000***	0.665
Observations	5,466	2,583	2,883	799	4,476	2,360	3,106
R-squared	0.057	0.065	0.052	0.082	0.051	0.072	0.058
Number of pupils	215	102	113	29	179	93	122

*notes: Robust standard errors clustered at school level are included in parentheses; * sig at 10%, ** sig at 5%, *** sig at 1%. Wild P-Values are shown in brackets and are estimated following Cameron, Gelbach, Miller (2008). Column [2] includes only female and column [3] contains only males. Column [4] includes only pupils who are eligible for free school meals (FSM) and column [5] contains those pupils not eligible for FSM; there are 15 pupils for whom we are missing FSM status. Column [6] contains only pupils in Year 2 and column [7] contains only pupils in Year 5.*

Table B3: Effect on Attendance On Overall Sample and Its Subgroups

	Dependent Variable (=1) if Student Attended School						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Competition (=1) * Week 2-5	0.017	0.002	0.037*	0.029	0.015	0.021	0.016
	(0.014)	(0.021)	(0.018)	(0.051)	(0.016)	(0.017)	(0.021)
	[0.276]	[0.897]	[0.068]	[0.621]	[0.396]	[0.304]	[0.559]
Competition (=1) * Week 6	-0.009	-0.023	0.014	-0.014	-0.006	-0.011	-0.004
	(0.017)	(0.026)	(0.020)	(0.061)	(0.015)	(0.027)	(0.028)
	[0.655]	[0.412]	[0.474]	[0.811]	[0.675]	[0.645]	[0.833]
Individual Incentive (=1) * Week 2-5	0.023	0.009	0.040*	0.002	0.029	0.015	0.032
	(0.022)	(0.029)	(0.023)	(0.042)	(0.026)	(0.018)	(0.037)
	[0.414]	[0.783]	[0.116]	[0.931]	[0.306]	[0.444]	[0.482]
Individual Incentive (=1) * Week 6	-0.022	-0.031	-0.007	-0.061*	-0.007	-0.007	-0.035
	(0.048)	(0.050)	(0.050)	(0.032)	(0.049)	(0.020)	(0.099)
	[0.733]	[0.581]	[0.937]	[0.104]	[0.865]	[0.717]	[0.809]
Constant	0.945***	0.945***	0.946***	0.971***	0.938***	0.956***	0.934***
	(0.007)	(0.009)	(0.008)	(0.014)	(0.008)	(0.007)	(0.013)
Sample	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
P-value for Competition=Individual for Wks 2-5	0.800	0.814	0.877	0.411	0.551	0.790	0.634
P-value for Competition=Individual for Wks 2-5 (wild)	0.831	0.859	0.917	0.482	0.579	0.837	0.689
Observations	16,472	8,548	7,917	2,843	13,200	8,596	7,876
R-squared	0.003	0.002	0.004	0.009	0.002	0.001	0.007
Number of pupils	643	331	312	115	513	345	298

*notes: Robust standard errors clustered at school level are included in parentheses; * sig at 10%, ** sig at 5%, *** sig at 1%. Wild P-Values are shown in brackets and are estimated following Cameron, Gelbach, Miller (2008). Column [2] includes only female and column [3] contains only males. Column [4] includes only pupils who are eligible for free school meals (FSM) and column [5] contains those pupils not eligible for FSM; there are 15 pupils for whom we are missing FSM status. Column [6] contains only pupils in Year 2 and column [7] contains only pupils in Year 5.*

Table B4: Effect on Attendance for Sample with Week 1 less than 100% Choice and Its Subgroups

	Dependent Variable (=1) if Student Attended School						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Competition (=1) * Week 2-5	-0.015	-0.063	0.030	0.046**	-0.025	-0.032	0.011
	(0.023)	(0.037)	(0.027)	(0.019)	(0.028)	(0.038)	(0.034)
	[0.563]	[0.322]	[0.380]	[0.076]	[0.424]	[0.424]	[0.785]
Competition (=1) * Week 6	-0.062**	-0.130***	0.010	-0.003	-0.067**	-0.081*	-0.034
	(0.022)	(0.041)	(0.034)	(0.036)	(0.029)	(0.042)	(0.036)
	[0.034]	[0.04]	[0.765]	[0.777]	[0.070]	[0.054]	[0.394]
Individual Incentive (=1) * Week 2-5	0.062	0.041	0.078**	0.040***	0.065	0.057	0.063
	(0.040)	(0.060)	(0.035)	(0.005)	(0.044)	(0.070)	(0.048)
	[0.204]	[0.533]	[0.066]	[0.124]	[0.208]	[0.440]	[0.386]
Individual Incentive (=1) * Week 6	0.045	-0.020	0.091**	-0.100	0.059	0.028	0.053
	(0.041)	(0.071)	(0.042)	(0.059)	(0.044)	(0.096)	(0.034)
	[0.266]	[0.823]	[0.014]	[0.507]	[0.206]	[0.789]	[0.240]
Constant	0.909***	0.901***	0.915***	0.980***	0.894***	0.931***	0.892***
	(0.010)	(0.014)	(0.011)	(0.007)	(0.012)	(0.016)	(0.014)
Sample	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
P-value for Competition=Individual for Wks 2-5	0.0443	0.0496	0.256	0.757	0.0324	0.163	0.233
P-value for Competition=Individual for Wks 2-5 (wild)	0.130	0.228	0.306	0.785	0.136	0.150	0.430
Observations	6,085	2,870	3,210	838	5,047	2,582	3,503
R-squared	0.008	0.016	0.006	0.014	0.010	0.006	0.011
Number of pupil	220	105	115	30	183	95	125

*notes: Robust standard errors clustered at school level are included in parentheses; * sig at 10%, ** sig at 5%, *** sig at 1%. Wild P-Values are shown in brackets and are estimated following Cameron, Gelbach, Miller (2008). Column [2] includes only female and column [3] contains only males. Column [4] includes only pupils who are eligible for free school meals (FSM) and column [5] contains those pupils not eligible for FSM; there are 15 pupils for whom we are missing FSM status. Column [6] contains only pupils in Year 2 and column [7] contains only pupils in Year 5.*

Table B5: Effect on Bringing Packed Lunch On Overall Sample and Its Subgroups

	Dependent Variable (=1) if Student Brought a Packed Lunch						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Competition (=1) * Week 2-5	0.000	0.000	-0.001	0.014	0.001	0.008	-0.014
	(0.021)	(0.032)	(0.023)	(0.039)	(0.025)	(0.034)	(0.025)
	[0.993]	[0.995]	[0.957]	[0.737]	[0.951]	[0.849]	[0.635]
Competition (=1) * Week 6	-0.038	-0.065	-0.003	0.008	-0.042	-0.063	-0.020
	(0.030)	(0.046)	(0.033)	(0.044)	(0.038)	(0.043)	(0.036)
	[0.220]	[0.176]	[0.923]	[0.883]	[0.332]	[0.202]	[0.621]
Individual Incentive (=1) * Week 2-5	-0.013	-0.001	-0.022	-0.038*	0.004	-0.014	-0.014
	(0.025)	(0.035)	(0.020)	(0.021)	(0.024)	(0.033)	(0.037)
	[0.569]	[1.02]	[0.394]	[0.200]	[0.827]	[0.681]	[0.815]
Individual Incentive (=1) * Week 6	-0.041	-0.037	-0.040	-0.057	-0.021	-0.078*	-0.008
	(0.036)	(0.052)	(0.029)	(0.042)	(0.036)	(0.043)	(0.055)
	[0.256]	[0.509]	[0.268]	[0.258]	[0.587]	[0.128]	[0.919]
Constant	0.499***	0.489***	0.511***	0.187***	0.566***	0.461***	0.539***
	(0.008)	(0.011)	(0.008)	(0.013)	(0.009)	(0.009)	(0.012)
Sample	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
P-value for Competition=Individual for Wks 2-5	0.525	0.968	0.421	0.255	0.919	0.0684	0.996
P-value for Competition=Individual for Wks 2-5 (wild)	0.583	1.035	0.482	0.306	0.865	0.092	0.957
Observations	14,575	7,622	6,953	2,501	11,671	7,348	7,227
R-squared	0.002	0.002	0.002	0.004	0.002	0.003	0.002
Number of pupils	623	322	301	110	498	329	294

*notes: Robust standard errors clustered at school level are included in parentheses; * sig at 10%, ** sig at 5%, *** sig at 1%. Wild P-Values are shown in brackets and are estimated following Cameron, Gelbach, Miller (2008). Column [2] includes only female and column [3] contains only males. Column [4] includes only pupils who are eligible for free school meals (FSM) and column [5] contains those pupils not eligible for FSM; there are 15 pupils for whom we are missing FSM status. Column [6] contains only pupils in Year 2 and column [7] contains only pupils in Year 5.*

Table B6: Effect on Bringing Packed Lunch for Sample with Week 1 less than 100% Choice and Its Subgroups

	Dependent Variable (=1) if Student Brought a Packed Lunch						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Competition (=1) * Week 2-5	0.007	-0.021	0.040	0.033	-0.000	0.020	-0.019
	(0.023)	(0.040)	(0.026)	(0.118)	(0.023)	(0.029)	(0.041)
	[0.719]	[0.641]	[0.124]	[0.783]	[0.991]	[0.543]	[0.657]
Competition (=1) * Week 6	-0.004	-0.076	0.080**	-0.006	0.003	-0.039	0.005
	(0.036)	(0.071)	(0.030)	(0.121)	(0.043)	(0.071)	(0.058)
	[0.957]	[0.348]	[0.032]	[0.985]	[0.971]	[0.515]	[0.925]
Individual Incentive (=1) * Week 2-5	0.036	0.054*	0.022	0.007	0.053*	0.060	0.027
	(0.025)	(0.030)	(0.038)	(0.005)	(0.028)	(0.054)	(0.022)
	[0.204]	[0.182]	[0.643]	[0.430]	[0.072]	[0.595]	[0.408]
Individual Incentive (=1) * Week 6	0.018	0.044	-0.003	-0.017	0.048	-0.039	0.050
	(0.046)	(0.076)	(0.041)	(0.014)	(0.041)	(0.072)	(0.057)
	[0.751]	[0.651]	[0.941]	[0.505]	[0.350]	[0.527]	[0.645]
Constant	0.532***	0.527***	0.536***	0.355***	0.564***	0.509***	0.549***
	(0.009)	(0.015)	(0.011)	(0.042)	(0.009)	(0.013)	(0.012)
Sample	All	Females	Males	FSM	Non-FSM	Year 2	Year 5
P-value for Competition=Individual for Wks 2-5	0.318	***0.0518***	0.646	0.825	***0.0749***	0.466	0.262
P-value for Competition=Individual for Wks 2-5 (wild)	0.384	0.112	0.697	0.821	0.100	0.781	0.302
Observations	5,376	2,555	2,821	771	4,412	2,195	3,181
R-squared	0.001	0.004	0.002	0.002	0.001	0.004	0.002
Number of pupils	214	102	112	29	178	93	121

*notes: Robust standard errors clustered at school level are included in parentheses; * sig at 10%, ** sig at 5%, *** sig at 1%. Wild P-Values are shown in brackets and are estimated following Cameron, Gelbach, Miller (2008). Column [2] includes only female and column [3] contains only males. Column [4] includes only pupils who are eligible for free school meals (FSM) and column [5] contains those pupils not eligible for FSM; there are 15 pupils for whom we are missing FSM status. Column [6] contains only pupils in Year 2 and column [7] contains only pupils in Year 5.*